Mycology
A Comprehensive Approach

Thomas Carrey

Mycology: A Comprehensive Approach

Mycology: A Comprehensive Approach

Editor: Thomas Carrey

Table of Contents

The fungal composition of natural biofinishes on oil-treated wood

Elke J. van Nieuwenhuijzen[1*], Jos A. M. P. Houbraken[1], Peter J. Punt[2,3], Guus Roeselers[2,4], Olaf C. G. Adan[5] and Robert A. Samson[1]

Abstract

Background: Biofinished wood is considered to be a decorative and protective material for outdoor constructions, showing advantages compared to traditional treated wood in terms of sustainability and self-repair. Natural dark wood staining fungi are essential to biofinish formation on wood. Although all sorts of outdoor situated timber are subjected to fungal staining, the homogenous dark staining called biofinish has only been detected on specific vegetable oil-treated substrates. Revealing the fungal composition of various natural biofinishes on wood is a first step to understand and control biofinish formation for industrial application.

Results: A culture-based survey of fungi in natural biofinishes on oil-treated wood samples showed the common wood stain fungus *Aureobasidium* and the recently described genus *Superstratomyces* to be predominant constituents. A culture-independent approach, based on amplification of the internal transcribed spacer regions, cloning and Sanger sequencing, resulted in clone libraries of two types of biofinishes. *Aureobasidium* was present in both biofinish types, but was only predominant in biofinishes on pine sapwood treated with raw linseed oil. Most cloned sequences of the other biofinish type (pine sapwood treated with olive oil) could not be identified. In addition, a more in-depth overview of the fungal composition of biofinishes was obtained with Illumina amplicon sequencing that targeted the internal transcribed spacer region 1. All investigated samples, that varied in wood species, (oil) treatments and exposure times, contained *Aureobasidium* and this genus was predominant in the biofinishes on pine sapwood treated with raw linseed oil. *Lapidomyces* was the predominant genus in most of the other biofinishes and present in all other samples. Surprisingly, *Superstratomyces*, which was predominantly detected by the cultivation-based approach, could not be found with the Illumina sequencing approach, while *Lapidomyces* was not detected in the culture-based approach.

Conclusions: Overall, the culture-based approach and two culture-independent methods that were used in this study revealed that natural biofinishes were composed of multiple fungal genera always containing the common wood staining mould *Aureobasidium*. Besides *Aureobasidium*, the use of other fungal genera for the production of biofinished wood has to be considered.

Keywords: Biofilm, Metagenomics, Mould, Wood protection, Wood staining

Background

Microbial growth causing discolouration on surfaces of outdoor situated materials is a common phenomenon [1–3]. Frequently these microbial stains are referred to as biofilm, although not all commonly accepted biofilm criteria might have been investigated [4]. Dark staining of painted and unpainted wood is mostly attributed to fungi and generally considered as unwanted discolouration [5, 6]. In contrast, the specific dark stain formation on wood called biofinish is considered to be a functional colouration [4] (Fig. 1). The colouration of a biofinish is, together with its presumed protection and self-healing properties, an important ingredient of a sustainable solution for a biocide free wood finish system [4, 7]. A

*Correspondence: e.nieuwenhuijzen@cbs.knaw.nl
[1] Applied and Industrial Mycology, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands
Full list of author information is available at the end of the article

Fig. 1 Oil-treated pine sapwood samples with natural formed biofinish and weathered samples without biofinish. Samples in the *left row*: treated with raw linseed oil. Samples in the *middle row*: treated with olive oil. Samples in the *right row*: untreated. *Scale bar* 10 mm

biofinish refers to a dark pigmented layer, that covers a wood surface almost entirely without exposing underlying wood structures, contains abundant microbial mass and that is irreversibly attached to the surface [4]. Biofinishes have been detected on wood impregnated with olive oil or raw linseed oil [4, 8]. The study described in this paper is focused on the characterization of the fungal composition of these biofinishes.

Although several fungal species are associated with outdoor wood staining [6, 9], little is known to which extent each taxon contributes to this staining. In some studies the fungal populations on timber surfaces were quantified [10, 11]. The study by Sailer et al. [7] provides data, particularly of interest for the biofinishes on wood. One specific sample, made of pine sapwood and impregnated with refined linseed oil dissolved in acetone, was used to study the fungal composition of a homogenous dark stained wood surface. Later, biofinishes were detected on other outdoor exposed oil-treated wood samples, including samples made of different wood species treated with olive oil. Analysis of these samples revealed the abundancy of the wood staining fungus *Aureobasidium*, but the results also indicated that this genus might not always dictate the fungal population of dense dark stained wood samples [4, 12]. The wood species, oil type and the geographical location could influence the fungal community composition of a biofinish. In order to manufacture a stable biofinish and eventually apply biofinished wood in practice, a detailed composition of fungi present at various stained wood surfaces is elucidated in this paper.

Different techniques are available to study a fungal community on the surface of an environmental sample such outdoor exposed wood. Each technique has its advantages and disadvantages. Microscopic examination of the surface can be an easy method to study the surface of a material [13]. The main disadvantage is that identification and enumeration is difficult in case of moulds on oil-treated wood surfaces [4]. A common way to identify and enumerate fungi is by culturing. A swab-based method can be used to analyse the culturable fungi that are present on a sample surface [4, 14]. This includes the determination of the number of colony forming units after incubation and identification based on macroscopic, microscopic and/or molecular analysis. Culture-based analysis only allows the detection of readily culturable species, which is effected by the media selection and overestimates the presence of abundantly sporulating species [13]. Culture-independent methods based on DNA analysis are frequently used and show complementary results [15–17]. Culture-independent approaches that rely on Next Generation Sequencing (NGS) methods have become state of the art to study microbial communities [16, 18, 19]. Albeit that these NGS methods provide advantages compared to earlier developed techniques, research to control and understand the biases occurring in all steps of a NGS method is still ongoing [18–20].

The objective of the present study was to analyse the fungal composition of various biofinishes on oil-treated wood surfaces. A culture-based swab method and two non-culturing methods based on either amplicon-cloning followed by Sanger sequencing or Illumina amplicon

sequencing were selected for the analysis. Wood samples without oil treatment and/or biofinish were used to compare the diversity and predominance of fungal genera.

Methods

Several natural biofinishes were studied with a culture-based and two DNA sequencing-based methods (Table 1). The biofinish containing samples varied in wood species, type of oil and origin. Wood samples made of the same wood species without a biofinish, were studied as well. The viable fungal composition was studied of all samples. Specific samples, which were exposed in the Netherlands, were selected for the culture-independent fungal profiling methods.

Wood samples

The different wood species tested were pine (*Pinus sylvestris*), spruce (*Picea abies*) and ilomba (*Pycnanthus angolensis)*. Pine samples were made totally of sapwood (sw) or a mixture of sapwood and heartwood (hw). No specific sapwood or heartwood selection was made for spruce and ilomba. Wood blocks were impregnated with raw linseed, stand linseed or olive oil. Sets of impregnated and untreated wood samples were exposed outdoors at different locations (Table 1). The sample dimensions, oil treatments, outdoor exposure and handling procedures were described in van Nieuwenhuijzen et al. [4]. All samples, except for the samples of set 3, have been analysed with the biofinish assessment method. This method consists of observations of the dark stained surface coverage at macroscopic and microscopic scale, and spectrophotometer measurements of the pigmentation [4]. In summary, a biofinish is assigned when more than 90% of the surface is stained and does not expose structures of the wood such as annual rings or wood fibres, and the pigmentation measurements, expressed by sRGB colour space triplets, meet specific criteria (R, G and B values are below 82 and the value difference within a single RGB triplet is below 20). The presence of biofinishes on the wood samples of set 3 was only estimated with visual observations of the stain coverage as described in the biofinish assessment method, which can overestimate biofinish identification.

Culturing colonies

The concentration of colony forming units (CFU) per cm^2 wood was determined for each specimen (Table 1). For this, biomass collected with a cotton swab was analysed as described by van Nieuwenhuijzen et al. [4]. Serial dilutions were plated in duplicate (set 1 and 3–5) or triplicate (sample set 2) on dichloran 18% glycerol agar (DG18) and malt extract agar (MEA) supplemented with penicillin and streptomycin (P/S). The agar media was prepared as described by Samson et al. [13]. The total number of colonies and those phenotypically resembling *Aureobasidium* was determined after seven and fourteen days of incubation at 25 °C [12]. Besides the *Aureobasidium* colonies, also the predominant colonies were counted. Two or more colonies of each predominantly present colony type were transferred to new MEA plates. The isolates were deposited in the working collection of the Applied and Industrial Mycology department (DTO) housed at the CBS-KNAW Fungal Biodiversity Centre, The Netherlands and subjected to molecular identification.

DNA extraction

DNA was extracted from cultures grown on MEA plates according to van Nieuwenhuijzen et al. [12]. With respect to the culture-independent methods, biomass was removed with a sterile scalpel from the upper surface of a mould stained wood sample, collected on sterile paper and subsequently used to extract DNA [4]. In both cases the Ultraclean Microbial DNA isolation kit (MoBio Laboratories, USA) was used according to manufacturer's instructions.

PCR and Sanger sequencing of isolates

The nuclear internal transcribed spacers including the 5.8S rRNA gene (ITS) of fungal isolates were amplified with the primer pair V9G [21] and LS266 [22]. In case additional sequence information was needed for a proper identification of a strain, the *LSU* gene was partially amplified using the forward primer LROR (Reh) GTACCGCTTGAACTTAAGC [23] or LROR (VilU) ACCCGCTGAACTTAAGC [Vilgalys, unpublished] and the reverse primer LR5 or LR7 [24]. The polymerase chain reaction (PCR) mixtures had final concentrations of: 4% DNA extract, 10% PCR buffer, 3% $MgCl_2$ (25 mM), 65.8% demineralised sterile water, 7.8% dNTP (1 mM), 5% DMSO, 2% forward primer (10 μM), 2% reverse primer and 0.4% Taq polymerase (5 U/mL, BioTaq, Bioline). The PCR program typically consisted of 1 cycle of 5 min denaturation at 95 °C; 35 cycles of 35 s denaturation at 95 °C, followed by ITS-primer annealing for 30 s at 55 °C or *LSU*-primer annealing at 54 °C for 50 s, and an extension for 1.5 min at 72 °C. The PCR-products were sequenced with the same primers as used for PCR amplification using the BigDye Terminator v. 3.1 Cycle Sequencing Kit (Applied Biosystems, USA). Sequence products were analysed on an ABI PRISM 3730XL genetic analyser (Applied Biosystems, USA) and traces were assembled using Seqman Pro v. 9.0.4 (DNAstar Inc.). The sequences were deposited in GenBank [25].

PCR, cloning and Sanger sequencing

ITS-specific clone libraries were made as described in van Nieuwenhuijzen et al. [12] of two types of biofinishes,

Table 1 Overview of the amount and type of wood samples used per type of fungal profiling method

Sample set	Wood species	Treatment	Presence biofinish	Location (exposure time)	Number of samples		
					Culture method	Cloning method	Illumina method
1	Spruce	Raw linseed oil	No	Utrecht, The Netherlands (1.5 year)	3		1
		Stand linseed oil	No		3		1
		Olive oil	Yes[a]		3		1
		No oil	No		3		1
	Pine sw	Raw linseed oil	Yes[a]		3	3	2
		Stand linseed oil	No		3		1
		Olive oil	Yes[a]		3	3	2
		No oil	No		3		1
	Ilomba	Raw linseed oil	No		3		1
		Stand linseed oil	No		3		1
		Olive oil	Yes[a]		3		1
		No oil	No		3		1
2	Pine sw	Raw linseed oil	Yes[a]	Utrecht, The Netherlands (1.8 year)	10		2
3	Spruce	Raw linseed oil	No	Utrecht, the Netherlands (1.5 year)	1		
	Ilomba	Raw linseed oil	No		1		
	Pine sw	Raw linseed oil	Yes[b]		1		
	Pine sw	Olive oil	Yes[b]		1		
	Pine sw	No oil	No		1		
	Pine hw	Raw linseed oil	Yes[b]		1		
4	Same materials as set 3		Biofinish only on olive oil	Johannesburg, South Africa (1.7 year)	1		
5	Same materials as set 3		No	Dover Gardens, Australia (1,5 year)	1		
6	Same materials as set 3		Biofinish on pine and r. lins.	Ås, Norway (2 year)	1		

[a] Biofinish assessment by Nieuwenhuijzen, van et al. [4]

[b] Determination based on stain coverage

formed on different substrates exposed in the Netherlands (Table 1, set 1): pine sapwood treated with raw linseed oil (libraries PRL.1, PRL.2 and PRL.3) and pine sapwood treated with olive oil (PO.1, PO.2 and PO.3). The ITS region was amplified with the primers V9G and LS266 and the GoTaq Long PCR Master Mix (Progema), while using the PCR-program as described above. Purified PCR products (QIAquick PCR purification kit) were ligated and cloned (pGEM®-T Easy Vector Systems) into an *Escherichia coli* plasmid library. Amplification and Sanger sequencing of DNA from ITS containing competent cells was performed as described in van Nieuwenhuijzen et al. [12]. The sequences were deposited in GenBank [25].

Illumina ITS1 amplicon sequencing
Internal transcribed spacer 1 region (ITS1) amplicon libraries were made of 16 wood samples that were all exposed at one test site (The Netherlands), but variated in wood species, treatment and the presence or absence of a biofinish (Table 1, set 1–2). In a preliminary study several ITS primer combinations and amplification approaches were tested for their suitability of generic detection of fungal genera (unpublished results). Based on these results barcoded ITS1 amplicons were generated using a two-step PCR approach. ITS1 regions were first amplified with the following primers: nex-ITS-BITS-F: TCGTCGGCAGCGT-CACCTGCGGARGGATCA and nex-ITS-B58S3-R: GTCTCGTGGGCTCGGGAGATCCRTTGYTRAAA-GTT (adapted from Bokulich and Mills [26]). Each reaction contained $300\times$ purified DNA, 1X hot start PCR master mix (Thermo Scientific) and nuclease free PCR grade water to a 50 µl final reaction volume. PCR reactions consisted of an initial denaturation step of 95 °C for 5 min and 30 amplification cycles (95 °C for 30 s, annealing 52 °C for 45 s and elongation 72 °C for 1 min) and a final extension step (72 °C for 10 min) followed by

Table 2 The predominantly cultured colony types of each wood sample set

Sample set	Wood species	Treatment	No. of samples	Biofinish	Number of wood samples with predominant colony types											
					Aureobasidium	Black yeasts	Cladosporium	Cryptococcus	Didymellaceae	Phacidiella	Pleurophoma	Pyrenochaeta	Cyanodermella	Superstratomyces	Sydowia	Taphrina
1	Spruce	R. lins. oil	3	No	3	1										
		St. lins. oil	3	No	3	3										
		Olive oil	3	Yes[a]	2									3		
		No oil	3	No	1									3		
	Pine sw	R. lins. oil	3	Yes[a]	1	1								2		1
		St. lins. oil	3	No		3								3		
		Olive oil	3	Yes[a]										3		
		No oil	3	No		2								3		
	Ilomba	R. lins. oil	3	No										2		
		St. lins. oil	3	No		2	1		1					2		
		Olive oil	3	Yes[a]										3		
		No oil	3	No			3				1			1		
2	Pine sw	R. lins. oil	10	Yes[a]	3	1				1		1		10		2
3	Spruce	R. lins. oil	1	No										1		
	Ilomba	R. lins. oil	1	No										1		
	Pine sw	R. lins. oil	1	Yes[b]	1									1		
	Pine sw	Olive oil	1	Yes[b]										1		
	Pine sw	No oil	1	No		1										
	Pine hw	R. lins. oil	1	Yes[b]	1									1		
4	Spruce	R. lins. oil	1	No	1	1										
	Ilomba	R. lins. oil	1	No	1											
	Pine sw	R. lins. oil	1	No	1											
	Pine hw	R. lins. oil	1	No	1	1										
	Pine sw	Olive oil	1	No				1								
	Pine sw	No oil	1	No	1											
	Pine hw	R. lins. oil	1	No	1								1			
5	Spruce	R. lins. oil	1	No	1											
	Ilomba	R. lins. oil	1	No	1	1										
	Pine sw	R. lins. oil	1	No	1											
	Pine sw	Olive oil	1	No	1											
	Pine sw	No oil	1	No	1											
	Pine hw	R. lins. oil	1	No	1	1										
6	Spruce	R. lins. oil	1	No											1	
	Ilomba	R. lins. oil	1	No		1									1	
	Pine sw	R. lins. oil	1	Yes[a]											1	
	Pine sw	Olive oil	1	No											1	
	Pine sw	No oil	1	No											1	
	Pine hw	R. lins. oil	1	Yes[a]	1	1									1	

[a] Biofinish detection by van Nieuwenhuijzen et al. [4]

[b] Determination based on stain coverage

cool down (10 min at 4 °C). A negative control (blank) was included for each 24 PCR reactions. Reactions were cleaned by solid-phase reversible immobilization (SPRI) using AMPure XP SPRI beads (Beckman Coulter, Inc.). Dual barcodes (8 bp) and Illumina Sequencing adapters were attached using the Nextera XT Index Kit (Illumina, San Diego, CA) according to manufacturer's protocols. Barcoded amplicons were quantified using the Caliper LabChip GX II system (Perkin Elmer, Hopkinton, USA), normalised to the same concentrations, pooled, and gel purified using the Qiaquick spin kit (Qiagen) and AMPure XP SPRI beads. Pooled amplicons were 250-bp paired-end sequenced using the MiSeq system (Illumina). Raw Illumina fastq files were deposited in the European Nucleotide Archive (accession PRJEB13755). The raw data was demultiplexed, quality filtered, and analysed using modules implemented in the Mothur software platform [27]. Quality filtering involved a minimal quality score of 25 in a window of 5 and removal of reads with a length below 100 and above 500 or with ambiguous bases. In addition, chimeras detected using UCHIME and the UNITE database (v6) [28] were removed. Filtered raw reads were merged into paired reads. Subsequently, the relative abundance of unique sequences were calculated for each sample by dividing the number of reads of a single unique sequence by the total number of reads of the sample. Unique sequences with a total sum of relative read abundances above 0.1% were used for further analysis.

DNA data analysis

ITS and ITS1 sequences were subjected to nucleotide BLAST searches [29] using the non-redundant database of GenBank [25], the Q-bank Fungi database [30] and an internal database of the CBS-KNAW Fungal Biodiversity Centre (Fungal Barcoding data). Identification was performed on genus level. ITS sequences obtained from non-cultured material, which resulted in hits in GenBank with an identity below 97% were marked as 'unidentified'. Also the ITS1 sequences with query coverages below 90% were marked as 'unidentified'. For the identification of culturable isolates with an unclear identity based on only ITS sequences, *LSU* sequences were also compared with the non-redundant nucleotide database of GenBank. ITS sequences of isolates that could not be identified on genus level were aligned with the sequences of the ITS cloning and ITS1 amplicon libraries using Nucleotide BLAST of GenBank(NCBI, Rockville Pike, USA).

Mathematical analysis

A Shannon's diversity test was used to index the genus diversity in the Illumina amplicon data.

Results

Culturable fungal composition

After culturing biomass of 70 wood samples, the predominantly colony types of 68 samples could be identified (Table 2). The total number CFU's per square centimetre biofinish varied between 6×10^2 and 4×10^5 CFU/cm^2 and the samples without biofinish, which all contained dark mould stains but less than the biofinish criteria prescribed, showed similar results (2×10^1 to 4×10^5 CFU/cm^2; Additional file 1: Table S1). Frequently, more than one predominant colony type could be determined in the plated biomass from a single wood sample (Table 2). The isolates and sequence data obtained to identify the predominant type of CFU are listed in Additional file 2: Table S2.

A recently described dark pycnidia producing coelomycete named *Superstratomyces* was detected as the most commonly occurring predominant colony type in the biofinishes (Table 2; Fig. 2) [31]. This genus was also predominantly present on outdoor exposed samples without a biofinish, including samples without oil (Table 2). Interestingly, the detection of *Superstratomyces* on wood was restricted to samples exposed in the Netherlands (sample set 1–3).

Aureobasidium was predominantly present on 8 of the 28 samples that showed a biofinish (Table 2; Fig. 2). Also samples without a biofinish frequently showed this genus to be one of the predominant colony types (Table 2; Fig. 2). *Aureobasidium* was isolated from samples existing of all combinations of wood species and (oil) treatments originating from all selected outdoor locations (Additional file 1: Table S1, Additional file 2: Table S2). The *Aureobasidium* contribution to the total cultured CFU varied largely for biofinish samples (0–97%), but the contribution per sample surface area only rarely (2 out of 28 samples) exceeded the 50% (Additional file 1: Table S1). The stained samples without a biofinish showed a similar range of percentages (0–97%), and the *Aureobasidium* contribution exceeded the 50% regurlary (13 out of the 42 samples; Additional file 1: Table S1).

Other colonies types which were predominantly isolated from samples with a biofinish were black yeasts (identfified as *Exophiala*, *Phaeococcomyces* or *Knufia*), *Taphrina*, *Sydowia*, *Phacidiella* and *Pyrenochaeta*. Black yeasts and *Sydowia* were also isolated as predominant colony types from the stained samples without a biofinish, expended by the genera *Cladosporium*, *Cryptococcus*, *Pleurophoma* and *Cyanodermella*, and a genus in the family *Didymellaceae* (Table 2).

Fungal composition of ITS clone libraries

Six ITS clone libraries were constructed from biofinish DNA obtained from pine sapwood samples treated with

Fig. 2 Identified colony types predominantly isolated from 28 outdoor exposed wood samples with a biofinish
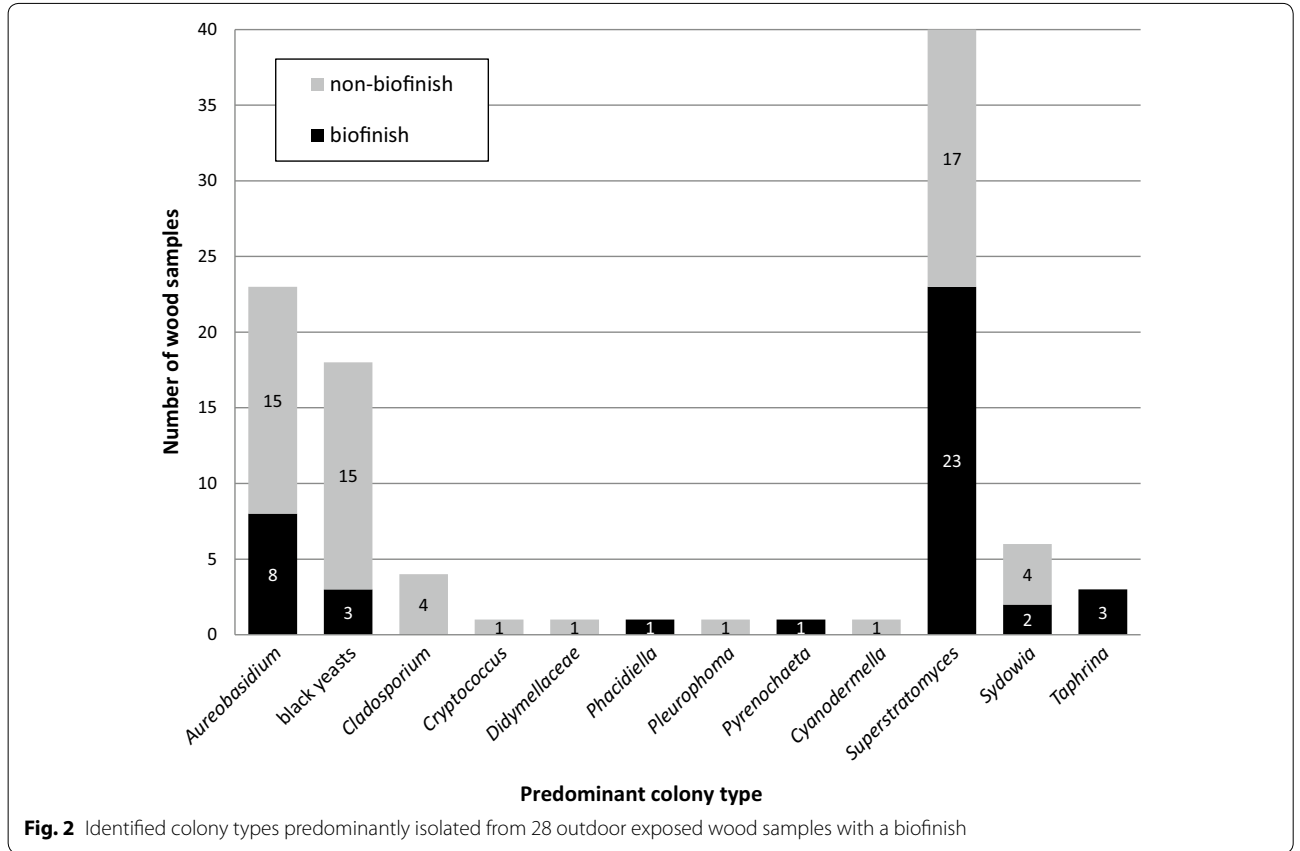
either raw linseed or olive oil. Each library contained 61–71 clones (Additional file 3: Table S3). In all libraries several genera were identified always including *Aureobasidium* (Fig. 3).

Aureobasidium was predominantly present in the cloned DNA of biofinishes on pine sapwood treated with raw linseed oil. In each of the three clone libraries more than 50% of the clones were identified as *Aureobasidium*. The investigated biofinishes on pine sapwood treated with olive oil did not show this predominance. Their clone libraries had a much lower *Aureobasidium* percentage, varying from 2 to 13%. Interestingly, the number of sequences which could not be assigned to the genus level was high in these samples. These divergent sequences all differed from the molecular identified strains obtained with the culturing method. Many of them had more than one best hit with sequences in the database while showing identity scores of 82% compared to sequences in GenBank that were named *Aureobasidium*, *Sarcinomyces* and *Rhizosphaera*. The other divergent sequences showed identity coverages varying from 80 to 96% compared to the best hits, which represented up to five genus names for each library (Additional file 3: Table S3). Furthermore, analysis of

Fig. 3 Genus composition in ITS clone libraries constructed from biofinish DNA. 'Unidentified': sequences resulting in best hits with an identity below 97%
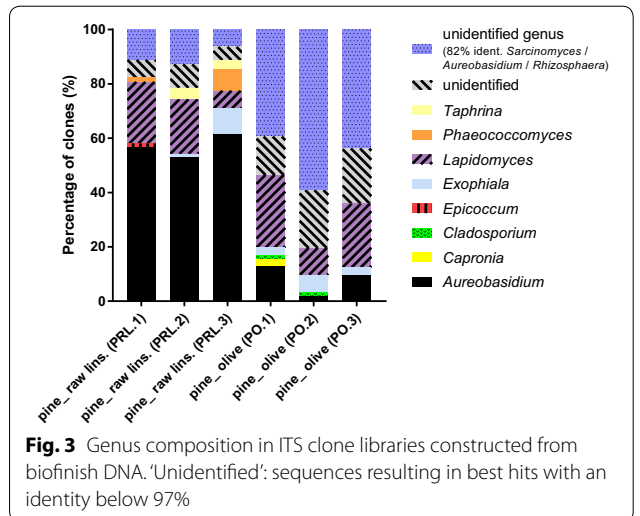
sequences obtained from all biofinish samples revealed the presence of other dark pigmented fungi, with *Lapidomyces* as a major contributor (6–27% per library). The in general more sparsely occurring genera were *Capronia*, *Cladosporium*, *Exophiala*, *Phaeococcomyces* and *Epicoccum*.

Fig. 4 Genus composition in Illumina ITS1 amplicon libraries constructed from DNA of fungal stained wood surfaces. 'Unidentified': sequences resulting in best hits with an identity below 97% or a query coverage below 90%
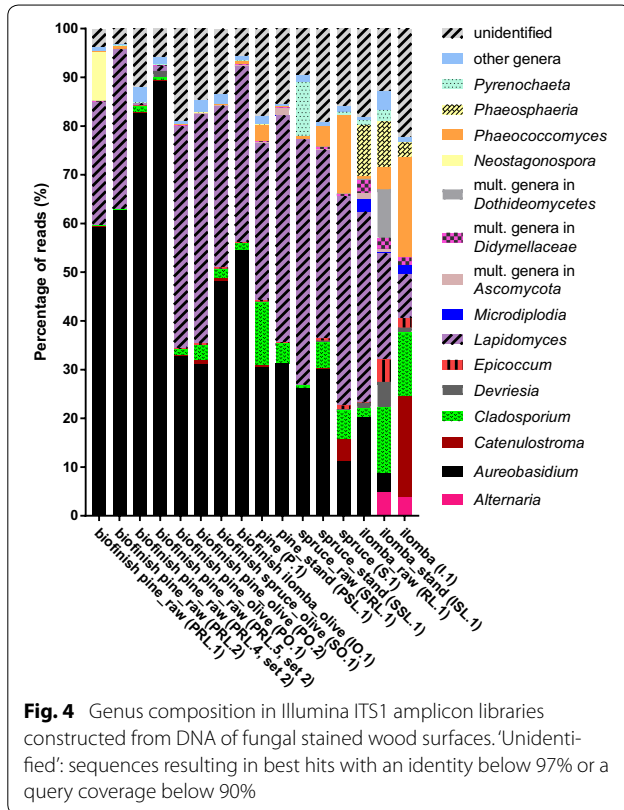
Fungal composition of Illumina ITS1 libraries

Application of the Illumina amplicon method to analyse the fungal biofinish community of the 16 selected wood samples resulted in a total of 2.17 million filtered ITS1 reads. The amount of reads passing the occurrence threshold was 2.02 million. Each wood sample had 6.5×10^4 up to 4.6×10^5 reads with a mean length of 171 nucleotides. In total 400 unique sequences were detected with no nucleotide variation among the reads of a single unique sequence (Additional file 4: Table S4). Most of these unique sequences could be identified to genus-level, but some sequences represented multiple genera or represented an unidentified genus (Fig. 4; Additional file 4: Table S4). The read percentage of this latter category varied for the biofinish samples between 3% till 22% of the total reads. Some of these sequences showed 100% similarity with the unidentified sequences obtained from the clone libraries, while none of these sequences were highly similar to the ITS sequences derived from any of the fungal isolates cultured from biofinishes. The number of different identified genera for the biofinish samples ranged from 26 to 34 and for the non-biofinish wood samples (that all contained dark mould stains but less than the biofinish criteria prescribed) from 27 to 35. The calculated Shannon's diversity indices were generally lower for the

biofinish samples (average 0.8) compared to the samples without a biofinish (average 1.4; Fig. 5).

The Illumina amplicon method revealed the presence of two predominant genera in the amplicon sequencing libraries of the eight samples that contained a biofinish: *Aureobasidium* and *Lapidomyces* (Fig. 4). Both genera were determined in the DNA extractions of all samples. In the amplicon sequence libraries of six biofinish containing samples the predominance of *Aureobasidium* was determined, including all four pine sapwood samples treated with raw linseed oil. These four samples had the highest contribution of *Aureobasidium* reads per sample (more than 56% for each sample). However, the predominance of this genus was only determined for half of the olive oil-treated samples that contained a biofinish. The other half of the samples contained *Lapidomyces* as the predominant genus. With respect to the analysed wood samples without biofinish, not one sample showed the predominance of *Aureobasidium* in the amplicon library, while most of them had *Lapidomyces* as predominantly present genus. In line with these results the percentage of *Aureobasidium* reads was higher for the biofinish samples [average 58%, standard deviation (SD) 20%] than for a non-biofinish sample (average 19%, SD 12%), despite the variation in substrates (Fig. 4). The two types of oil that were used for biofinish substrates differed in the percentages of *Aureobasidium* sequences. The biofinishes on wood treated with olive oil showed a lower percentage for *Aureobasidium* (31–55%) than the biofinishes on raw linseed oil (59–89%).

Although *Cladosporium* was detected in all amplicon sequencing libraries of the biofinish samples, the results showed a relative low contribution (average 1%, SD 1%) of this genus to the total reads of a library. The contribution of *Cladosporium* to the amplicon sequence libraries of samples without a biofinish was higher (average 7%, SD 5%). Similar results were found for the black yeast *Phaococcomyces* (biofinishes: average 0.2%, SD 0.2%; non-biofinishes: average 6%, SD 7%).

Discussion
The fungal composition of biofinishes
The fungal compositions generated with the culture-based, cloning and Illumina sequencing approach showed overlapping and partly complementary results. For example *Aureobasidium* was detected in biofinishes with all three techniques. In contrast, the genus *Superstratomyces*, detected by culturing as a predominant colony type, was absent in the data generated with the culturing-independent approaches. The monotypic genus *Lapidomyces*, represented by the rock-inhabiting species *L. hispanicus* [32], was detected in high numbers in the cloning and Illumina libraries, while it was absent
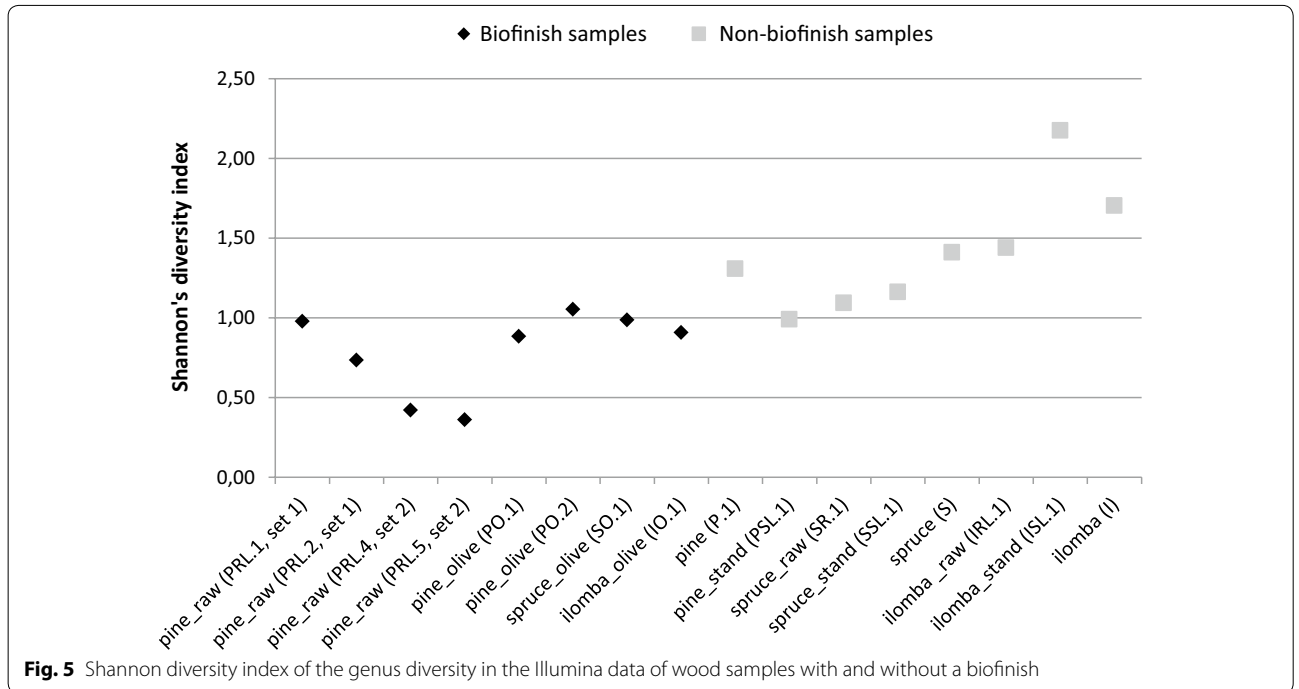
Fig. 5 Shannon diversity index of the genus diversity in the Illumina data of wood samples with and without a biofinish

in the predominantly cultured isolates. The culturing-independent approaches generated overlapping results, since *Aureobasidium* was detected with both techniques as predominant genus in the biofinishes on pine treated with raw linseed oil from set 1, with *Lapidomyces* as the secondly abundant genus. However, the results generated by these two methods of the biofinishes on pine sapwood treated with olive oil from set 1 were more complementary. Although both methods determined the presence of *Lapidomyces*, it was only predominant in the Illumina amplicon libraries, while the clone libraries had a remarkable high number of unidentified genera. In general the Illumina approach showed a larger diversity of genera (up to 30) compared to the clone approach (up to 7). Despite the complementary results of the three used techniques, each technique showed that biofinishes contain several genera including *Aureobasidium*.

The variation in the composition of the fungal community on a specific habitat with culture-dependent and culture-independent methods has been frequently reported [17, 33, 34]. Because each method is selective, variation in results seems inevitable. Firstly, only viable fungal propagules that are able to grow in specific lab conditions are identified in the culturing method, while in the case of the methods based on direct DNA extractions also non-culturable fungi can be detected. This explains why *Lapidomyces*, detected as one of the predominant genera in biofinishes, could only be found with the culture-independent techniques. One of the characteristics of this genus is the slow growth at low temperatures, such as 6 or

15 °C, and its inability to grow at 24 °C [32], while the incubation temperature used in this study was above 24 °C. Secondly, although a DNA-based method seems more complete than a culture-based, selection may happen already during the DNA extraction, since there is no equal efficiency of DNA extraction between all fungal species and/or cell structures [35, 36]. Also primers and PCR programs are known to selectively influence the profile of the microbial community [18, 19, 26]. Besides selection during DNA extraction and amplification, the possibility of variation in community composition due to low sampling numbers should be recognised in the case of the labour intensive cloning approach. Namely, the highest number of identified ITS-fragments represented not more than (71 clones/6,25 cm^2 sampling area =) 11 fungal units per cm^2 biofinish, while based on the CFU count the numbers in this study, only the culturable fungi can already be up to 4×10^5 fungal units per cm^2.

Fungal identification in this study was primarily based on ITS sequences and therefore restricted to genus level. The ITS locus, the formal fungal barcode [28, 37], is not necessarily unique for each species [38, 39] and particularly when only the ITS1 region is analysed. In case of *Aureobasidium*, the ITS sequences obtained from type strains [12] do show differences between species, but not when the ITS1 sequences are compared. Besides limited species discrimination, the taxonomic reliability of ITS sequences in a database can also be questioned [40], especially when updates of the data based on modern taxonomic revisions within, such as proposed for

Aureobasidium [41], are lacking. In order to obtain more accurate species identifications multi-locus sequencing should be applied [12].

The wood staining fungi *Aureobasidium*, *Lapidomyces* and *Superstratomyces* mainly contributed to the fungal biofinish composition

Aureobasidium was frequently detected with all tree techniques as predominant genus in natural biofinishes on oil-treated wood, which indicates the importance of this genus. Although the culturing method also revealed *Aureobasidium* among the predominant isolates of the stained wood samples that did not meet the specific biofinish criteria, the results of the Illumina approach used in this study showed that this predominance is not as easily detected as it may seem. All eight selected samples without a visible biofinish contained *Aureobasidium* in their amplicon library, but did not reveal its predominance , whereas six of the eight biofinishes did contain *Aureobasidium* as predominant detected genus. A few other fungal quantification studies of outdoor substrates, concerning the surface population of grapes, leaves of grapevines, apples and plasticised polyvinyl chloride, also indicate *Aureobasidium* to be predominantly present [42–45]. Other studied substrates, such as decomposing spruce logs [46], Scots pine needles [47], leaves/leaf litter [48], residential surfaces [49], public restroom floors [50] and the oral microbiome [51] contain *Aureobasidium,* but not as predominant fungus. In various substrates, even in fungal populations present in wood [52–55], *Aureobasidium* was not detected at all [33, 34, 56–59]. Apparently substrates and exposure conditions are selective and the surface of outdoor situated oil-treated wood samples that enables biofinish formation has favourable conditions for *Aureobasidium.*

The importance of genera other than *Aureobasidium* to outdoor biofinish formation has to be considered, especially of *Superstratomyces* and *Lapidomyces.* As expected from the previous results in the study by van Nieuwenhuijzen et al. [4], *Aureobasidium* was not always detected as the predominant genus of a biofinish. The culture-based method of the current study enabled the identification of at least five other predominant genera (Fig. 2) with *Superstratomyces* as the most commonly occurring predominant colony type isolated from biofinishes. The detection of *Superstratomyces* at outdoor exposed untreated wood samples showes that the presence of this genus is not limited to biofinishes or oil-treated substrates. In the clone libraries as well as the amplicon sequence libraries of biofinishes on pine sapwood treated with raw linseed oil, the predominance of *Aureobasidium* was determined. This is in line with the earlier finding of *Aureobasidium* as the only predominant genus in the fungal DNA extracted from a single specific sample with homogeneous dark staining published by Sailer et al. [7].

However, biofinishes on pine treated with olive oil contained more *Lapidomyces* than *Aureobasidium* sequences in their clone and Illumina amplicon libraries. The Illumina amplicon approach resulted in *Lapidomyces* as predominant genus, while sequences of an unidentified genus (top hits: *Sarcinomyces*/*Aureobasidium*/*Rhizosphaera*, maximum identity scores: 82%) were predominant in the clone libraries of these biofinishes.

In contrast to the potential importance of abundantly present genera in biofinishes, the importance of some of the detected genera can be questioned. For example, the results of the Illumina amplicon approach indicated a lower contribution of *Cladosporium* reads to the libraries of biofinish samples compared to the libraries of stained samples that did not meet the biofinish criteria(7%), but also a relative low abundance in total (1%). Also the black yeast *Phaeococcomyces* was far less represented in the libraries of samples with a biofinish than in the libraries of wood samples without a biofinish. These results are in line with the calculated Shannon's diversity indices that indicated a lower diversity in the fungal population of biofinish samples than of wood samples without a biofinish.

Competitive advantage of *Aureobasidium* in natural biofinishes

The detection of the structural presence and frequent predominance of *Aureobasidium* in biofinishes analysed in this study, confirms the important role of *Aureobasidium* in biofinish formation. As concluded from van Nieuwenhuijzen et al. [12] the development of *Aureobasidium* on oil-treated wood starts in the first few days of outdoor exposure. Since many parameters define the conditions on the surface of a material outdoors, it is likely that multiple factors are further responsible for the establishment of *Aureobasidium* in biofinishes on oil-treated wood.

In particular the production of melanin seems to be involved in the survival of *Aureobasidum* in biofinishes during outdoor exposure. Melanin plays a role in the protection against UV-light and other environmental stresses [60, 61]. Melanin production is observed in cultures of *Aureobasidium* [62, 63]. However, other genera are also known to produce melanins [60, 61]. Also the results in this study showed the presence of several other melanin producing fungi on weathered wood surfaces such as *Cladosporium* and *Exophiala.* Correlations between the pigmentation of fungal genera, the type and amount of melanins and their specific protective functions need to be investigated to understand the competitive advantage of *Aureobasidium.*

Also the presence of oil in wood, an essential ingredient for biofinish formation, is thought to play an important role at least due to its provision of carbon sources for *Aureobasidium* growth ([12], unpublished data). The results of the Illumina approach indicated that the

addition of oil to a specific wood species is related to an increase in the amount of *Aureobasidium* biomass on the wood surface. However, the use of oil or it derivatives as nutrient for growth is also known for species of other genera such as *Exophiala* (unpublished data) and *Malassezia* [64, 65]. To enhance the applicability of biofinishes in wood protection more studies are required on the availability of oil (components) at the wood surface and their role in fungal growth.

In addition, the water conditions on the surface of oil-treated wood samples that enables biofinish formation might favour growth of *Aureobasidium*. At first sight, the surfaces of wood samples treated with oils seem to be dry, unless it rains. However, dew drops have been noticed frequently on the hydrophobic surfaces during outdoor exposure. Although a few fungal species including *Aureobasidium* sp. and *Cladosporium* spp. are known to survive periods of low relative humidity [67, 68], fungal structures of *Aureobasidium* might benefit more than others from the considerably wet conditions. For example, this genus quickly formed visible colonies after inoculation of the biomass on MEA plates (water activity 0.99), while other genera needed more time to appear (data not shown). A detailed study on the water conditions on oil-treated wood samples and the impact of different water conditions on growth of wood-inhabiting genera should be performed.

Conclusions

The presence of the common wood stain fungus *Aureobasidium* in biofinishes on oil-treated wood with both culturing- and DNA-based techniques is demonstrated. Moreover, the frequent predominance of *Aureobasidium* emphasises the importance of this genus as biofinish component. The importance of other genera, such as *Lapidomyces* and *Superstratomyces*, in biofinish formation has to be considered, but is not recognized for all other detected wood inhibiting fungi.

Additional files

Additional file 1: Table S1. The total CFU concentration and the Aureobasidium CFU percentage of each wood sample (r. lins. oil = raw linseed oil, st. lins. oil = raw linseed oil, underlined numbers = estimated number, CFU count below 10).

Additional file 2: Table S2. Fungal isolates of each colony type with their culture collection numbers and GenBank accession numbers for the sequenced loci.

Additional file 3: Table S3. ITS specific clones inferred from biofinish DNA, their identification and GenBank accession number.

Additional file 4: Table S4. Illumina amplicon sequences identified by a BLAST search against the database of GenBank, Q-bank and CBS (type strains of the CBS Barcoding database were selected from data generated until September 2015). Sequences were identified to genus level when possible and marked as 'unidentified genus' when BLAST resulted in hits with an identity below 97%. The number of reads of each unique sequence is listed for each sample.

Author details
[1] Applied and Industrial Mycology, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands. [2] TNO, Microbiology and Systems Biology, Zeist, The Netherlands. [3] Present Address: Dutch DNA Biotech BV, Zeist, The Netherlands. [4] Present Address: Danone Nutricia Research, Utrecht, The Netherlands. [5] Department of Applied Physics, Section Transport in Permeable Media, University of Technology Eindhoven, Eindhoven, The Netherlands.

References
1. Gaylarde CC, Morton LHG, Loh K, Shirakawa MA. Biodeterioration of external architectural paint films: a review. Int Biodeterior Biodegrad. 2011;658:1189–98.
2. Gobakken LR, Vestol GI. Surface mould and blue stain fungi on coated Norway spruce cladding. Int Biodeterior Biodegrad. 2012;75:181–6.
3. Villa F, Stewart PS, Klapper I, Jacob JM, Cappitelli F. Subaerial biofilms on outdoor stone monuments: changing the perspective toward an ecological framework. Bioscience. 2016;664:285–94.
4. van Nieuwenhuijzen EJ, Sailer MF, Gobakken LR, Adan OCG, Punt PJ, Samson RA. Detection of outdoor mould staining as biofinish on oil treated wood. Int Biodeterior Biodegrad. 2015;105:215–27.
5. Bussjaeger S, Daisey G, Simmons R, Spindel S, Williams S. Joint Coatings Forest Prod C. Mildew and mildew control for wood surfaces. J Coat Technol. 1999;71890:67–9.
6. Gobakken LR, Westin M. Surface mould growth on five modified wood substrates coated with three different coating systems when exposed outdoors. Int Biodeterior Biodegrad. 2008;624:397–402.
7. Sailer MF, van Nieuwenhuijzen EJ, Knol W. Forming of a functional biofilm on wood surfaces. Ecol Eng. 2010;362:163–7.
8. van Nieuwenhuijzen EJ, Sailer MF, Gobakken LR, Adan OCG, Punt PJ, Samson RA, editors. Formation of biofinishes on outdoor exposed wood; the impact of wood-oil combination and geographical location Annual Meeting of the International Research Group of Wood Protection; 2016.

9. Schmidt O. Wood and tree fungi: biology, damage, protection, and use. Berlin Heidelberg: Springer; 2006.

10. Uzunovic A, Yang DQ, Gagne P, Breuil C, Bernier L, Byrne A, et al. Fungi that cause sapstain in Canadian softwoods. Can J Microbiol. 1999;4511:914–22.

11. Kelley J, Kennedy R, Kinsey G, Springle WR. Statistical methods applied to field colonisation of coatings by fungi. Part 3: ecology analysis. Surf Coat Int A Coat J. 2006;89(2):91–5.

12. van Nieuwenhuijzen EJ, Houbraken J, Meijer M, Adan OCG, Samson RA. *Aureobasidium melanogenum*: a native of dark biofinishes on oil treated wood. Antonie Van Leeuwenhoek. 2016;1095:661–83.

13. Samson RA, Houbraken J, Thrane U, Frisvad JC, Andersen B. Food and Indoor fungi. Utrecht: CBS-KNAW Fungal Biodiversity Centre; 2010.

14. Pitt JI, Hocking AD. Fungi and food spoilage. New York: Springer; 2009.

15. Pitkaranta M, Meklin T, Hyvarinen A, Paulin L, Auvinen P, Nevalainen A, et al. Analysis of fungal flora in indoor dust by ribosomal DNA sequence analysis, quantitative PCR, and culture. Appl Environ Microbiol. 2008;741:233–44.

16. Su C, Lei LP, Duan YQ, Zhang KQ, Yang JK. Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. Appl Microbiol Biotechnol. 2012;933:993–1003.

17. Rämä T. Diversity of marine wood-inhabiting fungi in North Norway. Tromsø: Tromsø University Museum; 2014.

18. Lindahl BD, Nilsson RH, Tedersoo L, Abarenkov K, Carlsen T, Kjøller R, et al. Fungal community analysis by high-throughput sequencing of amplified markers: a user's guide. New Phytol. 2013;1991:288–99.

19. Tedersoo L, Anslan S, Bahram M, Polme S, Riit T, Liiv I, et al. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. Mycokeys. 2015;10:1–43.

20. van Dijk EL, Jaszczyszyn Y, Thermes C. Library preparation methods for next-generation sequencing: tone down the bias. Exp Cell Res. 2014;3221:12–20.

21. de Hoog GS, van den Ende A. Molecular diagnostics of clinical strains of filamentous Basidiomycetes. Mycoses. 1998;415–6:183–9.

22. Masclaux F, Gueho E, Dehoog GS, Christen R. Phylogenetic relationships of human-pathogenic *Cladosporium* (*Xylohypha*) species inferred from partial LS rRNA sequences. J Med Vet Mycol. 1995;335:327–38.

23. Rehner SA, Samuels GJ. Taxonomy and phylogeny of *Gliocladium* analyzed from nuclear large subunit ribosomal DNA-sequences. Mycol Res. 1994;98:625–34.

24. Vilgalys R, Hester M. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J Bacteriol. 1990;1728:4238–46.

25. Agarwala R, Barrett T, Beck J, Benson DA, Bollin C, Bolton E, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2016;44D1:D7–19.

26. Bokulich NA, Mills DA. Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. Appl Environ Microbiol. 2013;798:2519–26.

27. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;7523:7537–41.

28. Koljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, et al. Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol. 2013;2221:5271–7.

29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;2153:403–10.

30. Bonants P, Edema M, Robert V. Q-bank, a database with information for identification of plant quarantine plant pest and diseases. EPPO Bull. 2013;432:211–5.

31. van Nieuwenhuijzen EJ, Miadlikowska JM, Houbraken JAMP, Adan OCG, Lutzoni FM, Samson RA. Wood staining fungi revealed taxonomic novelties in Pezizomycotina: new order Superstratomycetales and new species *Cyanodermella oleoligni*. Stud Mycol. 2016;85:107–24.

32. Egidi E, de Hoog GS, Isola D, Onofri S, Quaedvlieg W, de Vries M, et al. Phylogeny and taxonomy of meristematic rock-inhabiting black fungi in the *Dothideomycetes* based on multi-locus phylogenies. Fungal Divers. 2014;651:127–65.

33. Pangallo D, Buckova M, Krakova L, Puskarova A, Sakova N, Grivalsky T, et al. Biodeterioration of epoxy resin: a microbial survey through culture-independent and culture-dependent approaches. Environ Microbiol. 2015;172:462–79.

34. Stefani FOP, Bell TH, Marchand C, de la Providencia IE, El Yassimi A, St-Arnaud M, et al. Culture-dependent and-independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. PLoS ONE. 2015;10(6):e0128272.

35. Karakousis A, Tan L, Ellis D, Alexiou H, Wormald PJ. An assessment of the efficiency of fungal DNA extraction methods for maximizing the detection of medically important fungi using PCR. J Microbiol Methods. 2006;651:38–48.

36. Fredricks DN, Smith C, Meier A. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. J Clin Microbiol. 2005;4310:5122–8.

37. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci USA. 2012;10916:6241–6.

38. Geiser DM, Klich MA, Frisvad JC, Peterson SW, Varga J, Samson RA. The current status of species recognition and identification in *Aspergillus*. Stud Mycol. 2007;59:1–10.

39. Woudenberg JHC, Groenewald JZ, Binder M, Crous PW. *Alternaria* redefined. Stud Mycol. 2013;75:171–212.

40. Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, Koljalg U. Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS ONE. 2006;1(1):e59.

41. Gostincar C, Ohm RA, Kogej T, Sonjak S, Turk M, Zajc J, et al. Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. BMC Genom. 2014;15(1):1.

42. Webb JS, Nixon M, Eastwood IM, Greenhalgh M, Robson GD, Handley PS. Fungal colonization and biodeterioration of plasticized polyvinyl chloride. Appl Environ Microbiol. 2000;668:3194–200.

43. Prakitchaiwattana CJ, Fleet GH, Heard GM. Application and evaluation of denaturing gradient gel electrophoresis to analyse the yeast ecology of wine grapes. FEMS Yeast Res. 2004;48:865–77.

44. Vero S, Garmendia G, Gonzalez MB, Garat MF, Wisniewski M. *Aureobasidium pullulans* as a biocontrol agent of postharvest pathogens of apples in Uruguay. Biocontrol Sci Technol. 2009;1910:1033–49.

45. Pinto C, Pinho D, Sousa S, Pinheiro M, Egas C, Gomes AC. Unravelling the diversity of grapevine microbiome. PLoS ONE. 2014;9(1):e85622.

46. Ottosson E, Kubartova A, Edman M, Jonsson M, Lindhe A, Stenlid J, et al. Diverse ecological roles within fungal communities in decomposing logs of Picea abies. FEMS Microbiol Ecol. 2015;91(3):fiv012.

47. Millberg H, Boberg J, Stenlid J. Changes in fungal community of Scots pine (*Pinus sylvestris*) needles along a latitudinal gradient in Sweden. Fungal Ecol. 2015;17:126–39.

48. Voriskova J, Baldrian P. Fungal community on decomposing leaf litter undergoes rapid successional changes. ISME J. 2013;73:477–86.

49. Adams RI, Miletto M, Taylor JW, Bruns TD. The diversity and distribution of fungi on residential surfaces. PLoS ONE. 2013;8(11):e78866.

50. Fouquier J, Schwartz T, Kelley ST. Rapid assemblage of diverse environmental fungal communities on public restroom floors. Indoor Air. 2016. doi:10.1111/ina.12279.

51. Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, Naqvi A, et al. Characterization of the oral fungal microbiome (Mycobiome) in healthy individuals. PLoS Pathog. 2010;6(1):e1000713.

52. Ding S, Hu H, Gu JD. Fungi colonizing wood sticks of Chinese fir incubated in subtropical urban soil growing with *Ficus microcarpa* trees. Int J Environ Sci Technol. 2015;1212:3781–90.

53. Rama T, Norden J, Davey ML, Mathiassen GH, Spatafora JW, Kauserud H. Fungi ahoy! Diversity on marine wooden substrata in the high North. Fungal Ecol. 2014;8:46–58.

54. Hoppe B, Purahong W, Wubet T, Kahl T, Bauhus J, Arnstadt T, et al. Linking molecular deadwood-inhabiting fungal diversity and community dynamics to ecosystem functions and processes in Central European forests. Fungal Divers. 2016;771:367–79.

55. Rajala T, Tuomivirta T, Pennanen T, Makipaa R. Habitat models of wood-inhabiting fungi along a decay gradient of Norway spruce logs. Fungal Ecol. 2015;18:48–55.

56. Yuan ZL, Chen YC, Yang Y. Diverse non-mycorrhizal fungal endophytes inhabiting an epiphytic, medicinal orchid (*Dendrobium nobile*): estimation and characterization. World J Microbiol Biotechnol. 2009;252:295–303.

57. Redou V, Navarri M, Meslet-Cladiere L, Barbier G, Burgaud G. Species richness and adaptation of marine fungi from deep-subseafloor sediments. Appl Environ Microbiol. 2015;8110:3571–83.

58. Langarica-Fuentes A, Fox G, Robson GD. Metabarcoding analysis of home composts reveals distinctive fungal communities with a high number of unassigned sequences. Microbiol SGM. 2015;161:1921–32.

59. Zhang T, Wang NF, Zhang YQ, Liu HY, Yu LY. Diversity and distribution of fungal communities in the marine sediments of Kongsfjorden, Svalbard (High Arctic). Sci Rep. 2015;5:14524.

60. Kogej T, Stein M, Volkmann M, Gorbushina AA, Galinski EA, Gunde-Cimerman N. Osmotic adaptation of the halophilic fungus *Hortaea werneckii*: role of osmolytes and melanization. Microbiol SGM. 2007;153:4261–73.

61. Pal AK, Gajjar DU, Vasavada AR. DOPA and DHN pathway orchestrate melanin synthesis in *Aspergillus* species. Med Mycol. 2014;521:10–8.

62. Gniewosz M, Duszkiewicz-Reinhard W. Comparative studies on pullulan synthesis, melanin synthesis and morphology of white mutant *Aureobasidium pullulans* B-1 and parent strain A: p.-3. Carbohydr Polym. 2008;723:431–8.

63. Hernandez VA, Evans PD. Technical note: melanization of the wood-staining fungus *Aureobasidium pullulans* in response to UV radiation. Wood Fiber Sci. 2015;471:120–4.

64 Satow MM, Attili-Angelis D, de Hoog GS, Angelis DF, Vicente VA. Selective factors involved in oil flotation isolation of black yeasts from the environment. Stud Mycol. 2008;61:157–63.

65 Nazzaro Porro MN, Passi S, Caprilli F, Nazzaro P, Morpurgo G. Growth requirements and lipid metabolism of *Pityrosporum orbiculare*. J Invest Dermatol. 1976;663:178–82.

66 Velegraki A, Cafarchia C, Gaitanis G, Iatta R, Boekhout T. *Malassezia* infections in humans and animals: pathophysiology, detection, and treatment. PLoS Pathog. 2015;11(1):e1004523.

67 Park D. Phylloplane fungi: tolerance of hyphal tips to drying. Trans Br Mycol Soc. 1982;79AUG:174–8.

68 Segers FJJ, van Laarhoven KA, Huinink HP, Adan OCG, Wosten HAB, Dijksterhuis J. The indoor fungus cladosporium halotolerans survives humidity dynamics markedly better than aspergillus niger and penicillium rubens despite less growth at lowered steady-state water activity. Appl Environ Microbiol. 2016;8217:5089–98.

Comparing the physiochemical parameters of three celluloses reveals new insights into substrate suitability for fungal enzyme production

Lara Hassan[1†], Manfred J. Reppke[1†], Nils Thieme[1], Steffen A. Schweizer[2], Carsten W. Mueller[2] and J. Philipp Benz[1*] 🆔

Abstract

Background: The industrial applications of cellulases are mostly limited by the costs associated with their production. Optimized production pathways are therefore desirable. Based on their enzyme inducing capacity, celluloses are commonly used in fermentation media. However, the influence of their physiochemical characteristics on the production process is not well understood. In this study, we examined how physical, structural and chemical properties of celluloses influence cellulase and hemicellulase production in an industrially-optimized and a non-engineered filamentous fungus: *Trichoderma reesei* RUT-C30 and *Neurospora crassa*. The performance was evaluated by quantifying gene induction, protein secretion and enzymatic activities.

Results: Among the three investigated substrates, the powdered cellulose was found to be the most impure, and the residual hemicellulosic content was efficiently perceived by the fungi. It was furthermore found to be the least crystalline substrate and consequently was the most readily digested cellulose in vitro. In vivo however, only RUT-C30 was able to take full advantage of these factors. When comparing carbon catabolite repressed and de-repressed strains of *T. reesei* and *N. crassa*, we found that *cre1/cre-1* is at least partially responsible for this observation, but that the different wiring of the molecular signaling networks is also relevant.

Conclusions: Our findings indicate that crystallinity and hemicellulose content are major determinants of performance. Moreover, the genetic background between WT and modified strains greatly affects the ability to utilize the cellulosic substrate. By highlighting key factors to consider when choosing the optimal cellulosic product for enzyme production, this study has relevance for the optimization of a critical step in the biotechnological (hemi-) cellulase production process.

Keywords: Microcrystalline cellulose, Powdered cellulose, Cellulase production, Cellulose crystallinity, *Neurospora crassa*, *Trichoderma reesei*, RUT-C30

Background

Due to their wide applicability, the demand for cellulases and hemicellulases is constantly increasing. Currently,

these enzymes are used in the processing of food and animal feed, in the textile and laundry industries, for pulping and paper production, as well as for the biofuels industry [1]. The overall technical enzymes market is projected to reach a value of 1.27 billion USD in 2021, with the bioethanol application predicted to be the fastest-growing section [2]. The goal here is to efficiently convert sustainably produced lignocellulosic feedstocks to fermentable sugars for the production of biofuels, but also other products

of the biorefinery. Due to the high recalcitrance of cellulose, this process requires high enzyme loadings, warranting research efforts aiming to increase enzyme yields and decrease the production costs.

Cellulose is composed of unbranched chains with repeating ß-1,4-linkages of only D-glucose units. Many parallel glucan chains form tight microfibrils held together by hydrogen bonds, rendering the surface of cellulose highly hydrophobic and recalcitrant to enzymatic attack [3–6]. Traditionally, the fine structure of cellulose is described in a simplistic two-phase model, in which highly ordered regions are classified as crystalline and less well-ordered regions as amorphous [7]. Moreover, cellulose in the natural setting is embedded in a matrix of hemicelluloses and lignin, adding structural support and protection [8, 9]. The major hemicelluloses in hardwoods and grasses are xylans and mixed-linkage glucans, while (galacto)glucomannans dominate in softwoods [10–12].

Cellulases are commonly produced by fermentation of lignocellulosic substrates with microorganisms, such as bacteria or filamentous fungi. Microcrystalline celluloses (MCCs) have been used as excipients in the pharmaceutical industry for decades, but are also used as cellulase-inducing substrates due to their purity, availability and ease of use. MCCs are usually prepared by treatment of cotton linters or wood pulp with dilute mineral acid to hydrolyze and extract the amorphous regions of cellulose as well as hemicelluloses, lignin and pectin [13, 14]. The result is a partially depolymerized cellulose with a limited degree of polymerization in the form of colloidal crystallites that can aggregate and agglomerate to particle sizes of usually between 20 and 200 μM [15]. MCCs are derived from various sources, such as hardwoods and softwoods. Various products from the international market have been shown to differ in their characteristics regarding crystallinity, monosaccharide composition, and particle size [16–19]. Moreover, batch-to-batch variability has been shown to have an equally strong impact on the MCC properties [20].

The filamentous ascomycete *T. reesei* (teleomorph *Hypocrea jecorina*) has become the preferred organism for the production of cellulases [21–23], one of the best known and publicly available strains being RUT-C30 of the Rutgers lineage derived from screens for hyper-cellulase production after rounds of classical mutagenesis [24, 25].

Trichoderma reesei has also been instrumental in the elucidation of the molecular factors underlying the perception and degradation of cellulose in filamentous fungi [26]. The general principle of induction and repression governing the response is conserved as in all microorganisms, but varies in its implementation between fungi (for a review, see [27]). In *T. reesei* (and species of the genus

Aspergillus), the transcription factor (TF) XYR1/XlnR is the major regulator of the cellulolytic and hemicellulolytic response, even though recently ACE3 was described in *T. reesei* as a novel master regulator of cellulase expression and a modulator of xylan degrading enzyme expression [28]. In other fungi, such as in the genetic model system *N. crassa*, the XYR1 homologs only modulate production of cellulases and are mainly required for the induction of hemicellulases (for recent reviews, see [29, 30]). Instead, two other conserved TFs in tandem govern the response to cellulose: CLR-1 and CLR-2 [31, 32]. The function of CLR-2 does not seem to be strictly conserved in *T. reesei* [28], but is in other fungi such as in *A. nidulans* [31].

Other than the induction pathways, carbon catabolite repression (CCR), a mechanism enabling microorganisms to prefer easily metabolizable carbon sources over polymeric or recalcitrant substrates, seems strictly conserved in filamentous fungi [27]. A central mediator of CCR is the zinc-finger TF CreA/Cre1 [33–36], which acts in a double-lock mechanism on both the target genes as well as the regulatory TFs [29]. In *T. reesei* RUT-C30, a truncated version of the *cre1* gene is present [37], leading to a cellulase de-repressed phenotype [35].

The production of cellulases in filamentous fungi is furthermore dependent on the presence of specific inducer molecules. In case of cellulose, the relevant signaling molecules are short cellodextrins such as cellobiose, which are released from cellulose by the action of cellulases, or metabolic derivatives, such as sophorose [38–40].

According to the aforementioned points it is clear, therefore, that multiple factors will affect the production of cellulases in microorganisms: (1) the composition of the substrate, (2) the accessibility of the cellulose to enzymatic attack, (3) the overall enzymatic complement produced by the organism, (4) the nature and amount of inducer molecules being released, and (5) the wiring of the regulatory networks integrating the perceived signals in the respective production organism employed. MCCs as more pure substrates might appear to be less complex in their applicability than plant biomass, but their effectiveness is subject to the same combination of physical, chemical and biological factors. A huge variety of sources, production methods, as well as batch-to-batch variations [16–18, 20] makes it highly demanding for the user to choose the best substrate and warrant studies to determine the most relevant factors. Despite a plethora of studies on the characteristics of cellulose and their effects on enzyme hydrolysis, the effects of central factors, such as crystallinity and fine structure (surface area; porosity), are still unclear and partly disputed [41–51]. However, with some notable exceptions (e.g. [48]) most

of these studies used isolated enzyme systems, which is helpful to focus, but is also simplifying, since it ignores the biology of the production organism. To extend our view, we therefore analyzed both the physiochemical and molecular biological aspects of cellulase production in two filamentous fungi when grown on different cellulosic substrates. We chose three representative celluloses: a hardwood- and a softwood-derived MCC as well as a hardwood-derived powdered cellulose, and tested their effectiveness as cellulase-inducing substrates on the hypercellulolytic *T. reesei* strain RUT-C30 and the laboratory model strain *N. crassa*. The physiochemical analyses of the substrates were done at several structural levels (acc. to [45]): fiber (surface area and morphology), fibril (composition, particle size), and microfibril (crystallinity). To assess the fungal performance, cellulase productivity as well as the molecular response were recorded.

Results

Substrate characteristics: surface area and morphology

For this study, three different cellulose products were chosen as cellulase-inducing growth substrates: a hardwood-derived MCC (Emcocel HD90), a softwood-derived MCC (Avicel PH-101) and a hardwood-derived purified cellulose (Alphacel) (see Methods; Table 2). The celluloses were initially observed by scanning electron microscopy (SEM) to visualize macromolecular substrate characteristics. In line with the manufacturer's specifications, Emcocel contained the largest particles in comparison to the MCC gold standard Avicel as well as the purified cellulose Alphacel (Fig. 1a–c). While Alphacel had the most fibrous appearance (Fig. 1b), Emcocel consisted mostly of particle agglomerates that could be broken up by additional ball-milling (Fig. 1c, d).

N_2-BET measurements showed an inverse correlation between the specific surface area of the celluloses and the average particle size. The specific surface area of Emcocel (0.80 m^2/g) was only about 2/3 the area of Avicel (1.28 m^2/g) and less than half the area of Alphacel (1.64 m^2/g). After additional ball-milling, however, the surface area of Emcocel doubled and was comparable to the other substrates (1.58 m^2/g).

Determination of hemicellulose content of the celluloses

To determine the purity of the different MCCs regarding hemicellulose contaminations, we performed a compositional analysis after total acid hydrolysis. Bacterial cellulose was used as a hemicellulose contamination-free standard for comparison. For Avicel, Alphacel and the bacterial cellulose, a 1 h swelling time in 72% H_2SO_4 was sufficient to achieve an almost complete hydrolysis. For Emcocel however, an undissolved residual mass remained after the hydrolysis, indicating that the process had been
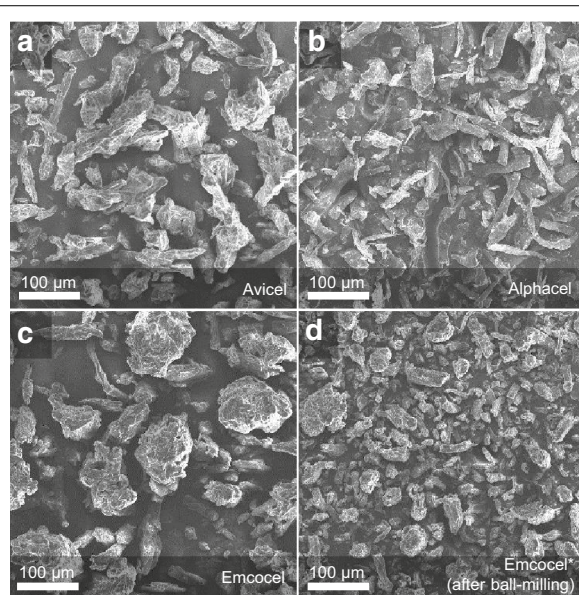
Fig. 1 Scanning electron micrographs of cellulose substrates obtained at 8.0 kV accelerating voltage and ×100 magnification (Jeol JSM-IT100). Images show a representative picture for each substrate out of several technical replicates

incomplete (data not shown). For that reason, the hydrolysis of Emcocel was repeated using longer swelling times (4 h) to give the sulfuric acid more time to completely react with the MCC.

As expected, the bacterial cellulose showed no trace of other monosaccharides other than glucose (Table 1). Alphacel on the other hand, was the most unpure cellulose with a particularly high content of xylan. The high xylan:mannan ratio is indicative of its source material being hardwood pulp (acc. to [16]). Avicel as a softwood-derived MCC presented a much more moderate xylan:mannan ratio, but still contained considerable amounts of both hemicelluloses. Emcocel proved to be the least hemicellulose-contaminated cellulose in our analysis with xylan and mannan contents of less than 1% each. This amount was more or less constant

Table 1 Results of sugar analysis of the celluloses after sulfuric acid hydrolysis (in %)

	Avicel[a]	Emcocel[b]	Alphacel[a]	bacterial cellulose[a]
D-glucan	93 ± 5.5	92 ± 0.4	84 ± 5.8	97 ± 4.2
D-xylan	3.5 ± 0.4	0.8 ± 0.08	14.7 ± 1.1	ND
D-mannan	1.8 ± 0.1	0.3 ± 0.01	1.3 ± 0.3	ND

ND none determined

[a] 1 h swelling

[b] 4 h swelling

at all different swelling times tested (not shown), while the amount of detected glucose increased considerably at 4 h, suggesting that the cellulosic fraction of Emcocel is extremely densely packed and recalcitrant to the hydrolysis. Even after 4 h of swelling, ~ 6–7% of the mass remained unaccounted for. We did not detect elevated amounts of lignin or extractables in Emcocel however (data not shown), indicating that the residual mass is mainly undissolved cellulose.

Cellulose crystallinity

Crystallinity has been widely used to describe celluloses and woods, since it is a good measure of the inherent degree of structural order and thus may have a major influence on the recalcitrance of the substrate to biochemical attack (e.g. [47, 52–54]). Due to the differences observed in bioavailability of the celluloses, we decided to measure also the crystallinity of all three substrates. To this end, samples of the celluloses were analyzed by solid-state ^{13}C nuclear magnetic resonance (NMR) and the crystallinity index (CrI) calculated by the NMR C4 peak separation method [47, 52, 55].

The NMR spectra displayed noticeable differences in height of the C4 peaks between the samples (Fig. 2). Alphacel showed the lowest crystallinity with 33.1% compared to Avicel with 54.4% and Emcocel with 56.7%. The calculated CrIs therefore allowed for a clear discrimination of the MCCs and the powdered cellulose product Alphacel. Additionally, the results demonstrate that Emcocel was the most crystalline cellulose product in our experiments.

In vitro digestibility of the cellulose substrates

Since the fungal cellulase induction will be highly dependent on the enzymatic digestibility for the liberation of inducer molecules, we next wanted to test this in an in vitro assay. To this end, *N. crassa* cellulases were incubated with the cellulose substrates. The liberated sugars were analyzed by HPAEC-PAD, and the residual cellulose harvested for SEM analysis (Fig. 3).

The chromatograms indicated that mainly glucose and higher cellodextrins accumulated in the assay supernatants after extensive digestion (8 h; Fig. 3a). The quantified amounts of glucose, xylose and mannose at more initial time points (1 h; Fig. 3a; inset) showed that Alphacel was the most readily digested substrate. Similarly, Emcocel was the least digested substrate again, corroborating that it is most recalcitrant of all three substrates to enzymatic or chemical attack.

When observed by SEM, all three celluloses showed a more particulate appearance after 24 h of enzymatic digestion. This was most prominent for the MCCs, which completely disintegrated into individual fibers
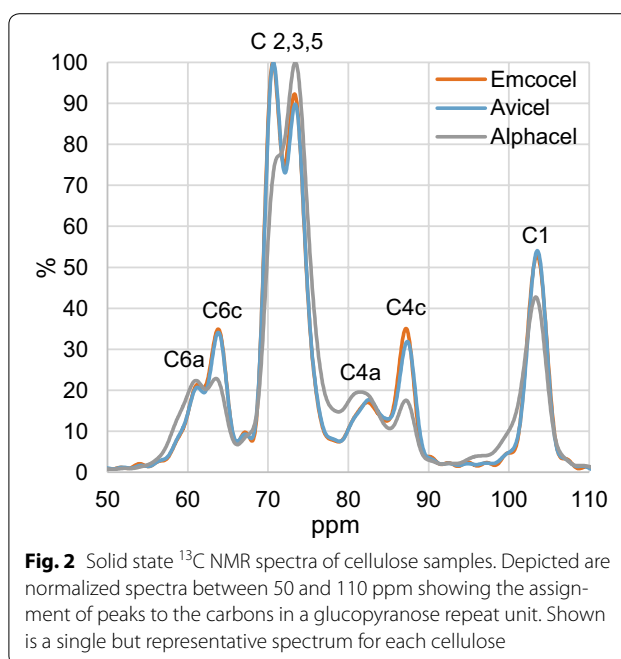
Fig. 2 Solid state ^{13}C NMR spectra of cellulose samples. Depicted are normalized spectra between 50 and 110 ppm showing the assignment of peaks to the carbons in a glucopyranose repeat unit. Shown is a single but representative spectrum for each cellulose

(Fig. 3b–g). At higher magnifications, it became evident that the surfaces of all digested substrates appeared smoother than the controls (Fig. 3h–m). This might indicate that the rough top layers were formed by more amorphous cellulose and/or hemicelluloses. It furthermore suggests that the flat layers exposed after the digest might represent more recalcitrant cellulose-rich areas. Particularly the parallel fibrillar structures visible on the surface of some fibers of Emcocel after the digest (Fig. 3i; arrow) seem to represent relatively pure, ordered cellulose fibrils, leaving little contact surface for enzymes to attack.

Determination of the potential to induce lignocellulolytic gene expression

With a molecular approach, we tested the ability of the cellulose substrates to induce the major cellulolytic and hemicellulolytic pathways on the level of gene expression. Moreover, this analysis was supposed to provide indirect insight into the early bioavailability of inducer molecules, and therefore to what extent fungi will be able to actually perceive the measured differences in the substrate composition. Since the molecular response pathways to cellulose and hemicellulose are more separated in *N. crassa* than in *T. reesei* (see background and reviewed in [27]), we only used *N. crassa* for these assays. Based on a survey of published transcriptomics analyses [31, 38, 56–59], we chose three genes that served as proxies for the induction of the cellulolytic, xylanolytic and mannanolytic pathways in *N. crassa*, since they had been shown
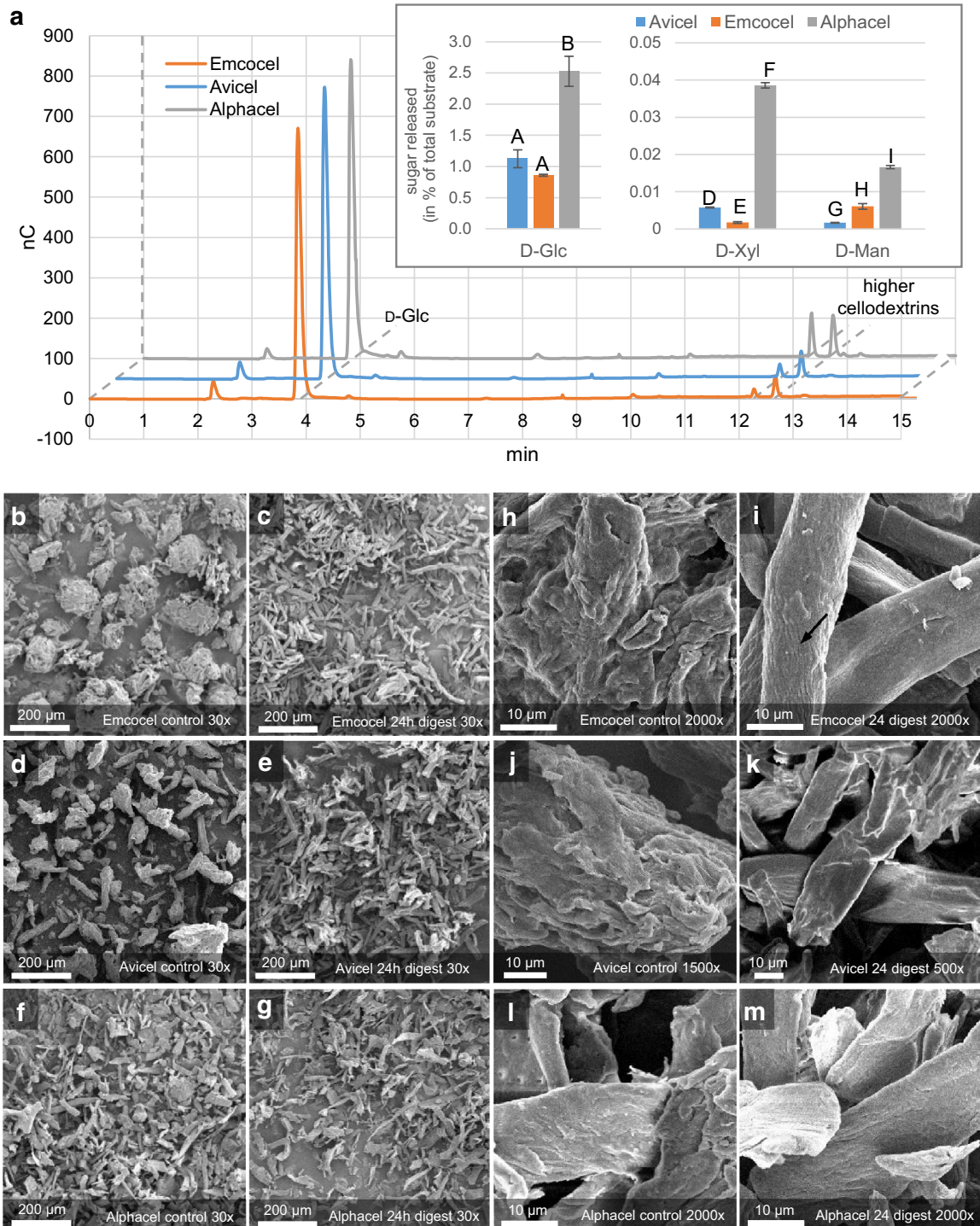
Fig. 3 Enzymatic digestion of the cellulose substrates in vitro. All celluloses were digested by a *N. crassa*-derived cellulase cocktail (filtered culture supernatant after 5 days growth on Avicel) for a total of 24 h. **a** Shown are representative HPAEC-PAD chromatograms of the reaction supernatants after 8 h as well as the quantification results for monosaccharides at an initial time point (1 h; inset in a). The peaks of D-glucose (D-Glc) and higher cellodextrins (not quantified) are indicated. Note: in this run (CarboPac® PA200 column), the other monosaccharides will also migrate at the same speed as D-Glc, but the amounts are substantially lower. The quantifications represent means of triplicate reactions. The error bars represent standard deviations. Letters indicate data groups that are significantly different (one-way ANOVA, *p*-values < 0.05 were considered significant). **b**–**m** Representative scanning electron micrographs out of technical triplicates for each cellulose substrate obtained at 8.0 kV accelerating voltage and × 30–× 2000 magnification (Jeol JSM-IT100)

to be robustly induced by their respective substrates and serve a specific function in those pathways: the cellobionic acid transporter gene *cbt-1/clp-1* (NCU05853), the D-xylulose kinase encoding gene NCU11353, and the putative acetylmannan esterase encoding gene *ce16-1* (NCU09416).

The initial cellulolytic response (as measured by *cbt-1/clp-1*) to Avicel, Emcocel and Alphacel was similar (Fig. 4a). Emcocel, however, showed a tendency to respond weaker. Even though it was found to have the lowest cellulose content of all tested substrates, the response to Alphacel tended to be strongest, indicating a better bioavailability of the cellulose microfibrils.

The xylanolytic response (as measured by the D-xylulose kinase gene expression) was clearly strongest in the case of Alphacel, which also had the highest content of xylan, and was found to be less intense in the MCC substrates (Fig. 4b). Interestingly however, *N. crassa*'s response to the xylan impurities in Avicel and Emcocel was not necessarily proportional to the overall content. Emcocel induced the xylanolytic pathway more strongly, even though the xylan content was lower than in Avicel, again indicating that the bioavailability is not directly proportional to the overall content.

In case of the mannanolytic pathway, the responses to the three substrates were very distinct (Fig. 4c). Avicel induced the acetylmannan esterase gene *ce16-1* > 7-fold stronger than Emcocel, which roughly reflects the difference in mannan content and might be expectable, since Avicel is derived from softwoods and Emcocel from hardwoods, in which xylans are the dominating hemicellulose. The fact that the strongest response was again detected on Alphacel, although the overall mannan content was found to be lower than in Avicel might have two

reasons: (1) the bioavailability of the plant cell wall sugars in the non-microcrystalline substrate is generally higher, and (2) there is some evidence that the cellulolytic and mannanolytic pathways cross-react in *N. crassa* (as well as in *Aspergillus oryzae*) [32, 60]. The strong cellulolytic response to Alphacel might therefore also co-induce the mannanolytic pathway more strongly.

Cellulase production in *N. crassa* and *T. reesei* RUT-C30

The effectiveness of the three celluloses as substrates for cellulase production in filamentous fungi was tested in 100 ml shake flask cultures. We used both the industrially optimized hypercellulolytic *T. reesei* strain RUT-C30, as well as the genetic model system *N. crassa*. The performance was evaluated after 3 and 6 days of growth by three main analyses of the culture supernatants: total secreted protein concentration, endo-glucanase activity and endo-xylanase activity (Fig. 5). Total and not specific activities (normalized to fungal biomass) are presented, since we aimed to study overall yields on each carbon source. These are therefore representative for the combined effects of bioavailability differences on induction, degradation, metabolism, secretion and growth.

Overall, the data showed that both the softwood-MCC (Avicel) as well as the powdered cellulose (Alphacel) outperformed the hardwood-MCC (Emcocel HD-90) in *N. crassa*, where total secreted protein as well as endo-glucanase and endo-xylanase activities were consistently lowest for Emcocel. While secreted protein, endo-glucanase and endo-xylanase activities were comparable on Avicel and Emcocel (Fig. 5), Alphacel was found to induce cellulases and hemicellulases more strongly in *T. reesei* (Fig. 5b, c). Moreover, protein secretion by *T. reesei* RUT-C30 had a longer lag phase (Fig. 5a; compare day 3 vs. day 6 data).

Fig. 4 Gene expression induction of selected genes used as proxies for the fungal cellulolytic and hemicellulolytic response. Sucrose pre-grown *N. crassa* cultures were exposed to the celluloses or no carbon source for 4 h before RNA was harvested. Gene induction was measured by RT-qPCR. Shown is the mean fold-induction over the no carbon (No C) starvation condition derived from biological and technical triplicates. Error bars denote standard deviation. Letters indicate data groups that are significantly different (one-way ANOVA, *p*-values < 0.05 were considered significant)

Fig. 5 Cellulase and hemicellulase production by *N. crassa* and *T. reesei* RUT-C30 on the cellulose substrates. Performance was measured by analysis of culture supernatant aliquots taken after 3 and 6 days, respectively. Secreted protein was measured by Bradford assay, endoglucanase activity by Azo-CMC assay and endo-xylanase activity by Azo-Xylanase assay as described in Methods. Values are the mean of biological triplicates. Error bars show standard deviation. Letters indicate data groups that are significantly different (separately for both fungi and sample days; one-way ANOVA, *p*-values < 0.05 were considered significant)

Since it seemed likely that the poor performance of Emcocel was at least partly due to the much smaller surface area of the substrate (see above), we also tested the performance of the ball-milled Emcocel (Emcocel*; *N. crassa* only) with an about doubled surface area

compared to the original (Additional file 1: Fig. S1). Surprisingly however, protein production was almost indistinguishable from the unmilled Emcocel, indicating that the specific surface area was not limiting for the performance in the fermentations.

WT versus industrially optimized strains
Combining the physiochemical properties of the celluloses with the fungal performance data, it appears like hypersecreting systems such as the industrially utilized *T. reesei* RUT-C30 strain are less sensitive to small increases at high crystallinities than non-engineered organisms such as *N. crassa* and are able to take full advantage of the much less crystalline PCs such as Alphacel. We hypothesized that differences in the genetic configuration in these fungi may play a decisive role to this end. Since one of the major modifications from WT to industrial strain is the reduction or removal of CCR, we tested the effect of this by comparing protein production on the three celluloses of the de-repressed *T. reesei* RUT-C30 strain to its WT strain QM6a as well as by comparing the *N. crassa* WT to a deletion strain of *cre-1*, encoding the major TF mediating CCR (Fig. 6). To limit the effect of differences in germination speed between the genotypes as much as possible [22, 61], we modified our experimental setup for these assays and shifted identical amounts of fungal biomass to cellulose cultures after pre-growth on sucrose/glucose (see Material and Methods). Since the starting material was mycelia in this experiment, we shortened the incubation time to two and three days for *N. crassa* and *T. reesei,* respectively.

Two observations could be made: First, the carbon catabolite de-repressed strains (*T. reesei* RUT-C30 with a truncated *cre1* gene [37] and *N. crassa* with a *cre-1* deletion) displayed strongly elevated protein levels in comparison to the respective WT strains (Fig. 6) which is in line to what was previously shown in both fungi [36, 62, 63], indicating that CCR has a repressing effect even on highly recalcitrant substrates with low glucose fluxes such as Emcocel (even though the effect there was smallest). And second, also the *T. reesei* WT strain (QM6a) was not able to take full advantage of the low-crystallinity substrates such as Alphacel (Fig. 6b, light blue), just as we had found for *N. crassa* (*cf.* Figure 5). However, while protein production of *T. reesei* RUT-C30 was roughly inversely proportional to the substrate crystallinities again (Fig. 6b, dark blue), the deletion of *cre-1* in *N. crassa* was not sufficient to phenocopy this, and protein production on the PC Alphacel was only marginally stronger than on the MCC Avicel (Fig. 6a, darkorange), mainly due to stronger xylanase expression (Fig. 7b) while cellulase expression was similar on the two aforementioned substrates. These observations indicate that other genetic

Fig. 6 Total protein secreted by *N. crassa* (WT and Δ*cre-1*) and *T. reesei* (QM6a and RUT-C30) on the cellulose substrates. Performance was measured by analysis of culture supernatant aliquots taken after 2 and 3 days, respectively. Protein concentration was measured by Bradford assay as described in Methods. Values are the mean of biological triplicates. Error bars show standard deviation. Letters indicate data groups that are significantly different (separately for each strain; one-way ANOVA, *p*-values < 0.05 were considered significant)

determinants are present in addition to the absence of a fully functional CRE-1 that allows *T. reesei* RUT-C30 to utilize the PC Alphacel better than the other substrates.

Discussion

In our study, we compared three commercial cellulose powders that might be used as substrates for fungal fermentations regarding their effectiveness as inducers for the production of cellulases and hemicellulases. We chose representative substrates that differed either in their source material (i.e. softwood vs. hardwood) or their production process (i.e. hydrolytically degraded MCC vs. mechanically processed powdered cellulose). Analytical comparisons have shown that powdered celluloses have an overall higher hemicellulose content, lower crystallinity and a broader molecular weight distribution than MCCs [16]. The nature of the source material, on the other hand, can influence properties such as the hemicellulose composition [10, 12, 64]. Depending on the application, one or more of these properties become more relevant than others (e.g. [15, 16]). Consequently, all of these properties might have a strong influence on

the performance of the cellulose products in cellulase fermentations. While these points are not well understood and were to be tested here, batch-wise variations were not considered since all chosen substrates were used from a single batch.

The hemicellulosic content of the cellulose products was found to be relevant in the context of the cellulase fermentations. The fungi are able to perceive their presence early on, as demonstrated by their effect on gene expression in *N. crassa* after only 4 h of induction. In agreement and extension of what was proposed by Baehr et al. [16], even the relatively "pure" MCCs can thus be considered to represent heterogeneous systems in which the non-cellulosic parts affect both structure (and thus enzymatic kinetics/accessibility) as well as the fungal response.

The powdered cellulose (Alphacel) was found to be the most readily digested substrate. In agreement with earlier studies, the analytical results confirmed that the crystallinity of the powdered cellulose was substantially lower than that of the MCCs, and that it was furthermore much less pure, with a hemicellulose content of ~ 15% [16]. Also among the MCCs, the softwood-derived product (Avicel) outperformed the hardwood-derived one (Emcocel) when *N. crassa* was used and was found to have a lower crystallinity as well as a higher hemicellulose content. These two factors therefore stand out and appear to be dominant for the effectiveness in our study. Crystallinity and hemicellulose-content might also be linked, since the non-removed hemicelluloses are likely tightly associated with the cellulose by intercalation between the cellulose strands [65], thereby lowering the overall crystallinity.

Intriguingly, enzyme production was inversely proportional to the differences in substrate crystallinity for *T. reesei* RUT-C30, but not for *N. crassa*, where production on Avicel and Alphacel were similar, despite this difference. We hypothesized that in absence of a fully functional CRE-1/Cre1 [24], the higher flux of signaling molecules derived from the more readily digested Alphacel would be directly turned into stronger expression. Indeed, the *T. reesei* WT strain QM6a (with functional CCR) was not able to utilize Alphacel equally well as inducing substrate. However, only xylanase production was found to benefit substantially by deletion of the CCR-mediating TF *cre-1* in *N. crassa*, indicating that particularly the XLR-1-dependent hemicellulase production is carbon catabolite repressed in the WT. In the de-repressed mutants, the effect of higher inducing sugar concentrations is then dependent on the regulatory networks present in the fungi [27]. In *T. reesei*, the dominant XYR1 system regulates both cellulase and hemicellulase induction, which might explain why both cellulase and

Fig. 7 Cellulase and hemicellulase production by *N. crassa* (WT and Δ*cre-1*) and *T. reesei* (QM6a and RUT-C30) on the cellulose substrates. Performance was measured by analysis of culture supernatant aliquots taken after 2 and 3 days, respectively. Endo-glucanase activity was measured by Azo-CMC assay and endo-xylanase activity by Azo-Xylanase assay as described in Methods. Values are the mean of biological triplicates. Error bars show standard deviation. Letters indicate data groups that are significantly different (separately for each strain; one-way ANOVA, *p*-values < 0.05 were considered significant). ND indicates that no values could be determined

Table 2 Major physical characteristics of the used substrates (manufacturer's information)

	Avicel™ PH-101	EMCOCEL® HD90	Alphacel
Brand	Fluka	JRS Pharma GmbH & Co. KG	ICN Biomedicals, Inc.
Type	Softwood MCC	Hardwood MCC	Hardwood PC
Average particle size (μm)	~ 50	124*	35**
Bulk density (g/cm³)	0.26–0.31	0.41	0.44–0.50

* By Malvern d50

** Avg. fiber length

xylanase production are benefiting from the availability of abundant inducing molecules on Alphacel. In *N. crassa* on the other hand, these pathways are much less linked and their regulation is decoupled in a way that CLR-1 and CLR-2 are responsible for the major cellulolytic response while the ortholog of XYR1, XLR-1, is responsible for the xylanolytic response and only modulates the expression of some genes in the cellulose degradation pathway (cf. [31, 58, 66] and reviewed in: [27]). The cross-induction of cellulases by xylan is therefore comparably weak. It is thus feasible that in *N. crassa* the high amount of available xylan in Alphacel will only hyperinduce the XLR-1-dependent hemicellulases, but the inducing effect of the low-crystallinity cellulose is compensated by the lower overall cellulose abundance in Alphacel.

It was also interesting to note that the slightly higher crystallinity of Emcocel compared to Avicel does not seem to allow continuous maximal production capacity despite the similarly strong gene induction measured at 4 h. This seems to indicate that even small differences in measured crystallinity in celluloses manifest themselves in substantial performance differences over time due to a positive feedback loop between substrate bioavailability, metabolism and fungal growth—a fact that has to be considered when choosing a substrate for enzyme fermentations.

Besides the cellular systems, the activity of the enzymes on the substrate for the release of inducer molecules and products will be a bottleneck step. Factors that have been described to affect enzyme activity include: reversible and irreversible adsorption (to cellulose or lignin, respectively), end product inhibition, synergism and others (e.g. [46, 67, 68]). Many studies have looked at the interaction of cellulose structure and enzymatic hydrolysis, and particularly structural characteristics such as crystallinity and surface area have been shown to be determining factors (e.g. [41–46, 48, 49, 51]). It has been pointed out, however, that the effects of crystallinity and surface area are difficult to be evaluated individually, since they are often highly interrelated (at least when celluloses are mechanically (pre-) treated) [45]. In our experiments, surface area did not appear to be limiting, since the performance of Emcocel could not be improved by further ball-milling. This observation is in line with other studies that did not detect any impact of particle size on the rate and extent of sugar release or enzyme binding [42, 45, 69]. However, depending on the fermentation condition and molecular crowding, this might change [68].

Overall, crystallinity seemed to be a major determinant of the effectivity of the celluloses in our study. Digestibility was found to be lower for more crystalline substrates in almost all assays (particularly visible for enzyme production in *T. reesei*). Many previous studies have found that crystallinity critically impacts the enzymatic cellulose hydrolysis (e.g. [41, 43, 44, 49, 51]). We are aware of only one study that was not able to support this finding [50]. Nevertheless, Park et al. cautioned to correlate small differences in CrI, as found between Avicel and Emcocel, with changes in cellulose digestibility, since many other factors might play a role as well, such as for example the distribution of crystalline, para-crystalline and "amorphous" regions in the particles and in relation to the surface area [47]. Still, the resistance of Emcocel towards acid hydrolysis, which was considerably stronger than found for Avicel, confirmed that this product is substantially more recalcitrant and suggests that small differences in CrI might have a strong effect at high crystallinities. The CrI of MCCs might therefore not necessarily correlate linearly with digestibility as previously shown [42] and other variables need to be considered. More studies are needed to determine this in more detail.

Conclusions

We found that both the inherent cellulose crystallinity and the hemicellulose content are major factors determining the suitability of celluloses as substrates for the expression of cellulases in filamentous fungi as production hosts. Crystallinity restricts the enzymatic digestibility and thus also the release of inducer molecules. Moreover, even in MCCs, that are considered relatively "pure", the residual hemicellulose content is efficiently perceived by the fungi and—depending on the molecular signaling pathways—translated into a corresponding cellulolytic and hemicellulolytic response, which is not necessarily analogous to the overall composition, indicating differences in bioavailability which cannot be resolved analytically.

The cellulolytic response profiles of *N. crassa* and *T. reesei* were similar, but showed also important differences. For example, only *T. reesei* RUT-C30 was able to take full advantage of the most amorphous cellulose substrate. This effect was not solely due to carbon catabolite de-repression, but is dependent on the species-specific regulatory network. The application of *N. crassa* as an indicator of the activated response pathways therefore seems promising, due to a rapid response time, and since the cellulolytic and hemicellulolytic pathways show a much more limited cross-talk than in *T. reesei* and can thus be individually evaluated.

In conclusion, the presented results provide a set of criteria that can theoretically be applied to broader ranges of substrates and thereby aid in the rational decision for cellulose substrates to be used in enzyme fermentations. Understanding the factors for the performance of cellulose substrates will help to optimize the manufacturing process of lignocellulolytic enzymes from fungi for a number of biotechnological applications, such as for food, feed, textile and biorefinery.

Methods

Substrates

Each used substrate is from a single batch preparation. The major physical characteristics of these substrates as indicated by the manufacturers are summarized in Table 2.

Strains and growth conditions

N. crassa wild type (FGSC #2489) and the Δ*cre-1* strain (kind gift of N. L. Glass, UC Berkeley, USA) were grown on 2% sucrose Vogel's minimal medium slants in the dark at 30 °C for 2 days, and transferred to constant light conditions at 25 °C for conidiation. *T. reesei* QM6a and RUT-C30 strains (kind gift of M. Schmoll, AIT, Austria) were propagated on malt extract agar plates in the dark at 28 °C and then switched to 25 °C in constant light for conidiation. The first set of growth experiments (Fig. 5) was performed in flasks containing 100 ml 1% (w/v) cellulose with 1× Vogel's (at 25 °C and in constant light) or 1× Mandels-Andreotti medium (at 30 °C and in constant light) and at 200 rpm for *N. crassa* and *T. reesei,* respectively [70, 71]. The second set of growth experiments (Figs. 6, 7) was carried out through a two-step cultivation procedure. *N. crassa* and *T. reesei* strains were first grown in 50 ml 2% (w/v) sucrose with 1× Vogel's for 20 h (at 25 °C and in constant light) or 50 ml 2% (w/v) glucose with 1× Mandels-Andreotti medium for 3 days (at 30 °C and in constant light), respectively. The cultures were collected through vacuum filtration and 0.5 g of *N. crassa* mycelia or 0.42 g of *T. reesei* mycelia were added to 100 ml 1% (w/v) cellulose with 1× Vogel's or 1× Mandels-Andreotti medium and the cultures were incubated at 25 °C and 200 rpm for 2 days (*N. crassa*) or at 30 °C and 200 rpm for 3 days (*T. reesei*), respectively.

For inoculation, generally a respective volume of conidial suspension was added after optical density measurements in order to achieve a starting concentration of 10^6 conidia/ml. All assays were done with biological triplicates for each strain per each condition.

Compositional analyses

Initially, the dried cellulose samples (~ 5 mg) were swollen in 50 μl of 72% H_2SO_4 for 1 h (Avicel, Alphacel, bacterial cellulose) or 4 h (ball-milled Emcocel) at RT. After addition of 1.45 ml water, the material was autoclaved for 1 h at 121 °C. Potential residual solids were subsequently removed by centrifugation, and the supernatant analyzed after appropriate dilution using an ICS-3000 liquid chromatography system (Dionex, Thermo Scientific) equipped with a pulsed amperometric detector. The columns used were a CarboPac® PA20 3 × 150 mm (Dionex, Thermo Scientific) and a CarboPac® PA200 3 × 250 mm (Dionex, Thermo Scientific), with column temperature

of 30 °C and a flow rate of 0.4 mL/min. Mobile phase for detection of monosaccharides (PA20) was 5 mM NaOH isocratic for 15 min. For detection of cellodextrins (PA200), mobile phases were 0.1 M NaOH (A) and 0.1 M NaOH/1 M NaAc (B). Gradient used was: 2 min 0% B, 10 min 0–10% B and 3 min 0% B. The analyses were done with technical triplicates for each substrate.

Solid state nuclear magnetic resonance spectroscopy

Solid-state ^{13}C NMR spectra were obtained on a Bruker Avance™ III 200 spectrometer (Bruker BioSpin GmbH, Karlsruhe, Germany). Cross-polarization magic angle spinning (CPMAS) was applied with a ^{13}C-resonance frequency of 50.32 MHz and a spinning speed of 5 kHz. Contact time was 1 ms and recycle delay was 2 s. Approximately 5000 scans were accumulated and no line broadening was applied. For calibration of the ^{13}C chemical shifts, tetramethylsilane was used and set to 0 ppm. Spectral analysis were performed using the spectrometer software. The crystallinity index (CrI) was then calculated by the NMR C4 peak separation method, that assigns peaks at about 87 and 82 ppm in the NMR spectra to the C4 carbons in ordered cellulose structures ("crystalline"; C4c) and non-crystalline domains ("amorphous"; C4a), respectively [47, 52, 55].

We would like to note that the CrI might have been underestimated for Alphacel, since we used a simple drying process instead of solvent-exchange drying, which might have given numbers more representative of the swollen state in the liquid broth. However, the measured differences have found to be less of an issue at high CrIs [41]. The CrIs of the MCCs were thus probably less affected, and that of Avicel was well in line with previously reported values [58].

Surface area measurements

The specific surface area of the substrates was determined by multi-point BET [72] with an Autosorb-1 analyzer (Quantachrome, Syosset, USA) using nitrogen gas as adsorbate at 77 K. The samples were outgassed before analysis in vacuum under helium flow at 60 °C for 12 h.

Scanning electron microscopy

All substrates were dried over immobilized on metal stubs using a double sided sticky tape and then sputter coated with gold. The scanning electron micrographs were taken with a JSM-IT100 (JEOL, Freising, Germany) at 8 kV accelerating voltage. The substrates were visualized in triplicates.

Enzymatic assays

Azo-CMCase and Azo-xylanase activity assays were carried out according to the protocols of the manufacturer

Table 3 Sequences of primers used for RT-qPCR

Gene	Forward primer	Reverse primer
actin (NCU04173)	CATCGACAATGGTTCGGGTATGTG	CCCATACCGATCATGATACCATGATG
cbt-1/clp-1 (NCU05853)	CGCCCTGACCTACACCTAC	GGCCAAACGACCAAAGAGC
D-xylulose kinase (NCU11353)	GGCACATCATCGCTTCACTG	CAAGGGAGATGCGCGAGG
ce16-1 (NCU09416)	GGTGCTGCTCCCATCTACTAC	GTACTTGGCGATGGCGAC

(Megazyme, Ireland) (S-ACMC and S-AXBL), slightly modified, since the reaction mixture was reduced to a quarter of the original volume. Assays were done with biological triplicates for each substrate per each strain.

RNA-extraction and RT-qPCR

Quantification of gene expression was done by quantitative real-time PCR (RT-qPCR) performed on RNA samples that were harvested after a 4 h induction phase on the respective celluloses as described in Benz et al. [49]. The RNA extraction was performed according to the TRIzol Reagent protocol (Fisher Scientific, Schwerte, Germany). RNA was then treated with DNase I (RNase-Free) according to manufacturer's recommendations (New England Biolabs, Frankfurt am Main, Germany) and subsequently cleaned up with the GeneJET RNA Purification Kit (Fisher Scientific, Schwerte, Germany). A 96-well plate reader (Infinite 200 PRO, Tecan) was used to check RNA purity and concentration. cDNA was obtained following instructions of the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems; Fisher Scientific, Schwerte, Germany). Finally, RT-qPCR was performed with the sensiFAST SYBR No-ROX Kit (Bioline, Luckenwalde, Germany) on a Mastercycler ep realplex^2 (Eppendorf, Wesseling-Berzdorf, Germany) and analyzed using the realplex 2.2 software. Actin gene (NCU04173) was used as a reference gene. Primers used are shown in Table 3. Expression analyses were done with biological and technical triplicates for each condition.

Statistical analyses

Statistical analyses were done by applying analysis of variance (ANOVA) followed by a Tukey test using the statistical computing software R [73].

Authors' contributions

JPB conceived the study, designed and supervised it. LH, MR, NT, SS, and CWM performed the experiments and acquired the data. LH, MR, NT, CWM and JPB analyzed and interpreted the data. JPB drafted the manuscript, which was critically revised by LH, MR, NT, SS, and CWM. All authors read and approved the final manuscript.

Author details

[1] HFM, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany. [2] Chair of Soil Science, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany.

References

1. Kuhad RC, Gupta R, Singh A. Microbial cellulases and their industrial applications. Enzyme Res. 2011;2011:280696.
2. Report. Technical enzymes market by type (cellulases, amylases, proteases, lipases, other enzymes), application (bioethanol, paper and pulp, textile and leather, starch processing, other applications), and by region—global forecasts to 2021. Research and markets. 2016.
3. Habibi Y, Lucia LA, Rojas OJ. Cellulose nanocrystals: chemistry, self-assembly, and applications. Chem Rev. 2010;110:3479–500.
4. Himmel ME, Ding S-Y, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 2007;315:804–7.
5. Matthews JF, Skopec CE, Mason PE, Zuccato P, Torget RW, Sugiyama J, Himmel ME, Brady JW. Computer simulation studies of microcrystalline cellulose Iβ. Carbohydr Res. 2006;341:138–52.
6. Pérez S, Samain D. Structure and engineering of celluloses. In: Derek H, editor. Advances in carbohydrate chemistry and biochemistry, vol. 64. New York: Academic Press; 2010. p. 25–116.
7. Nisizawa K. Mode of action of cellulases. J Ferment Technol. 1973;51:267–304.
8. Ding S-Y, Himmel ME. The maize primary cell wall microfibril: a new model derived from direct visualization. J Agric Food Chem. 2006;54:597–606.
9. Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, et al. Toward a systems approach to understanding plant cell walls. Science. 2004;306:2206–11.
10. Harada H, Côté WA. Structure of wood. In: Higuchi T, editor. Biosynthesis and biodegradation of wood components. Orlando: Academic Press, Inc.; 1985. p. 1–42.
11. Pettersen RC. The chemical composition of wood. In: Rowell R, editor. The chemistry of solid wood, vol. 207. Washington: American Chemical Society; 1984. p. 57–126.
12. Timell TE. Recent progress in the chemistry of wood hemicelluloses. Wood Sci Technol. 1967;1:45–70.
13. Battista OA, Smith PA. Level-off d.p. cellulose products. U.S. Patent No. 2,978,446. 1961.
14. Hanna M, Biby G, Miladinov V. Production of microcrystalline cellulose by reactive extrusion. U.S. Patent No. 6,228,213. 2001.
15. Adel AM, Abd El-Wahab ZH, Ibrahim AA, Al-Shemy MT. Characterization of microcrystalline cellulose prepared from lignocellulosic materials. Part II: physicochemical properties. Carbohydr Polym. 2011;83:676–87.

16. Baehr M, Führer C, Puls J. Molecular weight distribution, hemicellulose content and batch conformity of pharmaceutical cellulose powders. Eur J Pharm Biopharm. 1991;37:136–41.

17. Landín M, Martínez-Pacheco R, Gómez-Amoza JL, Souto C, Concheiro A, Rowe RC. Effect of batch variation and source of pulp on the properties of microcrystalline cellulose. Int J Pharm. 1993;91:133–41.

18. Landín M, Martínez-Pacheco R, Gómez-Amoza JL, Souto C, Concheiro A, Rowe RC. Effect of country of origin on the properties of microcrystalline cellulose. Int J Pharm. 1993;91:123–31.

19. Newman RH. Crystalline forms of cellulose in softwoods and hardwoods. J Wood Chem Technol. 1994;14:451–66.

20. Rowe RC, McKillop AG, Bray D. The effect of batch and source variation on the crystallinity of microcrystalline cellulose. Int J Pharm. 1994;101:169–72.

21. Paloheimo M, Haarmann T, Mäkinen S, Vehmaanperä J. Production of industrial enzymes in Trichoderma reesei. In: Schmoll M, Dattenböck C, editors. Gene expression systems in fungi: advancements and applications. Cham: Springer; 2016. p. 23–57.

22. Seiboth B, Ivanova C, Seidl-Seiboth V. Trichoderma reesei: a fungal enzyme producer for cellulosic biofuels. In: Dos Santos Bernardes MA, editor. Biofuel production—recent developments and prospects. Rijeka: InTech; 2011.

23. Viikari L, Vehmaanperä J, Koivula A. Lignocellulosic ethanol: from science to industry. Biomass Bioenergy. 2012;46:13–24.

24. Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, et al. Tracking the roots of cellulase hyperproduction by the fungus Trichoderma reesei using massively parallel DNA sequencing. Proc Natl Acad Sci U S A. 2009;106:16151–6.

25. Peterson R, Nevalainen H. Trichoderma reesei RUT-C30—thirty years of strain improvement. Microbiology. 2012;158:58–68.

26. Kubicek CP. Systems biological approaches towards understanding cellulase production by Trichoderma reesei. J Biotechnol. 2013;163:133–42.

27. Glass NL, Schmoll M, Cate JH, Coradetti S. Plant cell wall deconstruction by ascomycete fungi. Annu Rev Microbiol. 2013;67:477–98.

28. Häkkinen M, Valkonen MJ, Westerholm-Parvinen A, Aro N, Arvas M, Vitikainen M, Penttilä M, Saloheimo M, Pakula TM. Screening of candidate regulators for cellulase and hemicellulase production in Trichoderma reesei and identification of a factor essential for cellulase production. Biotechnol Biofuels. 2014;7:14.

29. Huberman LB, Liu J, Qin L, Glass NL. Regulation of the lignocellulolytic response in filamentous fungi. Fungal Biol Rev. 2016;30:101–11.

30. Seibert T, Thieme N, Benz JP. The renaissance of Neurospora crassa: how a classical model system is used for applied research. In: Schmoll M, Dattenböck C, editors. Gene expression systems in fungi: advancements and applications. Cham: Springer; 2016. p. 59–96.

31. Coradetti ST, Craig JP, Xiong Y, Shock T, Tian C, Glass NL. Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. Proc Natl Acad Sci U S A. 2012;109:7397–402.

32. Craig JP, Coradetti ST, Starr TL, Glass NL. Direct target network of the Neurospora crassa plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. MBio. 2015;6:e01452-15.

33. Bailey C, Arst HN. Carbon catabolite repression in Aspergillus nidulans. Eur J Biochem. 1975;51:573–7.

34. de la Serna I, Ng D, Tyler BM. Carbon regulation of ribosomal genes in Neurospora crassa occurs by a mechanism which does not require Cre-1, the homologue of the Aspergillus carbon catabolite repressor, CreA. Fungal Genet Biol. 1999;26:253–69.

35. Ilmén M, Thrane C, Penttilä M. The glucose repressor gene cre1 of Trichoderma: isolation and expression of a full-length and a truncated mutant form. Mol Gen Genet MGG. 1996;251:451–60.

36. Sun J, Glass NL. Identification of the CRE-1 cellulolytic regulon in Neurospora crassa. PLoS ONE. 2011;6:e25654.

37. Mello-de-Sousa TM, et al. A truncated form of the carbon catabolite repressor 1 increases cellulase production in Trichoderma reesei. Biotechnol Biofuels. 2014;7(1):129.

38. Sternberg D, Mandels GR. Induction of cellulolytic enzymes in Trichoderma reesei by sophorose. J Bacteriol. 1979;139:761–9.

39. Zhou Q, Xu J, Kou Y, Lv X, Zhang X, Zhao G, Zhang W, Chen G, Liu W. Differential involvement of β-glucosidases from Hypocrea jecorina in rapid induction of cellulase genes by cellulose and cellobiose. Eukaryot Cell. 2012;11:1371–81.

40. Znameroski EA, Coradetti ST, Roche CM, Tsai JC, Iavarone AT, Cate JH, Glass NL. Induction of lignocellulose-degrading enzymes in Neurospora crassa by cellodextrins. Proc Natl Acad Sci U S A. 2012;109:6012–7.

41. Fan LT, Lee YH, Beardmore DR. The influence of major structural features of cellulose on rate of enzymatic hydrolysis. Biotechnol Bioeng. 1981;23:419–24.

42. Fan LT, Lee Y-H, Beardmore DH. Mechanism of the enzymatic hydrolysis of cellulose: effects of major structural features of cellulose on enzymatic hydrolysis. Biotechnol Bioeng. 1980;22:177–99.

43. Hall M, Bansal P, Lee JH, Realff MJ, Bommarius AS. Cellulose crystallinity—a key predictor of the enzymatic hydrolysis rate. FEBS J. 2010;277:1571–82.

44. Li L, Zhou W, Wu H, Yu Y, Liu F, Zhu D. Relationship between crystallinity index and enzymatic hydrolysis performance of celluloses separated from aquatic and terrestrial plant materials. BioResources. 2014;9:3993–4005.

45. Mansfield SD, Mooney C, Saddler JN. Substrate and enzyme characteristics that limit cellulose hydrolysis. Biotechnol Prog. 1999;15:804–16.

46. Mooney CA, Mansfield SD, Beatson RP, Saddler JN. The effect of fiber characteristics on hydrolysis and cellulase accessibility to softwood substrates. Enzyme Microb Technol. 1999;25:644–50.

47. Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. Biotechnol Biofuels. 2010;3:10.

48. Peciulyte A, Anasontzis GE, Karlström K, Larsson PT, Olsson L. Morphology and enzyme production of Trichoderma reesei Rut C-30 are affected by the physical and structural characteristics of cellulosic substrates. Fungal Genet Biol. 2014;72:64–72.

49. Peng H, Li H, Luo H, Xu J. A novel combined pretreatment of ball milling and microwave irradiation for enhancing enzymatic hydrolysis of microcrystalline cellulose. Bioresour Technol. 2013;130:81–7.

50. Puri VP. Effect of crystallinity and degree of polymerization of cellulose on enzymatic saccharification. Biotechnol Bioeng. 1984;26:1219–22.

51. Rollin JA, Zhu Z, Sathitsuksanoh N, Zhang YHP. Increasing cellulose accessibility is more important than removing lignin: a comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia. Biotechnol Bioeng. 2011;108:22–30.

52. Horii F, Hirai A, Kitamaru R. CP/MAS carbon-13 NMR study of spin relaxation phenomena of cellulose containing crystalline and noncrystalline components. J Carbohydr Chem. 1984;3:641–62.

53. Newman RH. Homogeneity in cellulose crystallinity between samples of Pinus radiata wood. Holzforschung. 2004;58:91–6.

54. Sterk H, Sattler W, Janosi A, Paul D, Esterbauer H. Einsatz der Festkörper 13C-NMR-Spektroskopie für die Bestimmung der kristallinität in Cellulosen. Das Papier. 1987;41:664–7.

55. Newman RH, Hemmingson JA. Determination of the degree of cellulose crystallinity in wood by carbon-13 nuclear magnetic resonance spectroscopy. Holzforschung. 1990;44:351–5.

56. Benz JP, Chau BH, Zheng D, Bauer S, Glass NL, Somerville CR. A comparative systems analysis of polysaccharide-elicited responses in Neurospora crassa reveals carbon source-specific cellular adaptations. Mol Microbiol. 2014;91:275–99.

57. Ogawa M, Kobayashi T, Koyama Y. ManR, a novel Zn(II)2Cys6 transcriptional activator, controls the beta-mannan utilization system in Aspergillus oryzae. Fungal Genet Biol. 2012;49:987–95.

58. Sun J, Tian C, Diamond S, Glass NL. Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in Neurospora crassa. Eukaryot Cell. 2012;11:482–93.

59. Tian C, Beeson WT, Iavarone AT, Sun J, Marletta MA, Cate JH, Glass NL. Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. Proc Natl Acad Sci U S A. 2009;106:22157–62.

60. Ogawa M, Kobayashi T, Koyama Y. ManR, a transcriptional regulator of the beta-mannan utilization system, controls the cellulose utilization system in Aspergillus oryzae. Biosci Biotechnol Biochem. 2013;77:426–9.

61. Dashtban M, Buchkowski R, Qin W. Effect of different carbon sources on cellulase production by Hypocrea jecorina (Trichoderma reesei) strains. Int J Biochem Mol Biol. 2011;2:274–86.

62. Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sándor E, Le Crom S. The CRE1 carbon catabolite repressor of the fungus Trichoderma reesei: a

master regulator of carbon assimilation. BMC Genom. 2011;12(1):269.

63. Tangnu SK, Blanch HW, Wilke CR. Enhanced production of cellulase, hemi-cellulase, and β-glucosidase by *Trichoderma reesei* (Rut C-30). Biotechnol Bioeng. 1981;23(8):1837–49.

64. Willför S, Sundberg A, Pranovich A, Holmbom B. Polysaccharides in some industrially important hardwood species. Wood Sci Technol. 2005;39:601–17.

65. O'Sullivan AC. Cellulose: the structure slowly unravels. Cellulose. 1997;4:173–207.

66. Mach-Aigner AR, Pucher ME, Steiger MG, Bauer GE, Preis SJ, Mach RL. Transcriptional regulation of xyr1, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. Appl Environ Microbiol. 2008;74:6554–62.

67. Fox JM, Levine SE, Clark DS, Blanch HW. Initial- and processive-cut products reveal cellobiohydrolase rate limitations and the role of companion enzymes. Biochemistry. 2012;51:442–52.

68. Levine SE, Fox JM, Blanch HW, Clark DS. A mechanistic model of the enzymatic hydrolysis of cellulose. Biotechnol Bioeng. 2010;107:37–51.

69. Peters LE, Walker LP, Wilson DB, Irwin DC. The impact of initial particle size on the fragmentation of cellulose by the cellulase of *Thermomonospora fusca*. Bioresour Technol. 1991;35:313–9.

70. Mandels M, Andreotti R. Problems and challenges in the cellulose to cellulase fermentation. Proc Biochem. 1978;13:6–13.

71. Vogel HJ. A convenient growth medium for Neurospora (medium N). Microb Genet Bull. 1956;13:42–3.

72. Brunauer S, Emmett PH, Teller E. Adsorption of gases in multimolecular layers. J Am Chem Soc. 1938;60:309–19.

73. Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

ATNT: an enhanced system for expression of polycistronic secondary metabolite gene clusters in *Aspergillus niger*

Elena Geib and Matthias Brock*

Abstract

Background: Fungi are treasure chests for yet unexplored natural products. However, exploitation of their real potential remains difficult as a significant proportion of biosynthetic gene clusters appears silent under standard laboratory conditions. Therefore, elucidation of novel products requires gene activation or heterologous expression. For heterologous gene expression, we previously developed an expression platform in *Aspergillus niger* that is based on the transcriptional regulator TerR and its target promoter P*terA*.

Results: In this study, we extended this system by regulating expression of *terR* by the doxycycline inducible Tet-on system. Reporter genes cloned under the control of the target promoter P*terA* remained silent in the absence of doxycycline, but were strongly expressed when doxycycline was added. Reporter quantification revealed that the coupled system results in about five times higher expression rates compared to gene expression under direct control of the Tet-on system. As production of secondary metabolites generally requires the expression of several biosynthetic genes, the suitability of the self-cleaving viral peptide sequence P2A was tested in this optimised expression system. P2A allowed polycistronic expression of genes required for Asp-melanin formation in combination with the gene coding for the red fluorescent protein tdTomato. Gene expression and Asp-melanin formation was prevented in the absence of doxycycline and strongly induced by addition of doxycycline. Fluorescence studies confirmed the correct subcellular localisation of the respective enzymes.

Conclusion: This tightly regulated but strongly inducible expression system enables high level production of secondary metabolites most likely even those with toxic potential. Furthermore, this system is compatible with polycistronic gene expression and, thus, suitable for the discovery of novel natural products.

Keywords: Asp-melanin, P2A, Polycistronic mRNA, Tet-on system, Doxycycline, Terrein biosynthetic gene cluster

Background

Genome mining has revealed that fungal genomes contain a large number of yet unexplored secondary metabolite biosynthetic gene clusters [1]. Due to next generation sequencing approaches the number of available fungal genomes is steadily increasing as can be seen from the growing number of genomes in the 1000 fungal genomes project [2]. Interestingly, even highly related fungal species contain at least a few unique secondary metabolite biosynthetic gene clusters [3] and it has frequently been observed that more than one metabolite is produced from a single biosynthetic gene cluster [4]. Therefore, the potential of producing metabolites with interesting pharmaceutical characteristics appears nearly unlimited. However, as secondary metabolites are frequently produced in response to distinct biotic or abiotic stress factors [5], a large number of the respective biosynthetic gene clusters remains silent under laboratory conditions and, thus, their products unexplored. To exploit the full potential of fungal secondary metabolite production different strategies have been applied [6, 7].

---

*Correspondence: Matthias.brock@nottingham.ac.uk
Fungal Genetics and Biology, School of Life Sciences, University
of Nottingham, University Park, Nottingham NG7 2RD, UK

One approach that can be directly applied to cultivable fungal species is the addition of epigenetic modifiers [8] or co-cultivation with other microbes, which may result in the specific induction of biosynthetic gene clusters [9]. However, while this strategy may lead to the production of novel metabolites, a direct correlation between biosynthetic gene cluster and metabolite product remains difficult. Another strategy is the overexpression of a transcriptional regulator controlling a specific biosynthetic gene cluster [10]. Unfortunately, not all secondary metabolite biosynthetic gene clusters contain a transcriptional activator in direct proximity to their biosynthetic genes [11], which may hamper this approach. In addition, global transcriptional regulators may overrule the activation from a cluster specific transcription factor as shown for the dihydroisoflavipucine biosynthesis in *Aspergillus terreus* [12]. While this biosynthetic gene cluster contains a specific transcriptional activator that is indispensable for its activation, the activating effect is overruled in the presence of glucose through the carbon catabolite repressor CreA [12].

The strategy of targeted activation of cluster specific transcription factors additionally requires the ability for genetic modification of the natural producer strain and may not be suitable for many fungal species. Therefore, recent approaches used the generation of fungal artificial chromosomes (FAC) to clone and transfer whole fungal gene clusters into genetically amenable fungal expression platform strains [13]. In a previous study, 56 gene cluster containing FACs with yet uncharacterised biosynthetic genes from *Aspergillus wentii*, *Aspergillus aculeatus* and *A. terreus* were transferred to *Aspergillus nidulans*, which resulted in the identification of 17 novel metabolites from 15 different FACs [13]. However, not all gene clusters were successfully activated in the recombinant host, which may be due to the lack of transcriptional activators, repressing conditions or the lack of the correct starter metabolites in the heterologous host.

It has also been shown that induction of secondary metabolite biosynthetic gene clusters in a heterologous host can be achieved by regulating the expression of the global regulator of secondary metabolism LaeA in *Aspergillus* species [14]. In this respect, a transfer of the biosynthetic gene clusters for monacolin K from *Monascus pilosus* and terrequinone A from *A. nidulans* resulted in successful product formation after overexpression of *leaA* in *Aspergillus oryzae* [15]. However, induction of several biosynthetic gene clusters appears independent from LaeA control and a specific transcriptional activator in direct proximity to the biosynthetic gene cluster may be lacking. Therefore, a different strategy for gene activation was successfully applied to *A. nidulans*, in which a serial promoter exchange of each individual gene of a

biosynthetic gene cluster was performed. This strategy resulted in the identification of the proteasome inhibitor fellutamide B and its resistance conferring gene *inpE* [16]. Although successful, this strategy required several rounds of metabolite screening, marker regeneration and subsequent transformation and appears prohibitively time consuming for routine applications.

Due to these challenges it remains difficult to recommend an expression system that allows for high throughput screening for all yet uncharacterised secondary metabolite biosynthetic gene clusters. Heterologous gene expression generally aims for high product yields to elucidate the structure of the metabolite and to analyse its biological activity. A prerequisite for this is the high level expression of target genes, which can be achieved by generating multiple copy integrations, selection of strong promoters or a combination of both [17, 18]. Recently, we introduced a heterologous expression system that uses an *Aspergillus niger* strain as expression platform that contains regulatory elements from *A. terreus* [18]. These regulatory elements consist of the terrein biosynthetic gene cluster specific transcriptional activator TerR and its target promoter P*terA*. When expression of *ter*R is controlled by the *A. oryzae* amylase promoter and a reporter gene is expressed under P*terA* control the induction level of the amylase promoter gets amplified through this coupled system [18]. In addition, a SM-Xpress vector has been constructed that allows easy generation of expression plasmids by in vitro recombination with the target gene. This expression system had been successfully applied for the identification of lecanoric acid as product from the *A. nidulans orsA* gene [18], has enabled the heterologous in vivo reconstruction of the *A. terreus* Asp-melanin biosynthetic pathway in *A. niger* [19] and was recently successfully used for identification of basidioferrin, which is a novel siderophore produced from a non-ribosomal peptide synthetase (NRPS) that is widely distributed among basidiomycetes [20].

Another challenge in heterologous expression of secondary metabolite biosynthetic gene clusters derives from possible toxicity of resulting metabolites. Therefore, a tight regulation of gene expression is favoured as it allows for the formation of fungal biomass prior to induction of the expression of target genes. In this respect a tuneable Tet-on/Tet-off expression system has been adapted for use in *Aspergillus* species [21, 22]. The Tet-on system uses a reverse tetracycline-controlled transactivator that enables titratable induction of gene expression by the addition of the tetracycline derivative doxycycline. In the absence of doxycycline gene expression remains at low background levels, but expression gets strongly induced by addition of doxycycline [21, 22].

As most secondary metabolites are produced from biosynthetic gene clusters, production of the final metabolite generally requires the heterologous expression of more than only one single gene. While a strategy of subsequent transformations with isolated genes accompanied by a marker recycling technique works for clusters comprising only a small number of genes, this procedure is extremely time consuming and probably not suitable for larger clusters containing five or more genes. Therefore, another strategy is the use of self-splicing viral peptide sequences such as the 2A peptide that separates proteins from a polycistronic messenger in different viruses such as the porcine teschovirus-1 (P2A). P2A and similar sequences have been successfully used to separate individual proteins in a range of different eukaryotic organisms [23, 24] among them yeasts such as *Saccharomyces cerevisiae* [25, 26] and *Pichia pastoris* [27]. A recent study also used a 2A peptide in *Trichoderma reesei*, in which the gene coding for the cellobiohydrolase Cel7A from *Penicillium funiculosum* was combined in a single transcript with the eGFP coding gene to ease screening of cellobiohydrolase positive transformants [28]. Importantly, this technique has also been applied for heterologous production of penicillin in the filamentous fungus *A. nidulans* by genetic engineering of a synthetic *Penicillium chrysogenum* penicillin biosynthetic gene cluster [29]. Despite low yields, penicillin K was successfully produced by *A. nidulans* transformants expressing the polycistronic penicillin biosynthetic gene cluster [29], indicating that this strategy is suitable for use in fungal secondary metabolite biosynthesis. The suitability of P2A was further confirmed in a recent study on enniatin biosynthesis in *A. niger*, in which two genes required for enniatin biosynthesis and a luciferase were separated by P2A sequences [30]. While a positioning effect in dependence of the gene order in the polycistronic messenger was observed, all strains produced enniatin and displayed light emission from luciferase activity. Positioning effects were also observed in the cellobiohydrolase expression in *T. reesei* [28] and murine cells [31], indicating that despite polycistronic gene expression the amount of individual proteins may vary depending on the gene order in the expression construct.

Here, we aimed to generate an optimised fungal heterologous expression system by combining the three latter aspects of heterologous secondary metabolite production in *A. niger*: (1) using the expression amplification system of TerR/P*terA* under (2) fine-tuneable control of the Tet-on system for expression of (3) polycistronic mRNA of the Asp-melanin biosynthetic genes combined with a fluorescent reporter to study correct subcellular localisation of enzymes.

Results

Integration of Tet-on control into the coupled TerR/PterA expression system

We previously developed a heterologous expression system in *A. niger* that bases on the transcriptional activator TerR from the *A. terreus* terrein biosynthetic pathway and its *terA* (P*terA*) target promoter [18]. In this combination the induction level of P*terA* directly depends on the transcriptional level of the *terR* gene [18]. Furthermore, the activity of the promoter controlling *terR* expression gets amplified at the target promoter P*terA* as reporter expression in the coupled system was significantly higher than direct expression of reporter genes [18]. In this first version of the expression system, we controlled *terR* expression by either the glyceraldehyde-3-phosphate dehydrogenase promoter P*gpdA* or the amylase promoter P*amyB*. As both of these promoters derive from primary metabolism, their use may interfere with fungal metabolic physiology. In addition, both promoters are difficult to silence and P*amyB* shows significant background activity even when *A. niger* is grown on casamino acids in the absence of any sugars. As this background promoter activity may hamper the production of toxic metabolites, we replaced P*amyB* in the control of *terR* by the reverse tetracycline-controlled transactivator (Fig. 1a) containing the *fraA* promoter sequence for improved cassette stability. The *fraA* gene encodes a putative ribosomal subunit and had been identified from microarray analyses showing a similar expression pattern as the glyceraldehyde-3-phosphate dehydrogenase and is assumed to be constitutively expressed [22]. The Tet-on:*terR* construct was used for transformation of the *A. niger* A1144 strain (Fungal Genetics Stock Center, Kansas, USA) and resulting transformants were analysed for full length single copy integration into the genome (Additional file 1). The resulting expression platform strain ATNT16 (ATNT = **A**1144 **T**et-o**n**:**t**erR) was analysed for its performance in gene expression.

Analysis of β-galactosidase reporter gene expression

To elucidate the performance of the new Tet-on-controlled expression system in *A. niger* ATNT16, we generated β-galactosidase reporter strains. Two different constructs were made that both contained the *lacZ* gene from *Escherichia coli* as reporter. The first construct contained a fusion of P*terA* with the *lacZ* gene (P*terA:lacZ*) for transformation of the ATNT16 strain (Fig. 1a). The second construct contained a fusion of the Tet-on promoter system directly with the *lacZ* gene (Tet-on:*lacZ*) for transformation of the parental *A. niger* strain A1144 (Fig. 1b). This enabled the comparison of doxycycline dependent gene activation in the coupled amplification system of TerR/P*terA* under control of Tet-on against

Fig. 1 Analysis of β-galactosidase reporter activity from different expression systems. **a**–**c** Schematic representation of expression systems. **a** The ATNT system, in which expression of the transcriptional regulator *terR* is controlled by Tet-on. The reverse transactivator is constitutively expressed from the *fraA* promoter and transcription is terminated by the *ergA* terminator (T). In the presence of doxycycline the reverse transactivator binds to the *tet-on* responsive element (*1*), which leads to transcription of *terR* from the minimal *gpdA* promoter (*2*). TerR specifically recognises its P*terA* target promoter, which results in transcription of the β-galactosidase gene *lacZ*. **b** Direct control of *lacZ* expression by the Tet-on promoter system. **c** The P2 expression system. The *amyB* promoter controls expression of *terR* and is activated in the presence of sugars. TerR controls expression of *lacZ* by binding to the P*terA* promoter. **d** β-Galactosidase activity from expression systems shown in **a**–**c**. ATNT and Tet-on *lacZ* strains were grown in the presence of different doxycycline concentrations, whereas the P2 *lacZ* strain was cultivated without doxycycline. The positions of activities derived from parental strains are highlighted by arrows. Three individual transformants from each expression system were measured in biological and technical duplicates. Bars represent mean values of all measurements and standard deviations are indicated by error bars. Multiple t-tests using the Holm-Sidak method were applied, which confirmed significantly different expression levels in all groups in dependence of the doxycycline concentration. (****$p < 0.0001$; *n. d.* not determined)

the direct reporter gene induction by the Tet-on system. After transformation of the respective *A. niger* strains, transformants with a single copy integration of the respective reporter construct were identified by Southern blot analysis (Additional file 2) and three independent transformants from each construct were selected for downstream investigation of reporter activities. In addition, three reporter strains from the original P*amyB:terR* expression platform (P2 strain; *terR* gene under control of the amylase promoter) with single copy integration of the *lacZ* gene under control of P*terA* [18] were included (Fig. 1c). This allowed comparison of expression properties of the new ATNT16 expression platform with that of the previous platform strain P2. All strains were cultivated for 24 h on 100 mM glucose containing minimal media with 20 mM glutamine as nitrogen source and 1% talc to avoid the formation of cell pellets [32]. For the Tet-on-containing strains parallel cultures were supplemented with various amounts of doxycycline in a range between 0 and 30 μg/ml. All strains were cultivated in two biological replicates and β-galactosidase activity was determined from cell-free extracts in technical duplicates (Fig. 1d). The average specific β-galactosidase activity of

the P2 reporter strain on this glucose containing medium was about 6 U/mg, which was in agreement with previous determinations under this growth condition [18]. Both, the ATNT16 reporter strains as well as the Tet-on:*lacZ* strains only revealed very low background activity when cultivated in the absence of doxycycline (< 0.05 U/mg). Addition of doxycycline to the Tet-on:*lacZ* strains resulted in a titratable induction of reporter activity, reaching a maximum of 3.4 U/mg at 30 μg/ml of doxycycline. The ATNT16 strain with the P*terA:lacZ* reporter construct showed significant reporter activity of 2.5 U/mg already at 0.75 μg/ml, which further increased to 17.3 U/mg at 30 μg/ml of doxycycline. This latter activity is about five times higher than the maximum activity obtained from the uncoupled system at 30 μg/ml in which Tet-on directly induces the expression of the target gene. Thus, the ATNT16 expression platform with the *terR* gene under Tet-on control is tightly regulated in the absence of doxycycline and strongly induced by its addition. However, accompanied with high reporter gene expression, biomass formation in the presence of 30 μg/ml doxycycline in the ATNT16 reporter strains was significantly reduced. This may be due to the high reporter

protein production during initiation of germination and seems independent from high levels of activated trans-activator protein as the ATNT16 strain without *lacZ* reporter construct showed no growth defects in the presence of 30 µg/ml doxycycline and β-galactosidase background activity in the ATNT16 strain did not increase by the addition of different doxycycline concentrations (not shown).

Production of aspulvinone E in the ATNT16 expression platform

Our reporter gene analyses indicated that the Tet-on-controlled TerR/P*terA* system is tightly regulated and allows high level gene expression in the presence of doxycycline. To test whether this also transfers to secondary metabolite production, we used the *melA* gene, which encodes the aspulvinone E synthetase from *A. terreus* [19] under control of P*terA* and transferred this construct into the Tet-on:*terR* strain ATNT16. After selection for single copy integration (Additional file 3), strains were cultivated for 48 h in glucose minimal medium either in the absence or presence of 15 µg/ml doxycycline. In accordance with a light yellow colour of aspulvinone E, the cultures grown in the presence of doxycycline turned yellow (Fig. 2a) and the main proportion of the coloured substance solved in the ethyl acetate phase during extraction of culture filtrates (Fig. 2b). In contrast, no obvious yellow colouration of the culture or the ethyl acetate phase was observed in the control cultures without doxycycline (Fig. 2a, b). To confirm that aspulvinone E was produced only under inducing conditions samples were analysed by HPLC using reversed phase chromatography on a C_{18} column. As shown in Fig. 2c the induced culture revealed a strong signal for aspulvinone E and a minor signal of its stereoisomer isoaspulvinone E [19]. By contrast, only extremely weak background signals were detected in the control culture. These results are in agreement with the β-galactosidase reporter studies and confirm that (1) metabolite production is suppressed in the absence of doxycycline and (2) high yields of metabolites can be achieved under inducing conditions even in strains only carrying a single copy integration of the gene of interest.

Model gene cluster expression from polycistronic mRNA using the P2A peptide

In the next step we aimed in the expression of multiple genes in the Tet-on-controlled TerR/P*terA* system by engineering polycistronic mRNAs. For a proof-of-concept, the Asp-melanin pathway combined with a fluorescent reporter was used [19]. Asp-melanin is the conidial pigment produced by *A. terreus* and is distinct from the dihydroxynaphtalene melanin found in conidia of other
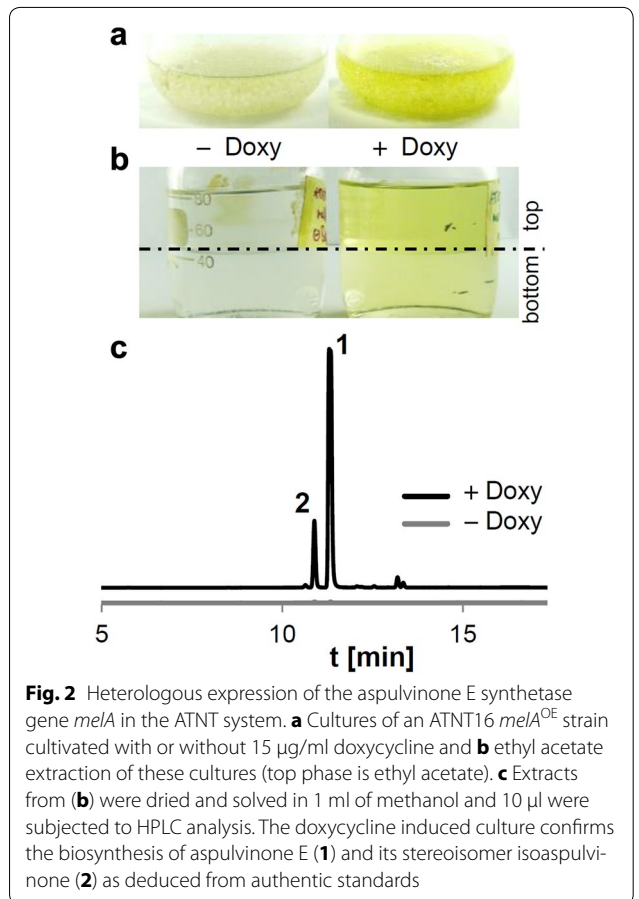
Fig. 2 Heterologous expression of the aspulvinone E synthetase gene *melA* in the ATNT system. **a** Cultures of an ATNT16 *melA*[OE] strain cultivated with or without 15 µg/ml doxycycline and **b** ethyl acetate extraction of these cultures (top phase is ethyl acetate). **c** Extracts from (**b**) were dried and solved in 1 ml of methanol and 10 µl were subjected to HPLC analysis. The doxycycline induced culture confirms the biosynthesis of aspulvinone E (**1**) and its stereoisomer isoaspulvinone E (**2**) as deduced from authentic standards

Aspergillus species. This melanin pigment does not derive from a naphtopyrone precursor that is produced by a polyketide synthase rather than an aspulvinone E synthetase, which is a non-ribosomal peptide synthetase-like (NRPS-like) protein [19]. This pigment biosynthesis pathway appeared most suitable as: (1) Asp-melanin is produced from only two proteins, which are the aspulvinone E synthetase MelA and the tyrosinase TyrP; (2) co-expression of individually controlled genes in the *A. niger* P2 strain resulted in brown mycelium due to the formation of Asp-melanin, which is easy to visualise; (3) Asp-melanin formation requires the correct subcellular localisation of both enzymes as MelA requires the reducing environment of the cytoplasm and TyrP the oxidising environment of Golgi or ER (4) protein localisation and cleavage efficiency can be visualised by using the red fluorescent protein tdTomato as a reporter.

For the separation of individual proteins during ribosomal translation, the 22 amino acid 2A peptide (P2A, GSGATNFSLLKQAGDVEENPGP) sequence from porcine teschovirus-1 was used [24], whereby codon sequences of individual P2A peptides were varied on DNA level to allow directed in vitro recombination into

the SM-Xpress expression vector that contains the *terA* promoter, the *trpC* terminator sequence and a resistance gene for selection of transformants [18]. Two different polycistronic constructs consisting of the *melA* gene, the *tyrP* gene and the gene coding for tdTomato were generated to test the efficiency of P2A cleavage and protein localisation in the ATNT16 expression platform (Fig. 3a). For the first construct all three genes were separated by a P2A coding sequence (P2A_P2A construct), which was assumed to result in three individual functional proteins under inducing conditions that lead to brown mycelium and a cytoplasmic localisation of tdTomato as this reporter does not contain a subcellular localisation signal. The second construct only contained a single P2A sequence (P2A construct) separating the *melA* and *tyrP* genes, whereby the gene coding for tdTomato was fused in frame with the *tyrP* gene [19]. Here, we expected the formation of brown mycelium under inducing conditions, but a fluorescence localisation in subcellular organelles of ER and Golgi, which would confirm the correct targeting of TyrP. ATNT16 was transformed with the respective constructs and resulting transformants were analysed by Southern blot analysis (Additional file 4) for single copy integration. For analysis of selected transformants split plates were prepared with glucose minimal medium containing 0 or 10 μg/ml of doxycycline. On these plates the control strain ATNT16 as well as strains containing either the P2A or the P2A_P2A construct were spotted and pictures taken after 72 h of incubation. As shown in Fig. 3b, all strains showed similar growth and conidia formation in the top view of plates. However, the bottom view shows that mycelium of strains
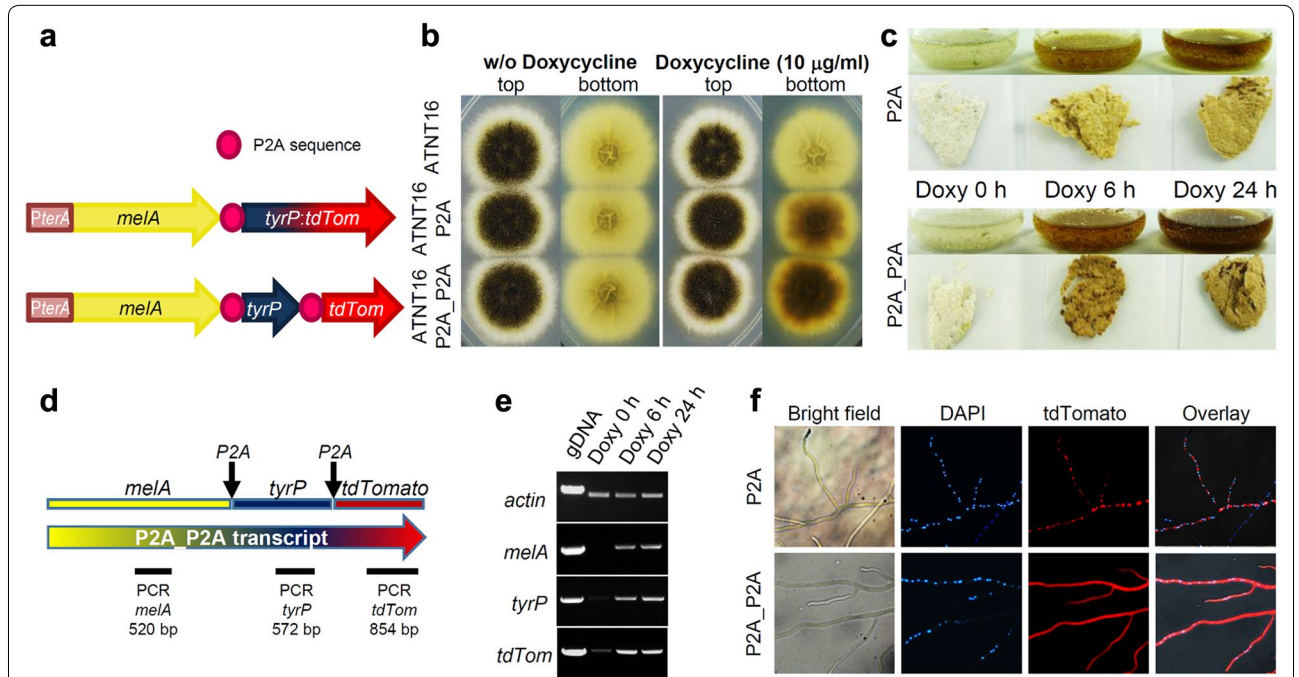
Fig. 3 Asp-melanin formation and subcellular protein localisation from polycistronic gene expression in the ATNT system. **a** Schematic presentation of polycistronic expression constructs separated by P2A sequences. *tyrP:tdTom* denotes an in frame fusion of the *tyrP* gene with the gene coding for the red fluorescent protein tdTomato. **b** Colonies in top and bottom view of the parental strain ATNT16 and strains carrying the expression construct with one or two P2A separations grown in the absence and presence of doxycycline. Addition of doxycycline induces the formation of Asp-melanin, which is indicated by brown colouration of mycelium in the bottom view. **c** Liquid cultures of ATNT16 strains carrying the expression construct with one or two P2A sequences. Mycelium was harvested after 24 h of incubation. Cultures were grown without doxycycline (Doxy 0 h) or were induced with doxycycline for the last 6 h of total incubation time (Doxy 6 h) or for the whole 24 h (Doxy 24 h). A stronger colouration of mycelium is observed when TyrP and tdTomato are separated by an additional P2A peptide. **d** Scheme of the polycistronic P2A_P2A mRNA. Localisation of the individual gene sequences are indicated above and localisation and size of PCR products for verification of transcription are shown below the transcript. **e** Semiquantitative RT-PCR on cDNA derived from cultures in **c**. The actin gene was used for normalisation of cDNA. Amplification from genomic DNA (gDNA) is shown as a control with a decrease in fragment size of the actin gene due to intron splicing. Full length-transcription of the polycistronic messenger is confirmed by PCR products from all genes when grown in the presence of doxycycline. **f** Fluorescence analysis for subcellular localisation of proteins produced from the two polycistronic expression constructs. Nuclei are shown in blue by DAPI staining. Red fluorescence indicates localisation of tdTomato. In the P2A construct the fusion of TyrP with tdTomato reflects a punctuated fluorescence consistent with ER and Golgi. When tdTomato is separated by P2A in the P2A_P2A construct, tdTomato localises to the cytoplasm

with the P2A and the P2A_P2A construct turned dark brown. Similarly, liquid cultures inoculated with conidia of the respective transformants were incubated for a total of 24 h in absence (Doxy 0 h) or presence of 15 μg/ml doxycycline. Thereby, the inducer doxycycline was added either directly at the start of cultivation (Doxy 24 h) or after a pre-cultivation for 18 h to allow for conidia germination and hyphae formation prior to induction resulting in an induction time of 6 h (Doxy 6 h). In the absence of doxycycline mycelia of cultures containing either of the two different constructs remained uncoloured, whereas mycelia turned brown under inducing conditions even when induced for only 6 h (Fig. 3c). Therefore, both constructs produce functional proteins that produce Asp-melanin and regulation of gene expression is active on solid and in liquid media.

Expression of full length polycistronic mRNAs

To confirm that all genes from the polycistronic mRNA were expressed with high efficiency only under inducing conditions, we aimed in semiquantitative RT-PCR analyses on the P2A_P2A transcript. Total RNA was isolated from liquid cultures containing the ATNT strain with the P2A_P2A construct. The cultures were grown for 24 h either without doxycycline or were induced for 6 or 24 h. cDNAs were generated with anchored oligo(dT) primers and cDNA levels from the different cultivation conditions were normalised against the *A. niger* actin gene. Oligonucleotides were deduced that amplify regions of the three individual genes that are contained in the polycistronic transcript (Fig. 3d). As shown in Fig. 3e, in the absence of doxycycline (Doxy 0 h) no amplification was observed on the 5′ and middle region of the polycistronic transcript encoding MelA and TyrP and only a weak band was detected for the gene sequence coding for tdTomato. This is in agreement with the low basal expression observed in *lacZ* reporter assays and in analysis of aspulvinone E metabolite production. However, strong induction was observed from both induced cultures (Doxy 6 h and Doxy 24 h) with PCR products on all three gene regions from the polycistronic messenger. This indicates that the full-length mRNA is efficiently transcribed only under doxycycline inducing conditions.

Subcellular localisation of proteins

Colouration of the mycelium indicated functional production and separation of MelA and TyrP and transcript analyses showed that all three genes encoded on the single transcript were efficiently transcribed. However, these analyses did not confirm the correct separation of proteins from the second P2A peptide, which should result in a cytoplasmic localisation of tdTomato, nor confirmed the correct subcellular localisation of any of the proteins. Therefore, fluorescence microscopy analyses were performed. While the supplementation of liquid media with talc avoids pellet formation, it hampers microscopic analyses of hyphae due to the attached talc particles. On the other hand, individual hyphae are difficult to visualise once fungal pellets are formed. As our analyses showed that Asp-melanin is formed either on liquid or solid media, strains were grown on glucose minimal media containing agar plates and coverslips coated with doxycycline containing glucose minimal media were placed around the colonies. Once hyphae grew on the edges of the coverslips they were removed, embedded in a DAPI-containing mounting solution and analysed by fluorescence microscopy. As shown in Fig. 3f hyphae of the strain containing the single P2A separator, which means a fusion of the *tyrP* gene with the gene coding for the red fluorescent protein tdTomato, exclusively showed red fluorescence in subcellular organelles most likely resembling Golgi and ER [19]. This indicates that TyrP is quantitatively transported into the correct subcellular compartment after P2A cleavage from MelA and a functional TyrP-tdTomato fusion protein is produced. In contrast, hyphae of the strain containing the P2A_P2A construct showed very strong cytoplasmic red fluorescence, indicating that both P2A cleavage sites were correctly recognised and full length tdTomato had been produced. Observation of colonies grown in the absence of doxycycline only revealed DAPI staining of nuclei but no red fluorescence signal (not shown).

Discussion

The aim of this study was the generation of an optimised fine-tuneable expression system in *A. niger* that produces high expression rates when fully induced and which is suitable to express polycistronic genes for recombinant expression of fungal secondary metabolite biosynthetic gene clusters. Previous studies investigated the suitability of *A. niger* as expression platform for the production of secondary metabolites such as the non-ribosomal peptide enniatin [33]. Thereby, the use of the Tet-on expression system combined with optimised fermentation conditions resulted in 4.5 g/l enniatin confirming both, the suitability of the Tet-on system to induce secondary metabolite production and the suitability of *A. niger* as expression platform. Similarly, we have previously shown that a range of different metabolites such as polyketides [18], non-ribosomal peptides [20] and products from NRPS-like enzymes [19] can be successfully produced in *A. niger*.

The combination of our TerR/P*terA*-system with the Tet-on-system resulted in an expression system with exceptionally high transcription rates which are still titratable. The coupling of Tet-on with the highly specific

transcription factor TerR resulted in an amplification of gene expression by more than 5 times compared to direct gene expression from Tet-on. Therefore, even single copy integrations result in high transcript levels, which makes a selection for multi copy integration strains dispensable [17] and reduces adverse effects on growth and physiology caused by multiple random genome integration events. However, as our approach did not target a specific gene locus, some positioning effect from independent single copy integration strains was observed, which is resembled in the standard deviations from β-galactosidase activity determinations.

Another advantage of using the coupled Tet-on controlled TerR/P*terA* expression system compared to direct expression from the Tet-on promoter system is the low concentration of doxycycline required for induction of gene expression. Significant expression rates were already observed at 0.75 μg/ml doxycycline. This activity from the coupled system was similar to that obtained from the direct Tet-on controlled gene expression at 15 μg/ml. Low amounts of doxycycline reduce the risk of co-extraction of the inducer when aiming for purification of secondary metabolites. However, as highest expression rates were observed at 15 to 30 μg doxycycline, these should be used especially when producing toxic metabolites that require significant biomass production prior to high level induction of gene expression. As a proof-of-concept we showed that Asp-melanin is efficiently produced when strains were pre-grown for 18 h and induced by doxycycline for only 6 h.

The tight regulation of the Tet-on controlled TerR/P*terA* expression system combined with its high induction rate makes it also superior to our original P*amyB* controlled TerR/P*terA* expression system, which is constitutively active on sugar containing media [18]. As glucose containing medium is generally used in the regeneration of protoplasts in fungal transformations, the production of toxic natural products may prevent growth of positive transformants. In contrast, regeneration of protoplasts of ATNT16 strain in the absence of doxycycline prevents expression of heterologous genes. In this respect, when we expressed the aspulvinone E synthetase gene *melA* in the *A. niger* P2 strain (P*amyB* control of *terR* expression) we suffered from the reduced ability of fungal colonies to produce conidia [19]. In this study, ATNT16 *melA*[OE] strains grown in the absence of doxycycline were indistinguishable from the parental control strain, unless induced by doxycycline.

Importantly, the P2A polycistronic gene expression was compatible with our high level expression system. Asp-melanin was efficiently produced from both, the single P2A construct that contained the fusion of TyrP with tdTomato as well as the construct in which all three

proteins were separated by P2A sequences. While we were not able to quantify the production of the insoluble pigment Asp-melanin, the brown colouration of the mycelium in the single P2A construct appeared less pronounced compared to the P2A_P2A construct (Fig. 3b, c). Positioning effects due to the order of genes in the polycistronic messenger as described previously [28, 30] that may in part be due to a drop-off of the ribosome after the translational skipping event [31] cannot account for this difference as the position of the *tyrP* gene was identical in both constructs. Therefore, it is likely that the fusion of TyrP with tdTomato affects activity of the tyrosinase. Nevertheless, this fusion unambiguously showed that the TyrP protein is correctly and quantitatively targeted to the ER and Golgi as (1) TyrP is inactive in the reducing environment of the cytoplasm [19] and therefore needs to be transported to the oxidative environment of Golgi and ER and (2) no cytoplasmic background fluorescence from tdTomato remaining in the cytoplasm was observed. Therefore, the recognition of the *N*-terminal subcellular localisation sequence has not been affected by the proline residue added to TyrP from the P2A peptide [24] when cleaved from the Aspulvinone E synthetase MelA. Whether *C*-terminal localisation sequences such as the peroxisomal PTS1 tripeptide import sequence SKL or AKL, as found in proteins of fungal siderophore biosynthesis [34], may be masked by the *C*-terminal addition of a P2A sequence needs to be tested in future studies.

This study also showed that the second P2A sequence in the P2A_P2A construct is efficiently cleaved as this construct showed extremely bright red fluorescence from the cytoplasm. However, due to the extremely bright fluorescence from this construct, which even leads to a reddish appearance of the edges of colonies on plates, we cannot exclude that some uncleaved protein may still be transported into the ER and Golgi. Nevertheless, due to the high fluorescence intensity from the cytoplasm combined with the high activity of the tyrosinase from the P2A_P2A construct, the majority of P2A peptides has efficiently been cleaved.

Conclusion

The combination of tightly controlled Tet-on induction with the highly specific TerR/P*terA* expression system resulted in a well-regulated fine-tuneable and very strong gene expression. Therefore, the system appears suitable for high-level production of metabolites, even those with antifungal properties. In addition, the system is compatible with the use of self-cleaving peptides such as P2A. Cleavage sites are efficiently recognised and at least *N*-terminal secretion signals seem to remain unaffected. Therefore, this system can be used for the discovery

of metabolites from yet unexplored fungal secondary metabolite biosynthetic gene clusters.

Methods

Media, fungal cultivation and transformation

Conidia suspensions were obtained by growing *Aspergillus niger* strains in slope cultures containing *Aspergillus* minimal medium with 50 mM glucose as carbon and 10 mM glutamine as nitrogen source [19] denoted as GG10 medium. For solid media 2% agar was added. Slopes were incubate for 4 days at 28 °C and overlaid with 6 ml phosphate-buffered saline (PBS) containing 0.01% Tween 20. Conidia were scraped into suspension using sterile cotton swaps. Suspensions were filtered over a 40 μm cell strainer (Greiner BioOne) to remove hyphae and clumps of conidia. After centrifugation the supernatant was discarded, conidia suspended in PBS and conidia concentrations determined using an improved Neubauer chamber. If not indicated otherwise, GG10 liquid cultures were inoculated with 1×10^6 conidia/ml with or without the addition of doxycycline (final concentration 0–30 μg/ml) and incubated at 28 °C on a rotary shaker at 150 rpm. Mycelia and culture supernatants were separated by filtration over Miracloth filter gauze (Merck, Calbiochem). Mycelia were pressed dry between tissue paper and frozen in liquid nitrogen for subsequent analyses. Fungal transformation was performed as described previously [20] with some minor modifications. Mycelia for protoplast formation were generated by inoculating YEPD medium (20 g peptone, 10 g yeast extract, 5 g glucose per litre) with spores of the *A. niger* wild-type strain A1144 (Fungal Genetics Stock Center, Kansas, USA) or the expression platform strain ATNT16. After 22 h mycelia were washed and incubated for 60 min in 90 mM citrate–phosphate buffer pH 7.3 containing 10 mM dithiothreitol. Protoplasts were generated by using 1.3 g/20 ml sterile filtered VinoTaste Pro (Novozymes) in osmotic medium with 0.6 M KCl as osmotic stabiliser. After transformation protoplasts were regenerated on solid GG10 media containing 1.2 M sorbitol and either 40 μg/ml phleomycin, 140 μg/ml hygromycin B or 0.1 μg/ml pyrithiamine as selectable marker. Genomic DNA was isolated using the MasterPure Yeast DNA purification kit (Epicenter).

Generation of the ATNT16 expression platform strain

All oligonucleotides used in this study are listed in Additional file 5: Table 1. All PCR reactions were performed in a SpeedCycler^2 (Analytic Jena) in 10 μl volumes using either Phusion (fragment cloning, Thermo Scientific) or Phire Hot Start II DNA polymerase (colony PCR, Thermo Scientific) for DNA amplification. PCR fragments and digested plasmids for cloning purposes

were gel purified using the GeneJet Gel Extraction Kit (Thermo Scientific). To generate the *A. niger* ATNT16 strain a plasmid was generated that contained the gene of the transcriptional activator *terR* under control of the Tet-on reverse transactivator system [22]. The construct was cloned into the *Hind*III linearized pUC19-ble [18] vector containing a phleomycin resistance cassette for selection of transformants. The Tet-on system was amplified with oligonucleotides 1 and 2 from plasmid pFW22.1 (kindly provided by V. Meyer, Berlin) and contained overhangs to the *Hind*III site of pUC19-ble. Subsequently, the *terR* gene including its own terminator sequence was amplified with oligonucleotides 3 and 4 from genomic DNA of *Aspergillus terreus* SBUG844 and cloned into the *Nco*I linearized TetOn_ble_pUC19 vector [12]. The 5′-end contained an overhang to the Tet-on system and at the 3′-end to the *Hind*III site of pUC19-ble. Linearized plasmid and gel-purified PCR products were mixed and assembled by in vitro recombination using the InFusion HD cloning kit (Takara/Clonetech) resulting in plasmid Tet-on:*terR*_ble_pUC19. The assembled plasmid was amplified in *Escherichia coli* DH5α using Mix & Go competent cells (Zymo Research). Positive clones were selected by colony PCR using oligonucleotides 5 and 6. Plasmids were isolated by use of the NucleoSpin Plasmid Miniprep kit (Macherey-Nagel) and correct assembly was confirmed by restriction analyses. The plasmid was used for transformation of *A. niger* A1144 and phleomycin resistant transformants were checked for single copy integration of the construct by Southern blot analysis using a dig labelled probe amplified with oligonucleotides 7 and 8. Transformant ATNT16 was selected for subsequent studies.

Generation of *lacZ* reporter and aspulvinone E synthetase gene expressing strains

To generate the *lacZ* reporter strains in the ATNT16 background, plasmid *hph_tdTomato*:*lacZ*:*trpC*^{T}_pJET1.2 [18] containing the *lacZ* reporter under control of the *terA* promoter and a hygromycin B resistance cassette were used for transformation of ATNT16. A fusion of the Tet-on transactivator with the *lacZ* gene was assembled in the *Pst*I restricted pUC19_*ptrA* plasmid [35] containing the pyrithiamine resistance cassette as selectable marker. Tet-on was amplified from plasmid pFW22.1 with oligonucleotide 9 containing an overhang to the *Pst*I site of pUC19_*ptrA* and oligonucleotide 10. The *lacZ* gene including a *trpC* terminator was amplified from plasmid *hph_tdTomato*:*lacZ*:*trpC*^{T}_pJET1.2 with oligonucleotide 11 containing an overhang to the 3′-end of Tet-on and oligonucleotide 12 with overhang to the *Pst*I site of pUC19_*ptrA*. Linearized plasmid and the two gel-purified PCR fragments

were mixed, assembled by in vitro recombination and transferred to *E. coli* as described above. Positive clones were selected by colony PCR using oligonucleotides 13 and 14. Isolated plasmids were checked by restriction analyses and used for transformation of A1144. Genomic DNA of ATNT16 and A1144 transformants was restricted with *Sma*I and analysed by Southern blot with a probe against the *lacZ* gene (oligonucleotides 13 and 14). At least three strains with a single copy integration of the reporter construct were used for expression analyses. For expression of the *A. terreus* aspulvinone E synthetase gene *melA* in the ATNT16 background, plasmid his_*melA*-SM-Xpress [19] was used as it contains a fusion of the *terA* promoter with the *melA* gene. The phleomycin resistance cassette of this plasmid was excised by *Not*I restriction and replaced by the pyrithiamine resistance cassette (*ptrA*) for transformation of ATNT16. Transformants were analysed by Southern blot with a probe against the *melA* gene (oligonucleotides 15 and 16) and strains with single copy integration were selected (Additional file 3).

Generation of model gene cluster expressing strains

Two different polycistronic expression constructs were generated for gene expression in ATNT16. The first construct contained the *melA* gene and a fusion of the *tyrP* gene and the gene coding for tdTomato. The *melA* and *tyrP* genes were separated by a P2A coding sequence. The *melA* gene was amplified from genomic DNA of *A. terreus* SBUG844 with oligonucleotide 17 that contained an overlap to the *Nco*I restricted SM-Xpress2 vector [19] and oligonucleotide 18 with an overhang coding for a the P2A sequence. The gene fusion of *tyrP* with the tdTomato gene was amplified from plasmid *tyrP:tdTomato*_SM-Xpress2 [19] with oligonucleotide 19 possessing an overhang to the P2A sequence and oligonucleotide 20 with an overlap to the *Nco*I restricted SM-Xpress2 vector. In the second construct all three genes were separated by P2A sequences. The *melA* gene was amplified with the same oligonucleotides as for the first construct. The *tyrP* gene was amplified from genomic DNA of *A. terreus* SBUG844 with oligonucleotide 19 containing the P2A sequence overhang towards *melA* and oligonucleotide 21 with a P2A sequence overhang towards the tdTomato gene. Finally, the tdTomato gene was amplified from plasmid *tyrP:tdTomato*_SM-Xpress2 with oligonucleotide 22 possessing the complementary overhang to the *tyrP* 3′ P2A sequence and oligonucleotide 20 with a compatible overhang to the SM-Xpress2 vector. Constructs were assembled by in vitro recombination and transferred to *E. coli* DH5α. Clones were checked by colony PCR using oligonucleotides 23 and 24 to test for correct gene assembly. Plasmid DNA was isolated and used for transformation

of ATNT16 with hygromycin B as selectable marker. Transformants were analysed by Southern blot with a probe against the *melA* gene (oligonucleotides 15 and 16) and strains with single copy integration were analysed further (Additional file 4).

β-Galactosidase reporter assays

To study β-galactosidase reporter activities, fungi were inoculated at 2×10^6 conidia per ml and grown for 24 h in 100 mM glucose containing minimal media with 20 mM glutamine as nitrogen source and 1% talc to avoid the formation of cell pellets. Mycelia were harvested over Miracloth, pressed dry and frozen in liquid nitrogen. Mycelia were ground to a fine powder under liquid nitrogen and suspended in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.7% β-mercaptoethanol). After centrifugation for 5 min at $16,000 \times g$ and 4 °C the cell-free supernatant was removed and used for determination of β-galactosidase activity as previously described [36] using *o*-nitrophenyl-β-D-galactopyranoside (ONPG; $\varepsilon = 3.5 \text{ mM}^{-1} \text{ cm}^{-1}$) as substrate. Protein concentrations were determined by using the Bradford Protein Assay (BioRad) with bovine serum albumin as standard. All spectrophotometric assays were carried out using an Evolution 220 UV–VIS spectrophotometer (ThermoFisher Scientific). From each construct three independent strains were grown in biological duplicates and activity determinations were made in technical duplicates.

Analysis of aspulvinone E production

To test production of aspulvinone E in ATNT16 strains carrying a single copy integration of the P*terA:melA* construct, GG10 medium was inoculated with 1×10^6 conidia/ml and one culture was supplemented with 15 μg/ml of doxycycline, whereas the other was left untreated. Incubation was performed for 48 h at 28 °C on a rotary shaker at 150 rpm. Mycelium was removed by filtration over Miracloth and the culture filtrate was extracted twice with an equal volume of ethyl acetate. After evaporation of the solvent under reduced pressure the residue was solved in 1 ml of methanol and subjected to HPLC analysis using a Dionex UltiMate3000 (ThermoFisher Scientific) and Eclipse XDB-C18 column, 5 μm, 4.6 × 150 mm (Agilent) that was kept at 40 °C. A binary solvent system consisting of water acidified with 0.1% formic acid (solvent A) and methanol (solvent B) was applied. The following gradient at a flow rate of 1 ml/ min was used: 0–0.5 min 10% B, 0.5–15 min 10–90% B, 15–17 min 90% B, 17–17.5 min 90–100% B, 17.5–22 min 100% B, 22–23 min 100–10% B, 23–25 min 10% B. An authentic sample containing a mixture of aspulvinone E and isoaspulvinone E served as reference.

Semiquantitative RT-PCR analyses

To analyse transcription of genes from the polycistronic messenger RNA of the ATNT16 P2A_P2A strain RNA was isolated using the MasterPure-Yeast RNA Purification Kit (Epicentre) from mycelium cultivated for 24 h in the absence or presence of 15 μg/ml doxycycline or pre-grown for 18 h without doxycycline and further cultivated for 6 h after addition of doxycycline. After a DNase treatment (TURBO DNase; ThermoFisher) RNA was transcribed into cDNA as previously described [19]. For normalisation of cDNA levels in the respective samples, serial dilutions were used for amplification of the *A. niger* actin gene using oligonucleotides 25 and 26. These primers span an intron region, which allows visualisation of a band shift from cDNA compared to genomic DNA (gDNA) and confirms the absence of contaminating gDNA in cDNA samples. For amplification of the *melA* gene oligonucleotides 15 and 16, for *tyrP* oligonucleotides 27 and 28 and for the tdTomato gene oligonucleotides 20 and 29 were used. PCRs of 30 cycles were performed in a SpeedCycler^2 (Analytik Jena) using Phire Hot Start II polymerase (Thermo Scientific).

Fluorescence microscopy

Fluorescence microscopy was performed as described previously [19] with some minor modifications. Strains were spotted on GG10 agar plates and pre-grown at 28 °C for one day, after which GG10 agarose coated coverslips containing 10 μg/ml doxycycline were placed next to the growing colony. 12 to 16 h later the coverslips were removed and placed on an object slide, overlaid with a droplet of mounting solution containing DAPI (ProLong Gold Antifade with DAPI, Thermo Scientific) and covered with a large coverslip. A GXML3201LED microscope (GX microscopes) was used for picture acquisition. Overlays of images were assembled by using the GIMP 2 software.

Statistical analyses

Comparison of expression levels from β-galactosidase activity determinations were analysed using GraphPad Prism (GraphPad Software) by applying multiple t-tests using the Holm-Sidak method, with α = 0.05. Each row was analysed individually, without assuming a consistent standard deviation.

Additional files

Additional file 1. Southern blot analyses and plasmid map of construct used for generation of ATNT strains. (**A**) Southern blot for identification of single copy integration strains. A digoxygenin labelled probe was used for hybridisation. Plasmid control and genomic DNA of parental strains and transformants were restricted with *Apa*I, which cuts once in the respective plasmid. The transformant used in subsequent analyses is numbered. (**B**) Plasmid map of the transformation construct. Position of oligonucleotides used in this study (P + number) as well as the position of the probe generated for Southern blot analysis and position of the restriction enzyme are shown. *ble* = phleomycin resistance cassette. TetOn = Tet-on promoter system. terR = *terR* gene including its native terminator sequence.

Additional file 2. Southern blot analyses and plasmid maps of constructs used for generation of *lacZ* reporter strains. (**A, C**) Southern blot for identification of single copy integration strains. Digoxigenin labelled probes were used for hybridisation. Transformants used in subsequent analyses are numbered. (**A**) A1144 strains with integration of the *tet-on:lacZ* construct. Plasmid control and genomic DNA of parental strains and transformants were restricted with *Ahd*I, which cuts once in the respective plasmid. (**C**) ATNT16 strain transformed with the P*terA:lacZ* construct. Plasmid control and genomic DNA of parental strains and transformants were restricted with *Hind*III, which cuts once in the respective plasmid. (**B, D**) Plasmid maps of the transformation constructs. Position of oligonucleotides used in this study (P + number) as well as the position of the probe generated for Southern blot analyses and position of the restriction enzyme are shown. ptrA = pyrithiamine resistance cassette. hph = hygromycin resistance cassette. PterA = *terA* promoter from *Aspergillus terreus*. lacZ = β-galactosidase gene from *Escherichia coli*. TtrpC = *trpC* terminator sequence from *Aspergillus terreus*.

Additional file 3. Southern blot analysis and plasmid map of construct used for generation of ATNT *melA* strains. (**A**) Southern blot for identification of single copy integration strains. A digoxygenin labelled probe was used for hybridisation. Plasmid control and genomic DNA of parental strains and transformants were restricted with *Bgl*II, which cuts once in the respective plasmid. The transformant used in subsequent analyses is numbered. (**B**) Plasmid map of the transformation construct. Position of oligonucleotides used in this study (P + number) as well as the position of the probe generated for Southern blot analysis and position of the restriction enzyme are shown. ptrA = pyrithiamine resistance cassette. PterA = *terA* promoter from *Aspergillus terreus*. TtrpC = *trpC* terminator sequence from *Aspergillus terreus*. melA = Aspulvinone E synthetase gene *melA* from *Aspergillus terreus*.

Additional file 4. Southern blot analyses and plasmid maps of constructs used for generation of ATNT16 P2A_P2A and P2A strains. (**A**) Southern blot for identification of single copy integration strains. A digoxygenin labelled probe was used for hybridisation. Plasmid control and genomic DNA of parental strains and transformants were restricted with *Xba*I, which cuts once in the respective plasmids. Transformants used in subsequent analyses are numbered. (**B, C**) Plasmid maps of the transformation constructs. Position of oligonucleotides used in this study (P + number) as well as the position of the probe generated for Southern blot analysis and position of the restriction enzyme are shown. *hph* = hygromycin B resistance cassette. PterA = *terA* promoter from *Aspergillus terreus*. TtrpC = *trpC* terminator sequence from *Aspergillus terreus*. melA = Aspulvinone E synthetase gene *melA* from *Aspergillus terreus*. tyrP = tyrosinase gene *tyrP* from *Aspergillus terreus*. tdTomato = codon optimised *tdTomato* gene. tyrP:tdTom = fusion of *tyrP* and *tdTomato* genes. P2A = sequence coding for the 2A peptide from porcine teschovirus-1.

Additional file 5: Table 1. Oligonucleotides used in this study.

References

1. Bok JW, Hoffmeister D, Maggio-Hall LA, Murillo R, Glasner JD, Keller NP. Genomic mining for *Aspergillus* natural products. Chem Biol. 2006;13:31–7.

2. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, et al. MycoCosm portal: gearing up for 1000 fungal genomes. Nucl Acids Res. 2014;42:D699–704.

3. de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, Anderluh G, Asadollahi M, Askin M, Barry K, et al. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol. 2017;18:28.

4. Wasil Z, Pahirulzaman KAK, Butts C, Simpson TJ, Lazarus CM, Cox RJ. One pathway, many compounds: heterologous expression of a fungal biosynthetic pathway reveals its intrinsic potential for diversity. Chem Sci. 2013;4:3845–56.

5. Gressler M, Meyer F, Heine D, Hortschansky P, Hertweck C, Brock M. Phytotoxin production in *Aspergillus terreus* is regulated by independent environmental signals. eLife. 2015;4:e07861.

6. Alberti F, Foster GD, Bailey AM. Natural products from filamentous fungi and production by heterologous expression. Appl Microbiol Biotechnol. 2017;101:493–500.

7. Anyaogu DC, Mortensen UH. Heterologous production of fungal secondary metabolites in Aspergilli. Front Microbiol. 2015;6:77.

8. Gonzalez-Menendez V, Perez-Bonilla M, Perez-Victoria I, Martin J, Munoz F, Reyes F, Tormo JR, Genilloud O. Multicomponent analysis of the differential induction of secondary metabolite profiles in fungal endophytes. Molecules. 2016;21(2):234.

9. Netzker T, Fischer J, Weber J, Mattern DJ, Konig CC, Valiante V, Schroeckh V, Brakhage AA. Microbial communication leading to the activation of silent fungal secondary metabolite gene clusters. Front Microbiol. 2015;6:299.

10. Brakhage AA, Schroeckh V. Fungal secondary metabolites–strategies to activate silent gene clusters. Fungal Genet Biol. 2011;48:15–22.

11. Brakhage AA. Regulation of fungal secondary metabolism. Nat Rev Microbiol. 2013;11:21–32.

12. Gressler M, Zaehle C, Scherlach K, Hertweck C, Brock M. Multifactorial induction of an orphan PKS-NRPS gene cluster in *Aspergillus terreus*. Chem Biol. 2011;18:198–209.

13. Clevenger KD, Bok JW, Ye R, Miley GP, Verdan MH, Velk T, Chen C, Yang K, Robey MT, Gao P, et al. A scalable platform to identify fungal secondary metabolites and their gene clusters. Nat Chem Biol. 2017;13:895–901.

14. Bok JW, Keller NP. LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell. 2004;3:527–35.

15. Sakai K, Kinoshita H, Nihira T. Heterologous expression system in *Aspergillus oryzae* for fungal biosynthetic gene clusters of secondary metabolites. Appl Microbiol Biotechnol. 2012;93:2011–22.

16. Yeh HH, Ahuja M, Chiang YM, Oakley CE, Moore S, Yoon O, Hajovsky H, Bok JW, Keller NP, Wang CC, Oakley BR. Resistance gene-guided genome mining: serial promoter exchanges in *Aspergillus nidulans* reveal the biosynthetic pathway for fellutamide B, a proteasome inhibitor. ACS Chem Biol. 2016;11:2275–84.

17. Fleissner A, Dersch P. Expression and export: recombinant protein production systems for *Aspergillus*. Appl Microbiol Biotechnol. 2010;87:1255–70.

18. Gressler M, Hortschansky P, Geib E, Brock M. A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. Front Microbiol. 2015;6:184.

19. Geib E, Gressler M, Viedniernikova I, Hillmann F, Jacobsen ID, Nietzsche S, Hertweck C, Brock M. A non-canonical melanin biosynthesis pathway protects *Aspergillus terreus* conidia from environmental stress. Cell Chem Biol. 2016;23:587–97.

20. Brandenburger E, Gressler M, Leonhardt R, Lackner G, Habel A, Hertweck C, Brock M, Hoffmeister D. A highly conserved basidiomycete peptide synthetase produces a trimeric hydroxamate siderophore. Appl Environ Microbiol. 2017;83:e01478-17.

21. Meyer V, Wanka F, van Gent J, Arentshorst M, van den Hondel CA, Ram AF. Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. Appl Environ Microbiol. 2011;77:2975–83.

22. Wanka F, Cairns T, Boecker S, Berens C, Happel A, Zheng X, Sun J, Krappmann S, Meyer V. Tet-on, or Tet-off, that is the question: advanced conditional gene expression in *Aspergillus*. Fungal Genet Biol. 2016;89:72–83.

23. Tang X, Liu X, Tao G, Qin M, Yin G, Suo J, Suo X. "Self-cleaving" 2A peptide from porcine teschovirus-1 mediates cleavage of dual fluorescent proteins in transgenic *Eimeria tenella*. Vet Res. 2016;47:68.

24. Kim JH, Lee SR, Li LH, Park HJ, Park JH, Lee KY, Kim MK, Shin BA, Choi SY. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLoS One. 2011;6:e18556.

25. Beekwilder J, van Rossum HM, Koopman F, Sonntag F, Buchhaupt M, Schrader J, Hall RD, Bosch D, Pronk JT, van Maris AJ, Daran JM. Polycistronic expression of a beta-carotene biosynthetic pathway in *Saccharomyces cerevisiae* coupled to beta-ionone production. J Biotechnol. 2014;192(Pt B):383–92.

26. Efimova VS, Isaeva LV, Makeeva DS, Rubtsov MA, Novikova LA. Expression of cholesterol hydroxylase/lyase system proteins in yeast *S. cerevisiae* cells as a self-processing polyprotein. Mol Biotechnol. 2017;59:394–406.

27. de Amorim Araujo J, Ferreira TC, Rubini MR, Duran AG, De Marco JL, de Moraes LM, Torres FA. Coexpression of cellulases in *Pichia pastoris* as a self-processing protein fusion. AMB Express. 2015;5:84.

28. Subramanian V, Schuster LA, Moore KT, Taylor LE 2nd, Baker JO, Vander Wall TA, Linger JG, Himmel ME, Decker SR. A versatile 2A peptide-based bicistronic protein expressing platform for the industrial cellulase producing fungus, *Trichoderma reesei*. Biotechnol Biofuels. 2017;10:34.

29. Unkles SE, Valiante V, Mattern DJ, Brakhage AA. Synthetic biology tools for bioprospecting of natural products in eukaryotes. Chem Biol. 2014;21:502–8.

30. Schuetze T, Meyer V. Polycistronic gene expression in *Aspergillus niger*. Microb Cell Fact. 2017;16:162.

31. Liu Z, Chen O, Wall JBJ, Zheng M, Zhou Y, Wang L, Ruth Vaseghi H, Qian L, Liu J. Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. Sci Rep. 2017;7:2193.

32. Driouch H, Sommer B, Wittmann C. Morphology engineering of *Aspergillus niger* for improved enzyme production. Biotechnol Bioeng. 2010;105:1058–68.

33. Richter L, Wanka F, Boecker S, Storm D, Kurt T, Vural O, Sussmuth R, Meyer V. Engineering of *Aspergillus niger* for the production of secondary metabolites. Fungal Biol Biotechnol. 2014;1:4.

34. Grundlinger M, Yasmin S, Lechner BE, Geley S, Schrettl M, Hynes M, Haas H. Fungal siderophore biosynthesis is partially localized in peroxisomes. Mol Microbiol. 2013;88:862–75.

35. Fleck CB, Brock M. *Aspergillus fumigatus* catalytic glucokinase and hexokinase: expression analysis and importance for germination, growth, and conidiation. Eukaryot Cell. 2010;9:1120–35.

36. Ebel F, Schwienbacher M, Beyer J, Heesemann J, Brakhage AA, Brock M. Analysis of the regulation, expression, and localisation of the isocitrate lyase from *Aspergillus fumigatus*, a potential target for antifungal drug development. Fungal Genet Biol. 2006;43:476–89.

Production of *Aspergillus niger* biomass on sugarcane distillery wastewater: physiological aspects and potential for biodiesel production

Graziella Chuppa-Tostain[1,2], Julien Hoarau[1], Marie Watson[1], Laetitia Adelard[2], Alain Shum Cheong Sing[1], Yanis Caro[1,6], Isabelle Grondin[1], Isabelle Bourven[3], Jean-Marie Francois[4], Elisabeth Girbal-Neuhauser[5] and Thomas Petit[1,6]*

Abstract

Background: Sugarcane distillery waste water (SDW) or vinasse is the residual liquid waste generated during sugarcane molasses fermentation and alcohol distillation. Worldwide, this effluent is responsible for serious environmental issues. In Reunion Island, between 100 and 200 thousand tons of SDW are produced each year by the three local distilleries. In this study, the potential of *Aspergillus niger* to reduce the pollution load of SDW and to produce interesting metabolites has been investigated.

Results: The fungal biomass yield was 35 g L^{-1} corresponding to a yield of 0.47 g of biomass/g of vinasse without nutrient complementation. Analysis of sugar consumption indicated that mono-carbohydrates were initially released from residual polysaccharides and then gradually consumed until complete exhaustion. The high biomass yield likely arises from polysaccharides that are hydrolysed prior to be assimilated as monosaccharides and from organic acids and other complex compounds that provided additional C-sources for growth. Comparison of the size exclusion chromatography profiles of raw and pre-treated vinasse confirmed the conversion of humic- and/or phenolic-like molecules into protein-like metabolites. As a consequence, chemical oxygen demand of vinasse decreased by 53%. Interestingly, analysis of intracellular lipids of the biomass revealed high content in oleic acid and physical properties relevant for biodiesel application.

Conclusions: The soft-rot fungus *A. niger* demonstrated a great ability to grow on vinasse and to degrade this complex and hostile medium. The high biomass production is accompanied by a utilization of carbon sources like residual carbohydrates, organic acids and more complex molecules such as melanoidins. We also showed that intracellular lipids from fungal biomass can efficiently be exploited into biodiesel.

Keywords: Sugarcane distillery wastewater, Vinasse, Distillery spent wash, *Aspergillus niger*, Biomass production, Bioremediation, Biodiesel, Lipids

*Correspondence: thomas.petit@univ-reunion.fr
[1] Antenne sud du laboratoire de chimie des Substances Naturelles et des Sciences des Aliments (LCSNSA), EA 2212, Université de la Réunion, UFR des Sciences et Technologies, 15 Avenue René Cassin, CS 92003, 97744 Saint-Denis Cedex 9, France
Full list of author information is available at the end of the article

Background

Sugarcane molasses fermentation and distillation into rum lead to the production of wastewater called stillage, vinasse, distillery wastewater or distillery spent wash. Every produced litre of ethanol brings about from 10 to 18 litres of sugarcane distillery wastewater (SDW) depending on distillation process and waste treatment [1]. SDW is a dark brown effluent characterized by a specific obnoxious odour, a high chemical oxygen demand (COD) and a total organic carbon (TOC) that can reach up to 120 and 17 g L^{-1} respectively [2, 3]. According to Wilkie et al. [4], COD is 4–5 times higher in sugarcane molasse stillage as compared to sugarcane juice stillage. Depending on the sugarcane origin and the industrial process for ethanol production, intrinsic composition of SDW can vary significantly. They generally have acidic pH (from 3.8 to 5) due to the presence of organic acids produced by the yeasts during the alcoholic fermentation process [5]. A high mineral load was also reported due to the presence of sulphur, potassium, phosphate, calcium and sodium [6, 7]. The high organic load of SDW is mainly composed of melanoidins which are produced through Maillard reactions between sugars and proteins and caramels from overheated sugars that are responsible for their colour and odour. Vinasse also contains other refractory materials such as phenolic compounds, anthocyanins, tannins and furfurans (for example hydroxyl methyl furfural) which can reach up to 10 g L^{-1} [8–10]. The colloidal nature of caramels makes them resistant to decomposition and toxic to microflora [11]. SDW also contains residual sugars and soluble proteins generated by the fermenting yeasts [12].

All these characteristics combined with the high volume of SDW produced worldwide cause significant environmental issues. Over the last decades and due to their high inorganic loads, SDW have been widely used as agricultural fertilizer [13] but spreading is made now statutory difficult due to their low pH, dark colour and chemical content which may be responsible for groundwater contamination and soil compaction [14]. Their high polluting loads lead to a modification of the soil composition and can cause eutrophication of the waterways because of the presence of proteins residues and furfurals [4, 15]. Moreover, melanoidins cause reduction of sunlight penetration, of photosynthetic activity and of dissolved oxygen concentration in natural aqueous environment, whereas on land, they cause reduction of soil alkalinity and inhibition of seed germination. In consequence, phenolic compounds and melanoidins may inhibit the activity of microorganisms contained in soils and aquatic environments [9, 10, 15].

Several methods have been described in literature for the use and disposal of SDW [10, 15, 16]. Among them, aerobic treatment of SDW has been proposed for decolourisation and COD reduction purposes. A number of microorganisms, such as yeast and fungi were found to be able to degrade melanoidins and to significantly decrease the COD vinasse [10]. Preliminary experiments performed in the lab (unpublished data) showed that only a few molds are capable of growing on crude SDW, such as *Aspergillus* strains and anamorphs. *Aspergillus niger* is able to grow on a large variety of substrates, a wide range of temperatures (6–47 °C) and pH (1.4–9.8), explaining the ubiquitous occurrence of this species that is encountered with a higher frequency in warm and humid environments [17]. *A. niger* is also known to be a good producer of extracellular enzymes with significant industrial importance, including amylases, proteases, pectinases, lipases as well as valuable molecules with industrial interest such as citric, oxalic or gluconic acids [18, 19]. *A. niger* is also used for organic waste enhancement [20] and its capacity to grow on diluted or supplemented SDW was observed [21–23]. However, the physiological growth characteristics of this micro-organism cultured in crude sugarcane distillery spent wash has not yet been reported. In addition, bioremediation and potential valorisation of crude SDW were estimated through the production of *A. niger* biomass as a valuable source for biofuel.

Methods

Fungal strains, growth conditions and culture media

The strain used in this study was *Aspergillus niger* MUCL 28820 from BCCM (Brussels, Belgium) strain collection. The strain was maintained routinely on potato dextrose agar plates (PDA). A suspension of *A. niger* spores was prepared as follow: spores, grown on PDA and incubated at 28 °C for 72 h, were harvested using a glass loop and suspended in sterile physiological water (NaCl 0.8%). Cellular concentration was calculated using a Thoma-Zeiss counting chamber. Growth experiments were performed during 10 days, after inoculation with 100 µL of spore suspension. Ten flasks containing 50 mL of sterile SDW liquid medium at a starting concentration of 10^5 spores mL^{-1} were plugged with sterile cotton carded and placed on a rotary shaker at 150 rpm at 28 °C. Assays were performed in three independent biological experiments. Every day, at the same hour, the biomass of three flasks was harvested by filtration for further analysis (see below) and this was repeated until day 10.

SDW medium was prepared as follows: raw SDW (85 °C) was harvested in decontaminated barrel directly from the output of the distillation column from distillery "Rivière du Mât" (Saint-Benoit, Reunion Island) and cooled to room temperature. SDW from the distillery still contains the residual inactivated yeast biomass used

during alcohol fermentation. SDW samples were harvested during the sugar production period (i.e. between July and December) in 2012 and in 2014 and were frozen and stored at − 20 °C until use. For each experiment, a new batch of frozen DSW was thawed and then sterilized by autoclaving for 20 min at 121 °C. Such autoclaved DSW medium was microbiologically stable over time (Additional file 1).

After filtration of 50 mL of SDW (through a cellulose filter paper Whatman No. 1—porosity 11 μm), the filtrates and filters containing the total suspended solids (see example on Fig. 1) were both dried during 24 h in an oven at 105 °C. The obtained dried masses were reported to 50 mL allowing to determinate the corresponding concentrations in total dissolved solids (TDS) and total suspended solids (TSS), respectively. Mineral matters present in the SDW filtrates were measured according to Analytical Procedure of National Renewable Energy Laboratory by incineration of 10 mL of SDW filtrates at 550 °C for 3 h [24]. The pH of the SDW filtrate was measured using a pH-meter Denver Instrument. Soluble COD and Total Nitrogen (TN) determinations were carried out on the SDW filtrate using a DR 2800 spectrophotometer (Hach Lange, Dusseldorf) and the appropriate analytical kits [25]. Samples were adequately diluted with sterile deionized water and analysed according to manufacturer's instructions. SDW filtrates were diluted to 1/100 and their optical density was measured at 475 nm using a spectrophotometer Genesys 10 UV Deionised water was used as blank.

Lipid accumulation medium (LAM) contained 30 g L^{-1} glucose, 1.5 g L^{-1} yeast extract, 0.5 g L^{-1} NH_4Cl, 5.0 g L^{-1} Na_2HPO_4 ($12H_2O$), 7.0 g L^{-1} KH_2PO_4, 1.5 g L^{-1} $MgSO_4$ ($7H_2O$), 0.1 g L^{-1} $CaCl_2$ ($2H_2O$), 0.01 g L^{-1} $ZnSO_4$ ($7H_2O$), 0.08 g L^{-1} $FeCl_3$ ($6H_2O$), 0.1 mg L^{-1} $CuSO_4$ ($5H_2O$), 0.1 mg L^{-1} $Co(NO_3)_2$ ($6H_2O$), 0.1 mg L^{-1} $MnSO_4$ ($5H_2O$) and pH was adjusted to 5.5 according to [26].

Fungal biomass determination

The concentration of total suspended solids (TSS) in the broth medium of each culture flask was determined by filtration of 50 mL of SDW (treated or not b*y A. niger*) through a cellulose filter paper Whatman No. 1 (porosity 11 μm) previously dried for 24 h at 105 °C. The insoluble suspended solids kept on the filter (see example on Fig. 1) were dried during 24 h in an oven at 105 °C and the obtained dry mass was weighed to provide TSS concentration. Therefore, TSS contained the fungal biomass produced during growth of *A. niger* as well as the initial suspended yeast biomass contained in raw SDW. Fungal biomass was thus estimated by subtracting the total suspended matter of raw SDW to the total mass harvested on the filter.

Analytical methods
Determination of carbohydrates and organic acids from filtrates of crude and pre-treated SDW
The carbohydrate concentration of the filtrates collected from crude SDW and SDW treated with *A. niger,* were analysed by High-Pressure Liquid Chromatography (HPLC) (Dionex Ultimate 3000) using an Evaporative Light Scattering (ELS) detector (VARIAN) and a Hi-plex Ca column (Varian, C18 bound—7.7 mm of diameter × 300 mm of length). A mobile phase of ultrapure water with a flow of 0.4 mL min^{-1} was used. The oven temperature was programmed at 80 °C. Alternatively, High-Pressure Ion Chromatography (HPIC) (Dionex) using a pulsed amperometric detector and a CarboPack PA1 column was used. A mobile phase of NaOH (150 mM) at 1.5 mL min^{-1} was used at an oven
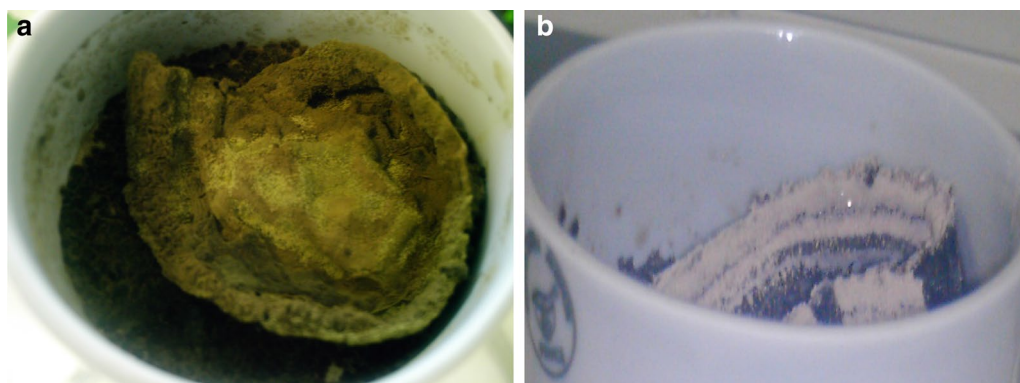

Fig. 1 Photographs of fungal biomass (**a**, **b**) produced during growth on SDW (day 10). Filtrated *A. niger* cell pellets were harvested under vacuum on Whatman No. 1 paper using Büchner funnel (see "Methods" section)

temperature of 30 °C. Analysis of organic acids was also performed by HPLC (Dionex Ultimate 3000), using a UV detector at 214 nm and an OA Acclaim column (Varian, Silica, C18 bound, reverse phase, 4.6 mm of diameter × 150 mm of length). The mobile phase was composed of 100 mM Na_2SO_4 set at pH of 2.65 with methanesulfonic acid (Sigma-Aldrich, CAS number 75-75-2) and the flow rate was 0.6 mL min^{-1}. For all analyses, 20 µL of samples diluted 100-fold for organic acids and 50-fold for carbohydrates in water were injected using an automatic autosampler. The identification and the quantification of carbohydrates (mannitol, glucose, fructose, sucrose) and organic acids (itaconic acid, trans-aconitic acid, citric acid, isocitric acid, oxalic acid) were made by determination of retention time of the commercial standards and establishment of calibration curves using external standard method. Treatment of the results was done using *Chromeleon* 7.2 Chromatography Data System (Dionex).

SEC profiles obtained from filtrates of crude and pre-treated SDW

A 5-days fermented SDW was chosen for this experiment because at this stage of the growth (mid-exponential phase), most of the sugars and organic acids remains unchanged while biomass already reached more than 10 g L^{-1} DW, suggesting that others classes of molecules were used preponderantly for growth of the cells. The filtrates of crude SDW and 5-days treated SDW with *A. niger* were ten times diluted with phosphate buffer (pH 7.0) and filtered (through a 0.45 µm filter) before injection in a HPLC system Äkta-Purifier (GE Healthcare). As previously described by [27], two SEC columns were connected in series in order to obtain a wide resolving range: the Superdex peptide 10/300 GL column with a resolving range from 0.1 to 7 kDa was placed before the second Superdex 200 10/300 column (GE Healthcare) with a resolving range from 10 to 600 kDa. A similar volume of 0.1 ml of the two samples previously diluted in PBS was injected and elution of the molecules was performed at room temperature using a 50 mM potassium phosphate buffer (pH 7.0) as the mobile phase at a flow-rate of 0.4 mL min^{-1} and fractions of 2 mL each were collected. The Unicorn 5.1 software (GE Healthcare) delivered on the Akta purifier allows to either multiply or divide the chromatograms with a constant factor: depending on the total COD concentration of the sample, the obtained chromatogram can be thus standardized per mg of COD. Peak area integration of the standardized chromatograms was performed by the Unicorn 5.1 software. The SEC columns were calibrated for molecular weight determination using a mixture of standard proteins of known molecular weight between 12 and 669 kDa (HMW and LMW calibration kits, GE Healthcare). Calibration showed a linear relationship between the log of molecular weight (MW) and the elution volume (Ve) of the standards according to the following equation:

$$\mathrm{Log\,(MW)} = -0.1536\,\mathrm{Ve} + 8.5794\,(\mathbf{1}) \quad \left(\mathrm{R}^2 = 0.9976\right)$$

with MW expressed in Da and Ve in mL.

EEM profiles obtained from filtrates of crude and pre-treated SDW

A three-dimensional excitation emission matrix (3-D EEM) was determined on the SDW filtrates (raw or treated with *A. niger*) and on the SEC fractions, using a spectrofluorophotometer (Shimadzu RF-5301 PC) with a 150-W Xenon lamp as the excitation source. Excitation scans were performed from 220 to 450 nm at 10 nm increments; emission scans were collected from 220 to 500 nm. The fluorescence data was processed using the Panorama Fluorescence 3.1 software (LabCognition, Japan). Prior to measurements, fractions of SEC samples were diluted by 3–100 times using 50 mM phosphate buffer (pH 7.0 ± 0.1) to avoid fluorescence signal saturation. However, due to the impact of water noise, only emissions obtained at excitation wavelengths exceeding 275 nm were considered for a wavelength emission exceeding 375 nm. Gallic acid (Sigma), used as polyphenols standard [28] was also diluted in phosphate buffer for analysis. Fluorescence was measured using a 1.0 cm quartz cell.

Lipid extraction from *A. niger* biomass and conversion into biodiesel

Intracellular lipids were extracted using a pressurized liquid extraction method (PLE). 200 mg of lyophilized biomass was mixed with Fontainebleau sand to fill a 10 mL stainless steel vial suitable for PLE. The extraction was carried out using chloroform/methanol (2/1) at 100 °C during 10 min (three times), then 10 mL of water was added to the extract and thoroughly mixed. Two phases were obtained after overnight separation. The organic phase was dried over anhydrous $MgSO_4$, filtered and concentrated using a rotative evaporator. Finally, the concentrate was suspended in 3 mL $CHCl_3$, transferred to a pre-weighed bottle and evaporated overnight. The bottle was weighted to determine the mass of extracted lipids. Transesterification was performed according to a procedure described by [29]. Briefly, 5 mL of 2% H_2SO_4/CH_3OH (v/v) was added to the extracted lipids and the mixture was reflux heated at 70 °C during 1 h under constant stirring. The flasks were then cooled at room temperature. Next, 2 mL of hexane and 0.75 mL of distilled water were added to the flasks and mixed. The two

phases were allowed to separate and the upper hexane layer was recovered and dried over anhydrous magnesium sulphate.

Analysis of the fatty acid composition was carried out on a CP3800 Gas chromatograph (Varian) equipped with a SG BPX-70 capillary column (50 m × 0.22 mm × 0.25 μm) and a flame ionization detector. The operating conditions were 240 °C injector temperature, 260 °C detector temperature, 1.3 mL min^{-1} flow rate and oven temperature programmed from 120 to 230 °C at 3 °C min^{-1} then 230 °C for 17 min. 0.5 μL of transesterification product was injected and subjected to a split ratio of 5 at 0.5 min then 50 at 5 min. The percentage of the peak area was assumed to be the percentage content of the corresponding compounds.

Results and discussion
Physico-chemical characteristics of SDW from Reunion Island
Physico-chemical parameters of raw SDW
To characterize our raw materials, main physico-chemical parameters of the collected SDW samples were assayed. Results are presented in Table 1. pH value of raw SDW (4.6 pH units) was comparable to average pH values (3.8–4.6) reported by [30] for SDW from others countries. Similarly, COD (107 g L^{-1}) and TDS (114 g L^{-1}) of SDW from Reunion Island were in the same range of order than the one reported for SDW from different origins that ranged from 42 to 121 and from 80 to 100 g L^{-1} respectively, with the outstanding exception of TDS of Brazilian SDW that reached up to 152 g L^{-1} [30–32]. Chemical composition of SDW filtrate

Table 1 Comparison of physico-chemical parameters of raw SDW and treated SDW filtrates obtained after 10 days of aerobic fermentation by *A. niger*

Physico-chemical parameters	Laboratory data	
	Raw SDW (day 0)	Fermented SDW (day 10)
pH	4.6	5.4
COD (g L^{-1})	107	50
TDS (g L^{-1})	114 ± 12.8	89 ± 7.07
TSS[a] (g L^{-1})	8.13 ± 1.41	43.42 ± 1.2
TN (g L^{-1})	2.32	1.7
Ashes (g L^{-1})	38.5 ± 2.33	43.2 ± 1.94
C/N	11.8	11.3
OD$_{475nm}$	34.5	25.2

SDW was incubated aerobically during 10 days with *A. niger* as explained in Methods section

TDS total dissolved solids, *TSS* total suspended solids, *COD* chemical oxygen demand, *TN* total nitrogen, *C/N* carbon/nitrogen, *OD$_{475nm}$* optical density measured at 475 nm

[a] Except for TSS that were measured on insoluble suspended solids

showed a TN content of 2.32 g L^{-1} and a total mineral content of 38.5 g L^{-1}. The first parameter was globally in good agreement with data of SDW from different southern countries, i.e. 1.23–4.8 g TN L^{-1} whereas the latter was higher than literature data namely 10.7–28.9 g L^{-1} [30, 33]. Overall, these physico-chemical parameters confirmed that SDW from Reunion Island are industrial wastes with high polluting organic and mineral loads that can be responsible for dangerous environmental disorders.

Physico-chemical parameters of SDW after treatment with *A. niger*
To assess the bioremediation potential of *A. niger*, the physico-chemical parameters of SDW were measured 10 days after the inoculation of the fungal spores in SDW. As shown in Table 1, a pH increase (from 4.6 to 5.4) and a decrease in OD$_{475nm}$ (linked to decolourisation) were observed during aerobic fermentation of SDW. TDS were significantly reduced from 114 to 89 g L^{-1} and this essentially concerns organic matter reduction since the mineral load was not significantly modified. A reduction of COD and TN by 53 and 27% respectively were observed, indicating a significant decrease of the organic pollutant load of SDW. The pH increase could result from the degradation of organic substances with peptidic moieties or with amino group like humic substances, melanoidins, peptides or amino acids initially contained in SDW medium. The carbon/nitrogen (C/N) ratio remained globally unchanged indicating that the fertilizing potential of SDW remained the same after the fermentation process.

Bioremediation potential of *A. niger* on SDW was partially described in literature. A maximal colour elimination of 69% and a maximal COD removal of 75% were obtained when MgSO$_4$, KH$_2$PO$_4$, NH$_4$NO$_3$ and a carbon source were added to SDW [34]. Also, immobilized *A. niger* resulted in a 80% decolourisation of previously anaerobically biodigested SDW [35]. Finally, the observed COD and colour decrease suggested that refractory molecules like melanoidins and other aromatic compounds were hydrolysed into simple ones. Such hydrolysis of some refractory compounds may contribute to the strong decrease of the measured COD (− 53%) because of the release of acidic moieties impacting on the oxidation degree of the polymers. In this way, qualitative characteristics of organic matter from raw and pre-treated SDW were investigated.

Physiology of *A. niger* cultured on SDW
Global biomass production
Concomitantly to the modification of some physico-chemical parameters, important biomass production was

observed in the SDW medium after 10 days of *A. niger* aerobic growth (Fig. 1 and Table 1). Fungal growth was evaluated by measurement of the total suspended solids in the broth medium that reached 43.4 g L^{-1} after 10 days. Given that the residual yeast biomass contained in raw SDW amounted to 8.1 g L^{-1}, a net production of 35.3 g L^{-1} of fungal biomass in SDW after 10 fermentation days could then be estimated. In addition to carbon containing substrates, residual dead yeasts contained in raw SDW (inactivated during the distillation process and newly sterilised before use) are most likely to play a role during growth such as bringing important nitrogen source. Consequently, SDW was considered as an interesting growth medium for *A. niger* biomass production. In their study, Oshoma et al. [21] demonstrated that the final concentration of *A. niger* biomass could be increased from 1.63 to 2.75 g L^{-1} Dry weight (DW) after nitrogen supplementation of cassava whey by yeast extract (2 g L^{-1}). By comparison, growth of *A. niger* on SDW from Brazil distilleries in which the yeast biomass was removed led to a biomass production of 8–13 g L^{-1} DW [22]. Here, the biomass production was much higher since until 35 g L^{-1} of *A. niger* biomass can be produced after 10 days on raw sugarcane vinasse without any supplementation. Considering that total organic matter of raw vinasse corresponds to TDS without ashes (75.5 g L^{-1}), a high biomass yield of 0.47 g g^{-1} on initial organic compounds can be reached. This yield is similar to that obtained by [36] that investigate the capability of *A. niger* to utilize lignocellulose-derived compounds after thermochemical pre-treatment of spruce wood chips. However, because of the presence of fermentation inhibitors, the pre hydrolysate medium had to be diluted 2 or 4 times to allow *A. niger* growth with a maximal volumetric biomass yield of 7 g L^{-1} and a biomass yield on initial carbon source of 0.46 g g^{-1}.

Sugar consumption

To explain the physiological behaviour of *A. niger* on raw SDW, carbohydrate content was monitored in the medium during the 10 days of fermentation process (Fig. 2). During the first 48 h, residual concentration of glucose, fructose and mannitol was increased by a factor of 2 and a maximal concentration of 7 g L^{-1} of fructose, 1.6 g L^{-1} of glucose and 4 g L^{-1} of mannitol were measured 2 days after inoculation of the fungal spores in the SDW medium. In the meantime, total fungal biomass increased slightly up to 4.37 g L^{-1}. Accumulation of these monosaccharides in the early stage of the exponential growth phase strongly suggested that some complex compounds were readily released by hydrolytic enzymes secreted by *A. niger*. In a second period, from 48 h (day 2) to 192 h (day 8), fructose, glucose and mannitol were
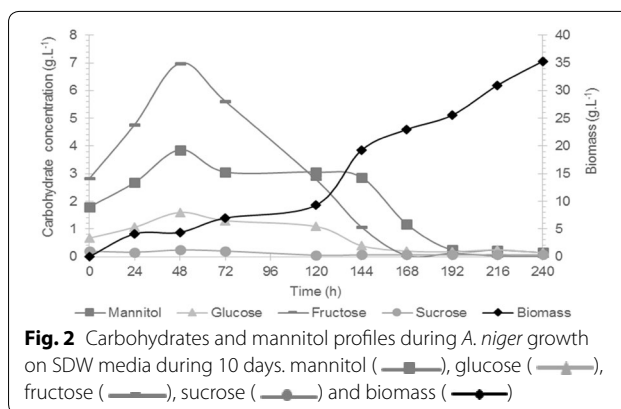
Fig. 2 Carbohydrates and mannitol profiles during *A. niger* growth on SDW media during 10 days. mannitol (—■—), glucose (—▲—), fructose (—▬—), sucrose (—●—) and biomass (—◆—)

gradually consumed, until complete exhaustion that occurred at 168 h (day 7) for glucose and fructose, and at 192 h (day 8) for mannitol. Fungal biomass that increased very weakly in the first period then suddenly increased during the period of sugars assimilation (from day 2 to 8, it increased from 0.8 to 25 g L^{-1}) and yet increased even after complete sugars exhaustion to reach 35.29 g L^{-1} at day 10 (Fig. 2). Jin et al. [37] also observed that monocarbohydrates were initially accumulated before being taken up for conversion into mycelial biomass (from 7.5 to 9.2 g L^{-1}) during aerobic fermentation of a raw starch processing wastewater, with either *Aspergillus oryzae* or *Rhizopus oligosporus*. This accumulation is most likely occurring when the rate of complex polymers hydrolysis is higher than the rate of carbohydrate uptake for cell growth.

When looking more carefully at the biomass production profile (Fig. 2), initial growth occurring during the first 120 h (day 5) did not appear to occur exponentially, but rather linearly. This observation would strengthen the hypothesis that the growth is mainly limited by the availability of fermentable sugars which are slowly and linearly produced through the activity of specific hydrolases from *A. niger* acting on complex polymers.

Organic acids utilization

It is known that SDW naturally contains large amount of organic acids [5]. Some of them were assayed in raw SDW filtrate and concentration of 5.7 ± 0.51 g L^{-1} for trans-aconitic acid, 2.8 ± 0.76 g L^{-1} for citric acid, 2.5 ± 0.47 g L^{-1} for isocitric acid, 0.7 ± 0.25 g L^{-1} for itaconic acid and 0.6 ± 0.18 g L^{-1} for oxalic acid were measured (Table 2). When sugars are being consumed by the cells, the concentration of itaconic, isocitric and oxalic acids tended to increase in the culture medium (+ 26, + 12 and + 136% respectively). An inverse relationship between consumption of sugars and organic acid production was already observed by [38] who reported that the

Table 2 Concentration (g L^{-1}) of organic acids and pH measured in SDW filtrates after 0, 7 and 10 days of aerobic fermentation by *A. niger*

Organic acids	Concentrations			
	Day 0	Day 5	Day 7	Day 10
Itaconic acid	0.70 ± 0.25	0.63 ± 0.26	0.88 ± 0.22	0.87 ± 0.28
Trans-aconitic acid	5.71 ± 0.51	4.54 ± 1.4	4.32 ± 0.77	1.59 ± 0.37
Citric acid	2.84 ± 0.76	3.37 ± 0.28	1.36 ± 0.91	Bd
Isocitric acid	2.47 ± 0.46	2.52 ± 0.17	2.77 ± 0.41	Bd
Oxalic acid	0.61 ± 0.18	0.52 ± 0.14	1.45 ± 0.53	0.38 ± 0.14
pH	4.6 ± 0.1	5.07 ± 0.49	5.93 ± 1.21	5.37 ± 0.13

Each value is a mean of at least three independent experiments

Bd below detection level

maximum acid production was found for 6 days old *A. niger* cultures. Opposite tendency was noticed for citric and trans-aconitic acids that were slightly consumed during the first 7 days of growth. However, except for itaconic acid for which concentration remained stable, all organic acids were consumed partially or completely in the remaining 3 days (Table 2). These data indicated that organic acids were preferentially consumed during the second period of fermentation after complete exhaustion of monosaccharides. These carbon sources could be the result of hydrolysis of melanoidins, polyphenols or proteins present in crude SDW [39].

Taken together, these results showed that growth of *A. niger* on SDW is a complex process. Free carbohydrates initially present in the media (namely glucose, fructose and mannitol) and other fermentable sugars probably released from complex polymers through hydrolytic activity of the fungal enzymes, are first consumed during the early growth phase. When free sugars disappeared from the medium (after 7–8 days of culture), growth continued on the free organic acids accumulated in the medium as well as other sugars released by *A. niger* hydrolases.

SDW biochemical fingerprints
Global EEM profiles of raw and pre-treated SDW
Three-dimensional excitation emission matrixes (EEM) were determined on the SDW filtrates in order to detect potential modification of complex dissolved organic matter like melanoidins or phenolic acids during *A. niger* fermentation (Fig. 3). EEM is a widely used non-degradative method for qualitative characterization of the soluble substances of many effluents [43]. As proposed by [40], for typical wastewater spectra, the EEM can be divided in at least two regions: (1) the region with emission wavelength $\lambda_{Em} < 380$ nm which is associated with fluorescent molecules types A and B containing a limited number of aromatic rings like phenols, indole moiety and free tryptophan; and (2) the region with $\lambda_{Em} > 380$ nm which is associated with polycyclic aromatic fluorophores (types C and D) such as Humic acid, flavonoid and quinone. In addition, EEM of pure gallic acid was performed to locate more precisely its associated zone: EEM showed a fluorescent peak ($250 < \lambda_{Ex} < 275$ nm and
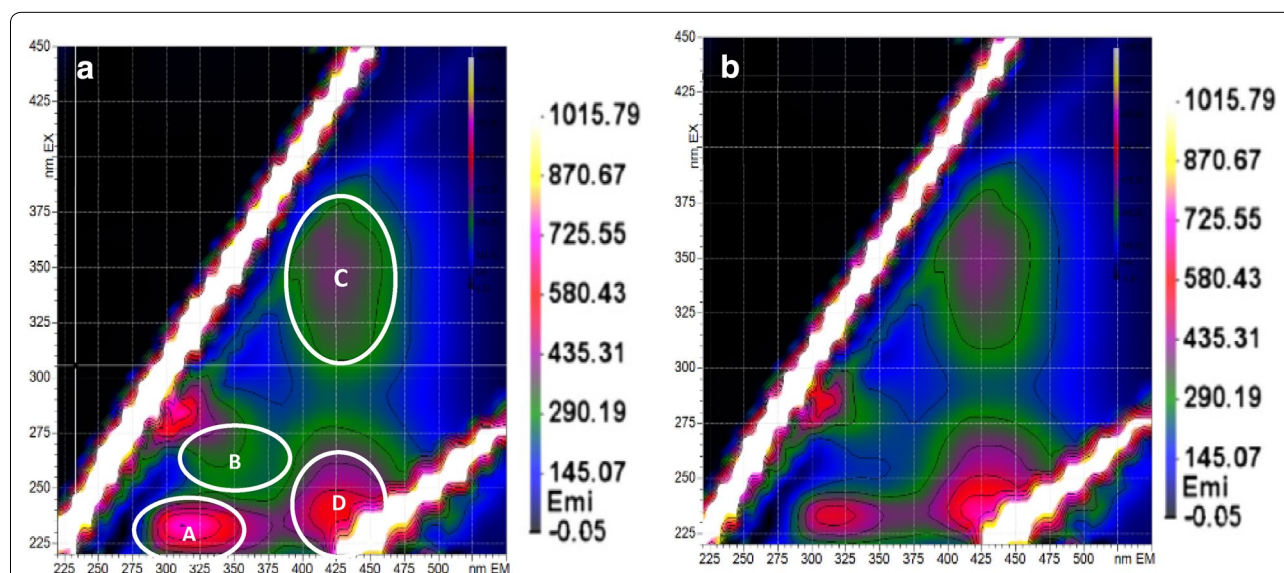

Fig. 3 Analysis of the fluorescent matter in raw SDW filtrate (**a**) and in SDW filtrate treated for 5 days with *A. niger* (**b**) according to classification provided for wastewater: peak (A) corresponds to protein-like (PN-like) substances [40] and peak (B) to phenolic acid-like (PA-like) compounds [41]; peaks (C) and (D) can be related to humic acid-like (HA-like) substances [42]

325 < λ_{Em} < 370 nm) that was included in the phenolic acid-like (PA-like) region (type B). The determination of PA-like area zone in SDW is in accordance with [44] who worked on partially degraded food waste and undigested dietary fibres.

Concerning the SDW media, both matrixes were composed by four peaks with similar excitation/emission wavelengths ($\lambda_{Ex}/\lambda_{Em}$) position and intensity (Fig. 3a, b): (1) for λ_{Em} < 380 nm, peak A and peak B were located in the regions corresponding to protein-like (PN-like) and phenolic acid-like (PA-like) compounds respectively [41] and peak A was much more intense than peak B (2) for λ_{Em} > 380 nm, peaks D and C were associated with quinine-like components and could be related to humic acid-like (HA-like) substances [42]. These results were in good accordance with the results obtained by [45] which highlighted these groups of fluorophores (A, B and C-D areas) in sugarcane vinasse. In this way, EEM determined on the soluble SDW fractions (treated or not) did not allow to clearly show EEM modifications pattern related to *A. niger* metabolism (four independent replicates were analysed; only one replicate was shown). This can be explained by the complexity of the SDW medium that contains highly fluorescent molecules possibly covering the detection of other ones. Moreover, only specific molecules with aromatic ring are detected by EEM.

EEM profiles after size fractionation of raw and fermented SDW

In order to evaluate whether the SDW has been altered by *A. niger* treatment, size fractionation of raw and 5-days fermented SDW was chosen to provide a synthetic view of their composition and size distribution. SEC chromatograms were first monitored by absorbance detection at 210 nm and 280 nm but raw and fermented SDW filtrates displayed similar profiles (data not shown). Regarding the EEM spectra of the two SDW samples (Fig. 3a, b), high fluorescence intensities could be noticed on the PN-like region (peak A). One common couple of wavelengths ($\lambda_{Ex}/\lambda_{Em}$ = 221/350 nm) that was previously described by [43] for detection of tryptophan containing PN-like molecules was then selected for SEC monitoring. Fluorescent compounds detected in this region (peak A) were reported by [46] as soluble microbial products associated to microbial activity or to cellular material.

Fractionation of the raw and fermented SDW filtrates was performed by SEC and could be divided in seven fractions from F1 to F7 corresponding to increasing elution volume and to decreasing apparent molecular size (Fig. 4a). Quantitative repartition of each fraction among all the eluted molecules was also evaluated after peak area integration (Table 3). It can be noticed that fermented SDW showed some early eluted molecules in the

Fig. 4 Size exclusion profiles of raw SDW (bold lines) or SDW treated for 5 days with *A. niger* (dotted lines) monitored at the $\lambda_{Ex}/\lambda_{Em}$ = 221 nm/350 nm and corresponding to the injection of 1 mg of soluble COD (**a**). Seven fractions (F1 to F7) were collected and performed through EEM fluorescence at λ_{Ex} comprised between 220 and 450 nm and λ_{Em} from 220 to 500 nm (**b**)
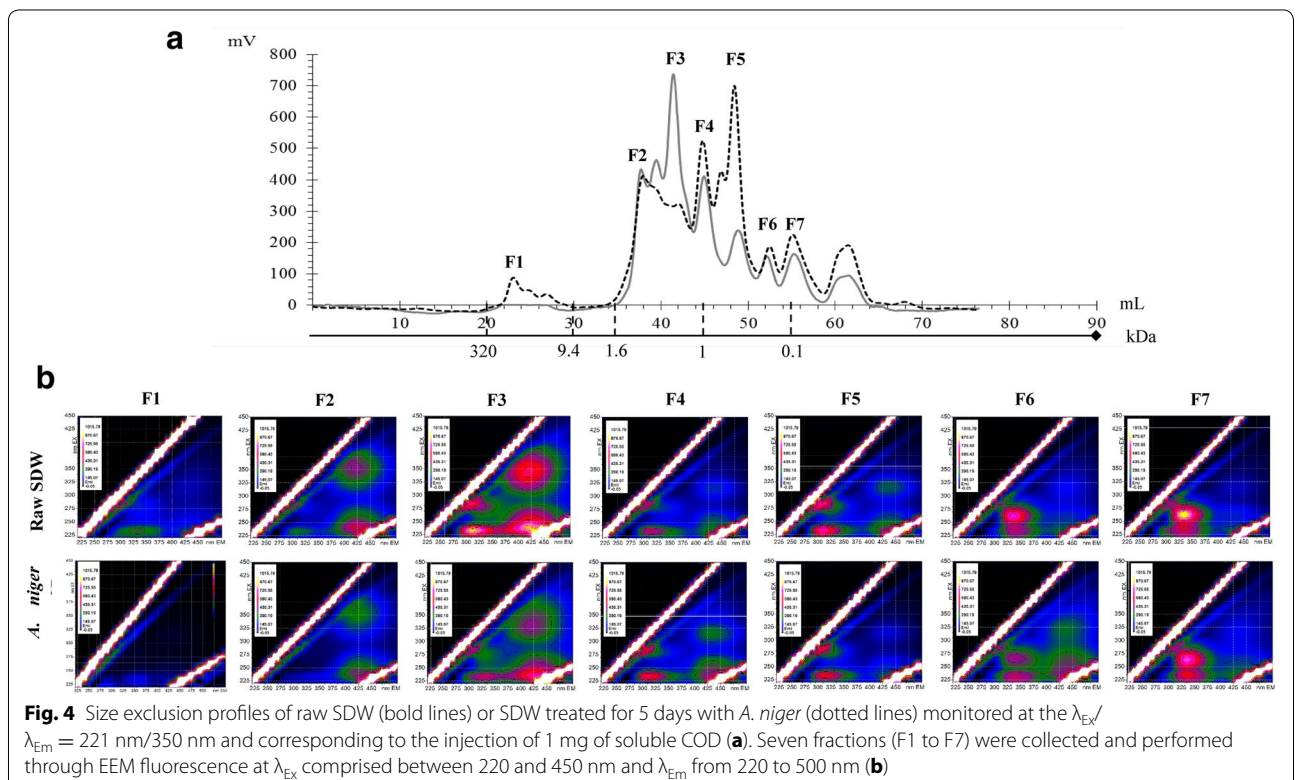
Table 3 Biochemical properties of the fractions eluted after SEC fractionation of raw and fermented SDW: quantitative distribution of each fraction and ratio of maxima fluorescence intensity for the three A, B and C peak areas detected in their EEM

Fractions	Peak (mL)	Repartition of peak area compared to total area (%)		A/B EEM peak area		A/C EEM peak area	
		Raw SDW	Fermented SDW	Raw SDW	Fermented SDW	Raw SDW	Fermented SDW
F1	24 ± 2	–	5.31	–	2.0	–	3.9
F2	38 ± 1	12.1	25.67	1.9	1.8	0.6	0.4
F3	41.5 ± 1.5	31.6	8.73	1.8	1.8	1.2	3.3
F4	45 ± 1	15.3	14.32	2.2	2.2	2.5	2.2
F5	49 ± 1.5	7.7	18.63	1.7	1.7	2.6	4.4
F6	52.25 ± 1.25	2.7	4.02	0.7	1.1	3.5	1.8
F7	55.75 ± 2.25	7.1	7.77	0.6	0.8	5.1	4.0

A corresponds to the maximum intensity of the peak area related to protein-like (PN-like) substances, B to phenolic gallic acid-like (PA-like) molecules and C to the humic acid-like (HA-like) substances. SEC and EEM profiles were obtained from supernatants of SDW pre-treated by *A. niger* during 5 days

F1 fraction that were not present in raw SDW. According to the calibration curve (see "EEM profiles obtained from filtrates of crude and pre-treated SDW" section), these PN-like molecules that eluted around 24 mL had high apparent molecular weight around 100,000 Da. Also, molecules with very small molecular weight, eluted in the F6 and F7 fractions, were found in both profiles with a similar repartition. The PN-like molecules included in the F2 fraction (around 1000 Da) were not fully digested during *A. niger* fermentation since they still represent 25.7% of the total molecules (Table 3). On the other hand, a drastic diminution of F3 peak, and an increase of the F5 peak were observed in fermented compared to raw SDW filtrate (Fig. 4a, Table 3). It is possible that the decrease of F3 from 31.6% in the untreated SDW to 8.73% after *A. niger* is recovered in the F5 peak that has increased from 7.7 to 18.63% of the total SEC area. These observations suggest that molecules with intermediate size (F3) might have been partially hydrolysed in small molecules (F5) after 5 days of *A. niger* fermentation in SDW. This is in agreement with the fact that after a first growth period of *A. niger* on released monosaccharides, other complex polymeric molecules need to be hydrolysed to provide additional carbon sources. The high apparent molecular weight molecules detected in the fermented SDW (F1) might thus correspond to enzymes secreted by the biomass for hydrolysis of SDW carbon-containing polymers.

To further investigate the effect of *A. niger* fermentation on the biochemical characteristics of vinasse, F1 to F7 fractions were collected and their EEM were determined (Fig. 4b). For better specificity towards PN-like detection, the ratios A/B and A/C of maxima fluorescence intensity for these different zones were calculated (Table 3). As shown in Fig. 4b, EEM fingerprints were quite similar for raw and fermented SDW. Globally, F2 fraction was more enriched in HA-like substances (ratio peak C/A more important for F2 fraction than for others) whereas F7 fraction was especially enriched in PA-like molecules (ratio peak B/A more important for F7 fraction and especially for SDW filtrate). EEM fingerprints of F3, F5 and F6 fractions were slightly impacted by fermentation. For F3 and F5, the A/C ratio was increased by a factor of 2.8 and 1.7 after fermentation respectively whereas A/B ratio was unchanged. That might be linked to the increase in PN-like and/or the hydrolysis of HA-like molecules during *A. niger* fermentation. On the other hand, concerning F6 fractions, the A/C ratio was reduced by a factor 1.9, decreasing from a value of 3.5–1.8 after fermentation. So, fermentation has decreased the level of HA-like substances in fractions F3 and F5 whereas these substances were recovered in a higher amount in F6 fractions.

According to these results, some physiological aspects of *A. niger* fermentation of raw SDW can be proposed: (1) production of high apparent molecular weight (F1 fraction) and hydrolysed (F3 and F5 fractions) PN-like molecules (2) hydrolysis of HA-like substances (F3 and F5 fractions) in smaller HA-like molecules, (F6 fraction) inducing vinasse decolourisation. This approach also demonstrated that SEC coupled with fluorescence monitoring at $\lambda_{Ex}/\lambda_{Em} = 221/350$ nm is a good alternative for determination of vinasse biochemical fingerprints. This strategy was previously used to show the impact of biological aggregate sludge and origin of aggregate on

exopolymeric substances fingerprint for which number of peaks and their intensity were easily identified with the specific PN-like fluorescence detection [43].

Lipid extraction from *A. niger* biomass and total Single Cell Oil yield

In an attempt to explore the potential of fungal biomass for biodiesel production, the lipid content of the biomass produced on SDW was measured and compared with the one produced on a lipid accumulation medium (Table 4). With almost 3 times more biomass produced on SDW as compared to LAM, the lipid content of the fungal biomass grown on SDW (6.94%) was slightly higher as compared to LAM (5.89%). Comparatively, Zheng et al. [47] showed that *A. niger* grown on glucose or xylose as sole carbon source led to a production of 5.8 and 4.6 g L^{-1} of biomass with a lipid content of 9.6 and 8% respectively. Similarly, *A. niger* grown on bagasse led to a fungal biomass of about 1.9 g L^{-1} with a lipid content of 13.6% [48]. Also, André et al. [49] showed that two different *A. niger* strains cultivated in crude glycerol could produce up to 8.2 g L^{-1} of biomass with a lipid content of about 50% (meaning about 3 g L^{-1} of lipids). Although lipid content of fungal biomass produced on SDW is rather low (circa 7%), the high *A. niger* biomass yield on this medium suggested that SDW can therefore constitute a good alternative and cheap medium for biodiesel production.

The composition of the lipids extracted from biomass produced on LAM and SDW was respectively 18.2 and 24.9% for palmitic acid (16:0), 28.2 and 17.2% for oleic acid (18:1, n-9) and 39.4 and 42.7% for linoleic acid (18:2, n-6). Stearic (18:0) and α-linolenic (18:3, n-3) acids were produced to a lesser extent by *A. niger* on both media (Table 4). Singh [48] reported that *A. niger* biomass grown on glucose medium contained mostly linoleic acid (50%) and to a lesser extent, palmitic, stearic and linolenic acids (8.3, 5.2 and 6% respectively). Whatever the medium used, linolenic acid appeared as the major intracellular lipid of *A. niger* biomass; however, it can be noticed that *A. niger* grown on glucose medium and on LAM were richer in oleic acid than biomass grown on SDW (23.5 and 28.19 against 17.23%) [48]. By comparison, biodiesel from *Yarrowia lipolytica* [50] contained twice higher oleate esters but less than three times linoleate esters than biodiesel from *A. niger* grown on SDW. This suggests that lipids produced from *A. niger* could be an interesting alternative to the ones produced by microorganisms such as yeasts [51].

Finally, the main relevant physical properties to assess fuel quality of biodiesel from *A. niger* are presented in Table 5. Whatever the growth media used (SDW or LAM), the biodiesel derived from *A. niger* showed similar properties for all the tested physical parameters such as cetane number (φ), viscosity (η), density (ρ), higher

Table 4 Biomass production, lipid content and lipid composition of *A. niger* grown on LAM and SDW media during 10 days

Medium	Biomass	Lipid content[a]	Lipid composition				
	(g L^{-1})	(% of DW)	16:0	18:0	18:1 (n-9)	18:2 (n-6)	18:3 (n-3)
LAM	8.523	5.889	18.19	6.84	28.19	39.38	7.4
SDW	24.060	6.940	24.94	5.25	17.23	42.66	9.92

16:0: palmitic acid; 18:0: stearic acid; 18:1 (n-9): oleic acid; 18:2 (n-6): linoleic acid; 18:3 (n-3): linolenic acid

[a] Lipid content expressed in gram of lipids per 100 g of dry weight biomass

Table 5 Most relevant physical characteristics of biodiesel extracted and converted from *A. niger* biomass grown on LAM and SDW media during 10 days

	CN (φ)	Viscosity (η) (mm^2 s^{-1})	Density (ρ) (g cm^{-3})	HHV (δ) (MJ kg^{-1})	CFPP (°C)
SDW	57.65	3.47	0.87	40.01	− 0.39
LAM	58.88	3.52	0.87	40.05	− 0.02
EN 14,214[a]	> 51[b]	3.5–5[c]	0.86–0.9[b]	–	< − 15[d,1]; < 0[d,2]
ASTM D 6751-08[a]	> 47[b]	1.9–6[c]	n.a	–	n.a

Each data is the mean of three independent biological experiments

CN cetane number, *HHV* higher heating value, *CFPP* cold filter plugging point ([1]in winter, [2]in summer), *n.a* not available

[a] According to European and American specifications biodiesel fuel blendstocks (B100), standard specifications EN 14,214 and D 6751-08 for biodiesel fuel blendstocks (B100) established respectively by the European Committee for Standardization (CEN) and American Society for Testing and Materials (ASTM). Data from [b] [52], [c] [53], [d] [54]

heating value (HHV—δ) and cold filter plugging point (CFPP). Cetane numbers of biodiesel produced from SDW and LAM media (respectively 57.65 and 58.88) were both at least 13% better than the minimal requirement of biodiesel proposed by the European and American standards. For comparison, the biodiesel derived from *A. niger* had similar cetane numbers to biodiesels produced from coconut, tallow or yellow grease (respectively 59.3, 58.9 and 56.9) [51]. In comparison, cetane number of biodiesel from *Y. lipolityca* was 64.37 [50]. Also viscosities of biodiesel produced from *A. niger* (3.52 mm^2 s^{-1} on LAM and 3.47 mm^2 s^{-1} on SDW) were globally in the range of values suggested by the European Standards (between 3.5 and 5 mm^2 s^{-1}). Although there are no European or American specification for this parameter, HHV of biodiesel from *A. niger* grown on SDW (40.01 MJ kg^{-1}) is considered as acceptable given that biodiesel from all kind of sources are generally 10% less energetic than diesel from petroleum (49.65 MJ kg^{-1}) [51]. Finally, CFPP value of biodiesel obtained from *A. niger* grown on SDW was lower than 0 °C, meaning that this biodiesel could be used at low temperature.

Conclusions

This study demonstrated that raw SDW contains suitable organic substrates for growth of *A. niger* including monosaccharides, organic acids and complex polymers. The growth reached up to 35.29 g L^{-1} DW fungal biomass with a biomass yield of 0.47 g per g of SDW organic compounds. Aerobic fermentation of raw SDW led to vinasse decolourization with pH increase and COD decrease, dropping thus significantly the pollutant load. Biochemical fingerprints revealed that high molecular weight PN-like components were secreted by *A. niger* during growth while some PA and/or HA-like molecules were consumed. Intracellular lipids from biomass showed good physical characteristics for use as biofuel giving new insights for concomitant bioremediation and carbon reuse of SDW medium.

Authors' contributions

GCT, JH, MW, IG, IB, EGN and TP performed the experimental and laboratory work. GCT, JH, ASCS, YC, IG, IB, JMF, EGN and TP worked on the analysis and interpretation of the data and contributed with valuable discussions. GCT, JH, LA, IB, JMF, EGN and TP conceived the project, worked on the structure and wrote the paper. All authors read and approved the final manuscript.

Author details

[1] Antenne sud du laboratoire de chimie des Substances Naturelles et des Sciences des Aliments (LCSNSA), EA 2212, Université de la Réunion, UFR des Sciences et Technologies, 15 Avenue René Cassin, CS 92003, 97744 Saint-Denis Cedex 9, France. [2] Laboratoire de Physique et Ingénierie Mathématique pour l'Energie et l'Environnement (PIMENT), EA 4518, Université de la Réunion, UFR Sciences de l'Homme et de l'Environnement, 117 rue Général Ailleret, 97430 Le Tampon, France. [3] Groupement de Recherche Eau Sol Environnement (GRESE), EA 4330, Université de Limoges, Faculté des Sciences et Techniques, 123 Avenue A. Thomas, 87060 Limoges Cedex, France. [4] LISBP, UMR INSA-CNRS &/INRA 792, 135 Avenue de Rangueil, 31077 Toulouse Cedex 4, France. [5] Laboratoire de Biotechnologies Agroalimentaire et Environnementale (LBAE), EA 4565, Université de Toulouse III, Institut Universitaire de Technologie, 24 Rue d'Embaquès, 32000 Auch, France. [6] Present Address: Département Hygiène Sécurité Environnement (HSE), Institut Universitaire de Technologie, Université de La Réunion, 40 Avenue de Soweto, 97410 Saint-Pierre, France.

Competing interests

The authors declare that they have no competing interests.

References

1. Lima AM, Souza RRD. Use of sugar cane vinasse as substrate for biosurfactant production using Bacillus subtilis PC. Chem Eng Trans. 2014;37:673–8.
2. Bhattacharyya A, Pramanik A, Maji S, Haldar S, Mukhopadhyay U, Mukherjee J. Utilization of vinasse for production of poly-3-(hydroxybutyrate-co-hydroxyvalerate) by *Haloferax mediterranei*. AMB Express. 2012;2:34.
3. Tewari PK, Batra VS, Balakrishnan M. Water management initiatives in sugarcane molasses based distilleries in India. Resour Conserv Recycl. 2007;52:351–67.
4. Wilkie AC, Riedesel KJ, Owens JM. Stillage characterization and anaerobic treatment of ethanol stillage from conventional and cellulosic feedstocks. Biomass Bioenergy. 2000;19:63–102.
5. Fuess LT, Garcia ML. Implications of stillage land disposal: a critical review on the impacts of fertigation. J Environ Manag. 2014;145:210–29.
6. Baez-Smith C. Anaerobic digestion of vinasse for the production of methane in the sugar cane distillery. In: SPRI Conference on Sugar Processing Research, Águas de São Pedro, S.P., Brazil. 2006.
7. Rajagopal V, Paramjit SM, Suresh KP, Yogeswar S, Nageshwar RDVK, Avinash N. Significance of vinasses waste management in agriculture and environmental quality—review. Afr J Agric Res. 2014;9:2862–73.
8. Acharya BK, Mohana S, Madamwar D. Anaerobic treatment of distillery spent wash—a study on upflow anaerobic fixed film bioreactor. Bioresour Technol. 2008;99:4621–6.
9. España-Gamboa E, Mijangos-Cortes J, Barahona-Perez L, Dominguez-Maldonado J, Hernández-Zarate G, Alzate-Gaviria L. Vinasses: characterization and treatments. Waste Manag Res. 2011;29:1235–50.
10. Mohana S, Acharya BK, Madamwar D. Distillery spent wash: treatment technologies and potential applications. J Hazard Mater. 2009;163:12–25.
11. Bustamante MA, Paredes C, Moral R, Moreno-Caselles J, Pérez-Espinosa A, Pérez-Murcia MD. Uses of winery and distillery effluents in agriculture: characterization of nutrient and hazardous components. Water Sci Technol. 2005;51:145–51.
12. Khairnar P, Chavan F, Diware VR. Generation of energy from distillery waste water. Int J Sci Spiritual Bus Technol. 2013;2:30–5.
13. Biswas AK, Mohanty M, Hati KM, Misra AK. Distillery effluents effect on soil organic carbon and aggregate stability of a Vertisol in India. Soil Tillage Res. 2009;104:241–6.
14. Ansari F. Environmental impact of distillery effluent on vertical soil horizon due to leaching effect: an experimental approach. Int J Chem Environ Eng. 2014;5:223.
15. Pant D, Adholeya A. Biological approaches for treatment of distillery wastewater: a review. Bioresour Technol. 2007;98:2321–34.
16. Kanimozhi R, Vasudevan N. An overview of wastewater treatment in distillery industry. Int J Environ Eng. 2010;2:159–84.

17. Palacios-Cabrera H, Taniwaki MH, Hashimoto JM, Menezes HC. Growth of *Aspergillus ochraceus*, *A. carbonarius* and *A. niger* on culture media at different water activities and temperatures. Braz J Microbiol. 2005;36:24–8.

18. Schrickx JM, Raedts MJH, Stouthamer AH, Vanverseveld HW. Organic acid production by *Aspergillus niger* in recycling culture analyzed by capillary electrophoresis. Anal Biochem. 1995;231:175–81.

19. Quintanilla D, Hagemann T, Hansen K, Gernaey KV. Fungal morphology in industrial enzyme production—modelling and monitoring. Adv Biochem Eng Biotechnol. 2015;149:29–54.

20. Schuster E, Dunn-Coleman N, Frisvad JC, Van Dijck PWM. On the safety of *Aspergillus niger*—a review. Appl Microbiol Biotechnol. 2002;59:426–35.

21. Oshoma CE, Imarhiagbe EE, Ikenebomeh MJ, Eigbaredon HE. Nitrogen supplements effect on amylase production by *Aspergillus niger* using cassava whey medium. Afr J Biotechnol. 2010;9:682–6.

22. Rosalem P, Tauk S, Santos MCN. Efeito da temperatura, pH, tempo de cultivo e nutrientes no crescimento de fungos imperfeitos em vinhaca. Rev microbiologia. 1985;16:299–304.

23. Silveira Ruegger MJ, Tauk-Tornisielo SM. Biomass production by filamentous fungi in sugar cane vinasse medium supplemented with molasses. Arq Biol Tecnol. 1996;39:323–32.

24. Sluiter A, Hames B, Scarlata C, Sluiter J, Templeton D. Determination of ash in biomass (No. NREL/TP-510-42622). National Renewable Energy Laboratory of U.S. Department of Energy, Golden, US. 2008.

25. Janke L, Leite A, Nikolausz M, Schmidt T, Liebetrau J, Nelles M, Stinner W. Biogas production from sugarcane waste: assessment on kinetic challenges for process designing. Int J Mol Sci. 2015;16:20685–703.

26. Suutari M, Priha P, Laakso S. Temperature shifts in regulation of lipids accumulated by *Lipomyces starkeyi*. J Am Oil Chem Soc. 1993;70:891–4.

27. Simon S, Païro B, Villain M, D'Abzac P, Van Hullebusch E, Lens P, Guibaud G. Evaluation of size exclusion chromatography (SEC) for the characterization of extracellular polymeric substances (EPS) in anaerobic granular sludges. Bioresour Technol. 2009;100:6258–68.

28. Cropotova J, Popel S, Parshacova L, Colesnicenco A. Effect of 1-year storage time on total polyphenols and antioxidant activity of apple fillings. J Food Packag Sci Tech Technol. 2015;4:44–9.

29. Hoarau J, Caro Y, Petit T, Grondin I. Evaluation of direct wet transesterification methods on yeast and fungal biomass grown on sugarcane distillery spent wash. Chem Eng Process Technol. 2016;2:1032.

30. Sangave PC, Pandit AB. Enhancement in biodegradability of distillery wastewater using enzymatic pretreatment. J Environ Manag. 2006;78:77–85.

31. España-Gamboa EI, Mijangos-Cortés JO, Hernández-Zárate G, Maldonado JAD, Alzate-Gaviria LM. Methane production by treating vinasses from hydrous ethanol using a modified UASB reactor. Biotechnol Biofuels. 2012;5:82.

32. Ferreira LFR, Aguiar MM, Messias TG, Pompeu GB, Lopez AMQ, Silva DP, Monteiro RT. Evaluation of sugarcane vinasse treated with *Pleurotus sajor-caju* utilizing aquatic organisms as toxicological indicators. Ecotoxicol Environ Saf. 2011;74:132–7.

33. Sheehan GJ, Greenfield PF. Utilization, treatment and disposal of distillery wastewater. Water Res. 1980;14:257–77.

34. Miranda MP, Benito GG, Cristobal NS, Nieto CH. Color elimination from molasses wastewater by *Aspergillus niger*. Bioresour Technol. 1996;57:229–35.

35. Patil PU, Kapadnis BP, Dhamankar VS. Decolorisation of synthetic melanoidin and biogas effluent by immobilised fungal isolate of *Aspergillus niger* UM2. Int Sugar J. 2003;105:10–3.

36. Cavka A, Jönsson LJ. Comparison of the growth of filamentous fungi and yeasts in lignocellulose-derived media. Biocatal Agric Biotechnol. 2014;3:197–204.

37. Jin B, Yan XQ, Yu Q, van Leeuwen JH. A comprehensive pilot plant system for fungal biomass protein production and wastewater reclamation. Adv Environ Res. 2002;6:179–89.

38. Khosravi-Darani K, Zoghi A. Comparison of pretreatment strategies of sugarcane baggase: experimental design for citric acid production. Bioresour Technol. 2008;99:6986–93.

39. Agarwal R, Lata S, Gupta M, Singh P. Removal of melanoidin present in distillery effluent as a major colorant: a review. J Environ Biol (India). 2010;31:521–8.

40. Li WT, Chen SY, Xu ZX, Li Y, Shuang CD, Li AM. Characterization of dissolved organic matter in municipal wastewater using fluorescence PARAFAC analysis and chromatography multi-excitation/emission scan: a comparative study. Environ Sci Technol. 2014;48:2603–9.

41. Lakowicz JR, editor. Principles of fluorescence spectroscopy. Boston: Springer; 2006.

42. Huang M, Li Y, Gu G. Chemical composition of organic matters in domestic wastewater. Desalination. 2010;262:36–42.

43. Bhatia D, Bourven I, Simon S, Bordas F, van Hullebusch ED, Rossano S, Lens PNL, Guibaud G. Fluorescence detection to determine proteins and humic-like substances fingerprints of exopolymeric substances from biological sludges performed by size exclusion chromatography. Bioresour Technol. 2013;131:159–65.

44. Pokhrel D, Viraraghavan T. Treatment of pulp and paper mill wastewater—a review. Sci Total Environ. 2004;333:37–58.

45. Soobadar A. Agronomic and environmental impacts of application of coal/bagasse ash and vinasse to sugarcane fields in Mauritius (PhD thesis). Université d'Avignon, Avignon, France. 2009.

46. Bridgeman J, Baker A, Carliell-Marquet C, Carstea E. Determination of changes in wastewater quality through a treatment works using fluorescence spectroscopy. Environ Technol. 2013;34:3069–77.

47. Zheng Y, Yu X, Zeng J, Chen S. Feasibility of filamentous fungi for biofuel production using hydrolysate from dilute sulfuric acid pretreatment of wheat straw. Biotechnol Biofuels. 2012;5:50.

48. Singh A. Lipid accumulation by a cellulolytic strain of *Aspergillus niger*. Experientia. 1992;48:234–6.

49. André A, Diamantopoulou P, Philippoussis A, Sarris D, Komaitis M, Papanikolaou S. Biotechnological conversions of bio-diesel derived waste glycerol into added-value compounds by higher fungi: production of biomass, single cell oil and oxalic acid. Ind Crops Prod. 2010;31:407–16.

50. Katre G, Joshi C, Khot M, Zinjarde S, RaviKumar A. Evaluation of single cell oil (SCO) from a tropical marine yeast *Yarrowia lipolytica* NCIM 3589 as a potential feedstock for biodiesel. AMB Express. 2012;2:36.

51. Hoekman SK, Broch A, Robbins C, Ceniceros E, Natarajan M. Review of biodiesel composition, properties, and specifications. Renew Sustain Energy Rev. 2012;16:143–69.

52. Ramírez-Verduzco LF, Rodríguez-Rodríguez JE, del Rayo Jaramillo-Jacob A. Predicting cetane number, kinematic viscosity, density and higher heating value of biodiesel from its fatty acid methyl ester composition. Fuel. 2012;91:102–11.

53. Ramos MJ, Fernández CM, Casas A, Rodríguez L, Pérez Á. Influence of fatty acid composition of raw materials on biodiesel properties. Bioresour Technol. 2009;100:261–8.

54. Su YC, Liu YA, Diaz-Tovar CA, Gani R. Selection of prediction methods for thermophysical properties for process modeling and product design of biodiesel manufacturing (thesis). Virginia Tech. 2011.

Heterologous and endogenous *U6* snRNA promoters enable CRISPR/Cas9 mediated genome editing in *Aspergillus niger*

Xiaomei Zheng[1,2] , Ping Zheng[1,2]*, Jibin Sun[1,2]* , Zhang Kun[1,2,3] and Yanhe Ma[1]

Abstract

Background: *U6* promoters have been used for single guide RNA (sgRNA) transcription in the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein (CRISPR/Cas9) genome editing system. However, no available *U6* promoters have been identified in *Aspergillus niger,* which is an important industrial platform for organic acid and protein production. Two CRISPR/Cas9 systems established in *A. niger* have recourse to the RNA polymerase II promoter or in vitro transcription for sgRNA synthesis, but these approaches generally increase cloning efforts and genetic manipulation. The validation of functional RNA polymerase II promoters is therefore an urgent need for *A. niger*.

Results: Here, we developed a novel CRISPR/Cas9 system in *A. niger* for sgRNA expression, based on one endogenous *U6* promoter and two heterologous *U6* promoters. The three tested *U6* promoters enabled sgRNA transcription and the disruption of the polyketide synthase *albA* gene in *A. niger*. Furthermore, this system enabled highly efficient gene insertion at the targeted genome loci in *A. niger* using donor DNAs with homologous arms as short as 40-bp.

Conclusions: This study demonstrated that both heterologous and endogenous *U6* promoters were functional for sgRNA expression in *A. niger*. Based on this result, a novel and simple CRISPR/Cas9 toolbox was established in *A. niger,* that will benefit future gene functional analysis and genome editing.

Keywords: *Aspergillus niger*, CRISPR/Cas9 system, *U6* snRNA promoters, Genome editing

Background

Aspergillus niger has attracted great attention due to its biotechnological value as a platform for producing organic acids, such as citric acid, gluconic acid and oxalic acid [1], as well as producing homologous and heterologous enzymes, including glucoamylases, amylases, acid protease, cellulase, glucose oxidase, pectinases, and xylanases [2]. *A. niger* can be used to create a promising, versatile cell factory for producing more low-priced bulk chemicals because of its remarkable unique features, including extreme acid resistance, significant robustness

and powerful polymer hydrolytic enzymes. Despite its industrial importance, efficient genetic tools are generally unavailable, hampering the fundamental study and industrial improvement of this species.

The clustered regularly interspaced short palindromic repeats/CRISPR associated protein (CRISPR/Cas9) system is a powerful and revolutionary genome editing tool [3, 4]. In the CRISPR/Cas9 system, the endonuclease Cas9 is guided to a specific locus by a single guide RNA (sgRNA) where it generates a double strand break (DSB) in the genome. The DSB is usually repaired by either non-homologous end joining (NHEJ) to allow NHEJ-mediated gene disruption with base-pair insertions or deletions or homologous recombination (HR), which allows HR-mediated precise genome editing with appropriate donor DNA. The HR frequency is very low (less than 5%) in filamentous fungi [5, 6]. In traditional gene editing

*Correspondence: zheng_p@tib.cas.cn; sunjibin@tib.cas.cn;
sun_jb@tib.cas.cn
[1] Tianjin Institute of Industrial Biotechnology, Chinese Academy
of Sciences, Xiqidao 32, Tianjin Airport Economic Area, Tianjin 300308,
China
Full list of author information is available at the end of the article

methods, efficiency is typically improved by increasing the homologous arm length. For example, the gene deletion efficiency was enhanced from 4 to 29%, when the homologous arm was increased from 100 to 1500-bp [6]. However, this approach had disadvantages, such a tedious donor DNA construction and onerous transformant screening. Double-stranded DNA breaks caused by Cas9 were reported to improve the HR frequency with shortened donor DNA homologous arms in *T. reesei* [7], *A. fumigatus* [8] and *P. chrysogenum* [9], whereas the CRISPR/Cas9 systems established in *A. niger* still used the donor DNA with the long homologous arms.

CRISPR/Cas9 systems have been established in *A. niger* using different strategies for sgRNA synthesis. They either depend on RNA polymerase II promoters, such as the strong promoter P*gpdA* [10] or P*mbfA* [11], or in vitro transcription is performed [12, 13]. When RNA polymerase II promoters are used, self-cleavage ribozymes, such as hepatitis delta virus (HDV) or Hammerhead (HH), are required to be added at the 5′-end and 3′-end of sgRNA, whereas the sgRNA conformation may be affected by reading-through of RNA polymerase II [10]. However, this strategy usually requires more effort when constructing sgRNA expression cassettes. As an alternative approach, Kuivanen et al. [12, 13] utilized in vitro synthesis for sgRNA expression. However, gRNA uptake and stability may influence genome editing efficiency [9]. RNA Pol III promoters for the spliceosomal *U6* snRNA have been widely used for sgRNA transcription in the CRISPR/Cas9 system. Some *U6* promoters have been used for efficient sgRNA transcription in fungi [8, 9, 14–17]. However, no *U6* promoter has been identified and validated in *A. niger*.

In this study, we aimed to establish a simple CRISPR/Cas9 system based on the *U6* promoter in *A. niger*. One endogenous *U6* promoter was identified. This endogenous promoter and two reported heterologous *U6* promoters (P*hU6* and P*yU6*) from humans and yeast were tested in *A. niger*. To enhance the simplicity of this CRISPR/Cas9 system, donor DNAs with short homologous arms (40-bp) were tested for gene insertion at DSBs induced by Cas9.

Materials and methods
Strains and media
Escherichia coli DH5α (Transgene, Beijing, China) was used for plasmid construction and cultured at 37 °C in Luria−Bertani broth containing ampicillin (100 μg/mL). The *A. niger* strains and plasmids used in this study are indicated in Additional file 1: Tables S1 and S2. *A. niger* G1 (*amdS⁻*, Δ*glaA*, and Δ*pepA*) was derived from *A. niger* NRRL3112 and presented by the Institute of Microbiology, CAS; this strain was used as the recipient strain for

genome editing. *A. niger* strains were cultivated on minimal medium (MM) [18] containing 1% glucose, 70 mM NaNO$_3$, 110 mM KH$_2$PO$_4$, 70 mM KCl, 2 mM MgSO$_4$, and trace element solution or on complete medium (CM) consisting of MM supplemented with 0.5% yeast extract and 0.1% casamino acids. When using *amdS* as a selection marker, NaNO$_3$ in MM was replaced by 10 mM acetamide and 15 mM caesium chloride (MMSA). For growth on solid plates, 1.5% agar was supplemented. If necessary, 150 μg/mL of hygromycin was added.

DNA constructions
All primers used in this study are listed in Additional file 1: Table S3. The *cas9* gene from Streptococcus pyogenes was codon-optimized for expression in *A. niger*. The nuclear localization signals (NLSs) of SV40 (PKK-KRKV) and nucleoplasmin (KRPAATKKAGQAK-KKK) were attached into the N-termini and C-termini of codon-optimized *cas9*, which was then synthesized by Life Science Research Services Company (Genewiz, Suzhou, China). After amplification with Cas9-Fm and Cas9-Rm, *cas9* was cloned into the *XhoI*I site of the *A. niger* expressing vector pGm via the ClonExpressTM one step cloning kit (Vazyme, C113), create the Cas9 expressing plasmid pCas9. To monitor the subcellular location of Cas9, enhanced green fluorescent protein (eGFP, S65T) was fused to the C-terminus of Cas9. The (G$_4$S)$_3$-linker-*egfp* was amplified using pMF272 as a template with Linker-eGFP-Fm and eGFP-Rm and then was assembled into the reverse PCR product of pCAS9 (amplified using pCas9-rev-F and pCas9-rev-R) via the ClonExpress™ one step cloning kit, thus yielding pCas-9GFP. The DNA sequences of codon-optimized *cas9* and *cas9gfp* are shown in Additional file 1: Table S4.

sgRNA expression constructs were synthesized containing the *Homo sapiens U6*, yeast *U6*, and *A. niger U6* promoter and sgRNA scaffolds with the *Bbs*I site inserted into the plasmid pEASY-Blunt to yield psgRNA1.0, psgRNA2.0 and psgRNA3.0, respectively. Then, targeting sgRNA constructs were built by digesting these sgRNA expression plasmids with *Bbs*I and ligating double stranded oligonucleotides with the protospacer of the *albA* gene to yield psgRNA1.1, psgRNA2.1 and psgRNA3.1. The linear sgRNA-target fragments were amplified from corresponding plasmids by PCR with M13-F and M13-R and used directly for transformation. The DNA sequences of sgRNA constructs are shown in Additional file 1: Table S5.

The donor DNA MHi-albA-hph with micro-homologous arms was synthesized by PCR amplification of the selection marker hph using the with primers MHi-albA-Fm and MHi-albA-Rm containing 40-bp homologous arms, which were homogenous to the flanking region

of the *albA* sequence to be targeted. After purification, PCR products were used directly for transformation. The DNA sequences of donor DNA are shown in Additional file 1: Table S6.

DNA transformation and analysis

Aspergillus niger transformation protocols, selection procedures, *A. niger* genomic DNA isolation and diagnostic PCR were performed as described in Meyer et al. [18]. The standard protocol of this novel CRISPR/Cas9 system for target gene editing is established. The construction of sgRNA with different targets and donor DNA with micro-homologous arms was followed by the co-transformation of the Cas9 expression plasmid, sgRNA, and donor DNA into the protoplasts of *A. niger* G1. The transformants were streaked on selective media at least once. After cultivating in CM rich media, genomic DNA isolation and diagnostic PCR were performed to confirm correct transformants.

For *Cas9* and *Cas9-eGFP* expression, 5 μg of the expression plasmid pCas9 and pCas9GFP, respectively, were transformed into *A. niger* G1 protoplasts by PEG/CaCl$_2$-mediated transformation. Colonies grown on MMSA for 5 days at 30 °C were screened for the *amdS* selection marker, and diagnostic PCR was performed. The positive transformants from each construct were named *A. niger* XM1 and *A. niger* XM2, respectively (Additional file 1: Table S1).

For *albA* disruption, 5 μg of linear sgRNA PCR products and the Cas9 expression plasmid pCas9 were co-transformed into *A. niger* G1 protoplasts. Colonies grown on MMSA for 5 days at 30 °C were screened for the *amdS* selection marker, and spore phenotype statistics and sequential identification via diagnostic PCR analysis with the primers albA-g-F and albA-g-R were performed.

For *albA* gene insertion directed by donor DNA with micro-homologous arms, 5 μg linear sgRNAs PCR products, 5 μg pCas9 and 5 μg dDNA MHi-albA-hph PCR fragments were co-transformed into the protoplasts of *A. niger* G1. Transformants grown on MMSA with 100 μg/mL hygromycin B (Sigma-Aldrich, St. Louis, MO, USA) for 5 days at 30 °C were screened for *amdS* and *hph* selection markers, and spore phenotype statistics and sequential identification via diagnostic PCR and sequencing analysis with primers albA-g-F/hph-R and hph-F/albA-g-R were performed.

Microscopic analysis of the subcellular localization of the Cas9-GFP fused protein

To analyse the subcellular localization of the Cas9-GFP fused protein, the hyphae cultivation and microscopic analysis were performed as described in Wanka et al. [19]. Briefly, two disinfected coverslips were placed onto the bottom of a small petri dish, and then 5 mL of liquid MM was supplemented with 0.003% yeast extract. After inoculation with 10^6 spores of *A. niger*, the petri dishes were incubated for 8 h at 30 °C. After incubated with 4′,6-diamidino-2-phenylindole (DAPI) at the final concentration of 1 mg/mL for 15 min, coverslips with adherent hyphae were placed upside down on an object slide and analysed by microscopy. Differential interference contrast (DIC) and green fluorescent images of the cells were captured with a 40× objective using a Leica DM5000B and the results were assembled in Adobe Photoshop 7.0 (Adobe, San Jose, CA).

Quantitative reverse transcription PCR (qRT-PCR)

Aspergillus niger transformants were grown in triplicate in CM liquid media as described above. After 24 h of growth, mycelia were harvested, and total RNA was extracted using the RNAprep pure Plant Kit (Tiangen, Beijing). For relative RT-qPCR, total RNA was first reverse transcribed for first-strand cDNA synthesis using Fast-Quant RT Super Mix KR108 (Tiangen, Beijing). A 1-μL sample from the 10-fold dilution of the cDNA synthesis mix was subjected to qPCR with SuperReal PreMix Plus (SYBR Green). Real time amplification was performed using an ABI 7500 real-time PCR system. The primers sgRNA-qPCR-F and sgRNA-qPCR-R were used for the amplification of sgRNA. The relative expression level was calculated using the ΔΔC$_T$ method. 18S rRNA was used as an internal control gene and was amplified using the primers 18S rRNA-qPCR-F and 18S rRNA-qPCR-R.

Results and discussion
Nuclear localization of codon-optimized Cas9

To ensure the nuclear localization of Cas9 in *A. niger*, the SV40 NLS (PKKKRKV) and the nucleoplasmin NLS (KRPAATKKAGQAKKKK) were fused onto the N-terminal and C-terminal, respectively, of Cas9 that was originally from the bacterium *Streptococcus pyogenes* but was codon-optimized for *A. niger*. Then, a Cas9 expression plasmid was constructed with the strong induced promoter P*glaA* and a universal fungal transcription terminator T*trpC* (Fig. 1a). To monitor the subcellular location of Cas9, eGFP was fused to the C-terminal of Cas9 (Fig. 1b). The plasmids pCas9 and pCas9GFP were transformed into *A. niger* G1, to generate XM1 and XM2, respectively. The localization of Cas9GFP in *A. niger* XM2 was detected by fluorescence microscopy. Compared with *A. niger* G1, the green fluorescence spots were detected in the mycelia of *A. niger* XM2, which overlapped the with the DAPI stained nuclei (Fig. 1c). This result demonstrated that Cas9 successfully localized to the nucleus with the aid of NLSs from SV40 and nucleoplasmin.
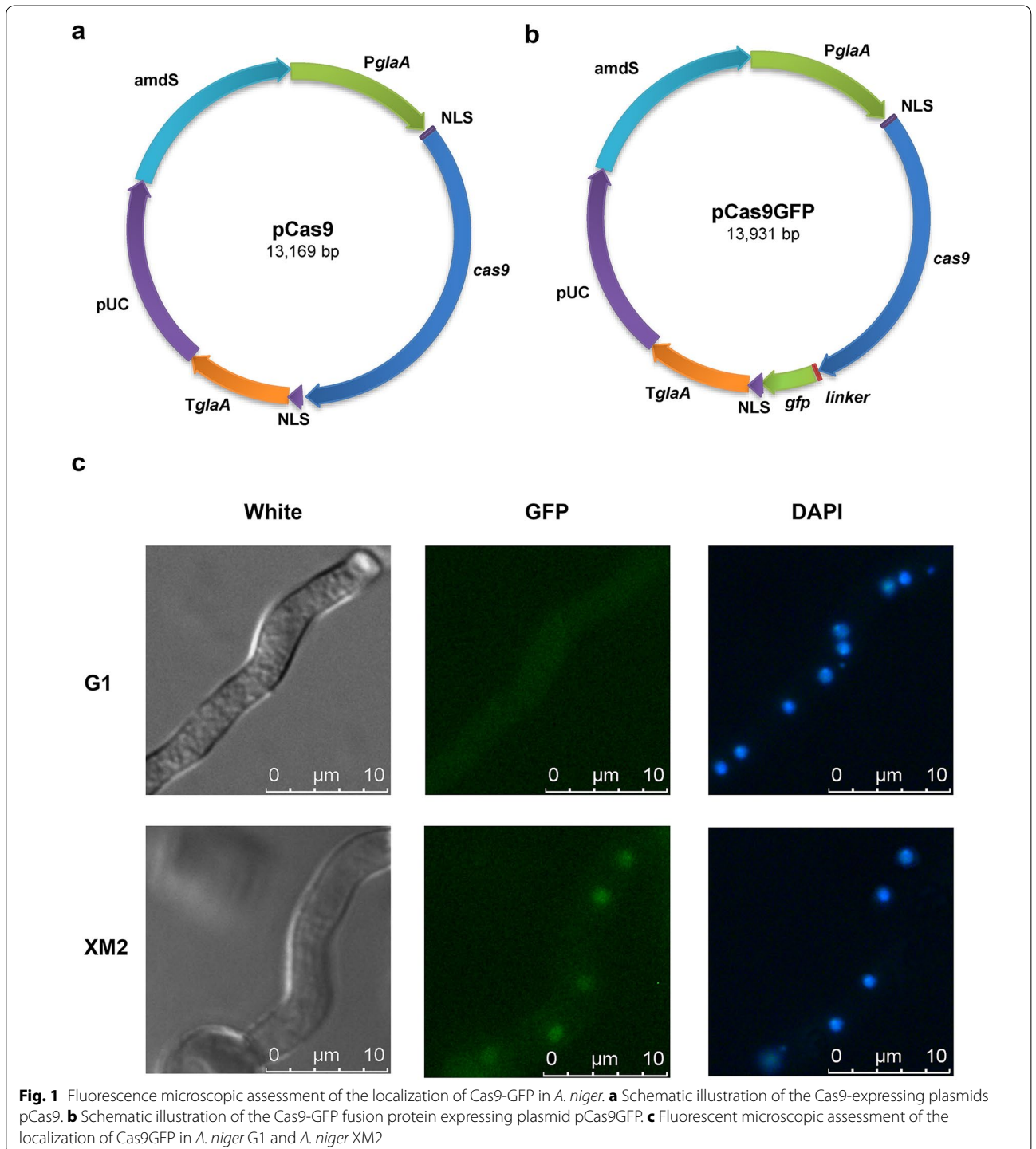
Fig. 1 Fluorescence microscopic assessment of the localization of Cas9-GFP in *A. niger*. **a** Schematic illustration of the Cas9-expressing plasmids pCas9. **b** Schematic illustration of the Cas9-GFP fusion protein expressing plasmid pCas9GFP. **c** Fluorescent microscopic assessment of the localization of Cas9GFP in *A. niger* G1 and *A. niger* XM2

Different *U6* promoters efficiently initiated sgRNA transcription for genomic *albA* disruption

To establish a simple CRISPR/Cas9 system based on the *U6* promoter in *A. niger*, we tested three *U6* promoters from different species for sgRNA expression. First, one *A. niger U6 snRNA* gene (AY136823.1) was retrieved form the NCBI GenBank database. To identify the transcription start site and promoter of this *U6 snRNA* gene, it was aligned with the *Homo sapiens RNU6* gene [20] (NR_004394) and yeast *RNU6* gene [21] (X12565.1). The 412-bp upstream of *A. niger U6 snRNA* was identified as the promoter, which showed approximately 79% identity

to yeast *RNU6* promoter sequence. This *A. niger U6* promoter included some key regulatory elements, such as the TATA-like box and proximal and distal sequence elements (Fig. 2a). Interestingly, this *U6 snRNA* gene could not be amplified using the *A. niger* CBS513.88 and G1 genome as templates. Thus, the 412-bp upstream of the *A. niger U6* promoter, the *Homo sapiens U6* promoter from pX330 and the 540-bp upstream of yeast *RNU6* were synthesized and tested for sgRNA expression in *A. niger* G1.

One putative polyketide synthase (PKS) gene, *albA* (An09g05730), was chosen as the target gene, because it is involved in black spore pigmentation synthesis and its mutation leads to a visible conidial albino phenotype [10]. *albA* targeting sgRNA was constructed by the *Bbs*I digestion of sgRNA expression plasmids and ligation with double stranded oligonucleotides with a protospacer of the *albA* gene. Then, the targeting sgRNAs PCR products under the control of each *U6* promoter were co-transformed with pCas9 into the protoplasts of *A. niger* G1 (Fig. 2b). After cultivation for 5 days, conidial pigmentless colonies were brought out on the primary transformation plates. The mutants with the disrupted *albA* gene mediated by the three sgRNA constructs were designated
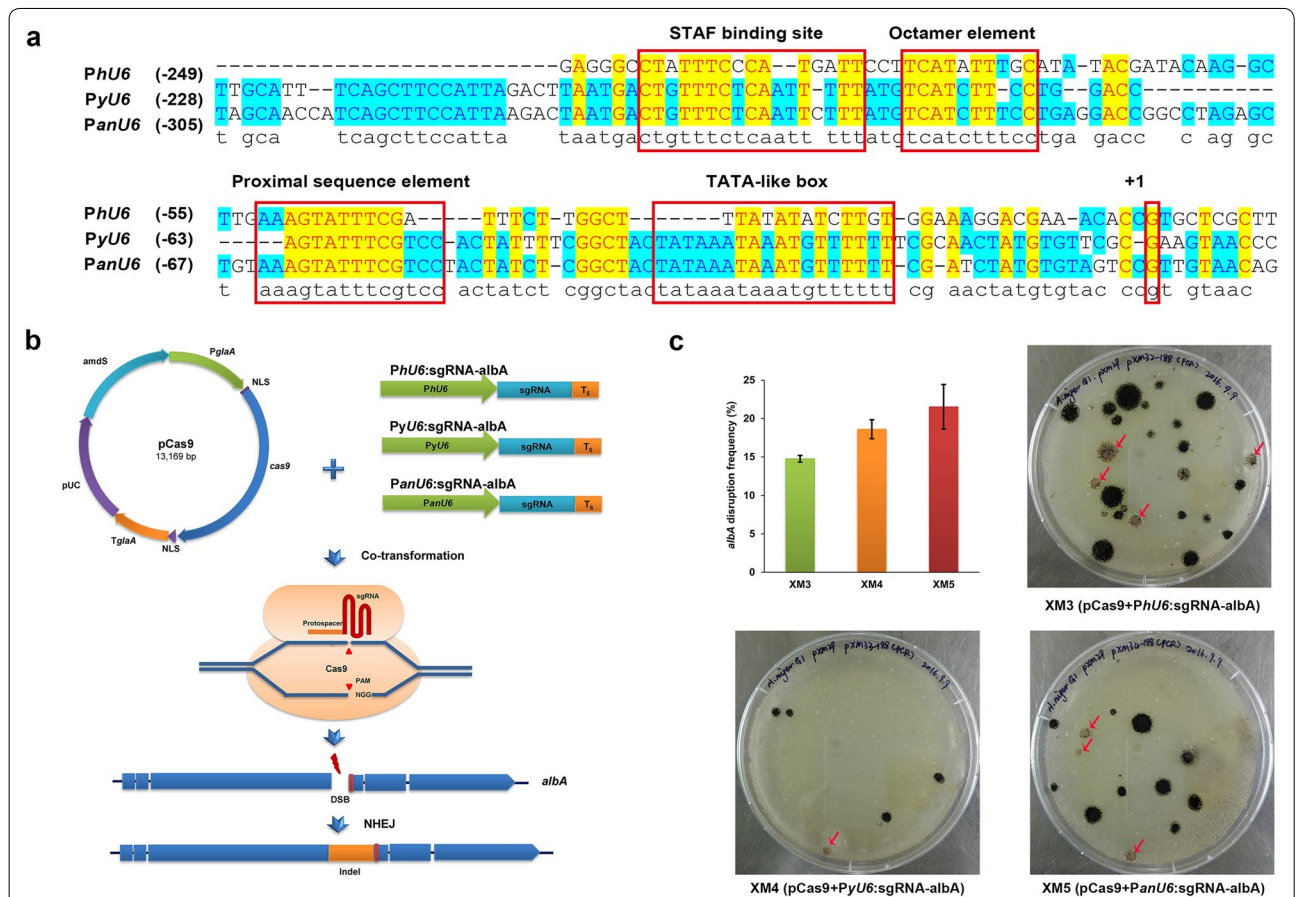
Fig. 2 Different RNA polymerase III-based promoters for CRISPR/Cas9 systems mediated *albA* gene disruption in *A. niger*. **a** Sequence alignment of the promoter sequences of Homo sapiens RNU6-1, yeast RNU6, and *A. niger* RNU6. + 1 represents the transcription start; the TATA-like box and proximal and distal sequence elements are represented by a red box. **b** Schematic diagram of *albA* disruption mediated by NHEJ using the CRISPR/Cas9 system based on the *U6* promoter. *hU6* promoter represents the promoter of the human *RNU6-1* gene (NR_004394); the *yU6* promoter represents the promoter of the yeast *RNU6* gene (X12565.1); the *anU6* promoter represents the 412-bp upstream of *A. niger RNU6* gene (AY136823.1). T_6 represents a string of six thymines serving as an RNA polymerase III terminator. Linear sgRNA constructs and Cas9 expression plasmid pCas9 were co-transformed into the protoplast. Without the donor DNA, the DSBs induced by Cas9 were repaired by the error-prone NHEJ system, which resulted in *albA* disruption. **c** Transformants growing on the primary transformation plates after 5 days incubation after being co-transformed with pCas9 and sgRNA expression cassettes. If *albA* was disrupted, the conidia of transformants turned pigmentless, forming albino colonies, as the red arrows indicate. The histogram shows the *albA* gene disruption efficiency of the transformants with sgRNA constructs under the control of different *U6* promoters. Bars represent the percentages of albino colonies that showed the *albA* disruption phenotype on the primary transformation plates (mean ± SD; n = 3)

A. niger XM3, XM4 and XM5. To ensure the reliability of the obtained data, we performed three transformations for each sgRNA cassette driven by these three *U6* promoters. We obtained dozens of transformants on the primary transformation plates with several albino colonies (Fig. 2c). The ratio of albino colonies was 15% (4/27 of primary transformants) for *A. niger* XM3 and 20% (1/5 of primary transformants) for *A. niger* XM4, and a slight higher ratio 23% (3/13 of primary transformants) was observed for *A. niger* XM5 (Fig. 2c). We performed diagnostic PCR for 12 purified clones isolated from the independent albino transformants to assess *albA* mutagenesis. No PCR product was obtained from these isolated albino colonies using the primers albA-g-F/albA-g-R, which spanned the PAM site, indicating that unpredicted large DNA deletion or insertion may occur in the targeted locus, as was seen in *A. fumigatus* [5, 8].

The intracellular sgRNA levels in cells with the three different *U6* promoters had obvious differences, even though the difference in the *albA* disruption ratio among the promoters was not very significant. Compared to the promoter from human P*hU6*, the endogenous promoter P*anU6* and the promoter from the yeast P*yU6* produced 43.76-fold and 6.09-fold more sgRNA, respectively (Additional file 1: Fig.

S1), indicating that the *U6* endogenous promoters achieved more efficient sgRNA expression.

Our results also suggested that *U6* promoters, even from distant evolutionary species, can be used to develop the CRISPR/Cas9 system in *A. niger*. Compared to previous studies [10, 11, 13], the CRISPR/Cas9 system based on the *U6* promoter is more feasible for sgRNA expression cassette construction (Table 1) without requiring any ribozymes or in vitro synthesis of sgRNA. Moreover, it is worth mentioning that in our study, the albino colonies grew directly on primary transformation plates, rendering their isolation easier than that in previous studies reported with *Aspergilli* [8, 10]. In other studies, repeated streaking was necessary to obtain albino colonies when the target gene was disrupted by the NHEJ repair pathway. This difference between our results and those of previous studies may be caused by the genetic background of host strains, the sgRNA expression efficiency or the time of DSB generation induced by the Cas9-sgRNA complex.

Precise gene insertion mediated by donor DNA with short homologous arms

To enhance the simplicity of this novel CRISPR/Cas9 system, the donor DNAs with short-homologous arms were

Table 1 Comparison on CRISPR/Cas9 system for *A. niger*

Cas9 expression		sgRNA expression		Targeted gene	Donor DNA (SM/homology arm size, bp)	Gene editing efficiency	Notes	References
Promoter	Terminator	Promoter	Terminator					
P*tef1*	T*tef1*	P*gpdA*	T*trpC*	*albA*	–	Some	Requiring to add HH and HDV for processing sgRNA	[10]
P*coxA*	T*tef1*	P*mbfA*	T*trpC*	*pyrG*	–	Obtaining 25 colonies on MM with 5′-FOA	Requiring to add self-cleaving ribozymes for processing sgRNA	[11]
				MttA	*pyrG*/690 and 834[b]	100% (7/7)		
P*tef1*	T*tef1*	In vitro synthesis		1090836	*pyrG*/1500	28% (11/40)	3% (1/30)[a]	[12]
				1117792		100% (8/8)	43% (13/30)[a]	
				1141260		100% (8/8)	0% (0/30)[a]	
				1121140		38% (3/8)	2% (1/60)[a]	
				1146483		88% (7/8)		
				1170646		63% (5/8)		
P*glaA*	T*glaA*	P*hU6*	Ploy(T)$_6$	*albA*	–	15% (4/27)	Without any selection pressure for targeted gene editing	This study
		P*yU6*			–	20% (1/5)		
		P*anU6*			–	23% (3/13)		
		P*anU6*		*albA*	*hph*/40	36% (5/14)	79% (11/14)[c]	

[a] The gene deletion efficiency without CRISPR/Cas9 system

[b] This donor DNA for gene integration was flanked by a 5′ flanking sequence of 690 bp homologous to the promoter region of the *pyrG* gene, while the 3′ flanking sequence was a mutated and truncated *pyrG* CDS of 834 bp (*pyrG*$^{m2, trunc}$)

[c] After co-transformation of donor DNA MHi-albA-hph, sgRNA3.1 and pCas9, the outgrown albino colonies with *albA* disruption reached up to 79% (11/14). In some albino colonies, some unexpected base pair errors were mediated by NHEJ at the 5′-junction or 3′-junction of DBSs. Therefore, the precise gene integration was only 36% (5/14)
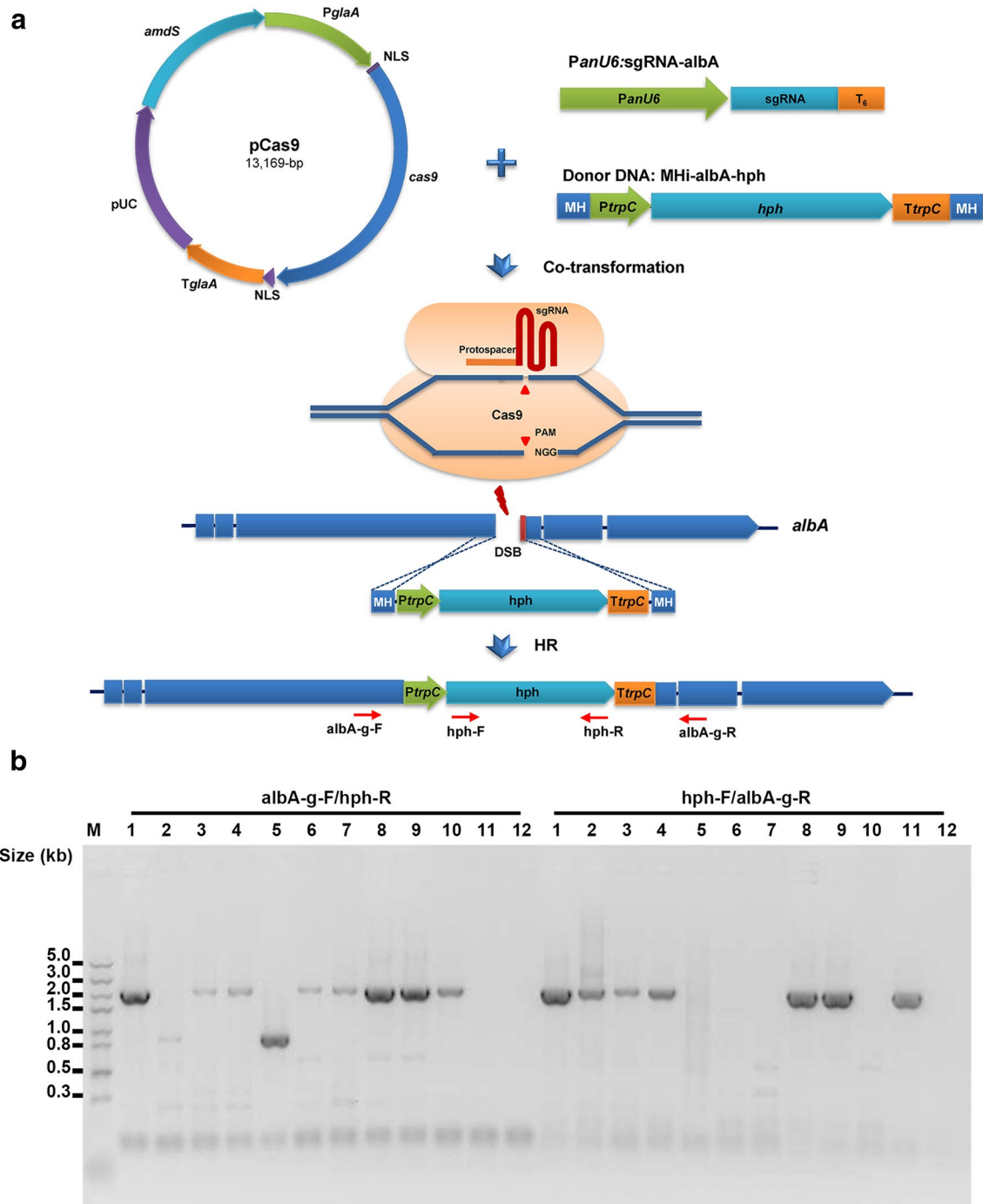
Fig. 3 Donor DNA with 40-bp short homologous arms mediated albA gene insertion by the novel CRISPR/Cas9 system in *A. niger*. **a** Schematic diagram of *albA* gene insertion mediated by integrating the donor DNA with 40-bp micro-homologous arms. The donor DNA MHi-albA-hph was co-transformed with linear sgRNA constructs and the Cas9-expressing plasmid pCas9 into wild-type *A. niger* G1. DSBs were generated by Cas9 under the guide of the sgRNA and were then repaired by HR with the integration of MHi-albA-hph. **b** Diagnostic PCR analysis of the genetic context of DSBs in albino colonies. Correct integration of the *hph* marker at the *albA* locus: 1697 bp (albA-g-F/hph-R) and 1934 bp (hph-F/albA-g-R). M, DNA ladder; 1-11, 11 albino colonies XM6.1-6.11; 12, one black colony XM6.12 without *albA* insertion

used to mediate the homology-directed recombination at the double-strand DNA breaks induced by Cas9. Donor DNAs of MHi-albA-hph were designed and constructed containing 40-bp short homologous arms located next to the PAM of the *albA* protospacer. The gene insertion was accomplished by the co-transformation of pCas9, two PCR fragments of MHi-albA-hph, and sgRNA3.1 into *A. niger* G1 protoplasts (Fig. 3a). In the negative controls, when the donor DNA MHi-albA-hph was only co-transformed with pCas9 or sgRNA3.1, no albino colonies grew on the primary transformation plates (Additional file 1: Fig. S2). However, after the co-transformation of pCas9, MHi-albA-hph, and sgRNA3.1, the outgrown albino colonies accounted for 79% (11/14) of primary transformants, dramatically increasing the *albA* gene disruption efficiency (Additional file 1: Fig. S2). This result indicated that the CRISPR/Cas9 system based on the *U6* promoter improved the gene editing efficiency and allowed the usage of donor DNA containing short homologous arms. Kuivanen et al. [12] also found that the gene editing efficiency was significantly increased with the assistance of the CRISPR/Cas9 system (Table 1).

The genetic context for the DSBs in 11 albino colonies and one black colony as a negative control were determined via PCR and DNA sequencing using two pairs of primers (Fig. 3b and Additional file 1: Fig. S3). Among 11 albino colonies, the expected PCR products were amplified in five colonies, i.e., XM6.1, 6.3, 6.4, 6.8, and 6.9, suggesting they carried the correct *hph* insertion at both expected cleavage sites (Fig. 3b, *lanes* 1, 3, 4, 8, and 9; Additional file 1: Fig. S3). For the other albino colonies, only one correct PCR product was amplified, indicating that the *hph* cassette was inserted at only the 5′-junction or 3′-junction (Fig. 3b, *lanes* 6, 7, 10, and 11; Additional file 1: Fig. S3). For the albino colony XM6.5, only one smaller PCR product was detected, indicating an 800-bp-deletion when the *hph* cassette was inserted at the 3′-junction (Fig. 3b, *lane* 5; Additional file 1: Fig. S3). The albino colony XM6.2 had a mixed genotype, similar to XM6.5 at the 3′-junction lane 5 and correct insertion at the 5′-junction (Fig. 3b, *lane* 2; Additional file 1: Fig. S3). Zhang et al. [8] found that donor DNAs with 39-bp or 28-bp homology arms were sufficient to precisely induce mutagenesis in *A. fumigatus* in a NHEJ system-independent manner. These differences could be caused by the high activity of error-prone NHEJ or by micro-homology-mediated end joining (MMEJ) [22] in the *A. niger* wild type strain.

Clearly, we have confirmed that, combined with the CRISPR/Cas9 system, short homologous arms as short 40-bp are sufficient for mediating targeted gene insertion, which facilitates the construction of donor DNA in *A. niger*. Moreover, due to the complicated genomic repair

outcomes at the DSBs in wild-type strains, it is recommended that more attention should be paid to mutant genotypes for precise editing.

Conclusions

In conclusion, we established a simple CRISPR/Cas9 system based on the *U6* promoter in *A. niger*. Two heterologous (P*hU6* and P*yU6*) and one endogenous *U6* promoter were capable of driving the transcription of sgRNA, which guided Cas9 to the target site for generating DSBs. Donor DNAs with short homologous arms (40-bp) were sufficient for insertion at DSBs induced by Cas9, simplifying and increasing the convenience of genetic manipulation in *A. niger*.

Additional file

Additional file 1. Figure S1. qPCR results of sgRNA expression levels from each promoter. Total RNA was isolated, converted to cDNA, and sgRNA expression level was quantified. sgRNA expression levels were normalized to the amount of sgRNA generated by P*hU6* promoter. 18S rRNA was used as internal control. Bars represent the fold change of sgRNA level under the control of different U6 promoters (mean ± SD; n = 3). **Figure S2.** Transformants with *albA* disruption by inserted the donor DNA with short homologous arms. Transformants XM6 grew on the primary transformation plates after co-transformed pCas9, sgRNA3.1 and donor DNA MHi-albA-hph. Transformants NC1 grew on the primary transformation plates after only co-transformed pCas9 and donor DNA MHi-albA-hph. Transformants NC2 grew on the primary transformation plates after only co-transformed sgRNA3.1 and donor DNA MHi-albA-hph. **Figure S3.** DNA sequencing analyses for genetic context at the DSBs in albA gene inserted transformants XM6. DNA sequencing results of PCR products amplified by albA-g-F/hph-R (a) and hph-F/albA-g-R (b) using the genomic DNA of albino colonies XM6 as templates. The red letters represent the protospacer sgRNA-albA1, and the yellow shaded red letters represent the PAM site. The green letters represent the to-be-inserted hph cassette, and blue letters represent the homology arms in the donor DNA MHi-albA-hph. XM6.1-6.11 represent the selected albino colonies. **Table S1.** A. niger strains used in this study. **Table S2.** Plasmids used in this study. **Table S3.** Primers used in this study. Restriction sites are underlined. Fm represents forward primer with modification and Rm represents reverse primer with modification. The modified additional sequences were represented in lowercase letters. **Table S4.** DNA sequences of codon optimized *cas9* used in this study. Black letters indicate the codon-optimized *cas9* gene. Purple letter indicate the NLS sequences of SV40 at 5′-termini and nucleoplasmin at 3′-termini. Green letters indicate the *gfp* (S65T) gene from pMF272. Orange letters indicate the (G$_4$S)$_3$ linker sequence. **Table S5.** DNA sequences of sgRNA constructs used in this study. Green letters indicate the promoter region for sgRNA expression. Orange letters indicate the transcription start of *U6* promoters. Blue letters indicate the sgRNA scaffold. Blue underlined letters indicate *BbsI* restriction sites. Red letters indicate the terminator of *RNU6* gene. Red underlined letters indicate genetic targets. **Table S6.** DNA sequences of donor DNA used in this study. Blue letters indicate the homogenous arms located at the 5′and 3′ flanking region of the genetic target sites. Black lowercase letters indicate the selection marker cassettes.

Abbreviations

CRISPR: clustered regularly interspaced short palindromic repeats; Cas9: CRISPR associated protein 9; sgRNA: single guide RNA; DSB: double-strand DNA break; PAM: protospacer adjacent motif; HR: homologous recombination; NHEJ: non-homologous end joining.

Authors' contributions

ZP and SJ conceived the project. ZX designed and carried out the experiments. ZK assisted to carry out the experiments. ZX wrote the manuscripts. SJ, ZP and MY revised the manuscript. All authors read and approved the final manuscript.

Author details

[1] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Xiqidao 32, Tianjin Airport Economic Area, Tianjin 300308, China. [2] Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China. [3] University of Chinese Academy of Sciences, Beijing 100049, China.

Competing interests

The authors declare that there is no competing interests.

Funding

References

1. Papagianni M. Advances in citric acid fermentation by *Aspergillus niger*: biochemical aspects, membrane transport and modeling. Biotechnol Adv. 2007;25:244–63.
2. Meyer V, Wu B, Ram AF. *Aspergillus* as a multi-purpose cell factory: current status and perspectives. Biotechnol Lett. 2011;33:469–76.
3. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157:1262–78.
4. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32:347–55.
5. Fuller KK, Chen S, Loros JJ, Dunlap JC. Development of the CRISPR/Cas9 system for targeted gene disruption in *Aspergillus fumigatus*. Eukaryot Cell. 2015;14:1073–80.
6. Meyer V, Arentshorst M, El-Ghezal A, Drews AC, Kooistra R, et al. Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. J Biotechnol. 2007;128:770–5.
7. Liu R, Chen L, Jiang Y, Zhou Z, Zou G. Efficient genome editing in filamentous fungus *Trichoderma reesei* using the CRISPR/Cas9 system. Cell Discov. 2015;1:15007.
8. Zhang C, Meng X, Wei X, Lu L. Highly efficient CRISPR mutagenesis by microhomology-mediated end joining in *Aspergillus fumigatus*. Fungal Genet Biol. 2016;86:47–57.
9. Pohl C, Kiel JA, Driessen AJ, Bovenberg RA, Nygard Y. CRISPR/Cas9 based genome editing of *Penicillium chrysogenum*. ACS Synth Biol. 2016;10:1021.
10. Nodvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS ONE. 2015;10:e0133085.
11. Sarkari P, Marx H, Blumhoff ML, Mattanovich D, Sauer M, et al. An efficient tool for metabolic pathway construction and gene integration for *Aspergillus niger*. Biores Technol. 2017;245:1327–33.
12. Kuivanen J, Wang YMJ, Richard P. Engineering *Aspergillus niger* for galactaric acid production: elimination of galactaric acid catabolism by using RNA sequencing and CRISPR/Cas9. Microb Cell Fact. 2016. https://doi.org/10.1186/s12934-016-0613-5.
13. Kuivanen J, Arvas M, Richard P. Clustered genes encoding 2-keto-l-gulonate reductase and l-idonate 5-dehydrogenase in the novel fungal d-glucuronic acid pathway. Front Microbiol. 2017;8:225.
14. Arazoe T, Miyoshi K, Yamato T, Ogawa T, Ohsato S, et al. Tailor-made CRISPR/Cas system for highly efficient targeted gene replacement in the rice blast fungus. Biotechnol Bioeng. 2015;112:2543–9.
15. Katayama T, Tanaka Y, Okabe T, Nakamura H, Fujii W, et al. Development of a genome editing technique using the CRISPR/Cas9 system in the industrial filamentous fungus *Aspergillus oryzae*. Biotechnol Lett. 2016;38:637–42.
16. Schuster M, Schweizer G, Reissmann S, Kahmann R. Genome editing in *Ustilago maydis* using the CRISPR-Cas system. Fungal Genet Biol. 2016;89:3–9.
17. Liu Q, Gao R, Li J, Lin L, Zhao J, et al. Development of a genome-editing CRISPR/Cas9 system in thermophilic fungal *Myceliophthora* species and its application to hyper-cellulase production strain engineering. Biotechnol Biofuels. 2017;10:1.
18. Carvalho ND, Arentshorst M, Jin Kwon M, Meyer V, Ram AF. Expanding the ku70 toolbox for filamentous fungi: establishment of complementation vectors and recipient strains for advanced gene analyses. Appl Microbiol Biotechnol. 2010;87:1463–73.
19. Wanka F, Cairns T, Boecker S, Berens C, Happel A, et al. Tet-on, or Tet-off, that is the question: advanced conditional gene expression in *Aspergillus*. Fungal Genet Biol. 2016;89:72–83.
20. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31:827–32.
21. Brow DA, Guthrie C. Spliceosomal Rna *U6* Is remarkably conserved from yeast to mammals. Nature. 1988;334:213–8.
22. Nakade S, Tsubota T, Sakane Y, Kume S, Sakamoto N, et al. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat Commun. 2014;5:5560.

Aspergillus niger is a superior expression host for the production of bioactive fungal cyclodepsipeptides

Simon Boecker[1,2], Stefan Grätz[1], Dennis Kerwat[1], Lutz Adam[1], David Schirmer[1], Lennart Richter[1], Tabea Schütze[2], Daniel Petras[1], Roderich D. Süssmuth[1*] and Vera Meyer[2*]

Abstract

Background: Fungal cyclodepsipeptides (CDPs) are non-ribosomally synthesized peptides produced by a variety of filamentous fungi and are of interest to the pharmaceutical industry due to their anticancer, antimicrobial and anthelmintic bioactivities. However, both chemical synthesis and isolation of CDPs from their natural producers are limited due to high costs and comparatively low yields. These challenges might be overcome by heterologous expression of the respective CDP-synthesizing genes in a suitable fungal host. The well-established industrial fungus *Aspergillus niger* was recently genetically reprogrammed to overproduce the cyclodepsipeptide enniatin B in g/L scale, suggesting that it can generally serve as a high production strain for natural products such as CDPs. In this study, we thus aimed to determine whether other CDPs such as beauvericin and bassianolide can be produced with high titres in *A. niger*, and whether the generated expression strains can be used to synthesize new-to-nature CDP derivatives.

Results: The beauvericin and bassianolide synthetases were expressed under control of the tuneable Tet-on promoter, and titres of about 350–600 mg/L for bassianolide and beauvericin were achieved when using optimized feeding conditions, respectively. These are the highest concentrations ever reported for both compounds, whether isolated from natural or heterologous expression systems. We also show that the newly established Tet-on based expression strains can be used to produce new-to-nature beauvericin derivatives by precursor directed biosynthesis, including the compounds 12-hydroxyvalerate-beauvericin and bromo-beauvericin. By feeding deuterated variants of one of the necessary precursors (D-hydroxyisovalerate), we were able to purify deuterated analogues of beauvericin and bassianolide from the respective *A. niger* expression strains. These deuterated compounds could potentially be used as internal standards in stable isotope dilution analyses to evaluate and quantify fungal spoilage of food and feed products.

Conclusion: In this study, we show that the product portfolio of *A. niger* can be expanded from enniatin to other CDPs such as beauvericin and bassianolide, as well as derivatives thereof. This illustrates the capability of *A. niger* to produce a range of different peptide natural products in titres high enough to become industrially relevant.

Keywords: *Aspergillus niger*, Natural products, Non-ribosomal peptide synthetase, Tet-on, Enniatin, Beauvericin, Bassianolide, Precursor directed biosynthesis, Stable isotope dilution analysis

*Correspondence: roderich.suessmuth@tu-berlin.de;
vera.meyer@tu-berlin.de
[1] Department Biological Chemistry, Institute of Chemistry, Technische
Universität Berlin, Straße des 17. Juni 124, 10623 Berlin, Germany
[2] Department Applied and Molecular Microbiology, Institute
of Biotechnology, Technische Universität Berlin, Gustav-Meyer-Allee 25,
13355 Berlin, Germany

Background

Fungal cyclodepsipeptides (CDPs), such as enniatins, beauvericins, bassianolide or PF1022 (Fig. 1b), comprise a class of secondary metabolites produced by (mostly pathogenic) filamentous fungi including *Fusarium oxysporum*, *Beauveria bassiana* or *Rosellinia* spp. [1]. They exhibit a variety of different pharmaceutically relevant bioactivities, including antibiotic, anthelmintic, cytotoxic, phytotoxic, insecticidal and anti-retroviral activities [1–4]. Additionally, some fungal CDPs are promising lead structures for new anti-cancer drugs [5–9]. These molecules are symmetric and consist of *N*-methylated L-amino acids and D-hydroxycarboxylic acids, which are alternatively linked to each other by amide and ester bonds. These CPDs are produced by highly homologous iteratively working bi-modular non-ribosomal peptide synthetases (NRPSs), i.e. enniatin synthetase (ESyn), beauvericin synthetase (BeauvSyn),

Fig. 1 Biosynthesis and structures of fungal CDPs. **a** Proposed biosynthesis of fungal CDPs. The D-hydroxycarboxylic acid is activated by module 1, while the L-amino acid is activated and methylated by module 2. Both substrates are coupled until a hexa- or octapeptidol is formed. Release from the enzyme at the C_3 domain occurs under cyclization. According to the "parallel" model, elongation of the depsipeptide chain occurs by the addition of dipeptidol units. In the "linear" or "looping" model, single building blocks are attached to the growing depsipeptide chain, which is shuttled between the PCP_1 and $PCP_{2a/b}$ domains. **b** Structures of enniatin B, beauvericin, bassianolide and PF1022

bassianolide synthetase (BassSyn) and PF1022 synthetase (PFSyn) [1, 2, 10]. Each module is composed of a condensation (C), an adenylation (A), and a peptidyl carrier protein (PCP) domain, whereby module 2 is linked to a terminal PCP_{2b}-C_3 bidomain (Fig. 1a).

At module 1, a D-hydroxycarboxylic acid is activated at the adenylation domain (A_1) and covalently bound to the peptidyl carrier protein domain (PCP_1). The L-amino acid is activated at the A-domain of module 2 (A_2), which is bound to the adjacent PCP_{2a} domain and methylated at the methylation domain (Mt). For substrate elongation, two different models have been proposed. At the "parallel" model, the depsipeptide chain grows by the addition of a dipeptidol, consisting of a hydroxy acid and an *N*-methyl amino acid, previously coupled in the C_2 domain. In this model, either PCP_{2a} or PCP_{2b} act as a so-called waiting position until the next dipeptidol is formed. Ester bond formation as well as macrocyclization occurs at the C_3 domain. However, based on recent results by Yu et al. [11] and from our group [12], experimental evidence points to the "linear" or "looping" model: the elongation occurs by the attachment of a single building block (hydroxy acid or *N*-methyl amino acid), while the growing depsipeptide chain is passed between the PCP_1 and the $PCP_{2a/b}$ domains. Peptide bond formation is catalysed in the C_2 domain, while the C_3 domain catalyses ester bond formation and macrocyclization. In this model, the role of the double $PCP_{2a/b}$ domains remains unclear, as either one of the domains is sufficient for biosynthesis of the final product [11, 12]. It was proposed that C_1 has no direct catalytic function, because truncated CDP synthetases missing the C_1 domain are still functional. Thus, C_1 could rather have a stabilizing or supportive role during the catalytic cycle [12]. Currently, it is not clear which of these two models accurately represents NRPs activity, and more investigations are required to fully understand the underlying mechanism of CDP biosynthesis. Due to their high degree of similarity, fungal CDP synthetases are ideal systems for combinatorial biosynthesis approaches, such as module and domain swapping, to obtain novel 'new-to-nature' compounds [11–14].

One option for obtaining fungal CDPs is via chemical synthesis, and several such strategies have thus far been described. However, *N*-methylation of amino acids, racemisation during the coupling of hydroxyl acids, as well as the final cyclization step, severely limit the effectiveness of these approaches [2]. For instance, an improved protocol for the total synthesis of enniatin established by Ley and co-workers requires nine steps and results in an overall yield of 15% [15]. Recently, the same group established a protocol based on flow chemistry and were able to synthesize different natural and unnatural CDPs with higher

yields (32–52%) [16]. However, high amounts of solvents and costly catalysts make this process uneconomical. A novel chemical synthesis approach using salt additives to support ring formation has been described to synthesize bassianolide, its closely related CDP verticilide [17], and a number of unnatural CDPs with varying ring sizes [18]. Although yields of bassianolide (9%) were almost twice as high as for the first total synthesis published (5.9%) [19], these overall yields are comparably low for production purposes. An alternative and more sustainable way to produce CDPs (and more generally natural products) is by using a biotechnological approach. Here, it is advantageous to transfer the biosynthetic pathway of the natural product of interest from a microbiologically challenging, genetically intractable, or even pathogenic organism into a safe, genetically amendable and industrially established heterologous production host. In the case of CDPs, natural production strains have been established and the highest titres reported for beauvericin production by *Fusarium oxysporum* KFCC 11353P and *Fusarium redolens* Dzf2, which range between 400 and 420 mg/L, respectively [20, 21]. However, not many tools for their genetic modification are available. Production of fungal CDPs in heterologous bacterial hosts has been established, but only low titres were achieved. In the case of beauvericin biosynthesis in *Escherichia coli*, only 8 mg/L were produced [22]. Additionally, enniatin production using *Bacillus subtilis* yielded titres which were also only in the mg/L range [23]. Encouragingly, when *Saccharomyces cerevisiae* was used as heterologous host, higher CDP titres were reported: 74.1 mg/L for beauvericin and 26.7 mg/L for bassianolide [24]. Recently, we were able to show that the industrial fungus *Aspergillus niger*, well-known for its high level production of organic acids and secreted proteins [25], is a promising host for heterologous production of enniatin. In this study, the ESyn encoding gene was put under control of the inducible Tet-on expression system [26] allowing high enniatin titres up to 4.5 g/L upon addition of the inducer doxycycline (Dox) [27]. This strain relies on feeding with the substrate D-hydroxy isovalerate, as it lacks the ketoisovalerate reductase gene *kivR* responsible for the generation of D-Hiv from 2-ketoisovalerate [28]. Autonomous expression strains of *A. niger* independent of D-Hiv feeding were additionally established. In these strains, the *kivR* gene was either monocistronically or polycistronically co-expressed with the ESyn gene [27, 29].

In the present study, we determined whether the *Beauveria bassiana* CDPs beauvericin and bassianolide can also be produced in *A. niger* with high titres. Furthermore, we aimed to test whether the *A. niger* production strains, which lack the ketoisovalerate reductase gene *kivR*, can be used to generate new-to-nature beauvericin

derivatives by precursor directed biosynthesis, which ultimately generated CDP variants that are accessible for further downstream chemical modifications.

Results and discussion

Generation of *A. niger* strains expressing BeauvSyn and BassSyn

Aspergillus niger is an excellent production organism for the synthesis of the hexamer enniatin, which consists of the two building blocks L-valine (L-Val) and D-hydroxy isovalerate (D-Hiv). To show that other CDPs relying on different precursor compositions or different ring sizes can be produced with high titres, our aim was to establish new production strains in an analogous fashion. Therefore, the Tet-on driven expression plasmids pDS8.2 (harbouring *bbBeas* encoding BeauvSyn), and pSB22.3 (harbouring *bbBsls* encoding BassSyn), were constructed and transformed into the *A. niger* strain AB1.13 (see Methods). This isolate is a useful production platform due to reduced protease activities [30]. Transformants carrying a single copy of the expression constructs integrated at the *pyrG* locus were verified by PCR and Southern blot (Additional file 1: Figure S1). Positive strains were cultivated as previously described [27], specifically in 20 mL media in shake flasks, which were then tested for production of the respective CDP. The metabolites were extracted from the dried biomass of the transformants and analysed by LC–MS. The identity of beauvericin and bassianolide was verified by tandem mass spectrometry (Additional file 1: Figures S2 and S3). The relative amounts of produced beauvericin and bassianolide were quantified by multiple reaction monitoring mass spectrometry and the beauvericin-producing strains DSc1.4 (single integration) and DSc1.5 (tandem integration), as well as the bassianolide-producing strain SB19.23 (single integration), which were each selected for further analysis.

Medium optimization and CDP purification

As indicated above, *A. niger* lacks the *kivR* gene, and consequently feeding of the precursor D-Hiv is necessary. To investigate the impact of precursor concentration on the product titres, D-Hiv as well as the respective amino acid were added to the culture broth in different concentrations. For better comparability, DSc1.4 and SB19.23 were chosen for these studies, as they have a single copy of the respective expression construct integrated at the *pyrG* locus. As shown in Table 1, titres of beauvericin and bassianolide are significantly increased by the addition of the respective precursors. Titres of beauvericin increase from 0.45 ± 0.13 mg/L to 293.62 ± 186.46 mg/L ($n = 4$) and of bassianolide from 1.04 ± 0.33 mg/L to 378.77 ± 59.74 mg/L ($n = 4$). Precursor concentrations

Table 1 Titres of beauvericin and bassianolide obtained in shake flask cultivations of *A. niger*

Concentration of amino acid and hydroxy acid precursor (mM)	Titre (mg/L)*	
	Beauvericin	**Bassianolide**
0	0.45 ± 0.13 (DSc1.4)	1.04 ± 0.33 (SB19.23)
2.5	83.42 ± 5.24 (DSc1.4)	45.90 ± 11.05 (SB19.23)
15	293.62 ± 186.46 (DSc1.4)	378.77 ± 59.74 (SB19.23)
	628.4 ± 211.1 (DSc1.5)	

Titres of beauvericin and bassianolide produced in the transformants DSc1.4, DSc1.5 and SB19.23 with and without precursor addition (D-Hiv and L-Phe for DSc1.4/DSc.15 and D-Hiv and L-Leu for SB19.23)

Respective strains analysed are given in brackets. *: Titres are given in mean ± SD ($n = 4$ for strains DSc1.4 and SB19.23, $n = 5$ for strain DSc1.5)

higher than 15 mM did not further increase the titres (data not shown). Apparently *A. niger* is able to synthesize D-Hiv in small amounts, as both metabolites are formed, even without addition of any precursor. This was also observed for the production of enniatin in *A. niger* [27] and could be due to a relaxed substrate specificity of other endogenous reductases of *A. niger* such as the 2-dehydropantoate 2-reductase (An11g09950), which shows a 28% similarity to KivR from *F. oxysporum* (BLASTP, [31, 32]).

To investigate biomass and beauvericin/bassianolide accumulation over time, 20 mL shake flask cultivations of DSc1.4 and SB19.23 were performed with 15 mM addition of the respective precursors as described above. Routinely, micro talc particles were added to the shake flask cultures of *A. niger* in order to obtain a homogeneous macroscopic growth morphology [27]. Samples were taken 0, 8, 32, 56 and 80 h after induction with 20 µg/mL of Dox. Usually, induction of secondary metabolite production in fungi coincides with sporulation [33]; however, a Tet-On based expression approach allows uncoupling of CDP synthesis from natural secondary metabolite kinetics. As shown in Fig. 2, beauvericin and bassianolide were indeed synthesized during the exponential growth phase, but achieve their highest titres during late/post exponential growth phase (Fig. 2). To investigate whether copy number of the expression construct affects the amount of synthesized product, strain DSc1.5 (harbouring a tandem copy of the construct at *pyrG*) was cultivated using the same setting as described above. With 628.4 ± 211.1 mg/L ($n = 5$), the highest titre of beauvericin was achieved after 80 h of cultivation which is about two-fold higher compared to the single copy strain DSc1.4 (Table 1).

For purification of beauvericin and bassianolide, each 5×200 mL shake flask cultivations of strains DSc1.5 and

Fig. 2 Biomass, beauvericin and bassianolide accumulation over time in shake flask cultivations. **a** Shake flask cultivations of strain DSc1.4. **b** Shake flask cultivations of strain SB19.23. Titres and insoluble matter (micro talc particles and biomass) concentrations are given in means from biological triplicates, SD is indicated by error bars. Note that biomass accumulation in the beauvericin-producing strain DSc1.4 (**a**) is delayed compared to the bassianolide-producing strain SB19.23 (**b**)

SB19.23 were performed. Gene expression and CDP biosynthesis was induced 16 h post inoculation by the addition of 20 μg/mL Dox and each 15 mM of the respective precursors. Biomass was harvested after an overall cultivation time of 96 h, and beauvericin and bassianolide purified as described in the Methods section. The fractions of the HPLC runs containing only the respective CDP (Additional file 1: Figure S4), were pooled, acetonitrile evaporated, and the residues freeze dried. Overall, 306 mg of beauvericin and 172 mg of bassianolide could be purified from each 1 L culture medium. ^{1}H-NMR spectra were recorded for both compounds and verified their purity. The signals obtained (Additional file 1: Figure S7) are in full accordance with data from the literature [7, 34].
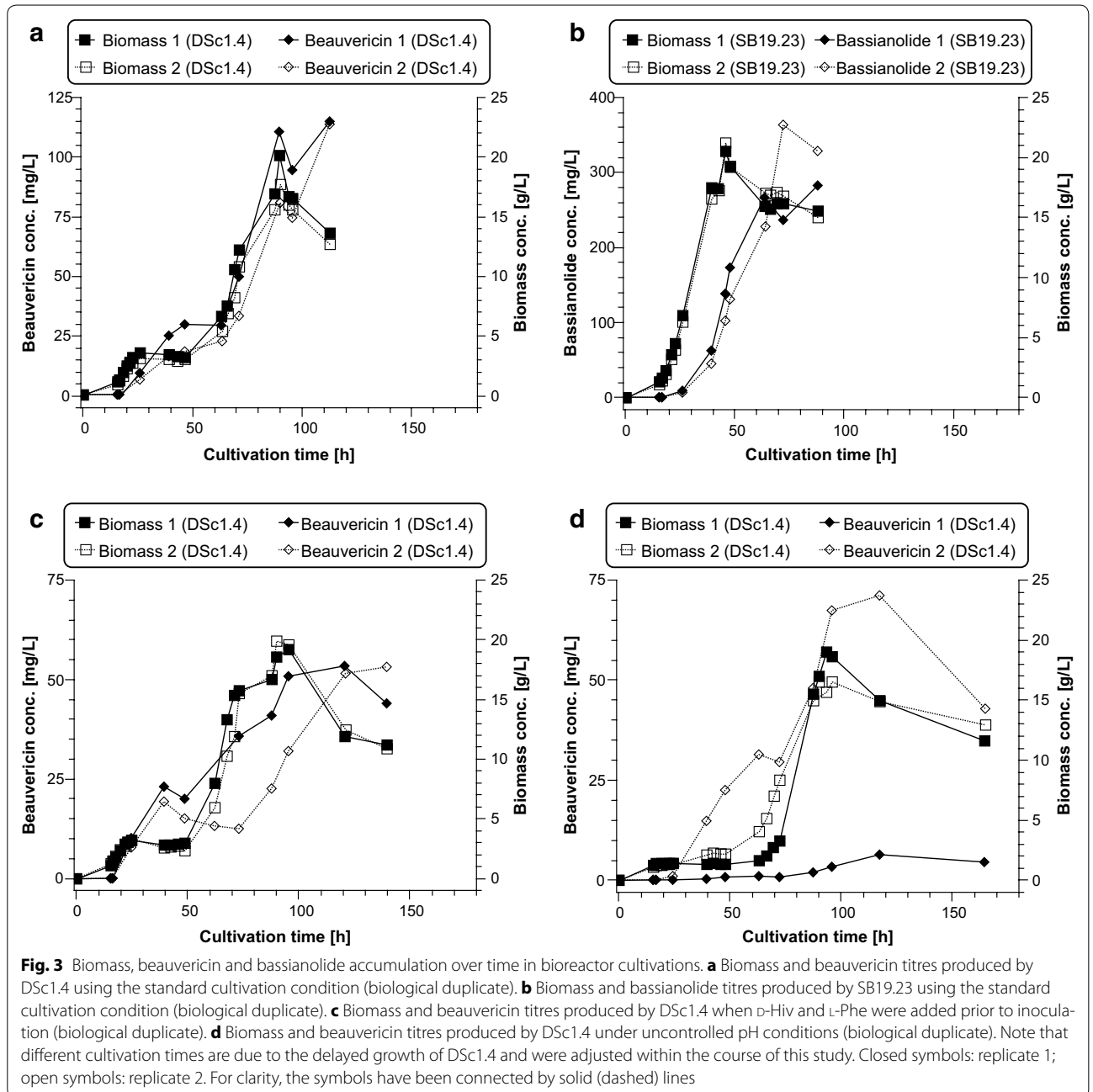
Bioreactor scale production of CDPs

Bioreactor cultivations allow tight control of culture conditions (e.g. temperature, pH, dissolved oxygen), and a better nutrient uptake compared to shake flask cultivations, and are thus better suited to perform highly reproducible fermentations. We thus carried out eight independent bioreactor runs in order to analyse the performance and productivity of the single copy beauvericin-producing strain DSc1.4 and the single copy bassianolide-producing strain SB19.23 in more detail.

For the first set of fermentations, expression was induced with 20 μg/mL of Dox and the respective precursors (amino acid and hydroxy acid) were each added to a final concentration of 15 mM 16 h post inoculation (standard cultivation conditions). The two DSc1.4 bioreactor experiments show an unusual growth behaviour

compared to strain SB19.23 (Fig. 3a, b). DSc1.4 stops growing after approximately 24 h post inoculation for about 36 h and resumes growth after nearly 60 h until it reaches its maximum biomass concentration (20.0 and 17.6 g/L) after 90 h of cultivation. Such a growth profile is reminiscent of a diauxic shift, a known growth phenomenon that occurs when different C- or N-sources are simultaneously present in a medium [35]. In contrast, strain SB19.23 grows continuously as expected with maximal growth rates of $\mu_{max} = 0.178$ and 0.174 h^{-1}, respectively, until it reaches its maximum biomass concentration (20.5 and 21.2 g/L) after 46 h. Notably, DSc1.4 also showed a delayed growth behaviour during shake flask cultivations (Fig. 2). Most interestingly, the bassianolide titres of strain SB19.23 obtained from bioreactor runs are in the same range as measured for shake flask cultivations (~ 300 mg/L), whereas the beauvericin titres of DSc1.4 obtained in bioreactor runs (~ 100 mg/L) were 3–4 fold lower compared to shake flask cultivations. We speculated that the diauxic shift observed in DSc1.4 cultivations might be causatively linked to the addition of L-Phe to the medium 16 h post inoculation. L-Phe is known to serve as an alternative nitrogen as well as carbon source to fungi [36]. When D-Hiv and L-Phe were added at the beginning of the bioreactor cultivations, the diauxic growth curve was still observed, and beauvericin titres even dropped to 50 mg/L (Fig. 3c).

To further investigate the phenomenon of decreased beauvericin titres obtained from bioreactor runs compared to shake flask cultivations, a third set of experiments was performed. Another parameter that is different between shake flask and bioreactor cultivations

Fig. 3 Biomass, beauvericin and bassianolide accumulation over time in bioreactor cultivations. **a** Biomass and beauvericin titres produced by DSc1.4 using the standard cultivation condition (biological duplicate). **b** Biomass and bassianolide titres produced by SB19.23 using the standard cultivation condition (biological duplicate). **c** Biomass and beauvericin titres produced by DSc1.4 when D-Hiv and L-Phe were added prior to inoculation (biological duplicate). **d** Biomass and beauvericin titres produced by DSc1.4 under uncontrolled pH conditions (biological duplicate). Note that different cultivation times are due to the delayed growth of DSc1.4 and were adjusted within the course of this study. Closed symbols: replicate 1; open symbols: replicate 2. For clarity, the symbols have been connected by solid (dashed) lines

is ambient pH: The pH of the medium used in shake flask cultivations is set to pH 5.6 and uncontrolled during cultivation, whereas the pH of the fermentation medium in bioreactor cultivations is adjusted to pH 3.0 and kept constant during fermentation. Thus, the pH of the medium used for the final two bioreactor runs was set to pH 5.6 and the pH control system switched off. Precursors were added at the same time point as in the shake flask cultivations (after 16 h). The growth behaviour of DSc1.4 still followed a pronounced diauxic shift pattern

and beauvericin titres were again lower in comparison to shake flask cultivations (Fig. 3d). Also, the overall growth of the strain in these two runs was slower (especially of replicate 1) compared to the previous runs (Fig. 3a, c) and the beauvericin titre was, with 6.3 mg/L for replicate 1, the lowest of all runs (Fig. 3d). From this experiment, we concluded that a controlled, low ambient pH is in favour of growth of *A. niger* and thus (because Tet-on driven expression couples CDP synthesis with growth) high beauvericin titres. Additionally, growth as well as

beauvericin production showed a high deviation under uncontrolled pH conditions. An overview of all titres obtained from bioreactor cultivations is summarised in Table 2.

While all medium parameters (except pH) were comparable in shake flask and bioreactor cultivations, oxygen supply differed considerably. In bioreactor cultivations, *A. niger* is constantly supplied with 2 L air/min while *A. niger* likely encounters hypoxic conditions in shake flask cultivations, as the cotton plugs do not guarantee optimal gas transfer into the medium [37]. The nitrogen source in the medium used for both bioreactor and shake flask cultivations is nitrate. Nitrate is known as a secondary nitrogen source: under normoxic conditions, it is reduced to nitrite and ammonium before assimilated into the biomass [38], whereas under hypoxic conditions, it can additionally serve as a terminal electron acceptor during energy generation, a process called nitrate respiration [39]. As under normoxic conditions nitrate needs first to be reduced to ammonium, it could be energetically more feasible for *A. niger* to use L-Phe
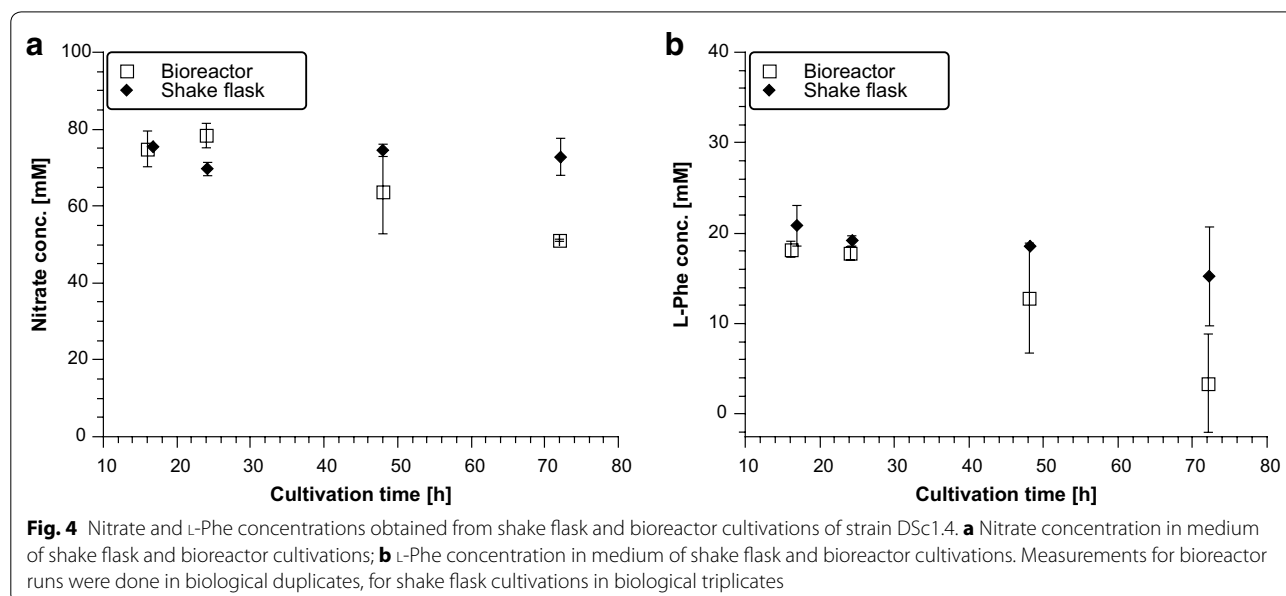
as nitrogen source, as shown for the model fungus *Neurospora crassa* [40]. Thus, more L-Phe could be used for biomass production rather than beauvericin biosynthesis, which would offer an explanation for reduced beauvericin titres during bioreactor cultivations. In any case, one would hypothesize that both nitrate and L-Phe are consumed faster during shake flask cultivations in comparison to bioreactor cultivations. This is what we could indeed observe (Fig. 4). From the data obtained, it cannot be deduced how much L-Phe is used for beauvericin synthesis and how much for biomass accumulation. Clearly, most of the added L-Phe is used for biomass accumulation in both cultivation conditions, as the L-Phe concentration decreased from 18.2 to 3.4 mM in shake flasks and from 20.8 to 15.2 mM in bioreactors, which would correspond to theoretical beauvericin titres of 3.86 g/L in shake flasks and 1.47 g/L in bioreactors respectively. Recently, Kniemeyer and co-workers studied the importance of oxygen on secondary metabolism in *A. fumigatus*. They showed that the production of some secondary metabolites depends on both the developmental stage of

Table 2 Titres of beauvericin and bassianolide obtained from bioreactor cultivations

Cultivation condition	Maximum titre (mg/L)	
	Beauvericin (run 1/run 2)	Bassianolide (run 1/run 2)
Standard cultivation condition	113.9/112.9	281.8/362.5
Precursor addition after 0 h	53.1/52.9	n.d.
No pH control	70.7/6.3	n.d.

Titres of beauvericin and bassianolide produced in transformants DSc1.4 and SB19.23 under standard and modified cultivation conditions (for details see text). Bioreactor runs were performed in duplicates (run 1 and run 2)

n.d. not done

Fig. 4 Nitrate and L-Phe concentrations obtained from shake flask and bioreactor cultivations of strain DSc1.4. **a** Nitrate concentration in medium of shake flask and bioreactor cultivations; **b** L-Phe concentration in medium of shake flask and bioreactor cultivations. Measurements for bioreactor runs were done in biological duplicates, for shake flask cultivations in biological triplicates

the fungus and on the available oxygen level [41–43]. The PKS/NRPS product pseurotin A, for example, is strongly up-regulated under hypoxia [41]. In the Tet-on based *A. niger* expression strains used in this study, transcription of the BeauvSyn is likely not affected by the concentration of dissolved oxygen in the medium, however, expression levels of other enzymes involved in the synthesis of beauvericin [e.g. phosphopantetheinyl transferases, CoA synthesis, *S*-adenosyl methionine synthesis] could be positively linked to a lower oxygen supply in shake flask cultures. Interestingly, many genes involved in L-Phe biosynthesis of *A. fumigatus* are higher expressed under hypoxic in comparison to normoxic conditions [43]. Also genes involved in the metabolism of secondary products derived from L-phenylalanine, and in the *S*-adenosyl-methionine-homocystein cycle were up-regulated under hypoxic conditions [43]. Overall, these observations support the hypothesis that higher beauvericin titres in shake flask cultivations of *A. niger* might be mechanistically linked with reduced oxygen levels and thus altered L-Phe metabolism. Further analysis to prove/disprove this speculation goes beyond the scope of this article; however, it clearly demonstrates that further optimization studies regarding beauvericin overproduction in *A. niger* should also consider metabolic engineering aspects that address precursor availability and consumption.

Generation of non-natural CDP derivatives

We next tested whether the beauvericin and bassianolide producing *A. niger* strains DSc1.4 and SB19.23 can be used to produce new-to-nature beauvericin and bassianolide derivatives by exploiting the relaxed substrate specificity of A_1 domains towards D-Hiv [44–46]. In non-autonomous production strains, substrate analogues do not compete with the natural substrates, and a precursor directed biosynthesis approach can therefore be applied [47, 48]. This technique, also called mutasynthesis or mutational biosynthesis, has been previously applied to obtain non-natural beauvericins from the heterologous host *E. coli*, as well as from a *kivR* deletion mutant of the natural producer *B. bassiana*. However, the titres for most of the non-natural beauvericin analogues stayed in the low mg/L range and only beauvericin analogues could be isolated when the alternative hydroxy acid displayed similar properties as D-Hiv (aliphatic side chains) [45, 47].

Here, we tested seven different D-Hiv analogues, which were fed to strains DSc1.4 and SB19.23 in shake flask cultivations in final concentrations of 7.5 mM. We decided to use lower feeding concentrations of hydroxy acids to decrease the risk of toxic effects of the precursors on *A. niger*. As a consequence, the corresponding amino acid was also added at 7.5 mM to the cultures. The strains

were cultivated in 20 mL scale as described for the natural beauvericin and bassianolide derivatives, the biomass extracted, and the extracts analysed by LC–MS. From all tested compounds, only 2-hydroxyvalerate and 3-bromo-lactate were incorporated by DSc1.4 into the beauvericin backbone, leading to the novel beauvericin analogues 2-hydroxyvalerate-beauvericin and bromo-beauvericin (Fig. 5b and Additional file 1: Figure S8). Interestingly, a beauvericin derivative with incorporated propargyl-lactate could not be detected, although the hydroxy acid had been accepted in in vitro and in vivo experiments in *E. coli* [45]. This could be due to a metabolization of the acid by *A. niger* as also observed in *Corynebacterium glutamicum* [49]. Surprisingly, SB19.23 was not able to synthesize any artificial bassianolide derivative, suggesting that the substrate specificity of the BassSyn is tightly controlled compared to BeauvSyn. Alternatively one may assume that the altered side chains with their sterically
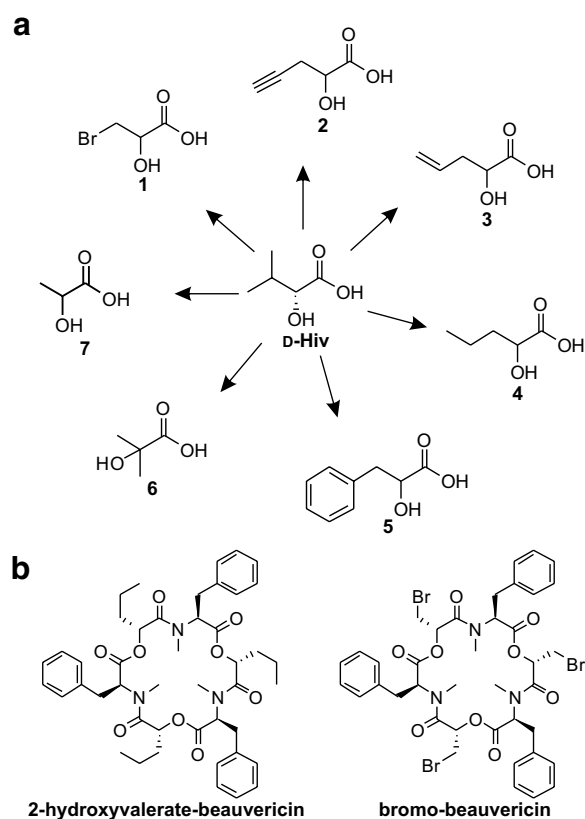
Fig. 5 Hydroxy acids fed to strains DSc1.4 and SB19.23 in order to synthesize new-to-nature beauvericins. **a** Seven artificial hydroxy acid analogues (1–7) were fed to strains DSc1.4 and SB19.23, varying in the length of the side chain (4, 6, 7), bearing a halogen (1), aromatic (5), alkene (3) or alkyne (2) functionality. Final precursor concentration was 7.5 mM. **b** Non-natural beauvericins (2-hydroxyvalerate-beauvericin and bromo-beauvericin) which were successfully produced by strain DSc1.4

more demanding substituents interfere with ring formation of an cyclooctadepsipeptide.

Because feeding of bromo-lactate resulted in significant amounts of bromo-beauvericin, we wanted to assess the capacity of this approach for production and isolation on a larger scale. As a preliminary experiment, different concentrations of the racemic precursor D/L-bromolactate (5, 10 and 15 mM) were added to DSc1.4 in small scale cultivations (20 mL scale). The addition of 10 mM of D/L-bromo-lactate gave higher titres than the addition of 5 mM, while no bromo-beauvericin could be detected in the cultures supplemented with 15 mM of the hydroxy acid. This coincided with significant less biomass formation, suggesting that high concentrations of bromo-lactate have toxic effects on *A. niger*. Based on these results, a concentration of 10 mM of D/L-bromolactate was chosen for large scale cultivation in shake flasks. 12 mg of pure bromo-beauvericin (preparative HPLC) could be successfully obtained as a colourless powder from a 1.1 L of culture of DSc1.4, the purity of which was proven by mass spectrometry and NMR (Additional file 1: Figures S9 and S10). The purified bromo-beauvericin was tested for antimicrobial and antiparasitic activity together with purified enniatin B, beauvericin and bassianolide. While bromo-beauvericin did not show any improved antimicrobial or antiparasitic activity compared to the other compounds, it interestingly did not show any cytotoxic effects against a mammalian cell line at a concentration of 100 μg/mL, whereas the IC_{50} value of natural beauvericin is 1.52 μg/mL. It is thus worth studying bromo-beauvericin further as a potential future antiparasitic drug (Additional file 1: Table S4).
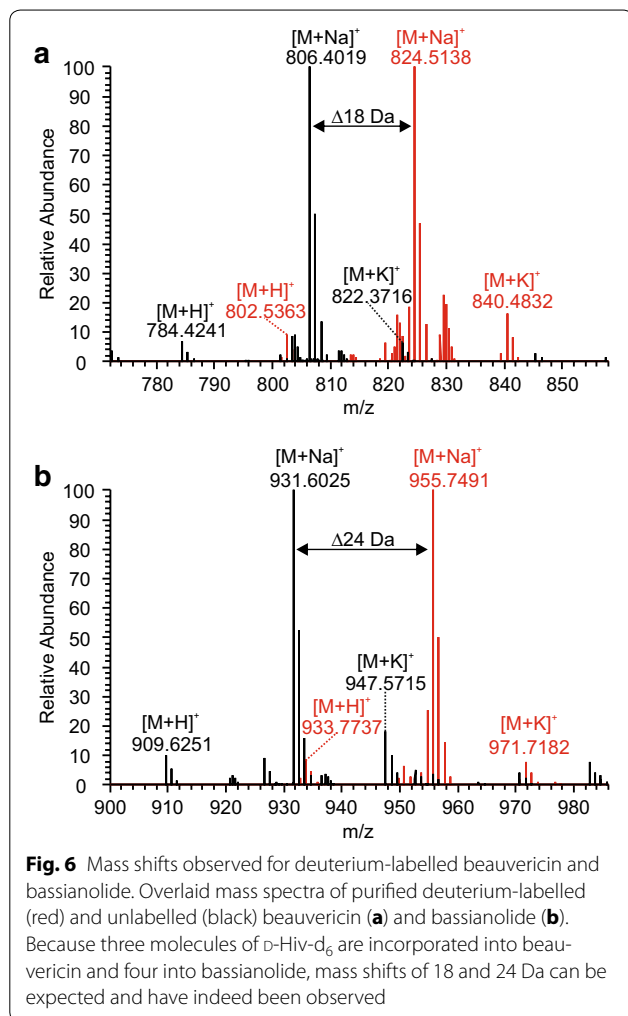
Generation of deuterated CDP standards
Fungal CDPs are not only of pharmaceutical interest as lead structures, but are also prominent contaminants (especially enniatins and beauvericin) of food and feed, as most of their natural producers are plant pathogenic fungi [50–52]. Thus, robust, fast, and exact analytical methods are needed to detect and quantify these compounds, even in trace amounts, in both food and feed products suspected to be spoiled by fungi. Most described protocols are based on LC–MS measurements in combination with an external standard calibration curve [53–56]. However, these methods are only exact to a certain degree as they do not consider effects of the matrix which can lead to ion suppression or ion enhancement [57, 58]. Furthermore, the recovery rates of the analytes from biological samples may vary, which would also lead to altered results [53]. Stable isotope dilution assays are superior to methods using external standards as they guarantee exact quantifications also of mycotoxins in grain products [58–61]. The biosynthesis of $^{15}N_3$-labelled

standards of enniatins and beauvericin in *F. sambucinum* (enniatin producer) and *F. fujikuroi* (beauvericin producer) grown on $Na^{15}NO_3$ as sole nitrogen source has been reported [62]. With 430 μg (enniatin A), 450 μg (enniatin A1), and 1460 μg (beauvericin) of ^{15}N-labelled compound purified from 500 mL of culture, titres are low while the price of the medium is relatively high.

We thus tested the transformants DSc1.4 and SB19.23 for their suitability to synthesize labelled beauvericin and bassianolide. Instead of ^{15}N, we used deuterium (^2H) to label D-Hiv. D/L-Hiv-d_6 was synthesized using acetone-d_6 as starting material. The respective expression strains were cultivated in shake flasks (100 mL scale) and 15 mM of D/L-Hiv-d_6 together with 7.5 mM of the respective L-amino acid were added. Because a racemic mixture of D/L-Hiv-d_6 was fed to the cultures, lower titres of beauvericin and bassianolide were expected compared to the purification of unlabelled metabolites. Cultivation and purification of both compounds were carried out as described above. From 100 mL of culture, 8.2 mg of beauvericin-d_{18} and 5.7 mg of bassianalide-d_{24} were isolated in analytically pure form (Fig. 6). Retention times between natural and deuterated variants of beauvericin and bassianolide differed slightly on a C18 reverse phase column. This effect has also been described earlier and is due to subtle differences in the polarity of labelled and unlabelled compounds (Additional file 1: Figure S6) [63].

Because sample preparation for LC–ESI–MS analysis from complex matrices (e.g. fungal biomass or grain products) can be laborious and time-consuming, especially if many samples need to be tested, it was evaluated whether the deuterated beauvericin and bassianolide standards could also be used for quantification on a MALDI-TOF instrument. In contrast to LC–ESI–MS, where the samples have to be pre-purified in order to keep the ion source clean, crude extracts can be directly applied to MALDI-TOF. To test this, defined amounts of deuterated beauvericin and bassianolide were added to a dilution series of the unlabelled compounds (Additional file 1: Table S5). The ratio of the peak areas of the sodium adducts (most abundant peaks) of the respective labelled and unlabelled compounds were plotted against the ratio of the concentration of the compounds to determine if a linear relation was observed. The ratios of the concentrations and the peak areas indeed show a linear correlation (Additional file 1: Figure S11). However, as pointed out by the coefficients of determination ($R^2 = 0.946$ for the beauvericin measurements and $R^2 = 0.988$ for the bassianolide measurements), quantification of both compounds is not exact. One problem of the MALDI-TOF measurements is that only a direct MS is being recorded and that MS/MS of labelled and unlabelled compound cannot be concomitantly measured. Thus, any compound showing

Fig. 6 Mass shifts observed for deuterium-labelled beauvericin and bassianolide. Overlaid mass spectra of purified deuterium-labelled (red) and unlabelled (black) beauvericin (**a**) and bassianolide (**b**). Because three molecules of D-Hiv-d_6 are incorporated into beauvericin and four into bassianolide, mass shifts of 18 and 24 Da can be expected and have indeed been observed

the same or a very similar m/z value to the labelled or unlabelled analytes would alter the results. Nevertheless, the MALDI-TOF measurement is an interesting alternative for high-throughput applications where an approximate estimation of CDP concentrations can be tolerated (e.g. for screening many samples or strains). For exact measurements however, LC–MS/MS remains the method of choice as reviewed in [61].

Conclusions

Cyclodepsipeptides are potential new drugs for various medicinal applications. Enniatin and beauvericin are discussed as anticancer drugs [8, 9], the PF1022A derivative emodepside is currently in phase I clinical trials and tested as an anthelmintic drug for use in humans [64]. Hence, sustainable approaches to synthesize these compounds at a large scale and in a commercially viable way have become a recent focus in biotechnology. In this and our previous study [27], we have shown that *A. niger* is an

excellent expression host for CDPs, including enniatin, beauvericin and bassianolide. The CDP titres reached are, to our knowledge, the highest ever reported for heterologous producers, as well as for natural producing organisms, and are summarised in Table 3. The fact that *A. niger* tolerates and accepts unnatural building blocks makes it furthermore an attractive platform for the production of new-to-nature CDPs by mutasynthesis, as we have demonstrated for the beauvericin derivative bromo-beauvericin. The amounts isolated were indeed high enough for bioactivity testing, which showed comparable antiparasitic activity and, interestingly, markedly reduced cytotoxicity. We have also demonstrated that *A. niger* can be genetically trimmed to synthesize deuterated CDPs which can be exploited as internal standards to evaluate mycotoxin burden of food by stable-isotope dilution assays.

Methods

Strains and general cloning procedures

Plasmids, primers and strains used in this study are summarized in Additional file 1: Tables S1–S3. Molecular techniques for *E. coli* followed protocols described earlier [65]. *A. niger* transformation and genomic DNA extraction from selected transformants was done according to [66]. The BeauvSyn and BassSyn encoding genes *bbBeas* (GenBank accession number EU886196) and *bbBsls* (GenBank accession number FJ439897) were amplified from the genomic DNA of *B. bassiana* ATCC 7159 using the primer pairs Beauv_InFusion1_fw/Beauv_InFusion3_rv for *bbBeas* and Bass_InFusion1_fw/Bass_InFusion3_rv for *bbBsls*, respectively. The amplicons were ligated into the cloning vector pJET2.1 (Thermo Fisher Scientific Inc.), resulting in pDS2.1 (harbouring *bbBeas*) and pDS1.9 (harbouring *bbBsls*) and verified by restriction analysis and sequencing. Direct cloning of the full-length genes into the *A. niger* Tet-on expression vector pVG2.2 was not successful. Thus, the genes were split into three parts of approximately 3 kbp length and 15 bp overhangs to each other and the *Pme*I-linearized vector pVG2.2. pDS2.1 and pSB1.9 were used as templates and primer pairs Beauv_InFusion1_fw/Beauv_InFusion1_rv, Beauv_InFusion2_fw/Beauv_InFusion2_rv, Beauv_InFusion3_fw/Beauv_In-Fusion3_rv for *bbBeas* and Bass_InFusion1_fw/Bass_InFusion1_rv, Bass_In-Fusion2_fw/Bass_InFusion2_rv, Bass_InFusion3_fw/Bass_InFusion3_rv for *bbBsls* to amplify the respective gene fragments. The amplicons were ligated and assembled into the *Pme*I-linearized Tet-on expression plasmid pVG2.2 via the In-Fusion® HD Cloning Kit (Clontech), resulting in plasmids pDS8.2 (harbouring *bbBeas*) and pSB22.3 (harbouring *bbBsls*).

Table 3 Summary of highest CDP titres ever reported for bacterial and fungal expression hosts

CDP	Native producer	Heterologous producer	
Enniatin B	5 g/L (enniatin B and other derivatives) *F. oxysporum* Shake flask cultivation [70]	1.1 mg/L *B. subtilis* Shake flask cultivation [23]	950 mg/L *A. niger* Shake flask cultivation [27]
			4.5 g/L *A. niger* Bioreactor cultivation [27]
Beauvericin	420 mg/L *F. oxysporum* Shake flask cultivation [20]	74.1 ± 0.3 mg/L *S. cerevisiae* Shake flask cultivation [24]	628.4 ± 211.1 mg/L *A. niger* Shake flask cultivation [This study]
	22.3 ± 1.5 mg/L *B. bassiana* Shake flask cultivation [24]	8 mg/L *E. coli* Shake flask cultivation [22]	113.4 mg/L *A. niger* Bioreactor cultivation [This study]
Bassianolide	18.2 ± 0.6 mg/L *B. bassiana* In shake flask cultivation [24]	26.7 ± 2.8 mg/L *S. cerevisiae* Shake flask cultivation [24]	378.77 ± 59.74 mg/L *A. niger* Shake flask cultivation [This study]
			322.12 mg/L *A. niger* Bioreactor cultivation [This study]

For simplicity, respective strain names, genetic modifications and details on cultivation conditions (medium composition, cultivation time, feeding conditions, etc.) are not indicated for the different cell factories. This data can be extracted from the references given

Shake flask cultivations of *A. niger*

For production of CDPs, transformers were cultivated in 20 mL or 200 mL enniatin production medium (EM) as described in [27] if not indicated otherwise. Hydroxy and amino acid precursors were added in the range of 0–25 mM. Cultures were inoculated with 5×10^6 spores/mL and Tet-On driven expression induced with 20 μg/mL doxycycline 16 h after inoculation.

Bioreactor cultivations of *A. niger*

Submerged cultivations were performed with Biostat bioreactors (Sartorius, Göttingen, Germany, 4 L working volume) as described before [27]. Glucose-limited batch cultivation was initiated by inoculation of fermentation medium (CM with 5% of glucose: 7 mM KCl, 11 mM KH_2PO_4, 70 mM $NaNO_3$, 2 mM $MgSO_4$, 1x trace element solution [67], 0.1% casamino acids, 0.5% yeast extract, 5% glucose) with conidial suspension of *A. niger* transformants to give 10^9 conidia L^{-1}. Glucose was sterilized separately from the fermentation medium. Temperature of 26 °C and pH 3 were kept constant if not stated otherwise, the latter by computer controlled addition of 2 M NaOH or 1 M HCl. Computer-controlled base addition to the culture broth was used as an indirect growth measurement [68]. When the culture reached the early exponential growth phase (about 16 h after inoculation, corresponds to 1 g biomass dry weight kg^{-1}), Dox (20 μg/

ml), D-Hiv (15 mM) and L-Phe or L-Leu (15 mM) were added.

Purification of CDPs

Purification of CDPs from *A. niger* biomass was adapted from [27]. In brief, the mycelium from a 1 L culture was harvested by suction filtration and lyophilized. The dried mycelium was ground in a mortar and extracted three times with 300 mL of EtOAc. The solvent was evaporated and the brownish residue filtered over a short silica column (*n*-hexanes/EtOAc = 50:50). The solvents were evaporated and the residues resolved in methanol. Insoluble residues were removed by filtration and the solvent evaporated. For beauvericin and bassianolide purification, the residues were dissolved in acetonitrile/water (80:20) and the solution was centrifuged at $10,000 \times g$ for 15 min to remove insoluble particles. The supernatant was subjected to reversed phase chromatography using a GROM-Sil 120 ODS-5 HE (10 μm, 250 × 20 mm) column on an Agilent 1100 series preparative HPLC system running isocratically on acetonitrile (+ 0.1% formic acid)/water (+ 0.1% formic acid) (70:30) for beauvericin or with a linear gradient (70–100% acetonitrile over 15 min) for bassianolide with a flow rate of 15 mL/min. For bromo-beauvericin purification, the residues were resolved in MeOH and subjected to reversed phase chromatography using a GROM-Sil 120 ODS-5 HE (10 μm, 250 × 20 mm)

column on an Agilent 1100 series preparative HPLC system running isocratically on MeOH (+ 0.1% formic acid)/water (+ 0.1% formic acid) (81.5/18.5) with a flow rate of 15 mL/min. Fractions containing the respective CDP were pooled, acetonitrile and MeOH were evaporated and water was removed by freeze drying.

Analysis and quantification of produced CDPs

Biomass (which included in the case of shake flask cultivations also insoluble parts, i.e. talc particles) of a defined amount of culture broth was harvested by suction filtration and lyophilized and weighed. The biomass was ground and 25 mg were transferred to a 2 mL test tube and extracted with 1 mL of EtOAc, shaking overnight. The tubes were centrifuged at $13,000\times g$ and 700 µL of the extract were transferred to a new 1.5 mL test tube and evaporated. The residues were dissolved in 1 mL of water/isopropanol (50:50), diluted if necessary and the amount of CDPs quantified in MRM mode on an ESI-Triple-Quadrupol-MS 6460 Series (Agilent Technologies) coupled to an Agilent 1290 Infinity HPLC system (Agilent Technologies) equipped with an Agilent Poroshell 120 EC-C18 (3.0 × 50 mm) column (Agilent Technologies), heated to 50 °C. The mobile phases were H_2O (A) and isopropanol (B). The injection volume was set to 3 µl and the flow rate was 0.4 ml/min. The applied gradient was: 50–100% (0.0–3.2 min), 100% (3.2–4.5 min), 100–5% (4.5–4.6 min), 5% (4.6–5.6 min), 5–50% (5.6–5.7 min), 50% B (5.7–7.0 min). For beauvericin quantification, the *m/z* value for the precursor ion was set to 806.4 ([M + Na]$^+$ adduct) and for the fragment ion to 384.1 as quantifier, for bassianolide quantification, the *m/z* value for the precursor ion was set to 931.6 ([M + Na]$^+$ adduct) and for the fragment ion to 350.1 as quantifier. For every set of measurements, a new calibration curve was made using beauvericin or bassianolide isolated from *A. niger* transformants as an external standard. Peak areas were determined by manual integration using MassHunter Workstation Qualitative Analysis (Agilent Technologies). Exact masses of purified CDPs were recorded on an ESI-LTQ-Orbitrap-MS, Orbitrap XL (Thermo Fisher Scientific). Samples were dissolved in MeOH and measured by direct injection. Analysis was performed with the Xcalibur 2.2 software (Thermo Fisher Scientific). Retention times of labelled and unlabelled beauvericin and bassianolide were determined on an ESI-Orbitrap-MS, Exactive (Thermo Fisher Scientific) coupled to an Agilent 1260 Infinity HPLC system (Agilent Technologies) equipped with an Agilent Poroshell 120 EC-C18 (2.1 × 50 mm) column (Agilent Technologies). The mobile phases were H_2O + 0.1% formic acid (A) and acetonitrile + 0.1% formic acid (B). The injection volume was set to 2 µl and the flow rate was 0.4 ml/min.

The applied gradient was: 40% (0.0–0.5 min), 40–100% (0.5–12.0 min), 100% (12.0–13.5 min), 100–40% (13.5–13.6 min), 40% B (13.6–15.5 min). 1:1 mixtures of labelled and unlabelled beauvericin and bassianolide were injected and retention times of the sodium adducts of each compound determined. Analysis was performed with the Xcalibur 2.2 software (Thermo Fisher Scientific).

For MALDI-TOF analysis, 1 µL of purified CDPs or crude extracts of *A. niger* transformants, solved in MeOH or acetonitrile, were either mixed with 1 µL of saturated 2,5-dihydroxybenzoic acid (DHB) or α-cyano-4-hydroxycinnamic acid (CHCA) solution [dissolved in an acetonitrile–water mixture (1:1), acidified with formic acid (1%)]. 1 µL of the mixture was spotted onto a ground steel MALDI target plate and allowed to dry and crystallize. Measurements were carried out on a Bruker ultrafleXtreme MALDI-TOF–MS, equipped with a smartbeam II laser. The intensity of the laser was set to 50% with a frequency of 2 kHz. Calibration was done with the peptide calibration standard (Bruker). Analysis was performed with the Compass for flexSeries 1.4 software (Bruker).

For NMR analysis, purified CDPs were solved in $CDCl_3$ or MeOH-d_4 and ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker Avance III 700 MHz NMR spectrometer or Bruker Avance II 400 MHz NMR spectrometer. The signals of the non-deuterated solvent rests were used as standards. Chemical shifts are given in δ-units (ppm) relative to the solvent signal.

Synthesis of α-hydroxy acid precursors

Synthesis of D-Hiv and α-hydroxy acid analogues (Fig. 5) followed the procedures described in [46]. The synthesis of D/L-Hiv-d_6 followed the procedure described in [69].

Quantification of nitrate and L-Phe

Quantification of nitrate and L-Phe concentrations in the cultivation medium was performed with the Nitrate/Nitrite Colorimetric Assay Kit (Cayman Chemical) and the Phenylalanine Assay Kit (Sigma-Aldrich) according to the manufacturers' protocols.

Antimicrobial and antiparasitic test assays

Bioactivity assays were performed as described in [12].

Authors' contributions

SB, VM and RDS designed the experiments, SB and DS conducted the plasmid construction and generation of the *A. niger* transformants, SB purified beauvericin and bassianolide and deuterated variants thereof from *A. niger* and performed the quantification and analysis of generated CDPs, SB and TS

References
1. Süssmuth R, Müller J, von Döhren H, Molnár I. Fungal cyclooligomer depsipeptides: from classical biochemistry to combinatorial biosynthesis. Nat Prod Rep. 2011;28:99–124. https://doi.org/10.1039/c001463j.
2. Sivanathan S, Scherkenbeck J. Cyclodepsipeptides: a rich source of biologically active compounds for drug research. Molecules. 2014;19:12368–420. https://doi.org/10.3390/molecules190812368.
3. Sy-Cordero AA, Pearce CJ, Oberlies NH. Revisiting the enniatins: a review of their isolation, biosynthesis, structure determination and biological activities. J Antibiot (Tokyo). 2012;65:541–9. https://doi.org/10.1038/ja.2012.71.
4. Shekhar-Guturja T, Gunaherath GMKB, Wijeratne EMK, et al. Dual action antifungal small molecule modulates multidrug efflux and TOR signaling. Nat Chem Biol. 2016;12:867–75. https://doi.org/10.1038/nchembio.2165.
5. Chen B-F, Tsai M-C, Jow G-M. Induction of calcium influx from extracellular fluid by beauvericin in human leukemia cells. Biochem Biophys Res Commun. 2006;340:134–9. https://doi.org/10.1016/j.bbrc.2005.11.166.
6. Jow G-M, Chou C-J, Chen B-F, Tsai J-H. Beauvericin induces cytotoxic effects in human acute lymphoblastic leukemia cells through cytochrome c release, caspase 3 activation: the causative role of calcium. Cancer Lett. 2004;216:165–73. https://doi.org/10.1016/j.canlet.2004.06.005.
7. Zhan J, Burns AM, Liu MX, et al. Search for cell motility and angiogenesis inhibitors with potential anticancer activity: beauvericin and other constituents of two endophytic strains of *Fusarium oxysporum*. J Nat Prod. 2007;70:227–32. https://doi.org/10.1021/np060394t.
8. Dornetshuber-Fleiss R, Heilos D, Mohr T, et al. The naturally born fusariotoxin enniatin B and sorafenib exert synergistic activity against cervical cancer in vitro and *in vivo*. Biochem Pharmacol. 2015;93:318–31. https://doi.org/10.1016/j.bcp.2014.12.013.
9. Heilos D, Rodríguez-Carrasco Y, Englinger B, et al. The natural fungal metabolite beauvericin exerts anticancer activity in vivo: a pre-clinical pilot study. Toxins (Basel). 2017;9:258. https://doi.org/10.3390/toxins9090258.
10. Süssmuth RD, Mainz A. Nonribosomal peptide synthesis—principles and prospects. Angew Chem Int Ed Engl. 2017;56:3770–821. https://doi.org/10.1002/anie.201609079.
11. Yu D, Xu F, Zhang S, Zhan J. Decoding and reprogramming fungal iterative nonribosomal peptide synthetases. Nat Commun. 2017;8:15349. https://doi.org/10.1038/ncomms15349.
12. Steiniger C, Hoffmann S, Mainz A, et al. Harnessing fungal nonribosomal cyclodepsipeptide synthetases for mechanistic insights and tailored engineering. Chem Sci. 2017;8:7834–43. https://doi.org/10.1039/C7SC03093B.
13. Yu D, Xu F, Gage D, Zhan J. Functional dissection and module swapping of fungal cyclooligomer depsipeptide synthetases. Chem Commun (Camb). 2013;49:6176–8. https://doi.org/10.1039/c3cc42425a.
14. Zobel S, Boecker S, Kulke D, et al. Reprogramming the biosynthesis of cyclodepsipeptide synthetases to obtain new enniatins and beauvericins. ChemBioChem. 2016;17:283–7. https://doi.org/10.1002/cbic.201500649.
15. Hu DX, Bielitza M, Koos P, Ley SV. A total synthesis of the ammonium ionophore, (-)-enniatin B. Tetrahedron Lett. 2012;53:4077–9. https://doi.org/10.1016/j.tetlet.2012.05.110.
16. Lücke D, Dalton T, Ley SV, Wilson ZE. Synthesis of natural and unnatural cyclooligomeric depsipeptides enabled by flow chemistry. Chem (A Eur J). 2016;22:4206–17. https://doi.org/10.1002/chem.201504457.
17. Monma S, Sunazuka T, Nagai K, et al. Verticilide: elucidation of absolute configuration and total synthesis. Org Lett. 2006;8:5601–4. https://doi.org/10.1021/ol0623365.
18. Batiste SM, Johnston JN. Rapid synthesis of cyclic oligomeric depsipeptides with positional, stereochemical, and macrocycle size distribution control. Proc Natl Acad Sci USA. 2016;113:14893–7. https://doi.org/10.1073/pnas.1616462114.
19. Kanaoka M, Isogai A, Suzuki A. Synthesis of bassianolide. Tetrahedron Lett. 1977;18:4049–50. https://doi.org/10.1016/S0040-4039(01)83423-7.
20. Lee H-S, Song H-H, An J-H, et al. Statistical optimization of growth medium for the production of the entomopathogenic and phytotoxic cyclic depsipeptide beauvericin from *Fusarium oxysporum* KFCC 11363P. J Microbiol Biotechnol. 2008;18:138–44.
21. Xu L-J, Liu Y-S, Zhou L-G, Wu J-Y. Enhanced beauvericin production with in situ adsorption in mycelial liquid culture of Fusarium redolens Dzf2. Process Biochem. 2009;44:1063–7. https://doi.org/10.1016/j.procbio.2009.05.004.
22. Xu Y, Orozco R, Wijeratne EMK, et al. Biosynthesis of the cyclooligomer depsipeptide beauvericin, a virulence factor of the entomopathogenic fungus *Beauveria bassiana*. Chem Biol. 2008;15:898–907. https://doi.org/10.1016/j.chembiol.2008.07.011.
23. Zobel S, Kumpfmüller J, Süssmuth RD, Schweder T. *Bacillus subtilis* as heterologous host for the secretory production of the non-ribosomal cyclodepsipeptide enniatin. Appl Microbiol Biotechnol. 2015;99:681–91. https://doi.org/10.1007/s00253-014-6199-0.
24. Yu D, Xu F, Zi J, et al. Engineered production of fungal anticancer cyclooligomer depsipeptides in *Saccharomyces cerevisiae*. Metab Eng. 2013;18:60–8. https://doi.org/10.1016/j.ymben.2013.04.001.
25. Meyer V, Andersen MR, Brakhage AA, et al. Current challenges of research on filamentous fungi in relation to human welfare and a sustainable bio-economy: a white paper. Fungal Biol Biotechnol. 2016;3:6. https://doi.org/10.1186/s40694-016-0024-8.
26. Meyer V, Wanka F, van Gent J, et al. Fungal gene expression on demand: an inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. Appl Environ Microbiol. 2011;77:2975–83. https://doi.org/10.1128/AEM.02740-10.
27. Richter L, Wanka F, Boecker S, et al. Engineering of *Aspergillus niger* for the production of secondary metabolites. Fungal Biol Biotechnol. 2014;1:4. https://doi.org/10.1186/s40694-014-0004-9.
28. Lee C, Görisch H, Kleinkauf H, Zocher R. A highly specific D-hydroxyisovalerate dehydrogenase from the enniatin producer *Fusarium sambucinum*. J Biol Chem. 1992;267:11741–4.
29. Schuetze T, Meyer V. Polycistronic gene expression in *Aspergillus niger*. Microb Cell Fact. 2017;16:162. https://doi.org/10.1186/s12934-017-0780-z.
30. Mattern IE, van Noort JM, van den Berg P, et al. Isolation and characterization of mutants of *Aspergillus niger* deficient in extracellular proteases. Mol Gen Genet. 1992;234:332–6.
31. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.
32. Altschul SF, Wootton JC, Gertz EM, et al. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 2005;272:5101–9. https://doi.org/10.1111/j.1742-4658.2005.04945.x.
33. Calvo AM, Wilson RA, Bok JW, Keller NP. Relationship between secondary metabolism and fungal development. Microbiol Mol Biol Rev. 2002;66:447–59. https://doi.org/10.1128/MMBR.66.3.447.
34. Kanaoka M, Isogai A, Murakoshi S, et al. Bassianolide, a new insecticidal cyclodepsipeptide from *Beauveria bassiana* and *Verticillium lecanii*. Agric Biol Chem. 1978;42:629–35. https://doi.org/10.1271/bbb1961.42.629.

35. Monod J. The growth of bacterial cultures. Annu Rev Microbiol. 1949;3:371–94.

36. Kishore G, Sugumaran M, Vaidyanathan CS. Metabolism of DL-(±)-phenylalanine by *Aspergillus niger*. J Bacteriol. 1976;128:182–91.

37. Schultz JS. Cotton closure as an aeration barrier in shaken flask fermentations. Appl Microbiol. 1964;12:305–10.

38. Krappmann S, Braus GH. Nitrogen metabolism of *Aspergillus* and its role in pathogenicity. Med Mycol. 2005;43:31–40. https://doi.org/10.1080/13693780400024271.

39. Zhou Z, Takaya N, Nakamura A, et al. Ammonia fermentation, a novel anoxic metabolism of nitrate by fungi. J Biol Chem. 2002;277:1892–6. https://doi.org/10.1074/jbc.M109096200.

40. Sikora LA, Marzluf GA. Regulation of L-phenylalanine ammonia-lyase by L-phenylalanine and nitrogen in *Neurospora crassa*. J Bacteriol. 1982;150:1287–91.

41. Vödisch M, Scherlach K, Winkler R, et al. Analysis of the *Aspergillus fumigatus* proteome reveals metabolic changes and the activation of the pseurotin A biosynthesis gene cluster in response to hypoxia. J Proteome Res. 2011;10:2508–24. https://doi.org/10.1021/pr1012812.

42. Barker BM, Kroll K, Vödisch M, et al. Transcriptomic and proteomic analyses of the *Aspergillus fumigatus* hypoxia response using an oxygen-controlled fermenter. BMC Genom. 2012;13:62. https://doi.org/10.1186/1471-2164-13-62.

43. Kroll K, Pähtz V, Hillmann F, et al. Identification of hypoxia-inducible target genes of *Aspergillus fumigatus* by transcriptome analysis reveals cellular respiration as an important contributor to hypoxic survival. Eukaryot Cell. 2014;13:1241–53. https://doi.org/10.1128/EC.00084-14.

44. Feifel SC, Schmiederer T, Hornbogen T, et al. *In vitro* synthesis of new enniatins: probing the alpha-D-hydroxy carboxylic acid binding pocket of the multienzyme enniatin synthetase. ChemBioChem. 2007;8:1767–70. https://doi.org/10.1002/cbic.200700377.

45. Matthes D, Richter L, Müller J, et al. *In vitro* chemoenzymatic and in vivo biocatalytic syntheses of new beauvericin analogues. Chem Commun. 2012;48:5674–6. https://doi.org/10.1039/c2cc31669b.

46. Müller J, Feifel SC, Schmiederer T, et al. *In vitro* synthesis of new cyclodepsipeptides of the PF1022-type: probing the alpha-D-hydroxy acid tolerance of PF1022 synthetase. ChemBioChem. 2009;10:323–8. https://doi.org/10.1002/cbic.200800539.

47. Xu Y, Wijeratne EMK, Espinosa-Artiles P, et al. Combinatorial mutasynthesis of scrambled beauvericins, cyclooligomer depsipeptide cell migration inhibitors from *Beauveria bassiana*. ChemBioChem. 2009;10:345–54. https://doi.org/10.1002/cbic.200800570.

48. Boecker S, Zobel S, Meyer V, Süssmuth RD. Rational biosynthetic approaches for the production of new-to-nature compounds in fungi. Fungal Genet Biol. 2016;89:89–101. https://doi.org/10.1016/j.fgb.2016.02.003.

49. Huddleston JP, Johnson WH, Schroeder GK, Whitman CP. Reactions of Cg10062, a cis-3-chloroacrylic acid dehalogenase homologue, with acetylene and allene substrates: evidence for a hydration-dependent decarboxylation. Biochemistry. 2015;54:3009–23. https://doi.org/10.1021/acs.biochem.5b00240.

50. EFSA Panel on Contaminants in the Food Chain (CONTAM). Scientific opinion on the risks to human and animal health related to the presence of beauvericin and enniatins in food and feed. EFSA J. 2014;12:3802. https://doi.org/10.2903/j.efsa.2014.3802.

51. European Food Safety Authority. Occurrence data on beauvericin and enniatins in food. 2017. https://doi.org/10.5281/ZENODO.571179.

52. Jestoi M. Emerging *Fusarium*-mycotoxins fusaproliferin, beauvericin, enniatins, and moniliformin: a review. Crit Rev Food Sci Nutr. 2008;48:21–49. https://doi.org/10.1080/10408390601062021.

53. Santini A, Ferracane R, Meca G, Ritieni A. Overview of analytical methods for beauvericin and fusaproliferin in food matrices. Anal Bioanal Chem. 2009;395:1253–60. https://doi.org/10.1007/s00216-009-3117-x.

54. Sørensen JL, Nielsen KF, Rasmussen PH, Thrane U. Development of a LC-MS/MS method for the analysis of enniatins and beauvericin in whole fresh and ensiled maize. J Agric Food Chem. 2008;56:10439–43. https://doi.org/10.1021/jf802038b

55. Uhlig S, Ivanova L. Determination of beauvericin and four other enniatins in grain by liquid chromatography-mass spectrometry. J Chromatogr A. 2004;1050:173–8. https://doi.org/10.1016/j.chroma.2004.08.031.

56. Sewram V, Nieuwoudt TW, Marasas WF, et al. Determination of the *Fusarium* mycotoxins, fusaproliferin and beauvericin by high-performance liquid chromatography-electrospray ionization mass spectrometry. J Chromatogr A. 1999;858:175–85.

57. Annesley TM. Ion suppression in mass spectrometry. Clin Chem. 2003;49:1041–4. https://doi.org/10.1373/49.7.1041.

58. Panuwet P, Hunter RE, D'Souza PE, et al. Biological matrix effects in quantitative tandem mass spectrometry-based analytical methods: advancing biomonitoring. Crit Rev Anal Chem. 2016;46:93–105. https://doi.org/10.1080/10408347.2014.980775.

59. Berg T, Strand DH. ^{13}C labelled internal standards-A solution to minimize ion suppression effects in liquid chromatography-tandem mass spectrometry analyses of drugs in biological samples? J Chromatogr A. 2011;1218:9366–74. https://doi.org/10.1016/j.chroma.2011.10.081.

60. Varga E, Glauner T, Köppen R, et al. Stable isotope dilution assay for the accurate determination of mycotoxins in maize by UHPLC-MS/MS. Anal Bioanal Chem. 2012;402:2675–86. https://doi.org/10.1007/s00216-012-5757-5.

61. Rychlik M, Asam S. Stable isotope dilution assays in mycotoxin analysis. Anal Bioanal Chem. 2008;390:617–28. https://doi.org/10.1007/s00216-007-1717-x.

62. Hu L, Rychlik M. Biosynthesis of ^{15}N$_3$-labeled enniatins and beauvericin and their application to stable isotope dilution assays. J Agric Food Chem. 2012;60:7129–36. https://doi.org/10.1021/jf3015602.

63. Stokvis E, Rosing H, Beijnen JH. Stable isotopically labeled internal standards in quantitative bioanalysis using liquid chromatography/mass spectrometry: necessity or not? Rapid Commun Mass Spectrom. 2005;19:401–7. https://doi.org/10.1002/rcm.1790.

64. Boyce M, Monnot F. First in man clinical trial of emodepside (BAY 44-4400). In: Clin. NCT02661178. 2017. https://clinicaltrials.gov/show/NCT02661178. Accessed 25 Oct 2017.

65. Sambrook J, Russel DW. Molecular cloning—a laboratory manual. 3rd ed. New York: Cold Spring Harbor Laboratory Press; 2001.

66. Meyer V, Ram AF, Punt PJ. Genetics, genetic manipulation, and approaches to strain improvement of filamentous fungi. In: Baltz RH, Davies JE, Demain AL, editors. Man. Ind. Microbiol. Biotechnol. 3rd ed. American Society for Microbiology (ASM); 2010. pp 318–30.

67. Vishniac W, Santer M. The thiobacilli. Microbiol Mol Biol Rev. 1957;21:195–213.

68. Iversen JJL, Thomsen JK, Cox RP. On-line growth measurements in bioreactors by titrating metabolic proton exchange. Appl Microbiol Biotechnol. 1994;42:256–62. https://doi.org/10.1007/BF00902726.

69. Baldwin JE, Adlington RM, Crouch NP, Pereira IAC. The enzymatic synthesis of isotopically labelled penicillin Ns with isopenicillin N synthase. J Label Compd Radiopharm. 1998;41:1145–63. https://doi.org/10.1002/(SICI)1099-1344(199812)41:12<1145::AID-JLCR159>3.0.CO;2-2.

70. Madry N, Zocher R, Kleinkauf H. Enniatin production by *Fusarium oxysporum* in chemically defined media. Appl Microbiol Biotechnol. 1983;17:75–9.

CoIN: co-inducible nitrate expression system for secondary metabolites in *Aspergillus nidulans*

Philipp Wiemann[1,3†] , Alexandra A. Soukup[1,4†], Jacob S. Folz[1,5], Pin-Mei Wang[1,6], Andreas Noack[1] and Nancy P. Keller[1,2*]

Abstract

Background: Sequencing of fungal species has demonstrated the existence of thousands of putative secondary metabolite gene clusters, the majority of them harboring a unique set of genes thought to participate in production of distinct small molecules. Despite the ready identification of key enzymes and potential cluster genes by bioinformatics techniques in sequenced genomes, the expression and identification of fungal secondary metabolites in the native host is often hampered as the genes might not be expressed under laboratory conditions and the species might not be amenable to genetic manipulation. To overcome these restrictions, we developed an inducible expression system in the genetic model *Aspergillus nidulans*.

Results: We genetically engineered a strain of *A. nidulans* devoid of producing eight of the most abundant endogenous secondary metabolites to express the sterigmatocystin $Zn(II)_2Cys_6$ transcription factor-encoding gene *aflR* and its cofactor *aflS* under control of the nitrate inducible *niiA/niaD* promoter. Furthermore, we identified a subset of promoters from the sterigmatocystin gene cluster that are under nitrate-inducible AflR/S control in our production strain in order to yield coordinated expression without the risks from reusing a single inducible promoter. As proof of concept, we used this system to produce β-carotene from the carotenoid gene cluster of *Fusarium fujikuroi*.

Conclusion: Utilizing one-step yeast recombinational cloning, we developed an inducible expression system in the genetic model *A. nidulans* and show that it can be successfully used to produce commercially valuable metabolites.

Keywords: Yeast recombinational cloning, Secondary metabolism, Genetic engineering, Carotenes, *Aspergillus nidulans*, *Fusarium fujikuroi*, AflR, Sterigmatocystin, Nitrate, Biotechnology, Synthetic biology

Background

Natural products or secondary metabolites (SMs) have been invaluable as platforms for developing front-line drugs. Between 1981 and 2010, 5% of the 1031 new chemical entities approved as drugs by the Food and Drug Administration (FDA) were natural products or derivatives, including 48.6% of cancer medications [36]. In addition, SMs are major sources of innovative therapeutic agents for both bacterial and fungal infectious diseases, lipid disorders, and immunomodulation [16]. Fungal SMs have proven to be a particularly important source of new leads with useful pharmaceutical activities. A literature survey of fungal metabolites, covering 1500 fungal SMs that were isolated and characterized between 1993 and 2001, showed that more than half of the molecules had antibacterial, antifungal or antitumor activity [41]. However, the full metabolic potential of the majority of existing fungal species has not been investigated. Major roadblocks in this endeavor are that some species are not cultivable under laboratory conditions and/or their SM gene clusters are silent. Previous strategies on activating fungal SMs have focused mainly on (1)

*Correspondence: npkeller@wisc.edu
†Philipp Wiemann and Alexandra A. Soukup contributed equally to this work
[1] Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Full list of author information is available at the end of the article

activating endogenous gene clusters by over-expressing the pathway-specific transcription factor [13, 43, 54]), (2) manipulating global regulators [11, 28], and (3) expressing the entire gene cluster in a heterologous host [8]. Although successful in some cases, these strategies have significant disadvantages. As not all fungal species are easily amenable to genetic manipulation, strategies that focus on endogenous activation are impossible in these species. This prevents the option of over-expressing a cluster-specific transcription factor, which has been the most successful approach to activating cryptic clusters thus far (reviewed in [51]). In addition, not all SM clusters contain transcription factors and although some clusters have been activated by overexpressing every gene in the cluster [14, 53], this adds labor and time to the process and may be hard to achieve with clusters containing many genes.

Previous approaches expressing fungal gene clusters in heterologous hosts (mainly *Saccharomyces cerevisiae* or *Aspergillus* spp.) focused on amplification of the entire gene cluster including native promoters. Although these approaches lead to expression of the targeted gene clusters in some cases [55], the use of native promoters cannot guarantee controlled activation of the genes. Therefore, the identification and use of defined promoters presents an alternative means to activate clusters. Cloning of entire gene clusters can be achieved by PCR-based amplification of the desired DNA region and subsequent yeast recombination-based cloning [55]. Promoter exchanges using this technique rely on the identification of different promoter regions as the use of identical promoter sequences is impossible due to the homologous recombination among promoters [14].
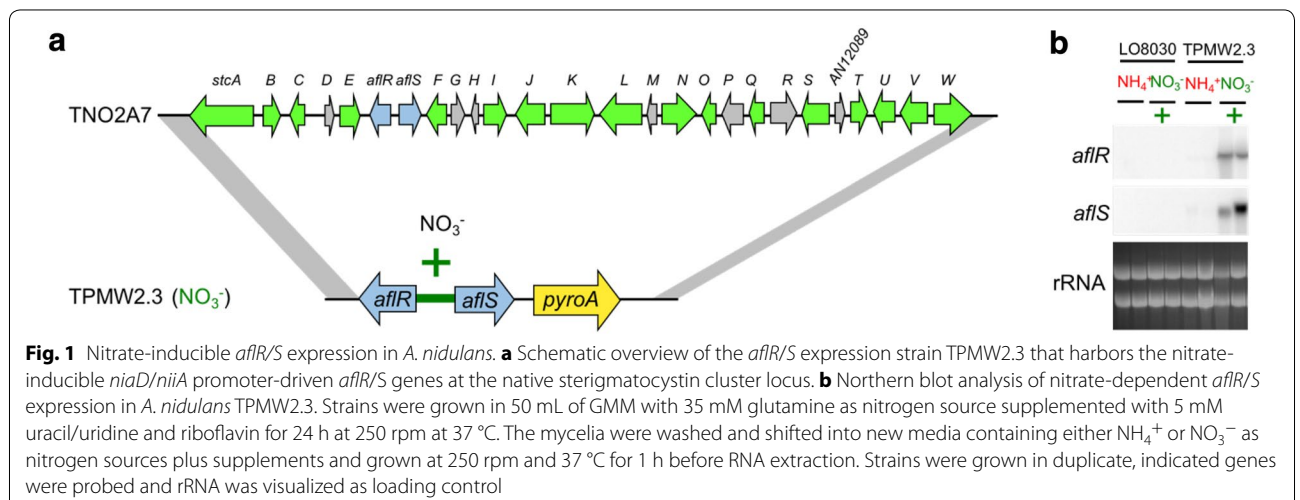
Our goal was to identify a series of distinct promoters that could be activated in one step and, furthermore,

activated under an inducible system as many SMs exhibit antifungal properties that could be toxic to the heterologous host [12, 46]. Thus, we designed a strain of genetic model organism *Aspergillus nidulans* that contains an inducible genetic construct which allows for expression of the positive acting transcriptional elements of the sterigmatocystin (ST) gene cluster, *aflR* and *aflS* (formerly *aflJ*, [20]). The ST gene cluster contains 25 distinct genes and it is known that the transcription factor AflR and its cofactor AflS are responsible for ST production [9]. We constructed a strain of *A. nidulans* with its endogenous ST cluster removed but with *aflR*/*aflS* placed back into the strain under the control of a nitrate inducible divergent promoter (*niiA*(p)/*niaD*(p)) [44, 45], thereby allowing controlled *aflR*/*aflS* expression based on culture conditions. We tested the expression of all 25 ST promoters by AflR/AflS in this strain and identified eight ST promoters specifically regulated by nitrate induction of *aflR/S*. To test the system for expression of a fungal secondary metabolite, we cloned the carotenoid gene cluster from *Fusarium fujikuroi* and placed it under control of these inducible ST promoters. We show that the derived *A. nidulans* transformants produce β-carotene in competitive levels to existing systems using our technology.

Methods
Fungal strains and culture conditions
Aspergillus nidulans strains used in this study are listed in Additional file 1: Table S1. *Fusarium fujikuroi* IMI58289 [51] was used for *carRA*, *carB*, and *ggs1* amplification as a reference for carotenoid production. Strains were maintained as glycerol stocks and activated on solid glucose minimal medium (GMM) at 37 °C with appropriate supplements [48]. For experiments in Fig. 1, nitrate was replaced with equimolar ammonium tartrate and

Fig. 1 Nitrate-inducible *aflR/S* expression in *A. nidulans*. **a** Schematic overview of the *aflR/S* expression strain TPMW2.3 that harbors the nitrate-inducible *niaD/niiA* promoter-driven *aflR*/S genes at the native sterigmatocystin cluster locus. **b** Northern blot analysis of nitrate-dependent *aflR/S* expression in *A. nidulans* TPMW2.3. Strains were grown in 50 mL of GMM with 35 mM glutamine as nitrogen source supplemented with 5 mM uracil/uridine and riboflavin for 24 h at 250 rpm at 37 °C. The mycelia were washed and shifted into new media containing either NH_4^+ or NO_3^- as nitrogen sources plus supplements and grown at 250 rpm and 37 °C for 1 h before RNA extraction. Strains were grown in duplicate, indicated genes were probed and rRNA was visualized as loading control

supplemented with 5 mM uracil and uridine, respectively. For solidified media, Noble Agar (Difco™, BD, USA) was added at 16 g/L. For *pyrG* auxotrophs, the growth medium was supplemented with 5 mM uridine and uracil. For *riboB* auxotrophs, the growth medium was supplemented with 5 mM riboflavin. For *pyroA* auxotrophs, the growth medium was supplemented with 5 mM pyridoxine. Conidia were harvested in 0.01% Tween 80 and enumerated using a hemocytometer. For RNA analysis, indicated strains were inoculated into 50 mL of liquid GMM with 35 mM glutamine as nitrogen source at 5×10^6 conidia/mL in duplicate and grown at 37 °C and 250 rpm for 24 h in ambient light conditions. The cultures were shifted into new GMM media either containing 70 mM nitrate or 35 mM glutamine as nitrogen source for 1 h. The mycelium was harvested and lyophilized before RNA extraction. For carotenoid production 5×10^6 conidia of indicated strains were inoculated on 20 mL liquid stationary GMM with either containing 70 mM nitrate or 35 mM glutamine as nitrogen source for 3 days at 37 °C (*A. nidulans*) or 29 °C (*F. fujikuroi*) in the dark. To distinguish between carotenoid production in the mycelia and spores, the strains were grown on liquid stationary GMM media (described above) for 3 days at 37 °C in the light to induce sporulation. Mycelia and spores were resuspended in 0.01% (v/v) Tween80, vortexed to separate spores from mycelia. Spores were separated from mycelia by filtration.

Yeast recombinational cloning

Yeast strain BJ5464 (*MATalpha, ura3-52, trp1, leu2-Δ1, his3-Δ200, pep4::HIS3, prb1-Δ1.6R, can1, GAL*) was inoculated into 25–50 mL of appropriate media (2× YPDA) and incubated at 30 °C at 200 rpm overnight. The concentration of overnight culture was determined using OD 600 with a 1×10^7 cells/mL set to an OD reading of 1.0. Then, 1.25×10^9 cells were centrifuged at $3000 \times g$ for 5 min. Fresh media was added to the pelleted cells and added to a baffled flask containing 250 mL of 2× YPAD to a final concentration of 5×10^6 cells/mL and incubated at 30 °C at 200 rpm until the cell titer reached 2×10^7 cells/mL. Cells were harvested by centrifugation at $3000 \times g$ for 5 min. Supernatant was removed and the cells were washed with double distilled H$_2$O (ddH$_2$O). The cells were transferred to one 50 mL falcon tube and washed an additional time with ddH$_2$O before they were centrifuged at $3000 \times g$ for 5 min. Cells were resuspended in 5% glycerol and 10% DMSO to a final concentration of 2×10^9 cells/mL aliquoted in 100 µL. These cells can be frozen at −80 °C for several weeks. Before transformation, cells were pelleted and the supernatant removed. For transformation 250 ng of the digested backbone vector and 500 ng of each DNA PCR product (see below)

were added and adjusted to a final volume of 14 µL with ddH$_2$O. The DNA mixture was added to the yeast along with 260 µL of a 50% (w/v) polyethyleneglycol 3600, 36 µL 1 M lithium acetate and 50 µL of denatured sheared salmon sperm DNA (2 mg/mL). The mixture was vortexed and incubated at 42 °C for 45 min. Cells were centrifuged at $13,000 \times g$ for 30 s and the supernatant was removed. The cells were carefully resuspended in 1 mL ddH$_2$O and 200–500 µL were spread on solidified synthetic drop out media containing all necessary supplements without uracil for selection. Plates were incubated at 30 °C for 3–5 days.

Plasmid isolation from yeast

All colonies from a transformation plate were scraped and incubated overnight in liquid synthetic yeast drop out solution containing the appropriate supplements without uracil at 200 rpm at 30 °C. One mL was pelleted and the supernatant removed. 200 µL of STC buffer (50 mM Tris–HCl pH 7.5, 1.2 M sorbitol, 50 mM CaCl$_2$) including 3 µL Zymolase was added and incubated at 37 °C for 1 h. To the mixture, 200 µL of 1% (w/v) sodium dodecyl sulfate (SDS) in 200 mM NaOH were added and inverted. The solution was neutralized by adding 240 µL of 3 M potassium acetate, pH 5.5 and inverted. The mixture was centrifuged and the supernatant mixed with 600 µL isopropanol, inverted and centrifuged at maximum speed for 10 min. The supernatant was removed and the pellet was washed with 70% (v/v) ethanol. The pellet was air dried and resuspended in 30 µL ddH$_2$O.

Transformation of *Escherichia coli* and plasmid conformation

Following standard techniques [23], 10 µL of the yeast plasmid extract were transformed into *E. coli* and positive colonies were selected on media containing ampicillin. Plasmids from colonies were isolated using standard procedures [23]. Plasmids were restriction digested with appropriate enzymes to confirm correct insertion. Additional confirmation was achieved using PCR amplification of fused DNA products. To ensure correct DNA sequences for expression plasmids, Sanger sequencing was performed. The correct plasmids were then grown in a 50 mL culture and plasmids were isolated using the Quantum Prep® Plasmid Midiprep Kit (Biorad) according to the manufacturers' instructions. Before fungal transformation, the plasmids were linearized using *Asc*I.

Plasmid construction and fungal transformation

Expression fragments were created by yeast recombinational cloning as described above. All primers used are listed in Additional file 1: Table S2 and all plasmids are listed in Additional file 1: Table S3. For assembling

the nitrate inducible *aflR/S* construct, six fragments total were amplified and eventually cloned into the *Asc*I digested plasmid backbone of pYHC-yA-riboB [57]. The 3′ flanks of *stcA* and *stcW* were amplified from *A. nidulans* LO8030 DNA using primer pairs *stcA*3′-F/-R and *stcW*3′-F/-R with the -R primers containing 5′ overlaps to the respective site of *Asc*I digested plasmid backbone and the -F primers having overlaps to the *aflS* terminator and *pyroA* cassette, respectively. The bidirectional *niaD/niiA* promoter region was amplified from *A. nidulans* LO8030 [15, 38] with overlaps to the open reading frames of *aflR* and *aflS* using primer pairs nitrate-F/-R. The open reading frame of *aflR* including 500 bp of terminator was amplified from *A. nidulans* FGSC 4A DNA using primer pair *aflR*-F/-R where the -R primer had a 5′ overhang to the terminator region of the *A. fumigatus pyroA* gene. The open reading frame of *aflS* was amplified from *A. nidulans* FGSC 4A DNA using primer pairs *aflS*-F/-R with the -R primer having a 5′ overlap to the -F primer used to amplify the *stcW* flank. The *pyroA* cassette was retrieved through *Pst*I restriction digest of pJMP61 [7]. After yeast recombinational cloning, the plasmid pPMW1 was created. For sterigmatocystin promoter studies the entire bidirectional promoter region between two open reading frames was cloned or, in the case of monodirectional promoters 500 bp upstream of the open reading frame was amplified using primer pairs *stc* "gene name"-pF/*stc* "gene name"-pR including 5′ overlaps to the *wA* 5′ flank and the open reading frame of *pyrG* gene from *A. fumigatus* CEA10. Plasmid pYHC-wA-pyrG [55] was linearized using *Nhe*I. After yeast recombineering, plasmid pAN "stcGene" were yielded. For constructing the carotenoid expression plasmid pJSF1, the bidirectional promoter region between *stcA* and *stcB* was amplified using primer pair *stcAB*-cF/-cR with overlaps to the carotenoid cluster genes *carRA* and *carB* from *F. fujikuroi* IMI58289. The open reading frame of *carRA* including 500 bp terminator region was amplified from *F. fujikuroi* IMI58289 DNA using primer pairs *carRA*-F/-R with the -R primer including an overlap to the *wA* 3′ flank of plasmid pYHC-wA-pyrG. The open reading frame of *carB* including 500 bp terminator region was amplified from *F. fujikuroi* IMI58289 DNA using primer pairs *carB*-cF/-cR with the -cR primer including an overlap to the *wA* 5′ flank of plasmid pYHC-wA-pyrG. All fragments were assembled using yeast recombinational cloning into *Eco*RI/*Xho*I linearized pYHC-wA-pyrG resulting in pJSF1. pJSF2 was assembled in a similar process using EcoRI/XhoI linearized pYHC-yA-riboB and PCR amplicons of the *stcM* promoter (amplified with *stcM-cF/-cR*) and *ggs1* (amplified with *ggs1-cF/-cR*).

Transformation of *A. nidulans* was performed as previously described [40]. For selection of nitrate inducible

aflR/S strains, *A. nidulans* LO8030 was used as the recipient strain. pPMW1 was linearized using *Asc*I, transformed into LO8030, and transformants were selected on media where pyridoxine was omitted and uracil/uridine and riboflavin were supplemented yielding strain TPMW2.3. For selection of *stc* promoter test strains, TPMW2.3 was used as the recipient strain. pAN "stc-Gene" plasmids were linearized using *Sbf*I and transformants were selected on media were riboflavin was omitted and uracil/uridine was supplemented yielding strains TANx and TAASx (see Additional file 1: Table S1). To create a riboflavin prototrophic strain, TPMW2.3 was transformed with *Sbf*I linearized pYHC-yA-riboB and selected on media where riboflavin was omitted and uracil/uridine was supplemented yielding strain TPMW7.2. For selection of *car* expression strains, TPMW7.2 was used as the recipient strain and *Sbf*I linearized pTJSF1 was transformed into TPMW7.2 and transformants selected on media omitting uracil/uridine yielding strain TJSF1.1. An auxotrophic control strain was generated by using TPMW7.2 as recipient strain and *Sbf*I linearized pYHC-wA-pyrG was transformed and selected on media without supplements yielding strain TPMW8.2. For DNA isolation, all fungal strains were grown for 24 h at 37 °C (*Aspergillus*) or 29 °C (*Fusarium*) in steady state liquid GMM, supplemented appropriately as described by Shimizu and Keller [48]. Single integration was confirmed by Southern analysis as described by [23] using P³²-labelled probes created by amplification of the indicated DNA fragment in Additional file 2: Figures S1–S4.
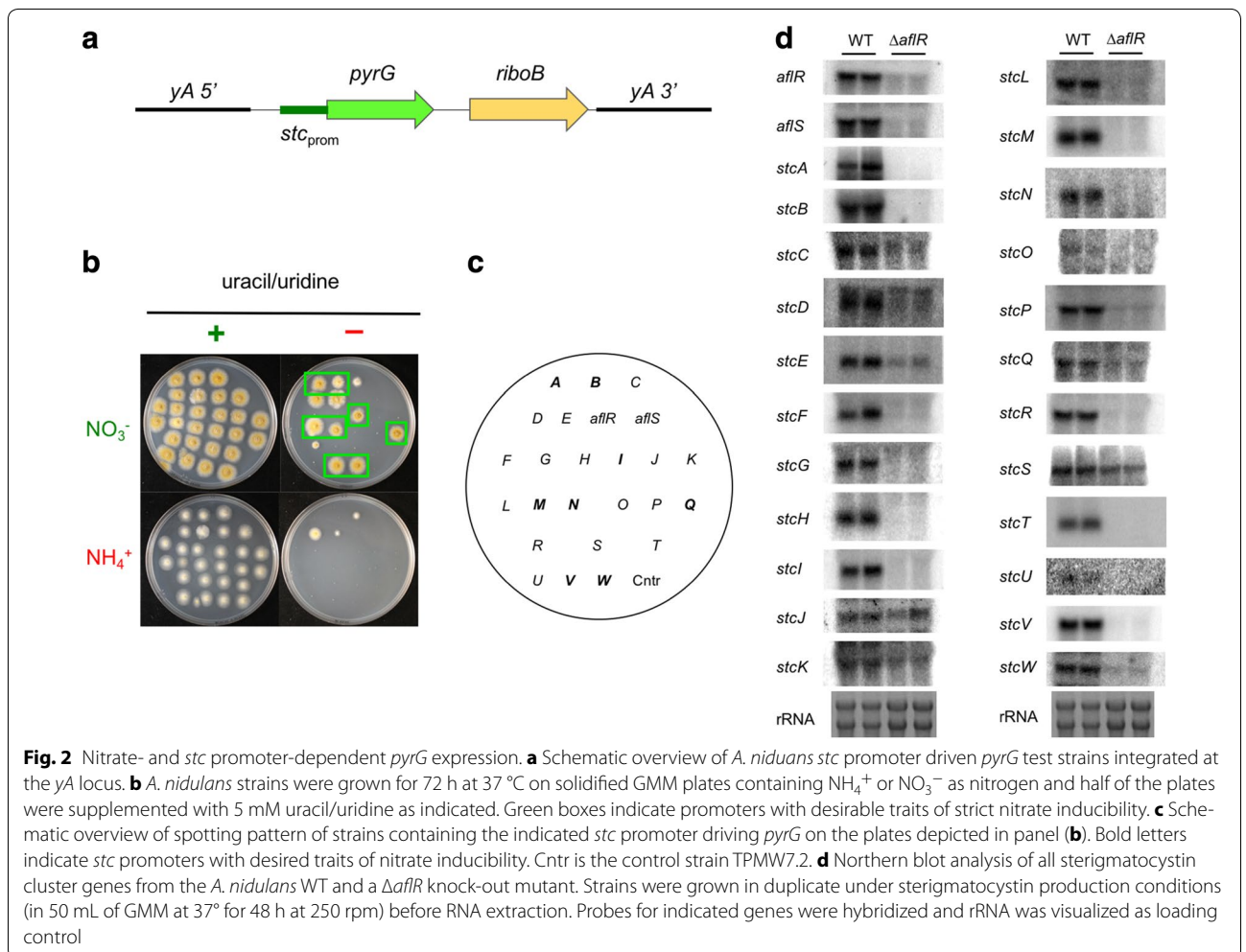
Carotenoid analysis

Carotenoids were extracted and analyzed as previously described [18, 19]. Briefly, carotenoids were extracted with acetone from freeze dried mycelia and were separated by thin layer chromatography developed in light petroleum/diethyl ether/acetone (4:1:1; v/v/v). The bands were scraped out and dissolved in acetone. High-performance liquid chromatography (HPLC) was used to analyze the β-carotene content by comparison to an authentic standard. HPLC separation was performed on using a ZORBAX Eclipse XDB-C18 column (Agilent, 4.6 mm by 150 mm with a 5 μm particle size) by using a binary gradient of methanol/t-butylmethyl ether (1:1) (v/v) as solvent A and methanol/t-butylmethyl ether/water (5:1:1) (v/v/v) as solvent B using a Flexar Binary Liquid Chromatography (LC) Pump (PerkinElmer) coupled to a Flexar LC Autosampler (Perkin Elmer) and a Flexar PDA Plus Detector (PerkinElmer). The binary gradient started with a linear step from 0 A to 57% A in 45 min and an additional linear gradient to 100% A in 0.5 min and hold for 25 min at a flow rate of 2 mL/min. Identification and relative quantification of secondary

metabolites was performed using Chromera Manager (PerkinElmer) by comparison to an authentic standard (Sigma Aldrich).

Results

The sterigmatocystin (ST) gene cluster of *A. nidulans* is known to harbor 25 genes involved in biosynthesis of sterigmatocystin [9]. While environmental regulation of the ST gene cluster is complex and not well understood, it was the first cluster that identified a gene product encoded within the cluster itself to function as a cluster-specific $Zn(II)_2Cys_6$ transcription factor, called AflR [21]. The gene encoding AflR shares a bidirectional promoter with *aflS* encoding a transcriptional cofactor of AflR [20]. We replaced the native promoter of *aflR/S* with the well characterized *niaD/niiA* promoter which is induced by the presence of nitrate in the absence of other nitrogen sources [10]. We confirmed nitrate-dependent expression of *aflR/S* by northern blot analysis (Fig. 1b).

Next, we set out to test which of the 25 *stc* gene promoters would be nitrate-inducible in our production strain (TPMW2.3). Since TPMW2.3 is a uracil/uridine and riboflavin auxotroph, we designed plasmids that contain each of the 25 *stc* gene promoters, respectively, driving expression of the *A. fumigatus pyrG* gene along with a *riboB* selectable marker flanked by bordering regions of the *yA* locus (Fig. 2a; Additional file 1: Table S3; Additional file 2: Fig. S2). Using a minimalized promoter selection strategy, we chose promoter regions as follows: For unidirectional *stc* genes, the promoter region was amplified from the first base after the stop codon of the first gene to the start codon of the second gene, but not exceeding 1 kb. In cases of bidirectional promoters, the entire region between the two start codons was chosen, not exceeding 1 kb. We selected 25 strains for each *stc* promoter for riboflavin prototrophy, exhibiting yellow spore color, and a control strain that did not include a *stc* promoter. To test for the ability of AflR/S to induce Af*pyrG* expression driven by each of the *stc*

Fig. 2 Nitrate- and *stc* promoter-dependent *pyrG* expression. **a** Schematic overview of *A. niduans stc* promoter driven *pyrG* test strains integrated at the *yA* locus. **b** *A. nidulans* strains were grown for 72 h at 37 °C on solidified GMM plates containing NH_4^+ or NO_3^- as nitrogen and half of the plates were supplemented with 5 mM uracil/uridine as indicated. Green boxes indicate promoters with desirable traits of strict nitrate inducibility. **c** Schematic overview of spotting pattern of strains containing the indicated *stc* promoter driving *pyrG* on the plates depicted in panel (**b**). Bold letters indicate *stc* promoters with desired traits of nitrate inducibility. Cntr is the control strain TPMW7.2. **d** Northern blot analysis of all sterigmatocystin cluster genes from the *A. nidulans* WT and a Δ*aflR* knock-out mutant. Strains were grown in duplicate under sterigmatocystin production conditions (in 50 mL of GMM at 37° for 48 h at 250 rpm) before RNA extraction. Probes for indicated genes were hybridized and rRNA was visualized as loading control

promoters, we grew them on media containing either nitrate (induces *aflR/S* expression; Fig. 1b) or ammonium, and supplemented with or without uracil/uridine (Fig. 2b, c). The growth assay showed that eight of the tested promoters (*stcA*, *stcB*, *stcI*, *stcM*, *stcN*, *stcQ*, *stcV*, and *stcW*) exhibited the desired ability to grow on plates without uracil/uridine supplementation on nitrate containing media only (Fig. 2b, c), demonstrating specific expression under induction conditions. Three of the promoters tested exhibited leaky expression (*stcC*, *stcD*, and *stcE*) as we observed colony growth of strains on media containing ammonium (Fig. 2b, c) where *aflR/S* should not be induced (Fig. 1b). In order to confirm control of the identified promoters by AflR, we investigated expression of all *stc* cluster genes in an *A. nidulans* WT and an isogenic Δ*aflR* knock-out strain [56] under sterigmatocystin production conditions. We found that in addition to the eight promoters identified in our plate assay, most of the remaining cluster genes were also expressed in an AflR-dependent manner (Fig. 2d). We speculate that either the length of the chosen promoters, the difference in culture conditions, or the insufficient expression level could be responsible for the observed discrepancies of *stc* activation of the *AfpyrG* reporter gene.

To test the functionality of our expression system genes responsible for carotenoid production from *Fusarium fujikuroi* [2–4] were expressed in our *A. nidulans* nitrate-inducible *aflR/S* strain TPMW2.3. In *F. fujikuroi*, the geranylgeranyl diphosphate (GGDP) synthase gene *ggs1* is responsible for production of GGDP [34], which is a substrate for CarRA and CarB, encoded by two of the clustered carotenoid biosynthetic genes needed for β-carotene production [31]. We inserted the *ggs1* gene driven by the *stcM* promoter and 0.5 kb of the native terminator, a riboflavin selectable marker flanked by the *yA* border regions (Additional file 1: Table S3; Additional file 2: Fig. S3), and the two carotenoid cluster genes *carRA* and *carB* including 0.5 kb of the native terminator regions, responsible for β-carotene production [31] under control of the bidirectional *stcA/B* promoter flanked by the *wA* border regions (Additional file 1: Table S3; Additional file 2: Fig. S3). Both plasmids were linearized and transformed into TPMW2.3 consecutively, yielding strain TJSF3.1 (Fig. 3a; Additional file 1: Table S1; Additional file 2: Fig. S4). Nitrate-inducible expression of *ggs1*, *carRA*, and *carB* was confirmed by northern blot analysis compared to a prototroph control strain that produced white spores (TPMW8.2) (Fig. 3b). When the two strains were grown on nitrate containing media, TJSF3.1 exhibited a characteristic orange color that was absent in the control (Fig. 3c). Characterization of carotene production by HPLC showed that the strain TJSF3.1 produced 125 μg β-carotene per gram mycelial dry

weight in our experimental setting (Fig. 3d, e). The production of β-carotene was significantly higher on nitrate induction media than on non-induction media containing glutamine (Fig. 3d, e). The control strain TPMW8.2 did not show any carotene production (Fig. 3d).

As carotenoid production in *N. crassa* and *Fusarium* spp. occurs in both mycelia and spores [5], we asked whether a similar distribution would occur in our production strain. Therefore, mycelia and spores were assessed individually for β-carotene content. Carotenoids were only produced in the mycelia and not in the spores (Fig. 4a). Since the GGDP produce by Ggs1 is also utilized for ergosterol production in *F. fujikuroi* [34] we set out to investigate if the homolog of *ggs1* in *A. nidulans* (AN0654, 54% identity, *e*-value: 4.0^{-118}) would be sufficient for β-carotene production. A strain was constructed that only expressed *carRA* and *carB* called TJSF1.1 (Additional file 2: Fig. S4). When carotenoid production between TJSF1.1 (*carRA* and *carB*) and TJSF3.1 (*ggs1*, *carRA*, and *carB*) was compared no significant difference under inducing conditions was observed (Fig. 4b), suggesting that AN0654 is sufficient to provide the maximum amount of GGDP that can be funneled into carotenoid production. As it is known that one of the bottlenecks during carotenogenesis is the production of mevalonate (a GGDP precursor) by the 3-hydroxyl-3-methyl-glutaryl-conenzyme A reductase (HMG CoA reductase) [1], we grew the two production strains on nitrate media supplemented with mevalonate before carotenoid quantification. However, we did not find any difference in production levels between the strains grown with or without mevalonate (Fig. 4c).

Discussion

Many efforts have been made to increase expression of fungal natural products [6, 35]. Apart from increasing production in the native host, a major focus has been on developing heterologous expression systems. Heterologous systems have the advantage that they can be carried out in a safe host system without toxic byproducts, that is easily amenable to genetic manipulation and preferably inducible [50]. Traditional approaches are laborious as they are mainly based on over-expressing each natural product cluster individually or sequentially, thereby relying on multiple selection markers that limit the number of genes expressed and subsequently reduce the chemical complexity of the natural product produced [55]. There have been successful reports on marker recycling to overcome this issue [14, 38], but these approaches involve multiple time-consuming transformation steps. Additionally, construction of the expression plasmids or cassettes has been achieved by labor intensive restriction enzyme- or fusion PCR-mediated methods [22, 25].
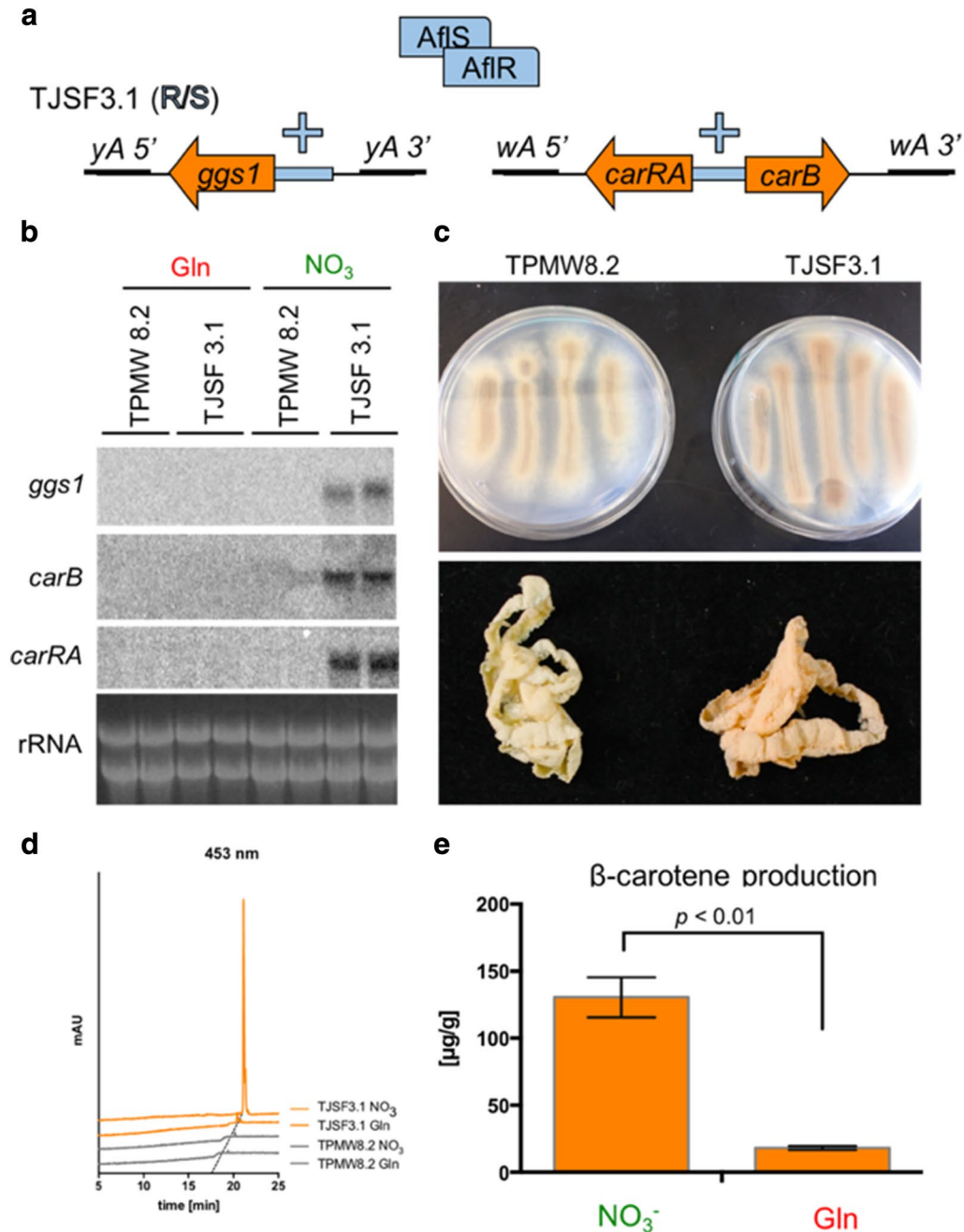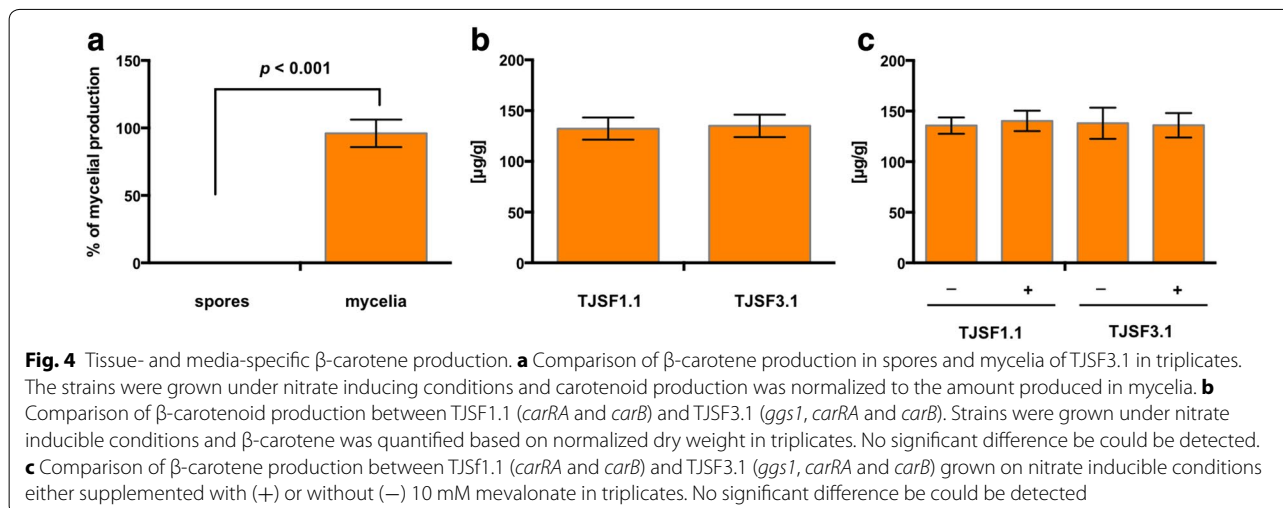
Fig. 3 Nitrate-dependent β-carotene production in *A. nidulans*. **a** Schematic overview of *A. nidulans* carotenoid production strain TJSF3.1 that harbors the *stcM*-driven *F. fujikuroi ggs1* gene at the *yA* locus and the *stcA/B*-driven *F. fujikuroi carRA/B* genes at the *wA* locus. **b** Northern blot expression analysis of indicated genes in the indicated strains. The strains were grown in duplicate for 24 h in 50 mL of GMM with 35 mM glutamine (Gln) as nitrogen source for 24 h at 250 rpm at 37 °C. The mycelia were washed and shifted into new media containing either Gln or NO_3^- as nitrogen sources and grown at 250 rpm and 37 °C for 1 h before RNA extraction. RNA was visualized as loading control. **c** Growth on solid medium demonstrating carotene expression from strain TJSF3.1 and control strain TPMW8.2 grown for 72 h at 37 °C on solidified GMM media containing nitrate as nitrogen source. Bottom picture shows mycelia of the same strains collected from liquid stationary GMM media containing nitrate as nitrogen source grown for 72 h at 37 °C. **d** HPLC chromatograms at 453 nm of β-carotene extracted from indicate strains grown as described in panel (**c**). **e** Quantification of β-carotene produced by TJSF3.1 grown on either NO_3^- or glutamine (Gln)

Fig. 4 Tissue- and media-specific β-carotene production. **a** Comparison of β-carotene production in spores and mycelia of TJSF3.1 in triplicates. The strains were grown under nitrate inducing conditions and carotenoid production was normalized to the amount produced in mycelia. **b** Comparison of β-carotenoid production between TJSF1.1 (*carRA* and *carB*) and TJSF3.1 (*ggs1*, *carRA* and *carB*). Strains were grown under nitrate inducible conditions and β-carotene was quantified based on normalized dry weight in triplicates. No significant difference be could be detected. **c** Comparison of β-carotene production between TJSf1.1 (*carRA* and *carB*) and TJSF3.1 (*ggs1*, *carRA* and *carB*) grown on nitrate inducible conditions either supplemented with (+) or without (−) 10 mM mevalonate in triplicates. No significant difference be could be detected

The ease of yeast recombinational cloning [39] has been exploited for a wide range of molecular methods, including gene knock-out libraries [17] and expression systems [47] in filamentous fungi as well as yeast itself [8]. Technically, yeast recombinational cloning allows for the assembly of multiple PCR fragments up to a vector size of several ten thousand kilo bases [39]. One of the major hurdles to overcome during yeast recombineering, is undesired recombination among multiple identical DNA regions. Recently, a study in *A. terreus* demonstrated the requirement of an AflR-like transcription factor, TerR, for expression of all twelve terrein cluster genes [24] similar to our expression data for AflR-dependency of all 25 *stc* genes. The system was subsequently used to control one of the terrein promoters in a heterologous expression system in *A. niger* to demonstrate activation of *orsA* from *A. nidulans* [24]. Here, we have demonstrated the specific induction of eight of the 25 *stc* promoters to control the expression of a reporter gene (*pyrG*). Subsequently, we have utilized three of the eight promoters to successfully express three *Fusarium* spp. derived genes responsible for carotenoid production in *A. nidulans* and confirm functionality of their gene products by detection of β-carotene.

As a precursor of vitamin A, β-carotene has long been in the focus of biotechnology. The most prominent example of heterologous gene expression leading to the production of β-carotene is the development of Golden Rice by Monsanto [37]. However, β-carotenoid production was also achieved in baker's yeast and bacteria. The amounts of carotenoids produced by the production strain constructed in this study are equivalent to the first production strain engineered in yeast [52]. Recent advances in manipulating metabolic pathways and genetic elements have increased the production in

baker's yeast [8, 30], and similar approaches could be undertaken to increase production in *A. nidulans*.

Notably, production of β-carotene is restricted to the mycelia in the *A. nidulans* production strains investigated in this study, whereas in other fungal species like *Neurospora crassa* and *F. fujikuroi*, carotenes are predominantly produced in asexual spores [26, 31]. One explanation might be the developmentally controlled expression pattern of the *niaD* promoter, as it was shown to be only transiently expressed at early stages of conidiation, but not at later time points [33]. Sterigmatocystin and aflatoxin in *Aspergillus* species are predominantly found in the mycelial fraction and a complex fusion network of vesicles containing different precursors and biosynthetic enzymes that ensure correct cellular localization of these secondary metabolites is being unveiled by several studies [32, 42, 49]. Additionally, pioneering work in *A. fumigatus* has demonstrated that certain natural products are predominantly produced in the asexual spores (conidia) [27, 29]. These findings suggest that conidial directed cellular pathways in the native host (*Fusarium*) may differ significantly from *Aspergillus* as location of β-carotene is not the same in the native and heterologous host. Determining which factors control the direction of fungal natural products to certain developmental structures in different fungal species will be a fascinating future task.

Conclusions

This study presents a new heterologous expression system for fungal natural products in the genetic model organism *A. nidulans*. The system described here makes use of the ability to co-express, minimally, eight promoters by a fungal-specific $Zn(II)_2Cys_6$ transcription factor, AflR, and its cofactor AflS. By replacing the intrinsic bidirectional *aflR/S* promoter with a nitrate bidirectional

inducible promoter, all eight identified genes can be simultaneously activated and repressed. As all eight identified promoters differ in their DNA sequence, the system has the potential to utilize one-step yeast recombinational cloning for assembly of entire secondary metabolite gene clusters. Here, we demonstrate the production of β-carotene by heterologous expression of three genes from *F. fujikuroi*. The inducibility of the system also is useful for production of toxic metabolites at a stage when the host strain has accumulated a significant biomass.

Abbreviations
SM: secondary metabolite; FDA: Food and Drug Administration; ST: sterigmatocystin.

Authors' contributions
PW, AAS, and NPK conceived and designed the experiments. PW, AAS, JSF, PMW, and AN performed the experiments. NPK provided reagents and materials. PW, AAS, and NPK wrote the manuscript. All authors read and approved the final manuscript.

Author details
[1] Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA. [2] Department of Bacteriology, University of Wisconsin, Madison, WI 53706, USA. [3] Present Address: Hexagon Bio, Menlo Park, CA 94025, USA. [4] Present Address: Department of Cell and Regenerative Biology, University of Wisconsin, Madison, WI 53705, USA. [5] Present Address: Davis Genome Center – Metabolomics, University of California, 451 Health Science Drive, Davis, CA 95616, USA. [6] Present Address: Ocean College, Zhejiang University, Hangzhou 310058, Zhejiang Province, People's Republic of China.

Competing interests
The expression system described in this study has been filed as patent in the United States under Number P150029US01 entitled "Methods and Systems for Producing Fungal Secondary Metabolites" for the Wisconsin Alumni Research Foundation (WARF).

References
1. Albermann S, Linnemannstöns P, Tudzynski B. Strategies for strain improvement in *Fusarium fujikuroi*: overexpression and localization of key enzymes of the isoprenoid pathway and their impact on gibberellin biosynthesis. Appl Microbiol Biotechnol. 2013;97:2979–95.
2. Avalos J, Carmen Limón M. Biological roles of fungal carotenoids. Curr Genet. 2015;61:309–24.
3. Avalos J, Estrada AF. Regulation by light in *Fusarium*. Fungal Genet Biol. 2010;47:930–8.
4. Avalos J, Prado-Cabrero A, Estrada AF. Neurosporaxanthin production by *Neurospora* and *Fusarium*. Methods Mol Biol. 2012;898:263–74.
5. Baima S, Carattoli A, Macino G, Morelli G. Photoinduction of albino-3 gene expression in *Neurospora crassa* conidia. J Photochem Photobiol B. 1992;15:233–8.
6. Boecker S, Zobel S, Meyer V, Süssmuth RD. Rational biosynthetic approaches for the production of new-to-nature compounds in fungi. Fungal Genet Biol. 2016;89:89–101.
7. Bok JW, Soukup AA, Chadwick E, Chiang YM, Wang CC, Keller NP. VeA and MvlA repression of the cryptic orsellinic acid gene cluster in *Aspergillus nidulans* involves histone 3 acetylation. Mol Microbiol. 2013;89:963–74.
8. Bond C, Tang Y, Li L. Saccharomyces cerevisiae as a tool for mining, studying and engineering fungal polyketide synthases. Fungal Genet Biol. 2016;89:52–61.
9. Brown DW, Yu JH, Kelkar HS, Fernandes M, Nesbitt TC, Keller NP, Adams TH, Leonard TJ. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. Proc Natl Acad Sci USA. 1996;93:1418–22.
10. Burger G, Tilburn J, Scazzocchio C. Molecular cloning and functional characterization of the pathway-specific regulatory gene *nirA*, which controls nitrate assimilation in *Aspergillus nidulans*. Mol Cell Biol. 1991;11:795–802.
11. Butchko RA, Brown DW, Busman M, Tudzynski B, Wiemann P. Lae1 regulates expression of multiple secondary metabolite gene clusters in *Fusarium verticillioides*. Fungal Genet Biol. 2012;49:602–12.
12. Carberry S, Molloy E, Hammel S, O'Keeffe G, Jones GW, Kavanagh K, Doyle S. Gliotoxin effects on fungal growth: mechanisms and exploitation. Fungal Genet Biol. 2012;49:302–12.
13. Chiang YM, Szewczyk E, Davidson AD, Keller N, Oakley BR, Wang CC. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. J Am Chem Soc. 2009;131:2965–70.
14. Chiang YM, Oakley CE, Ahuja M, Entwistle R, Schultz A, Chang SL, Sung CT, Wang CC, Oakley BR. An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. J Am Chem Soc. 2013;135:7720–31.
15. Chiang YM, Ahuja M, Oakley CE, Entwistle R, Asokan A, Zutz C, Wang CC, Oakley BR. Development of genetic dereplication strains in *Aspergillus nidulans* results in the discovery of Aspercryptin. Angew Chem Int Ed Engl. 2016;55:1662–5.
16. Clardy J, Walsh C. Lessons from natural molecules. Nature. 2004;432:829–37.
17. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc Natl Acad Sci USA. 2006;103:10352–7.
18. Díaz-Sánchez V, Estrada AF, Trautmann D, Al-Babili S, Avalos J. The gene *carD* encodes the aldehyde dehydrogenase responsible for neurosporaxanthin biosynthesis in *Fusarium fujikuroi*. FEBS J. 2011;278:3164–76.
19. Díaz-Sánchez V, Limón MC, Schaub P, Al-Babili S, Avalos J. A RALDH-like enzyme involved in *Fusarium verticillioides* development. Fungal Genet Biol. 2016;86:20–32.
20. Ehrlich KC. Predicted roles of the uncharacterized clustered genes in aflatoxin biosynthesis. Toxins (Basel). 2009;1:37–58.
21. Fernandes M, Keller NP, Adams TH. Sequence-specific binding by *Aspergillus nidulans* AflR, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Mol Microbiol. 1998;28:1355–65.
22. Fujii I, Ono Y, Tada H, Gomi K, Ebizuka Y, Sankawa U. Cloning of the polyketide synthase gene *atX* from *Aspergillus terreus* and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. Mol Gen Genet. 1996;253:1–10.
23. Green MR, Sambrook J. Molecular cloning: a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2012.
24. Gressler M, Hortschansky P, Geib E, Brock M. A new high-performance heterologous fungal expression system based on regulatory elements from the *Aspergillus terreus* terrein gene cluster. Front Microbiol. 2015;6:184.

25. Kealey JT, Liu L, Santi DV, Betlach MC, Barr PJ. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. Proc Natl Acad Sci USA. 1998;95:505–9.

26. Li C, Schmidhauser TJ. Developmental and photoregulation of *al-1* and *al-2*, structural genes for two enzymes essential for carotenoid biosynthesis in *Neurospora*. Dev Biol. 1995;169:90–5.

27. Lim FY, Keller NP. Spatial and temporal control of fungal natural product synthesis. Nat Prod Rep. 2014;31:1277–86.

28. Lim FY, Hou Y, Chen Y, Oh JH, Lee I, Bugni TS, Keller NP. Genome-based cluster deletion reveals an endocrocin biosynthetic pathway in *Aspergillus fumigatus*. Appl Environ Microbiol. 2012;78:4117–25.

29. Lim FY, Ames B, Walsh CT, Keller NP. Co-ordination between BrlA regulation and secretion of the oxidoreductase FmqD directs selective accumulation of fumiquinazoline C to conidial tissues in *Aspergillus fumigatus*. Cell Microbiol. 2014;16:1267–83.

30. Lin YJ, Chang JJ, Lin HY, Thia C, Kao YY, Huang CC, Li WH. Metabolic engineering a yeast to produce astaxanthin. Bioresour Technol. 2017;245:899–905.

31. Linnemannstöns P, Prado MM, Fernández-Martín R, Tudzynski B, Avalos J. A carotenoid biosynthesis gene cluster in *Fusarium fujikuroi*: the genes *carB* and *carRA*. Mol Genet Genomics. 2002;267:593–602.

32. Linz JE, Chanda A, Hong SY, Whitten DA, Wilkerson C, Roze LV. Proteomic and biochemical evidence support a role for transport vesicles and endosomes in stress response and secondary metabolism in *Aspergillus parasiticus*. J Proteome Res. 2012;11:767–75.

33. Marcos AT, Ramos MS, Marcos JF, Carmona L, Strauss J, Cánovas D. Nitric oxide synthesis by nitrate reductase is regulated during development in *Aspergillus*. Mol Microbiol. 2016;99:15–33.

34. Mende K, Homann V, Tudzynski B. The geranylgeranyl diphosphate synthase gene of *Gibberella fujikuroi*: isolation and expression. Mol Gen Genet. 1997;255:96–105.

35. Meyer V, Nevoigt E, Wiemann P. The art of design. Fungal Genet Biol. 2016;89:1–2.

36. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod. 2012;75:311–35.

37. Normile D. Agricultural biotechnology. Monsanto donates its share of golden rice. Science. 2000;289:843–5.

38. Oakley CE, Ahuja M, Sun WW, Entwistle R, Akashi T, Yaegashi J, Guo CJ, Cerqueira GC, Russo Wortman J, Wang CC, Chiang YM, Oakley BR. Discovery of McrA, a master regulator of *Aspergillus* secondary metabolism. Mol Microbiol. 2017;103:347–65.

39. Oldenburg KR, Vo KT, Michaelis S, Paddon C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. Nucleic Acids Res. 1997;25:451–2.

40. Palmer JM, Perrin RM, Dagenais TR, Keller NP. H3K9 methylation regulates growth and development in *Aspergillus fumigatus*. Eukaryot Cell. 2008;7:2052–60.

41. Peláez F. The historical delivery of antibiotics from microbial natural products—can history repeat. Biochem Pharmacol. 2006;71:981–90.

42. Pfannenstiel BT, Zhao X, Wortman J, Wiemann P, Throckmorton K, Spraker JE, Soukup AA, Luo X, Lindner DL, Lim FY, Knox BP, Haas B, Fischer GJ, Choera T, Butchko RAE, Bok JW, Affeldt KJ, Keller NP, Palmer JM. Revitalization of a forward genetic screen identifies three new regulators of fungal secondary metabolism in the genus *Aspergillus*. MBio. 2017;8:e01246-17.

43. Porquier A, Morgant G, Moraga J, Dalmais B, Luyten I, Simon A, Pradier JM, Amselem J, Collado IG, Viaud M. The botrydial biosynthetic gene cluster of *Botrytis cinerea* displays a bipartite genomic structure and is positively regulated by the putative Zn(II)2Cys6 transcription factor BcBot6. Fungal Genet Biol. 2016;96:33–46.

44. Punt PJ, Greaves PA, Kuyvenhoven A, van Deutekom JC, Kinghorn JR, Pouwels PH, van den Hondel CA. A twin-reporter vector for simultaneous analysis of expression signals of divergently transcribed, contiguous genes in filamentous fungi. Gene. 1991;104:119–22.

45. Punt PJ, Strauss J, Smit R, Kinghorn JR, van den Hondel CA, Scazzocchio C. The intergenic region between the divergently transcribed *niiA* and *niaD* genes of *Aspergillus nidulans* contains multiple NirA binding sites which act bidirectionally. Mol Cell Biol. 1995;15:5688–99.

46. Qiao J, Kontoyiannis DP, Wan Z, Li R, Liu W. Antifungal activity of statins against *Aspergillus* species. Med Mycol. 2007;45:589–93.

47. Schumacher J. Tools for *Botrytis cinerea*: New expression vectors make the gray mold fungus more accessible to cell biology approaches. Fungal Genet Biol. 2012;49:483–97.

48. Shimizu K, Keller NP. Genetic involvement of a cAMP-dependent protein kinase in a G protein signaling pathway regulating morphological and chemical transitions in *Aspergillus nidulans*. Genetics. 2001;157:591–600.

49. Soukup AA, Fischer GJ, Luo J, Keller NP. The *Aspergillus nidulans* Pbp1 homolog is required for normal sexual development and secondary metabolism. Fungal Genet Biol. 2017;100:13–21.

50. Sun WW, Guo CJ, Wang CCC. Characterization of the product of a non-ribosomal peptide synthetase-like (NRPS-like) gene using the doxycycline dependent Tet-on system in *Aspergillus terreus*. Fungal Genet Biol. 2016;89:84–8.

51. Wiemann P, Keller NP. Strategies for mining fungal natural products. J Ind Microbiol Biotechnol. 2013;41:301–13.

52. Yamano S, Ishii T, Nakagawa M, Ikenaga H, Misawa N. Metabolic engineering for production of beta-carotene and lycopene in *Saccharomyces cerevisiae*. Biosci Biotechnol Biochem. 1994;58:1112–4.

53. Yeh HH, Ahuja M, Chiang YM, Oakley CE, Moore S, Yoon O, Hajovsky H, Bok JW, Keller NP, Wang CC, Oakley BR. Resistance gene-guided genome mining: Serial promoter exchanges in *Aspergillus nidulans* reveal the biosynthetic pathway for fellutamide B, a proteasome inhibitor. ACS Chem Biol. 2016;11:2275–84.

54. Yin WB, Baccile JA, Bok JW, Chen Y, Keller NP, Schroeder FC. A nonribosomal peptide synthetase-derived iron(III) complex from the pathogenic fungus *Aspergillus fumigatus*. J Am Chem Soc. 2013;135:2064–7.

55. Yin WB, Chooi YH, Smith AR, Cacho RA, Hu Y, White TC, Tang Y. Discovery of cryptic polyketide metabolites from dermatophytes using heterologous expression in *Aspergillus nidulans*. ACS Synth Biol. 2013;29:549–55.

56. Yu JH, Butchko RA, Fernandes M, Keller NP, Leonard TJ, Adams TH. Conservation of structure and function of the aflatoxin regulatory gene *aflR* from *Aspergillus nidulans* and *A. flavus*. Curr Genet. 1996;29:549–55.

57. Zhang P, Wang X, Fan A, Zheng Y, Liu X, Wang S, Zou H, Oakley BR, Keller NP, Yin WB. A cryptic pigment biosynthetic pathway uncovered by heterologous expression is essential for conidial development in *Pestalotiopsis fici*. Mol Microbiol. 2017;105:469–83.

Regulation of plant cell wall degradation by light in *Trichoderma*

Monika Schmoll[*]

Abstract

Trichoderma reesei (syn. *Hypocrea jecorina*) is the model organism for industrial production of plant cell wall degradating enzymes. The integration of light and nutrient signals for adaptation of enzyme production in *T. reesei* emerged as an important regulatory mechanism to be tackled for strain improvement. Gene regulation specific for cellulase inducing conditions is different in light and darkness with substantial regulation by photoreceptors. Genes regulated by light are clustered in the genome, with several of the clusters overlapping with CAZyme clusters. Major cellulase transcription factor genes and at least 75% of glycoside hydrolase encoding genes show the potential of light dependent regulation. Accordingly, light dependent protein complex formation occurs within the promoters of cellulases and their regulators. Additionally growth on diverse carbon sources is different between light and darkness and dependent on the presence of photoreceptors in several cases. Thereby, also light intensity plays a regulatory role, with cellulase levels dropping at higher light intensities dependent in the strain background. The heterotrimeric G-protein pathway is the most important nutrient signaling pathway in the connection with light response and triggers posttranscriptional regulation of cellulase expression. All G-protein alpha subunits impact cellulase regulation in a light dependent manner. The downstream cAMP pathway is involved in light dependent regulation as well. Connections between the regulatory pathways are mainly established via the photoreceptor ENV1. The effect of photoreceptors on plant cell wall degradation also occurs in the model filamentous fungus *Neurospora crassa*. In the currently proposed model, *T. reesei* senses the presence of plant biomass in its environment by detection of building blocks of cellulose and hemicellulose. Interpretation of the respective signals is subsequently adjusted to the requirements in light and darkness (or on the surface versus within the substrate) by an interconnection of nutrient signaling with light response. This review provides an overview on the importance of light, photoreceptors and related signaling pathways for formation of plant cell wall degrading enzymes in *T. reesei*. Additionally, the relevance of light dependent gene regulation for industrial fermentations with *Trichoderma* as well as strategies for exploitation of the observed effects are discussed.

Keywords: *Trichoderma reesei*, *Hypocrea jecorina*, CAZymes, Light response, Signal transduction, Surface sensing, Carbon source utilization, EMSA, Genomic clusters

Background

The continuing alteration between light and darkness on earth caused by its rotation resulted in an evolutionary adaptation of the majority of living beings to day and night. The physiological changes connected to day and night are triggered by the circadian clock, which governs preparation to the upcoming day or night. Exposure

to light at unexpected times causes phase shifts in the clock controlled cycles and hence alters the connected gene regulation. Nevertheless, control by the circadian clock and response to changing light conditions are distinct processes [1]. This adaptation does not only concern obvious necessities such as the protection against harmful UV light during the day or dealing with higher temperatures and lower humidity levels connected to sunlight versus darkness. Also, the adjustment of metabolic processes to light and darkness has evolved—not only in humans, but also in fungi [2, 3],—which raises

*Correspondence: monika.schmoll@ait.ac.at
Center for Health and Bioresources, AIT Austrian Institute of Technology GmbH, Konrad Lorenz Straße 24, 3430 Tulln, Austria

the question whether fungi do have an equivalent to the physiological condition of sleep or the physiologically different situation in day and night in humans.

In fungi, light is highly relevant to diverse physiological regulation mechanisms and impacts many signaling pathways that integrate light response with metabolism, stress response and development [4]. Even though many of the fungi currently investigated in academia and industry never experience natural conditions of day and night, their gene expression levels and physiology still follow circadian rhythms [5], which also target metabolism [6–8]. Groundbreaking work in elucidation of the molecular machinery governing circadian rhythmicity and light response has been done in *Neurospora crassa* [9–11], which involves transcriptional cascades as well as epigenetic and posttranslational modifications [12, 13] and many of the studies on light effects in dthe fungi built on the discoveries from *N. crassa* thereafter.

Trichoderma reesei represents one of the most important filamentous fungi nowadays used in industry for production of homologous and heterologous enzymes—predominantly for biofuel production [14, 15]. *T. reesei* expresses diverse carbohydrate active enzymes (CAZymes), the most important being cellulases and hemicellulases [16]. Induction of these enzymes occurs on different carbon sources such as cellulose or lactose or in the presence of sophorose, but also on other carbon sources representing building blocks of plant cell wall material [17, 18]. Repression occurs on easily metabolizable carbon sources like glucose by the function of carbon catabolite repression. Secondary metabolism has not been studied in detail in *T. reesei* yet, but harmful mycotoxins are not known to be produced by this fungus [19]. *T. reesei* was the first industrially relevant filamentous fungus for which the method of sexual crossing became available. Sexual development is dependent on specific conditions of light, temperature and carbon source in the medium and numerous regulators, including the *T. reesei* photoreceptors and several signaling compounds [20, 21].

In this review, I will first give a short overview on composition of plant cell walls and their degradation followed by a general introduction to the light response pathway and known physiological effects of light on fungi in order to familiarize the reader with the two topics connected in this review. Thereafter I briefly describe the discovery of the influence of light on cellulase regulation along with some general findings on the topic later on. Subsequently, I explain the impact of light and the light response machinery on regulatory pathways influencing metabolic functions in *Trichoderma* with an emphasis on plant cell wall degradation and interconnections between nutrient and light signaling pathways (Fig. 1).

Individual regulatory factors as well as their genome/transcriptome wide effects in dependence of light will be discussed including alterations on different carbon sources, non random distribution of light regulated genes and the interplay of light response with carbon catabolite repression. Based on that, light dependent differences at the promoter level and in carbon source utilization are outlined. Finally, the relevance of light for regulation of specific, known factors crucial for highly efficient plant cell wall degradation as well as aspects of light influenced processes for research and industry (Fig. 2) are discussed.

The plant cell wall and the enzymes degrading it

Lignocellulosic plant biomass represents the most important carbon source on our planet and it is hence of high ecological importance as major component in the global carbon cycle. Plant biomass is composed of several polymers including the recalcitrant cellulose, hemicellulose and lignin, but also pectin and starch [22]. Filamentous fungi are highly efficient degraders of plant biomass and have evolved a complex yet very efficient machinery for degradation of plant cell walls [23, 24]. The involved Carbohydrate Active enZymes (CAZys) act on glycoside linkages in the plant cell wall polysaccharides and can have a broad range of activities [25, 26]. It is assumed that glycoside hydrolases, the most important enzymes for plant cell wall degradation, have evolved from a common ancestor. In the course of the subsequent specialization, saprotropic fungi like *T. reesei* appear to have lost a significant number of genes including glycoside hydrolases [27]. Since cellulases are usually found coregulated, the cellobiohydrolases CBH1/Cel7a and CBH2/CEL6a are frequently used as representatives for cellulase expression. In recent years, the oxidation of cellulose by polysaccharide monooxygenases (previously assigned to glycoside hydrolase family 61) was shown to significantly contribute to plant cell wall degradation [28, 29]. Interestingly, an involvement of light and photosynthetic pigments in oxidation of polysaccharides was postulated [30].

Regulation of plant cell wall degrading enzymes is complex and occurs at the transcriptional level involving a plethora of transcription factors and regulators, some of the most important being XYR1, ACE1, ACE2, ACE3, VIB1 and the HAP complex [23, 31, 32]. Recently, also a post-transcriptional section of cellulase regulation was shown [33]. Thereby, a considerable number of substrates cause induction of expression of plant cell wall degrading enzymes [16–18]. Importantly, the mechanism of carbon catabolite repression, with its major transcription factor CRE1 serves to avoid biosynthesis of enzymes in the presence of easily metabolizable carbon sources [23, 34]. By application of their sophisticated regulatory machineries,

Fig. 1 Schematic overview of pathways involved in light regulation of plant cell wall degradation
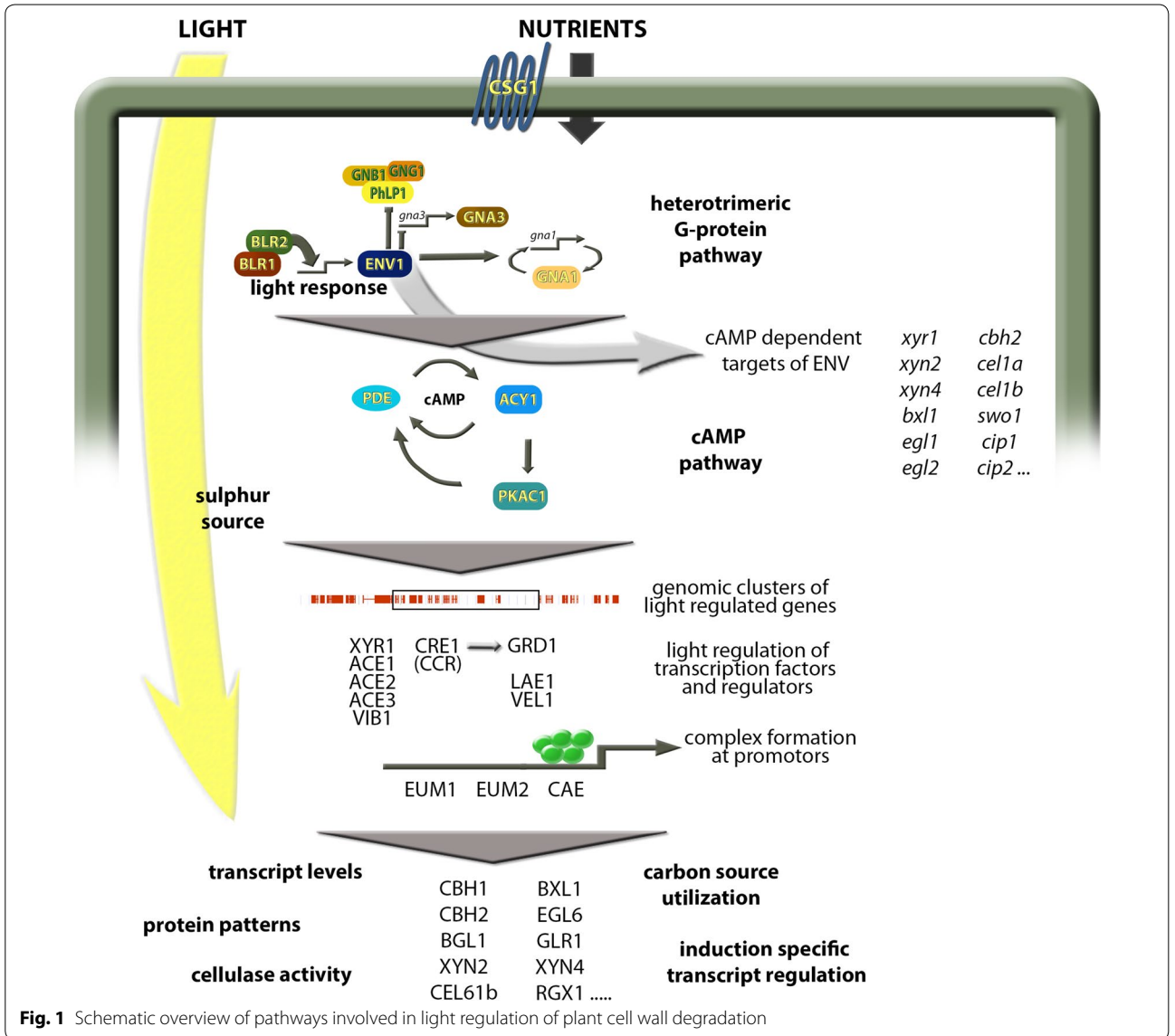
fungi adjust the amount and type of enzymes to produce to their environment—be it a tropical forest or an industrial fermentor.

The light response pathway in *Trichoderma* spp.

Trichoderma spp. have a long tradition of research towards fungal light responses and their physiological consequences [35, 36]. Eight proteins are considered to be responsible for light perception [36] including BLR1 and BLR2 (blue light regulators 1 and 2): two GATA-type transcription factors predominantly acting as a photoreceptor complex, the photoreceptor ENV1, two photolyases, a cryptochrome, a phytochrome and an opsin, which is however only present in the genome of *Trichoderma atroviride*, but not in *T. reesei* [36, 37].

The major blue light photoreceptors BLR1 and BLR2, which are the homologues of *N. crassa* WC-1 and WC-2 (white collar 1 and 2) [38], were first characterized in *T. atroviride*, where they are essential for blue light induced conidiation [39]. Thereafter, ENV1, a homologue of *N. crassa* VVD [40] and the third photoreceptor was detected in *T. reesei* and an unexpectedly broad range of physiological functions in light response, development, stress response and metabolism was shown for ENV1 that exceed those of BLR1 or BLR2 [41, 42]. In contrast to *T. atroviride*, deletion of *blr1* or *blr2* does not abolish conidiation in *T. reesei* [43, 44], while lack of ENV1 causes a severe growth phenotype in light [42], which does not occur on every carbon source [45, 46]. In both fungi, BLR1 and BLR2 exert common as well as independent
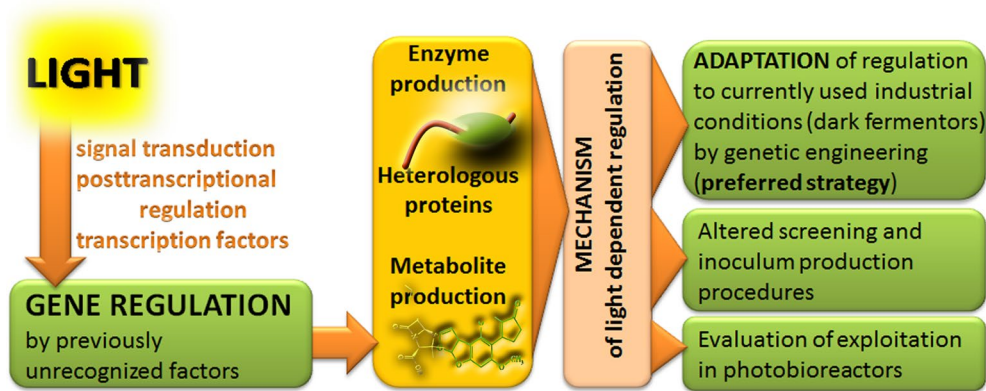
Fig. 2 Schematic representation of strategies for strain improvement by exploiting of light dependent effects. Gene regulation in *T. reesei* is considerably influenced by light, with the light signal coordinated with nutrient signals via the signal transduction pathways of heterotrimeric G-proteins and cAMP signaling. Posttranscriptional regulation of cellulase gene expression is triggered by a G-protein coupled receptor, whose signal is channeled through the G-protein pathway and its subunits. The consequences of light exposure include changes in normal enzyme production, production of heterologous proteins and secondary metabolites expressed from homologous or heterologous gene clusters. Once the mechanisms of light dependent regulation are understood, this information can serve to improve performance under currently applied industrial conditions (dark fermentors) by knowledge based genetic engineering. Additionally, screening procedures are recommended to be performed under controlled light conditions and inoculum production can be improved. For high value products, illumination or specifically applied light regimes in photobioreactors can be evaluated

functions and also influence gene expression in darkness [46, 47]. Consequently, while there is considerable overlap in functions, there are also some differences in the light response machinery in *Trichoderma* spp. which may cause altered responses.

BLR1 and BLR2 are GATA type zinc finger transcription factors and are assumed to act in a complex. BLR1 contains three PAS domains, and a PAS/LOV domain, which is responsible for reception of the light signal due to the flavin moiety bound in this domain. BLR2 contains only one PAS domain and ENV1 contains a PAS/LOV domain like BLR1 [36]. Light dependent induction of *env1* is strictly dependent on the presence of both BLR1 and BLR2 [43, 48, 49]. Nevertheless, phenotypic effects of deletion of *blr1* or *blr2* are much less severe than deletion of *env1* in light [41–44, 46].

The hierarchy in the light response cascade thereby depends on the function of the respective regulators in *T. reesei*. In many cases, ENV1 acts via BLR1 and BLR2, for example with respect to growth and sexual fertility [48]. Like BLR1 and BLR2, also ENV1 has individual functions [46], albeit often effects reminiscent of photoadaptation occur: If the negative effect of ENV1 is relieved by its deletion, the usually balancing positive effect of BLR1 and BLR2 becomes obvious [48]. This balanced function of ENV1, BLR1 and BLR2 further explains the rescue of the severe growth phenotype of strains lacking ENV1 in light in Δ*env1* double or triple mutants with *blr1* and/or *blr2* deletions [48].

In *T. atroviride*, enhanced expression of *blr2* causes increased photoconidiation and higher transcript levels of light induced genes, while the opposite effect is caused by overexpression of *blr1*. *Blr2* overexpression further causes higher sensitivity to blue light and complex formation of BLR1/BLR2 is assumed to be required for appropriate light perception in *T. atroviride* [50].

Physiological responses to light in *Trichoderma*
Life in soil—a mainly dark environment—and on the surface of a substrate are fundamentally different in terms of UV radiation, oxygen, temperature stability, humidity and levels of reactive oxygen species (ROS). Also during day and night many of these environmental properties change and hence, fungi evolved to use light as a fast indicator of the expected changes and prepare for daily alterations by applying circadian rhythms [51].

Blue light responses in *T. reesei* include enhanced conidiation [43] and sexual reproduction [21, 52], secondary metabolism [46, 53, 54], growth [45] and altered regulation of enzyme gene transcription [46, 53], expression and activity [55]. Genes of the light response machinery including *phr1* encoding a photolyase, the photoreceptor gene *env1* and *frq*, which is required for circadian rhythmicity in *N. crassa*, the MAPkinase gene *tmk3* [45] as well as several genes involved in sexual development are light dependently regulated irrespective of the carbon source [33].

The earliest response to light evaluated in *T. reesei* was after 15 min [45], albeit much shorter light pulses

are assumed to be sensed and can cause altered gene regulation. Thereby it has to be considered that not only the duration but also the light intensity (fluence rate) is important [55]. An effect of red light on the regulation of few genes was detected for *T. atroviride* [56], while in *T. reesei* no effect of red light was observed so far [43].

In *T. atroviride*, light stimulates the tolerance to osmotic stress through the Hog1-related MAPkinase TMK3. This stimulation of stress signaling pathways by light is considered a benefit for the cell [57]. Interestingly, *T. reesei* TMK3 was found to regulate cellulase production [58, 59], albeit an influence of light was not tested in this study. Together these findings indicate that the MAPkinase pathway contribute to light dependent modulation of cellulase gene expression.

In *T. reesei* an involvement of ENV1 in stress response was observed, which requires the conserved amino acids C96 and T101 [60, 61]. Thereby, C96 is evolutionarily conserved in Hypocreaceae, which in contrast to for example *N. crassa*, indicating the integration of stress and light responses via ENV1.

An unexpected effect of light and photoreceptors on cellulase gene expression

Traditionally, filamentous fungi were grown under random light conditions, except for investigation of circadian rhythms or photoreceptor functions in *N. crassa*. *Trichoderma* spp. also served as models for analysis of light responses in the 1970s and 80s [36], but this did not lead to cultivation under controlled light conditions thereafter because no connection to carbon metabolism was assumed.

A screening assay for genes differentially expressed between the cellulase negative mutant strain QM9978 and the early high production mutant QM9414 surprisingly yielded the photoreceptor gene *env1* as differentially regulated between the two strains under inducing conditions (cellulose and sophorose) [62]. ENV1 is a homologue of the *N. crassa* photoreceptor VIVID (VVD), which negatively acts on the photoreceptor complex established by WC-1 and WC-2, the homologues of BLR1 and BLR2 [63, 64]. However, *T. reesei env1* could not complement a *N. crassa vvd* non-functional mutant and is hence not considered a functional homologue [42]. Still, subsequent research revealed a considerable functional overlap of ENV1 and VVD.

Consequently, cultivation of *T. reesei* was performed under controlled light conditions and with the cellulase inducing carbon sources cellulose and lactose in order to assess a potential effect of light on enzyme expression [42]. Indeed, striking differences in cellulase transcript levels between growth in light and in darkness were observed upon growth on cellulose, which did not correlate with previous results in a conventional incubator with random light pulses—neither in light nor in darkness. Transcript levels of *cbh1* increased by roughly 50% upon growth in constant light on cellulose and an important function of ENV1 in this regulation became obvious, with decreased *cbh1* levels on lactose, transient increase on cellulose in darkness and decrease in light in the mutant strain [42]. Upon induction of cellulase gene expression by the natural inducer sophorose, we found that in darkness, cellulase transcript levels increase with time, while in light cellulase induction appears to be accelerated, but transient [65].

Functions of the light response machinery in cellulase regulation

Searching for the basis of light dependent regulation of cellulase genes, comparison of the promoters of *env1* and *vvd* revealed two potential DNA binding motifs, EUM1 (envoy upstream motif; 5′ CTGTGC 3′) and EUM2 (5′ ACCTTGAC 3′). EUM1 is also present in the promoters of the cellulase genes *cbh1* and *cbh2* as well as in the promoters of *blr1* and *blr2* suggesting a potential for co-regulation or feedback [42] (see also below). Additionally, the same study revealed that *env1* is not expressed in the cellulase negative strain QM9978, which has a mutation in the EUM1 promoter motif in the *env1* promoter. However, complementation with the entire *env1* gene that included this promoter region did not rescue cellulase expression in QM9978 [42]. EUM1, light responsive elements (LREs) [66] and GATA sites, which all might be predictive for binding of the photoreceptor complex or other light regulatory factors, were found in numerous promoters of known cellulose and hemicellulose degrading enzymes [67].

Later on, investigation of the genome of QM9978 and comparison to that of QM6a recently led to the finding that the PhoG homologue VIB1 is responsible for the defect in cellulase induction of this strain. A translocation of the gene abolished expression of *vib1* in QM9978 [68]. Interestingly, *vib1* is strongly down-regulated in light upon growth on cellulose in the wild-type strain QM6a (Stappler and Schmoll, unpublished). Downregulation of *vib1* in light is further observed upon growth on lactose or sophorose [33]. Transcriptome analysis upon growth on cellulose showed that *vib1* is strongly regulated by ENV1 in light and by the adenylate cyclase ACY1 in light [33, 49]. Consequently, *vib1* is a cAMP dependent target of the light response pathway [49], which is also in agreement with the finding of the gene encoding ENV1 as a regulator of *vib1* to be down-regulated in QM9978 compared to QM9414 under inducing conditions [42] [62].

BLR1 and BLR2 positively regulate transcript levels of *cbh1* in *T. reesei* [43] and in *T. atroviride* [69]. Interestingly, effects are obvious in light and darkness, indicating that these photoreceptors also have a function in darkness in *T. reesei* [43, 46]. Accordingly, effects of ENV1, BLR1 and BLR2 were also observed in largely dark fermentor cultivations in *T. reesei*. Under these conditions, Δ*env1* secretes a more efficient enzyme mixture and for BLR2 an effect on secretion capacity was found, which leads to increased cellulase activities as well. Δ*blr1* forms more biomass, but its lower secretion capacity counteracts its efficiency in production of cellulose degrading enzymes [67].

Abundance of secreted proteins is considerably altered in light [55] including such important enzymes as CBH1, BGL1, XYN2 or CEL61B. BLR1 and BLR2 have a major impact in this light dependent regulation. While initial analyses on light response of cellulase regulation were performed in the QM9414 strain background, we later found significant differences between light and darkness also for QM6a [55], RutC30 and industrial production strains as well (unpublished results). Thereby, QM6a shows a much lower light tolerance than QM9414 in terms of alteration of cellulase gene expression. In QM9414 *cbh1* transcript levels initially increase with increasing light intensities and only drop at 5000 lx, while in QM6a already low light intensities abolish *cbh1* transcription. In Δ*blr1* and Δ*blr2*, this drop at high light intensities does not occur [55].

In addition to the photoreceptors, further regulators involved in light dependent signaling were identified as cellulase regulators. In *Aspergillus nidulans*, VeA and LaeA coordinate the light signal with fungal development and secondary metabolism [70]. Their homologues in *T. reesei*, VEL1 and LAE1 are important regulators of cellulase gene expression [71, 72]. Unfortunately, the latter studies were done under uncontrolled light conditions and hence a light dependent relevance of this regulation is not known. Since the function of VEL1 in development shows light dependent differences and a connection to photoreceptors [73, 74], this can also be expected for cellulase regulation.

Signaling pathways involved in light dependent modulation of cellulase gene expression

For the presence of cellulose to be sensed in the environment, low molecular weight degradation products liberated from degradable plant material are likely signals that may be sensed by membrane bound receptors. Alternatively, an inducer synthesized outside the cell may be taken up via diffusion or a permease and initiate cellulase formation by an intracellular process. Both hypotheses are likely to describe contributions to the regulation of cellulase gene expression.

The heterotrimeric G-protein pathway

The prime candidate pathway for sensing and transmission of an extracellular cellulose related signals is the heterotrimeric G-protein pathway. Signal transduction via the heterotrimeric G-protein pathway starts with reception of a signaling compound by a G-protein coupled receptor (GPCR). Once such a ligand binds to the GPCR, the heterotrimeric G-protein complex composed of G-protein alpha, beta and gamma subunits, dissociates, GDP is exchanged for GTP on the G-alpha subunit for activation and all subunits then modulate their target pathways [37, 75, 76].

T. reesei has 50 GPCRs, 3 G-alpha subunits (GNA1, GNA2 and GNA3), one G-beta subunit (GNB1) and one G-gamma subunit (GNG1). Additionally, two phosducin like proteins (PhLP1 and PhlP2), which are assumed to act as co-chaperones for G-protein beta and gamma folding are encoded in the genome as well as 7 RGS (regulator of G-protein signaling) domain proteins [37]. The G-protein alpha subunits GNA1 and GNA3 impact cellulase gene expression and this function is dependent on light [77, 78]. Constitutive activation of GNA3 led to a strong increase in cellulase transcripts upon growth on cellulose (around tenfold), but only in light. In darkness, neither constitutive activation nor knock down altered cellulase transcript levels [77]. Interestingly, upregulation of *gna3* without constitutive activation caused somewhat increased cellulase levels in light, which however did by far not reach the strong up regulation seen for the constitutive activation of GNA3. It is therefore likely that the rate of inactivation of the alpha subunit by the intrinsic GTPase of GNA3 plays a role. This function is enhanced by RGS (regulator of G-protein signaling) proteins, which are consequently assumed to be involved in cellulase regulation in *T. reesei* [77]. Of those, *rgs1* represents a light independent regulatory target of GNB1, GNG1 and PhLP1 [53].

Lack of GNA1 causes a strong increase of cellulase gene expression in darkness and abolishment in light upon growth on cellulose. Constitutive activation of GNA1 caused—like for GNA3—an increase in cellulase transcript levels in light [78]. For both GNA1 and GNA3, constitutive activation did not enable inducer independent cellulase gene expression. Consequently, the signal for induction of cellulase gene expression is separate from that transmitted by GNA1 and GNA3 and may indeed involve intracellular detection of an inducer.

The G-protein beta and gamma subunits GNB1 and GNG1 as well as the phosducin-like protein PhLP1, which is assumed to act as a chaperone for complex

formation of GNB1 and GNG1, influence regulation of several glycoside hydrolases including *cbh1* and *cbh2* [53]. Cellulase activity increases in all three mutants, while transcript abundance of several genes encoding plant cell wall degrading enzymes decreases, indicating posttranscriptional regulation [53]. However, the light dependent effects of GNB1, GNG1 and PhLP1 are considerably less pronounced than the strongly positive impact of the G-alpha subunits GNA1 and GNA3 [53]. In summary, these findings are in line with a positive effect of phosducin-like proteins on the efficiency of G-protein signaling [79].

Only recently, a new aspect was added to the function of the heterotrimeric G-protein pathway in cellulase regulation. The class XIII of G-protein coupled receptors (GPCRs) was found to have only a minor influence on cellulase transcript levels (up to 50%), but both were required for normal specific cellulase activities on lactose. In the absence of CSG1, representing one of the two class XIII GPCRs of *T. reesei*, secreted activity levels even decreased to basal levels upon growth on cellulose and lactose [33]. Hence it is assumed that the signal received by CSG1 is responsible for posttranscriptional regulation of cellulase gene expression, a mechanism previously not known to be involved. Based on the light dependent function of GNA1 and GNA3 it can be assumed that these downstream subunits integrate the light dependent relevance with the signal received by CSG1 to a physiological response adapted to the change of day and night.

The cAMP pathway

Cyclic adenosine monophosphate (cAMP) is one of the most important secondary messengers in living organisms. cAMP levels function as coincidence detectors due to their regulation by diverse environmental cues that are integrated to a defined level, which determines the extent of the effect on downstream pathways. Adenylate cyclase synthesizes cAMP and phosphodiestereases, which are activated by the presence of cAMP [80, 81], degrade it [82], forming a feedback cycle for adjustment of levels [80, 81]. Protein kinase A is cAMP dependent and consists of catalytic and regulatory subunits [37]. Light dependent effects on cAMP levels and adenylate cyclase activity have been shown previously for *Trichoderma* [83–85].

The cAMP dependent protein kinase A plays an important role in regulation of light responses in *T. atroviride* [86]. In *T. reesei*, the cAMP pathway is one of the targets of the heterotrimeric G-protein pathway and cAMP levels are impacted by both GNA1 and GNA3 [77, 78]. Investigation of the function of adenylate cyclase (ACY1) and the catalytic subunit 1 of protein kinase A (PKAc1) showed a light dependent effect on cellulase

gene expression in both cases upon growth on lactose [87]. Deletion of *acy1* and *pkac1* caused increased light responsiveness of cellulase genes i.e. strongly increased differences between cellulase levels in light and darkness compared to the light dependent regulation in the wild-type. Regulation of the transcript levels of the transcription factors ACE1 and ACE2 that regulate transcription of cellulase genes show a response to light, but do not correlate well with cellulase levels. Modification of transcript levels of the cellulase regulator XYR1 strongly resembles the pattern of regulation seen for cellulase genes. Hence, it is assumed that the cAMP pathway does not directly target phosphorylation of XYR1, but rather causes phosphorylation of a regulator of XYR1, which by modification of XYR1 acts on cellulase transcription [87].

In this respect it is interesting that deletion of *acy1* or *pkac1* also causes decreased levels of the photoreceptor gene *env1* [87], which regulates cellulase levels [42]. Phosphorylation of the photoreceptor complex by protein kinase A was shown in *N. crassa* [88] and critically impacts the function of the circadian clock by establishing a negative feedback loop. A similar mechanism may be responsible for regulation of *env1* levels in *T. reesei*. In turn, strains lacking ENV1 show strongly decreased cAMP levels, which is attributed to an impact of ENV1 on the function of phosphodiesterases rather than adenylate cyclase [89]. Such a mechanism would be in agreement with a feedback mechanism for fine-tuning of the integration of light response with nutrient signaling. Only recently we could confirm the involvement of a phosphodiesterase in light dependent cellulase regulation and the connection to ENV1 (E. Stappler et al., manuscript in preparation).

Interconnections between light response machinery and nutrient signaling pathways

The finding that light and its photosensors are involved in regulation of cellulase gene expression, indicated that nutrient utilization is interrelated with light response. The findings of the influence of the light signaling pathway on carbon-compound and carbohydrate metabolism [53] further raised the question as to the interconnections between nutrient signaling and sensing light. More detailed investigation of the heterotrimeric G-protein pathway and the cAMP pathway indicated that the light signal must be integrated with the transmitted nutrient signals.

BLR1 and BLR2 do not have a major impact on the G-protein signaling pathway [49]. ENV1 negatively influences *blr1* and *blr2* transcript levels, which in turn are required for induction of *env1* in light. This mutual regulation causes a steady state level of transcripts for these three genes [49]. ENV1 is required for the positive

feedback cycle of GNA1, which increases *gna1* transcript levels upon constitutive activation. In turn, GNA1 negatively regulates *env1* transcript levels [89]. Concerning the regulation mechanism of GNA3/*gna3*, ENV1 is not involved in establishing the *gna3* feedback cycle, but negatively regulates *gna3* transcript levels in light [89]. During investigation of the mutual regulatory effects of BLR1, BLR2 and ENV1 as well as GNA1, GNA3, GNB1, GNG1 and PhLP1, a pair with regulatory interaction at early light response emerged: PhLP1 and ENV1 [49]. With its negative effect on transcript levels of *phlp1*, *gnb1* and *gng1*, ENV1 dampens G-protein signaling during early light response, presumably to provide resources for protection from light. Thereafter, PhLP1 can exert a positive effect on complex formation of GNB1 and GNG1 and G-protein signaling in general, which is likely to be important for metabolic adaptation to light [49]. The striking overlap in regulatory targets of photoreceptors, G-protein pathway components and ACY1 suggests that this interrelationship constitutes a core function in adaptation beyond transient effects in early light response.

Surprisingly, these interconnection studies revealed that it is not the photoreceptor complex comprising the two BLR proteins, but rather ENV1 that acts in a central position as a checkpoint, which is in line with its strong effect on gene regulation in light [46]. The more prominent role of ENV1 in comparison to BLR1 and BLR2 is also obvious in the role of ENV1 in the regulation of other physiological processes such as sexual development or stress response [41].

The strikingly low cAMP levels in mutants lacking ENV1 indicated that ENV1 could exert at least part of its function via the cAMP pathway adding another link between light response and nutrient signaling (see also above). The Δ*acy1* strain showed a striking overlap of regulated genes in light, but not in darkness [49]. These genes have in common their regulation in the presence of strongly decreased or abolished cAMP levels. Among these cAMP targets there are numerous glycoside hydrolases, auxiliary proteins of cellulose degradation as well as the cellulase and hemicellulase regulator gene *xyr1* [49]. Functional category analysis of these genes showed a strong enrichment in C-compound and carbohydrate metabolism. Consequently, the regulatory function of ENV1 in targeting regulation of enzyme expression is to a considerable extent mediated by its effect on cAMP levels [49]. Interestingly, many of the genes down-regulated in Δ*env1* and Δ*acy1* in the light are also down-regulated on soluble carbon sources compared to the insoluble carbon source cellulose [33]. Consequently, the cAMP dependent output of the regulatory cascade triggered by ENV1 could be targeted at sensing and/or reaction to a surface of specifically to cellulose, but not to a soluble carbon source [33].

In summary, ENV1 emerged as a major checkpoint between nutrient and light signaling [41]. However, it still remains to be shown whether this interrelationship is established directly by protein–protein interaction and an influence on activity of the targeted regulators or if the influence is indirect by an impact on regulation of abundance and/or activity of signaling factors, for example kinases, which in turn target regulatory factors. The considerable impact of deletion of *env1* on the transcriptome in light [46] makes the latter hypothesis more likely.

Light dependent regulation of cellulose degradation as influenced by the sulphur source

Genes involved in metabolic processes are major targets of light response in *T. reesei* [46, 53]. Amino acid metabolism and sulphur metabolism are among these targets [33, 46, 53] and these functional categories are further found correlated with altered levels of cellulase production and carbon catabolite repression [90–92].

Studies aimed at identification of regulators binding to CAE within the *cbh2* promoter revealed the sulphur related regulator protein LIM1, a homologue of *N. crassa* SCON2, as a candidate regulator [93]. Investigation of expression of *lim1* under cellulase inducing conditions (cellulose, sophorose) in light and darkness supported this hypothesis. Interestingly, in the absence of a sulphur source, the methionine content, which can be used as a sulphur source, was important for cellulase gene expression and high sulphur conditions (5 mM) were even deleterious for growth in light. In the presence of normal sulphur levels in the MA minimal medium on cellulose, the same amount of methionine causes an increase of cellulase gene expression in darkness, but abolishment of *cbh1* transcription in light. Therefore, methionine appears to represent a signal molecule with a light dependent impact on cellulase regulation [93].

Additionally, sulphate uptake upon growth on cellulose is regulated by light, while this is not the case with glucose as carbon source [93]. Hence, light dependence of sulfur metabolism is dependent on the carbon source.

Interestingly, we found a significant enrichment (p value > 0.05) of genes involved in amino acid metabolism among those specifically down-regulated under cellulase inducing conditions [33]. Thereby, significant enrichment in methionine metabolism was detected only in darkness [33]. Moreover, significant enrichment in functions in methionine metabolism and sulphur metabolism were detected for down-regulated genes in Δ*env1* on cellulose [46]. Only recently, metabolic modelling of *T. reesei* under chemostat conditions with controlled growth rate

revealed sulphur assimilation as a major limiting factor of protein production [94].

Light dependent gene regulation patterns in *Trichoderma* spp.—a genome wide view

First data on the influence of light on the transcriptome showed regulation of 2.8% of genes in response to light in *T. atroviride* [56]. In *T. reesei,* a comparable extent of light dependent gene regulation (2.7%) was detected in the early high cellulase expression mutant QM9414 upon growth on cellulose [53]. The gene set up-regulated upon growth in light on cellulose is significantly enriched in metabolic functions (p-value 1.39e-04), particularly C-compound and carbohydrate metabolism (p-value 7.74e-06) and polysaccharide metabolism (p-value 1.28e-04), but also in nitrogen, sulfur and selenium metabolism, carbohydrate transport and peptide transport. Among down-regulated genes, significant enrichment occurs for metabolism of polyketides and toxin transport [53].

The same study however also showed that in the absence of components of the heterotrimeric G-protein pathway, considerably higher differences in gene regulation between light and darkness occur, with up to 23% of differential regulation in mutants lacking the G-protein beta subunit GNB1. Thereby, a high number of genes with decreased transcript abundance in light compared to darkness was found in strains lacking PhLP1, GNB1 or GNG1. Many of these genes are not simply down-regulated in light compared to the wild-type, but rather not up-regulated anymore because the light signal is not transmitted correctly. Accordingly, the highest number of consistently regulated genes in Δ*gnb1*, Δ*gng1* and Δ*phlp1* is found in the set of genes downregulated in light (628 genes including 21 glycoside hydrolase genes). This gene set includes a high number of genes associated with metabolism including 89 genes involved in C-compound and carbohydrate metabolism, but significant enrichment was only detected for functions in secondary metabolism (p-value 1.43e-08), photoperception and related functions. The strongest enrichment was observed for unclassified functions, indicating that the major function of GNB1, GNG1 and PhLP1 upon growth on cellulose still remains to be determined [53]. Available data confirm that PhLP1, GNB1 and GNG1 act in the same pathway and suggest that GNB1 and GNG1 may adjust sensitivity of the fungus to environmental signals.

Investigation of the transcriptomes of strains with mutations in the *blr1*, *blr2* or *env1* genes showed the same tendency of increased differences between transcript levels in light and darkness compared to wild-type. Light responsiveness of transcript abundances in Δ*env1* even exceeds that of Δ*gnb1* [46]. This analysis showed that the major function of BLR1 and BLR2 is observed in light, but clearly their absence did not abolish light response as the mutant strains are not blind. The majority of the regulatory targets of BLR1 and BLR2 was found in light, most of them being down-regulated in Δ*blr1* and Δ*blr2* (769 in Δ*blr1* and 873 in Δ*blr2*) and only few genes being up-regulated. In light and darkness, BLR1 and BLR2 have a considerable number of regulatory targets with major functions in C-compound and carbon metabolism, polysaccharide metabolism and transport [46]. Thereby, BLR1 and BLR2 share many targets, but in addition their function extends to individual targets as well, indicating that BLR1 and BLR2 not only act as a transcription factor complex but have individual functions as well. Independent roles have also been shown for their homologues in *T. atroviride* and *N. crassa* [47, 92]. In contrast to BLR1 and BLR2, ENV1 also exerts considerable negative effects on gene regulation in light [46].

In darkness, there are no shared targets between BLR1, BLR2 and ENV1. Only 20 genes show consistent upregulation, while 564 genes are down-regulated in light in Δ*blr1*, Δ*blr2* and Δ*env1*. These genes are dominated by functions in metabolism, including C-compound and carbohydrate metabolism including 22 glycoside hydrolase encoding genes, sulphur metabolism and with enrichment in secondary metabolism and an unexpectedly high number of genes with no assigned function.

Hence, for the photoreceptors, like for GNB1, GNG1 and PhLP1 their major function lies in up-regulation of their target genes in the presence of light, which is aimed at the light dependent modulation of metabolic functions of primary and secondary metabolism [46, 53].

Considering light dependent regulation patterns in wild-type and the available mutants in BLR1, BLR2, ENV1, GNB1, GNG1 and PhLP1, members of all glycoside hydrolase (GH) families—a remarkably high 75% of all GH encoding genes—show potential light dependent regulation.

cAMP signalling represents one of the most important output pathways of the heterotrimeric G-protein signaling cascade [95] and is hence the prime candidate for a coincidence detector between nutrient and light signaling. Transcriptome analysis of the major components of the cAMP pathway in *T. reesei*, adenylate cyclase (*acy1*) and protein kinase A (catalytic subunit 1, *pkac1*) revealed that their deletion causes an increase in light dependent gene regulation ([49]; unpublished data on Δ*pkac1*). ACY1 targets a broad array of genes involved in metabolism, Particularly, significant enrichment in metabolic functions was found for its positive regulatory targets in light [49], which is also the condition under which BLR1, BLR2, ENV1, GNB1, GNG1 and PhLP1 show their major function [49]. Comparable with Δ*acy1*, cAMP levels in Δ*env1* are very low [89]. Of the 136 positive targets of

ACY1 in light, 114 overlap with those of ENV1 including 25 glycoside hydrolase encoding genes, for example *xyn2, xyn4, bxl1, cel3a, cel3b, egl1,egl2, egl6* and *cbh2,* the lytic polysaccharide monooxygenase encoding *cel61a* and *cel61b* as well as swollenin, *cip1, cip2* and the major transcription factor gene *xyr1*. Additionally 16 genes involved in sulphur metabolism are in this gene set [49].

These findings on altered numbers of light responsive genes lead to the conclusion that the signaling pathways for nutrient (heterotrimeric G-protein pathway, cAMP pathway) and light response balance gene expression and this balance is perturbed if single factors are removed. Since a sizable number of the targets of the investigated signaling factors show broad functions in plant cell wall degradation it can be assumed that this balance is aimed at adjustment of substrate utilization to day and night as well as growth within and on the surface of a substrate.

Gene regulation specific for induction of cellulases is light dependent

Cellulases are expressed on diverse carbon sources, mostly representing degradation products of plant cell walls [18, 23, 96, 97]. Cellulase gene expression is not consistently regulated by light on different carbon sources [43, 65, 87, 93], indicating a different relevance of light under different conditions. Therefore, of interestis the a core gene set that is specifically regulated under inducing conditions and whether this gene set is different in light and darkness [33]. This two-dimensional analysis (5 carbon sources including inducing and repressing ones, 2 light conditions) revealed 1324 induction specific genes, comprising a set of 530 genes to be cellulase induction specific in light and in darkness, but also a considerable number of genes, that show regulation specific for cellulase induction only in light or only in darkness. The gene set of cellulase induction specific genes only in darkness includes 16 CAZyme encoding genes and comprises such important genes as *egl3* (*cel12a*), the predicted oxidoreductase TR_56840 (more than eightfold regulation in both cases), *bgl1* (*cel3a*), *egl5* (*cel45a*), *cel5b* and *bgn1* [33]. Genes of the central pathways of primary metabolism, including the pentose phosphate pathway and glycolysis turned out to be targets of light responses, with a particularly important function of ENV1. Generally, photoreceptors clearly contribute to regulation of induction specific genes both in light and darkness [33].

In several cases, specific regulation for cellulase induction [33] and positive regulation by ENV1 [46] overlaps with negative correlation of the specific protein production rate in continuous fermentations controlled for growth rate [98]. Consequently, light dependent processes and their regulation is are to be considered in fermentations as well. Of the 1324 induction specific

genes, only 218 were neither light regulated nor photoreceptor targets and this gene set does not include the known cellulase regulators, glycoside hydrolases or proteins involved in oxidative degradation of cellulose [33]. Comparison of gene regulation on the insoluble cellulase inducing carbon source cellulose with the soluble inducing carbon sources lactose and sophorose [33] revealed indications that *T. reesei* may sense surfaces, particularly in light, due to cellulose specific expression of hydrophobin genes, swollenin, *cip1* and *cip2*, which play a role in the attack on cellulose [17, 99, 100]. Because of overlapping regulation with the cAMP dependent targets of ENV1 [33, 49, 89], surface sensing may be one of the outputs of the function of ENV1.

The pathways described above signal to the cell, among other information inputs, which carbon source is available in the surroundings of the cell. In this respect it is interesting that the number of light regulated genes varies on different carbon sources, like seen with different signaling mutants. In particular, the number of genes increases on lactose, sophorose and glycerol while upon growth on glucose, the number of genes downregulated in light increases [33]. This finding explains at least in part why perturbing signaling pathways leads to altered light response: the adjustment of gene regulation according to nutrient availability does not work properly anymore.

Genome wide transcription patterns as influenced by light

Genome wide transcriptome patterns are reflect the physiological state of the organism under the tested conditions. Therefore we wanted to assess general effects on deletion of certain signaling genes and how they relate to light response on a given carbon source. For that we re-analyzed transcriptomic data under growth on cellulose in wild-type and signaling mutants in light and darkness for the whole genome and for different gene groups [33]. Generally we found that light dependent gene regulation does not break carbon source dependent regulation [49]. This finding is also in agreement with the higher relevance of the carbon source versus the light status: in other words, light modulates gene expression, but induction or repression is not initiated or abolished by light.

Hierarchical clustering of transcript levels across different carbon sources and in the absence of nutrient and light signaling genes revealed that in mutant strains, the genome wide patterns of carbon source specific regulation are largely retained. Transcripts from cultures grown on cellulose still cluster separately from those grown on soluble carbon sources. Within soluble carbon sources i.e. comparing growth on glucose, glycerol, lactose or sophorose, also patterns on inducing and repressing

carbon sources are more similar within the group than with each other. Similarly, transcript profiles from light and dark grown cultures clustered together ([33], our unpublished results). The only exception to these groupings was gene regulation in the Δenv1 strain upon growth in light: although grown on cellulose, gene expression patterns clustered with soluble inducing carbon sources, indicating a certain malfunction in carbon sensing. Accordingly, 77% of the cAMP dependent targets of ENV1 [49] overlap with those potentially involved in cellulose/surface sensing [33].

With respect to environmental sensing, the clustering of genes encoding G-protein coupled receptors indicated perturbed carbon sensing in Δenv1 and also in Δgnb1 [33]. Although deletion of env1 causes a considerable growth defect in light [42], which is at least in part dependent on altered cAMP levels [49], deletion of gnb1 does not result in strongly altered growth or morphology in light [53]. Consequently, this strong increase in light responsiveness of gene regulation is not simply due to altered growth in light and darkness. Moreover, altered gene regulation in light is correlating with altered growth and reflects a significant physiological modification in response to light and indicates different relevance of the targeted pathways in light and darkness.

Non-random distribution of light regulated genes in the genome

Genes operative in the same pathway are often clustered in a genomic region, facilitating their enhanced regulation or co-regulation for example due to an open chromatin structure in this area as exemplified in secondary metabolite clusters [101]. In the genome of T. reesei, clustering was found for CAZyme encoding genes and in several cases genes associated with secondary metabolism were present in these genomic areas as well [102]. In T. reesei, light regulated genes are clustered in 15 genomic clusters (containing 66 genes) have clusters of light regulated genes when grown on cellulose, but only one cluster for growth on sophorose and 2 for growth on glucose. On lactose and glycerol no significant clusters were found [33]. In part, the light dependent clusters found on cellulose overlap with CAZyme clusters. One cluster contains the lytic polysaccharide monooxygenase encoding cel61a gene along with two polyketide synthase genes [33] and overlaps with a cluster positively correlated with the specific protein production rate of T. reesei under constant cultivation conditions [98]. This cluster represents a secondary metabolite cluster, called the "SOR cluster" which was recently characterized [54, 103]. It is responsible for production of the yellow pigment of T. reesei [103] and regulates biosynthesis of sorbicillin compounds [104] and their derivatives trichodimerol and dihydrotrichotetronine [54]. Interestingly, this cluster overlaps with a CAZyme cluster comprising a candidate chitinase, acetylxylanesterase axe1, the auxiliary protein encoding cip1 in addition to cel61a [102].

The most interesting of the clusters with respect to cellulase gene regulation comprises cbh2/cel6a, one of the two major cellulase encoding genes of T. reesei, cel5a/egl2, a predicted sugar transporter, a putative urea transporter and an FMN dependent oxidoreductase [33]. Normal transcript levels of this cluster require the presence of BLR1, BLR2 and ENV1 in light.

Another cluster contains the mannitol dehydrogenase gene lxr1 [105]. On mannitol, growth of the wild-type in light is slightly slower than in darkness [46]. A cluster of 9 genes around lxr1 is regulated negatively by ENV1 in light on cellulose (up to 40 fold). This cluster overlaps with a light regulated cluster on cellulose [33, 46].

Light influences carbon catabolite repression

Biosynthesis of enzymes is very energy consuming and hence only initiated by the fungus if the environmental conditions require these enzymes for survival [23, 34]. This is mainly the case for insoluble plant material that cannot be taken up by the cell without liberation of low molecular compounds that serve as carbon sources. The most important process for this regulation is carbon catabolite repression, which shuts down enzyme production when easily metabolizable carbon sources are sensed in the environment.

The carbon catabolite repressor CRE1 regulates cellulase gene expression not only upon growth on glucose [106], but also on cellulose [54]. Upon growth on cellulose the regulatory targets of CRE1 are different in light and darkness and few genes were consistently regulated irrespective of the light condition, and targeted functions of CRE1 are different in light and darkness as well [54]. The genes regulated by CRE1 are non-randomly distributed in the genome and form 36 clusters, several of which contain CAZyme encoding genes. Five clusters are found, of which one comprises two CAZyme encoding genes (TR_32243/carbohydrate esterase family 1 and TR_62182/glycosyltransferase family 1) that are consistently upregulated in darkness in Δcre1 [54]. One of the clusters shows light dependent regulation by CRE1 and photoreceptors i.e. the recently characterized SOR cluster [54, 103, 107].

Of the 1324 genes with cellulase induction specific regulation, 409 are regulated by CRE1. Of those, only 14 are downregulated in Δcre1 in light and darkness and only 6 (including the hydrophobin gene hfb5) are upregulated in light and darkness [54]. Additionally, genes regulated by CRE1 are in part non-randomly distributed in the genome. Interestingly, 142 genes are upregulated in

Δ*cre1* only in light, while 251 genes are upregulated in this strain only in darkness. Consequently, the relief from carbon catabolite repression by deletion of *cre1* has also light dependent effects [54].

Among the light dependent targets of CRE1, the predicted glucose/ribitol dehydrogenase gene *grd1* was analyzed in more detail [65]. *Grd1* is positively regulated by CRE1 in darkness (more than eightfold) and its transcript size increases upon onset of inducing conditions, indicating a role of alternative splicing in its regulation. Transcript levels of *grd1* are thereby strictly correlated with those of *cbh1* and also follow the transient pattern upon sophorose induction in light [65]. GRD1 acts on cellobiose as a substrate with the likely reaction product cellobiitol, which could inhibit beta glucosidase function. With respect to cellulase regulation, GRD1 influences transcript levels, activity and protein abundance of the major cellulases in a light dependent manner. These findings suggest that GRD1 acts in intracellular sensing of cellulase efficiency and adaptation of cellulase levels for optimization of gene regulation [65].

Regulation of plant cell wall degradation by light is conserved in fungi

After the finding that cellulases are regulated by light and photoreceptors in *T. reesei*, we were interested if this is a conserved phenomenon or specific to *Trichoderma*. Transcriptome analysis of *N. crassa* upon growth on cellulose showed that regulation of cellulases by photoreceptors is conserved [92]. Lack of photoreceptors causes initially elevated secreted cellulase activity upon growth on cellulose in *N. crassa*. After prolonged cultivation, the increased levels are only maintained in Δ*vvd*, while in the *white collar* mutants lower levels occur, which is likely due their function in regulation of energy metabolism [92]. Targets of the light response machinery in *N. crassa* include numerous genes encoding plant cell wall degrading enzymes, including several lytic polysaccharide monooxygenases as well as functions in amino acid metabolism and sulphur metabolism. 55 genes are consistently regulated in one or more of the photoreceptor mutants in *N. crassa* and *T. reesei* including the glycoside hydrolase encoding *xyn1*, *xyn3*, *cbh1* and *cbh2*. However, the important transcription factor encoding genes *xyr1*, *cre1*, *clr1* and *clr2* are not among these genes [92].

In contrast to *T. reesei,* gene regulation by the white collar proteins did not overlap with that of VVD, but indications for photoadaptation (contrasting regulation between white collars and VVD) were found. Also the striking difference of gene regulation in Δ*env1* compared to wild-type in *T. reesei* was not observed in Δ*vvd* compared to wild-type in *N. crassa* in the same extent. Targets of this photoadaptation include the carbon

catabolite repressor CRE1, which is also not the case in *T. reesei*. It was particularly obvious that ribosomal genes were among those regulated by photoreceptors, which supported the hypothesis of posttranscriptional regulation of cellulase expression.

In general, this study indicates that light dependent regulation of plant cell wall degradation is a conserved process in fungi, albeit some steps of individual regulation of degradation and photoadaptation are not [92]. Exploration of light dependent gene regulation in further fungi will reveal if this conservation is indeed general.

Light alters complex formation in promoters of cellulase genes and their regulators

Regulation of cellulase gene expression at the transcriptional level was shown to involve the CAE (*cbh2* activating element) motif in the *cbh2* promoter [108] and a similar motif in the *cbh1* promoter [109]. This DNA motif is bound by protein complexes both under inducing (sophorose) and repressing (glucose) conditions [108] and induction specific chromatin rearrangement occurs at this site [110]. Thereby, the CAE (cellulase activating element) is crucial for regulation of *cbh2* and at least in part similar proteins bind to a comparable motif in *cbh1* [111].

The findings that both *cbh1* and *cbh2* are subject to regulation by light [42, 43] indicated that complex formation within these promoters may be influenced by light. Light dependent promoter binding of the white collar complex (WCC) in *N. crassa* was shown previously and complex formation of the WCC on its target promoters is influenced by light even in cell free extracts [112].

In *T. reesei* complex formation at the CAE motif of the *cbh2* promoter is clearly different between light and darkness upon growth on lactose [87]. Additionally, one of the formed complexes at CAE is dependent on the presence of the protein kinase A catalytic subunit (PKAC1) [87].

Light dependent alterations in complex formation were also shown at the EUM1 motif, which was found as a conserved motif in the *T. reesei env1* promoter and the promoter of the *N. crassa* homologue *vvd* [42, 113]. This motif is also present in the *cbh1* and *cbh2* promoters. Cellulase gene expression is modulated by the G-protein alpha subunit GNA3 in light [77]. The EUM1 was detected in the *gna3* promoter as well. In order to investigate a possible connection or feedback between *gna3* and *env1*, we analyzed complex formation at the *gna3* promoter in darkness and after illumination upon growth on glycerol [113]. Indeed, we found light dependent changes in complex formation at the EUM1 motif in the promoters of *gna3* and *env1*. The patterns of the protein complexes strongly resembled each other in the two promoters and competition experiments confirmed that

similar proteins bind to EUM1 in *gna3* and *env1* [113]. Consequently, *env1* and *gna3* are at least in part regulated by similar regulators and due to the regulation of *gna3* by ENV1 a feedback mechanism mediated by an as yet unknown transcription factor is likely.

However, the studies on complex formation at EUM1 mentioned above were done under non-inducing conditions. For growth on cellulose, complex formation also occurs within the EUM1 motifs of the *cbh2* promoter as well as on the EUM1 motifs in the *env1*, *vvd* and *gna3* promoters in a light dependent manner. Preliminary data suggest that part of the proteins binding to the EUM1 motif in these promoters also bind to the EUM1 motifs in the *env1* and in the *gna3* promoter, particularly in light (Schmoll, unpublished results).

Carbon source utilization by *Trichoderma* is altered by light and regulated by photoreceptors

Early analysis of fungal light responses indicated that the influence of light on growth of a fungus is dependent on the carbon source it grows on [114]. Thereby, a carbon source close to its natural conditions would cause the least growth defect [115].

Using Biolog phenotype microarrays allows for simultaneous testing of growth on 95 different carbon sources [116]. Our analysis showed that there are considerable differences between growth on many of these carbon sources, particularly on substrates associated with plant cell wall degradation such as D-sorbitol, L-arabinose, D-fructose, D-galactose or xylitol [46]. Additionally, the photoreceptor ENV1 caused significantly different growth patterns on many carbon sources [45]. Correlating these results with gene regulation data from transcriptome analyses showed that the pentose-phosphate pathway responsible for utilization of substrates like xylose or lactose, is a target of light response and subject to regulation by photoreceptors. At several of the enzymatic steps in this pathway, modification of growth correlated with light dependent alteration of enzyme gene transcript levels [46].

In addition, the positive regulator of cellulase gene expression, GRD1, influences light dependent growth patterns, especially on several intermediates of lactose and D-galactose carabolism, like D-xylose, D-galactose, D-fructose, D-mannitol and D-xylose. This influence on growth on intermediates of a cellulase inducing carbon source (lactose) is in agreement with the function of GRD1 in cellulase regulation and potentially inducer formation [65].

In *T. atroviride*, stimulation of growth in light was observed on many carbon sources, particularly 17 carbon sources which are well utilized by this fungus [117]. These carbon sources comprise common building blocks of plant cell walls like fructose, mannose, galactose, glucose or xylose, while growth on the best carbon source for *T. atroviride*, glycerol, is not influenced by light and growth on sorbitol even decreased in light. The increase in growth rate on several carbon sources in light is dependent on the presence of BLR1 and BLR2. Lack of these factors even causes decreased growth, albeit some exceptions on certain carbon sources for Δ*blr1* and Δ*blr2* confirm that these strains are not totally blind, but rather show altered and diminished responses [117].

The clear connection between the cAMP pathway and cellulase regulation by light and photoreceptors indicated a role of BLR1 and BLR2 in introducing the light signal into the cAMP pathway upon regulation of utilization of diverse carbon sources. In *T. atroviride* stimulation of growth by addition of cAMP was observed only on a few carbon sources (glucose, gentiobiose, cellobiose and xylose) and this response was dependent on BLR1 and BLR2. On these carbon sources, stimulation of growth was also achieved by addition of menadione that causes oxidative stress, which should simulate oxidative stress caused by light. Also for this response functional BLR1 and BLR2 are required. However, a general correlation of responses to cAMP and oxidative stress with the presence or absence of BLR1 or BLR2 across all 95 carbon sources tested could not be confirmed [117].

Consequently, light influences growth and carbon utilization of *T. atroviride* and the cAMP pathway as well as oxidative stress response are involved in this regulation mechanism. Additionally, BLR1 and BLR2 play important roles in this connection particularly on carbon sources representing building blocks of plant cell walls, hence confirming carbohydrate degradation as a major target of light response.

Crucial factors for plant cell wall degradation as targets of light response

The long history of research towards improvement of cellulase gene expression yielded a considerable number of regulators enhancing or decreasing efficiency of plant cell wall degradation [23, 31]. Transcriptome analysis showed that carbon metabolism and degradation of polysaccharides are important targets of light dependent gene regulation in *T. reesei* and *N. crassa* [53, 92]. For *N. crassa* it has been shown that the photoreceptors themselves act on regulation of further transcription factors, constituting a flat hierarchy of regulation [118]. Consequently, the transcription factors involved in plant cell wall degradation are of particular interest as targets of light response.

Therefore we screened the recently published, comprehensive list of transcription factors involved in plant biomass utilization [31] for genes differentially regulated in light and darkness or by photoreceptors. Interestingly,

we found that among them, the transcript levels of *xyr1*, *ace1*, *ace2*, *ace3*, *vib1* and *amyR* are decreased in light compared to darkness in QM6a upon growth on cellulose (Table 1). Also the *T. reesei* homologues of the important *N. crassa* transcription factors *clr-1* and *clr-2* show regulation in QM6a by light, but they do not regulate cellulase expression in *T. reesei* ([119] and our unpublished results). *Xyr1* is regulated by BLR1 and ENV1. Interestingly, the mating type protein MAT1-2-1 interacts with XYR1 and is recruited to the *cbh1* promoter in a XYR1 and light dependent manner [120]. The finding that ENV1 considerably regulates *mat-2-1* [44], the hypothesis that this light dependent effect is due to the function of photoreceptors is supported. *Ace3*, *vib1* and *amyR* are regulated by ENV1 upon growth on cellulose (Table 1; [46]).

A further important parameter for plant cell wall degradation is the hydrolysis performance as determined by the secretome of *T. reesei* [121]. This study identified 12 proteins to be limiting for hydrolysis. Ten of them were found to be regulated in a cellulase induction specific manner in light or darkness or both [33] (Table 2). But most importantly, all of the genes encoding these limiting proteins were strongly downregulated in light in QM6a upon growth on cellulose and regulated by ENV1 under the same conditions. Additionally, all those genes except *xyn4* are regulated by BLR1 and *man1*, *xyn2* and *bgl1* are regulated by BLR2 as well [46].

These findings are in perfect agreement with polysaccharide degradation being a major target of light response in *T. reesei* [53]. Moreover, these results indicate that for stable and predictable performance of strains in research and industry, light conditions and random light pulses have to be considered of equal importance as for example pH and oxygen supply and a defined cultivation medium.

Implications for research, industrial application and strain improvement

Fungi sense their environment and adjust growth and enzyme production for efficient use of resources accordingly. They use complex signal transduction pathways to integrate these environmental signals for a defined response to the conditions they perceive—in their natural environment as well as in an industrial setting. Consequently, knowledge on signal transduction pathways is highly relevant for efficient and stable gene expression [122].

In order to draw reliable conclusions on gene regulation, we have to ensure that the growth conditions of the organisms we are testing remain to same. Reproducibility of results requires defined growth media in terms of carbon source, nitrogen source, pH, temperature etc. Often, different standard media and growth conditions for different fungi are discussed as problematic for comparability of data. The research reviewed here clearly shows that light is an important environmental cue not only in *T. reesei* where light can alter regulation of cellulase genes tenfold and more. Nevertheless, even in *T. reesei* there is still research towards gene regulation and cellulase gene expression performed under random light conditions. With transcriptional regulation, it is known that even a few seconds of light cause a response in gene regulation. Considering the broad metabolic regulations by light, this compares to using a random mixture of carbon sources in an assay, which may show a beneficial effect first and no effect at all upon repetition. Moreover, we recently showed that cellulase regulation additionally involves a posttranscriptonal section, which is triggered by the light sensitive heterotrimeric G-protein pathway [33]. Consequently, good laboratory practice has to include the use of controlled light conditions in order to draw any reasonable conclusions as to the function of a gene. Especially for transcriptome studies, the effect of random light pulses—for example during harvesting or due to an open shaker—can be dramatic and lead to wrong conclusions on the function of a gene in enzyme production.

The implementation of such data from uncontrolled laboratory experiments into industrial strain improvement may hence lead to costly failures either early in testing or later in upscaling. For example, if constitutive activation of the G-protein alpha subunit GNA3 [77] would be tested for improvement of cellulase fermentations: Shake flask cultivations on table top shakers in the lab (in light) and subsequently in a glass fermentor would lead to beneficial effects on cellulase expression, because GNA3 exerts its function in light. However, upon (cost intensive) upscaling in a dark steel fermentor, the strain would not show benefits anymore, because in darkness constitutive activation of GNA3 does not have an effect on cellulase regulation [77]. Research done under random conditions bears such costly risks for surprises, because based on current knowledge the difference in the function of a given regulator in light and darkness cannot be predicted reliably. Only if part of the pathway is already known, like for the heterotrimeric G-protein pathway, some perspectives can be suggested.

Exploitation of light dependent effects

Available data show that enzyme production is most efficient in darkness in *T. reesei*. Nevertheless, the strong differences between light and darkness indicate an efficient regulatory mechanism, that—due to previous negligence of light dependent phenomena—bears the chance to identify novel regulators that can be artificially

Table 1 Light dependent regulation of transcription factors involved in plant biomass utilization

Name	Protein ID	Regulatory function (related to plant biomass degradation) in	Cellulase induction specific			Regulation in QM6a in light	Regulated by BLR1	Regulated by BLR2	Regulated by ENV1
			In light	In darkness	In light and darkness				
XYR1	122208	(hemi-)cellulose utilization	×	×	×	--	×		×
CRE1	120117	Carbon catabolite repression				n			
ACE1	75418	Cellulose utilization				-			
ACE2	78445	Cellulose utilization		×		--			
ACE3	77513	Cellulose utilization	×	×	×	---			×
CLR1	27600	(hemi-)cellulose utilization (in *N. crassa*)		×		-			×
CLR2	26163	(hemi-)cellulose utilization (in *N. crassa*)		×		--	×		×
XPP1	122879	(hemi-)cellulose utilization				n			
VIB1	54675	Cellulose utilization				--			×
HAP2	124286	CAZy regulation				n			
HAP3	121080	CAZy regulation				n			
HAP5	124301	CAZy regulation				n			
AmyR	55105	Starch utilization				--			
MalR	21997	Maltose utilization				n			
BglR	52368	Sugar sensing				n			
RhaR1	79871	L-rhamnose utilization				n			
RhaR2	121107	L-rhamnose utilization	×	×	×	n			
McmA	42249	Cellulase regulation				n			
VEL1	122284	Cellulase regulation				++			
VEL2	40551	Metabolism				n			
AreA	76817	N-assimilation				n			
AreB	120127	Nitrogen metabolite repression				n			
NmrA	74375	Nitrogen metabolite repression				n			
NIT4	76705	Nitrate pathway				n			
PAC1	120698	pH response				n			
PACX	59740	pH response				-			

For the regulatory impact of the GATA-type transcription factors BLR1 and BLR2 see text on photoreceptors

n, no regulation; +, upregulation in light; -, downregulation in light; +/-, minor regulation; ++/--, moderate regulation (around 5–10 fold); +++/---, strong regulation

Table 2 Light dependent regulation of genes encoding proteins limiting for hydrolysis of pretreated corn stover

Name	Protein ID	Function	Cellulase induction specific			Regulation in QM6a in light	Regulated by BLR1	Regulated by BLR2	Regulated by ENV1
			In light	In darkness	In light and darkness				
	69276	Xylanase	×	×	×	——	×		×
CIP1	73638	Carbohydrate binding module containing	×	×	×	——	×		×
CIP2	123940	Glucuronyl esterase				——	×		×
MAN1	56996	Beta-mannanase	×	×	×	——	×	×	×
EGL2	120312	Endoglucanase CEL5A	×	×	×	——	×		×
XYN2	123818	Xylanase	×	×	×	——	×	×	×
EGL3	123232	Endoglucanase CEL12A		×		——	×		×
SWO1	123992	Swollenin	×	×	×	——	×		×
BGL1	76672	Beta-glucosidase		×		——	×	×	×
CBH2	72567	Cellobiohydrolase CEL6A	×	×	×	——	×		×
EGL6	49081	Xyloglucanase CEL74A				——	×		×
XYN4	111849	Xylanase	×	×	×	——			×

n, no regulation; +, upregulation in light; −, downregulation in light; +/−, minor regulation; ++/−−, moderate regulation (around 5–10 fold); +++/−−−, strong regulation

misregulated for dark specific cellulase enhancement in an industrial fermentor (Fig. 2).

Since light does not override the induction process [33, 49], the use of photofermentors for enzyme production in a mass fermentation is unlikely to be cost effective. However, for special high value products like secondary metabolites, the use of light dependent promoters for precisely triggered regulation should be kept in mind. Similarly, secondary metabolism can be altered by application of light which might be beneficial for co-production or production of valuable metabolites. Additionally, the striking differences between light and darkness can be used to select light responsive promoters to enable triggering or abolishing protein production using light pulses. Exemplary studies in yeast [123] and *T. reesei* [124] have already proven that this approach is successful.

Besides the actual production process, also inoculum production is crucial for stable and high level production. Production of spores is regulated by light in fungi. Moreover, conditions of spore production are crucial for physiology of the next generation of fungi grown from it, particularly with respect to stress response. Accordingly, our own research showed that reproducible results require constant conditions for inoculum generation (precultivation on plates) prior to cultivation in liquid culture and analysis (Schmoll and Tisch, unpublished). Although a circadian rhythmicity has not yet been demonstrated for *T. reesei*, our experience strongly suggests its operation. Therefore, random light pulses, which reset the molecular circadian clock should be avoided during spore production to ensure reliable production in the cultivation thereafter. It remains to be determined whether constant light or constant darkness or light cycles of defined length during preculture can be used to increase production capacity for fermentations.

Conclusions

Since the discovery that light influences regulation of cellulases and more generally carbohydrate active enzymes, several regulation mechanisms were shown to govern this influence. In some cases, regulators only show an effect under one light status but not another. Growth, enzyme regulation, carbon utilization, carbon catabolite repression and even cross talk with secondary metabolism and sulfur metabolism were shown. Therefore, in order to obtain scientifically meaningful results on gene regulation aimed at investigation of plant cell wall degradation, application of controlled light conditions is mandatory and even short random light pulses are to be avoided.

References
1. Dunlap JC, Loros JJ. Making time: conservation of biological clocks from fungi to animals. Microbiol Spectr. 2017;5(3). https://doi.org/10.1128/microbiolspec.FUNK-0039-2016.
2. Hurley JM, Loros JJ, Dunlap JC. The circadian system as an organizer of metabolism. Fungal Genet Biol. 2016;90:39–43.
3. Tisch D, Schmoll M. Light regulation of metabolic pathways in fungi. Appl Microbiol Biotechnol. 2010;85:1259–77.
4. Schmoll M. Assessing the relevance of light for fungi implications and insights into the network of signal transmission. Adv Appl Microbiol. 2011;76:27–78.
5. Hurley JM, Loros JJ, Dunlap JC. Circadian oscillators: around the transcription-translation feedback loop and on to output. Trends Biochem Sci. 2016;41:834–46.
6. Hurley JM, Dasgupta A, Emerson JM, Zhou X, Ringelberg CS, Knabe N, Lipzen AM, Lindquist EA, Daum CG, Barry KW, et al. Analysis of clock-regulated genes in *Neurospora* reveals widespread posttranscriptional control of metabolic potential. Proc Natl Acad Sci USA. 2014;111:16995–7002.
7. Montenegro-Montero A, Canessa P, Larrondo LF. Around the fungal clock: recent advances in the molecular study of circadian clocks in *Neurospora* and other fungi. Adv Genet. 2015;92:107–84.
8. Sancar C, Sancar G, Ha N, Cesbron F, Brunner M. Dawn- and dusk-phased circadian transcription rhythms coordinate anabolic and catabolic functions in *Neurospora*. BMC Biol. 2015;13:17.
9. Baker CL, Loros JJ, Dunlap JC. The circadian clock of *Neurospora crassa*. FEMS Microbiol Rev. 2012;36:95–110.
10. Brunner M, Kaldi K. Interlocked feedback loops of the circadian clock of *Neurospora crassa*. Mol Microbiol. 2008;68:255–62.
11. Schafmeier T, Diernfellner AC. Light input and processing in the circadian clock of *Neurospora*. FEBS Lett. 2011;585:1467–73.
12. Proietto M, Bianchi MM, Ballario P, Brenna A. Epigenetic and posttranslational modifications in light signal transduction and the circadian clock in *Neurospora crassa*. Int J Mol Sci. 2015;16:15347–83.
13. Sancar C, Ha N, Yilmaz R, Tesorero R, Fisher T, Brunner M, Sancar G. Combinatorial control of light induced chromatin remodeling and gene activation in *Neurospora*. PLoS Genet. 2015;11:e1005105.
14. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. Microb Cell Fact. 2016;15:106.
15. Paloheimo M, Haarmann T, Mäkinen S, Vehmaanperä J. Production of industrial enzymes in *Trichoderma reesei*. In: Schmoll M, Dattenböck C, editors. Gene expression systems in fungi: advancements and applications. Heidelberg: Springer International; 2016. p. 23–58.
16. Bazafkan H, Tisch D, Schmoll M. Regulation of glycoside hydrolase expression in *Trichoderma*. In: Gupta VK, Schmoll M, Herrera-Estrella A, Upadhyay RS, Druzhinina I, Tuohy MG, editors. Biotechnology and biology of *Trichoderma*. Oxford: Elsevier; 2014. p. 291–307.
17. Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, et al. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem. 2003;278:31988–97.
18. Margolles-Clark E, Ilmén M, Penttilä M. Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. J Biotechnol. 1997;57:167–79.
19. Blumenthal CZ. Production of toxic metabolites in *Aspergillus niger*, *Aspergillus oryzae*, and *Trichoderma reesei*: justification of mycotoxin testing in food grade enzyme preparations derived from the three fungi. Regul Toxicol Pharmacol. 2004;39:214–28.
20. Schmoll M, Wang TF. Sexual development in *Trichoderma*. In: Wendland J, editor. The Mycota (Vol I): growth, differentiation and sexuality. Cham: Springer; 2016. p. 457–74.

21. Seidl V, Seibel C, Kubicek CP, Schmoll M. Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci USA. 2009;106:13909–14.

22. Kowalczyk JE, Benoit I, de Vries RP. Regulation of plant biomass utilization in *Aspergillus*. Adv Appl Microbiol. 2014;88:31–56.

23. Glass NL, Schmoll M, Cate JH, Coradetti S. Plant cell wall deconstruction by ascomycete fungi. Annu Rev Microbiol. 2013;67:477–98.

24. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annu Rev Phytopathol. 2014;52:427–51.

25. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucl Acids Res. 2009;37:D233–8.

26. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol. 2011;91:1477–92.

27. Bech L, Busk PK, Lange L. Cell wall degrading enzymes in *Trichoderma asperellum* grown on wheat bran. Fungal Genom Biol. 2014;4:116.

28. Beeson WT, Vu VV, Span EA, Phillips CM, Marletta MA. Cellulose degradation by polysaccharide monooxygenases. Annu Rev Biochem. 2015;84:923–46.

29. Kracher D, Scheiblbrandner S, Felice AK, Breslmayr E, Preims M, Ludwicka K, Haltrich D, Eijsink VG, Ludwig R. Extracellular electron transfer systems fuel cellulose oxidative degradation. Science. 2016;352:1098–101.

30. Cannella D, Mollers KB, Frigaard NU, Jensen PE, Bjerrum MJ, Johansen KS, Felby C. Light-driven oxidation of polysaccharides by photosynthetic pigments and a metalloenzyme. Nat Commun. 2016;7:11134.

31. Benocci T, Aguilar-Pontes MV, Zhou M, Seiboth B, de Vries RP. Regulators of plant biomass degradation in ascomycetous fungi. Biotechnol Biofuels. 2017;10:152.

32. Stricker AR, Mach RL, de Graaff LH. Regulation of transcription of cellulases- and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*). Appl Microbiol Biotechnol. 2008;78:211–20.

33. Stappler E, Dattenböck C, Tisch D, Schmoll M. Analysis of light- and carbon-specific transcriptomes implicates a class of G-protein-coupled receptors in cellulose sensing. mSphere. 2017;2(3). https://doi. org/10.1128/mSphere.00089-17.

34. Kiesenhofer D, Mach-Aigner AR, Mach RL. Understanding the mechanism of carbon catabolite repression to increase protein production in filamentous fungi. In: Schmoll M, Dattenböck D, editors. Gene expression systems in fungi: advancements and applications. Cham: Springer; 2016. p. 275–88.

35. Casas-Flores S, Herrera-Estrella A. The influence of light on the biology of *Trichoderma*. In: Mukherjee P, Horwitz BA, Singh US, Mukherjee M, Schmoll M, editors. Trichoderma—biology and applications. Wallingford: CAB International; 2013. p. 43–66.

36. Schmoll M, Esquivel-Naranjo EU, Herrera-Estrella A. *Trichoderma* in the light of day—physiology and development. Fungal Genet Biol. 2010;47:909–16.

37. Schmoll M, Dattenböck C, Carreras-Villasenor N, Mendoza-Mendoza A, Tisch D, Aleman MI, Baker SE, Brown C, Cervantes-Badillo MG, Cetz-Chel J, et al. The genomes of three uneven siblings: footprints of the lifestyles of three *Trichoderma* species. Microbiol Mol Biol Rev. 2016;80:205–327.

38. Ballario P, Macino G. White collar proteins: PASsing the light signal in Neurospora crassa. Trends Microbiol. 1997;5:458–62.

39. Casas-Flores S, Rios-Momberg M, Bibbins M, Ponce-Noyola P, Herrera-Estrella A. BLR-1 and BLR-2, key regulatory elements of photoconidiation and mycelial growth in *Trichoderma atroviride*. Microbiology. 2004;150:3561–9.

40. Heintzen C, Loros JJ, Dunlap JC. The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting. Cell. 2001;104:453–64.

41. Schmoll M. Light, stress, sex and carbon—the photoreceptor ENVOY as a central checkpoint in the physiology of *Trichoderma reesei*. Fungal Biol. 2017. https://doi.org/10.1016/j.funbio.2017.10.007.

42. Schmoll M, Franchi L, Kubicek CP. Envoy, a PAS/LOV domain protein of *Hypocrea jecorina* (Anamorph *Trichoderma reesei*), modulates cellulase gene transcription in response to light. Eukaryot Cell. 2005;4:1998–2007.

43. Castellanos F, Schmoll M, Martinez P, Tisch D, Kubicek CP, Herrera-Estrella A, Esquivel-Naranjo EU. Crucial factors of the light perception machinery and their impact on growth and cellulase gene transcription in *Trichoderma reesei*. Fungal Genet Biol. 2010;47:468–76.

44. Seibel C, Tisch D, Kubicek CP, Schmoll M. ENVOY is a major determinant in regulation of sexual development in *Hypocrea jecorina* (*Trichoderma reesei*). Eukaryot Cell. 2012;11:885–90.

45. Schuster A, Kubicek CP, Friedl MA, Druzhinina IS, Schmoll M. Impact of light on *Hypocrea jecorina* and the multiple cellular roles of ENVOY in this process. BMC Genom. 2007;8:449.

46. Tisch D, Schmoll M. Targets of light signalling in *Trichoderma reesei*. BMC Genom. 2013;14:657.

47. Sanchez-Arreguin A, Perez-Martinez AS, Herrera-Estrella A. Proteomic analysis of *Trichoderma atroviride* reveals independent roles for transcription factors BLR-1 and BLR-2 in light and darkness. Eukaryot Cell. 2012;11:30–41.

48. Bazafkan H, Beier S, Stappler E, Böhmdorfer S, Oberlerchner JT, Sulyok M, Schmoll M. SUB1 has photoreceptor dependent and independent functions in sexual development and secondary metabolism in *Trichoderma reesei*. Mol Microbiol. 2017;106(5):742–59.

49. Tisch D, Schuster A, Schmoll M. Crossroads between light response and nutrient signalling: ENV1 and PhLP1 act as mutual regulatory pair in *Trichoderma reesei*. BMC Genom. 2014;15:425.

50. Esquivel-Naranjo EU, Herrera-Estrella A. Enhanced responsiveness and sensitivity to blue light by *blr-2* overexpression in *Trichoderma atroviride*. Microbiology. 2007;153:3909–22.

51. Rodriguez-Romero J, Hedtke M, Kastner C, Muller S, Fischer R. Fungi, hidden in soil or up in the air: light makes a difference. Annu Rev Microbiol. 2010;64:585–610.

52. Chen CL, Kuo HC, Tung SY, Hsu PW, Wang CL, Seibel C, Schmoll M, Chen RS, Wang TF. Blue light acts as a double-edged sword in regulating sexual development of *Hypocrea jecorina* (*Trichoderma reesei*). PLoS ONE. 2012;7:e44969.

53. Tisch D, Kubicek CP, Schmoll M. The phosducin-like protein PhLP1 impacts regulation of glycoside hydrolases and light response in *Trichoderma reesei*. BMC Genom. 2011;12:613.

54. Monroy AA, Stappler E, Schuster A, Sulyok M, Schmoll M. A CRE1-regulated cluster is responsible for light dependent production of dihydrotrichotetronin in *Trichoderma reesei*. PLoS One. 2017;12:e0182530.

55. Stappler E, Walton JD, Schmoll M. Abundance of secreted proteins of *Trichoderma reesei* is regulated by light of different intensities. Front Microbiol. 2017;8:2586. https://doi.org/10.3389/fmicb.2017.02586.

56. Rosales-Saavedra T, Esquivel-Naranjo EU, Casas-Flores S, Martinez-Hernandez P, Ibarra-Laclette E, Cortes-Penagos C, Herrera-Estrella A. Novel light-regulated genes in *Trichoderma atroviride*: a dissection by cDNA microarrays. Microbiology. 2006;152:3305–17.

57. Esquivel-Naranjo EU, Garcia-Esquivel M, Medina-Castellanos E, Correa-Perez VA, Parra-Arriaga JL, Landeros-Jaime F, Cervantes-Chavez JA, Herrera-Estrella A. A *Trichoderma atroviride* stress-activated MAPK pathway integrates stress and light signals. Mol Microbiol. 2016;100:860–76.

58. Wang M, Zhang M, Li L, Dong Y, Jiang Y, Liu K, Zhang R, Jiang B, Niu K, Fang X. Role of *Trichoderma reesei* mitogen-activated protein kinases (MAPKs) in cellulase formation. Biotechnol Biofuels. 2017;10:99.

59. Wang M, Zhao Q, Yang J, Jiang B, Wang F, Liu K, Fang X. A mitogen-activated protein kinase Tmk3 participates in high osmolarity resistance, cell wall integrity maintenance and cellulase production regulation in *Trichoderma reesei*. PLoS ONE. 2013;8:e72189.

60. Lokhandwala J, Hopkins HC, Rodriguez-Iglesias A, Dattenböck C, Schmoll M, Zoltowski BD. Structural biochemistry of a fungal LOV domain photoreceptor reveals an evolutionarily conserved pathway integrating light and oxidative stress. Structure. 2015;23:116–25.

61. Lokhandwala J. Silverman YdIVRI, Hopkins HC, Britton CW, Rodriguez-Iglesias A, Bogomolni R, Schmoll M, Zoltowski BD: a native threonine coordinates ordered water to tune Light-Oxygen-Voltage (LOV) domain photocycle kinetics and osmotic stress signaling in *Trichoderma reesei* ENVOY. J Biol Chem. 2016;291:14839–50.

62. Schmoll M, Zeilinger S, Mach RL, Kubicek CP. Cloning of genes expressed early during cellulase induction in *Hypocrea jecorina* by a rapid subtraction hybridization approach. Fungal Genet Biol. 2004;41:877–87.

63. Malzahn E, Ciprianidis S, Kaldi K, Schafmeier T, Brunner M. Photoadaptation in *Neurospora* by competitive interaction of activating and inhibitory LOV domains. Cell. 2010;142:762–72.

64. Schwerdtfeger C, Linden H. Blue light adaptation and desensitization of light signal transduction in *Neurospora crassa*. Mol Microbiol. 2001;39:1080–7.

65. Schuster A, Kubicek CP, Schmoll M. Dehydrogenase GRD1 represents a novel component of the cellulase regulon in *Trichoderma reesei* (*Hypocrea jecorina*). Appl Environ Microbiol. 2011;77:4553–63.

66. He Q, Liu Y. Molecular mechanism of light responses in *Neurospora*: from light-induced transcription to photoadaptation. Genes Dev. 2005;19:2888–99.

67. Gyalai-Korpos M, Nagy G, Mareczky Z, Schuster A, Reczey K, Schmoll M. Relevance of the light signaling machinery for cellulase expression in *Trichoderma reesei* (*Hypocrea jecorina*). BMC Res Notes. 2010;3:330.

68. Ivanova C, Ramoni J, Aouam T, Frischmann A, Seiboth B, Baker SE, Le Crom S, Lemoine S, Margeot A, Bidard F. Genome sequencing and transcriptome analysis of *Trichoderma reesei* QM9978 strain reveals a distal chromosome translocation to be responsible for loss of *vib1* expression and loss of cellulase induction. Biotechnol Biofuels. 2017;10:209.

69. Friedl M. Global and specific effects of light on *Hypocrea/Trichoderma* (master thesis). Vienna University of Technology, Vienna; 2006.

70. Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon NJ, Keller NP, Yu JH, Braus GH. VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science. 2008;320:1504–6.

71. Karimi Aghcheh R, Nemeth Z, Atanasova L, Fekete E, Paholcsek M, Sandor E, Aquino B, Druzhinina IS, Karaffa L, Kubicek CP. The VELVET A orthologue VEL1 of *Trichoderma reesei* regulates fungal development and is essential for cellulase gene expression. PLoS ONE. 2014;9:e112799.

72. Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP. The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol. 2012;84:1150–64.

73. Bazafkan H, Dattenböck C, Böhmdorfer S, Tisch D, Stappler E, Schmoll M. Mating type dependent partner sensing as mediated by VEL1 in *Trichoderma reesei*. Mol Microbiol. 2015;96:1103–18.

74. Bazafkan H, Dattenbock C, Stappler E, Beier S, Schmoll M. Interrelationships of VEL1 and ENV1 in light response and development in *Trichoderma reesei*. PLoS ONE. 2017;12:e0175946.

75. Cabrera-Vera TM, Vanhauwe J, Thomas TO, Medkova M, Preininger A, Mazzoni MR, Hamm HE. Insights into G protein structure, function, and regulation. Endocr Rev. 2003;24:765–81.

76. Li L, Borkovich KA. GPR-4 is a predicted G-protein-coupled receptor required for carbon source-dependent asexual growth and development in *Neurospora crassa*. Eukaryot Cell. 2006;5:1287–300.

77. Schmoll M, Schuster A, do Nascimento Silva R, Kubicek CP. The G-alpha protein GNA3 of *Hypocrea jecorina* (anamorph *Trichoderma reesei*) regulates cellulase gene expression in the presence of light. Eukaryot Cell. 2009;8:410–20.

78. Seibel C, Gremel G, Silva RD, Schuster A, Kubicek CP, Schmoll M. Light-dependent roles of the G-protein subunit GNA1 of *Hypocrea jecorina* (anamorph *Trichoderma reesei*). BMC Biol. 2009;7:58.

79. Willardson BM, Howlett AC. Function of phosducin-like proteins in G protein signaling and chaperone-assisted protein folding. Cell Signal. 2007;19:2417–27.

80. Hicks JK, Bahn YS, Heitman J. Pde1 phosphodiesterase modulates cyclic AMP levels through a protein kinase A-mediated negative feedback loop in *Cryptococcus neoformans*. Eukaryot Cell. 2005;4:1971–81.

81. Wang L, Griffiths K Jr, Zhang YH, Ivey FD, Hoffman CS. *Schizosaccharomyces pombe* adenylate cyclase suppressor mutations suggest a role for cAMP phosphodiesterase regulation in feedback control of glucose/cAMP signaling. Genetics. 2005;171:1523–33.

82. Houslay MD, Adams DR. PDE4 cAMP phosphodiesterases: modular enzymes that orchestrate signalling cross-talk, desensitization and compartmentalization. Biochem J. 2003;370:1–18.

83. Kolarova N, Haplova J, Gresik M. Light-activated adenyl cyclase from *Trichoderma viride*. FEMS Microbiol Lett. 1992;72:275–8.

84. Sestak S, Farkas V. Metabolic regulation of endoglucanase synthesis in *Trichoderma reesei*: participation of cyclic AMP and glucose-6-phosphate. Can J Microbiol. 1993;39:342–7.

85. Gresik M, Kolarova N, Farkas V. Membrane potential, ATP, and cyclic AMP changes induced by light in *Trichoderma viride*. Exp Mycol. 1988;12:295–301.

86. Casas-Flores S, Rios-Momberg M, Rosales-Saavedra T, Martinez-Hernandez P, Olmedo-Monfil V, Herrera-Estrella A. Cross talk between a fungal blue-light perception system and the cyclic AMP signaling pathway. Eukaryot Cell. 2006;5:499–506.

87. Schuster A, Tisch D, Seidl-Seiboth V, Kubicek CP, Schmoll M. Roles of protein kinase A and adenylate cyclase in light-modulated cellulase regulation in *Trichoderma reesei*. Appl Environ Microbiol. 2012;78:2168–78.

88. Huang G, Chen S, Li S, Cha J, Long C, Li L, He Q, Liu Y. Protein kinase A and casein kinases mediate sequential phosphorylation events in the circadian negative feedback loop. Genes Dev. 2007;21:3283–95.

89. Tisch D, Kubicek CP, Schmoll M. New insights into the mechanism of light modulated signaling by heterotrimeric G-proteins: ENVOY acts on *gna1* and *gna3* and adjusts cAMP levels in *Trichoderma reesei* (*Hypocrea jecorina*). Fungal Genet Biol. 2011;48:631–40.

90. Peter GJ, Düring L, Ahmed A. Carbon catabolite repression regulates amino acid permeases in *Saccharomyces cerevisiae* via the TOR signaling pathway. J Biol Chem. 2006;281:5546–52.

91. Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, Hartl L, Karaffa L, Druzhinina IS, Seiboth B, et al. The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. BMC Genom. 2011;12:269.

92. Schmoll M, Tian C, Sun J, Tisch D, Glass NL. Unravelling the molecular basis for light modulated cellulase gene expression—the role of photoreceptors in *Neurospora crassa*. BMC Genom. 2012;13:127.

93. Gremel G, Dorrer M, Schmoll M. Sulphur metabolism and cellulase gene expression are connected processes in the filamentous fungus *Hypocrea jecorina* (anamorph *Trichoderma reesei*). BMC Microbiol. 2008;8:174.

94. Pakula TM, Nygren H, Barth D, Heinonen M, Castillo S, Penttilä M, Arvas M. Genome wide analysis of protein production load in *Trichoderma reesei*. Biotechnol Biofuels. 2016;9:132.

95. Lengeler KB, Davidson RC, D'Souza C, Harashima T, Shen WC, Wang P, Pan X, Waugh M, Heitman J. Signal transduction cascades regulating fungal development and virulence. Microbiol Mol Biol Rev. 2000;64:746–85.

96. Aro N, Pakula T, Penttilä M. Transcriptional regulation of plant cell wall degradation by filamentous fungi. FEMS Microbiol Rev. 2005;29:719–39.

97. Znameroski EA, Glass NL. Using a model filamentous fungus to unravel mechanisms of lignocellulose deconstruction. Biotechnol Biofuels. 2013;6:6.

98. Arvas M, Pakula T, Smit B, Rautio J, Koivistoinen H, Jouhten P, Lindfors E, Wiebe M, Penttila M, Saloheimo M. Correlation of gene expression and protein production rate—a system wide study. BMC Genom. 2011;12:616.

99. Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyyssonen E, Bhatia A, Ward M, Penttilä M. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. Eur J Biochem. 2002;269:4202–11.

100. Li XL, Spanikova S, de Vries RP, Biely P. Identification of genes encoding microbial glucuronoyl esterases. FEBS Lett. 2007;581:4029–35.

101. Gacek A, Strauss J. The chromatin code of fungal secondary metabolite gene clusters. Appl Microbiol Biotechnol. 2012;95:1389–404.

102. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol. 2008;26:553–60.

103. Derntl C, Rassinger A, Srebotnik E, Mach RL, Mach-Aigner AR. Identification of the main regulator responsible for synthesis of the typical yellow pigment produced by *Trichoderma reesei*. Appl Environ Microbiol. 2016;82:6247–57.

104. Derntl C, Kluger B, Bueschl C, Schuhmacher R, Mach RL, Mach-Aigner AR. Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. Proc Natl Acad Sci USA. 2017;114:E560–9.

105. Metz B, de Vries RP, Polak S, Seidl V, Seiboth B. The *Hypocrea jecorina* (syn. *Trichoderma reesei*) *lxr1* gene encodes a D-mannitol dehydrogenase and is not involved in L-arabinose catabolism. FEBS Lett. 2009;583:1309–13.

106. Antonieto AC, dos Santos Castro L, Silva-Rocha R, Persinoti GF, Silva RN. Defining the genome-wide role of CRE1 during carbon catabolite

repression in *Trichoderma reesei* using RNA-Seq analysis. Fungal Genet Biol. 2014;73:93–103.

107. Salo O, Guzman-Chavez F, Ries MI, Lankhorst PP, Bovenberg RA, Vreeken RJ, Driessen AJ. Identification of a polyketide synthase involved in sorbi-cillin biosynthesis by *Penicillium chrysogenum*. Appl Environ Microbiol. 2016;82:3971–8.

108. Zeilinger S, Mach RL, Kubicek CP. Two adjacent protein binding motifs in the *cbh2* (cellobiohydrolase II- encoding) promoter of the fungus *Hypocrea jecorina* (*Trichoderma reesei*) cooperate in the induction by cellulose. J Biol Chem. 1998;273:34463–71.

109. Saloheimo A, Aro N, Ilmen M, Penttila M. Isolation of the *ace1* gene encoding a Cys(2)-His(2) transcription factor involved in regulation of activity of the cellulase promoter *cbh1* of *Trichoderma reesei*. J Biol Chem. 2000;275:5817–25.

110. Zeilinger S, Schmoll M, Pail M, Mach RL, Kubicek CP. Nucleosome transactions on the *Hypocrea jecorina* (*Trichoderma reesei*) cellulase pro-moter *cbh2* associated with cellulase induction. Mol Genet Genomics. 2003;270:46–55.

111. Leonhartsberger M. Untersuchungen an einem cis-wirkenden Element der Cellulaseinduktion (master thesis, german). Vienna University of Technology, Vienna; 1999.

112. Froehlich AC, Liu Y, Loros JJ, Dunlap JC. White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. Science. 2002;297:815–9.

113. Schuster A, Schmoll M. Heterotrimeric G-protein signaling and light response: two signaling pathways coordinated for optimal adjustment to nature. Commun Integr Biol. 2009;2:308–10.

114. Chovanec P, Hudecova D, Varecka L. Vegetative growth, aging- and light-induced conidiation of *Trichoderma viride* cultivated on different carbon sources. Folia Microbiol (Praha). 2001;46:417–22.

115. Carlile MJ. The photobiology of fungi. Ann Rev Plant Physiol. 1965;16:175–202.

116. Druzhinina IS, Schmoll M, Seiboth B, Kubicek CP. Global carbon utiliza-tion profiles of wild-type, mutant, and transformant strains of *Hypocrea jecorina*. Appl Environ Microbiol. 2006;72:2126–33.

117. Friedl MA, Schmoll M, Kubicek CP, Druzhinina IS. Photostimulation of *Hypocrea atroviridis* growth occurs due to a cross-talk of carbon metabolism, blue light receptors and response to oxidative stress. Microbiology. 2008;154:1229–41.

118. Smith KM, Sancar G, Dekhang R, Sullivan CM, Li S, Tag AG, Sancar C, Bredeweg EL, Priest HD, McCormick RF, et al. Transcription factors in light and circadian clock signaling networks revealed by genome wide mapping of direct targets for *Neurospora* white collar complex. Eukaryot Cell. 2010;9:1549–56.

119. Häkkinen M, Valkonen MJ, Westerholm-Parvinen A, Aro N, Arvas M, Viti-kainen M, Penttila M, Saloheimo M, Pakula TM. Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. Biotechnol Biofuels. 2014;7:14.

120. Zheng F, Cao Y, Wang L, Lv X, Meng X, Zhang W, Chen G, Liu W. The mating type locus protein MAT1-2-1 of *Trichoderma reesei* interacts with Xyr1 and regulates cellulase gene expression in response to light. Sci Rep. 2017;7:17346.

121. Lehmann L, Ronnest NP, Jorgensen CI, Olsson L, Stocks SM, Jorgensen HS, Hobley T. Linking hydrolysis performance to *Trichoderma reesei* cel-lulolytic enzyme profile. Biotechnol Bioeng. 2016;113:1001–10.

122. Stappler E, Rodriguez-Iglesias A, Bazafkan H, Li G, Schmoll M. Relevance of signal transduction pathways for efficient gene expression in fungi. In: Schmoll M, Dattenböck C, editors. Gene expression systems in fungi: advancements and applications. Cham: Springer; 2016. p. 309–34.

123. Salinas F, Rojas V, Delgado V, Agosin E, Larrondo LF. Optogenetic switches for light-controlled gene expression in yeast. Appl Microbiol Biotechnol. 2017;101:2629–40.

124. Zhang G, Liu P, Wei W, Wang X, Wei D, Wang W. A light-switchable bidi-rectional expression system in filamentous fungus *Trichoderma reesei*. J Biotechnol. 2016;240:85–93.

Improved microscale cultivation of *Pichia pastoris* for clonal screening

Alexander Eck[1,4], Matthias Schmidt[1,4], Stefanie Hamer[2,4], Anna Joelle Ruff[2,4], Jan Förster[3,4], Ulrich Schwaneberg[2,4], Lars M. Blank[3,4], Wolfgang Wiechert[1,4,5] and Marco Oldiges[1,2,4*]

Abstract

Background: Expanding the application of technical enzymes, e.g., in industry and agriculture, commands the acceleration and cost-reduction of bioprocess development. Microplates and shake flasks are massively employed during screenings and early phases of bioprocess development, although major drawbacks such as low oxygen transfer rates are well documented. In recent years, miniaturization and parallelization of stirred and shaken bioreactor concepts have led to the development of novel microbioreactor concepts. They combine high cultivation throughput with reproducibility and scalability, and represent promising tools for bioprocess development.

Results: Parallelized microplate cultivation of the eukaryotic protein production host *Pichia pastoris* was applied effectively to support miniaturized phenotyping of clonal libraries in batch as well as fed-batch mode. By tailoring a chemically defined growth medium, we show that growth conditions are scalable from microliter to 0.8 L lab-scale bioreactor batch cultivation with different carbon sources. Thus, the set-up allows for a rapid physiological comparison and preselection of promising clones based on online data and simple offline analytics. This is exemplified by screening a clonal library of *P. pastoris* constitutively expressing AppA phytase from *Escherichia coli*. The protocol was further modified to establish carbon-limited conditions by employing enzymatic substrate-release to achieve screening conditions relevant for later protein production processes in fed-batch mode.

Conclusion: The comparison of clonal rankings under batch and fed-batch-like conditions emphasizes the necessity to perform screenings under process-relevant conditions. Increased biomass and product concentrations achieved after fed-batch microscale cultivation facilitates the selection of top producers. By reducing the demand to conduct laborious and cost-intensive lab-scale bioreactor cultivations during process development, this study will contribute to an accelerated development of protein production processes.

Keywords: *Pichia pastoris*, High throughput, Screening, Bioprocess development, Microbioreactor, Fed-batch, Phytase

Background

The methylotrophic yeast *Pichia pastoris* is frequently used for the heterologous high-level expression of proteins for industrial and pharmaceutical applications [1, 2]. Advantages of this expression host are, among others, the availability of strong promotors enabling the heterologous production of proteins up to several grams per liter, growth to high cell densities in mineral medium in fed-batch cultivation, the possibility to target proteins for secretion and the low amount of secreted endogenous proteins [3].

As recently reviewed by Ahmad et al. [4], heterologous protein production in *P. pastoris* is affected by a number of biological and process parameters. For *P. pastoris*, the most important parameters are the selection of host strain, expression vector, promotor for target gene transcription, signal peptide for protein secretion, and clonal selection after transformation, which is necessitated by a pronounced variability, e.g., due to differences in the

*Correspondence: m.oldiges@fz-juelich.de
[1] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
Full list of author information is available at the end of the article

integration locus in the host genome, the gene copy number, or the orientation of multiple expression cassettes [5–11].

Further, protein production with *P. pastoris* is also affected by several process parameters, e.g., growth medium composition [12–14], carbon source selection [15, 16], temperature [17, 18], pH value [18, 19], and the dissolved oxygen concentration (DO) [20, 21].

Unfortunately, there is no straight forward way to combine these factors to achieve optimal production of different target proteins. With the advent of new tools for the generation of biological diversity (for a review see [22]), the bottleneck in bioprocess development is currently shifted from strain engineering to the evaluation of resulting strains. This creates a need for a repeated rapid testing of a large number of parameter combinations during bioprocess development, even if statistical tools for experimental design (DoE) are applied to reduce the experimental effort [23, 24].

High cultivation throughput for microbial screenings is currently achieved by the use of microtiter plates and shake flasks [25, 26], although drawbacks of these systems for aerobic cultivation are well documented [27, 28]. This can lead to results that are not reproducible in bioreactors and the false selection of production strains [19, 29]. On the other hand, bioreactor cultivation is laborious and expensive and its use is limited to later stages of process development [25, 26]. Parallel microbioreactor systems developed in recent years have evolved to promising alternative tools for screening and process development and have been successfully used for the cultivation of industrially important microorganisms [30, 31]. As reviewed earlier [27], especially the severe effect of oxygen limitation on cellular metabolism, resulting from low maximal oxygen transfer rates in microplates and shake flasks, may prevent the success of a screening experiment. To prevent oxygen limitation, low substrate and therefore low biomass concentrations are prevailing conditions during screening and represent striking differences to the conditions applied during protein production processes with *P. pastoris*. These are usually designed as carbon-limited fed-batch processes, in which the specific growth rate is determined by the applied feeding profile. The metabolic state of the cell at excess carbon conditions (e.g., carbon catabolite repression, overflow metabolism) clearly differs from carbon-limited conditions during fed-batch, as recently shown, e.g., on transcriptional and translational level [32]. In addition, only a limited insight into microbial physiology can be gained in screening experiments if growth is not monitored.

Taken together, insufficient data and inappropriate operational conditions can lead to the false selection of production strains in screening experiments [5, 25, 27].

Conditions in screening and early stages of process development should therefore already resemble those of the final production process [25, 33, 34]. Different mini- and microbioreactor systems have been developed in recent years to bridge this gap. Either shaken or stirred parallel vessels or plates, both allowing high oxygen transfer rates and mixing of the culture broth [26, 35], are used to perform parallel aerobic batch cultivation with culture volumes between < 1 and 100 mL [36]. The endowment with non-invasive optical monitoring allows the recording of important process parameters, biomass concentration and fluorescence as well as pH and DO with sensor spots. Thereby the informative content of parallel cultivation experiments is substantially increased.

Mini- and microscale cultivation have been applied for different microorganisms and results comparable to laboratory scale bioreactor cultivations have been reported in many cases [37–40]. Thus, such devices represent promising screening tools and can reduce the number of laborious and expensive bioreactor cultivations during process development [26, 30].

Considering both the potential of this technology and the importance of *P. pastoris* for protein production processes, surprisingly few application examples for microscale cultivation of this host have been published so far. Isett et al. established glycerol batch cultivation of *P. pastoris* in 4 mL medium in a 24 well plate miniature bioreactor system and showed successful control of temperature, DO and pH at increased cell densities after repeated manual substrate addition [41]. Using Applikon's M24 miniscale cultivation device with controlled temperature, DO and pH, Holmes et al. [19] report a DoE approach to optimize methanol inducible (P_{AOX1}) GFP production. In a further study, the influence of different feeding strategies on clone screening for inducible lipase production using P_{AOX1} was investigated in a RoboLector microscale cultivation platform and the obtained clone ranking could be confirmed in laboratory bioreactor cultivation [31, 42]. However, the realization of fed-batch processes in microscale cultivation systems is still in its infancy [31, 36, 43] and batch cultivation is prevailing for large clonal screenings [44].

We here adapt a BioLector microbioreactor system [39] for parallel cultivation of *P. pastoris*. 48-well baffled FlowerPlates used for cultivation ensure high oxygen transfer rates by high-frequency shaking. Online process parameters are recorded non-invasively by means of optical detection of scattered light as biomass-dependent signal and of online fluorescence for pH and DO using integrated optodes [39]. The aim of this study is to enable a facile and reliable screening of *P. pastoris* clonal libraries for protein production using the GAP-promotor. We therefore modified a high cell density cultivation medium

for the parallel cultivation of *P. pastoris* in the BioLector system with different carbon sources and investigated scalability to laboratory bioreactor cultivation in terms of process characteristics like yields and growth rates. The application of this powerful set-up is applied to screen a *P. pastoris* library constitutively expressing *Escherichia coli* AppA phytase for the best production strain under different process conditions. Finally, fed-batch-like process conditions were realized in microscale cultivation to allow screening under process-relevant conditions, i.e. increased biomass and product concentrations.

Results

Adjustment of medium composition for microscale cultivation

To accelerate bioprocess development for the eukaryotic protein production host *P. pastoris*, the aim of this study was to increase the cultivation throughput and to enable screening under process-relevant, defined conditions. Therefore, the BioLector microbioreactor system was adapted for miniaturized parallel cultivation of *P. pastoris*. In a first step, a basal salts medium (BSM) suitable for high cell density bioreactor cultivation was selected

[45]. Customization of the medium for microscale cultivation was necessary since it does not contain a sufficient pH buffer, which is not required during bioreactor cultivation with controlled pH. Further, nitrogen is usually supplied partly by the use of ammonia, which is not possible in the BioLector. Therefore, 75.7 mM $(NH_4)_2SO_4$ and 150 mM PIPPS pH 5.0 were added as nitrogen source and buffer, respectively. 4% D-glucose was used as carbon source. However, precipitation of medium components was observed at the standard cultivation pH 5.0, which caused a high background of the scattered light signal in the BioLector and thus interfered with online biomass detection (Fig. 1a). Therefore, further adjustment of the medium composition was necessary to meet the specific demands of BioLector microscale cultivation. Comparing the medium composition to growth media for other organisms, the iron concentration in BSM is remarkably high (0.92 mM $FeSO_4$). In order to prevent the likely precipitation of iron salts in the medium, a reduction of the $FeSO_4$ concentration and its influence on cellular growth was tested. When reducing the initial concentration to 0.46 mM (2-fold reduction), 92 µM (10-fold reduction), or 9.2 µM (100-fold reduction, Fig. 1b–d), the backscatter
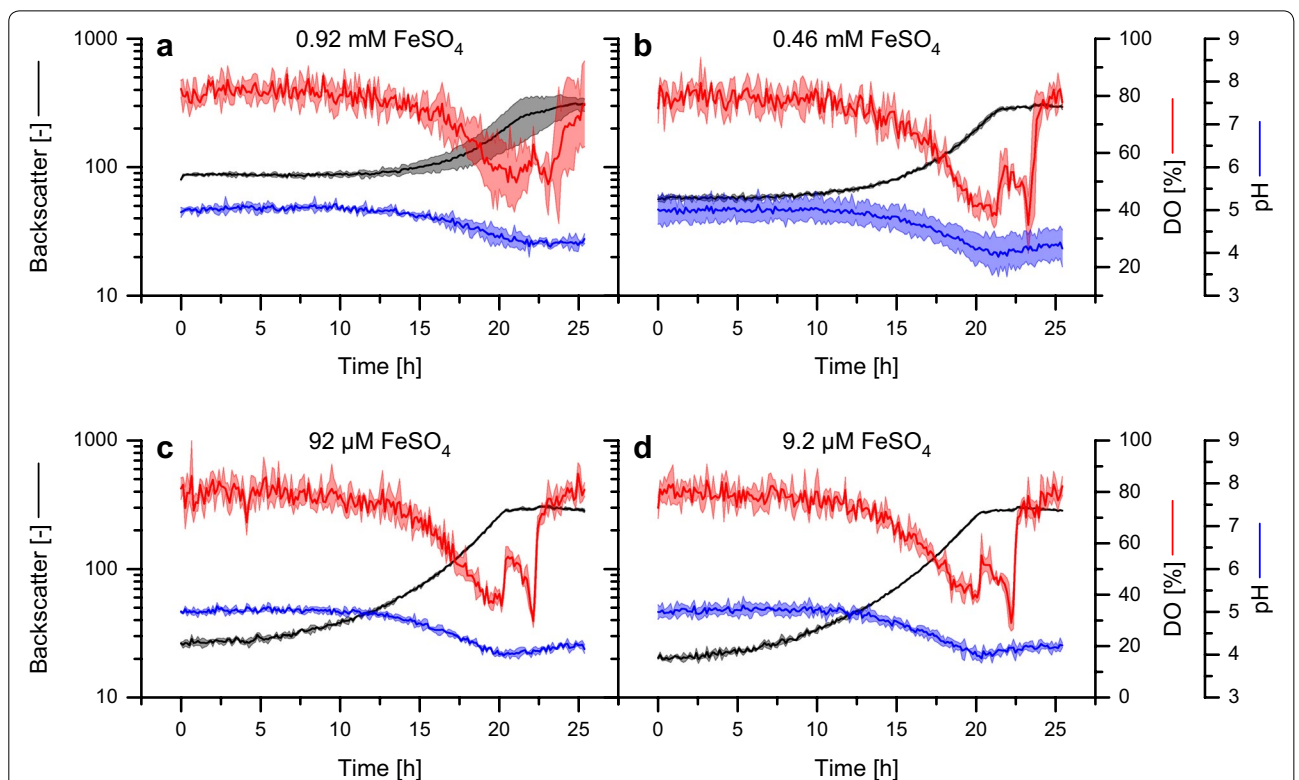

Fig. 1 Reduction of the iron concentration in BSM prevents precipitation. The effect of reduced $FeSO_4$ concentrations on medium stability and growth of *P. pastoris* in BSM was investigated in microscale cultivation (4% D-glucose, 150 mM PIPPS, pH 5.0, 0.8 mL, 1500 rpm, 30 °C). Different media were inoculated from the same pre-culture to an initial $OD_{600} = 0.3$ to ensure equal starting conditions. **a** 0.92 mM $FeSO_4$ (original concentration); **b** 0.46 mM $FeSO_4$; **c** 92 µM $FeSO_4$; **d** 9.2 µM $FeSO_4$. Light colors around solid lines show standard deviations from at least three individual wells
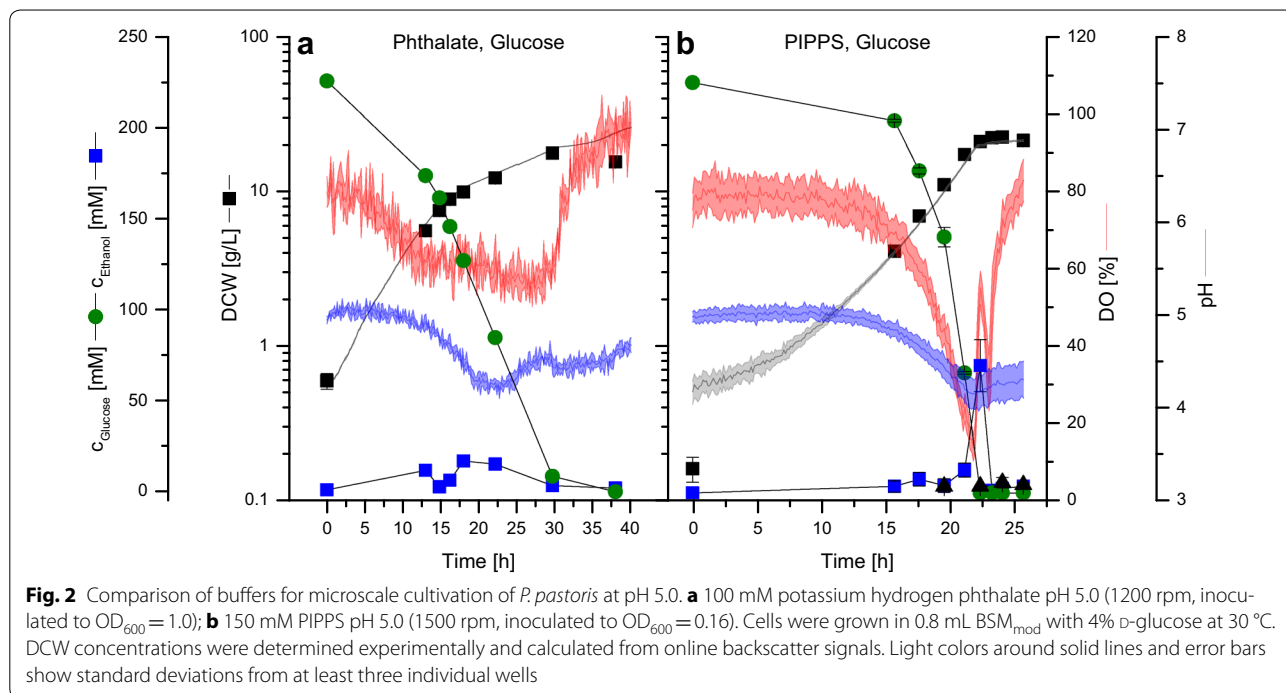
signal at the beginning of the cultivation, in case of the lowest concentration tested, decreased to a value similar to the backscatter of medium without $FeSO_4$, thus showing that iron ions were the reason for precipitation which was avoided by a 100-fold reduction of $FeSO_4$.

The online data for biomass, pH and DO (Fig. 1a–d) represent typical growth curves for batch cultivation of *P. pastoris*. After a lag phase, during which the backscatter, pH and DO remained constant for several hours, the increasing backscatter and the falling pH and DO reflect cellular growth. After the end of the growth phase, a second decrease of the DO signal occurred at t ~ 20–22 h, indicating the metabolization of by-products, which accumulated during batch cultivation. In accordance with a reduction in batch times, specific growth rates calculated from the backscatter signal (see methods) during the exponential growth phase slightly increased with the reduction of the $FeSO_4$ concentration. While the specific growth rate was 0.18 ± 0.01 h^{-1} for 0.92 mM, it increased significantly (p < 0.05) to 0.22 ± 0.00 h^{-1} (0.46 mM), 0.26 ± 0.01 h^{-1} (92 μM) and 0.24 ± 0.01 h^{-1} (9.2 μM), respectively. The final backscatter was comparable for all conditions. Thus, the reduction of the $FeSO_4$ content did not reduce the biomass yield and even increased the specific growth rate, showing sufficient supply of iron with no growth phenotype. A concentration of 9.2 μM $FeSO_4$ in BSM (BSM$_{mod}$) was used for all further cultivations.

A requirement for microscale cultivation in the BioLector is the addition of a buffer to the medium since the pH is not maintained otherwise. The choice of a buffer is crucial and can strongly influence the results of a cultivation experiment. We therefore selected two different buffers at pH 5.0. PIPPS was chosen because of its pK_a value ($pK_a = 3.7$). It is a tertiary amine sulfonic acid not showing weak organic acid characteristics and is thus assumed to be an inert buffer. Potassium hydrogen phthalate ($pK_{a,1} = 5.4$, $pK_{a,2} = 3.0$) is a common buffer for the cultivation of *P. pastoris* at low pH [45, 46]. Figure 2 shows growth curves for *P. pastoris* in BSM$_{mod}$ containing 4% D-glucose and either 100 mM potassium hydrogen phthalate or 150 mM PIPPS at pH 5.0 in the BioLector. Online data for dry cell weight (DCW) calculated from the backscatter signal with the help of a calibration function (Additional file 1: Figure S1), pH, DO and experimental values for DCW, substrate and by-product concentrations are also shown. For the cultivation with potassium hydrogen phthalate a biphasic growth was observed. The specific growth rate dropped from 0.17 ± 0.0 to 0.06 ± 0.0 h^{-1} after 15 h.

In the presence of PIPPS the DCW values reveal exponential growth with a significantly increased (p < 0.05) specific rate of 0.23 ± 0.02 h^{-1}. This is confirmed by the monotone decline of both the DO to a minimum of 14% and of the pH to 4.16 ± 0.16, which means no significant reduction compared to potassium hydrogen phthalate. Subsequently, in a second, short phase at t ~ 21–23 h the metabolization of ethanol, which was produced as by-product in the late

Fig. 2 Comparison of buffers for microscale cultivation of *P. pastoris* at pH 5.0. **a** 100 mM potassium hydrogen phthalate pH 5.0 (1200 rpm, inoculated to OD$_{600}$ = 1.0); **b** 150 mM PIPPS pH 5.0 (1500 rpm, inoculated to OD$_{600}$ = 0.16). Cells were grown in 0.8 mL BSM$_{mod}$ with 4% D-glucose at 30 °C. DCW concentrations were determined experimentally and calculated from online backscatter signals. Light colors around solid lines and error bars show standard deviations from at least three individual wells

exponential growth phase with maximal concentration of 70 ± 14 mM, was observed.

When comparing the results of cultivations with either PIPPS or potassium hydrogen phthalate as buffer, it is obvious that the latter exerted a negative effect on growth of *P. pastoris*. Therefore, all further cultivations in the BioLector were carried out using PIPPS.

Scalability of microscale to laboratory bioreactor cultivation

Cultivations in laboratory bioreactors were carried out to evaluate the scalability of the microscale cultivation set-up for *P. pastoris*. Cells were grown in a stirred tank reactor in 0.8 L BSM_{mod} with controlled pH (5.0) and DO (30%), respectively, representing an up-scale of the culture volume by a factor of thousand. Two different carbon sources, D-glucose and glycerol, were tested. Parameters for growth of *P. pastoris*::pGAPZαB_*appA* in the BioLector and the bioreactor are summarized in Table 1. Growth on 4% glycerol in the BioLector and bioreactor was exponential in both cases (Fig. 3). While the DO was controlled at 30% in the bioreactor, it steadily decreased during cultivation in the BioLector and was zero for the last 1.5 h. Nevertheless, no significant differences were measured for both the specific growth rate and the biomass yield, which were 0.23 ± 0.01 h^{-1} (bioreactor) and 0.24 ± 0.01 h^{-1} (BioLector), and 0.60 ± 0.02 g/g (both systems), respectively. No by-products were detected via HPLC-RI analysis in the culture supernatants in both cultivation setups.

In contrast, the biomass yield for 4% D-glucose in the BioLector (0.48 ± 0.03 g/g, Table 1) was increased compared to the bioreactor cultivation (0.30 ± 0.02 g/g, Table 1), while the specific growth rate remained unchanged. The analysis of culture supernatants revealed differences in by-product accumulation with D-glucose as substrate between the two systems. In the BioLector, a transient maximal accumulation of 70 ± 14 mM ethanol was the only detectable by-product during aerobic growth (Fig. 2b), while the ethanol titer during bioreactor cultivation peaked at 141 ± 39 mM. In the latter case, the ethanol peak was followed by a transient acetate accumulation, which was then further metabolized (Additional file 1: Figure S2).

Table 1 Specific growth rates and biomass yields for bioreactor and BioLector cultivations of *P. pastoris*

	D-Glucose		Glycerol	
	Bioreactor	**BioLector**	**Bioreactor**	**BioLector**
μ (h^{-1})	0.22 ± 0.01	0.23 ± 0.02	0.23 ± 0.01	0.24 ± 0.01
$Y_{x/s}$ (g/g)	0.30 ± 0.02	0.48 ± 0.03	0.60 ± 0.02	0.60 ± 0.02

4% D-glucose or 4% glycerol were used as substrate in BSM_{mod} pH 5.0 (BioLector: 150 mM PIPPS, 0.8 mL, 1500 rpm, 30 °C; bioreactor: 0.8 L, pH = 5.0 (NH_4OH/H_2SO_4), DO = 30%, 30 °C). Mean values ± standard deviations of at least 3 cultivations are shown

In conclusion, microscale cultivation of *P. pastoris* was successfully established using a modified high cell density growth medium. Reduction of the $FeSO_4$ content was necessary to prevent precipitation of media components. PIPPS was chosen over potassium hydrogen phthalate for growth at pH 5.0. Growth parameters with glycerol as carbon source were in good agreement with results from laboratory-scale bioreactor cultivations, providing evidence that comparable results were achieved in both systems. However, cultivations with D-glucose point out important differences between both systems. After validation, the established microscale cultivation protocol for *P. pastoris* was applied to investigate the influence of clonal variability on protein production, both under batch and fed-batch-like conditions, and the results are described in the following section.

Application of microscale cultivation for screening of clonal libraries under batch and fed-batch conditions
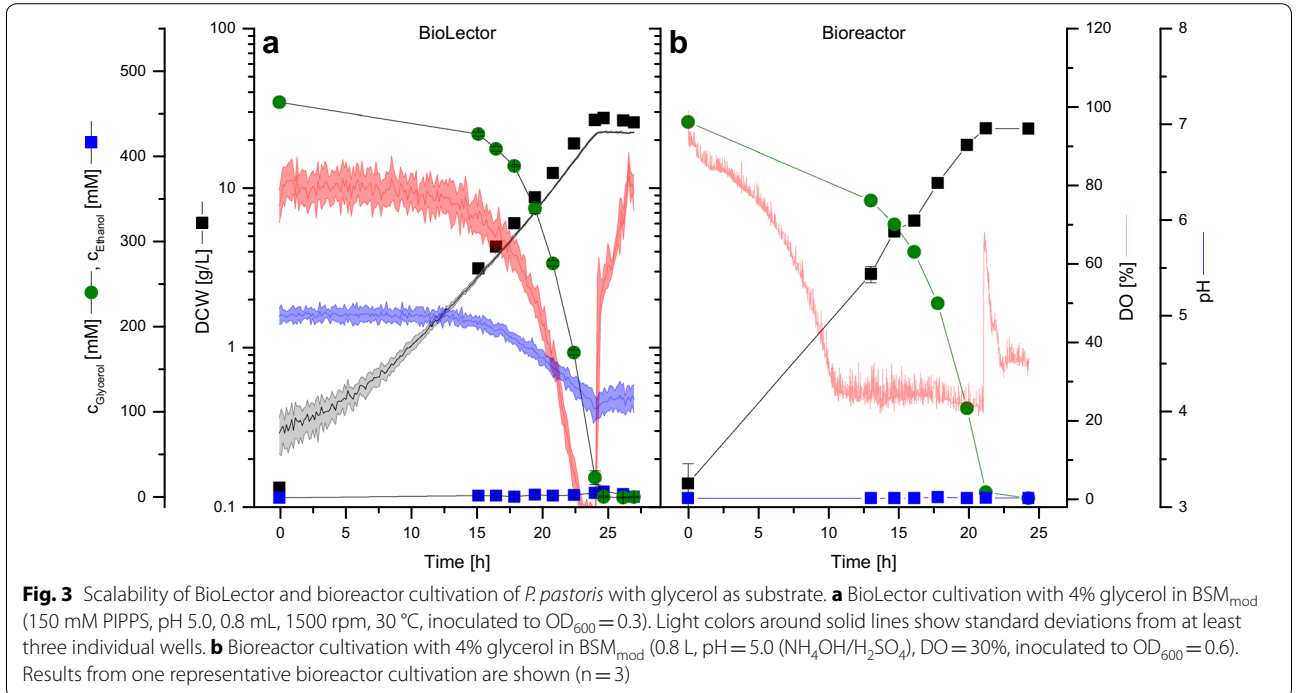### Initial selection of phytase secreting clones under batch conditions

Microscale cultivation in the BioLector was applied to screen a clonal library of *P. pastoris*::pGAPZαB_*appA* constitutively expressing and secreting AppA phytase from *E. coli* under control of P_{GAP}. 47 individual clones resulting from the transformation of *P. pastoris* X-33 with the plasmid pGAPZαB_*appA*, previously linearized to facilitate chromosomal integration of the expression cassette, were cultivated in triplicate in the BioLector with 4% D-glucose as substrate to compare growth of individual clones and to identify the best phytase secreting ones.

Online backscatter signals for all clones were recorded during cultivation and converted to dry cell weight concentrations (Additional file 1: Figure S3 A). The resulting biomass data were used to calculate specific growth rates while biomass yields were determined experimentally at the end of the cultivation. Results for individual clones are shown in Additional file 1: Figures S4 and S5. Specific growth rates ranged from 0.22 to 0.26 h^{-1}, while the mean biomass yield was 0.44 ± 0.02 g/g. A significant deviation ($p < 0.05$) of the specific growth rate from the parental strain was observed for 11 clones and for 9 clones with regard to the biomass yield, as determined by ANOVA and multiple comparison of means.

The online DO signal revealed two distinct phases during cultivation for all clones (Additional file 1: Figure S3 B), indicating transient accumulation of ethanol as a by-product in the presence of excess D-glucose. pH profiles for all clones were highly similar with the value decreasing steadily to a minimum of approximately 4.1 (Additional file 1: Figure S3 C).

The purpose of this screening was the identification of transformants secreting phytase to the culture

Fig. 3 Scalability of BioLector and bioreactor cultivation of *P. pastoris* with glycerol as substrate. **a** BioLector cultivation with 4% glycerol in BSM_mod (150 mM PIPPS, pH 5.0, 0.8 mL, 1500 rpm, 30 °C, inoculated to OD_600 = 0.3). Light colors around solid lines show standard deviations from at least three individual wells. **b** Bioreactor cultivation with 4% glycerol in BSM_mod (0.8 L, pH = 5.0 (NH₄OH/H₂SO₄), DO = 30%, inoculated to OD_600 = 0.6). Results from one representative bioreactor cultivation are shown (n = 3)

supernatant. Therefore, total protein concentrations in the supernatants collected after complete substrate and by-product consumption were initially used to identify protein secreting clones. On average, the supernatants of all tested transformants contained 11.5 ± 2.7 mg/L protein (Fig. 4a). As confirmed by ANOVA and multiple comparison of means, only clones 12 (20.3 ± 1.4 mg/L) and 18 (21.8 ± 0.8 mg/L) produced protein titers significantly ($p < 0.05$) exceeding the level of the parental strain (10.8 ± 0.3 mg/L). For these two as well as for clones 1, 23, 27, and 30 phytase activities were determined (Fig. 4b). The parental strain was used as negative control. Activities were highest for clone 18 (31.5 ± 1.0 U/L) and second highest for clone 12 (9.2 ± 1.6 U/L). Differences to the parental strain were found to be significant ($p < 0.05$) for all tested clones but clones 1, 23, 27 and 30 showed only very low phytase activities. Notably, values for protein concentrations in the supernatant do not fully correlate to phytase activities, since clone 12 showed protein concentrations similar to clone 18 but less than one third of its phytase activity.
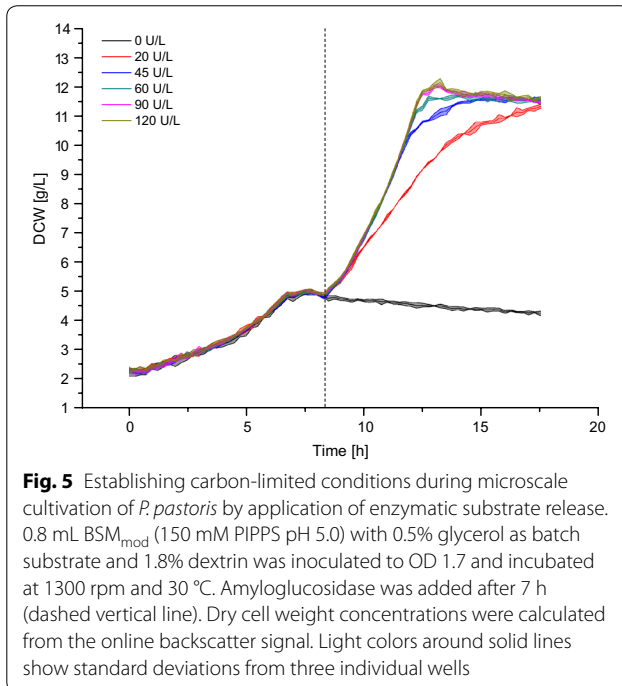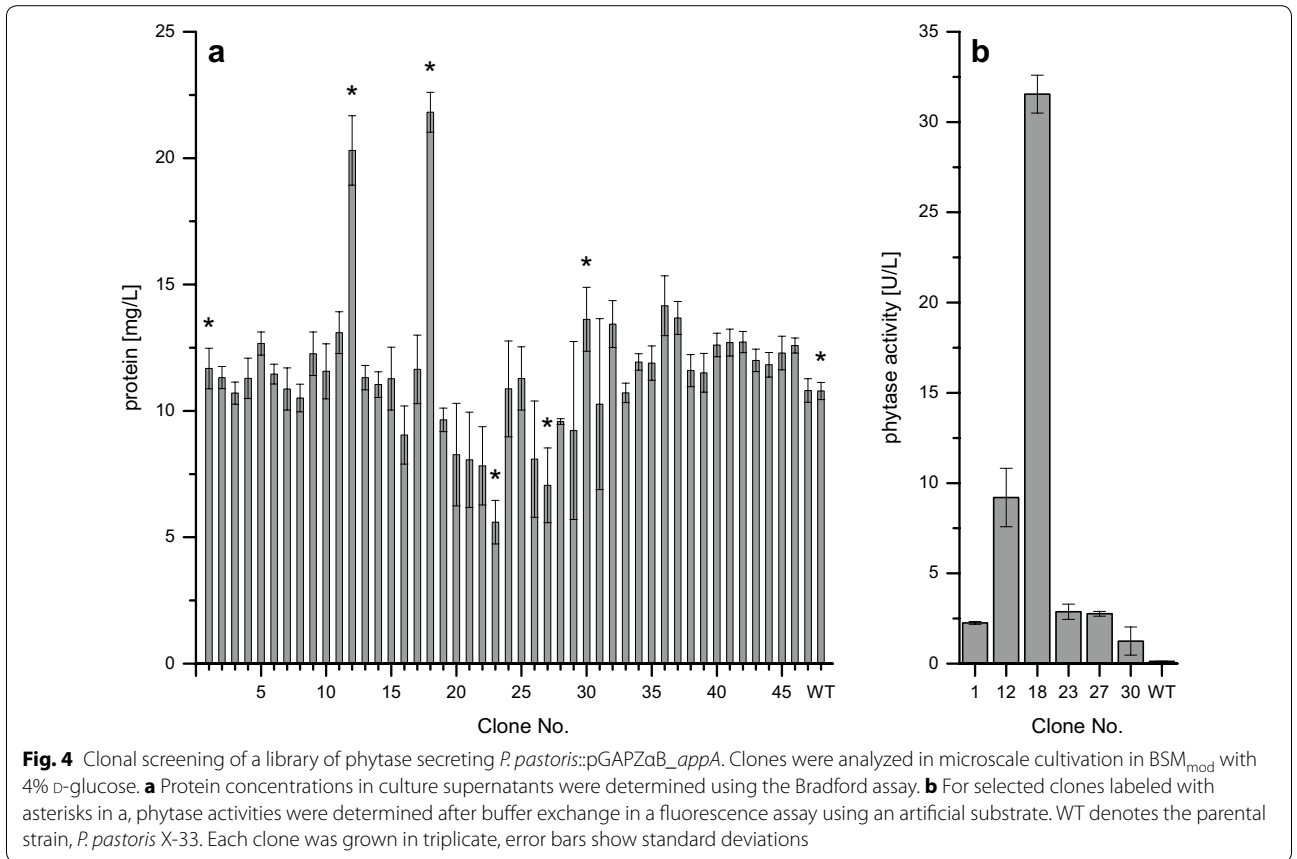
This example points out the feasibility of screening of clonal libraries applying microscale cultivation of *P. pastoris*::pGAPZαB_*appA* in the BioLector. Best performing clones in terms of phytase secretion were identified and additional information on physiological parameters for individual clones, and qualitative information on substrate consumption and by-product formation was obtained.

In a next step, the microscale cultivation protocol was extended to enable clonal screenings under fed-batch-like conditions.

Adjusting carbon-limited conditions during microscale cultivation enables screening under process-relevant fed-batch conditions

Limiting the specific growth rate by controlled feeding of substrate allows taking control of metabolic rates during fed-batch cultivation.

We enabled fed-batch-like conditions for microscale cultivation of *P. pastoris* by the use of dextrin as a non-metabolizable glucose polymer and highly stable amyloglucosidase from *Aspergillus niger* for enzymatic glucose release. In a first experiment, different volumetric amyloglucosidase activities were tested to identify the boundaries for carbon-limited growth. BSM_mod was supplemented with 0.5% glycerol and 1.8% dextrin. After consumption of the batch substrate after 7 h (5 g/L DCW), 0–120 U/L amyloglucosidase was added to provide glucose monomers as additional carbon source at a constant rate (Fig. 5). A constant glucose release rate leads to linear growth and a decrease of the specific growth rate over time. After enzyme addition, increasing biomass concentrations were observed except for the negative control, confirming that *P. pastoris* does not thrive on dextrin as substrate. For high amyloglucosidase activities from 60 to 120 U/L (12–24 U/g DCW), growth curves are nearly identical and show exponential and thus

Fig. 4 Clonal screening of a library of phytase secreting *P. pastoris*::pGAPZαB_*appA*. Clones were analyzed in microscale cultivation in BSM$_{mod}$ with 4% D-glucose. **a** Protein concentrations in culture supernatants were determined using the Bradford assay. **b** For selected clones labeled with asterisks in a, phytase activities were determined after buffer exchange in a fluorescence assay using an artificial substrate. WT denotes the parental strain, *P. pastoris* X-33. Each clone was grown in triplicate, error bars show standard deviations

Fig. 5 Establishing carbon-limited conditions during microscale cultivation of *P. pastoris* by application of enzymatic substrate release. 0.8 mL BSM$_{mod}$ (150 mM PIPPS pH 5.0) with 0.5% glycerol as batch substrate and 1.8% dextrin was inoculated to OD 1.7 and incubated at 1300 rpm and 30 ℃. Amyloglucosidase was added after 7 h (dashed vertical line). Dry cell weight concentrations were calculated from the online backscatter signal. Light colors around solid lines show standard deviations from three individual wells

non-limited growth. At 45 U/L (9 U/g DCW), growth started exponentially but became substrate-limited with increasing biomass concentration. In contrast, a linear increase of the DCW as expected for carbon limitation was observed after the addition of 20 U/L amyloglucosidase (4 U/g DCW). Thus, the specific amyloglucosidase activity of 4 U/g DCW was regarded as the upper limit of the specific glucose release rate in order to establish carbon-limited conditions in further cultivations.

Next, clones of *P. pastoris*::pGAPZαB_*appA* selected in batch screening (clones no. 1, 12, 18, 23, 27, 30, compare Fig. 4) were re-screened using the presented microscale fed-batch approach. Therefore, BSM$_{mod}$ was supplemented with 2% glycerol and 10% dextrin as initial batch substrate and as substrate for enzymatic glucose release, respectively. After consumption of glycerol, 2.6 U/g DCW amyloglucosidase was added (25 U/L). Since this specific glucose release rate is below the defined upper limit of 4 U/g DCW, carbon limited growth conditions were ensured in the experiment. Addition of 25 U/L amyloglucosidase was repeated after 41.8 h (2.1 U/g

DCW) and 66.0 h (1.8 U/g DCW) to re-adjust the glucose release rate in order to prevent a strong glucose limitation for the increasing biomass concentration.

Time profiles of the online values for biomass, DO and pH (Fig. 6a–c) are nearly identical for all tested clones and the parental strain. Final dry cell weight concentrations were determined experimentally after 72 h and reached between 44.9 and 46.4 g/L. The exponentially increasing biomass concentrations indicate unlimited cell growth during the batch phase with a specific growth rate of 0.20 ± 0.00 h^{-1}. At the same time, the DO decreased constantly and shortly dropped below 10% due to the high biomass concentrations at the end of the batch phase, before a DO peak reflected the consumption of the substrate. This oxygen limited phase could be avoided with lower batch substrate concentration. However, the likely short time by-product accumulation is expected to be rapidly consumed in the following carbon-limited phase. Then, the fed-batch phase was started by the addition of amyloglucosidase, resulting in a linear increase of the biomass signal over time. As expected, repeated addition of amyloglucosidase after 41.8 and 66.0 h lead to an accelerated but still limited growth due to the increase of the glucose release rate. Since the addition was performed offline under a sterile work bench, which took approximately 5 min, it caused short phases of oxygen limitation, but DO values recovered quickly. pH values decreased from 5.2 to 4.6 during batch cultivation but re-increased to > 5.0 at the end of the experiment.

The clonal ranking according to final protein titers and phytase activities (Fig. 6d) of the fed-batch screening confirmed batch results with one exception. In contrast to batch screening, protein titers and phytase activities were similar within the experimental error for clones 12 and 18 under substrate-limited conditions (153 ± 11 and 173 ± 18 mg/L, and 652 ± 43 and 630 ± 60 U/L). The background level of 59 ± 13 mg/L protein measured for the parental strain approximately corresponds to the amount of added amyloglucosidase, and no phytase activity was detected in this case. Clones 1, 23, 27, and 30 produced intermediate protein titers and phytase activities at similar levels (111 ± 12 to 128 ± 21 mg/L and
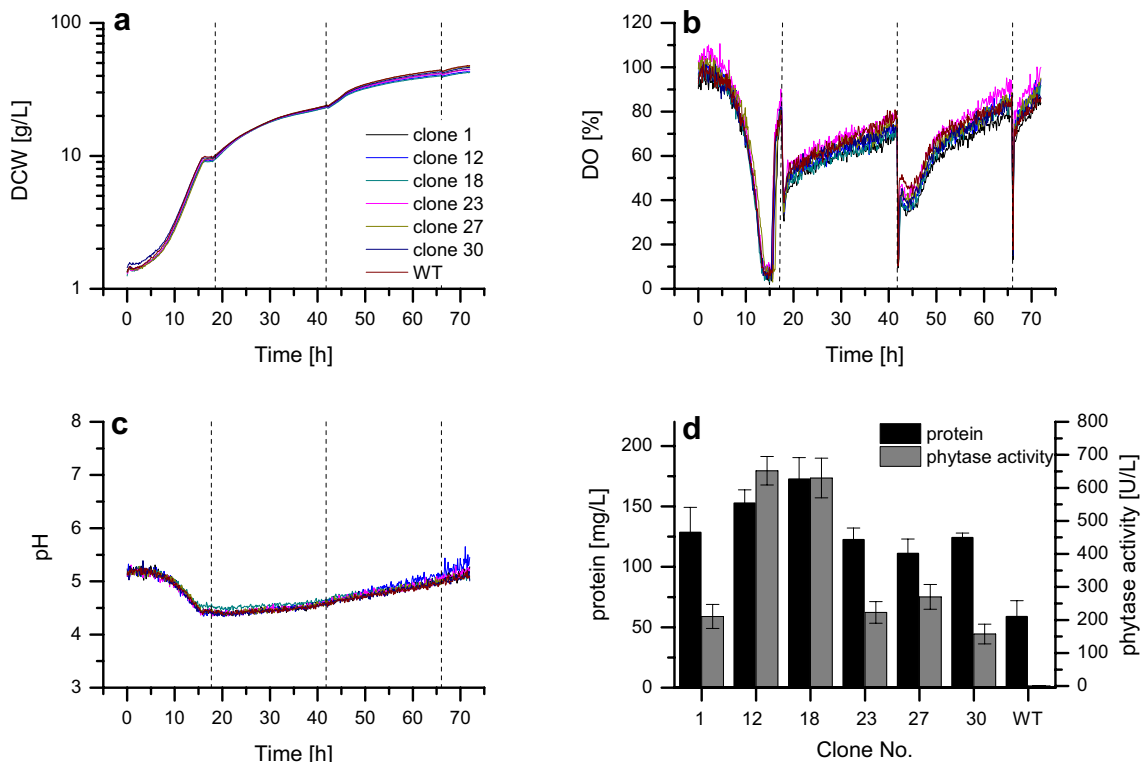

Fig. 6 Microscale cultivation of phytase secreting *P. pastoris*::pGAPZαB_*appA* under carbon-limited conditions. Clones selected in a batch-screening (compare Fig. 5) were analyzed in microscale cultivation in 0.8 mL BSM$_{mod}$ at pH 5.0 with 2% glycerol as batch substrate and 10% dextrin at 1300 rpm and 30 °C. 25 U/L amyloglucosidase was added after 17.7, 41.8, and 66.0 h (dashed vertical lines). **a** Dry cell weight concentrations calculated from the online backscatter signal; **b** Online measurement of dissolved oxygen concentrations; **c** Online measurement of pH; **d** Protein concentrations and phytase activities in culture supernatants. All values are corrected for evaporation. WT denotes the parental strain, *P. pastoris* X-33. Each clone was grown in quadruplicate, lines show mean values and error bars in **d** denote standard deviations

157 ± 30 to 270 ± 37 U/L, respectively). All culture supernatants were also analyzed via SDS-PAGE (Additional file 1: Figure S6). A protein band at 95 kDa visible after Coomassie staining represents amyloglucosidase, and one additional band is observed at 55 kDa. The native size of AppA phytase is 45 kDa, giving rise to the conclusion that AppA was secreted in non-native form and had been glycosylated by the posttranslational modification machinery in *P. pastoris* as observed earlier [47].

In conclusion, screening under batch- and substrate-limited conditions resulted in different clonal rankings since the performance of the second-best producer under batch conditions was identical to that of the best producer under carbon-limited conditions. Most strikingly, product titers were markedly increased after fed-batch cultivation. After correction for the background protein concentration of the parental strain the protein titer and phytase activity for clone 18 was increased 10-and 20-fold, respectively. Although batch-screening reliably distinguishes clones with low and increased protein secretion in shorter cultivation times, higher product titers produced during screening under carbon-limited conditions obviously facilitate the selection of best performing clones. The low increase of protein titers even for best phytase producers compared to the background level caused by the secretion of host proteins (parental strain) in batch cultivation presents a possible source of error. Therefore, screening under fed-batch conditions has the advantage of resulting in more reliable and process-relevant results, especially for proteins that are produced at low levels in *P. pastoris* only.

Discussion

In this study, we applied the 48 well plate based BioLector microbioreactor system [39] for the cultivation of *P. pastoris*, which is an important protein production host. By tailoring a defined high cell density medium for batch and carbon-limited parallel growth experiments, we show that process-relevant and scalable screening conditions can be achieved. Although simple in its application, successfully setting up microscale BioLector cultivation for *P. pastoris* required to consider several aspects. The selection of a suitable pH buffer is a crucial requirement since the BioLector does not allow pH control. *P. pastoris* grows at a broad pH range between 3 and 7 [48], but changes in pH alter the specific growth rate, can activate host proteases [49], and were shown to affect GFP production [19]. We thus compared growth at pH 5.0 using two different buffers, PIPPS and potassium hydrogen phthalate. It was obvious that phthalate severely inhibits growth. This was probably attributed to an uncoupling of the proton motive force due to its weak organic acid properties combined with aromatic carbon backbone,

facilitating membrane diffusion. Additionally, a biochemical inhibition of glycolytic enzymes from *Saccharomyces cerevisiae* by phthalic acid has been noted earlier in vitro [50]. In contrast, no growth inhibition after the addition of PIPPS was observed.

Precipitation of media components is a frequent drawback of high cell density growth media. This has been recognized as a problem potentially leading to nutrient starvation and interfering with optical process analyses [26, 51]. Reliable process information via optical online detection of biomass, pH, and DO in the BioLector was mandatory for the success of this study, which made stabilizing the medium by adjusting its composition a prime task. It was found that reducing the original $FeSO_4$ concentration (0.92 mM) by a factor of 100 prevented precipitation (Fig. 1). High iron ion concentrations above 1 mM are also found in other *Pichia* media described in literature [52, 53] and possibly are needed to support growth to very high biomass concentrations in fed-batch cultivation. In a study on medium optimization, Isidro et al. report that iron ions inhibit metabolic activity of *P. pastoris* even after 10-fold reduction to 0.1 mM [54]. In our experiments, a slight increase of the specific growth rate was measured upon $FeSO_4$ reduction to 9.2 μM. Further, the biomass yield was not altered up to the maximum tested substrate concentration of 4%, thus showing that cells did not face iron limitation at reduced $FeSO_4$ concentrations.

When testing the scalability of physiological parameters for *P. pastoris* between BioLector and laboratory scale bioreactor cultivations, specific growth rates and biomass yields were in good agreement for glycerol. Surprisingly, growth yields on D-glucose were lower for bioreactor cultivations. This effect was observed reproducibly in biological triplicate bioreactor cultivation. Main differences between the two cultivation systems are the absence of pH and DO control in the BioLector as well as in aeration. As a consequence of the high and constant oxygen transfer rate at the adjusted setting of the device (110 mM/h as specified by the manufacturer), cells were exposed to high DO levels in the BioLector for the major part of the cultivation. According to earlier observations [15], the lower and constant DO setpoint of 30% in the bioreactor probably accounted for the formation of ethanol and acetate as by-products and the reduced biomass yield, pointing towards the necessity of increased DO levels to fully saturate the respiratory chain under the applied conditions. Further, oxygen is supplied by surface aeration in the BioLector while the medium is supplied with a constant air stream in the bioreactor. In the latter case, this might lead to an increased loss of volatile by-products like ethanol via evaporation and might thereby prevent its re-assimilation. We also observed a reduced

by-product accumulation during bioreactor cultivation at lower constant pH values of 3 and 4 (own unpublished results), indicating a combined effect of both DO and pH on the physiology of *P. pastoris*.

The presented microscale cultivation set-up was applied to screen a library of *P. pastoris*::pGAPZαB_*appA* transformants secreting AppA phytase from *E. coli*.

By screening of 47 *P. pastoris*::pGAPZαB_*appA* transformants in BSM$_{mod}$ with 4% D-glucose in microscale cultivation, significantly increased protein concentrations and phytase activities in the supernatant compared to the parental strain were measured for two clones. Protein titers still were only low, which might be due to the fact that we did not include steps for the selection of multicopy clones during strain construction. Clones carrying multiple copies of the expression cassette often show higher expression levels than those with only one or few copies [55]. This correlation not always holds true, especially in the case of secreted proteins, where the secretion apparatus may become a bottleneck in protein production [10, 56].

Specific growth rates and biomass yields were very similar. Nevertheless, it is common knowledge that protein overproduction can lead to reduced biomass yields and growth rates [57, 58]. The possibility for parallel phenotyping with the help of online process data already during screenings makes it possible to select those that promise to perform well after scale-up to bioreactor cultivation. Non-invasively monitoring DO and pH during screening further assures reproducible conditions and is an advantage compared to cultivation in standard microplates.

Some solutions to transfer carbon-limited fed-batch cultivation to microscale have been established in the past as recently reviewed [59]. For example, repeated dosing of a substrate bolus to the culture has been used to increase biomass and product concentrations in microscale cultivation [31]. However, this approach cannot be regarded as a true fed-batch process because it leads to alternating conditions of excess substrate (maximum specific growth rate) and substrate limitation (zero growth). Continuous feeding of small volumes required for microscale fed-batch cultivation is currently still limited by the availability of suitable micropumps, although first prototypic examples have been described [43, 59, 60]. This problem has been circumvented by providing substrate via diffusion [61, 62]. Another possibility is the enzymatic substrate release from oligo- or polymeric inert compounds like starch [31, 60, 63] or sucrose [43] at a constant rate. Contrary to exponential substrate feeding, which can be applied during lab-scale bioreactor cultivation to adjust a constant specific growth rate, a constant substrate release rate will lead to a decreasing specific growth rate over time. Since an interdependence

of specific growth rate and protein production rate is well documented [15, 64, 65], the development of more flexible feeding strategies suitable for microscale cultivation remains an important issue for further optimization.

A thorough characterization of cultivation profiles in batch-mode was performed in this study to enable a direct comparison of results from microscale and laboratory scale to ensure scalability of the developed cultivation setup before advancing to the development of a microscale fed-batch cultivation protocol. For the latter, we employed enzymatic substrate release, since it is easily applicable and well suited to infer carbon-limited conditions during microscale cultivation. When screening a clonal library of phytase producing *P. pastoris* under carbon-limited conditions, biomass concentrations, protein titers and phytase activities in the supernatant were strongly increased compared to batch cultivation. In batch screening, it was possible to identify clones with increased phytase production in a shorter time, but the clonal ranking for fed-batch-like screening was different concerning the two best producers. Even more drastic changes in clonal rankings for different operational modes were reported by Scheidle et al. for GFP production in *Hansenula polymorpha* [33]. Further, clones with a reduced maximum growth rate, e.g., due to a metabolic burden caused by protein overproduction, might be deselected after batch screening, although they could be well suited for protein production in a carbon-limited process Taken together, the possibility to perform clonal selections under fed-batch conditions clearly presents a major advantage over screening in batch mode.

Application of the microscale cultivation protocol could comprise not only screening, but also the rapid initial testing of process parameters during process optimization. Variables like C- and N-source, media composition, temperature and initial pH could be tested, e.g., as part of DoE studies, to reduce the number of bioreactor cultivations. Further, the developed microscale cultivation set-up for *P. pastoris* also allows an easy integration into liquid handling stations [31, 42]. Miniaturization and parallelization of microbial cultivation will eventually contribute to an acceleration of bioprocess development.

Conclusions

In this study, we developed, validated and applied a setup for the screening of clonal libraries of protein producing *P. pastoris* employing the constitutive GAP promotor in a parallel microbioreactor system. First, a defined high cell density cultivation medium was modified by the addition of N-source, the selection of an inert buffer for cultivation at pH 5 and the FeSO$_4$ concentration was reduced to prevent precipitation. Scalability of cultivation results from microscale to laboratory bioreactor cultivation was

ensured by the comparison of physiological parameters ($Y_{X/S}$, μ) for growth on D-glucose and glycerol as substrates in both scales and for D-glucose even an improvement of $Y_{X/S}$ was found in microscale. In addition to laborious offline measurements, online growth signals provided by the cultivation system were used for the parallel determination of physiological parameters, e.g., the calculation of specific growth rates from online backscatter measurements, and online monitoring of DO and pH ensured reproducible conditions throughout the conducted screening experiments.

Enzymatic substrate release was adopted using a glucose polymer and glucose liberating amyloglucosidase to impose carbon-limited growth conditions, allowing screening under process relevant fed-batch conditions in microscale cultivation. Therefore, different glucose release rates were adjusted by the addition of amyloglucosidase and suitable cultivation conditions could be identified.

The established microscale cultivation was applied for the screening of a clonal library of *P. pastoris*::pGAPZαB_*appA* constitutively secreting AppA phytase from *E. coli* both in batch and fed-batch mode. Although the former allowed a rapid selection of clones producing increased extracellular protein titers and phytase activities, protein concentrations and phytase activities were increased 10- and 20-fold, respectively, in fed-batch microscale cultivation. DCW concentrations > 45 g/L were reached at fully aerobic conditions, exceeding maximum DCW concentrations during batch cultivation more than 2-fold. Most importantly, differences in the clonal ranking were found between the two operational modes, i.e. batch and fedbatch. This confirms the necessity of reliable advanced screening technologies such as is presented in this study to accelerate the future development of bioprocesses.

Methods
Chemicals and growth media
All chemicals and enzymes were obtained from Sigma Aldrich (Steinheim, Germany) or Carl Roth (Karlsruhe, Germany) if not mentioned otherwise and were of analytical grade. PIPPS was purchased from Merck (Darmstadt, Germany).

YPD complex medium consisted of 10 g/L yeast extract, 20 g/L peptone and 20 g/L D-glucose. For solid YPD medium, 16 g/L agar–agar was added. Basal salts medium (BSM) [45] was used with modifications, consisting of 81.1 mM ortho-phosphoric acid, 9.4 mM $MgSO_4 \times 7\ H_2O$, 1 mM $CaSO_4 \times 2\ H_2O$, 16.4 mM K_2SO_4, and 75.7 mM $(NH_4)_2SO_4$. The pH was adjusted to 5.0 with KOH. A suitable buffer was added for cultivations without pH control (see results). After sterilization, 4 mL/L trace salts solution was added containing 24 mM

$CuSO_4 \times 5\ H_2O$, 0.53 mM NaI, 19.87 mM $MnSO_4 \times H_2O$, 0.83 mM $Na_2MoO_4 \times 2\ H_2O$, 0.32 mM H_3BO_4, 2.1 mM $CoCl_2 \times 6\ H_2O$, 150 mM $ZnCl_2$, 230 mM $FeSO_4 \times 7\ H_2O$ and 0.82 mM biotin. For BSM_{mod}, the concentration of $FeSO_4 \times 7\ H_2O$ in the trace salts solution was reduced to 2.30 mM.

Strain construction
Escherichia coli DH5α [66] was used for general cloning, whereas *P. pastoris* X-33 (Invitrogen, Carlsbad, USA) was used for gene expression. The gene *appA* (gene ID 946206) from *E. coli* K-12 substrain MG1655 was codon optimized for *P. pastoris* and synthesized by GeneArt™ (Regensburg, Germany) and inserted, after digestion with *Pst*I and *Not*I, into the vector pGAPZαB (Invitrogen). The resulting plasmid pGAPZαB_*appA* was transformed into *P. pastoris* X-33 by electroporation using the method described by Wu and Letchworth [67], resulting in the strain *P. pastoris*::pGAPZαB_*appA*.

Cultivation of *P. pastoris*
Strains were preserved as cryocultures in 30% glycerol at − 80 °C. For cultivation of *P. pastoris* strains, cells were plated on YPD agar containing 100 μg/mL Zeocin (Invivogen, Toulouse, France) if required and incubated at 30 °C for 2 days. Then, 10 mL YPD liquid medium were inoculated and incubated in a rotary shaker (Multitron Standard, Infors HT, Einsbach, Germany) at 250 rpm and 30 °C for 8 h. A second pre-culture in 10–100 mL basal salts medium supplemented with 150 mM PIPPS (pH 5.0) and 1% of the same carbon source as in the main culture was inoculated and cultivated under the same conditions for 16 h. Pre-cultures were grown in baffled shake flasks ten times the volume of the culture broth.

Microscale cultivation of *P. pastoris* in a BioLector (m2p-labs, Baesweiler, Germany) was performed using MTP-48-BOHL FlowerPlates, which are equipped with pre-calibrated optodes for online DO (dynamic range 0–100% saturation) and pH (dynamic range pH 2.25–5.75) measurement (m2p-labs). The plates were sealed with gas-permeable F-GP-10 sealing foils (m2p-labs). Cultivation settings were 30 °C, 1500 rpm, and atmospheric humidity control ≥ 85%. 800 μL BSM or BSM_{mod} were supplemented with 150 mM PIPPS pH 5.0 and 4% of either D-glucose or glycerol. For cultivation under carbon-limited conditions, 10% dextrin from potato starch and amyloglucosidase from *A. niger* (see "Results") were added. Media were inoculated from a pre-culture to an initial OD_{600} between 0.1 and 1.7. Growth was monitored online by backscatter measurement (620 nm, gain = 15).

Bioreactor cultivations were carried out in 0.8 L BSM_{mod} with 75.7 mM $(NH_4)_2SO_4$ without buffer at 30 °C

in DASGIP Bioblock stirred tank reactors (DASGIP, Jülich, Germany) equipped with electrodes for pH and DO and off-gas analysis for O_2 and CO_2. The pH was maintained at 5.0 by addition of 18% (v/v) NH_4OH and 8% (v/v) H_2SO_4. The DO was controlled by applying a stirrer cascade (400–1200 rpm), while the airflow was constant at 60 L/h sterile ambient air (0.2 μm PTFE filter, DIAFIL, Wieliczka, Poland).

Analytical procedures

DCW concentrations were determined gravimetrically. 0.5–2 mL sample were collected in dried Eppendorf tubes (2 days at 80 °C, 1 day in a desiccator) and centrifuged for 5 min at 13,000 rpm (Biofuge pico, Heraeus, Hanau, Germany). The pellet was washed in 0.9% NaCl solution and dried as described above.

Protein concentrations in culture supernatants were determined using Bradford reagent (Sigma Aldrich) according to the manufacturer's instructions with bovine serum albumin as standard. SDS-PAGE for analysis of culture supernatants was performed using 12% TruPAGE™ (Sigma Aldrich) polyacrylamide gels and Coomassie staining according to the manufacturer's instructions.

A fluorescence assay was applied to determine phytase activities in culture supernatants using 4-MUP as substrate [68]. Proteins excreted to the culture supernatant were separated from low molecular weight compounds via ultrafiltration (Nanosep Omega 10 kDa, Pall, Dreieich, Germany) using 3 volumes of 50 mM sodium acetate (pH 5.0) for washing and as exchange buffer. 50 μL of the resulting sample was mixed with 50 μL 1 mM 4-MUP in 50 mM sodium acetate buffer (pH 5.0) in a transparent 96 well plate and the increase of fluorescence (excitation 336 nm, emission 448 nm, 25 °C) was recorded for 15 min in an Enspire microplate reader (Perkin Elmer, Rodgau, Germany). Enzyme activities were calculated from the resulting slope using a calibration curve of the reaction product.

Sugars and organic acids in the culture supernatant were analyzed on an Agilent 1100 series HPLC system equipped with RI and DAD detector (Agilent Technologies, Waldbronn, Germany). As stationary phase, an organic acid resin (300 × 8 mm, CS-Chromatographie Service, Langerwehe, Germany) was used at 80 °C and isocratic elution with 0.6 mL/min 100 mM H_2SO_4. Samples were analyzed quantitatively by external calibration.

Calculation of physiological parameters and rates

Biomass yields were calculated from final DCW concentrations and initial substrate concentrations. For the calculation of specific growth rates from the backscatter measurement in the BioLector, the signal was converted to DCW concentrations with the help of a non-linear calibration (Additional file 1: Figure S1) and by fitting an exponential growth model to the progression of the DCW over time. For the calculation of all rates and yield coefficients for BioLector cultivations, the culture volumes were corrected for 9.09% evaporation per day, which was determined gravimetrically and assumed to be linear over time.

Statistical analysis

ANOVA and multiple comparison of means tests were performed with Matlab R2016b (The MathWorks GmbH, Ismaning, Germany).

Author details

[1] Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany. [2] Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany. [3] iAMB - Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, 52074 Aachen, Germany. [4] Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany. [5] Computational Systems Biotechnology (AVT.CSB), RWTH Aachen University, 52074 Aachen, Germany.

References

1. Spohner SC, Muller H, Quitmann H, Czermak P. Expression of enzymes for the usage in food and feed industry with Pichia pastoris. J Biotechnol. 2015;202:118–34.
2. Weinacker D, Rabert C, Zepeda AB, Figueroa CA, Pessoa A, Farias JG. Applications of recombinant Pichia pastoris in the healthcare industry. Braz J Microbiol. 2013;44:1043–8.
3. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the Pichia pastoris expression system. Yeast. 2005;22:249–70.

4. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol. 2014;98:5301–17.

5. Barnard GC, Kull AR, Sharkey NS, Shaikh SS, Rittenhour AM, Burnina I, Jiang Y, Li F, Lynaugh H, Mitchell T, et al. High-throughput screening and selection of yeast cell lines expressing monoclonal antibodies. J Ind Microbiol Biotechnol. 2010;37:961–71.

6. Schwarzhans JP, Wibberg D, Winkler A, Luttermann T, Kalinowski J, Friehs K. Integration event induced changes in recombinant protein productivity in *Pichia pastoris* discovered by whole genome sequencing and derived vector optimization. Microb Cell Fact. 2016;15:84.

7. Brooks CL, Morrison M, Joanne Lemieux M. Rapid expression screening of eukaryotic membrane proteins in *Pichia pastoris*. Protein Sci. 2013;22:425–33.

8. Sygmund C, Gutmann A, Krondorfer I, Kujawa M, Glieder A, Pscheidt B, Haltrich D, Peterbauer C, Kittl R. Simple and efficient expression of *Agaricus meleagris* pyranose dehydrogenase in *Pichia pastoris*. Appl Microbiol Biotechnol. 2012;94:695–704.

9. Sunga AJ, Tolstorukov I, Cregg JM. Posttransformational vector amplification in the yeast *Pichia pastoris*. FEMS Yeast Res. 2008;8:870–6.

10. Zhu T, Guo M, Tang Z, Zhang M, Zhuang Y, Chu J, Zhang S. Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. J Appl Microbiol. 2009;107:954–63.

11. Aw R, Polizzi KM. Can too many copies spoil the broth? Microb Cell Fact. 2013;12:128.

12. Liu M, Potvin G, Gan YR, Huang ZB, Zhang ZS. Medium optimization for the production of phytase by recombinant *Pichia pastoris* grown on glycerol. Int J Chem React Eng. 2011;9:1–15.

13. Ghosalkar A, Sahai V, Srivastava A. Optimization of chemically defined medium for recombinant *Pichia pastoris* for biomass production. Bioresour Technol. 2008;99:7906–10.

14. Freier L, Hemmerich J, Schöler K, Wiechert W, Oldiges M, von Lieres E. Framework for Kriging-based iterative experimental analysis and design: optimization of secretory protein production in *Corynebacterium glutamicum*. Eng Life Sci. 2016;16:538–49.

15. Garcia-Ortega X, Ferrer P, Montesinos JL, Valero F. Fed-batch operational strategies for recombinant Fab production with *Pichia pastoris* using the constitutive GAP promoter. Biochem Eng J. 2013;79:172–81.

16. Varnai A, Tang C, Bengtsson O, Atterton A, Mathiesen G, Eijsink VGH. Expression of endoglucanases in *Pichia pastoris* under control of the GAP promoter. Microb Cell Fact. 2014;13:57.

17. Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D. The effect of temperature on the proteome of recombinant *Pichia pastoris*. J Proteome Res. 2009;8:1380–92.

18. Shi XZ, Karkut T, Chamankhah M, Alting-Mees M, Hemmingsen SM, Hegedus D. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. Protein Expr Purif. 2003;28:321–30.

19. Holmes WJ, Darby RAJ, Wilks MDB, Smith R, Bill RM. Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime. Microb Cell Fact. 2009;8:35.

20. Charoenrat T, Ketudat-Cairns M, Stendahl-Andersen H, Jahic M, Enfors SO. Oxygen-limited fed-batch process: an alternative control for *Pichia pastoris* recombinant protein processes. Bioprocess Biosyst Eng. 2005;27:399–406.

21. Baumann K, Maurer M, Dragosits M, Cos O, Ferrer P, Mattanovich D. Hypoxic fed-batch cultivation of *Pichia pastoris* increases specific and volumetric productivity of recombinant proteins. Biotechnol Bioeng. 2008;100:177–83.

22. Ruff AJ, Dennig A, Schwaneberg U. To get what we aim for-progress in diversity generation methods. FEBS J. 2013;280:2961–78.

23. Mandenius CF, Brundin A. Bioprocess optimization using design-of-experiments methodology. Biotechnol Prog. 2008;24:1191–203.

24. Harms J, Wang XY, Kim T, Yang XM, Rathore AS. Defining process design space for biotech products: case study of *Pichia pastoris* fermentation. Biotechnol Prog. 2008;24:655–62.

25. Bareither R, Pollard D. A review of advanced small-scale parallel bioreactor technology for accelerated process development: current state and future need. Biotechnol Prog. 2011;27:2–14.

26. Wewetzer SJ, Kunze M, Ladner T, Luchterhand B, Roth S, Rahmen N, Kloss R, Silva ACE, Regestein L, Büchs J. Parallel use of shake flask and microtiter plate online measuring devices (RAMOS and BioLector) reduces the number of experiments in laboratory-scale stirred tank bioreactors. J Biol Eng. 2015;9:9.

27. Büchs J. Introduction to advantages and problems of shaken cultures. Biochem Eng J. 2001;7:91–8.

28. Gupta A, Rao G. A study of oxygen transfer in shake flasks using a non-invasive oxygen sensor. Biotechnol Bioeng. 2003;84:351–8.

29. Singh S, Gras A, Fiez-Vandal C, Ruprecht J, Rana R, Martinez M, Strange PG, Wagner R, Byrne B. Large-scale functional expression of WT and truncated human adenosine A2A receptor in *Pichia pastoris* bioreactor cultures. Microb Cell Fact. 2008;7:28.

30. Unthan S, Radek A, Wiechert W, Oldiges M, Noack S. Bioprocess automation on a Mini Pilot Plant enables fast quantitative microbial phenotyping. Microb Cell Fact. 2015;14:32.

31. Hemmerich J, Adelantado N, Barrigon JM, Ponte X, Hormann A, Ferrer P, Kensy F, Valero F. Comprehensive clone screening and evaluation of fed-batch strategies in a microbioreactor and lab scale stirred tank bioreactor system: application on *Pichia pastoris* producing *Rhizopus oryzae* lipase. Microb Cell Fact. 2014;13:36.

32. Prielhofer R, Cartwright SP, Graf AB, Valli M, Bill RM, Mattanovich D, Gasser B. *Pichia pastoris* regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics. 2015;16:167.

33. Scheidle M, Jeude M, Dittrich B, Denter S, Kensy F, Suckow M, Klee D, Büchs J. High-throughput screening of *Hansenula polymorpha* clones in the batch compared with the controlled-release fed-batch mode on a small scale. FEMS Yeast Res. 2010;10:83–92.

34. Weis R, Luiten R, Skranc W, Schwab H, Wubbolts M, Glieder A. Reliable high-throughput screening with *Pichia pastoris* by limiting yeast cell death phenomena. FEMS Yeast Res. 2004;5:179–89.

35. Kirk TV, Szita N. Oxygen transfer characteristics of miniaturized bioreactor systems. Biotechnol Bioeng. 2013;110:1005–19.

36. Lattermann C, Büchs J. Microscale and miniscale fermentation and screening. Curr Opin Biotechnol. 2015;35:1–6.

37. Hortsch R, Weuster-Botz D. Milliliter-scale stirred tank reactors for the cultivation of microorganisms. Adv Appl Microbiol. 2010;73:61–82.

38. Islam RS, Tisi D, Levy MS, Lye GJ. Scale-up of *Escherichia coli* growth and recombinant protein expression conditions from microwell to laboratory and pilot scale based on matched $k_L a$. Biotechnol Bioeng. 2008;99:1128–39.

39. Kensy F, Zang E, Faulhammer C, Tan RK, Büchs J. Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates. Microb Cell Fact. 2009;8:31.

40. Rohe P, Venkanna D, Kleine B, Freudl R, Oldiges M. An automated workflow for enhancing microbial bioprocess optimization on a novel microbioreactor platform. Microb Cell Fact. 2012;11:144.

41. Isett K, George H, Herber W, Amanullah A. Twenty-four-well plate miniature bioreactor high-throughput system: assessment for microbial cultivations. Biotechnol Bioeng. 2007;98:1017–28.

42. Huber R, Ritter D, Hering T, Hillmer AK, Kensy F, Müller C, Wang L, Büchs J. Robo-Lector—a novel platform for automated high-throughput cultivations in microtiter plates with high information content. Microb Cell Fact. 2009;8:42.

43. Faust G, Janzen NH, Bendig C, Römer L, Kaufmann K, Weuster-Botz D. Feeding strategies enhance high cell density cultivation and protein expression in milliliter scale bioreactors. Biotechnol J. 2014;9:1293–303.

44. Wilming A, Bahr C, Kamerke C, Büchs J. Fed-batch operation in special microtiter plates: a new method for screening under production conditions. J Ind Microbiol Biot. 2014;41:513–25.

45. Hellwig S, Emde F, Raven NP, Henke M, van Der Logt P, Fischer R. Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations. Biotechnol Bioeng. 2001;74:344–52.

46. Heyland J, Fu J, Blank LM, Schmid A. Carbon metabolism limits recombinant protein production in *Pichia pastoris*. Biotechnol Bioeng. 2011;108:1942–53.

47. Stahl CH, Wilson DB, Lei XG. Comparison of extracellular *Escherichia coli*

AppA phytases expressed in *Streptomyces lividans* and *Pichia pastoris*. Biotechnol Lett. 2003;25:827–31.

48. Cregg JM, Vedvick TS, Raschke WC. Recent advances in the expression of foreign genes in *Pichia pastoris*. Nat Biotechnol. 1993;11:905–10.

49. Sinha J, Plantz BA, Inan M, Meagher MM. Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-τ. Biotechnol Bioeng. 2005;89:102–12.

50. Neal AL, Weinstock JO, Lampen JO. Mechanisms of fatty acid toxicity for yeast. J Bacteriol. 1965;90:126–31.

51. Cos O, Ramon R, Montesinos JL, Valero F. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: a review. Microb Cell Fact. 2006;5:17.

52. Maurer M, Kühleitner M, Gasser B, Mattanovich D. Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*. Microb Cell Fact. 2006;5:37.

53. Invitrogen Life Technologies: Pichia fermentation process guidelines version B. http://tools.thermofisher.com/content/sfs/manuals/pichia-ferm_prot.pdf. Last accessed 25/01/2017.

54. Isidro IA, Ferreira AR, Clemente JJ, Cunha AE, Oliveira R. Analysis of culture media screening data by projection to latent pathways: the case of *Pichia pastoris* X-33. J Biotechnol. 2016;217:82–9.

55. Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K, Romanos MA. High-Level expression of tetanus toxin fragment-C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. Nat Biotechnol. 1991;9:455–60.

56. Inan M, Aryasomayajula D, Sinha J, Meagher MM. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol Bioeng. 2006;93:771–8.

57. Mattanovich D, Gasser B, Hohenblum H, Sauer M. Stress in recombinant protein producing yeasts. J Biotechnol. 2004;113:121–35.

58. Ramon R, Ferrer P, Valero F. Sorbitol co-feeding reduces metabolic burden caused by the overexpression of a *Rhizopus oryzae* lipase in *Pichia pastoris*. J Biotechnol. 2007;130:39–46.

59. Krause M, Neubauer A, Neubauer P. The fed-batch principle for the molecular biology lab: controlled nutrient diets in ready-made media improve production of recombinant proteins in *Escherichia coli*. Microb Cell Fact. 2016;15:110.

60. Grimm T, Grimm M, Klat R, Neubauer A, Palela M, Neubauer P. Enzyme-based glucose delivery as a high content screening tool in yeast-based whole-cell biocatalysis. Appl Microbiol Biot. 2012;94:931–7.

61. Ruth C, Buchetics M, Vidimce V, Kotz D, Naschberger S, Mattanovich D, Pichler H, Gasser B. *Pichia pastoris* Aft1—a novel transcription factor, enhancing recombinant protein secretion. Microb Cell Fact. 2014;13:120.

62. Jeude M, Dittrich B, Niederschulte H, Anderlei T, Knocke C, Klee D, Büchs J. Fed-batch mode in shake flasks by slow-release technique. Biotechnol Bioeng. 2006;95:433–45.

63. Panula-Perälä J, Vasala A, Karhunen J, Ojamo H, Neubauer P, Mursula A. Small-scale slow glucose feed cultivation of *Pichia pastoris* without repression of AOX1 promoter: towards high throughput cultivations. Bioprocess Biosyst Eng. 2014;37:1261–9.

64. Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, Mattanovich D. In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. Biotechnol J. 2014;9:511–25.

65. Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar K. Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. Biotechnol Adv. 2015;33:1177–93.

66. Grant SG, Jessee J, Bloom FR, Hanahan D. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc Natl Acad Sci USA. 1990;87:4645–9.

67. Wu S, Letchworth GJ. High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. Biotechniques. 2004;36:152–4.

68. Garrett JB, Kretz KA, O'Donoghue E, Kerovuo J, Kim W, Barton NR, Hazlewood GP, Short JM, Robertson DE, Gray KA. Enhancing the thermal tolerance and gastric performance of a microbial phytase for use as a phosphate-mobilizing monogastric-feed supplement. Appl Environ Microbiol. 2004;70:3041–6.

Gene regulation associated with sexual development and female fertility in different isolates of *Trichoderma reesei*

Christoph Dattenböck[1], Doris Tisch[2], Andre Schuster[2], Alberto Alonso Monroy[1], Wolfgang Hinterdobler[1] and Monika Schmoll[1*]

Abstract

Background: *Trichoderma reesei* is one of the most frequently used filamentous fungi in industry for production of homologous and heterologous proteins. The ability to use sexual crossing in this fungus was discovered several years ago and opens up new perspectives for industrial strain improvement and investigation of gene regulation.

Results: Here we investigated the female sterile strain QM6a in comparison to the fertile isolate CBS999.97 and back-crossed derivatives of QM6a, which have regained fertility (FF1 and FF2 strains) in both mating types under conditions of sexual development. We found considerable differences in gene regulation between strains with the CBS999.97 genetic background and the QM6a background. Regulation patterns of QM6a largely clustered with the backcrossed FF1 and FF2 strains. Differential regulation between QM6a and FF1/FF2 as well as clustering of QM6a patterns with those of CBS999.97 strains was also observed. Consistent mating type dependent regulation was limited to mating type genes and those involved in pheromone response, but included also *nta1* encoding a putative N-terminal amidase previously not associated with development. Comparison of female sterile QM6a with female fertile strains showed differential expression in genes encoding several transcription factors, metabolic genes and genes involved in secondary metabolism.

Conclusions: Evaluation of the functions of genes specifically regulated under conditions of sexual development and of genes with highest levels of transcripts under these conditions indicated a relevance of secondary metabolism for sexual development in *T. reesei*. Among others, the biosynthetic genes of the recently characterized SOR cluster are in this gene group. However, these genes are not essential for sexual development, but rather have a function in protection and defence against competitors during reproduction.

Keywords: *Trichoderma reesei*, *Hypocrea jecorina*, Sexual development, Female fertility, Secondary metabolism, Mating type

Background

The parental strain of *Trichoderma reesei* (syn. *Hypocrea jecorina*) strains that dominated research and industry, QM6a, has been isolated during WWII in the tropics [1–3]. For decades it was considered asexual, which was a drawback for genetics research and industrial

applications [4]. In 2009, sexual development of *T. reesei* under laboratory conditions was achieved [5, 6], albeit at the same time, female sterility of QM6a was discovered. Later on, a mutation in the gene encoding the WD-40 protein IDC1/HAM-5 [7, 8] was identified to cause female sterility in QM6a [9, 10].

Sexual development is dependent on the presence of a functional pheromone system in fungi [11] as well as on precisely defined environmental conditions [12, 13]. In *T. reesei*, in contrast to many other fungi, sexual development is initiated upon growth on complex media such as

malt extract agar (MEA) or potato dextrose agar (PDA) and occurs preferentially in light [14]. *T. reesei* requires one of the two mating type associated pheromone precursor—pheromone receptor pairs (*hpp1–hpr1* or *ppg1–hpr2*) to be functional in order to undergo mating successfully [15]. Interestingly, *T. reesei* has no conventional a-type peptide pheromone precursor, but employs a novel type of pheromones, with HPP1 as the first representative of h-type pheromones [16].

The blue light photoreceptor ENV1 is crucial for light dependent balancing of regulation of the pheromone system. Its abolishment causes female sterility in light due to deregulation of the expression of pheromone receptor and precursor genes, predominantly in the mating type MAT1-2 [17]. This female sterility is conditional and can be overcome by application of an altered light regime [18]. In contrast to ENV1, the BLR1-BLR2 photoreceptor is not essential for sexual development as mutation of the genes causes only minor modulations in its efficiency as well as some morphological alterations in the fruiting body [17, 19].

Besides the light signaling pathway also protein kinase A and adenylate cyclase, the major components of the cAMP pathway of *T. reesei,* influence efficiency of sexual development [20]. For the heterotrimeric G-protein pathway, functions in sexual development are known for the G-protein beta and gamma subunits GNB1 and GNG1 [21].

Chemical communication via the secretion of secondary metabolites was shown to be important for sexual development in *T. reesei.* The pattern of secondary metabolites secreted into the medium changes if a compatible mating partner is sensed. Thereby, VEL1 was found to be crucial for triggering recognition associated signaling [22]. Secretion of secondary metabolites is regulated by light in *T. reesei* and a connection to carbon catabolite repression has been shown [23]. Moreover, among the genes regulated in a cellulase induction specific manner, several secondary metabolism associated genes were found, including the polyketide synthase *pks4* [24], which is responsible for the green coloration of spores of *T. reesei* among other functions [25].

The genome of the natural isolate CBS999.97 [6, 26] was published recently [10] and showed a particularly low occurrence of non-synonymous SNPs within genesets enriched in functions of metabolism, signal transduction and stress response, while genes comprising a high number of SNPs are involved in secondary metabolism or photoperception [10]. Comparative analysis of CBS999.97, QM6a and FF1/FF2, female fertile strains backcrossed from CBS999.97 to gain the QM6a phenotype revealed different carbon utilization characteristics between the two strain backgrounds. Additionally,

secondary metabolite profiles were different between CBS999.97 and QM6a and regulatory differences associated with female fertility and female sterility were detected, which include regulation of CAZyme and transporter encoding genes [10].

In this study we investigated the transcriptome of QM6a representing strains applied in research and industry as well as those of the female fertile isolate CBS999.97 and backcrossed strains (FF1, FF2) under conditions of sexual development. Besides differential gene regulation between different mating types and strain backgrounds, we also found altered regulation between female fertile strains and QM6a. Moreover, a relevance of secondary metabolism for sexual development in *T. reesei* became obvious.
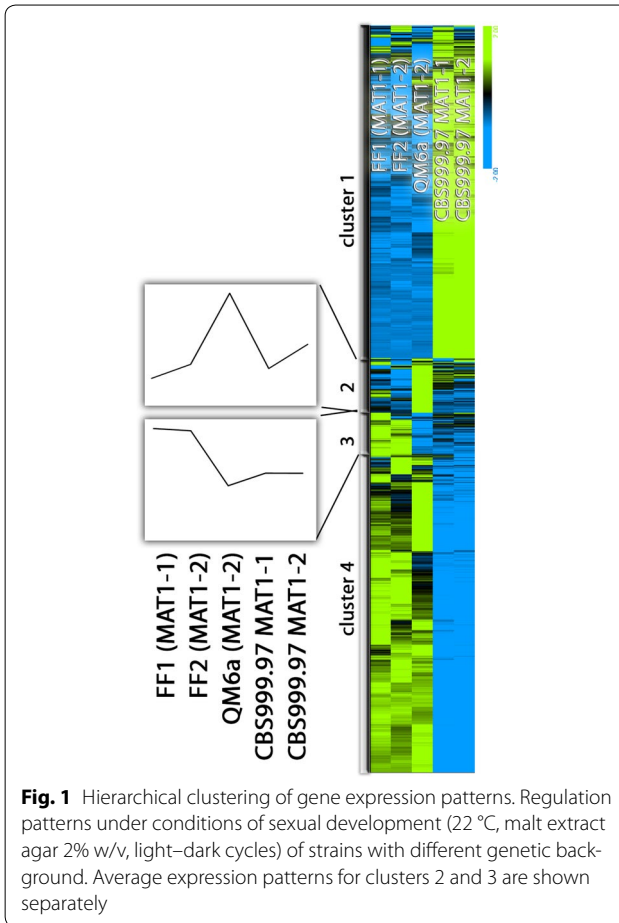
Results

Gene expression patterns in strains with QM6a background versus CBS999.97

We performed transcriptome analysis under conditions facilitating sexual development and enable the associated chemical communication using the wild-isolate CBS999.97 in both mating types [6], the female sterile strain QM6a and female fertile derivatives of QM6a in both mating types (FF1 and FF2), which were prepared by repeated backcrossing [10, 22]. Hierarchical cluster analysis of gene expression revealed 4 clusters (Fig. 1). This analysis clearly showed that under conditions of sexual development, QM6a is more similar to FF1 and FF2 than to CBS999.97 and hence confirms that the backcrossing procedure largely restored the QM6a phenotype. A similar result was also achieved for conditions of cellulase gene expression [10].

Of the identified 4 clusters, clusters 2 and 3 show the most interesting patterns (Fig. 1). Cluster 2 comprises genes upregulated in QM6a compared to FF1/FF2 as well as CBS999.97. Functional category analysis of these genes showed numerous metabolic genes in this cluster, but significant enrichment (p-values below $5E-03$) was only found for functions in RNA processing, sesquiterpenes metabolism, cellular import and defense. In cluster 3 there are genes for which QM6a clusters with CBS999.97 rather than FF1/FF2. This gene set was enriched for functions in detoxification as well as C-1 compound metabolism. Interestingly, this gene set also contained an unusually high proportion of unclassified proteins suggesting that yet unknown functions may be shared by QM6a and CBS999.97.

Hierarchical clustering of the subset of genes annotated as involved in sexual development [3] showed a comparable distribution, and a clear similarity between QM6a and FF1/FF2 (data not shown).

Fig. 1 Hierarchical clustering of gene expression patterns. Regulation patterns under conditions of sexual development (22 °C, malt extract agar 2% w/v, light–dark cycles) of strains with different genetic background. Average expression patterns for clusters 2 and 3 are shown separately

Gene regulation in female fertile strains compared to female sterile QM6a

We were interested whether the difference between female fertility and female sterility is reflected in differential gene expression between crossings of QM6a and those of female fertile strains. Therefore we first compared gene regulation in FF2 strains versus QM6a. We found 210 genes to be up-regulated and 170 genes to be downregulated in FF2 strains compared to QM6a (Additional file 1). In order to get a more robust evaluation of potential alterations in QM6a, we checked how many of these genes are consistently regulated in CBS999.97 MAT1-2. Only 93 of the upregulated genes are also upregulated in CBS999.97 MAT1-2 and 74 of the down regulated genes show the same regulation in this strain (Additional file 1). The gene set of upregulated genes in female fertile MAT1-2 strains was significantly enriched in functions of drug/toxin transport (*p* value 1.39e−04), type I protein secretion (p-value 1.20e−03), disease, virulence and defense (p-value 3.28e−03) and detoxification by export (p-value 1.80e−03).
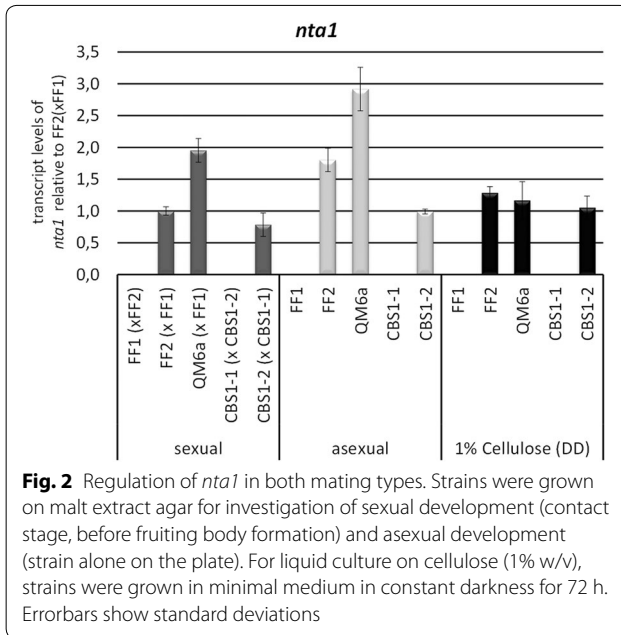
Interestingly, also the gene encoding SOR4/TR_43701, which was recently shown to influence production of sorbicillin derivatives in *T. reesei* [23], was among these genes. Its deletion does not impact sexual development of *T. reesei* (A. Monroy, unpublished results). Further upregulated genes include a *ccg-13* homologue potentially involved in asexual development, three CAZyme encoding genes, 6 transcription factor genes and several predicted transporter genes (Additional file 1).

Genes down-regulated in female fertile MAT1-2 compared to QM6a were significantly enriched in functions in metabolism (p-value 5.89e−04), particularly secondary metabolism (p-value 2.46e−05). However, also functions in disease, virulence and defense were enriched, indicating a consistent shift in regulation of genes within similar functional groups between female fertile MAT1-2 strains and QM6a. This gene set comprises moreover three transcription factors, several transporters as well as the polyketide synthase gene *pks6g* and the terpenoid synthase encoding *tps7*. Interestingly, also *lae1*, which encodes a putative methyltransferase and impacts secondary metabolism in several fungi [27–29] is up to three-fold downregulated compared to QM6a (Additional file 1).

Few genes only are consistently differentially regulated in different mating types

Differences in gene expression between different mating types have been reported previously, also in *T. reesei* [10]. Therefore we were interested whether such differences are detectable in *T. reesei* under conditions of sexual development and if they are consistent in different strain backgrounds (CBS999.97 vs. FF1/FF2 which have the QM6a genetic background).

In CBS999.97 we found 18 genes to be differentially regulated between MAT1-1 and MAT1-2 (at least two-fold, p-value 0.01), 12 were downregulated and 6 were upregulated in MAT1-1 (Additional file 2). For FF1/FF2 we detected differential regulation for 39 genes, with 13 genes upregulated in FF1 and 26 genes downregulated. Of those genes, only 6 were consistently regulated in CBS999.97 and FF1/FF2 and can hence be considered consistently mating type regulated in *T. reesei*. These genes comprise the two peptide pheromone receptor genes *hpr1* and *hpr2*, as well as the mating type genes *mat1-2-1*, *mat1-1-1* and *mat1-1-3*. Thereby, *hpr1* is upregulated in MAT1-1 and *hpr2* is upregulated in MAT1-2, as would be expected due to the associate mating types [15]. The genes of the mating type locus can be considered to be regulated above background, as they are not present in the opposite mating types and show up-regulation in their cognate mating types (*mat1-2-1* in MAT1-2; *mat1-1-1* and *mat1-1-3* in MAT1-1). In contrast, *mat1-1-2* was not found to be regulated and only very

Fig. 2 Regulation of *nta1* in both mating types. Strains were grown on malt extract agar for investigation of sexual development (contact stage, before fruiting body formation) and asexual development (strain alone on the plate). For liquid culture on cellulose (1% w/v), strains were grown in minimal medium in constant darkness for 72 h. Errorbars show standard deviations

Fig. 3 Comparision between cellulose and sexual development. **a** PCA analysis of gene expression patterns of different strains (CBS999.97 MAT1-1, CBS999.97 MAT1-2, FF1, FF2, QM6a) and upon growth on cellulose in liquid media as well as under conditions of sexual development. **b** Venn diagram of the comparison between genes differentially regulated between female fertile strains and QM6a versus the same gene set upon sexual development. **c** Venn diagram of the comparison between genes differentially regulated between cellulose and sexual development in CBS999.97 versus FF1/FF2 upon growth on cellulose compared to sexual development. Overlapping areas in the Venn diagrams show differentially regulated genes under both conditions or in both strains

low transcript levels we observed, indicating that this gene may be relevant for a different stage of sexual development than investigated here.

Only one further gene showed mating type specific regulation, TR_121284, with more than 200-fold downregulation in MAT1-1 in both strain combinations. This gene is related to the *Saccharomyces cerevisiae* N-terminal amidase NTA1 (domain accession: cd07566; p-value 4.99e−141), which functions in the N-end rule protein degradation pathway. Hence, *T. reesei* NTA1 may be involved in mating type specific regulation of protein stability. RTqPCR confirmed the MAT1-2 specific regulation of *nta1* under conditions of sexual development, asexual development and growth on cellulose in liquid culture (Fig. 2). Since transcript levels of nta1 in MAT1-1 were under the detection level of our assay, the difference between MAT1-1 and MAT1-2 is at least 50000fold and signals for MAT1-1 in our transcriptome data can be considered background.

Female sterility related gene expression under different conditions

In order to assess a more general defect in QM6a, we compared gene regulation in QM6a with that in the female fertile strains of both mating types under conditions of sexual development with differential regulation in the same strains (female fertile vs. QM6a), but upon growth on cellulose [10]. Principal component (PCA) analysis of these gene sets shows that QM6a patterns are closely related to those of FF1 and FF2 strains, which have largely the same background, and are distinct from CBS999.97, both
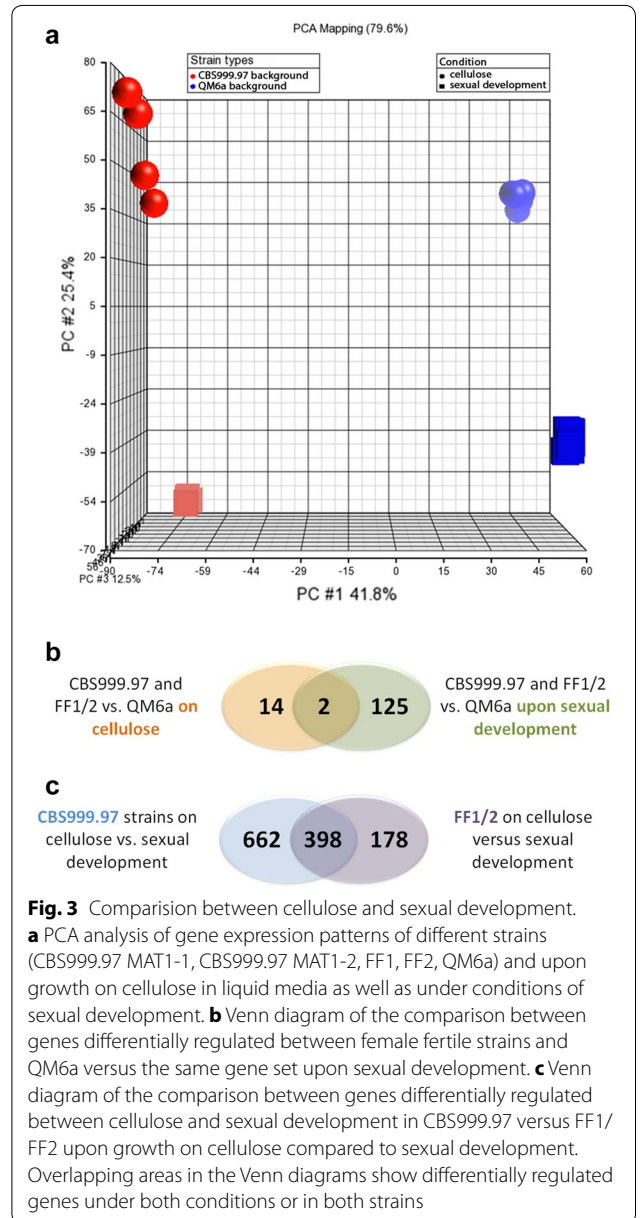
with respect to growth on cellulose and under conditions of sexual development (Fig. 3a). 16 genes were differentially regulated between female fertile strains and QM6a on cellulose and 127 under conditions of sexual development (Fig. 3b). Only two genes showed significant regulation in this comparison, TR_35534 (roughly three-fold upregulation), a gene potentially involved in diterpene metabolism and TR_105242 (6- to 16-fold upregulation in female fertile strains), a putative SAM dependent methyltransferase. Therefore we designated TR_105242 as FFR1 (female fertility related 1) and deleted the encoding gene in FF1b in order to assess its function in sexual development

Fig. 4 Relevance of FFR1 for sexual development. **a** Strains were crossed under standard conditions and showed normal sexual development. **b** If *ffr1* is lacking in both mating partners, fruiting body formation is delayed and fruiting bodies of this cross are not yet mature after 18 days. After 25 days, ascospore discharge commences as in the wildtype in the Δ*ffr1* crosses. Microscopic pictures show 400x magnification

(Fig. 4). Δ*ffr1* showed normal fruiting body formation and ascospore discharge in crosses with wild-type or QM6a (Fig. 4a). Upon crossing of Δ*ffr1* MAT1-1 with Δ*ffr1* MAT1-2, we found that fruiting body formation was delayed by 2 days and ascospore discharge by 3 days. After 18 days we found that the fruiting bodies of this cross were less mature than the wild-type, while after 25 days strains this difference was not visible anymore (Fig. 4b). Hence, although the absence of *ffr1* from the genome is relevant for sexual development, *ffr1* is neither essential for mating nor for female fertility.

Gene regulation specific for sexual development

The availability of datasets of the same strains under conditions of sexual development and cellulase gene expression enabled us to narrow down the gene set specific for sexual development. Genes with highly regulated expression under sexual development conditions compared to cellulase conditions are expected to be more specific for sexual development. We compared gene regulation in CBS999.97 strains upon growth on cellulose versus under sexual development conditions and made the same comparison separately for FF1/FF2 strains to consider the different strain backgrounds and increase the significance of results (Additional file 3). 388 genes were more than fivefold (p-value threshold 0.01) differentially regulated under these conditions in CBS999.97 and FF1/FF2 strains (Fig. 3c). Of those, 168 were up-regulated on cellulose versus sexual development and 220 were down-regulated (Additional file 3).

No significant functional enrichment was detected in the gene set downregulated upon sexual development (up on cellulose, 168 genes), with the exeption of an enrichment in unclassified proteins (p-value 6.80e−09). However, this gene set contains 10 CAZyme encoding genes including *rgx1* (rhamnogalacturonase) and *xyn5* (xylanase; [30]), three genes involved in asexual development, the cellulose specific gene *ooc1* [31], the hydrophobin genes *hfb2*, *hfb3* and *hfb5*, the ceratoplatanin encoding gene *epl1*, involved in elicitation of plant responses [32], as well as the polyketide synthase encoding *pks4* gene (Additional file 3).

The gene set of upregulated genes upon sexual development (down on cellulose) is enriched in functions in metabolism (p-value 3.88e−07), particularly nitrogen, sulphur and selenium metabolism (p-value 3.34e−05), secondary metabolism (p-value 1.47e−13), C-compound and carbohydrate transport, amino acid transport and peptide transport (p-values below 3.5e−03), electron transport (p-value 8.61e−05) and fruit body development (p-value 4.15e−03) (Additional file 3). Twenty CAZyme encoding genes including several alpha- and beta-glycosidases and a chitinase were in this gene set as well as 8 PTH11 like G-protein coupled receptor encoding genes, 8 genes involved in secondary metabolism including several polyketide synthases, and 9 transcription factor genes. Interestingly, the whole SOR cluster with exception of the transcription factor gene *ypr2* [23, 33], which was recently found to be responsible for biosynthesis of the sorbicillin compounds trichodimerol and dihydrotrichotetronine [23], is upregulated on sexual development compared to growth on cellulose with fold regulations of around 20- up to 70-fold (Additional file 3).

Ten genes showed contrasting regulation between sexual development specific genes in CBS999.97 and strains

with QM6a background. They include a candidate alpha xylosidase and *cel3d*, the alcohol oxidase gene *aox1* as well as two transporter genes (Additional file 3).

Resource distribution specific to sexual development

Increased transcript levels under a certain condition represent the first step to high level expression of the respective genes i.e. biosynthesis of the gene products. This investment of resources can be considered a preparation to exert the associated functions if translation and processing continues. However, a considerable number of highly expressed genes are unspecific housekeeping genes. Therefore we selected the 1000 genes with the highest transcript levels of CBS999.97 upon growth under conditions of sexual development and removed those that are among the 1000 most highly transcribed ones on cellulose. From this gene set we selected those genes that fulfilled the same criteria in FF1/2. Ninety four genes remained, which are likely to represent the most strongly expressed genes under sexual development conditions (Additional file 4). This gene set contains 5 CAZyme encoding genes, 2 PTH11-type G-protein coupled receptors, a protein phosphatase, 7 transcription factors and three transporters (Additional file 4). Among the transcription factor genes, a homologue of the *N. crassa* grainy head like transcription factor encoding *csp-2*, which is involved in conidial separation, development and cell wall remodeling [34], was found. However, the most interesting finding was the high level transcription of the three biosynthetic genes of the recently described SOR cluster [23, 33, 35] with TR_73618/*sor2/pks10s*, TR_73621/*sor1/pks11s* and TR73623/*sor5* being among the 10 genes with highest overall transcript levels in CBS999.97. We therefore tested transcript abundance of *sor1* under conditions of sexual development (contact stage, before fruiting body formation) compared to cellulose (Fig. 5a). The strong overexpression of *sor1* upon growth under conditions favouring sexual development compared to liquid culture on cellulose was confirmed. However, testing transcript abundance in the absence of a mating partner showed similarly high transcript levels (Fig. 5a). Hence, for *sor1* a specific significance for sensing of a mating partner is not supported, albeit a relevance for growth under conditions favouring sexual development cannot be excluded.

Analysis of the function of the SOR cluster in sexual development

As the biosynthetic genes of the SOR cluster show particularly high transcript levels, we were interested in its relevance for sexual development. Also the corresponding up-regulation obtained for CBS999.97 and FF1/2 strains upon sexual development supports an importance for sexual development.

This cluster is responsible for light modulated production of the yellow compounds trichodimerol and dihydrotrichotetronin [23]. For regulators, similar expression levels as the regulated genes have been reported previously in *T. reesei* [20, 36] and were expected also in this case. Interestingly, the most important regulators of the cluster, *ypr1* and *ypr2* [33], are not among those with the highest transcript levels in CBS999.97, although their expression level upon sexual development is still higher than on cellulose. Additionally, differential regulation between CBS999.97 and FF1/FF2 occurs, which was not the case for the SOR biosynthetic genes. Therefore we consider it possible that other transcription factors contribute to regulation of the SOR cluster under conditions of sexual development. In order to identify candidates for such a function we performed a coregulation analysis. Coregulated genes are enriched in aromate metabolism (p-value 3.07e−03) and secondary metabolism (p-value 2.72e−04), supporting the hypothesis that secondary metabolism is highly important under sexual development conditions. We found several transcription factor encoding genes with a similar regulation pattern as TR_73618/*sor2/pks10s*, TR_73621/*sor1/pks11s* and TR73623/*sor5* with consistently high regulation in CBS999.97 and FF1/FF2 under sexual development versus low levels on cellulose. However, none of them reached comparably high transcript levels as these genes. Promising candidates for a contribution to regulation of *sor1*, *sor2* and *sor5* upon sexual development are TR_71823, TR_1941, TR_60761, TR_56141 and TR_3449.

Due to the high expression levels of *sor1*, *sor2* and *sor5*, we checked whether deletion of these genes, which abolishes or strongly reduces production of trichodimerol and dihydrotrichotetronin [23], would be essential for sexual development. Crossings with mutants in these genes showed that *sor1*, *sor2* and *sor5* are not essential for fruiting body formation and ascospore discharge of *T. reesei* (Fig. 5b). However, we observed that in the absence of *sor5* fruiting body formation is somewhat delayed (Fig. 5b). Hence neither these genes nor their biosynthetic products are essential for sexual development in *T. reesei*, but may have a beneficial influence.

Discussion

Since the discovery of sexual development in *T. reesei*, the environmental conditions supporting this process as well as gene regulation required for mating to happen are subject of ongoing investigations [14]. Here we studied gene regulation patterns under conditions of sexual development in strains of different genetic background and compared them to the female sterile isolate QM6a. While a number of interesting targets for investigating female
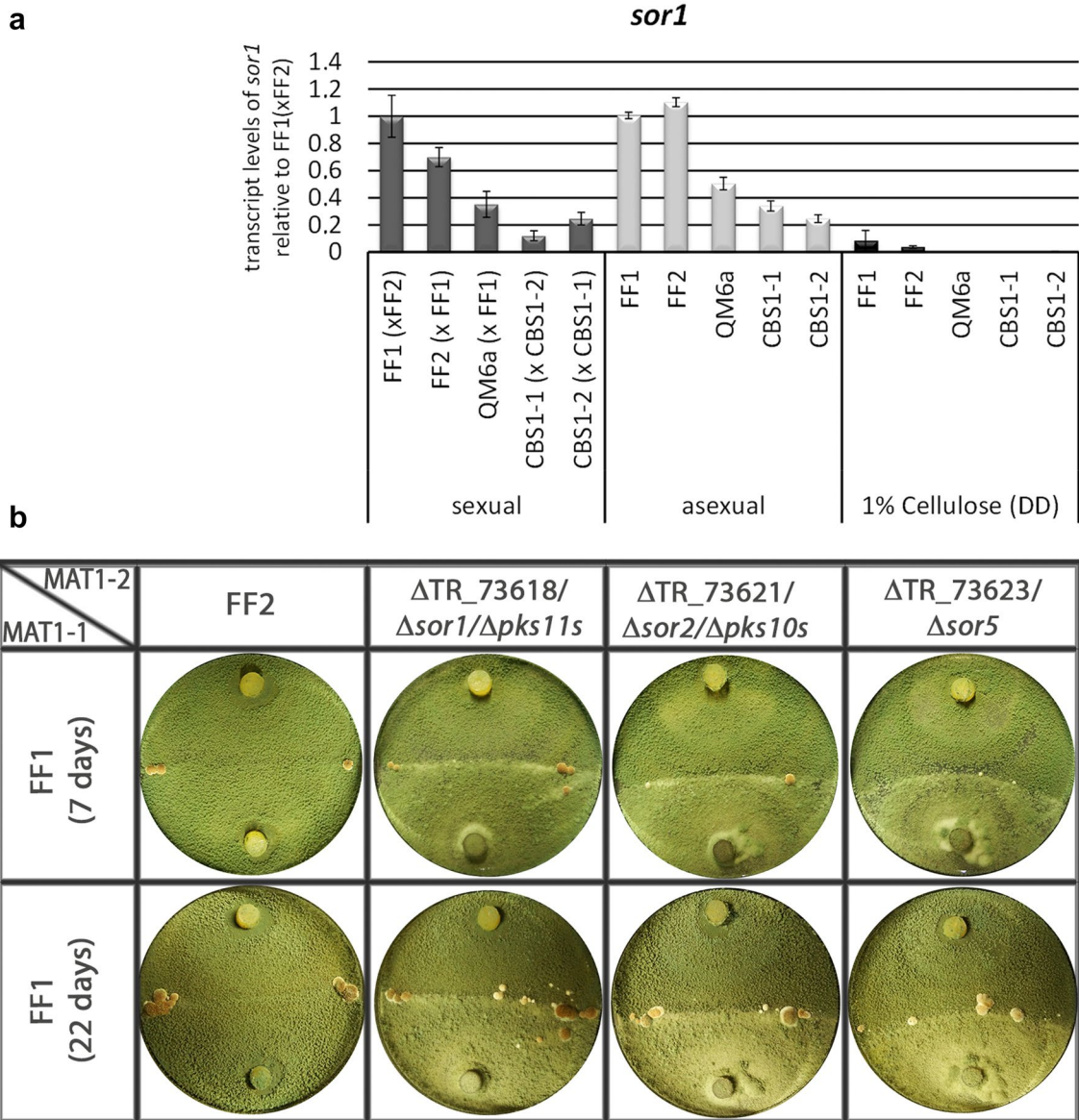
Fig. 5 Relevance of SOR cluster genes for sexual development. **a** Transcript levels of *sor1* under conditions of sexual or asexual development or on cellulose. Strains were grown on malt extract agar for investigation of sexual development (contact stage, before fruiting body formation) and asexual development (strain alone on the plate). For liquid culture on cellulose (1% w/v), strains were grown in minimal medium in constant darkness for 72 h. Errorbars show standard deviations. **b** Analysis of sexual development in strains lacking biosynthetic genes of the SOR cluster. Strains grown in light cycles at 22 °C are shown after 7 or 22 days. Fruiting body formation started after 7 days, no changes were observed after 22 days. Ascospore formation was indistinguishable from wild-type

sterility of QM6a emerged from this study, the most striking finding was the differential regulation of genes involved in secondary metabolism in several of our evaluations. Hence, our analysis indicated an importance of secondary metabolism for sexual development both due to a strong enrichment in most highly expressed genes upon sexual development as well as consistent elevated transcript levels in a comparison of sexual development conditions with growth on cellulose.

Connections between development and secondary metabolism have been subject to intense research in fungi for decades [37–39] with the velvet family of proteins as important connecting factors [40]. Additionally, the *A. nidulans* transcription factor NsdD [41, 42] and its homologues, like SUB-1 in *N. crassa* [43], emerged as regulators of secondary metabolism and development [44–46].

In *T. reesei*, the light dependent transcription factor SUB1, a homologue of NsdD, was only recently shown to be required for female fertility and plays a role in regulation of the pheromone system and secondary metabolism [47]. SUB1 thereby influences abundance of a product of the SOR cluster, trichodimerol [23] in darkness, but not in light upon growth on cellulose [47]. In general, deletion of *sub1* causes altered abundance of secondary metabolites under different growth conditions, including sexual development albeit the nature of these metabolites remains to be determined in most cases [47].

LAE1 is a further known regulator of secondary metabolism and required for formation of the yellow pigment in *T. reesei* and at the same time essential for cellulase gene expression [48]. Therefore it can be assumed that LAE1 targets the SOR cluster as well. In *A. nidulans*, lack of laeA negatively influences sexual development, and the defect becomes even more severe if the *T. reesei* homologue is expressed in such a strain [49]. Deletion of *lae1* abolishes sporulation [48] and although a function in mating is likely, it is not yet known whether LAE1 influences sexual development in *T. reesei* as well.

Functions in both development and secondary metabolism have also been shown for VELVET, which controls the ratio of sexual/asexual development in response to light in *Aspergillus nidulans* [50]. However, VeA only has a small effect on the SUB1 homologue NsdD [50]. Also in *T. reesei*, the VeA homologue VEL1 has functions in sexual development as well as in secondary metabolism [22]. There, we could show a specific change in secondary metabolite patterns if a mating partner was present on the same plate. Consequently, VEL1 regulates chemical communication in *T. reesei* [22]. The nature of the secondary metabolites crucial for this communication is not yet known.

Our study revealed a considerable importance of genes involved in secondary metabolism during sexual development in *T. reesei*. Interestingly, the products of the SOR cluster, with the associated genes being highly expressed under conditions of sexual development in *T. reesei*, appear to be not essential for sexual development, because deletion of the biosynthetic genes shows that they are not required for mating (Fig. 5). Consequently, it is also unlikely that the sorbicillin derivatives biosynthesized by the SOR cluster enzymes are essential for development associated chemical communication.

The importance of the high level transcription of the SOR cluster genes upon sexual development can therefore be rather attributed to a protective mechanism, which can protect fruiting bodies against predators or competitors as suggested previously [51]. Our results are in agreement with earlier data showing that deletion of polyketide synthase genes does not abolish sexual development [52].

In our analysis, effects of the carbon source as well as the difference between cultivation on solid or in liquid media have to be kept in mind. However, available transcriptome data do not support a regulation of the SOR cluster genes here other than specific to development: *sor1*, *sor2* and *sor5* are up-regulated upon growth on cellulose and glucose compared to glycerol, lactose or sophorose [24]. Therefore the specific up-regulation under mating conditions cannot be attributed simply to an altered carbon source, because expression levels on carbon sources other than cellulose are similar or lower than on cellulose [24], but not strongly elevated as the results for sexual development shows.

Upon growth on cellulose, the SOR cluster genes are down-regulated upon growth in light compared to darkness [23, 53]. In contrast, in our study, these genes are strongly up-regulated under conditions of sexual development (light, malt extract medium) compared to growth on cellulose (darkness). We conclude that indeed conditions of sexual development, but not merely illumination or a specific carbon source are responsible for the elevated transcript levels of *sor1*, *sor2* and *sor5*.

Conclusions

In summary we found that generally, the strain background (CBS999.97 versus QM6a) is more relevant for gene regulation than the mating type. While in different strain backgrounds a number of genes are regulated according to mating type, consistent regulation in the different strains and hence in general in T. reesei appears limited to the pheromone system and mating type genes. Our data support the role of secondary metabolism for chemical communication as postulated earlier. Hence the interrelationship between secondary metabolism and sexual development warrants further investigation.

Methods

Strains and cultivation conditions

QM6a (ATCC13631; [2]), CBS999.97 MAT1-1, CBS999.97 MAT1-2 [6], FF1 and FF2 [22, 24] along with sister strains from different crossing lines of FF1 and FF2, (FF1a, FF1b, FF2a, FF2b) were used in this study. FF1 and FF2 were prepared by backcrossing the female fertile CBS999.97 MAT1-1 with female sterile QM6a 10 times in order to acquire sexual competence while retaining the QM6a phenotype [22]. For testing the influence of SOR1, SOR2 and SOR5 on sexual development, the respective deletion strains [23] were used (Table 1). Strains were propagated on malt extract medium. For inoculum preparation, strains were grown in constant darkness for 10 days, thereby avoiding an influence of random light pulses or circadian rhythmicity on gene regulation.

Table 1 Strains used in this study

Strain	Code	Characteristics	Source/reference
CBS999.97 MAT1-1	CBS1-1	Wild-type MAT1-1, female fertile	[6]
CBS999.97 MAT1-2	CBS1-2	Wild-type MAT1-2, female fertile	[6]
FF1a, FF1b	FF1	Backcrossed wild-type strain MAT1-1	[22]
FF2a, FF1b	FF2	Backcrossed wild-type strain MAT1-2	[22]
QM6a		Wildt-type MAT1-2, female sterile	[2]
QM6a Δsor1		Δsor1Δku80::hph + MAT1-2	[23]
QM6a Δsor2		Δsor2Δku80::hph + MAT1-2	[23]
QM6a Δsor5		Δsor5Δku80::hph + MAT1-2	[23]
FF1 Δffr1		Δffr1Δku80::hph + MAT1-1	This study
FF2 Δffr1		Δffr1Δku80::hph + MAT1-2	This study

Crossings were done on malt extract medium (2% w/v) in light–dark cycles at 22 °C as described previously [6].

For transcriptome analysis, strains were grown under conditions facilitating sexual development as described previously [6, 22]. Strains with similar genetic background were combined in crosses (CBS999.97 MAT1-1 × CBS999.97 MAT1-2; QM6a MAT1-2 × QM10a MAT1-1 [6]; FF1a × FF2a; FF1b × FF2b). 40 plates per combination were used, strains were harvested separately at subjective noon at the contact stage, separated in two groups per combination and pooled resulting in two biological replicates based on 20 individual plates each. Contamination of samples by the other strain of different mating type on the same plate was tested as described previously [15] and was generally below 0.1%.

Transcriptome analysis
Isolation of total RNA and quality control using the Bio-Rad Experion system (Hercules, CA, USA) was done as described previously [54]. The quality threshold for use of samples in transcriptome analysis was set to a RIN (RNA integrity number) value of 9. The NimbleGen (Madison, WI, USA) gene expression full service was used as described previously [10, 21] using custom arrays for QM6a and CBS999.97. Data are available at NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE89104.

Bioinformatic analysis was performed using the Partek Genomics Suite 6.5 (Partek Inc., St. Louis, MO, USA), which applies ANOVA analysis for identification of statistically significant gene regulation. Datasets were grouped according to the specific scientific question and treated as replicates (for example, all strains with the CBS999.97 background were compared to all strains with the QM6a background).

Hierarchical clustering analysis and analysis of expression patterns was performed using the open source software HCE3.5 [55]. The online analysis platform at MIPS (http://mips.helmholtz-muenchen.de/funcatDB/) [56] was used for functional category analysis of gene sets with its latest version of May 2014. The p-values shown with this analysis indicate the extent of significant enrichment of a given gene group within a gene set of regulated genes.

Annotation of genes listed in additional files was done using the comparative genome study on *T. reesei*, *T. atroviride* and *T. virens* [3] and complemented by data provided in [57].

Quantitative reverse transcription PCR
RTqPCR was performed as described previously [23, 54] using *rpl6e* and *sar1* as reference genes, which were shown to be appropriate for the conditions we used here [47]. For analysis of *sor1* primers RT_73621_F (5′ GCAACCTCGTCGATTTGGCTGC 3′) and RT_73621_R (5′ AAGTGTCTCGAGAAGGACGCGC 3′) [23] and for *nta1*, primers 121284RTq1F (5′ ACTCTC ATGCTGAATGTTCAC 3′) and 121284RTq1R (5′ TGGAGGCAGAGTAGCTCAC 3′) were used. Data were evaluated using the CFX Maestro software (Bio-Rad, Hercules, USA).

Gene deletion
The gene encoding TR_105242, *ffr1*, was deleted in the female fertile strain FF1b. Therefore, the vector pDEL105242 was constructed by yeast recombination cloning as described in [58] for selection using hygromycin. Transformation and selection of deletion strains was performed by protoplasting as described previously [59]. The 5′ flanking region was amplified using primers 105242_5F (5′ GTAACGCCAGGGTTTTCCCAGTC ACGACGGCGTAGGCTACTCAGTCTGC 3′) and 105242_5R (5′ ATCCACTTAACGTTACTGAAATCTCC AACATCCTGTGTCACTCCTATCC 3′) and the 3′ flanking region was amplified using primers 105242_3F (5′ CTCCTTCAATATCATCTTCTGTCTCCGACATATGG AGGTCGAGGAAACC 3′) and 105242_3R (5′ GCGGA TAACAATTTCACACAGGAAACAGCCTCCGAGTT GCAATAGTAGC 3′). Removal of the open reading frame was confirmed by PCR using primers 105242_qF (5′ ATTCGCACGACCACTCTCAC 3′) and 105242_qR (5′ CGCCATGCTTGGAGATTGTG 3′).

Author details
[1] Center for Health and Bioresources, AIT Austrian Institute of Technology GmbH, Konrad Lorenz Straße 24, 3430 Tulln, Austria. [2] Institute of Chemical Engineering, Research Area Molecular Biotechnology, TU Wien, 1060 Vienna, Austria.

References
1. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. Microb Cell Fact. 2016;15:106.
2. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol. 2008;26:553–60.
3. Schmoll M, Dattenböck C, Carreras-Villasenor N, Mendoza-Mendoza A, Tisch D, Aleman MI, Baker SE, Brown C, Cervantes-Badillo MG, Cetz-Chel J, et al. The genomes of three uneven siblings: footprints of the lifestyles of three *Trichoderma* species. Microbiol Mol Biol Rev. 2016;80:205–327.
4. Kuhls K, Lieckfeldt E, Samuels GJ, Kovacs W, Meyer W, Petrini O, Gams W, Borner T, Kubicek CP. Molecular evidence that the asexual industrial fungus *Trichoderma reesei* is a clonal derivative of the ascomycete *Hypocrea jecorina*. Proc Natl Acad Sci USA. 1996;93:7755–60.
5. Schmoll M. Sexual development in *Trichoderma*—scrutinizing the aspired phenomenon. In: Mukherjee PK, Horwitz BA, Singh US, Mukherjee M, Schmoll M, editors. Trichoderma: biology and applications. Wallingford: CAB International; 2013. p. 67–86.
6. Seidl V, Seibel C, Kubicek CP, Schmoll M. Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci USA. 2009;106:13909–14.
7. Jamet-Vierny C, Debuchy R, Prigent M, Silar P. IDC1, a pezizomycotina-specific gene that belongs to the PaMpk1 MAP kinase transduction cascade of the filamentous fungus *Podospora anserina*. Fungal Genet Biol. 2007;44:1219–30.
8. Jonkers W, Leeder AC, Ansong C, Wang Y, Yang F, Starr TL, Camp DG 2nd, Smith RD, Glass NL. HAM-5 functions as a MAP kinase scaffold during cell fusion in *Neurospora crassa*. PLoS Genet. 2014;10:e1004783.
9. Linke R, Thallinger GG, Haarmann T, Eidner J, Schreiter M, Lorenz P, Seiboth B, Kubicek CP. Restoration of female fertility in *Trichoderma reesei* QM6a provides the basis for inbreeding in this industrial cellulase producing fungus. Biotechnol Biofuels. 2015;8:155.
10. Tisch D, Pomraning KR, Collett JR, Freitag M, Baker SE, Chen CL, Hsu PW, Chuang YC, Schuster A, Dattenböck C, et al. Omics analyses of *Trichoderma reesei* CBS999.97 and QM6a indicate the relevance of female fertility to carbohydrate-active enzyme and transporter levels. Appl Environ Microbiol. 2017;83:e01578-17.
11. Ni M, Feretzaki M, Sun S, Wang X, Heitman J. Sex in fungi. Annu Rev Genet. 2011;45:405–30.
12. Debuchy R, Berteaux-Lecellier V, Silar P. Mating systems and sexual morphogenesis in ascomycetes. In: Borkovich KA, Ebbole DJ, editors. Cellular and molecular biology of filamentous fungi. Washington: ASM Press; 2010. p. 501–35.
13. Moore-Landecker E. Physiology and biochemistry of ascocarp induction and deveopment. Mycol Res. 1992;96:705–16.
14. Schmoll M, Wang TF. Sexual development in trichoderma. In: Wendland J, editor. The Mycota, (vol I): growth, differentiation and sexuality. Cham: Springer; 2016. p. 457–74.
15. Seibel C, Tisch D, Kubicek CP, Schmoll M. The role of pheromone receptors for communication and mating in *Hypocrea jecorina* (*Trichoderma reesei*). Fungal Genet Biol. 2012;49:814–24.
16. Schmoll M, Seibel C, Tisch D, Dorrer M, Kubicek CP. A novel class of peptide pheromone precursors in ascomycetous fungi. Mol Microbiol. 2010;77:1483–501.
17. Seibel C, Tisch D, Kubicek CP, Schmoll M. ENVOY is a major determinant in regulation of sexual development in *Hypocrea jecorina* (*Trichoderma reesei*). Eukaryot Cell. 2012;11:885–90.
18. Bazafkan H, Dattenbock C, Stappler E, Beier S, Schmoll M. Interrelationships of VEL1 and ENV1 in light response and development in *Trichoderma reesei*. PLoS ONE. 2017;12:e0175946.
19. Chen CL, Kuo HC, Tung SY, Hsu PW, Wang CL, Seibel C, Schmoll M, Chen RS, Wang TF. Blue light acts as a double-edged sword in regulating sexual development of *Hypocrea jecorina* (*Trichoderma reesei*). PLoS ONE. 2012;7:e44969.
20. Schuster A, Tisch D, Seidl-Seiboth V, Kubicek CP, Schmoll M. Roles of protein kinase A and adenylate cyclase in light-modulated cellulase regulation in *Trichoderma reesei*. Appl Environ Microbiol. 2012;78:2168–78.
21. Tisch D, Kubicek CP, Schmoll M. The phosducin-like protein PhLP1 impacts regulation of glycoside hydrolases and light response in *Trichoderma reesei*. BMC Genom. 2011;12:613.
22. Bazafkan H, Dattenböck C, Böhmdorfer S, Tisch D, Stappler E, Schmoll M. Mating type dependent partner sensing as mediated by VEL1 in *Trichoderma reesei*. Mol Microbiol. 2015;96:1103–18.
23. Monroy AA, Stappler E, Schuster A, Sulyok M, Schmoll M. A CRE1-regulated cluster is responsible for light dependent production of dihydrotrichotetronin in *Trichoderma reesei*. PLoS ONE. 2017;12:e0182530.
24. Stappler E, Dattenböck C, Tisch D, Schmoll M. Analysis of light- and carbon-specific transcriptomes implicates a class of G-protein-coupled receptors in cellulose sensing. mSphere. 2017;2:e00089-00017.
25. Atanasova L, Knox BP, Kubicek CP, Druzhinina IS, Baker SE. The polyketide synthase gene *pks4* of *Trichoderma reesei* provides pigmentation and stress resistance. Eukaryot Cell. 2013;12:1499–508.
26. Lieckfeldt E, Kullnig CM, Samuels GJ, Kubicek CP. Sexually competent, sucrose- and nitrate-assimilating strains of *Hypocrea jecorina* (*Trichoderma reesei*) from South American soils. Mycologia. 2000;92:374–80.
27. Bok JW, Keller NP. LaeA, a regulator of secondary metabolism in *Aspergillus* spp. Eukaryot Cell. 2004;3:527–35.
28. Karimi Aghcheh R, Druzhinina IS, Kubicek CP. The putative protein methyltransferase LAE1 of *Trichoderma atroviride* is a key regulator of asexual development and mycoparasitism. PLoS ONE. 2013;8:e67144.
29. Wu D, Oide S, Zhang N, Choi MY, Turgeon BG. ChLae1 and ChVel1 regulate T-toxin production, virulence, oxidative stress response, and development of the maize pathogen *Cochliobolus heterostrophus*. PLoS Pathog. 2012;8:e1002542.
30. Herold S, Bischof R, Metz B, Seiboth B, Kubicek CP. Xylanase gene transcription in *Trichoderma reesei* is triggered by different inducers representing different hemicellulosic pentose polymers. Eukaryot Cell. 2013;12:390–8.
31. Schmoll M, Kubicek CP. *ooc1*, a unique gene expressed only during growth of *Hypocrea jecorina* (anamorph: *Trichoderma reesei*) on cellulose. Curr Genet. 2005;48:126–33.
32. Seidl V, Marchetti M, Schandl R, Allmaier G, Kubicek CP. Epl1, the major secreted protein of *Hypocrea atroviridis* on glucose, is a member of a strongly conserved protein family comprising plant defense response eliciters. FEBS J. 2006;273:4346–59.

33. Derntl C, Rassinger A, Srebotnik E, Mach RL, Mach-Aigner AR. Identification of the main regulator responsible for synthesis of the typical yellow pigment produced by *Trichoderma reesei*. Appl Environ Microbiol. 2016;82:6247–57.

34. Pare A, Kim M, Juarez MT, Brody S, McGinnis W. The functions of grainy head-like proteins in animals and fungi and the evolution of apical extracellular barriers. PLoS ONE. 2012;7:e36254.

35. Druzhinina IS, Kubicek EM, Kubicek CP. Several steps of lateral gene transfer followed by events of 'birth-and-death' evolution shaped a fungal sorbicillinoid biosynthetic gene cluster. BMC Evol Biol. 2016;16:269.

36. Portnoy T, Margeot A, Seidl-Seiboth V, Le Crom S, Ben Chaabane F, Linke R, Seiboth B, Kubicek CP. Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. Eukaryot Cell. 2011;10:262–71.

37. Amare MG, Keller NP. Molecular mechanisms of *Aspergillus flavus* secondary metabolism and development. Fungal Genet Biol. 2014;66:11–8.

38. Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon NJ, Keller NP, Yu JH, Braus GH. VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science. 2008;320:1504–6.

39. Macheleidt J, Mattern DJ, Fischer J, Netzker T, Weber J, Schroeckh V, Valiante V, Brakhage AA. Regulation and role of fungal secondary metabolites. Annu Rev Genet. 2016;50:371–92.

40. Bayram O, Braus GH. Coordination of secondary metabolism and development in fungi: the velvet family of regulatory proteins. FEMS Microbiol Rev. 2012;36:1–24.

41. Han KH, Han KY, Yu JH, Chae KS, Jahng KY, Han DM. The nsdD gene encodes a putative GATA-type transcription factor necessary for sexual development of *Aspergillus nidulans*. Mol Microbiol. 2001;41:299–309.

42. Lee MK, Kwon NJ, Choi JM, Lee IS, Jung S, Yu JH. NsdD is a key repressor of asexual development in *Aspergillus nidulans*. Genetics. 2014;197:159–73.

43. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. Proc Natl Acad Sci USA. 2006;103:10352–7.

44. Schumacher J, Simon A, Cohrs KC, Viaud M, Tudzynski P. The transcription factor BcLTF1 regulates virulence and light responses in the necrotrophic plant pathogen *Botrytis cinerea*. PLoS Genet. 2014;10:e1004040.

45. Niehaus EM, Schumacher J, Burkhardt I, Rabe P, Spitzer E, Munsterkotter M, Guldener U, Sieber CMK, Dickschat JS, Tudzynski B. The GATA-type transcription factor Csm1 regulates conidiation and secondary metabolism in *Fusarium fujikuroi*. Front Microbiol. 2017;8:1175.

46. Cary JW, Harris-Coward PY, Ehrlich KC, Mack BM, Kale SP, Larey C, Calvo AM. NsdC and NsdD affect *Aspergillus flavus* morphogenesis and aflatoxin production. Eukaryot Cell. 2012;11:1104–11.

47. Bazafkan H, Beier S, Stappler E, Böhmdorfer S, Oberlerchner JT, Sulyok M, Schmoll M. SUB1 has photoreceptor dependent and independent functions in sexual development and secondary metabolism in Trichoderma reesei. Mol Microbiol. 2017;106(5):742–59.

48. Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP. The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol. 2012;84:1150–64.

49. Karimi-Aghcheh R, Bok JW, Phatale PA, Smith KM, Baker SE, Lichius A, Omann M, Zeilinger S, Seiboth B, Rhee C, et al. Functional analyses of *Trichoderma reesei* LAE1 reveal conserved and contrasting roles of this regulator. G3 (Bethesda). 2013;3:369–78.

50. Kato N, Brooks W, Calvo AM. The expression of sterigmatocystin and penicillin genes in *Aspergillus nidulans* is controlled by veA, a gene required for sexual development. Eukaryot Cell. 2003;2:1178–86.

51. Calvo AM, Cary JW. Association of fungal secondary metabolism and sclerotial biology. Front Microbiol. 2015;6:62.

52. Gaffoor I, Brown DW, Plattner R, Proctor RH, Qi W, Trail F. Functional analysis of the polyketide synthase genes in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). Eukaryot Cell. 2005;4:1926–33.

53. Tisch D, Schmoll M. Targets of light signalling in *Trichoderma reesei*. BMC Genom. 2013;14:657.

54. Tisch D, Kubicek CP, Schmoll M. New insights into the mechanism of light modulated signaling by heterotrimeric G-proteins: ENVOY acts on *gna1* and *gna3* and adjusts cAMP levels in *Trichoderma reesei* (*Hypocrea jecorina*). Fungal Genet Biol. 2011;48:631–40.

55. Seo J, Gordish-Dressman H, Hoffman EP. An interactive power analysis tool for microarray hypothesis testing and generation. Bioinformatics. 2006;22:808–14.

56. Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 2004;32:5539–45.

57. Druzhinina IS, Kopchinskiy AG, Kubicek EM, Kubicek CP. A complete annotation of the chromosomes of the cellulase producer *Trichoderma reesei* provides insights in gene clusters, their expression and reveals genes required for fitness. Biotechnol Biofuels. 2016;9:75.

58. Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP, Schmoll M. A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. Biotechnol Biofuels. 2012;5:1.

59. Gruber F, Visser J, Kubicek CP, de Graaff LH. The development of a heterologous transformation system for the cellulolytic fungus *Trichoderma reesei* based on a *pyrG*-negative mutant strain. Curr Genet. 1990;18:71–6.

Evolutionary freedom in the regulation of the conserved itaconate cluster by Ria1 in related Ustilaginaceae

Elena Geiser[1,2†] [id], Hamed Hosseinpour Tehrani[1†], Svenja Meyer[1], Lars M. Blank[1*] and Nick Wierckx[1]

Abstract

Background: Itaconate is getting growing biotechnological significance, due to its use as a platform compound for the production of bio-based polymers, chemicals, and novel fuels. Currently, *Aspergillus terreus* is used for its industrial production. The Ustilaginaceae family of smut fungi, especially *Ustilago maydis*, has gained biotechnological interest, due to its ability to naturally produce this dicarboxylic acid. The unicellular, non-filamentous growth form makes these fungi promising alternative candidates for itaconate production. Itaconate production was also observed in other Ustilaginaceae species such as *U. cynodontis*, *U. xerochloae*, and *U. vetiveriae*. The investigated species and strains varied in a range of 0–8 g L^{-1} itaconate. The genes responsible for itaconate biosynthesis are not known for these strains and therefore not characterized to explain this variability.

Results: Itaconate production of 13 strains from 7 species known as itaconate producers among the family Ustilaginaceae were further characterized. The sequences of the gene cluster for itaconate synthesis were analyzed by a complete genome sequencing and comparison to the annotated itaconate cluster of *U. maydis*. Additionally, the phylogenetic relationship and inter-species transferability of the itaconate cluster transcription factor Ria1 was investigated in detail. Doing so, itaconate production could be activated or enhanced by overexpression of Ria1 originating from a related species, showing their narrow phylogenetic relatedness.

Conclusion: Itaconate production by Ustilaginaceae species can be considerably increased by changing gene cluster regulation by overexpression of the Ria1 protein, thus contributing to the industrial application of these fungi for the biotechnological production of this valuable biomass derived chemical.

Keywords: Activation of silent cluster, (S)-2-hydroxyparaconate, (S)-2-hydroxyparaconic acid, Itaconic acid, Itatartarate, Secondary metabolites, Transcription factors, Basidiomycota, *Ustilago maydis*

Background

Secondary metabolites are organic, naturally produced, bioactive compounds with a low molecular weight, that are produced by fungi, bacteria, and plants via pathways not belonging to the primary metabolism of this organism [1, 2]. In 2000, a literature survey identified more than

23,000 already discovered secondary metabolites mainly from the fungal kingdom [1, 3]. Closely related species usually produce related compounds and each compound is produced in a highly-narrowed taxonomy [2, 4]. Genes coding for the biosynthesis of secondary metabolites are usually co-localized in a gene cluster with a size of approximately over 10,000 bp depending on the complexity of the metabolite and regions of non-coding base pairs of up to 2000 bp between the coding genes [2, 5, 6]. In cases of polyketide synthases these regions are more extended [7]. These clusters contain genes coding for corresponding biosynthesis enzymes and transporters, regulatory proteins like transcription factors, and optionally modifying

---

*Correspondence: lars.blank@rwth-aachen.de
†Elena Geiser and Hamed Hosseinpour Tehrani have contributed equally to this work
[1] iAMB – Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany
Full list of author information is available at the end of the article

enzymes. Secondary metabolite clusters are often controlled by a complex regulatory network [8]. Several levels of regulation exist, which allow the organism to respond to various environmental influences. Transcription of these clusters can be regulated either by specific/narrow-domain or by global/broad-domain transcription factors or regulators or a combination thereof. Alternatively, regulation can be chromatin-mediated by histone acetylation or methylation [8].

Itaconate and its lactone (*S*)-2-hydroxyparaconate are examples of secondary metabolites. Itaconate is produced by fungi like *Aspergillus terreus* and *Ustilago maydis*, but also by less well-known Ustilaginaceae species, such as *Ustilago cynodontis, Ustilago vetiveriae,* and *Ustilago xerochloae* [9–12]. Itaconate has industrial applications as a co-monomer, for example in the production of acrylonitrile–butadiene–styrene and acrylate latexes in the paper and architectural coating industries [13]. According to an independent evaluation report of the U.S. Department of Energy (DoE) in 2004 [14], itaconate was assigned to be among the top 12 building blocks with

a high biotechnological potential, enabling a conversion into a range of new interesting molecules such as 2-or 3-methyltetrahydrofuran with applications as novel biofuels [15, 16]. Recent studies showed that genes for the biosynthesis of itaconate are co-localized in the genome and co-regulated in *U. maydis* [17], and therefore fulfilling the main criteria to be a secondary metabolite. *U. maydis*' itaconate cluster (GenBank: KT852988.1) contains two itaconate biosynthesis genes *UMAG_tad1* and *UMAG_adi1* encoding a *trans*-aconitate decarboxylase (Tad1) and an aconitate-Δ-isomerase (Adi1), and two transporter genes *UMAG_itp1* and *UMAG_mtt1* encoding an itaconate transport protein (Itp1) and a mitochondrial tricarboxylate transporter (Mtt1), respectively (Fig. 1). Their expression is co-regulated by the transcriptional regulator Ria1, also encoded in this cluster, which is considered as an itaconate cluster specific/narrow domain transcription factor, triggering the transcription of the itaconate biosynthesis genes [17]. Overexpression of *UMAG_ria1* upregulated the expression of biosynthesis core-cluster genes and transporters [17].
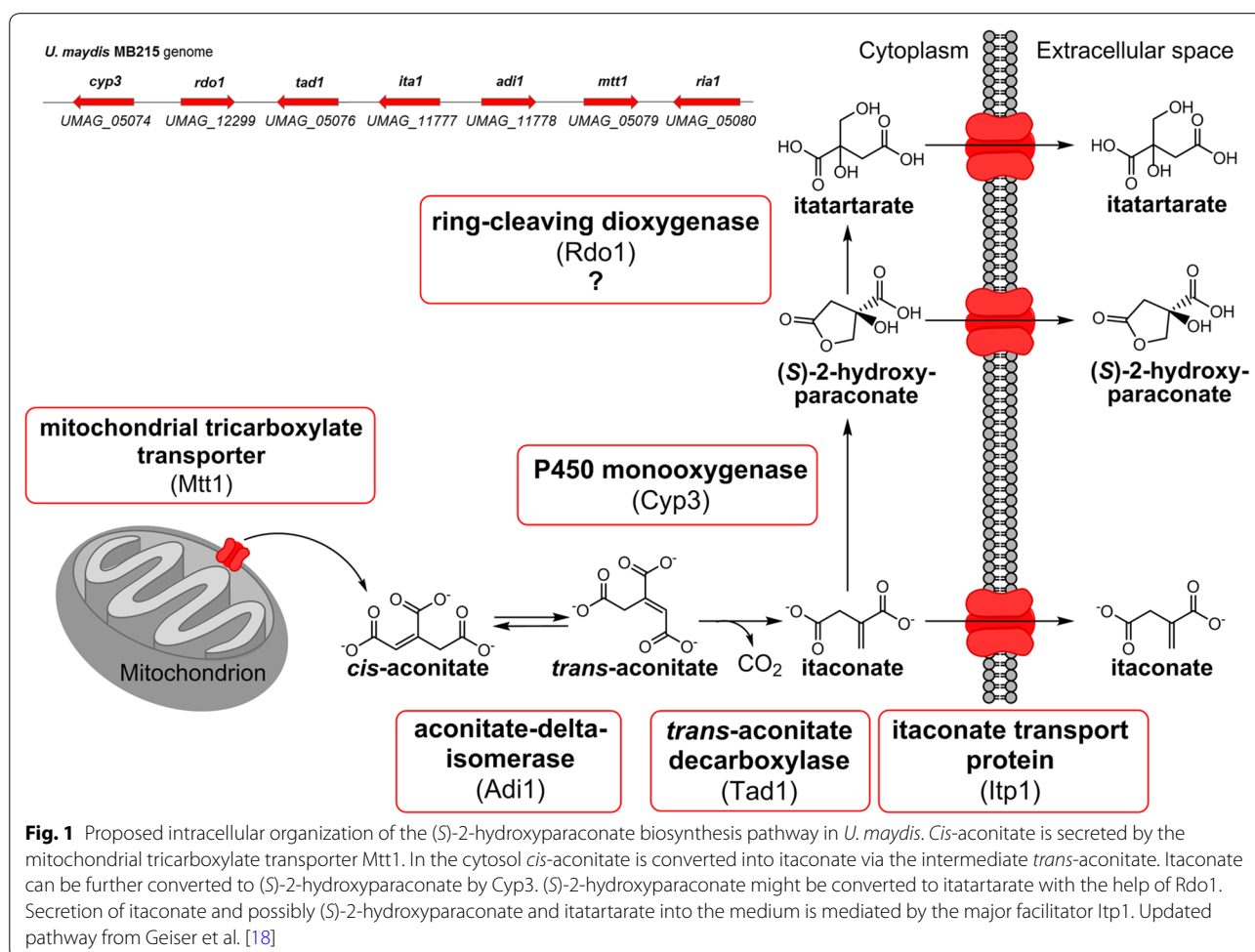

Fig. 1 Proposed intracellular organization of the (*S*)-2-hydroxyparaconate biosynthesis pathway in *U. maydis*. *Cis*-aconitate is secreted by the mitochondrial tricarboxylate transporter Mtt1. In the cytosol *cis*-aconitate is converted into itaconate via the intermediate *trans*-aconitate. Itaconate can be further converted to (*S*)-2-hydroxyparaconate by Cyp3. (*S*)-2-hydroxyparaconate might be converted to itatartarate with the help of Rdo1. Secretion of itaconate and possibly (*S*)-2-hydroxyparaconate and itatartarate into the medium is mediated by the major facilitator Itp1. Updated pathway from Geiser et al. [18]

Additionally, the (*S*)-2-hydroxyparaconate biosynthesis gene *UMAG_cyp3* encoding the cytochrome P450 family 3 monooxygenase Cyp3 and *UMAG_rdo1* encoding a putative ring cleaving dioxygenase are adjacent to the itaconate gene cluster of *U. maydis*, the former of which converts itaconate to (*S*)-2-hydroxyparaconate [18]. Further, it was reported by Guevarra and Tabuchi that (*S*)-2-hydroxyparaconate is converted to itatartarate by a lactonase [11, 19]. *UMAG_cyp3* and *UMAG_rdo1* are not part of the core cluster and not directly upregulated by overexpression of *UMAG_ria1* [17]. However, all itaconate cluster genes including the two adjacent to the core cluster, *UMAG_cyp3* and *UMAG_rdo1*, are strongly upregulated during teliospore formation in the late biotrophic growth stage during plant colonization [20–22].

The number of so far undiscovered secondary metabolites produced by enzymes encoded by cryptic or orphan gene clusters are innumerably high [23, 24]. However, the availability of numerous whole fungal genome sequences and in silico gene prediction by bioinformatic algorithms, such as SMURF [25], MiBiG [26], antiSMASH [27], and FungiFun [28], allow the identification of these cryptic gene clusters. These bioinformatic tools enable 'genome mining' via comparison of protein sequence and structure homology. Traditional ways of activating the expression of secondary metabolites clusters include the variation in the cultivation conditions, such as medium, pH, temperature, aeration, or light, or co-cultivation with other microbes to simulate the natural expression conditions [8]. Often these physiological or ecological triggers are not sufficient to activate these clusters, and therefore several strategies have been developed to induce undiscovered silent secondary metabolite cluster [8]. The most prominent strategies are genetic engineering approaches: the overexpression of a cluster-specific transcription factor gene allowing the increased expression of the whole cluster [24]. In this case, the overexpression of the enzymes encoded within the cluster leads to diverse products, a potential challenge for natural product production [18, 24]. In addition, the endogenous promoters of secondary metabolism biosynthesis genes can be exchanged for strong inducible or constitutive promoters or global regulators can be overexpressed or deleted. A prominent example of the activation of a silent gene cluster is the overexpression of the transcriptional regulator gene *apdR* in *Aspergillus nidulans*, which induced the expression of all cluster genes, leading to the discovery of the cytotoxic aspyridones [24].

Itaconate production is naturally induced by nitrogen limitation in *U. maydis* and was also observed in other related Ustilaginaceae species such as *U. cynodontis*, *U. xerochloae*, *U. vetiveriae* that show high potential to be promising and effective itaconate producers [10, 12, 29].

However, the investigated species and strains varied in their product spectra and the amount of secreted product. Among the species, individual strains of *U. maydis* differed highly in their itaconate and (*S*)-2-hydroxyparaconate production [10]. Some of the species investigated, for example *U. vetiveriae* strain CBS 131474, produced itaconate or (*S*)-2-hydroxyparaconate only with glycerol as carbon source. Also, itaconate production varied depending on extracellular pH. While in wild type *U. maydis* itaconate production is only possible in the pH-range of 5–7, *U. cynodontis* strains also produce itaconate at pH values below 3. The genes responsible for itaconate biosynthesis and how they are regulated to explain this variation in production levels and environmental inputs are not known for these specific strains.

In the current study, 13 itaconate producers of the Ustilaginaceae family were further characterized towards their itaconate cluster sequence-function relationship. The itaconate gene clusters of these strains were identified by genome sequencing [29] and comparison to the annotated itaconate cluster of *U. maydis* strain MB215. To explore the evolutionary conservation of regulation of the itaconate cluster in respect to itaconate production by members of the Ustilaginaceae family, the phylogenetic relationship and inter-species transferability of the itaconate cluster transcription factor Ria1 was investigated. Itaconate production could be activated or enhanced by overexpression of Ria1 originating from related species. This is the first time that activation of silent itaconate clusters by overexpression of a cluster-specific transcription factor in Ustilaginaceae species other than *U. maydis* is shown.

Results and discussion
Variation in itaconate and (S)-2-hydroxyparaconate production among Ustilaginaceae

Previous studies showed a high variation in natural itaconate production among related Ustilaginaceae species cultured on glucose and glycerol as carbon sources [10, 12]. Besides their varying amounts of product and product spectrum, they also differed in their efficiency of carbon utilization. Some of the species produced itaconate only on a single carbon source like glycerol or glucose. These differences motivated us to investigate itaconate and derivates production on glucose and glycerol in more detail (Fig. 2, Additional file 1: Fig. S1). *U. maydis* Δ*Umag_ria1* was used as a negative control, since the transcriptional regulator gene *ria1* is deleted and therefore itaconate production abolished. In *U. maydis* strain AB33P5Δ five extracellular proteases are deleted [30]. With these deletions, the strain is well suited for the secretion of heterologous or intrinsic extracellular biomass degrading CAZymes. This strain would be an
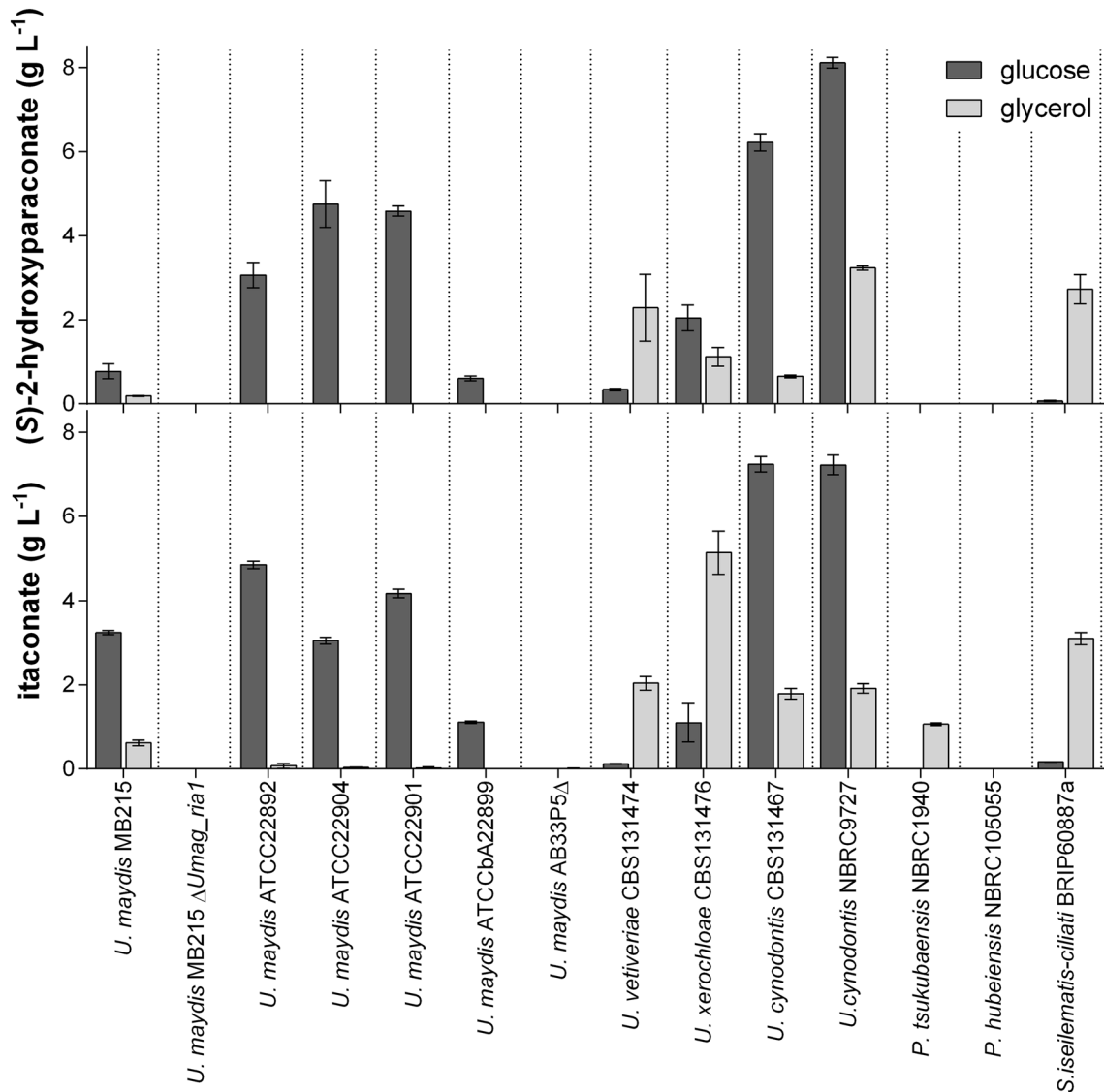
Fig. 2 Itaconate and (*S*)-2-hydroxyparaconate production by various species in the Ustilaginaceae cultivated on glucose and glycerol. Itaconate and (*S*)-2-hydroxyparaconate concentrations after 120 h or 384 h System Duetz® cultivations in screening medium with glucose or glycerol, respectively. The *U. maydis* ΔUmag_ria1 mutant derived from wild type strain MB215 was used as a negative control. Error bars indicate standard deviation from the mean (n = 3)

optimal candidate for the synthesis of itaconate or other valuable chemicals directly from biomass-derived substrates [31]. However, it does not produce itaconate and the lack of extracellular proteases significantly reduces the growth rate of this mutant.

All wild type strains consumed at least 50% of the glucose in the 120 h except of *U. maydis* AB33P5Δ, which utilized 40% (Additional file 2: Fig. S2). The growth on glycerol is slower in comparison to glucose, therefore samples were taken after 384 h. At this time point, all strains consumed at least 30% of the glycerol, except of

U. maydis AB33P5Δ, which used 13% (Additional file 2: Fig. S2). Most *U. maydis* strains produced itaconate only on glucose as the carbon source, whereas *U. vetiveriae, P. tsukubaensis,* and *S. iseilematis-ciliati* did so only on glycerol. *U. cynodontis* and *U. xerochloae* produced itaconate on both carbon sources. *U. maydis* AB33P5Δ, *U. maydis* ΔUmag_ria1, and *P. hubeiensis* did not produce itaconate at all. Since (*S*)-2-hydroxyparaconate and itatartarate are derivatives from itaconate [11], the production of these compounds was also investigated. (*S*)-2-hydroxyparaconate production of the tested strains

on glucose and glycerol was similar to itaconate production (Fig. 2), with the exception of *P. tsukubaensis*, which only produced itaconate. Also, estimated itatartarate production levels showed a similar trend compared to (*S*)-2-hydroxyparaconate production except for *S. iseilematis-ciliati*, which did not to produce itatartarate (Additional file 1: Fig. S1). Previous studies showed a negative correlation between itaconate and malate production [12], therefore malate production was also determined. All strains produced malate on glucose and glycerol except for *U. maydis* AB33P5Δ and *S. iseilematis-ciliati*, which produced malate only on glycerol (Additional file 1: Fig. S1). In general *U. maydis* strains showed the highest malate titers. These results are in accordance with our previous study [10].

A possible reason for these varying titers of itaconate and its derivatives could be differences in the sequences of the itaconate and (*S*)-2-hydroxyparaconate biosynthesis genes, or the genetic inventory of these genes. Furthermore, different regulation or relative expression levels of the biosynthesis genes could cause varying production [1, 8]. Due to the targeted disruption of the genes encoding its five proteases, *U. maydis* strain AB33P5Δ is a slow growing strain in comparison to wild type and other Ustilaginaceae strains, possibly caused by different timing of the strains concerning C- or N-source utilization or their growth rate [32]. To gain a deeper understanding of the sequence-function relationship between itaconate/(*S*)-2-hydroxyparaconate biosynthesis genes and production, the genomes of 13 Ustilaginaceae were analyzed and genes related to synthesis of these secondary metabolites annotated and characterized.
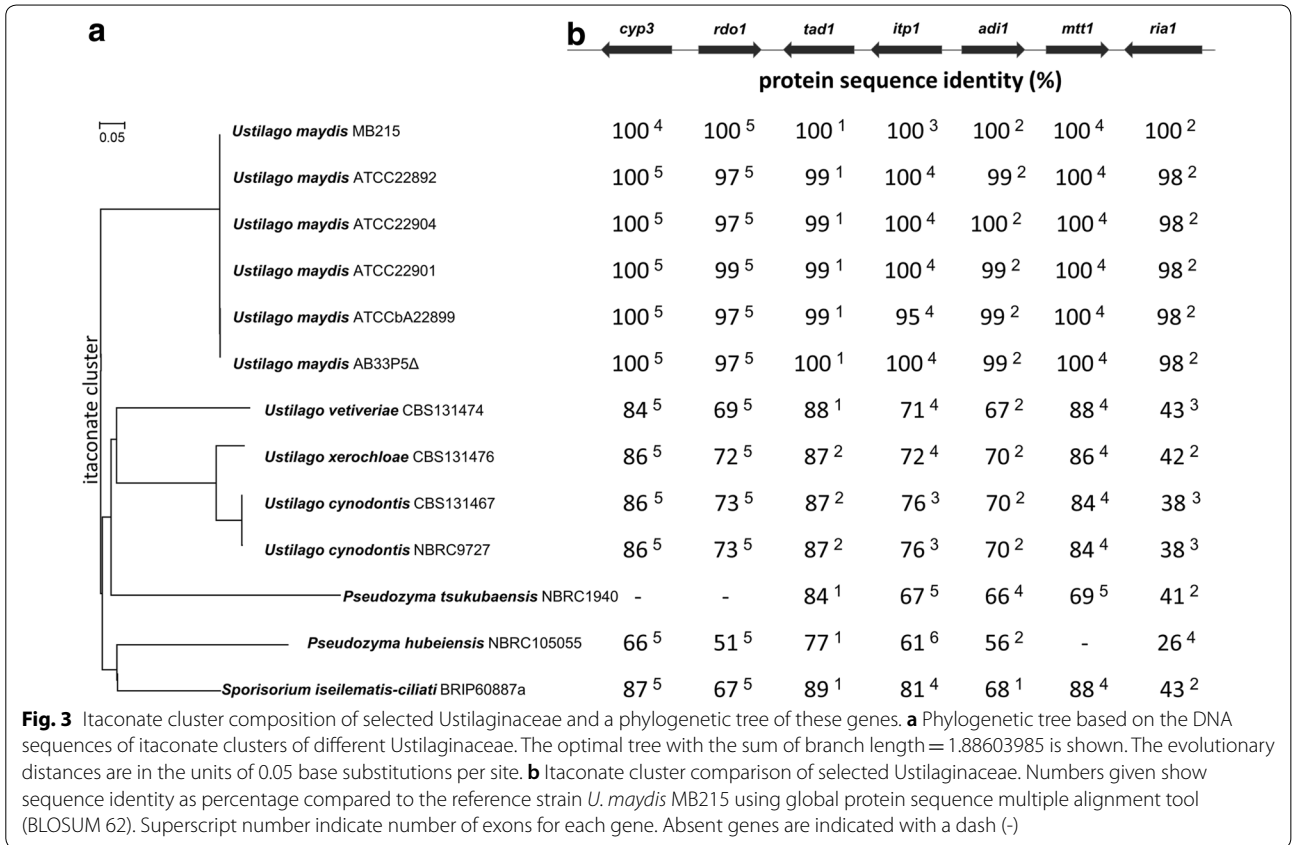
Genetic differences in the itaconate biosynthesis cluster
The Whole Genome Shotgun sequences of *Ustilago maydis* MB215 (DSM17144), *Ustilago maydis* ATCC 22892, *Ustilago maydis* ATCC22904, *Ustilago maydis* ATCC22901, *Ustilago maydis* ATCCbA22899, *Ustilago maydis* AB33P5Δ, *Ustilago vetiveriae* CBS131474, *Ustilago xerochloae* CBS131476, *Ustilago cynodontis* CBS131467, *Ustilago cynodontis* NBRC9727, *Pseudozyma tsukubaensis* NBRC1940, *Pseudozyma hubeiensis* NBRC105055, and *Sporisorium iseilematis-ciliati* BRIP60887a have been deposited in DDBJ/ENA/GenBank [29]. Their accession numbers are listed in "Methods". To find the genes responsible for itaconate and (*S*)-2-hydroxyparaconate biosynthesis in these sequenced strains, the protein sequences encoded in the *U. maydis* MB215 itaconate biosynthesis cluster (GenBank KT852988.1) were used as queries against the Whole Genome Shotgun sequences database using the tBLASTn algorithm [18, 33]. Multiple hits with neighboring genes were defined as putative itaconate clusters.

For cluster annotation, the highest resulting homologous sequences were further analyzed using the online tool "Augustus gene prediction" to identify start/stop codons and exons [34], followed by manual curation. Furthermore, protein sequences of *U. maydis* MB215 were compared to the predicted proteins of the investigated Ustilaginaceae using the global protein sequence multiple alignment tool (BLOSUM 62) [35] in Clone Manager 9 Professional Edition. The protein sequence identity of the investigated Ustilaginaceae proteins compared to the itaconate cluster of reference strain *U. maydis* MB215 is presented in Fig. 3b. Additionally, the phylogenetic tree based on the DNA sequences of itaconate clusters of different Ustilaginaceae indicates the phylogenetic relationship among the chosen strains (Fig. 3a).

Exact phylogenetic classification among Ustilaginaceae is challenging, with several species being renamed based on new analysis of indicator genes such as nuclear ribosomal RNA genes [39–41]. Wang et al. especially mentioned that strains in the genus *Pseudozyma* have an uncertain phylogenetic position due to the taxonomic confusion between their teleomorphic genera [39]. Therefore, a phylogenetic relation is shown based on the DNA sequence of the itaconate cluster (Fig. 3a).
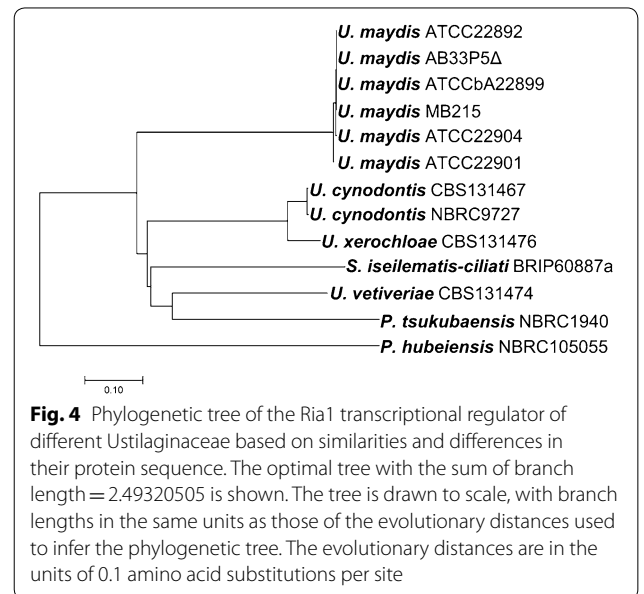
In all sequenced organisms except *P. tsukubaensis* and *P. hubeiensis*, the complete gene cluster for itaconate synthesis and conserved synteny of all genes (gene orientation and chromosome) were identified (Fig. 3b). The cluster in *P. tsukubaensis* does not contain *rdo1* and *cyp3*. For these two genes no likely homologous candidate was found elsewhere in genome, explaining the lack of (*S*)-2-hydroxyparaconate and itatartarate production in this strain (Fig. 2). In *P. hubeiensis*, *mtt1* is not present in the itaconate cluster or its direct surrounding DNA regions. In *U. maydis* MB215, deletion of *UMAG_mtt1* led to a strong decrease in itaconate production [17]. This transporter, which putatively shuttles malate and *cis*-aconitate between the mitochondria and the cytoplasm, is the rate-limiting step in itaconate biosynthesis in *U. maydis* MB215 [42]. Since itaconate formation was not completely abolished by *Umag_mtt1* deletion in *U. maydis*, most likely other less specialized, and therefore less efficient, transport proteins substituted its function, as most eukaryotic mitochondrial transporters have a diverse substrate spectrum with different affinities [43]. At least one similar mitochondria tricarboxylate transporter gene is present in the genome of *P. hubeiensis*, which could take over the function of Mtt1. This gene showed 54% sequence similarity on protein level in comparison to *Umag_mtt1* and 98% to *Umag_02365* upon tBLASTn analysis [33]. The latter gene, *Umag_02365*, is known to be one of two related mitochondrial citrate transporters in *U. maydis*, with redundant function to *Umag_mtt1*

	cyp3	rdo1	tad1	itp1	adi1	mtt1	ria1
		protein sequence identity (%)					
Ustilago maydis MB215	100 [4]	100 [5]	100 [1]	100 [3]	100 [2]	100 [4]	100 [2]
Ustilago maydis ATCC22892	100 [5]	97 [5]	99 [1]	100 [4]	99 [2]	100 [4]	98 [2]
Ustilago maydis ATCC22904	100 [5]	97 [5]	99 [1]	100 [4]	100 [2]	100 [4]	98 [2]
Ustilago maydis ATCC22901	100 [5]	99 [5]	99 [1]	100 [4]	99 [2]	100 [4]	98 [2]
Ustilago maydis ATCCbA22899	100 [5]	97 [5]	99 [1]	95 [4]	99 [2]	100 [4]	98 [2]
Ustilago maydis AB33P5Δ	100 [5]	97 [5]	100 [1]	100 [4]	99 [2]	100 [4]	98 [2]
Ustilago vetiveriae CBS131474	84 [5]	69 [5]	88 [1]	71 [4]	67 [2]	88 [4]	43 [3]
Ustilago xerochloae CBS131476	86 [5]	72 [5]	87 [2]	72 [4]	70 [2]	86 [4]	42 [2]
Ustilago cynodontis CBS131467	86 [5]	73 [5]	87 [2]	76 [3]	70 [2]	84 [4]	38 [3]
Ustilago cynodontis NBRC9727	86 [5]	73 [5]	87 [2]	76 [3]	70 [2]	84 [4]	38 [3]
Pseudozyma tsukubaensis NBRC1940	-	-	84 [1]	67 [5]	66 [4]	69 [5]	41 [2]
Pseudozyma hubeiensis NBRC105055	66 [5]	51 [5]	77 [1]	61 [6]	56 [2]	-	26 [4]
Sporisorium iseilematis-ciliati BRIP60887a	87 [5]	67 [5]	89 [1]	81 [4]	68 [1]	88 [4]	43 [2]

Fig. 3 Itaconate cluster composition of selected Ustilaginaceae and a phylogenetic tree of these genes. **a** Phylogenetic tree based on the DNA sequences of itaconate clusters of different Ustilaginaceae. The optimal tree with the sum of branch length = 1.88603985 is shown. The evolutionary distances are in the units of 0.05 base substitutions per site. **b** Itaconate cluster comparison of selected Ustilaginaceae. Numbers given show sequence identity as percentage compared to the reference strain *U. maydis* MB215 using global protein sequence multiple alignment tool (BLOSUM 62). Superscript number indicate number of exons for each gene. Absent genes are indicated with a dash (-)

[42]. This may explain why *P. hubeiensis* failed to produce itaconate.

In general, the conservation of a protein sequence could point to its evolutionary origin. The comparison showed that among the tested *U. maydis* strains the itaconate cluster is conserved. At the DNA level the clusters in different *U. maydis* strains are >98% similar and the clusters of the two *U. cynodontis* strains have 99% sequence identity on DNA level. For the other species, the sequence identity of proteins encoded by the itaconate and (*S*)-2-hydroxyparaconate biosynthesis (*cyp3*, *tad1* and *adi1*) and transporters (*itp1* and *mtt1*) genes were mostly conserved in a range of 56–89% compared to the *U. maydis* MB215 sequence. The most divergent protein of the itaconate cluster is Ria1, a transcription factor of approximately 380 amino acids. The annotated *Uc_ria1* of both *U. cynodontis* strains encode a transcription factor of 471 amino acids. A conserved helix-loop-helix structural motif could be found in all 13 regulators approximately in position 100-AA by SMART analysis, which is a characteristic DNA-binding motif for one of the largest families of dimerizing transcription factors [44, 45]. The phylogenetic tree of the predicted Ria1

transcriptional regulators is shown in Fig. 4. Ria1 proteins of *U. maydis* species are very closely related. *U. cynodontis* and *U. xerochloae* are closely related [46], which is reflected in the relatedness of their Ria1 proteins.

Fig. 4 Phylogenetic tree of the Ria1 transcriptional regulator of different Ustilaginaceae based on similarities and differences in their protein sequence. The optimal tree with the sum of branch length = 2.49320505 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances are in the units of 0.1 amino acid substitutions per site

However, *U. maydis* and *U. vetiveriae* are phylogenetically closely related as well [46], even though their Ria1 proteins are only 43% identical. This may indicate that the amino acid sequence of Ria1 proteins is evolving faster than its actual function, for which just the DNA-binding motif is essential. The Ria1 sequence of *P. hubeiensis* is phylogenetically the most distant of the species compared. In general, no accurate subcategorization of the transcription factor according the species is possible. A reason might be the aforementioned difficulties in categorization of the Ustilaginaceae.

In summary, all tested strains have the genetic inventory for itaconate biosynthesis, and the synteny of the itaconate cluster is preserved in most of the investigated Ustilaginaceae. *P. tsukubaensis* and *P. hubeiensis* do not possess the complete itaconate cluster, partly explaining the differences in product spectrum. However, the variable itaconate and (*S*)-2-hydroxyparaconate titer, especially among different *U. maydis* strains with highly similar clusters, could not be explained. As different regulation or expression levels might be responsible for these production differences, the Ria1 transcriptional regulator of the tested Ustilaginaceae were investigated in more detail.

Inter-species transferability of Ria1 regulator

The itaconate clusters of the tested strains are mostly conserved, while production levels of itaconate differ. As one example, the *U. maydis* AB33P5Δ gene cluster is 98% similar at the DNA level to that of *U. maydis* ATCCbA22899; however, strain AB33P5Δ does not produce itaconate or (*S*)-2-hydroxyparaconate while strain ATCCbA22899 does. Probably in some strains, like *U. maydis* AB33P5Δ, *ria1* is functional but not expressed. To test whether production differences are a result of different regulation, the inter-species transferability of Ria1 was investigated by overexpression of various *ria1* genes to activate the production of itaconate. We chose the itaconate cluster regulator genes *Umag_ria1*, *Uc_ria1*, *Pt_ria1*, and *Si_ria1* of *U. maydis* MB215, *U. cynodontis* NBRC9727, *P. tsukubaensis*, and *S. iseilematis-ciliati*, respectively, due to their considerable differences in the sequences of both the itaconate cluster and Ria1. These regulators were expressed under control of the constitutive promoter P_{etef} in *U. maydis* MB215, *U. maydis* AB33P5Δ, *U. vetiveriae*, *U. xerochloae*, *U. cynodontis* CBS131467, *U. cynodontis* NBRC9727, *P. tsukubaensis*, *P. hubeiensis*, and *S. iseilematis-ciliati*, as well as in the control strain *U. maydis* MB215 Δ*Umag_ria1*. Successful integration was verified by PCR.

All strains tested consumed at least 35% of the applied glucose after 120 h and 30% of the applied glycerol after 384 h except of *U. maydis* AB33P5Δ, which used 13%

glycerol (Additional file 2: Fig. S2). A summary of the activation experiments is shown in Fig. 5 and Additional file 3: Fig. S3. The itaconate and (*S*)-2-hydroxyparaconate production yield (gram product per gram substrate) of the activated strains was determined on both glucose and glycerol (Fig. 6) as well as the malate yield and the estimated relative itatartarate production (Additional file 4: Fig. S4). In *U. maydis* MB215 Δ*Umag_ria1*, itaconate production could be restored by expression of all tested regulators (*Umag_ria1*, *Uc_ria1*, *Pt_ria1*, and *Si_ria1*), demonstrating the functionality of this expression system, as well as their transferability of the genes between related species. It should be noted that quantitative differences in production level may be caused by different copy number, or by the random ectopic integration locus, of the integrated regulator, which were not determined in detail. Thus, these results should be viewed mostly in a qualitative manner. In strains that do not produce itaconate on glucose, such as *U. maydis* AB33P5Δ (derivative of *U. maydis* FB1), *U. vetiveriae*, *P. tsukubaensis*, *P. hubeiensis*, and *S. iseilematis-ciliati* itaconate production could be activated by expression of all tested regulators, except *Uc_ria1*. This suggests that in these wild type strains the itaconate cluster genes are silent, because the regulator gene *ria1* is silent and not transcribed. Constitutive expression of the itaconate regulator *ria1*, even originating from different species, activated the expression of the itaconate cluster genes, which resulted in itaconate production.

As already encountered for the wild type strains, (*S*)-2-hydroxyparaconate and itatartarate production correlated with itaconate production. In general, activation or enhancement of itaconate biosynthesis also activated or enhanced (*S*)-2-hydroxyparaconate and itatartarate biosynthesis (Fig. 5, Additional files 3 and 4: Fig. S3 and Fig. S4). An exception is *P. tsukubaensis* that does not possess the (*S*)-2-hydroxyparaconate biosynthesis genes *rdo1* and *cyp3*, and therefore (*S*)-2-hydroxyparaconate and itatartarate are not produced in the activated strains. Zambanini et al. showed a negative correlation of itaconate and malate biosynthesis after overexpression of *Umag_ria1* in *U. vetiveriae* CBS131474 on glycerol [12]. This is in line with our results. However, for the other tested Ustilaginaceae this negative correlation could not be shown, as in most activated strains malate production resembled the wild type level (Fig. 5, Additional file 4: Fig. S4).

Comparing the successfully activated or improved strains, those expressing *Uc_ria1* perform considerably less well than strains expressing the other regulators. Deletion of *Uc_ria1* in *U. cynodontis* NBRC9727 completely abolished itaconate production (data not shown), indicating that *Uc_ria1* is essential for itaconate production. However, *U. cynodontis* NBRC9727 Δ*Uc_ria1* could not be complemented by *Uc_ria1* under control of the

Fig. 5 Overview of the influence of overexpression of *Umag_ria1*, *Uc_ria1*, *Pt_ria1* and *Si_ria1*, on itaconate (ITA), (*S*)-2-hydroxyparaconate (HP), and malate (MAL) production. Differences in production were determined after 120 h or 384 h System Duetz® cultivations in screening medium containing either glucose or glycerol as the sole carbon source
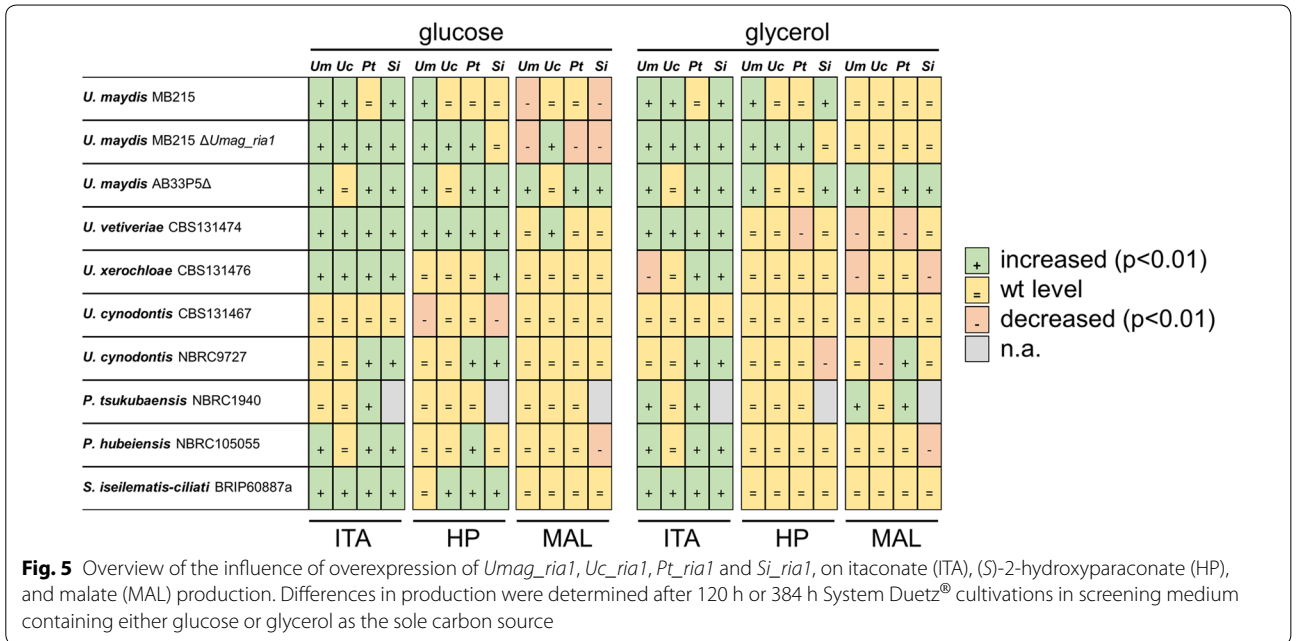
constitutive promoter P$_{etef}$, although complementation experiments under control of the native promotor P$_{Uc\_ria1}$ and terminator T$_{Uc\_ria1}$ and a random genome integration was successfully (data not shown). This indicates that the integration locus of genes under control of P$_{etef}$ plays a crucial role for heterologous expression in *U. cynodontis* and that the chosen expression cassette design may have affected the outcome of *ria1* overexpression in different hosts. In *U. cynodontis* strains itaconate production could not be considerably enhanced, even by overexpression of the native regulator *Uc_ria1*. In contrast, strains more closely related phylogenetically with a lower wild type production level, such as *U. xerochloae*, could still enhance itaconate production by overexpression of *Uc_ria1*. Since the *U. cynodontis* strains were the best performing wild types, it might be possible that the natural expression level of the cluster genes is already at a high level, and the rate limiting factor lies upstream of the itaconate production pathway. Alternatively, induction by Ria1 is in *U. cynodontis* already at its maximum.

Although the itaconate production of *P. tsukubaensis* and *S. iseilematis-ciliati* strains is comparatively low, the regulators *Pt_ria1* and *Si_ria1* seem to be the most universally applicable, since they improved itaconate production in 80% of the tested strains when cultured on glucose or glycerol. Therefore, *Pt_ria1* and *Si_ria1* might open new possibilities to activate itaconate production in other species through heterologous gene expression approaches.

In general, the differences in itaconate production in strains expressing the same regulator could have several explanations. The chosen constitutive promoter P$_{etef}$ is a modified *tef* promoter controlling transcription of the gene for the translation elongation factor 2 of *U. maydis* [47]. It may be less efficient in other Ustilaginaceae than in *U. maydis*. However, its functionality was verified in *U. trichophora* [48] and *U. vetiveriae* [12]. As mentioned before, different copy numbers of the integrated regulators can cause differences in transcription levels and therefore in production levels. Especially for results on glycerol showing an overall similar trend than on glucose, different growth kinetics, including, growth rates, and substrate uptake rates can cause differences in itaconate production. The growth rate on glycerol of Ustilaginaceae is lower in comparison to that on glucose, hence less nitrogen for biomass synthesis per time is required, which subsequently influences nitrogen limitation during cultivation. Nitrogen limitation is necessary for natural induction of itaconate production in *U. maydis* [49]. The maximum theoretical yield of itaconate production is directly related to the consumed C/N ratio, and thus poor growth (low growth rate) could result in a lower yield given the chosen cultivation time. Altogether, itaconate production could be activated or enhanced by overexpression of Ria1 originating from a related species, even though the chosen Ria1 protein sequences are very dissimilar. This is the first time that activation of silent itaconate clusters by overexpression of a cluster-specific transcription factor across species and even genus boundaries was shown.

Since overexpression of Ria1 upregulates all genes of the itaconate core cluster in *U. maydis* MB215 [17],
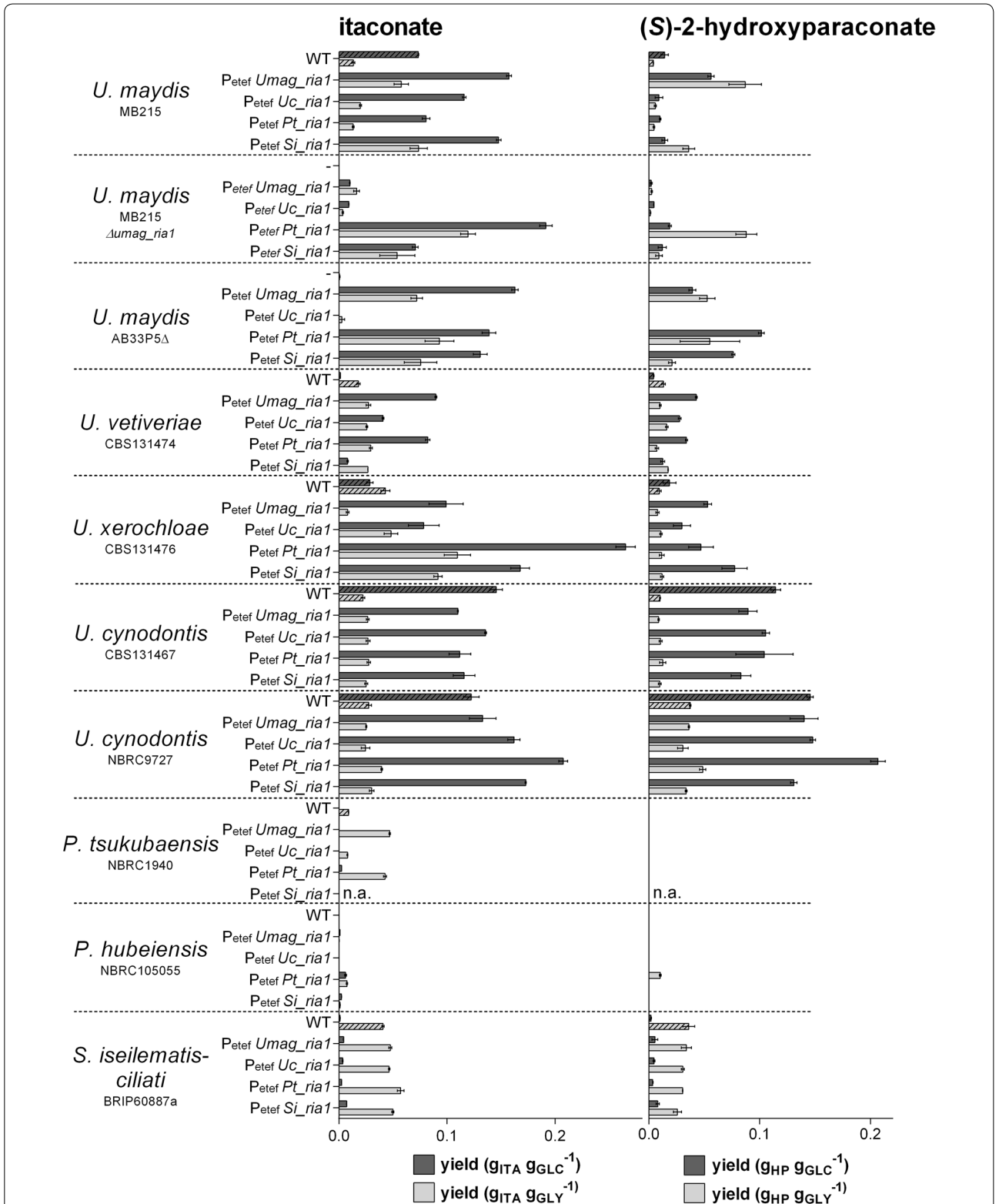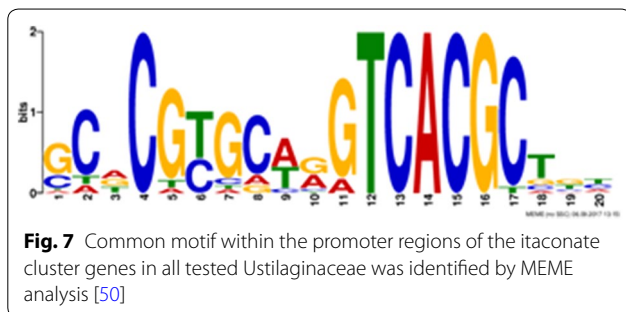
Fig. 6 Itaconate and (S)-2-hydroxyparaconate production by various *Ustilaginaceae* species and their mutants transformed with *Umag_ria1*, *Uc_ria1*, *Pt_ria1*, *Si_ria1*. Itaconate (g_{ITA} g_{GLC}^{-1}, g_{ITA} g_{GLY}^{-1}) and (S)-2-hydroxyparaconate (g_{HP} g_{GLC}^{-1} g_{HP} g_{GLY}^{-1}) yield after 120 h or 384 h System Duetz® cultivations in screening medium containing glucose (GLC) and glycerol (GLY), respectively. A dash (–) indicates the negative control without an overexpression construct. Error bars indicate standard deviation from the mean (n = 3)

Fig. 7 Common motif within the promoter regions of the itaconate cluster genes in all tested Ustilaginaceae was identified by MEME analysis [50]

promoter regions of the co-regulated genes will likely have a common conserved regulator binding domain. In this study the phylogenetic relatedness and the feasible inter-species transferability of Ria1 regulator originating from related species could be shown. To identify potential common regulatory sequences, in silico analysis for conserved sequence motifs was performed using the MEME algorithm Version 4.12.0 under standard settings [50]. This analysis revealed that promoters of Ria1-regulated genes share a putative conserved Ria1 binding domain with a short consensus sequence (CN[T/C]NNNN[G/A]TCACG[C/T]) (Fig. 7). This sequence can be found in all tested Ustilaginaceae in either orientation in the promoter regions of all annotated cluster genes in at least one copy with an average E-value of 1.8×10^{-83}. Interestingly, none of the *ria1* promoters themselves contain this element. Since most of the tested regulators do not seem to be very species-specific, this site likely binds regulators from multiple species. Although the role of this motif as the binding site for Ria1 needs to be confirmed by biochemical methods, its occurrence in the sequenced wild type strains (*U. maydis* MB215 (DSM17144), *U. vetiveriae* CBS131474, *U. xerochloae* CBS131476, *U. cynodontis* CBS131467, *P. tsukubaensis* NBRC1940 and *P. hubeiensis* NBRC105055, and *S. iseilematis-ciliati* BRIP60887a) strongly suggests that in spite of the relatively low amino acid sequence similarity of Ria1 in these species, the function of this regulator is the same.

Conclusion
This study indicates phenotypically that itaconate production differences among related Ustilaginaceae species are based on different transcriptional regulation of the itaconate cluster genes, governed in turn by the expression level of Ria1. All tested strains have the genetic equipment for itaconate production; also, itaconate non-producers. However, in some strains the itaconate clusters are silent, because the itaconate regulator *ria1* is silent. By overexpression of itaconate cluster-specific transcription factors Ria1 originating from related species, we could activate silent itaconate clusters, even

though the amino acid sequences of Ria1 regulators are relatively dissimilar. In additional to the silent itaconate clusters being activated, itaconate production in weak producers could be enhanced up to 4-fold. Especially, the activated form of *U. maydis* strain AB33P5Δ might be a promising candidate for the combination of biomass degradation and itaconate production in one strain [31]. As such, this study contributes to demonstrating the industrial applicability of Ustilaginaceae for the biotechnological production of itaconate, and also suggests that activation of silent secondary metabolite clusters can be achieved in a range of related species with reduced genetic engineering efforts.

Methods
Strains and culture conditions
All strains used in this work are listed in Table 1.

Ria1 overexpression constructs
To generate the overexpression construct, the backbone of the plasmid P_{etef}-*ria1-cbx* from *Escherichia coli* Top10+ *Petef Umag_ria1-Cbx* was amplified by PCR with the primer pair HT-212 and HT-213 (Table 1, Additional file 5: Table S1). The genes encoding the transcription factors Ria1 from *Pseudozyma tsukubaensis*, *Ustilago cynodontis* NBRC9727, and *Sporisorium iseilematis-ciliati* were amplified by PCR using the primer pairs HT-214/HT-215, HT-218/HT-219, and HT-216/HT-217, respectively (Additional file 5: Table S1). Gibson cloning with backbone and different *ria1* genes was conducted to obtain the plasmids P_{etef}Pt_ria1, P_{etef}Uc_ria1, and P_{etef}Si_ria1, respectively [51]. Enzyme digestion and PCR ensured correct assembly.

Overexpression of Ria1
For random integration of the *ria1* overexpression constructs into the genome of the different Ustilaginaceae, the different plasmids (P_{etef} *Pt_ria1*, P_{etef} *Si_ria1*, P_{etef} *Uc_ria1* and P_{etef} *Umag_ria1*-cbx) were linearized by SspI except of P_{etef} *Pt_ria1*, which was linearized by BsrgI. Integration of the linearized overexpression construct in the different Ustilaginaceae was conducted by protoplasts transformation according to Tsukuda et al. [52]. To confirm plasmid integration, the *ria1* was amplified by PCR using the primer Potef-fwd and Tnos-rev.

Shaking cultures were performed in the System Duetz® (24 well plates) with a filling volume of 1.5 mL (d = 50 mm, n = 300 rpm, T = 30 °C and Φ = 80%) [53]. The screening medium contained 50 g L^{-1} glucose or 100 g L^{-1} glycerol, 0.8 g L^{-1} NH$_4$Cl, 0.2 g L^{-1} MgSO$_4$·7H$_2$O, 0.01 g L^{-1} FeSO$_4$·7H$_2$O, 0.5 g L^{-1} KH$_2$PO$_4$, 1 mL L^{-1} vitamin solution, 1 mL L^{-1} trace element solution, and as buffer 132 g L^{-1} calcium carbonate [10].

Table 1 Strains used in this study

Strain designation	Resistance	Reference/ GenBank Accession number
Escherichia coli Top10 + P*etef* *Umag_ria1-Cbx*	Ampicillin	[54]
Escherichia coli NEB® 5a P*etef* *Uc_ria1-Cbx*	Ampicillin	This study
Escherichia coli NEB® 5a P*etef* *Pt_ria1-Cbx*	Ampicillin	This study
Escherichia coli NEB® 5a P*etef* *Si_ria1-Cbx*	Ampicillin	This study
Ustilago maydis DSM17144 (*Ustilago maydis* MB215)	Wild type	AACP00000000
Ustilago maydis DSM17144 P*etef* *Umag_ria1*	Carboxin	This study
Ustilago maydis DSM17144 P*etef* *Uc_ria1*	Carboxin	This study
Ustilago maydis DSM17144 P*etef* *Pt_ria1*	Carboxin	This study
Ustilago maydis DSM17144 P*etef* *Si_ria1*	Carboxin	This study
Ustilago maydis DSM17144 Δ*Umag_ria1*	Hygromycin	[54]
Ustilago maydis DSM17144 Δ*Umag_ria1* P*etef* *Umag_ria1*	Hygromycin, carboxin	[54]
Ustilago maydis DSM17144 Δ*Umag_ria1* P*etef* *Uc_ria1*	Hygromycin, carboxin	This study
Ustilago maydis DSM17144 Δ*Umag_ria1* P*etef* *Pt_ria1*	Hygromycin, carboxin	This study
Ustilago maydis DSM17144 Δ*Umag_ria1* P*etef* *Si_ria1*	Hygromycin, carboxin	This study
Ustilago maydis ATCC 22892	Wild type	LYOO00000000
Ustilago maydis ATCC22904	Wild type	LZQT00000000
Ustilago maydis ATCC22901	Wild type	LZNJ00000000
Ustilago maydis ATCCbA22899	Wild type	LYZD00000000
Ustilago maydis AB33P5Δ	Wild type	[30] /LZQU00000000
Ustilago maydis AB33P5Δ P*etef* *Umag_ria1*	Carboxin	This study
Ustilago maydis AB33P5Δ P*etef* *Uc_ria1*	Carboxin	This study
Ustilago maydis AB33P5Δ P*etef* *Pt _ria1*	Carboxin	This study
Ustilago maydis AB33P5Δ P*etef* *Si_ria1*	Carboxin	This study
Ustilago vetiveriae CBS131474	Wild type	MAIM00000000
Ustilago vetiveriae CBS131474 P*etef* *Umag_ria1*	Carboxin	[12]
Ustilago vetiveriae CBS131474 P*etef* *Uc_ria1*	Carboxin	This study
Ustilago vetiveriae CBS131474 P*etef* *Pt _ria1*	Carboxin	This study
Ustilago vetiveriae CBS131474 P*etef* *Si_ria1*	Carboxin	This study
Ustilago xerochloae CBS131476	Wild type	MAIN00000000
Ustilago xerochloae CBS131476 P*etef* *Umag_ria1*	Carboxin	This study
Ustilago xerochloae CBS131476 P*etef* *Uc_ria1*	Carboxin	This study
Ustilago xerochloae CBS131476 P*etef* *Pt _ria1*	Carboxin	This study
Ustilago xerochloae CBS131476 P*etef* *Si_ria1*	Carboxin	This study
Ustilago cynodontis CBS131467	Wild type	LZQV00000000
Ustilago cynodontis CBS131467 P*etef* *Umag_ria1*	Carboxin	This study
Ustilago cynodontis CBS131467 P*etef* *Uc_ria1*	Carboxin	This study
Ustilago cynodontis CBS131467 P*etef* *Pt_ria1*	Carboxin	This study
Ustilago cynodontis CBS131467 P*etef* *Si_ria1*	Carboxin	This study
Ustilago cynodontis NBRC9727	Wild type	LZZZ00000000
Ustilago cynodontis NBRC9727 P*etef* *Umag_ria1*	Carboxin	This study
Ustilago cynodontis NBRC9727 P*etef* *Uc_ria1*	Carboxin	This study
Ustilago cynodontis NBRC9727 P*etef* *Pt _ria1*	Carboxin	This study
Ustilago cynodontis NBRC9727 P*etef* *Si_ria1*	Carboxin	This study
Pseudozyma tsukubaensis NBRC1940	Wild type	MAIP00000000
Pseudozyma tsukubaensis NBRC1940 P*etef* *Umag_ria1*	Carboxin	This study

Table 1 (continued)

Strain designation	Resistance	Reference/ GenBank Accession number
Pseudozyma tsukubaensis NBRC1940 P$_{etef}$ *Uc_ria1*	Carboxin	This study
Pseudozyma tsukubaensis NBRC1940 P$_{etef}$ *Pt _ria1*	Carboxin	This study
Pseudozyma hubeiensis NBRC105055	Wild type	MAIO00000000
Pseudozyma hubeiensis NBRC105055 P$_{etef}$ *Umag_ria1*	Carboxin	This study
Pseudozyma hubeiensis NBRC105055 P$_{etef}$ *Uc_ria1*	Carboxin	This study
Pseudozyma hubeiensis NBRC105055 P$_{etef}$ *Pt_ria1*	Carboxin	This study
Pseudozyma hubeiensis NBRC105055 P$_{etef}$ *Si_ria1*	Carboxin	This study
Sporisorium iseilematis-ciliati BRIP60887a	Wild type	MJEU00000000
Sporisorium iseilematis-ciliati BRIP60887a P$_{etef}$ *Umag_ria1*	Carboxin	This study
Sporisorium iseilematis-ciliati BRIP60887a P$_{etef}$ *Uc_ria1*	Carboxin	This study
Sporisorium iseilematis-ciliati BRIP60887a P$_{etef}$ *Pt _ria1*	Carboxin	This study
Sporisorium iseilematis-ciliati BRIP60887a P$_{etef}$ *Si_ria1*	Carboxin	This study

Cultures were parallelly inoculated into multiple plates and for each sample point a complete plate was taken as sacrificial sample in order to ensure continuous oxygenation.

Analytical methods

Cell densities were measured by determining the absorption at 600 nm with a Unico spectrophotometer 1201.

Products and substrates in the supernatants were analyzed in a DIONEX UltiMate 3000 High Performance Liquid Chromatography System (Thermo Scientific, Germany) with an ISERA Metab AAC column 300×7.8 mm column (ISERA, Germany). As solvent 5 mM H_2SO_4 with a flow rate of 0.6 mL min^{-1} and a temperature of 40 °C was used. All samples were filtered with Acrodisc® Syringe Filters (GHP, 0.20 μm, Ø 13 mm). Itaconate, (*S*)-2-hydroxyparaconate, malate, and itatartarate were determined with a DIONEX UltiMate 3000 Variable Wavelength Detector set to 210 nm, glycerol and glucose with a refractive index detector SHODEX RI-101 (Showa Denko Europe GmbH, Germany). Itaconate, malate, (*S*)-2-hydroxyparaconate, glucose, and glycerol were identified via retention time and UV/RI quotient compared to corresponding standards. Synthesized (*S*)-2-hydroxyparaconate (purity ~70%) was used as the HPLC standard for quantification and hence the indicated

(*S*)-2-hydroxyparaconate values should be taken as rough estimates only [18]. Since no standards of itatartarate are commercially available this compound was analyzed relatively based on HPLC peak area (mAU*min) using the UV detector. All values are the arithmetic mean of at three biological replicates. Error bars indicate the standard deviation from the mean. Statistical analysis was performed using unequal variances t test with unilateral distribution (P values < 0.01 were considered significant and indicated in figures with *).

Genome sequencing

Genomic DNA was isolated by standard phenol–chloroform extraction [55]. Eurofins Genomics (Ebersberg, Germany) created the library using the NEBNext® Ultra DNA Library Prep Kit for Illumina® (Art No E7370), and sequenced the library using an llumina HiSeq 2500 machine with TruSeq SBS kit v3 both according to manufacturer's instructions. The sequencing mode was 1x100 and the processing used the HiSeq Control software 2.0.12.0 RTA 1.17.21.3 bcl2fastq-1.8.4. Quality check of the sequence data was performed with FastQC (Version 0.11.2). The SPAdes-3.7.0-Linux pipeline was used for de novo genome assembly of single-read libraries and read error or mismatch correction including BayesHammer, IOnHammer, SPAdes, MismatchCorrector, dipSPAdes, and truSPAdes. The k-mer size was determined to 55

using VelvetOptimiser Version 2.2.5. The Whole Genome Shotgun sequences have been deposited in DDBJ/ENA/GenBank. Their accession numbers are listed in Table 1.

Phylogenetic analyses

The evolutionary history of itaconate cluster DNA sequences was inferred using the Neighbor-Joining method [36] after alignment via ClustalW algorithm with the MEGA 7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets Alignment Explorer. The optimal tree with the sum of branch length = 1.88603985 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method [37] and are in the units of the number of base substitutions per site. The analysis involved 13 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 10874 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [38].

For the phylogenetic tree of Ria1, protein sequences were aligned via ClustalW (codon) algorithm with MEGA 7 [38]. The evolutionary history was inferred using the Neighbor-Joining method [36]. The optimal tree with the sum of branch length = 2.49320505 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [56] and are in the units of the number of amino acid substitutions per site. The analysis involved 13 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 269 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [38].

Additional files

Additional file 1: Fig. S1. Malate and itatartarate production by various Ustilaginaceae on glucose and glycerol. Malate concentrations and itatartarate UV area after 120 h or 384 h System Duetz® cultivations in screening medium with glucose or glycerol, respectively. *U. maydis* MB215 Δ*Umag_ria1* was used as a negative control. Error bars indicate standard deviation from the mean (n = 3).

Additional file 2: Fig. S2. Glucose and glycerol consumption by various Ustilaginaceae. Glucose and glycerol consumption in % after 120 h or 384 h System Duetz® cultivations in screening medium with glucose or glycerol, respectively. Error bars indicate standard deviation from the mean (n = 3).

Additional file 3: Fig. S3. Schematic overview of the influence of overexpression of *Umag_ria1*, *Uc_ria1*, *Pt_ria1* and *Si_ria1*, on itatartarate (**ITT**) production. Differences in itatartarate production were determined after 120 h or 384 h System Duetz® cultivations in screening medium containing glucose or glycerol, respectively.

Additional file 4: Fig. S4. Malate and itatartarate production by various *Ustilaginaceae* and their mutants transformed with *Umag_ria1*, *Uc_ria1*, *Pt_ria1*, *Si_ria1*. Malate (g_{Mal} g_{GLC}^{-1}, g_{ITA} g_{GLY}^{-1}) yield and itatartarate titer after 120 h or 384 h System Duetz® cultivations in screening medium containing glucose (GLC) and glycerol (GLY), respectively. A dash (–) indicates the negative control without an overexpression construct. Error bars indicate standard deviation from the mean (n = 3).

Additional file 5: Table S1. Oligonucleotides used for overexpression constructs.

Authors' contributions

All authors have contributed significantly to the work. EG and HHT contributed equally to this manuscript. NW, LMB conceived the project. EG and HHT designed experiments and analyzed results. EG wrote the manuscript with the help of HHT. HHT and SM generated overexpression strains. EG performed cultivation experiments and analytics. All authors read and approved the final manuscript.

Author details

[1] iAMB – Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany. [2] BioSC, c/o Forschungszentrum Jülich, 52425 Jülich, Germany.

References

1. Brakhage AA, Schroeckh V. Fungal secondary metabolites—strategies to activate silent gene clusters. Fungal Genet Biol. 2011;48:15–22.
2. Keller NP, Turner G, Bennett JW. Fungal secondary metabolism—from biochemistry to genomics. Nat Rev Microbiol. 2005;3:937–47.
3. Lazzarini A, Cavaletti L, Toppo G, Marinelli F. Rare genera of actinomycetes as potential producers of new antibiotics. Antonie Van Leeuwenhoek. 2000;78:399–405.
4. Verpoorte R. Secondary metabolism. In: Verpoorte R, Alfermann AW, editors. Metabolic engineering of plant secondary metabolism. Dordrecht: Springer; 2000. p. 1–29.
5. Smith DJ, Burnham MK, Bull JH, Hodgson JE, Ward JM, Browne P, Brown J, Barton B, Earl AJ, Turner G. Beta-lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. EMBO J. 1990;9:741–7.
6. Trail F, Mahanti N, Rarick M, Mehigh R, Liang SH, Zhou R, Linz JE. Physical and transcriptional map of an aflatoxin gene cluster in *Aspergillus parasiticus* and functional disruption of a gene involved early in the aflatoxin pathway. Appl Environ Microbiol. 1995;61:2665–73.

7. Blažič M, Starcevic A, Lisfi M, Baranasic D, Goranovič D, Fujs Š, Kuščer E, Kosec G, Petković H, Cullum J, et al. Annotation of the modular polyketide synthase and nonribosomal peptide synthetase gene clusters in the genome of *Streptomyces tsukubaensis* NRRL18488. Appl Environ Microbiol. 2012;78:8183–90.

8. Brakhage AA. Regulation of fungal secondary metabolism. Nat Rev Microbiol. 2013;11:21–32.

9. Kinoshita K. Über eine neue *Aspergillus* Art, Asp. *itaconicus* nov. spec. J Plant Res. 1931;45:45–60.

10. Geiser E, Wiebach V, Wierckx N, Blank LM. Prospecting the biodiversity of the fungal family Ustilaginaceae for the production of value-added chemicals. BMC Fung Biol Biotech. 2014;1:2.

11. Guevarra ED, Tabuchi T. Accumulation of itaconic, 2-hydroxyparaconic, itatartaric, and malic acids by strains of the genus *Ustilago*. Agric Biol Chem. 1990;54:2353–8.

12. Zambanini T, Hosseinpour Tehrani H, Geiser E, Merker D, Schleese S, Krabbe J, Buescher JM, Meurer G, Wierckx N, Blank LM. Efficient itaconic acid production from glycerol with *Ustilago vetiveriae* TZ1. Biotechnol Biofuels. 2017;10:131.

13. Okabe M, Lies D, Kanamasa S, Park EY. Biotechnological production of itaconic acid and its biosynthesis in *Aspergillus terreus*. Appl Microbiol Biotechnol. 2009;84:597–606.

14. Werpy T, Petersen G, Aden A, Bozell J, Holladay J, White J, Manheim A, Eliot D, Lasure L, Jones S. Top value added chemicals from biomass. Volume 1—results of screening for potential candidates from sugars and synthesis gas. In: DTIC Document; 2004.

15. Geilen FMA, Engendahl B, Harwardt A, Marquardt W, Klankermayer J, Leitner W. Selective and flexible transformation of biomass-derived platform chemicals by a multifunctional catalytic system. Angew Chem. 2010;122:5642–6.

16. Leitner W, Klankermayer J, Pischinger S, Pitsch H, Kohse-Hoinghaus K. Advanced biofuels and beyond: chemistry solutions for propulsion and production. Angew Chem Int Ed Engl. 2017;56:5412–52.

17. Geiser E, Przybilla SK, Friedrich A, Buckel W, Wierckx N, Blank LM, Bolker M. *Ustilago maydis* produces itaconic acid via the unusual intermediate *trans*-aconitate. Microb Biotechnol. 2016;9:116–26.

18. Geiser E, Przybilla SK, Engel M, Kleineberg W, Büttner L, Sarikaya E, Hartog Td, Klankermayer J, Leitner W, Bölker M, et al. Genetic and biochemical insights into the itaconate pathway of *Ustilago maydis* enable enhanced production. Metab Eng. 2016;38:427–35.

19. Guevarra ED, Tabuchi T. Production of 2-hydroxyparaconic and itatartaric acids by *Ustilago cynodontis* and simple recovery process of the acids. Agric Biol Chem Tokyo. 1990;54:2359–65.

20. Zheng Y, Kief J, Auffarth K, Farfsing JW, Mahlert M, Nieto F, Basse CW. The *Ustilago maydis* Cys_2His_2-type zinc finger transcription factor Mzr1 regulates fungal gene expression during the biotrophic growth stage. Mol Microbiol. 2008;68:1450–70.

21. Tollot M, Assmann D, Becker C, Altmuller J, Dutheil JY, Wegner CE, Kahmann R. The WOPR protein Ros1 Is a master regulator of sporogenesis and late effector gene expression in the maize pathogen *Ustilago maydis*. PLoS Pathog. 2016;12:e1005697.

22. Lanver D, Muller AN, Happel P, Schweizer G, Haas FB, Franitza M, Pellegrin C, Reissmann S, Altmuller J, Rensing SA, et al. The biotrophic development of *Ustilago maydis* studied by RNA-seq analysis. Plant Cell. 2018;30:300–23.

23. Hertweck C. Hidden biosynthetic treasures brought to light. Nat Chem Biol. 2009;5:450–2.

24. Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat Chem Biol. 2007;3:213–7.

25. Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND. SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol. 2010;47:736–41.

26. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, et al. Minimum information about a biosynthetic gene cluster. Nat Chem Biol. 2015;11:625–31.

27. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res. 2011;39:339–46.

28. Priebe S, Linde J, Albrecht D, Guthke R, Brakhage AA. FungiFun: a web-based application for functional categorization of fungal genes and proteins. Fungal Genet Biol. 2011;48:353–8.

29. Geiser E, Ludwig F, Zambanini T, Wierckx N, Blank LM. Draft genome sequences of itaconate-producing Ustilaginaceae. Genome Announc. 2016;4:e01291.

30. Sarkari P, Reindl M, Stock J, Müller O, Kahmann R, Feldbrügge M, Schipper K. Improved expression of single-chain antibodies in *Ustilago maydis*. J Biotechnol. 2014;191:165–75.

31. Geiser E, Reindl M, Blank LM, Feldbrugge M, Wierckx N, Schipper K. Activating intrinsic CAZymes of the smut fungus *Ustilago maydis* for the degradation of plant cell wall components. Appl Environ Microbiol. 2016;82:5174–85.

32. Sanchez S, Chavez A, Forero A, Garcia-Huante Y, Romero A, Sanchez M, Rocha D, Sanchez B, Avalos M, Guzman-Trampe S, et al. Carbon source regulation of antibiotic production. J Antibiot. 2010;63:442–59.

33. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

34. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004;32:309–12.

35. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA. 1992;89:10915–9.

36. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.

37. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA. 2004;101:11030–5.

38. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4.

39. Wang QM, Begerow D, Groenewald M, Liu XZ, Theelen B, Bai FY, Boekhout T. Multigene phylogeny and taxonomic revision of yeasts and related fungi in the Ustilaginomycotina. Stud Mycol. 2015;81:55–83.

40. McTaggart AR, Shivas RG, Boekhout T, Oberwinkler F, Vánky K, Pennycook SR, Begerow D. Mycosarcoma (Ustilaginaceae), a resurrected generic name for corn smut (*Ustilago maydis*) and its close relatives with hypertrophied, tubular sori. IMA Fungus. 2016;7:309–15.

41. McLaughlin D, Spatafora JW. Systematics and evolution. Berlin: Springer; 2014.

42. Geiser E. Itaconic acid production by *Ustilago maydis*, vol. 1. Aachen: Apprimus Verlag; 2015.

43. Palmieri F. The mitochondrial transporter family (SLC25): physiological and pathological implications. Pflugers Arch. 2004;447:689–709.

44. Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH. Structure and function of helix-loop-helix proteins. Biochim Biophys Acta. 1994;1218:129–35.

45. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 2015;43:257–60.

46. Stoll M, Piepenbring M, Begerow D, Oberwinkler F. Molecular phylogeny of *Ustilago* and *Sporisorium* species (Basidiomycota, Ustilaginales) based on internal transcribed spacer (ITS) sequences. Can J Bot. 2003;81:976–84.

47. Spellig T, Bottin A, Kahmann R. Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus *Ustilago maydis*. Mol Genet Genomics. 1996;252:503–9.

48. Zambanini T, Tehrani HH, Geiser E, Sonntag CS, Buescher JM, Meurer G, Wierckx N, Blank LM. Metabolic engineering of *Ustilago trichophora* TZ1 for improved malic acid production. Metab Eng Commun. 2017;4:12–21.

49. Maassen N, Panakova M, Wierckx N, Geiser E, Zimmermann M, Bölker M, Klinner U, Blank LM. Influence of carbon and nitrogen concentration on itaconic acid production by the smut fungus *Ustilago maydis*. Eng Life Sci. 2013;14:129–34.

50. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:202–8.

51. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6:343–5.
52. Tsukuda T, Carleton S, Fotheringham S, Holloman WK. Isolation and characterization of an autonomously replicating sequence from *Ustilago maydis*. Mol Cell Biol. 1988;8:3703–9.
53. Duetz WA, Ruedi L, Hermann R, O'Connor K, Büchs J, Witholt B. Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. Appl Environ Microbiol. 2000;66:2641–6.
54. Przybilla S. Genetische und biochemische Charakterisierung der Itacon-säure-Biosynthese in *Ustilago maydis*. *Dissertation.* Phillips University Marburg, Germany; 2014.
55. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987;162:156–9.
56. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, editors. Evolving genes and proteins. Cambridge: Academic Press; 1965. p. 97–166.

Translocated duplication of a targeted chromosomal segment enhances gene expression at the duplicated site and results in phenotypic changes in *Aspergillus oryzae*

Tadashi Takahashi[*] [ID], Masahiro Ogawa, Atsushi Sato and Yasuji Koyama

Abstract

Background: Translocated chromosomal duplications occur spontaneously in many organisms; segmental duplications of large chromosomal regions are expected to result in phenotypic changes because of gene dosage effects. Therefore, experimentally generated segmental duplications in targeted chromosomal regions can be used to study phenotypic changes and determine the functions of unknown genes in these regions. Previously, we performed tandem duplication of a targeted chromosomal segment in *Aspergillus oryzae*. However, in tandem chromosomal duplication, duplication of chromosomal ends and multiple chromosomal duplication are difficult. In this study, we aimed to generate fungal strains with a translocated duplication or triplication of a targeted chromosomal region via break-induced replication.

Results: Double-strand breaks were introduced into chromosomes of parental strains by treating protoplast cells with I-SceI meganuclease. Subsequently, strains were generated by nonreciprocal translocation of a 1.4-Mb duplicated region of chromosome 2 to the end of chromosome 4. Another strain, containing a triplicated region of chromosome 2, was generated by translocating a 1.4-Mb region of chromosome 2 onto the ends of chromosomes 4 and 7. Phenotypic analyses of the strains containing segmental duplication or triplication of chromosome 2 showed remarkable increases in protease and amylase activities in solid-state cultures. Protease activity was further increased in strains containing the duplication and triplication after overexpression of the transcriptional activator of proteases *prtT*. This indicates that the gene-dosage effect and resulting phenotypes of the duplicated chromosomal region were enhanced by multiple duplications, and by the combination of the structural gene and its regulatory genes. Gene expression analysis, conducted using oligonucleotide microarrays, showed increased transcription of a large population of genes located in duplicated or triplicated chromosomal regions.

Conclusion: In this study, we performed translocated chromosomal duplications and triplications of a 1.4-Mb targeted region of chromosome 2. Strains containing a duplication of chromosome 2 showed significant increases in protease and amylase activities; these enzymatic activities were further increased in the strain containing a triplication of chromosome 2. This indicates that segmental duplications of chromosomes enhance gene-dosage effects, and that the resulting phenotypes play important phenotypic roles in *A. oryzae*.

Keywords: Targeted chromosomal duplication, Translocated duplication, Break induced replication, *Aspergillus oryzae*, Phenotypic change, *prtT*, Screening, Molecular breeding, Biotechnology, Genome engineering, Gross genome editing, Protease
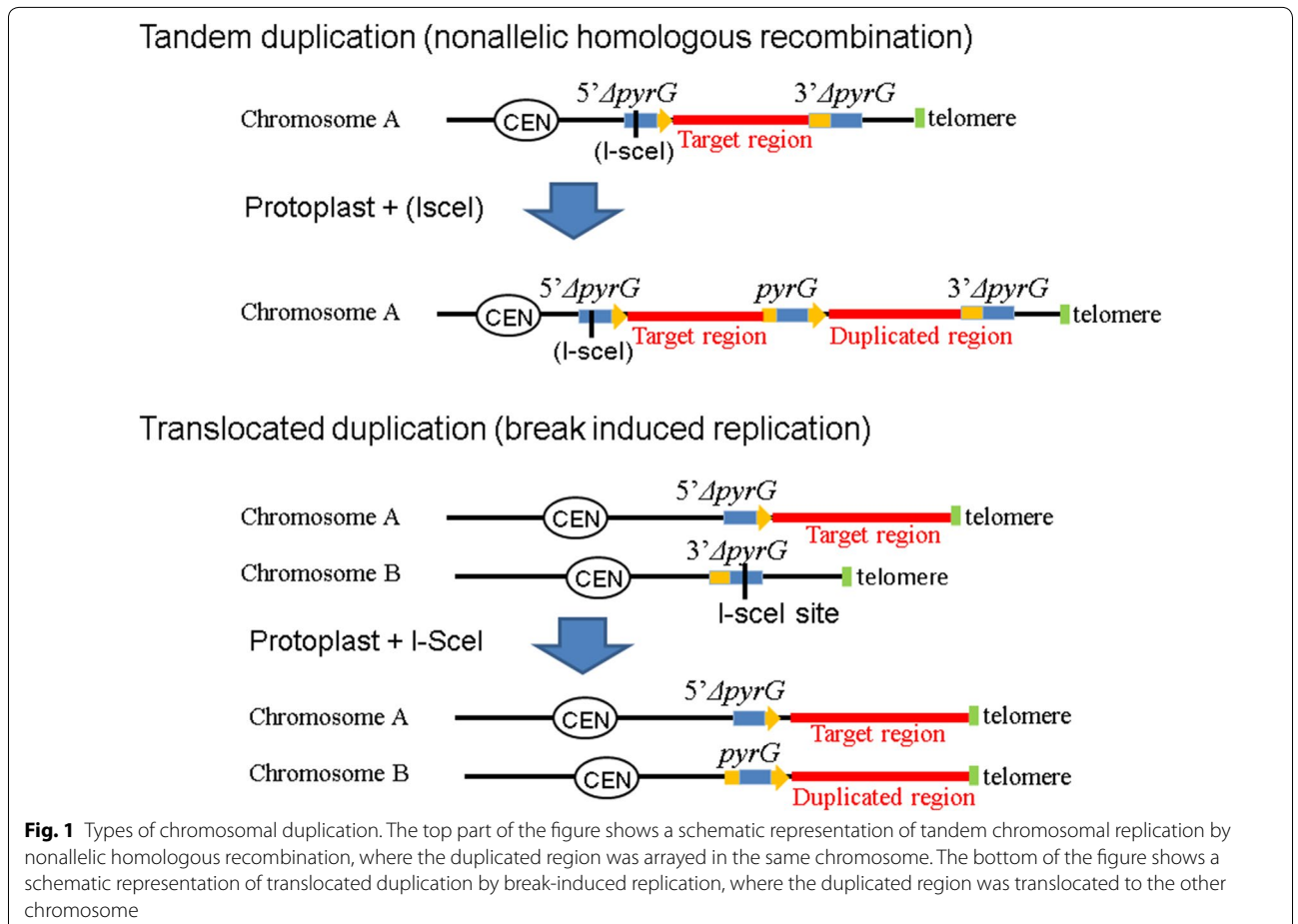
*Correspondence: ttakahashi@mail.kikkoman.co.jp
Noda Institute for Scientific Research, 399 Noda, Noda City, Chiba Pref
278-0037, Japan

Background

In eukaryotic organisms, duplications, translocations, reversions, and deletions of chromosomal regions occur spontaneously. For example, in *Saccharomyces cerevisiae*, duplication and translocation of chromosomes frequently occurs via Ty transposable element sequences, which are repeat sequences widely distributed in yeast chromosomes [1]. Similarly, previous studies of mammalian cells indicated that duplications, translocations, and deletions of chromosomes occur via repeat sequences such as Alu and L1 [2]. Changes in the chromosomal state often induce significant alterations in gene expression and result in phenotypic changes. In humans and other mammals, chromosome duplication or translocation events are reported to be among the causes of genetic diseases and cancers [3, 4].

The koji molds *Aspergillus oryzae* and *Aspergillus sojae* are filamentous fungi that are used in the fermentation of soy sauce and sake, and in production of industrial enzymes. The complete genome sequences of *A. oryzae* and *A. sojae* have been determined [5, 6]. Previously, genetic tools such as gene targeting [7, 8] and large chromosomal deletions [9, 10], have been developed to analyze the nature of *A. oryzae* and *A. sojae.* However, in koji molds, the functions of many genes remain unknown. Because of the presence of endogenous multiple orthologous genes, disrupting individual genes in strains of koji molds does not always generate clear phenotypic changes [5]. Overexpression of individual genes via a strong promoter can also cause reduced growth phenotypes [11]. However, strains with duplicated chromosomes can show phenotypic changes via increased dosage of unidentified genes [12]. By examining the genome of targeted chromosomal segments and phenotypic changes in the strains, we can determine the role of unknown genes corresponding to these phenotypic changes. In this study, we generated targeted segmental chromosomal duplications in *A. oryzae* and examined the phenotypic effects of these duplications. There are two common types of segmental chromosomal duplications: tandem duplication and translocated duplication (Fig. 1). In a tandem duplication, duplicated regions are arranged next to each other on the same chromosome by nonallelic homologous recombination (Fig. 1, top). In translocated duplications,

Fig. 1 Types of chromosomal duplication. The top part of the figure shows a schematic representation of tandem chromosomal replication by nonallelic homologous recombination, where the duplicated region was arrayed in the same chromosome. The bottom of the figure shows a schematic representation of translocated duplication by break-induced replication, where the duplicated region was translocated to the other chromosome

chromosomal duplications are translocated onto another chromosome by break-induced replication (Fig. 1, bottom). We previously generated *A. oryzae* strains containing targeted tandem chromosomal duplications using protoplasts of the parental strain, in which artificially generated consensus sequences (partly deleted selection marker) were introduced into both ends of the duplication target region [12]. If the copy number of the duplication target region is increased, such as in triplication, the phenotypic change in the strain may be further enhanced. However, to maintain stability of a tandem chromosomal duplication, a selection marker must be located at the junction between the two duplicated segments. In *A. oryzae*, this prevents removal of the duplicated region by recombination events occurring between the duplicated homologous sequences [12]. Therefore, it is technically difficult to increase the copy numbers of genes located in duplicated regions in tandem chromosomal duplications. Moreover, it is difficult to duplicate chromosomal ends in tandem duplications because a selection marker positioned at the chromosomal end is unstable and lost during regeneration (Takahashi et al. unpublished data). In contrast, generating multiple translocated chromosomal duplications is theoretically possible because the potential for recombination between the two duplicated regions is low in translocated chromosomal duplications, and selective pressures are not required to maintain the translocation. Furthermore, duplication of chromosomal ends is straightforward in translocated duplication because a selection marker is not required at chromosomal ends. However, no previous studies have generated translocated duplications of targeted chromosomal regions in filamentous fungi such as *A. oryzae*. Therefore, we generated a translocated duplication of a targeted chromosomal segment in *A. oryzae*.

The break-induced replication (BIR) mechanism causes translocated chromosomal duplications in yeast [13–15]. However, no previous studies have shown that the same mechanism can be used in filamentous fungi such as *A. oryzae*. To cause BIR in *A. oryzae*, it is necessary to artificially introduce double-strand breaks (DSBs) into the chromosomes (Fig. 1, bottom). In most studies on yeast, DSBs are introduced after the expression of genes encoding endogenous yeast homothallic switching endonuclease (*HO*) or I-SceI endonuclease (*SCEI*). In contrast, chromosome modifications have been performed in *A. oryzae* using polyethylene glycol (PEG)-mediated introduction of enzymes into protoplast cells [12, 16–18]. Therefore, we generated chromosomal duplications by directly treating protoplast cells with I-SceI meganuclease and monitoring the resulting chromosomal DSBs. This approach eliminated the need for I-SceI expression and produced a translocated chromosomal duplication

strain containing 1.4-Mb segment of the targeted chromosomal region. To the best of our knowledge, this is the first study to generate a targeted translocated chromosomal duplication using BIR and analyze the effect of chromosomal duplication on gene expression in filamentous fungi. Phenotypes of the strain resulting from this chromosomal duplication show increased activities of protease and amylase, indicating that this method can be used in functional analysis and molecular breeding of *Aspergillus* strains.

Results

Translocated duplication of targeted segment of chromosome 2 onto chromosome 4 and a strain bearing translocated triplication

We previously constructed strains containing targeted tandem chromosomal duplications [12]. Chromosome 2 of *A. oryzae* includes genes encoding alkaline protease and alpha-amylase, and their respective regulatory genes *prtT* and *amyR*, which are important for fermentation. Accordingly, strains containing a 700-kb tandem chromosomal duplication in chromosome 2 showed increased protease and amylase activities under solid-state culture conditions [12]. In the present study, we duplicated a 1.4-Mb region of chromosome 2 and translocated this region onto the end of chromosome 4 (Fig. 2). Successfully translocated duplications can increase the activities of protease and amylase in solid-state cultures, as shown previously by tandem duplication of chromosome 2 [12]. In subsequent analyses using BLASTN, no I-SceI recognition sequence was found in the chromosome of *A. oryzae* strain RIB40 [17]. To translocate the 1.4-Mb region of chromosome 2 onto the end of chromosome 4, we placed the *5′ΔpyrG* marker (targeted integration of the basic unit of 5′ΔpyrG and subsequent 5-fluoroorotic acid [5-FOA] selection) adjacent to the target donor chromosomal region. The *3′ΔpyrG* marker (targeted integration of the basic unit of 3′ΔpyrG and subsequent 5-FOA selection), which included the I-SceI recognition sequence, was placed adjacent to the acceptor chromosomal region (Fig. 2, top). The resulting parental strain was then used to generate translocated duplication of the targeted region of chromosome 2. For these procedures, protoplast cells of the parental strain were prepared and gently treated with I-SceI meganuclease and PEG. The treated protoplasts were then incubated at 30 °C on Czapek-Dox minimum medium (CZ) plates containing 1.2 M sorbitol. After more than 2 weeks of incubation, we observed a colony that was regenerated from treated protoplasts. The frequency of the number of regenerated colonies was approximately 10^{-8} per cell. To confirm that translocated duplication was achieved in the regenerated strain, genomic DNA of the regenerated strain was extracted

Fig. 2 A schematic showing translocated duplication of chromosome 2 in a strain of *A. oryzae*. A 1.4-Mb target donor region of chromosome 2 was duplicated and translocated to the acceptor region of chromosome 4 using break-induced replication
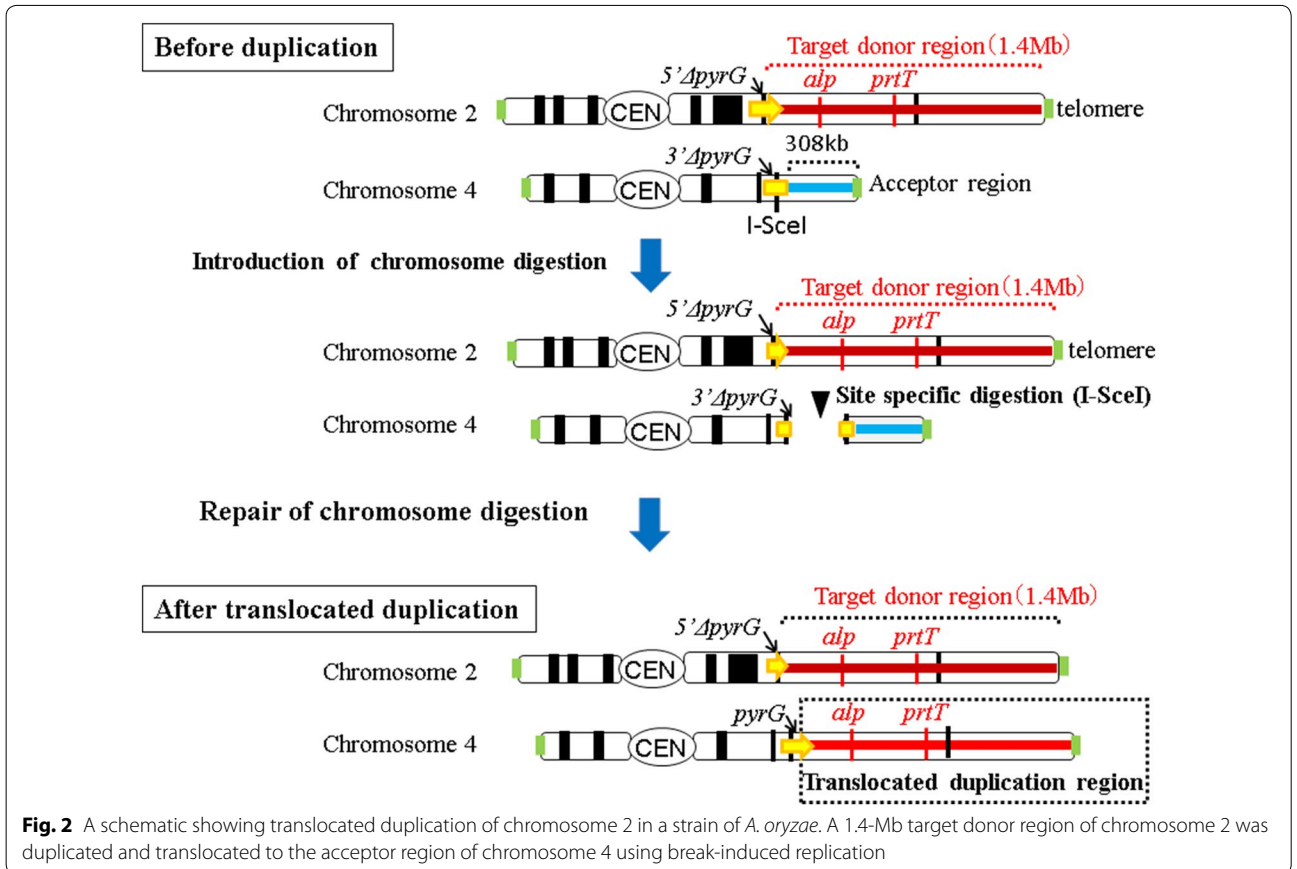
and analyzed by PCR using primers targeting the border of the translocated region (Fig. 3a).

We detected amplification of the 5-kb DNA fragment using combinations of primers targeting the translocated position (primers ct166-U and B-L in lane 4 of Fig. 3a) and original position (primers B-U and B-L in lane 1 of Fig. 3a) in the regenerated strain. In contrast, amplification of the 3-kb fragment was detected only by using the primer combination corresponding to the original position (primes B-U and B-L in lane 5 of Fig. 3a, and primers ct166-U and ct166-L in lane 6 of Fig. 3a) in the control strain (RIB40). These analyses confirmed that translocated duplication of the targeted chromosomal region

occurred in the regenerated strain; this strain was thereafter referred to as J4 and used in further analyses.

As described above, generating multiple translocated chromosomal duplications is theoretically possible. Therefore, we constructed a strain in which the targeted region of chromosome 2 was triplicated, and evaluated the effects of triplication on the phenotype of this strain. Initially, we removed the *pyrG* marker at the border of the translocated duplication region of chromosome 4 by homologous recombination between the chromosome and fragment used to remove *pyrG* (Additional file 1: Figure S3, top). Transformants were then selected on 1.2-M sorbitol-CZ plates containing 5-FOA, and the

(See figure on next page.)

Fig. 3 **a** Confirmation of translocated duplication by PCR. Targeted translocated duplication of a 1.4-Mb region of chromosome 2 to chromosome 4 was confirmed by PCR. *Lane 1*. B–U, B–L; *Lane 2*. ct166-U, ct166-L; *Lane 3*. B–U, ct166-L; *Lane 4*. ct166-U, B–L (*Lanes 1–4*. translocated duplication strain J4); *Lane 5*. B–U, B–L; *Lane 6*. ct166-U, ct166-L; *Lane 7*. B–U, ct166-L; *Lane 8*. ct166-U, B–L (*Lanes 5–8*. Wild-type strain). **b** Confirmation of translocated duplication by comparative genome hybridization (CGH) arrays. Vertical bars show ratios of signal intensities for probes of the strain bearing translocated duplication (J4) relative to those of the control strain (RIB40). Ratios of hybridization signals in the 1.4-Mb region of chromosome 2 were twice those of the other region, indicating that the 1380-kb region of chromosome 2, including AO090002001035 to AO090002001558, was duplicated. The ratio of hybridization signals in the 308-kb region of chromosome 4 was nearly zero, indicating that the 308-kb region of chromosome 4, including AO090166000010 to AO090166000123, was deleted
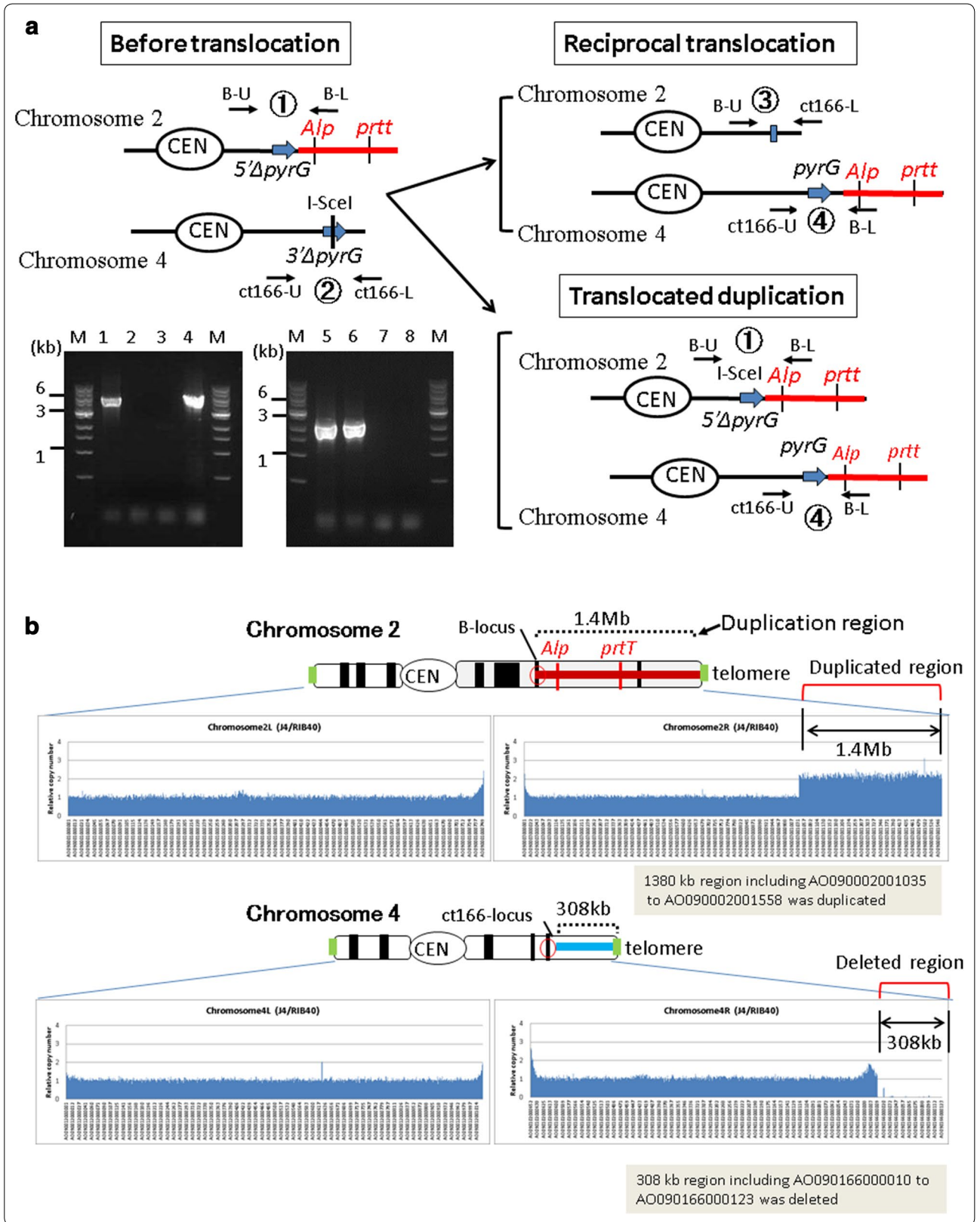
vector for introducing 3′Δ*pyrG* with the I-SceI recognition sequence was integrated near the end of chromosome 7 (Additional file 1: Figure S3, acceptor region). This was followed by selection using 5-FOA. The resulting triplication-containing parental strain (K1-IF5-2-5FOA) was cultured, and protoplasts were prepared for phenotypic analyses. Protoplasts were treated with I-SceI and PEG as described above. After incubation for 3 weeks at 30 ℃, regenerated colonies were analyzed by PCR. After single-colony isolation (details are described in Additional file 1: Fig. S4), the strain containing homokaryotic nuclei with the translocated chromosomes was thereafter referred to as I-8 strain.

Confirmation of translocated chromosomal duplications using array comparative genome hybridization (array CGH)

CGH was used to confirm the presence of translocated chromosomal duplication in the targeted chromosomal region. Briefly, genomic DNA isolated from RIB40 (control) and J4 strains (Δ*ku70* strain with translocated chromosomal duplication) was labeled with Cy dyes, hybridized on microarray slides at 65 ℃ for 24 h, and scanned with a laser-scanner. Comparison of the two strains revealed that the signal ratio of the probes in the 1.4-Mb region of chromosome 2 (AO090003001035–AO090003001558) was approximately two, and that of the probes located in the 308-kb region near the terminal end of chromosome 4 (AO090166000010–AO090166000123) was nearly zero (Fig. 3b). These observations confirmed the presence of translocated duplication of the targeted chromosomal region in the J4 strain (Fig. 3b and Additional file 1: Figure S1). The BN1-1 strain (*ku70*+strain with translocated chromosomal duplication) was also analyzed using CGH. Similar results were obtained (Additional file 1: Figure S2), indicating that translocated duplication of the targeted chromosomal region occurred successfully in the J4 and BN1-1 strains.

Next, genomic DNA isolated from RIB40 (control) and I-8 strains (Δ*ku70* strain with translocated chromosomal triplication) was labeled with Cy dyes, hybridized on microarray slides, and scanned using a laser-scanner. As shown in Fig. 4, comparisons of triplicate (I-8) and control strains revealed a signal ratio of approximately 3 for the probes located in the 1.4-Mb region of chromosome 2 (AO090003001035–AO090003001558). Moreover, the signal ratios of the probes located in the 308-kb near-telomeric region of chromosome 4 (AO090166000010–AO090166000123) and in the 13-kb near-telomeric region of chromosome 7 (AO090011000001–AO090011000008) were nearly zero, indicating successful translocation of chromosome 2 to the end of chromosomes 4 and 7 in the I-8 strain (Fig. 4).

These results indicate that multiple translocated duplications (such as translocated triplication) of the targeted chromosomal region were achieved in the I-8 strain.

Gene expression analysis of translocated chromosomal duplication and triplication strains in solid-state culture

Segmental duplication of large chromosomal regions may upregulate numerous genes in the duplicated chromosomal region and cause phenotypic changes via gene dosage effect. To demonstrate the effects of chromosomal duplication on gene expression, we examined gene expression levels of the strains under conditions of solid-state fermentation. The J4 strain with a translocated duplication of a 1.4-Mb region of chromosome 2, I-8 strain with a translocated triplication of the 1.4-Mb region of chromosome 2, and RIB40 strain (control) were cultivated for 65 h at 30 ℃ in wheat bran medium; then, RNA was extracted and analyzed by gene expression microarrays. The ratios of upregulated genes in duplicated chromosomal regions were remarkably higher in strains with duplicated and triplicated chromosomes. As shown in Tables 1 and 2, gene expression was increased by 11% in the whole chromosomal region (1293 vs. 12010). Gene expression was increased by 37 and 64% in the strains with duplicated (166 vs. 446) and triplicated regions (284 vs. 446), respectively. These data indicate that segmental duplication and triplication of target chromosomal regions effectively increased transcription of resident genes. These data were summarized according to Clusters of Orthologous Groups (COGs) classification [19]. Per COGs classifications, the influence gene dosage effect on signal transduction mechanisms (T) was low with 25% increases in the J4 (2 vs. 8) and 38% increases in the I-8 (3 vs. 8) strains. The expression of genes categorized into the nucleotide transport and metabolism (F) group, per COGs classification, was significantly increased (5 vs. 5) in the duplicated region of the strains (Tables 1 and 2). In contrast, gene expression was decreased by 4 and 2.5% in the duplicated (17 vs. 446) and triplicated chromosomal regions, respectively. For the chromosomes in the J4 (duplicated) strain, sites of upregulated genes were identified using probes (Additional file 1: Figure S5), which clearly showed increased gene transcription in the duplicated chromosomal region. Approximately 10% of the genes located outside of the duplicated region (1127 vs. 11,564) were upregulated; several of these genes were markedly upregulated in the duplicated strain (Additional file 1: Figure S5), indicating the presence of regulatory factors in the duplicated region. The genes upregulated in the J4 strain are summarized in Tables 3, 4, 5 and 6. Twenty-three proteolytic genes, 8 amylolytic genes, and 23 xylanolytic genes, which were located outside of the duplicated region, were

Fig. 4 Confirmation of translocated duplication by array CGH. Results of array CGH indicate that the 1.4-Mb region of chromosome 2 was triplicated and translocated to the ends of chromosomes 4 and 7
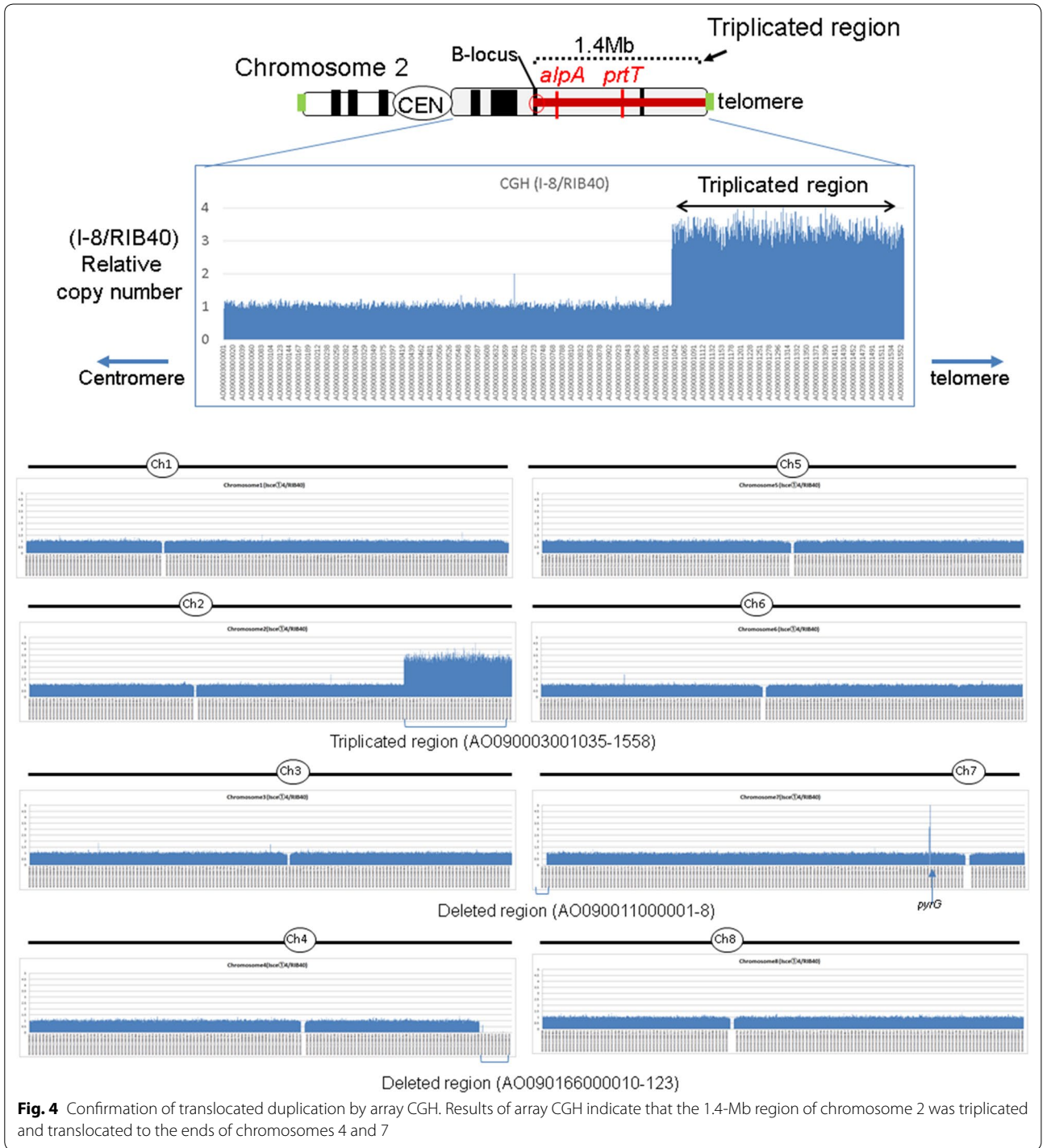
upregulated. Increased expression of proteolytic genes, located outside of the duplicated region, was observed in the J4 strain, suggesting that gene dosage effect exerted by *prtT* occurred in the duplicated region. Similarly, the upregulated expression of amylolytic genes indicated that gene dosage effect, exerted by *amyR*, occurred in the duplicated region. The expression of *xlnR* was not increased. However, we observed upregulated expression of genes related to xylanase, most of which were positively regulated by *xlnR* [20]. This indicate that *xlnR2* (AO090003001292) affected the expression of these genes. The expression of *prtT*, *amyR*, and *xlnR2* was

Table 1 Percent increase in the expression of upregulated genes in the duplicated chromosomal region of the J4 strain

COG description	Total	Up-regulated genes	%	Duplicated region	Up-regulated genes	%
[A] RNA processing and modification	176	3	2	8	1	13
[B] Chromatin structure and dynamics	38	2	5	4	2	50
[C] Energy production and conversion	371	56	15	9	3	33
[D] Cell cycle control, cell division, chromosome partitioning	101	10	10	5	2	40
[E] Amino acid transport and metabolism	453	65	14	18	9	50
[F] Nucleotide transport and metabolism	97	15	15	5	5	100
[G] Carbohydrate transport and metabolism	500	103	21	27	15	56
[H] Coenzyme transport and metabolism	133	13	10	4	3	75
[I] Lipid transport and metabolism	314	37	12	12	7	58
[J] Translation, ribosomal structure and biogenesis	277	5	2	6	3	50
[K] Transcription	180	4	2	9	3	33
[L] Replication, recombination and repair	172	11	6	7	2	29
[M] Cell wall/membrane/envelope biogenesis	98	11	11	5	1	20
[N] Cell motility	2	0	0	0	0	0
[O] Posttranslational modification, protein turnover, chaperones	390	19	5	12	6	50
[P] Inorganic ion transport and metabolism	195	21	11	9	2	22
[Q] Secondary metabolites biosynthesis, transport and catabolism	447	68	15	9	4	44
[R] General function prediction only	1370	186	14	51	18	35
[S] Function unknown	333	27	8	16	7	44
[T] Signal transduction mechanisms	310	26	8	8	2	25
[U] Intracellular trafficking, secretion, and vesicular transport	241	5	2	9	1	11
[V] Defense mechanisms	50	11	22	0	0	0
[W] Extracellular structures	3	0	0	0	0	0
[Y] Nuclear structure	6	0	0	0	0	0
[Z] Cytoskeleton	94	10	11	2	0	0
Unannotated	5659	585	10	211	70	33
Total	12,010	1293	11	446	166	37

*$p < 0.02$ and fold change > 1 indicate upregulated genes

% increase shows the percent ratio of upregulated genes to total genes in the whole genome or that of the genes in the duplicated region

upregulated in the duplicated region (Table 6). In contrast, 1252 genes located outside of the duplicated region were downregulated in the J4 strain, indicating that genes involved in negative regulation were also present in the duplicated region.

Enzymatic activity in translocated chromosomal duplication and triplication strains in solid-state culture

To examine how chromosomal duplication and triplication affect phenotypes, we assessed enzymatic activity in translocated chromosomal duplication strains under solid-state conditions. We have previously shown increased activity of protease, amylase, and acid carboxypeptidase in strains with the 700-kb region (AO090003001003–AO090003001258) of the tandem duplication of chromosome 2 under conditions of solid-state cultivation on wheat bran media [12]. Therefore, we measured the activities of these enzymes in translocated

duplication strains under solid-state culture conditions (Fig. 5a). Both translocated (J4) and tandem (700k-dup) chromosomal duplications led to increased activity of protease and amylase. Moreover, the activity of acid carboxypeptidase in the translocated duplication strain showed slightly higher levels than that in the wild-type strain (RIB40 control), but lower levels than that in the tandem duplication strain. A list of genes located in the overlapped region between J4 and 700k-dup is provided in Additional file 1: Table S2. We then examined enzymatic activity in the triplication strain I-8 under solid-state culture conditions. Protease activity was increased by more than fourfold in the triplication strain (I-8) compared to that in the control strain (RIB40), and was higher than those in the translocated and tandem duplication strains (J4 and 700k-dup) (Fig. 5a). The activity of amylase was also higher in the triplication strain (I-8) than in the duplication strains (J4 and 700k-dup), indicating

Table 2 Percent increase in upregulated genes in the triplicated chromosomal region of the I-8 strain

COG description	Total	Up-regulated genes	%	Duplicated region	Up-regulated genes	%
[A] RNA processing and modification	176	9	5	8	7	88
[B] Chromatin structure and dynamics	38	3	8	4	3	75
[C] Energy production and conversion	371	42	11	9	5	56
[D] Cell cycle control, cell division, chromosome partitioning	101	13	13	5	5	100
[E] Amino acid transport and metabolism	453	60	13	18	14	78
[F] Nucleotide transport and metabolism	97	11	11	5	5	100
[G] Carbohydrate transport and metabolism	500	70	14	27	24	89
[H] Coenzyme transport and metabolism	133	13	10	4	3	75
[I] Lipid transport and metabolism	314	39	12	12	9	75
[J] Translation, ribosomal structure and biogenesis	277	8	3	6	6	100
[K] Transcription	180	10	6	9	8	89
[L] Replication, recombination and repair	172	14	8	7	5	71
[M] Cell wall/membrane/envelope biogenesis	98	10	10	5	3	60
[N] Cell motility	2	0	0	0	0	0
[O] Posttranslational modification, protein turnover, chaperones	390	21	5	12	10	83
[P] Inorganic ion transport and metabolism	195	18	9	9	4	44
[Q] Secondary metabolites biosynthesis, transport and catabolism	447	54	12	9	5	56
[R] General function prediction only	1370	138	10	51	24	47
[S] Function unknown	333	29	9	16	14	88
[T] Signal transduction mechanisms	310	21	7	8	3	38
[U] Intracellular trafficking, secretion, and vesicular transport	241	10	4	9	7	78
[V] Defense mechanisms	50	11	22	0	0	0
[W] Extracellular structures	3	0	0	0	0	0
[Y] Nuclear structure	6	0	0	0	0	0
[Z] Cytoskeleton	94	12	13	2	2	100
Unannotated	5659	537	9	211	118	56
Total	12,010	1153	10	446	284	64

*$p < 0.02$ and fold change > 1 indicate upregulated genes

% increase shows the percent ratio of upregulated genes to total genes in the whole genome or that of the genes in the duplicated region

that gene dosage showed increased phenotypic effects after multiple duplication of the chromosomal region. The BP-B3 strain, a tandem duplication strain bearing a 9-kb region (AO090003001033-AO090003001036) that included *alp* (gene encoding alkaline protease, AO090003001036) [21], showed 1.7-fold increase in protease activity compared to that of the control strain (Fig. 5a). The expression of *alp* and *prtT* in the duplication strains were examined by real-time PCR (Table 7).

We then overexpressed the transcription factor *prtT* and examined its effects in duplicated and triplicated chromosomal translocation strains. PrtT regulates proteolytic enzymes in *A. oryzae* [22]. Our gene expression microarrays revealed moderate fivefold increases in *prtT* expression in duplication strains. These data suggest that *prtT* is rate-limiting, indicating increased protease activity in *prtT* over-expressing strains. The *prtT* over-expression vector pAPRT contains an *amyB* promoter and terminator, connected to the open reading frame

of the *prtT* gene, and a *pyrG* marker. pAPRT in circular state was used to transform the RP1 strain (*pyrG* deletion strain), del 1546K4 strain (*pyrG* deletion strain derived from the J4 strain), and TLTAs11B strain (*pyrG* deletion strain derived from the I-8 strain). Single transformants showing clear large halos on casein plates, which indicates high protease activity [12], were selected for each *prtT* overexpression strain; overexpression of *prtT* was confirmed by real-time PCR (Table 8).

The expression of *prtT* was increased more than 10-fold in the APRT, 1546 K-APRT, and TLTA-APRT strains compared to that in the control strain (Table 8). Accordingly, the activity of protease in solid-state cultures (Fig. 5a) was approximately threefold higher in the APRT-transformed strain than in the RIB40 strain (wild type). This indicates that a single copy of *alp* was rate-limiting for protease activity under condition of *prtT* overexpression. Moreover, protease activity was similar in the 1546 K-APRT and TLTA-APRT strains and nearly

Table 3 Upregulation of proteolytic genes located outside of the duplicated chromosomal region in the J4 strain

Systematic name	Relative gene[a] expression	Description
AO090001000135	2.7	Mep20-metalloproteinase
AO090003000354	3.2	LapII, transferrin receptor and related proteins
AO090009000148	2.1	OpsA;aspartyl protease
AO090009000171	3.5	SPRT-like metalloprotease
AO090009000593	4.4	Metal-dependent amidase/aminoacylase/carboxypeptidase
AO090010000493	4.5	NpII-neutral protease II
AO090010000534	4.4	Serine carboxypeptidases (lysosomal cathepsin A)
AO090010000540	3.9	Aminoacylase ACY1 and related metalloexopeptidases
AO090011000036	9.7	Np I-neutral protease I
AO090011000052	4.7	LapI-leucine aminopeptidase
AO090011000235	3.0	TppA-tripeptidyl peptidase A
AO090012000022	2.9	Metal-dependent amidase/aminoacylase/carboxypeptidase
AO090012000080	2.4	Metal-dependent amidase/aminoacylase/carboxypeptidase
AO090012000706	4.0	Carboxypeptidase C (cathepsin A)
AO090020000288	3.1	Aminoacylase ACY1 and related metalloexopeptidases
AO090020000351	4.0	Serine carboxypeptidases (lysosomal cathepsin A)
AO090023000382	2.6	Carboxypeptidase C (cathepsin A)
AO090026000083	4.5	AorO-aorsin
AO090026000680	3.8	Serine carboxypeptidases (lysosomal cathepsin A)
AO090103000264	3.9	Predicted molecular chaperone distantly related to HSP70-fold metalloproteases
AO090138000101	4.0	Putative intracellular protease/amidase
AO090138000114	7.1	Meltrins, fertilins and related Zn-dependent metalloproteinases of the ADAMs family
AO090701000220	2.3	Carboxypeptidase C (cathepsin A)

[a] Relative gene expression in the J4 strain compared to that in the RIB40 strain was measured by gene expression array. $p < 0.02$

Table 4 Upregulation of amylolytic genes located outside of the duplicated region in the J4 strain

Systematic name	Relative gene[a] expression	Description
AO090001000259	2.8	Beta-galactosidase
AO090001000492	2.6	Arabinogalactan endo-1,4-beta-galactosidase
AO090012000389	8.3	Beta-galactosidase/beta-glucuronidase
AO090023000944	2.5	Alpha-amylase
AO090038000471	2.5	Maltase glucoamylase and related hydrolases, glycosyl hydrolase family 31
AO090103000378	2.9	Alpha-amylase
AO090120000158	6.4	Beta-galactosidase
AO090701000558	2.2	Alpha-glucosidases, family 31 of glycosyl hydrolases

[a] Relative gene expression in the J4 strain compared to that in the RIB40 strain was measured by gene expression array. $p < 0.02$

sixfold higher than that in the RIB40 strain; this indicates that two copies of *alp* were sufficient for protease activity when *prtT* was overexpressed.

The APRT, 1546 K-APRT, and TLTA-APRT strains showed similarly increases in the activity of acid carboxypeptidase compared to that of the RIB40 strain. This suggests that acid carboxypeptidase genes are transactivated by *prtT*, but are not located in the duplicated chromosomal region. Single-copy integration of pAPRT in the APRT and TLTA-APRT strains, and two-copy integration of pAPRT in the 1546 K-APRT strain, were confirmed by quantitative PCR (Additional file 1: Figure S7).

The growth phenotypes of translocated duplication strains on CZ plates are presented in Fig. 5b. The strains were cultured on 1.2 M sorbitol-CZ plates at 30 °C for 7 days. The J4 strain showed a slight delay in growth compared to the growth of the RIB40 strain. The I-8 strain showed a slight delay in growth compared to that of the J4 strain. The growth rate of the strains gradually decreased as the copy number of the duplicated region from chromosome 2 increased. The deletion of a 308-kb region from chromosome 4, resulting from translocation in the J4 and I-8 strains, may have caused the delay in growth. The growth rates of the APRT and 1546 K-APRT strains were similar to those of the RIB40 and J4 strains, respectively. However, the TLTA-APRT strain showed a

Table 5 Upregulation of xylanolytic genes located outside of the duplicated region in the J4 strain

Systematic name	Relative gene[a] expression	Description
AO090001000164	2.3	Protein with predicted nucleotide binding activity*
AO090001000383	6.9	Extracellular catechol oxidase*
AO090005000476	2.8	Beta-xylosidase
AO090005000531	3.6	Endoglucanase B; predicted secretion signal peptide*
AO090005000698	3.7	Beta-xylosidase*
AO090010000063	5.5	Sugar transporter (MFS family); transcriptionally induced by growth on xylose*
AO090011000141	3.6	Putative exoarabinase*
AO090023000001	6.7	AbfB-alpha-L-arabinofuranosidase B
AO090023000401	3.3	PglB-polygalacturonase B*
AO090038000426	4.1	Dehydrogenases related to short-chain alcohol dehydrogenases*
AO090038000631	3.5	xdhA1-Xylitol dehydrogenase*
AO090103000087	3.0	Putative endoglucanase precursor*
AO090103000268	6.2	Beta-xylosidase*
AO090103000326	2.5	Beta-1,4-xylanase
AO090103000423	2.8	XynF1, beta-1,4-xylanase*
AO090103000426	4.0	Shikimate 5-dehydrogenase*
AO090124000023	3.1	AbfA,;alpha-L-arabinofuranosidase
AO090701000885	3.1	As Abf-alpha-L-arabinofuranosidase*
AO090001000649	3.4	Alpha-glucosidases, family 31 of glycosyl hydrolases*
AO090011000140	3.6	Beta-glucosidase-related glycosidases*
AO090011000715	2.7	Ak eglA-Endoglucanase
AO090023000056	2.1	Endoglucanase B
AO090038000175	2.5	Endoglucanase

*Positively regulated by XlnR [20]

[a] Relative gene expression in the J4 strain compared to that in the RIB40 strain was measured by gene expression array. $p < 0.02$

Table 6 Upregulation of proteolytic, amylolytic, and xylanolytic genes, and expression of their respective regulatory genes, located in the duplicated region of the J4 strain

Systematic name	Relative gene[a] expression	Description
AO090003001036	3.2	Alp, subtilisin-related protease
AO090003001208	2.0	AmyR
AO090003001209	2.5	AgdA, maltase glucoamylase and related hydrolases, glycosyl hydrolase family 31
AO090003001211	5.1	prtT
AO090003001258	2.4	Peptidase family M48
AO090003001292	3.0	XlnR2
AO090003001305	2.7	Alpha-D-galactosidase (melibiase)
AO090003001341	8.3	Endoglucanase
AO090003001497	2.1	Glycosidases
AO090003001511	2.3	Beta-glucosidase-related glycosidases

[a] Relative gene expression in the J4 strain compared to that in the RIB40 strain was measured by gene expression array. $p < 0.02$

Discussion

In this study, we generated strains of *A. oryzae* in which a targeted chromosomal region was translocated and duplicated. CGH and PCR were used to confirm that targeted translocation indeed occurred. Because CGH is commonly used to detect chromosomal rearrangement in yeast. It was confirmed that the original genome of *A. oryzae* did not contain an I-SceI recognition sequence [17], indicating a low possibility of unexpected chromosomal rearrangement.

The strain with the translocated duplication of the targeted chromosomal region was obtained by treating protoplast cells with the I-SceI enzyme. No strains containing translocated duplication were generated in the absence of I-SceI, indicating that translocated chromosomal duplication in *A. oryzae* requires artificially introduced DSBs and depends on the break-induced replication mechanism. We previously obtained tandem chromosomal duplications from protoplasted cells without I-SceI-derived chromosomal DSBs in *A. oryzae* [12] and hypothesized that the tandem chromosomal duplications were generated by a nonallelic homologous recombination mechanism instead of classical unequal sister chromatid exchange in diploid budding yeast [23]. In addition, the time required for regenerating cells and frequency of regenerated colonies differed between tandem and translocated chromosomal duplications in *A. oryzae*. Specifically, tandem chromosomal duplications were produced after incubating protoplast cells for 5–7 days on regeneration plates, and the frequency of regenerated
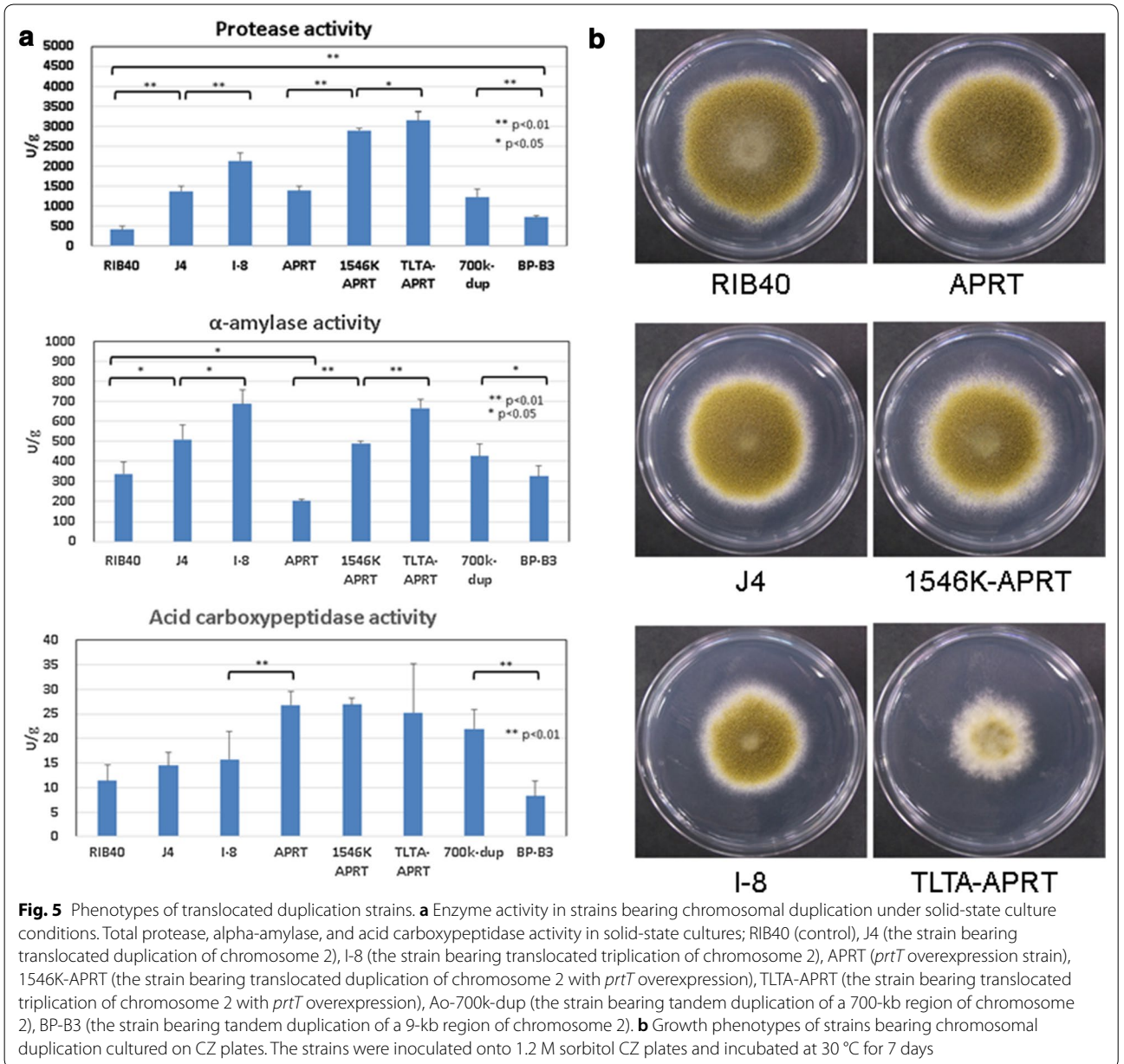
severe delay in growth compared to that of the I-8 strain. As shown in Table 8, the expression of *alp* was extremely high in the TLTA-APRT strain compared to that in the APRT or 1546 K-APRT strains. The expression of *alp* was originally high in the control strain (RIB40), suggesting that increased expression of *alp* caused growth delays in the TLTA-APRT strain.

Fig. 5 Phenotypes of translocated duplication strains. **a** Enzyme activity in strains bearing chromosomal duplication under solid-state culture conditions. Total protease, alpha-amylase, and acid carboxypeptidase activity in solid-state cultures; RIB40 (control), J4 (the strain bearing translocated duplication of chromosome 2), I-8 (the strain bearing translocated triplication of chromosome 2), APRT (*prtT* overexpression strain), 1546K-APRT (the strain bearing translocated duplication of chromosome 2 with *prtT* overexpression), TLTA-APRT (the strain bearing translocated triplication of chromosome 2 with *prtT* overexpression), Ao-700k-dup (the strain bearing tandem duplication of a 700-kb region of chromosome 2), BP-B3 (the strain bearing tandem duplication of a 9-kb region of chromosome 2). **b** Growth phenotypes of strains bearing chromosomal duplication cultured on CZ plates. The strains were inoculated onto 1.2 M sorbitol CZ plates and incubated at 30 ℃ for 7 days

Table 7 Quantitative expression of *alp* and *prtT* in strains bearing chromosomal duplication

Strain	*alp*			*prtT*		
	Relative[a] quantity (dR)	Upper[b] error bars	Lower[b] error bars	Relative[a] quantity (dR)	Upper[b] error bars	Lower[b] error bars
RIB40	1	0.58	0.37	1	0.26	0.21
J4	3.11	0.58	0.49	4.14	0.61	0.53
I-8	7.10	2.18	1.98	5.81	1.32	1.54
D2	2.84	1.16	0.71	4.28	1.75	1.08
BP-B3	1.84	1.47	0.70	2.19	1.67	0.92

[a] Relative quantity of gene expression was measured by real-time PCR

[b] Data are expressed as mean ± SD, and all experiments were conducted in triplicate

Table 8 Quantitative expression of *alp* and *prtT* in *prtT* overexpression strains

Strain	*alp*			*prtT*		
	Relative[a] quantity (dR)	Upper[b] error bars	Lower[b] error bars	Relative[a] quantity (dR)	Upper[b] error bars	Lower[b] error bars
RIB40	1	0.08	0.07	1	0.15	0.13
APRT	5.26	1.19	1.00	15.69	2.45	2.14
1546 K APRT	8.07	1.10	1.03	15.34	2.27	2.03
TLTA-APRT	17.31	1.52	1.35	18.33	2.34	2.05

[a] Relative quantity of gene expression was measured by real-time PCR

[b] Data are expressed as mean $\pm$ SD, and all experiments were conducted in triplicate

colonies from the protoplast cells was approximately 10^{-7} [12]. In contrast, strains with translocated chromosomal duplications were generated only after protoplast cell regeneration was conducted for more than 2 weeks, and the frequency of regenerated colonies from the protoplasted cell was approximately 10^{-8}. To obtain translocated chromosomal triplications in regenerated state, nearly 3 weeks were necessary to regenerate the colonies, and most regenerated cells were heterokaryons containing both the original chromosome and translocated chromosome (Additional file 1: Figure S4). Hence, in the present study, single-colony isolation was necessary to obtain a pure translocated chromosomal triplication strain, indicating that more time is required for recombination to form translocated chromosomal regions in *A. oryzae*. These results indicate that tandem duplication of a targeted chromosomal region achieved in our previous study [12], and translocated duplication of a targeted chromosomal region achieved in the present study, depend on completely different recombination mechanisms and techniques.

In the present study, we generated duplication and triplication strains of *A. oryzae* chromosome 2 including *alp* and *prtT*. The strains bearing duplicated and triplicated chromosome 2 showed significant increases in the activity of protease under solid-state culture conditions.

However, protease activity in the BP-B3 strain (generated using tandem duplication of the 9-kb region, which included *alp* but not *prtT*) was less than those of the J4 and 700k-dup strains (generated using duplication of a chromosomal region containing both *alp* and *prtT*) (Fig. 5a). This indicates that duplication of both *alp* and *prtT*, or a structural gene and its regulatory gene, enhances the activity of protease. The duplicated region of chromosome 2 included the alpha-amylase gene and *amyR*, suggesting that amylase activity in the strains was enhanced via similar mechanism.

Oligonucleotide microarrays revealed increased gene expression with increasing amounts of genetic material in the strains containing duplicated and triplicated chromosome regions as compared to that of the control strain (Tables 1 and 2). Thus, the transcription levels of genes in the duplicated chromosomal regions were predominantly affected, and the resulting phenotypes were enhanced as the copy number increased. This suggests that segmental chromosomal duplication can be used to identify unknown gene functions based on phenotypic observations.

Protease activity was used to confirm the duplication stability of chromosome 2. Genome stability of *A. oryzae* strains in which translocated duplication was generated by mutagenesis was described in a United States patent US8900647B2. The genomes of the strains bearing translocated duplication were stable after 10 generation of subculture. Additionally, in our previous report [12], we showed that targeted tandem duplication of chromosome 2 remained stable after 5 rounds of subculture. Translocated duplication is more stable than tandem duplication because the possibility of losing the duplicated region via recombination is low.

Translocated duplication strains were generated from both *ku70+* and *ku70-* strains; the BN1-1 strain was derived from *ku70+* and J4 strain was derived from *ku70−* (Table 9). This suggests that non-homologous end-joining pathways involving Ku70/80 heterodimers are not important in the mechanism of translocated duplication. As shown previously, undesirable deletions or illegitimate recombinations were not observed in the combination of the *Δku70* strain and 5FOA counter selection [17]. This indicates that the *Δku70* strain of *A. oryzae* is suitable for *pyrG*-mediated targeted chromosomal duplication used previously [12] and in our present study. Therefore, it is possible to identify interesting phenotypes by constructing a chromosome duplication strain library that covers the whole genome and to detect unknown regulatory genes by screening and analyzing the library in *A. oryzae*.

Table 9 Strains used in the present study

A. oryzae RIB40 strain	Genotype	Source or reference
RIB40	Wild type	
Ku70RC7-2	*Δku70, ΔpyrG*	[10]
RP-1	*ΔpyrG*	[7, 8]
Ao-700k-dup	Tandem duplication of 700kb region of chromosome 2	[12]
B1036p5'ΔCTp3'Δ	*5'ΔpyrG, 3'ΔpyrG* (parental strain of BN1-1)	Present study
ku70-B1036p5'ΔCTp3'Δ	*Δku70, 5'ΔpyrG, 3'ΔpyrG* (parental strain of J4)	Present study
BN1-1	Translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4	Present study
J4	*Δku70*, translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4	Present study
del1546K4	*Δku70*, translocated duplication of 1.4 Mb of region chromosome 2 to chromosome 4, *ΔpyrG*	Present study
K1-IF5-2-5FOA	*Δku70*, translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4, *5'ΔpyrG, 3'ΔpyrG* (parental strain of I-8)	Present study
I-8	*Δku70*, translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4 and chromosome 7	Present study
TLTAs11B	*Δku70*, translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4 and chromosome 7, *ΔpyrG*	Present study
APRT	pAPRT	Present study
1546 K-APRT	*Δku70*, translocated duplication of 1.4 Mb region of chromosome 2 to chromosome 4, pAPRT	Present study
TLTA-APRT	*Δku70*, translocated duplication of 1.4 Mb of chromosome 2 to chromosome 4 and chromosome 7, pAPRT	Present study
BP-B3	Tandem duplication of 9 kb region of chromosome 2	Present study

Conclusion

In this study, we achieved translocated chromosomal duplication and triplication of a 1.4-Mb targeted chromosomal region by directly introducing I-SceI meganuclease into *A. oryzae* protoplast cells. Strains with duplication and triplication of chromosome 2 showed substantial increases in the activity of protease and amylase. Gene dosage effects were enhanced by combining the structural gene and its regulatory gene, indicating that segmental duplications of chromosomes play important phenotypic roles in koji mold strains.

Methods

Strains, media, and transformation

Host strains included *A. oryzae* RIB40 (ATCC 42149), *A. oryzae* RP-1(*ΔpyrG*) [7, 8], and *A. oryzae* Ku70RC7-2 (*Δku70, ΔpyrG*) [10]. Positive selection of *pyrG*-deficient strains was performed on Czapek-Dox minimum medium (CZ) plates containing 20 mM uridine and 1.5 mg/mL 5-fluoroorotic acid (5-FOA). Solid-state cultivation was performed using medium containing 20% wheat bran and 80% water (wt/wt). Transformation of *Aspergillus* strains using protoplasts and PEG was conducted as previously described [17, 24]. Strains bearing translocated chromosomal duplication were obtained as follows. Approximately 1×10^8 protoplasts were prepared from the host strain and maintained on ice for 30 min in a solution containing 0.05 mL of 1.2 M sorbitol, 0.02 mL of PEG, and 50 U of I-SceI endonuclease (New England Biolabs, Ipswich, MA, USA). After the addition of 0.07 mL PEG, the cells were incubated at room temperature for 1 h. Then, the cells were spread on regeneration plates containing 1.2 M sorbitol-CZ and incubated at 30 °C for 2–3 weeks. Regenerated colonies on regeneration plates were transferred to CZ plates for further analysis. Information on the genome of *Aspergillus oryzae* was collected from the genome database at the National Institute of Technology and Evaluation (http://www.bio.nite.g.o.jp/dogan/Top) and AspGD (http://www.aspergillusgenome.org). All oligonucleotide primers used in this study are listed in Additional file 1: Table S1.

DNA techniques and expression analysis

Genomic DNA from *Aspergillus* strains was extracted as described previously [24]. RNA samples were prepared from mycelia that were inoculated on wheat bran medium and cultivated at 30 °C for 65 h. After cultivation, the mycelia were harvested and ground in liquid nitrogen using a mortar and pestle; RNA was extracted using Isogen reagent (Nippon Gene, Toyama, Japan). Further purification was performed using RNeasy Mini kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions. RNA was used in quantitative real-time PCR (RT-PCR) and gene expression microarrays.

PCR amplification was performed using TaKaRa Ex Taq DNA polymerase (TaKaRa, Shiga, Japan) in a T100 thermal cycler (Bio-Rad, Hercules, CA, USA). Quantitative RT-PCR was performed in a MxPro3000P (Agilent Technologies, Santa Clara, CA, USA) using PrimeScript™ RT reagent with gDNA Eraser (Perfect Real Time) and SYBR® *Premix Ex Taq*™ II (Tli RNase H Plus) (TaKaRa). The expression of *alp* and *prtT* were analyzed by comparative quantification using MxPro software version 4.10 (Agilent Technologies). AO090005000807 (ortholog of transcription factor TFIID) was used as normalizer, and primers are listed in Table S2. Oligonucleotide arrays were purchased from Agilent Technologies; the experimental protocol was detailed in a previous study [17, 25]. The CGH and gene expression array data obtained in the present study have been deposited into the NCBI Gene Expression Omnibus [26] and are accessible using GEO Series accession number GSE120604.

Construction of vector used for partial deletion of *pyrG*

Construction of vectors used to create the 5′*ΔpyrG* and 3′*ΔpyrG* constructs was carried out in a manner similar to a method described previously [12]. The schematic of vector construction is shown in Additional file 1: Figure S6. The 5′*ΔpyrG* construct was *pyrG*-truncated by removing 631 bp from the 5′ end of *pyrG*, which included the promoter region. The 3′*ΔpyrG* construct was *pyrG*-truncated by removing 1236 bp from the 3′ end of *pyrG*, which included the terminator region. A 462-bp consensus region was present between 5′*ΔpyrG* and 3′*ΔpyrG*, and 3′*ΔpyrG* included an I-SceI site in the consensus region. To construct the 5′*ΔpyrG* vector, a 5′*ΔpyrG* vector backbone was constructed using the primers P499L, P509U, P1130U-IF5, and P1592L-IF5. Then, 5.8- and 0.5-kb fragments were amplified via PCR from the pPB9 plasmid (a pUC-based plasmid bearing a 3.0-kb fragment containing *pyrG*) [12] using the primer pairs P499L/P509U and P1130U-IF5/P1592L-IF5. The fragments were then ligated using an In-Fusion cloning kit (TaKaRa), and the basic construct for 5′*ΔpyrG* BP1130 was generated. To construct a vector for integrating 5′*ΔpyrG* at the B1036 site of chromosome 2, a 3-kb fragment was PCR-amplified from the genome of the RIB40 strain using primers B1036-U and B1036-L. Cloning of this fragment using a TA cloning kit (TOYOBO, Osaka, Japan) produced pB1036T, and pB1036T and BP1130 were PCR-amplified using the primers B1036-IFU and B1036IFL, and P363U and P3297L, respectively. The resulting two fragments were treated with DpnI and ligated using In-Fusion Cloning Kits (TaKaRa) to generate the pB1036pyr5′d vector. To construct the 3′*ΔpyrG* vector including the I-SceI site, a 3′*ΔpyrG* vector backbone was constructed using the primers P2788L, P2828U, P1130U-IF3, and P1592L-IF3.

Then, a 5.8-kb fragment was amplified from the pPB9 plasmid [12] using the primer pair P2788L/P2828U, and a 0.5 kb-fragment was amplified from the pBP9-sceI (a pBP9-based vector bearing *pyrG*, including the I-SceI recognition site) [12] using the primer pair P1130U-IF3/P1592-IF3. The fragments were then ligated using an In-Fusion cloning kit (TaKaRa), and the basic construct for 3′*ΔpyrG* BP1130-I was generated. To construct the vector for integrating 3′*ΔpyrG* at the CT166 locus of chromosome 4, a 3-kb fragment of genomic DNA from the RIB40 strain was amplified using the primers ct166-U and ct166-L. This fragment was cloned using a TA cloning kit (TOYOBO) to prepare pCT166T. pCT166T was amplified via PCR using the primers ct166-IFU and ct166-IFL. The basic unit of 3′*ΔpyrG*, BP1130-I, was amplified using the primers P363U and P3297L. The resulting fragments were treated with DpnI and ligated into the vector using an In-Fusion Cloning Kits (TaKaRa) to obtain pCT166pyr3′d. The fragment for removing *pyrG* from the J4 strain was constructed as follows: a 2-kb amplification fragment and a 2.3-kb fragment were amplified from RIB40 genomic DNA using the primer pairs TL-B5626L and TL-B3616U-IF, and TL-ct3685L-IF and TL-ct1388U, respectively. Fragments were ligated using In-fusion cloning kits (TaKaRa) to obtain the vector for removing *pyrG* from the J4 strain. To construct the vector for integrating 3′*ΔpyrG* at the S11 locus of chromosome 7, a 1.3-kb fragment of genomic DNA from the RIB40 strain was amplified using the primers S11-U and S11-L. This fragment was cloned using a TA cloning kit (TOYOBO) to prepare pS11T. pS11T was amplified via PCR using the primers S11-IFU and S11-IFL. The basic unit of 3′*ΔpyrG*, BP1130-I, was amplified using the primers P363U and P3297L. The resulting fragments were treated with DpnI and ligated into the vector using In-Fusion Cloning Kits (TaKaRa) to obtain pS11pyr3′d. The fragment for removing *pyrG* from the I-8 strain was constructed as follows: a 2.8-kb fragment was amplified using the primers TL-B3616U-IF-S11 and B6217L, and a 2.7-kb fragment was amplified using the primers TL-S11-14389U-IF-B and S11-17060L. The two PCR products were then ligated using In-Fusion Cloning Kits (TaKaRa) to obtain the vector for removing *pyrG* from the I-8 strain. Details of *pyrG* removal using homologous fragments were described previously [10].

Construction of parental strain for strains bearing translocated chromosomal duplication

To construct parental strains for generating strains with translocated chromosomal duplication, RP-1 (*ΔpyrG*) and ku70RC7-2 (*Δku70*, *ΔpyrG*) were transformed with pB1036pyr5′d. After confirming vector integration at target sites, transformants were subjected to selection

using 5-FOA, and strains with $5'\Delta pyrG$ at the target site of chromosome 2 were isolated (Additional file 1: Figure S6B). The strains B1036pyrG5′Δ and ku70-B1036pyrG5′Δ were transformed with pCT166pyr3′d. Strains with pCT166pyr3′d integrated at the target site were selected using 5-FOA to obtain parental strains for generating the strains B1036p5′ΔCTp3′Δ and ku70-B1036p5′ΔCTp3′Δ containing translocated chromosomal duplication (Additional file 1: Figure S6C). Construction of parental strain for translocated triplication was conducted as follows. The del1546 K strain, obtained by removing *pyrG* from chromosome 4 of the J4 strain (Additional file 1: Figure S4), was transformed with pS11pyr3′d. After confirming vector integration of chromosome 7, transformants were subjected to selection using 5-FOA. The strains with $3'\Delta pyrG$ at the target site of chromosome 7 (K1-IF5-2-5FOA) were isolated (Additional file 1: Figure S4).

Construction of *prtT* overexpression strains

The *prtT* overexpression vector pAPRT was constructed as follows. A 3.7 kb *amyB* fragment, amplified from RIB40 genome DNA using the primers amyU and amyl, was cloned by a TA cloning kit (TOYOBO) to generate pAmyTA. A 6.7-kb fragment, amplified from pAmyTA using the primer pair Amy118UIF/TA816UIF, and a 2.9-kb *pyrG* fragment amplified from pBP9 [12] using the primer pair P363U/P3229L, were fused using an In-Fusion-Cloning Kit (TaKaRa) to obtain pAmyPYR. A 7.6-kb fragment amplified from pAmyPYR, and a 2.1 kb fragment containing *prtT* ORF and amplified from RIB40 genomic DNA using the primer pair prtTstU/prtTteL, were fused using an In-Fusion-Cloning Kit (TaKaRa) to obtain the *prtT* overexpression vector pAPRT. The RP1, del1546K4, and TLTAS11B strains were transformed with pAPRT in a circular state to obtain the APRT, 1546 K-APRT, and TLTA-APRT strains, respectively. The copy numbers of the pAPRT vector integrated in the strains were estimated by quantitative PCR (Additional file 1: Figure S7). The APRT and TLTA-APRT strains contained one copy of APRT. The 1546 K-APRT strain contained two copies of pAPRT. Determining the copy number of pAPRT in the strains via Southern blot was difficult. This is because *A. oryzae* strains originally contain three copies of amylase genes, and the duplicated chromosomal region contained *prtT* and amylase genes. The results of PCR suggested that one copy of pAPRT in the 1546 K-APRT strain was integrated at one of the amylase loci (data not shown). The other copy of pAPRT appeared to be randomly integrated in the strains.

Solid-state cultivation and measurements of enzyme activity

Solid-state cultures were generated by inoculating 1×10^8 conidiospores into 5 g of wheat bran medium in a 150-mL Erlenmeyer flask and incubating at 30 °C for 4 days. Water (50 mL) was then added to the flask and extracted after shaking. The liquid fraction was filtered through filter paper and used as sample extract. Protease activity was determined by mixing the substrate and 2% milk casein (pH 7) with sample extracts and incubating at 30 °C for 20 min. Reactions were stopped using trichloroacetic acid, and quantities of liberated amino acids were measured at 660 nm using a tyrosine standard. One unit (U) was defined as the amount of enzyme yielding 1 μg of tyrosine per min at 30 °C and at pH 7. The activity of alpha-amylase was determined using Alpha-Amylase Activity kits (Kikkoman Biochemifa, Minato-ku, Japan) according to the manufacturer's instructions. Briefly, reaction buffer containing 100 mM acetate (pH 5), 2 mM 2-chloro-4 nitrophenyl 6^5-azido-6^5-deoxy-beta-maltopentaocid, 50 U/mL glucoamylase, 6 U/mL beta-glucosidase, 50 mM NaCl, and 2 mM $CaCl_2$ was mixed with 300 volumes of diluted sample extract and incubated for 5 min at 37 °C. The reactions were stopped using trichloroacetic acid, and liberated 2-chloro-4 nitrophenol contents were determined at 400 nm. One unit (U) was defined as the amount of enzyme yielding 1 μmol of 2-chloro-4 nitrophenol per min.

The activity of acid carboxypeptidase was determined using Acid Carboxypeptidase Assay kits (Kikkoman Biochemifa) according to the manufacturer's instructions. The reaction buffer containing 50 mM acetate (pH 3), 0.5 mM Cbz-Tyr-Ala (carboxybenzoxy-L-tyrosyl-L-alanine), and 5 mM NAD was mixed with 30 volumes of diluted sample extract and incubated for 10 min at 37 °C. The reaction was stopped using 0.5 M Tris–HCl buffer (pH 8.5) containing 5 mM WST-8 and 13 U of alanine dehydrogenase, and the mixture was incubated at 37 °C for 20 min. Subsequently, 0.5 mM 1-methoxy-5-methyl-phenazinium methylsulfate was added to reaction mixture and incubated at 37 °C for 10 min. Quantities of liberated NADH were measured at 460 nm. One unit was defined as the amount of enzyme that liberated 1 μmol of L-alanine from Cbz-Tyr-Ala per min.

Competing interests
The authors declare that they have no competing interests.

References

1. Chan JE, Kolodner RD. A genetic and structural study of genome rearrangements mediated by high copy repeat Ty1 elements. PLoS Genet. 2011;7:e1002089.
2. Bailey JA, Liu G, Eichler EE. An *Alu* transposition model for origin and expansion of human segmental duplication. Am J Genet. 2003;73:823–34.
3. Deininger P. Genetic instability in cancer: caretaker and gatekeeper genes. Ochsner J. 1999;1:206–9.
4. Sadikovic B, Al-Romaih K, Squire JA, Zielenska M. Cause and consequences of genetic and epigenetic alternations in human cancer. Curr Genom. 2008;9:364–408.
5. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, et al. Genome sequencing and analysis of *Apergillus oryzae*. Nature. 2005;438:1157–61.
6. Sato A, Oshima K, Noguchi H, Ogawa M, Takahashi T, Oguma T, et al. Draft genome sequencing and comparative analysis of *Aspergillus sojae* NBRC4239. DNA Res. 2011;18:165–76.
7. Takahashi T, Masuda T, Koyama Y. Identification and analysis of *Ku70* and *Ku80* homologs in the *koji* molds *Aspergillus sojae* and *Aspergillus oryzae*. Biosci Biotechnol Biochem. 2006;70:135–43.
8. Takahashi T, Masuda T, Koyama Y. Enhanced gene targeting frequency in *ku70* and *ku80* disruption mutants of *Aspergillus sojae* and *Aspergillus oryzae*. Mol Genet Genomics. 2006;275:460–70.
9. Takahashi T, Jin F, Sunagawa M, Machida M, Koyama Y. Generation of large chromosomal deletions in koji molds *Aspergillus oryzae* and *Aspergillus sojae* via a loop-out recombination. Appl Environ Microbiol. 2008;74:7684–893.
10. Takahashi T, Jin F, Koyama Y. Non-homologous end-joining deficiency allows large chromosomal deletions to be produced by replacement-type recombination in *Aspergillus oryzae*. Fungal Genet Biol. 2009;46:815–24.
11. Sopoko R, Huang D, Preston N, Chua G, Papp B, Kafadar K, Snyder M, Oliver SG, Cyert M, Hughes TR, Boone C, Andrews B. Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell. 2006;21:319–30.
12. Takahashi T, Sato A, Ogawa M, Hanya Y, Oguma T. Targeted tandem duplication of a large chromosomal segment in *Aspergillus oryzae*. Appl Environ Microbiol. 2014;80:4547–58.
13. Bosco G, Haber JE. Chromosome break-induced DNA replication leads to nonreciprocal translocation and telomere capture. Genetics. 1998;150:1037–47.
14. Kraus E, Leung W, Haber JE. Break-induced replication: a review and an example in budding yeast. PNAS. 2001;98:8255–62.
15. Saini N, Ramarkrishnan S, Elango R, Ayyar S, Zhang Y, Deem A, et al. Migrating bubble during break-induced replication drives conservative DNA synthesis. Nature. 2013;502:389–92.
16. Yaver DS, Lamsa M, Munds R, Brown SH, Otani S, Franseen L, et al. Using DNA-tagged mutagenesis to improve heterologous protein production in *Aspergillus oryzae*. Fungal Genet Biol. 2000;29:28–37.
17. Takahashi T, Ogawa M, Koyama Y. Analysis of the functions of recombination-related genes in the generation of Large chromosomal deletions by loop-out recombination in *Aspergillus oryzae*. Eukaryot Cell. 2012;11:507–17.
18. Mizutani O, Masaki K, Gomi K, Iefuji H. Modified Cre-loxP recombination in *Aspergillus oryzae* by direct introduction of Cre recombinase for marker gene rescue. Appl Environ Microbiol. 2012;78:4126–33.
19. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 2000;28:33–6.
20. Noguchi Y, Sano M, Kanamaru K, Ko T, Takeuchi M, Kato M, Kobayashi T. Genes regulated by AoxlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol. 2009;85:141.
21. Murakami K, Ishida Y, Masaki A, Tatsumi H, Murakami S, Nakano E, Motai H, Kawabe H, Arimura H. Isolation and characterization of the alkaline protease gene of *Aspergillus oryzae*. Agric Biol Chem. 1991;55:2807–11.
22. Punt PJ, Schuren FH, Lehmbeck J, Christensen T, Hjort C, van den Hondel CA. Characterization of the *Aspergillus niger prtT*, a unique regulator of extracellular protease encoding genes. Fungal Genet Biol. 2008;45:1591–9.
23. Fasullo MT, Davis RW. Direction of chromosome rearrangements in *Saccharomyces cerevisiae* by use of his3 recombinational substrates. Mol Cell Biol. 1998;8:4370–80.
24. Takahashi T, Hatamoto O, Koyama Y, Abe K. Efficient gene disruption in the koji-mold *Aspergillus sojae* using a novel variation of the positive-negative method. Mol Gen Genomics. 2004;272:344–52.
25. Ogawa M, Tokuoka M, Jin F, Takahashi T, Koyama Y. Genetic analysis of conidiation regulatory pathways in koji-mold *Aspergillus oryzae*. Fungal Genet Biol. 2010;47:10–8.
26. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207–10.

Physiological characterization of secondary metabolite producing *Penicillium* cell factories

Sietske Grijseels[1] , Jens Christian Nielsen[2], Jens Nielsen[2,3], Thomas Ostenfeld Larsen[1], Jens Christian Frisvad[1], Kristian Fog Nielsen[1], Rasmus John Normand Frandsen[1]* and Mhairi Workman[1]

Abstract

Background: *Penicillium* species are important producers of bioactive secondary metabolites. However, the immense diversity of the fungal kingdom is only scarcely represented in industrial bioprocesses and the upscaling of compound production remains a costly and labor intensive challenge. In order to facilitate the development of novel secondary metabolite producing processes, two routes are typically explored: optimization of the native producer or transferring the enzymatic pathway into a heterologous host. Recent genome sequencing of ten *Penicillium* species showed the vast amount of secondary metabolite gene clusters present in their genomes, and makes them accessible for rational strain improvement. In this study, we aimed to characterize the potential of these ten *Penicillium* species as native producing cell factories by testing their growth performance and secondary metabolite production in submerged cultivations.

Results: Cultivation of the fungal species in controlled submerged bioreactors showed that the ten wild type *Penicillium* species had promising, highly reproducible growth characteristics in two different media. Analysis of the secondary metabolite production using liquid chromatography coupled with high resolution mass spectrometry proved that the species produced a broad range of secondary metabolites, at different stages of the fermentations. Metabolite profiling for identification of the known compounds resulted in identification of 34 metabolites; which included several with bioactive properties such as antibacterial, antifungal and anti-cancer activities. Additionally, several novel species–metabolite relationships were found.

Conclusions: This study demonstrates that the fermentation characteristics and the highly reproducible performance in bioreactors of ten recently genome sequenced *Penicillium* species should be considered as very encouraging for the application of native hosts for production via submerged fermentation. The results are particularly promising for the potential development of the ten analysed *Penicillium* species for production of novel bioactive compounds via submerged fermentations.

Keywords: *Penicillium*, Submerged fermentation, Physiology, Secondary metabolite, Cell factory

Background

Filamentous fungi are important producers of secondary metabolites: low-molecular-weight compounds that often have bioactive properties. The genus *Penicillium* currently includes more than 354 accepted species [1]

many of which are capable of producing a wide variety of secondary metabolites [2]. The most well-known secondary metabolite produced by *Penicillium* is the antibiotic penicillin, which was discovered by Fleming [3] and which is nowadays produced in large scale using *P. rubens*, following intense strain improvement programs aimed at increasing the titers. Other important pharmaceutical compounds produced by *Penicillium* species include the antifungal griseofulvin [4], the

*Correspondence: rasf@bio.dtu.dk
[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
Full list of author information is available at the end of the article

immunosuppressant mycophenolic acid [5] and the cholesterol lowering drug compactin/mevastatin [6–8]. These examples illustrate the great importance of *Penicillium* species as production hosts and sources of bioactive compounds with medical applications. On the other hand, *Penicillium* species can also produce mycotoxins such as citrinin, ochratoxin and patulin [2], which can pose a health risk to humans and animals.

The increasing prevalence of antibiotic resistance among pathogenic bacteria was in 2016 named by the UN General Assembly to be amongst the greatest and most urgent global risk factor for the health of humans [9]. However, the discovery of novel antibiotics has stagnated since the 1970's and only few novel classes of antibiotics have been discovered since then [10, 11]. This is most likely a consequence of the fact that drug development is time consuming and expensive, while market value and dominance for antibiotics is limited by the development of resistance. The broad metabolic diversity found within the fungal kingdom, however, continues to provide a rich source of novel drug leads, this coupled with advancements within process engineering and physiological characterization tools, provides a strong basis for exploiting the chemical and physiological diversity in fungi to establish novel fungal based bioprocesses which can meet the demands of modern society for novel antimicrobials and pharmaceuticals.

To facilitate the process from identification of a beneficial compound to the establishment of an industrially feasible bio-based production process, two major strategies can be followed; i.e. find and optimizing the best native producer or transfer the involved genes to an established heterologous host. The latter strategy is typically desirable if the native producer is not suitable for industrial fermentation processes or is genetically inaccessible. However, production of secondary metabolites in heterologous hosts is still in the developmental stage and expressing complex multistep pathways to commercially viable titers has proven highly challenging and typically requires substantial optimization [12]. The transfer of the given genetic trait in addition requires a thorough understanding of the given biosynthetic pathway and the involved genes. The other strategy, advancement of the native organism for the production of secondary metabolites, can be favourable as this eliminates the often tedious work of moving the biosynthetic pathway to a heterologous host. However, taming a uncharacterized filamentous fungus for industrial scale production is not trivial, as natural secondary metabolite yields are often low which combined with the complex interplay between physiology, physical environment and morphology often causes problems during fermentation such as nonhomogeneous mixing and nutrient limitation [13].

Therefore, in many instances, relatively unexploited fungi are not tested in submerged bioreactor cultivations, even though this would provide the necessary process performance data to evaluate whether to optimize the native producing organism or to transfer the pathway into a heterologous host.

The recent genome sequencing of ten *Penicillium* species [14, 15] has provided valuable novel insight into the rich diversity in secondary metabolite gene clusters present in these *Penicillium* species and has in addition, made the species amenable to rational strain improvement. In this study, we analyzed the potential of these ten recently genome sequenced *Penicillium* species as native producing cell factories. The species were cultivated in highly controlled bioreactors both in a defined medium that allowed for quantitative physiological evaluation and in a secondary metabolite inducing complex medium, to evaluate overall process performance, growth rates and reproducibility. Additionally, secondary metabolite profiles of the species during different time points in the fermentation were studied. To our knowledge, none of these species had been tested in a bioreactor before. This work demonstrates an implementation strategy for novel fungal cell factories based on highly controlled submerged bioreactors cultivations and quantification of key physiological parameters.

Methods

Organisms

The species used in this study were *P. coprophilum* (IBT31321), *P. nalgiovense* (IBT 13039), *P. polonicum* (IBT 4502), *P. antarcticum* (IBT31811), *P. vulpinum* (IBT 29486), *P. arizonense* (IBT 12289), *P. solitum* (IBT 29525), *P. decumbens* (IBT11843), *P. flavigenum* (IBT 14082), and *P. steckii* (IBT 24891). All strains are available from the IBT culture collection (Department of Biotechnology and Biomedicine, Technical University of Denmark).

Submerged bioreactor batch cultivations

With the aim of investigating the growth characteristics of ten potential secondary metabolite producing *Penicillium* species, submerged bioreactor cultivations were performed with each species in two different media in biological triplicates.

Media

Czapek yeast autolysate (CYA) medium was used for spore propagation, containing per liter of demineralized water: 30 g sucrose, 5 g yeast extract, 3 g NaNO₃, 1 g K₂HPO₄, 0.5 g MgSO₄·7H₂O, 0.5 g KCl, 0.01 g FeSO₄·7H₂O, 20 g agar and 1 mL trace metal solution containing 0.1 g/L ZnSO₄·7H₂O and 0.05 g/L CuSO₄·5H₂O. The pH was adjusted to 6.2 with NaOH

prior to autoclaving. All chemicals applied were obtained from Sigma-Aldrich (Sigma-Aldrich).

Batch cultivations were performed in CY medium (CM) and in a defined medium (DM) for *Penicillium.* The defined medium contained per liter of demineralized water: 15 g glucose, 3.5 g $(NH_4)_2SO_4$, 0.5 g $MgSO_4 \cdot 7H_2O$, 0.15 g Na_2-EDTA, 0.04 g $FeSO_4 \cdot 7H_2O$, 0.8 g KH_2PO_4, 5 mL trace solution containing 1 g/L $CuSO_4 \cdot 5H_2O$, 4 g/L $ZnSO_4 \cdot 7H_2O$, 4 g/L $MnSO_4 \cdot 7H_2O$, 1 g/L $CaCl_2 \cdot H_2O$. The CY medium contained the same components as stated above for spore propagation without the agar and 15 g sucrose per liter. The medium was heat sterilized by autoclaving at 121 °C for 20 min. After media sterilization, a sucrose solution (for CM) or glucose solution (for DM), separately autoclaved at 121 °C, was added.

Preparation of spore inoculum

Spores were propagated on CYA plates for 7 days and harvested with approximately 5 mL of cold MiliQ water containing 0.9% NaCl and 0.01% Tween solution, followed by filtration through Miracloth (Merck Millipore), centrifuged at 5000 g and the spore pellet washed again with 10 mL MiliQ water. Spores were counted using a Bürker–Türk counting chamber, followed by inoculation of bioreactors to a final spore concentration of 1×10^9 spores/L.

Bioreactors

All bioreactor batch cultivations were carried out in Sartorious 1 L bioreactors (Satorious Stedim Biotech) with a working volume of 0.9 L, equipped with 2 Rushton six-blade disc turbines. Throughout the cultivation the temperature was maintained at 25 °C and the pH was kept constant at 6.5 by automatic addition of 2 M NaOH or H_2SO_4. The bioreactors were sparged with sterile atmospheric air. For the first 1200 min, the airflow was linearly increased from 0.1 volume of air per volume of liquid per minute (vvm) to 0.9 vvm and the stirring rate from 100 to 600 rpm. After that, these parameters were kept constant. The off-gas concentrations of oxygen and carbon dioxide were measured with a Prima Pro Process Mass Spectrometer (Thermo-Fischer Scientific), calibrated monthly with gas mixtures containing 5% (v/v) CO_2, 0.04% (v/v) ethanol and methanol, 1% (v/v) argon, 5% (v/v) and 15% (v/v) oxygen all with nitrogen as carrier gas. The pH electrode (Mettler Toledo) was calibrated according to manufacturer's standard procedures.

Biomass dry weight determination

Growth rates were determined based on biomass dry weight samples collected during the exponential growth phase. First a known amount of cell culture (typically around 3 mL) was filtered through pre-dried and weighted 0.45 μm polyether sulfone filters (Sartorius Stedim Biotech) which was then washed with demineralized water. The filter was folded to lock the biomass inside and dried in a microwave at 150 W for 20 min. The dry weight was measured after a cooling period of approximately 30 min in a desiccator.

Glucose concentration determination

For quantification of the glucose concentration in the culture medium, fermentation broth was filtered through a Q-Max® Ca-Plus Filter (Frisenette ApS) with a pore size of 0.45 μm and stored at −20 °C until analysis. Separation and detection of the compounds was accomplished with a high performance liquid chromatography (HPLC) system equipped with a Bio-Rad Aminex HPX-87H column (BioRad) coupled to a RI detector. Elution was performed isocratically with H_2SO_4 (5 mM) as the mobile phase with a flow velocity of 0.6 mL/min at 60 °C. Quantification was performed using a six-level external calibration curve.

Calculations of physiological characteristics

The end of the lag phase was determined as the time point were the percentage of CO_2 in the off gas was consecutively over 6%.

Maximum specific growth rates were calculated via linear regression on a semi-logarithmic scale on at least three experimental biomass dry weight data points during the exponential phase and all R^2 values were at least 0.96. Maximum CO_2 production rates where determined similarly but with an R^2 of at least 0.99. A custom script was developed in the statistical programming language R to determine the duration of the non-exponential growth phase for all cultivations, defined as the time difference between the end of the exponential fit to the CO_2 exhaust, and the time point where the maximum overall CO_2 production was reached. The code is available at: https://github.com/JensChrNielsen/Fungal-Biol-Biotechnol-2017.

The overall biomass yield on glucose was calculated as the ratio between the biomass gain and the glucose consumption in the growth phase.

Secondary metabolite analysis

With the aim of identifying the secondary metabolite profiles in the fermentation broth of the ten *Penicillium* species grown in two different media, the filtered broth was extracted with ethyl acetate, analyzed with ultra-high performance liquid chromatography-diode array detection-quadrupole time of flight mass spectrometry UHPLC–DAD–QTOFMS, and known compounds were identified using an in-house compound library.

Extraction

The fermentation broth was filtered through a 0.45 μm cellulose acetate filter (Frisenette), for extraction of non-polar compounds 1.0 mL of ethyl acetate extraction solvent was added to 0.8 mL of filtrate. After ultra-sonification for 1 h and centrifugation for 1 min at 13,000g the organic (top) layer was transferred to a clean vial. For extraction of polar compounds 1.0 mL of acidic ethyl acetate (ethyl acetate + 0.5% formic acid) was added to the remaining filtrate. Following ultra-sonification and centrifugation the organic (top) layer was added to the extraction solution containing the extracted non-polar compounds. The combined solution was evaporated to dryness and re-dissolved in 125 μL HPLC grade methanol. After centrifugation for 5 min at 13,000g the supernatant was directly used for chemical analysis.

UHPLC–DAD–QTOFMS analysis

Secondary metabolite analysis was achieved by UHPLC–DAD–QTOFMS on a maXis HD orthogonal acceleration quadrupole time-of-flight mass spectrometer (Bruker Daltonics) equipped with an electrospray ionization (ESI) source and coupled to an Ultimate 3000 UHPLC system (Dionex, Thermo Scientific). Separation was achieved with a Kinetex 2.6 μm C$_{18}$, 100 × 2.1 mm (Phenomenex) column maintained at 40 °C with a flow rate of 0.4 mL/min. A linear gradient system composed of 20 mmol/L formic acid in water, and 20 mmol/L formic acid in acetonitrile was used, starting from 10% (v/v) acetonitrile and increased to 100% in 10 min, maintaining this rate for 3 min before returning to the starting conditions. The system was re-equilibrated before subsequent sample analysis. MS analysis was performed in ESI$^+$ with a data acquisition rate of 10 scans per second at m/z 100–1000, switching between 0 and 20 eV fragmentation energy.

Identification of secondary metabolites

Identification of secondary metabolites was performed using aggressive dereplication of the full scan high resolution (HR) MS data, and pseudo MS/MS data from the 20 eV fragmentation trace, using a search list of compounds based on former taxonomic identification and a manual search of major peaks in the internal library. This library consists of 1500 compounds of which 95% are fungal secondary metabolites [16]. Compounds were confirmed by comparison of HRMS, UV/Vis and MS/HRMS to a reference standard. The number of peaks was counted by making a list of compounds (dereplicated and unknown ones) based on molecular features in the latest time sample in the fermentation, with an absolute area higher than 500,000 counts and intensity higher than 10,000 counts. Subsequently, the presence of these compounds was searched for in all the other samples. The

total amount was determined by counting all compounds that were present in two or three of the triplicate samples, without using a threshold value.

Results

Physiology

Ten different recently genome sequenced wild type *Penicillium* species were cultivated in 1 L controlled bioreactors to evaluate their behaviour in submerged cultivations. The species were grown in both a defined medium for *Penicillium* (DM) and in the more industrially relevant medium Czapek Yeast Autolysate (CM). During the course of the fermentations the CO$_2$ off gas was continuously measured, while biomass dry weight was sampled until CO$_2$ off gas values started declining. Additionally, for DM the glucose concentration was determined at regular time intervals throughout the fermentations. All cultivations were conducted in biological triplicates to evaluate the reproducibility of the experiments. The work revealed that the ten tested *Penicillium* species displayed highly diverse growth characteristics and responded differently to the two tested nutritional regimes, as evident from the growth curves (Fig. 1).

The initial lag phase, in which minimal changes in the CO$_2$ levels were observed, lasted between 11 and 38 h on DM, while it was shorter for all species on CM lasting between 10 and 18 h, post inoculation (Table 1). In particular, *P. flavigenum* showed a short lag phase on both media where exponential growth was reached after 14 and 11 h for DM and CM, respectively. *Penicillium vulpinum* showed the longest lag phase in DM and CM of 38 and 18 h, respectively. Microscopic examination of culture broth from the ten species revealed that the lag phase was characterized by germination of spores followed by elongation of hyphae. After the lag phase, the exponential phase was initiated, where the growth rate reached its maximum. Here, the morphology of the ten species varied as a function of the species and media. In CM, all species initially grew as freely dispersed mycelium until the maximum CO$_2$ off gas was reached, after which the formation of pellets typically started. In the DM, five of the ten species showed pellet formation relatively early in the fermentation; *P. steckii* and *P. solitum* showed the formation of small dense pellets, *P. vulpinum* and *P. coprophilum* showed a mixture of pellets and clumps, and *P. flavigenum* showed a mixture of small clumps and dispersed growth. The remaining species and all species in CM showed dispersed growth at least until the time of maximum CO$_2$ off gas (Table 1).

The duration of the exponential phase proved to be correlated with the observed growth morphology, where the exponential growth transitioned into non-exponential growth at the same time as pellet or clump formation
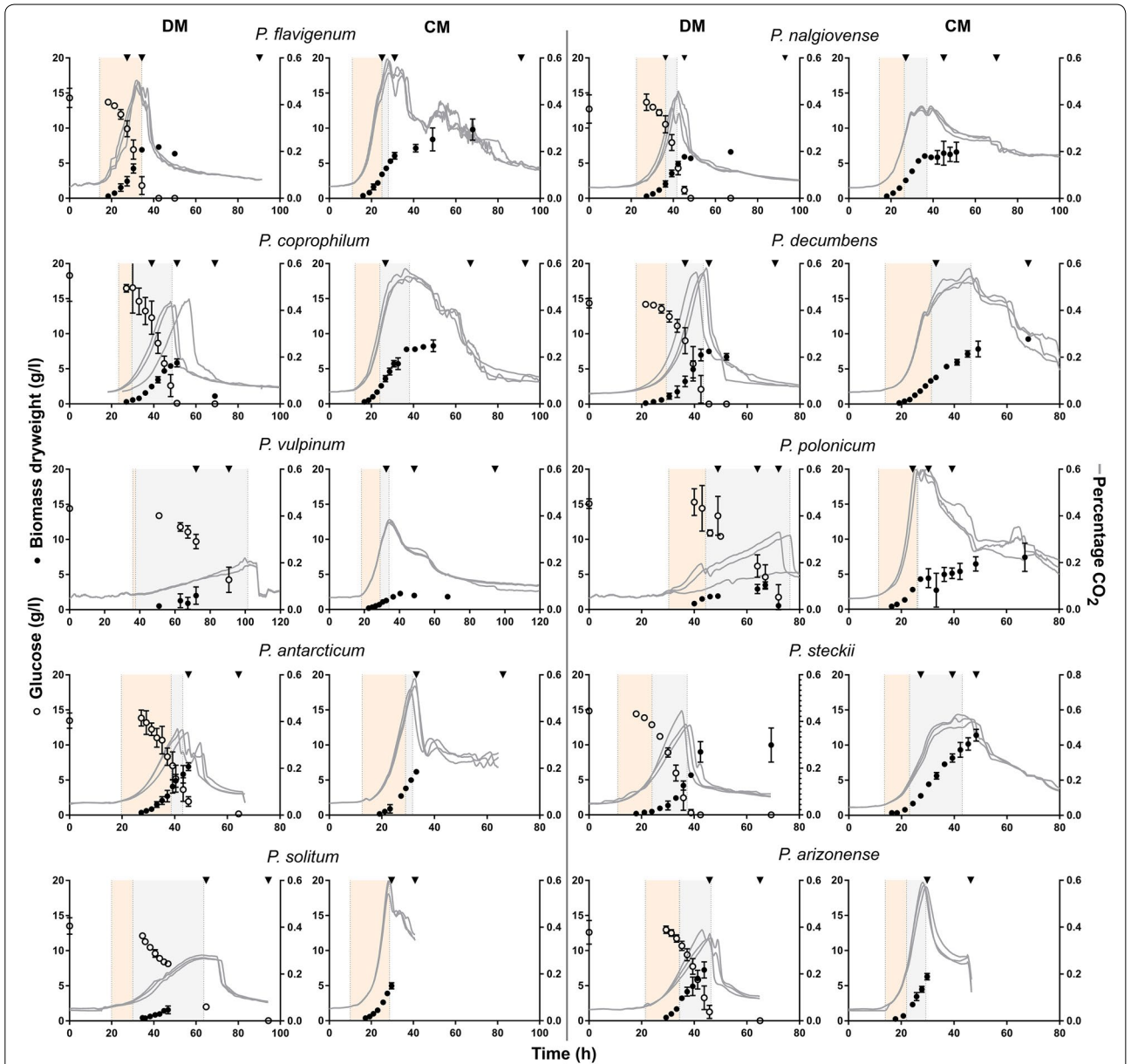
Fig. 1 Fermentation profiles of ten *Penicillium* species cultivated in 1 L bioreactors in DM and CM. All fermentations were performed in biological triplicates. The CO_2 exhaust values (%) are shown for all triplicate experiments separately as solid lines (right Y-axis), the dry weight (g/L) values are shown as a mean with standard deviations as black circles (left Y-axis) and for the defined medium the glucose concentrations are shown as mean with standard deviations as open circles (left Y-axis). Light orange shaded area shows the exponential phase and light grey shaded area shows the non-exponential phase. For *P. polonicum* one of the triplicate experiments showed a different growth curve, for this fermentation only the CO_2 profile is shown. Triangles are the time points were samples for secondary metabolite analysis were taken. The third sample point in *P. decumbens* was at 120 h and falls therefore outside the figure. For *P. steckii* in DM the samples for secondary metabolite analysis were taken from a different experiment with similar growth characteristics

occurred. For seven of the ten species, the non-exponential growth phase was more than ten h in DM, while this was the case for four of ten species in the CM (Fig. 2). Only *P. flavigenum* and *P. antarcticum* showed exponential growth and dispersed morphology during the entire

growth phase in both media (Fig. 2 and Additional file 1). In contrast, *P. vulpinum* showed dispersed growth and grew exponentially during the entire growth phase in CM, while in DM it had pelleted morphology and grew exponentially until 36 h, where after the CO_2 off gas

Table 1 Physiological characteristics of ten *Penicillium* species cultivated in 1 L bioreactors in DM and CM

Species	Biomass yield growth phase (Y_{sx}) DM (g DW/g glucose)	Lag phase DM (h)	Morphology DM	Lag phase CM (h)	Morphology CM
P. flavigenum	0.58 ± 0.05	14.3 ± 0.5	Small clumps/dispersed	10.9 ± 0.3	Dispersed
P. nalgiovense	0.60 ± 0.17	22.5 ± 0.8	Dispersed	14.5 ± 0.1	Dispersed
P. coprophilum	0.37 ± 0.03	23.3 ± 0.6	Pellet	12.2 ± 0.6	Dispersed
P. decumbens	0.55 ± 0.02	17.8 ± 0.9	Dispersed	13.8 ± 0.2	Dispersed
P. vulpinum	0.54 ± 0.14	37.6 ± 1.4	Pellet + clumps + wall	18.3 ± 0.2	Dispersed
P. polonicum	0.25 ± 0.01	30.3 ± 1.7	Wall growth	11.3 ± 0.1	Dispersed
P. antarcticum	0.58 ± 0.03	19.7 ± 0.2	Dispersed	12.5 ± 0.1	Dispersed
P. steckii	0.37 ± 0.02	10.9 ± 2.3	Pellet + wall growth	13.5 ± 0.2	Dispersed
P. solitum	0.29 ± 0.13	19.9 ± 1.2	Pellet + wall growth	9.9 ± 0.2	Dispersed
P. arizonense	0.67 ± 0.04	21.4 ± 0.4	Dispersed	14.0 ± 0.8	Dispersed

Biomass yield on glucose (g DW/g glucose), duration of lag phase (h) and observed morphology at the time point where the CO_2 was at its maximum

Fig. 2 Length of non-exponential growth phase of ten *Penicillium* species cultivated in 1 L bioreactors in DM and CM. Non-exponential growth phase was the time in hours between the end of exponential growth phase and the maximum CO_2 off gas value. Bars show average and standard deviation if triplicate submerged fermentations. *P. flavigenum* in DM, *P. polonicum* in CM and *P. solitum* in CM showed exponential growth until the maximum CO_2 off gas value

values were still increasing, but in a non-exponential manner, until after 110 h maximum values of exhausted CO_2 off gas were reached.

Maximum specific growth rates (μ^{max}) were calculated based on dry weight values during the exponential growth phase. For several species in DM it proved impossible to determine accurate growth rates based on biomass dry weight due to insufficient samples during the short exponential phase. As an alternative method for the growth rate, the rates of CO_2 production were determined, as this is directly proportional to the amount of metabolically active biomass present, assuming that the yield and specific growth rate remain constant [17]. The

species showed growth rates between 0.19 and 0.24 h^{-1} in DM and between 0.2 and 0.35 h^{-1} in CM (Fig. 3). The biomass yields on glucose during the growth phase (Y_{sx}) varied highly among species; between 0.25 g DW/g glucose for *P. polonicum* to 0.67 g DW/g glucose for *P. arizonense.* The species with the lowest Y_{sx} values, *P. polonicum*, *P. steckii*, *P. solitum* and *P. coprophilum* (all under 0.4 g DW/g glucose) all showed pellet formation and/or wall growth, while species with high yields, *P. nalgiovense* and *P. arizonense* (0.6 and 0.67 g DW/g glucose), showed dispersed growth.

Secondary metabolite analysis
Samples taken during different stages of the submerged cultivations were filtered, extracted, and analyzed for the presence of secondary metabolites in the fermentation medium using reverse phase UHPLC coupled to time-of-flight (QTOF) MS. Analysis of the biological triplicate experiments for each of the tested species in either of the two media were found to be highly reproducible based on the occurrence of base peak chromatogram (BPC) peaks with the same retention times and comparable peak areas in the normalized chromatograms (examples given in Fig. 4).

To allow for identification of the individual compounds found in the samples, the effluent from the UHPLC column was directly analysed by the attached QTOFMS which allowed for the recording of accurate mass. For each sample taken at the end of the cultivation the amount of molecular features (equivalent with the number of different compounds) above an area threshold of 500,000 and an intensity of 10,000 counts in the UHPLC–QTOFMS data was determined. The ten different species each displayed a unique metabolic profile in the two tested media (Table 2 and Additional file 2). A minimum of two compounds was observed in *P. flavigenum*

Fig. 3 Maximum specific growth rates and maximum CO_2 production rates for ten *Penicillium* species cultivated in 1 L bioreactors in DM and CM. Dark grey bars: maximum specific growth rates μ^{max} (h^{-1}) and standard deviations calculated by linear regression of minimum three biomass dry weight data points during the exponential phase. For *P. vulpinum*, *P. polonicum*, *P. steckii*, and *P. solitum* in DM there were less than three dry weight samples during exponential phase and therefore μ^{max} could not be calculated. Light grey bars: maximum production rate of CO_2 (h^{-1}) and standard deviations calculated by linear regression of accumulated CO_2 values during exponential growth. All values are based on independent triplicate fermentations, except for the ones with an asterix; there the values are based on duplicate fermentations

Fig. 4 Base peak chromatograms (BPCs) of extracted samples of biological triplicates of *P. nalgiovense* and *P. vulpinum* fermentations in CM and DM. The BPCs of the triplicates are shown in red, dark orange and light orange. BPCs show peaks with the similar retention times and comparable peak areas. Peak with asterisk showed a slight change of retention times but have the same molecular features

cultivated in DM, and a maximum of 34 compounds in *P. steckii* cultivated in CM (last column in Table 2). In eight out of the ten species a greater number of compounds were detected in the cultivations conducted in CM compared to DM; exceptions were *P. polonicum*, where cultivation in DM showed a greater number of peaks than in CM, and *P. solitum*, where the number of compounds was equal in DM and CM.

Additionally, metabolite profiling was performed to identify the known compounds produced during the fermentations. In total 34 different metabolites were identified, from which most of them could be confirmed by

comparison of HRMS, UV/Vis and MS/HRMS to a reference standard. Most compounds were unique to one of the ten species, with the exception of chrysogine identified in three species (*P. nalgiovense*, *P. flavigenum* and *P. antarcticum*), fungisporin in three species (*P. coprophilum*, *P. nalgiovense* and *P. flavigenum*) and andrastin A in two species (*P. vulpinum* and *P. decumbens*), while andrastin C was only identified in *P. decumbens*. Additionally, the roquefortine pathway was identified in three species (*P. coprophilum*, *P. vulpinum* and *P. flavigenum*), where depending on the media and species, different intermediates and end products of this pathway could be

Table 2 Metabolites produced in different growth phases of ten *Penicillium* species

Species	Media	Identified metabolites	Total amount of molecular features per sample (id number of metabolite)		
			Sample 1	Sample 2	Sample 3
P. coprophilum	DM	neoxaline[a] (1), meleagrin[a] (2), oxaline[a] (4), roquefortine C[a] (5), dechlorogriseofulvin[a] (6), griseofulvin[a] (7)	10 (1, 2, 4, 5)	13 (1, 2, 4, 5)	14 (1, 2, 4, 5, 6, 7)
	CM	neoxaline[a] (1), meleagrin[a] (2), N6-formyl-roquefortine-c (3), oxaline[a] (4), fungisporin (6), dechlorogriseofulvin[a] (7), griseofulvin[a] (8), dehydrogriseofulvin[a] (9)	10 (2, 3, 4)	18 (1, 2, 3, 4, 7, 8, 9)	19 (1, 2, 3, 4, 6, 7, 8, 9)
P. nalgiovense	DM	chrysogine[a] (10), citreo isocumarin[a] (11), citreo-isocumarin analog-5 (12), diaportinol[a] (13), diaportinic acid[a] (14), 6-methoxy-citreoisocoumarin[a] (15) fungisporin (9)	0	2 (10, 14)	10 (10, 11, 12, 13, 14, 15, 9)
	CM	chrysogine[a] (10), citreo isocumarin[a] (11), citreo-isocumarin analog-5 (12), diaportinol[a] (13), diaportinic acid[a] (14), 6-methoxy-citreoisocoumarin[a] (15) fungisporin (9)	5 (14, 15)	17 (10, 11, 12, 13, 14, 15, 9)	18 (10, 11, 12, 13, 14, 15, 9)
P. vulpinum	DM	desoxypatulinic acid (16), patulin[a] (17), neopatulin (18), meleagrin[a] (2), oxaline[a] (4), (-)-3-butyl-7-hydroxyphthalide (19), 3-[2'(R)hydroxybutyl]-7-hydroxyphthalide (20), andrastin A[a] (21)		17 (16, 17, 2, 4, 20, 21)	21 (16, 17, 2, 4, 19, 20, 21)
	CM	desoxypatulinic acid (16), patulin[a] (17), neopatulin (18), meleagrin[a] (2), oxaline[a] (4), (-)-3-butyl-7-hydroxyphthalide (20), andrastin A[a] (21)	23 (16, 17, 2, 4, 19, 20, 21)	27 (16, 17, 2, 4, 19, 20, 21)	28 (16, 17, 2, 4, 19, 20, 21)
P. polonicum	DM	l-tyrosin[a] (22), penicillic acid[a] (23), leucyltryptophanyl-diketopiperazine[a] (24), aspterric acid[a] (25), verrucofortine[a] (26),	10 (22, 24, 25, 26)	19 (22, 23, 24, 25, 26)	21 (22, 23, 24, 25, 26)
	CM	l-tyrosin[a] (22), leucyltryptophanyl-diketopiperazine[a] (24), aspterric acid[a] (25), verrucofortine[a] (26),	2 (26)	4 (25, 26)	16 (22, 24, 25, 26)
P. decumbens	DM	andrastin A[a] (27), andrastin C (28)	3 (27)	4 (27, 28)	16 (27, 28)
	CM	calbistrin A[a] (26), andrastin A[a] (27), andrastin C (28)	5	24 (26, 27, 28)	24 (26, 27, 28)
P. flavigenum	DM	chrysogine[a] (10), meleagrin[a] (2)	0	1 (10)	2 (10, 2)
	CM	chrysogine[a] (10), meleagrin[a] (2), N6-formylroquefortine (3), roquefortine C[a] (5), fungisporin (6)	0	1 (2)	5 (10, 2, 3, 5, 6)
P. steckii	DM	trichodermamide C[a] (30)	3	21	25 (30)
	CM	trichodermamide C[a] (30)	13	29	34 (30)
P. antarcticum	DM	chrysogine[a] (10), 2-Pyruvoylamino+benzamide[a] (31), 2-Acetyl-4(3H)-quinazolone[a] (32)		7 (10, 31, 32)	8 (10, 31, 32)
	CM	chrysogine[a] (10), 2-Pyruvoylamino+benzamide[a] (31), 2-Acetyl-4(3H)-quinazolone[a] (32)		6 (10, 31, 32)	9 (10, 31, 32)
P. arizonense	DM	pseurotin A (33)		4	8 (33)
	CM	pseurotin A (33)		6 (33)	8 (33)
P. solitum	DM	atlantinone A (34)		5 (34)	5 (34)
	CM	(no known compounds)		4	4

Column 3: metabolites (and their id numbers) identified with dereplication, in order of retention time. Column 4, 5 and 6: per sample the total amount of molecular features (identified and unknown metabolites), in between brackets the id number of the identified compounds in each sample. A heat map from yellow to dark orange follows the total amount of molecular features in the samples

[a] Confirmed with a reference standard

identified; in *P. flavigenum* only meleagrin and precursors were identified while both in *P. coprophilum* and *P. vulpinum* the end product of the pathway, oxaline was identified. Three species–metabolite relationships have, to our knowledge, not been described in literature before: *P. vulpinum* and *P. decumbens* producing andrastin A and *P. flavigenum* producing chrysogine.

To identify in which stages of the fermentation the various metabolites were produced, we examined the individual metabolite production profiles covering the three analyzed stages of the fermentation, one during the growth phase, one a few hours after the CO_2 values started decreasing and one in the stationary phase (depicted as triangles in Fig. 1 and exact values in Additional file 2). This analysis revealed production of specific metabolites was associated with specific growth phases of the cultivations. For example for *P. coprophilum* in DM, the metabolites from the roquefortine/oxaline pathway were already present in the first sample, while griseofulvin and dechlorogriseofulvin could only be detected in the last sample. Another example of this stage depending production was *P. steckii* grown in CM where the vast majority of the metabolites were identified at the second sample point, while trichodermamide C was only first observed at the third sample point. Many of the metabolites in *P. steckii* were identified as belonging to the family of hynapenes and arohynapenes (based on UV and accurate mass) but because of a lack of reference standards could not be assigned to specific peaks and are therefore not shown in Table 2. In some cases, metabolites were produced only in CM or DM; penicillic acid was only produced in *P. polonicum* in CM, calbistrin A was only detected in *P. decumbens* in CM, and atlantinone A was only detected in *P. solitum* in DM. This was also seen for unidentified metabolites (Additional file 2). In 17 out of the 20 cultivations the metabolite levels in the third sample were higher than in the second sample, showing that part of the metabolites started to be produced only after the growth phase.

Discussion

The fungal kingdom contains a huge reservoir of bioactive secondary metabolites; however, the development from discovery of a relevant metabolite to application remains a challenge. The recent genome sequencing of ten secondary metabolite rich *Penicillium* species facilitates exploitation of their genome sequences and/or their use as native producing cell factories. The aim of this study was to determine the suitability of these ten Penicillia as novel cell factories for native compounds by testing their growth performance and secondary metabolite production in submerged cultivations.

Cultivation of the fungal species in controlled submerged bioreactors proved that the ten wild type *Penicillium* species all could be cultivated using the same conditions. The cultivations proved to be highly reproducible, both in defined medium (DM) and rich medium (CM). In 18 out of the 20 cultivations, replicate values for CO_2 off gas and biomass dry weight concentrations were reproducible with maximum CO_2 production rates having standard deviations between 1 and 27%. The cultivations performed in CM showed dispersed growth until the maximum CO_2 off gas was reached, while five out of the ten species cultivated in DM showed clump/pelleted growth. The fungal morphology can have an enormous impact on the production of enzymes and primary or secondary metabolites [18]. For example, micro-colonies are required for the production of citric acid by *A. niger* [19]. However, there is not a defined formula on how morphology affects productivity and it should be tested and optimized for the relevant species and metabolites. The work conducted in this study can be used as a starting point in optimization of process parameters for the production of a specific secondary metabolite.

Maximum specific growth rates in DM were calculated to be between 0.14 and 0.22 h^{-1} depending on the species, and in CM between 0.17 and 0.29 h^{-1}. In case of pellet formation the cultures grew only exponentially in the beginning of the cultivation prior to pellet formation. The obtained growth rates are reasonable starting points for industrial fermentations, for example the penicillin producing *P. chrysogenum*, has a maximum growth rate under comparable conditions of around 0.2 h^{-1} [20, 21]. An interesting observation is that the species with the highest growth rates, *P. decumbens* and *P. coprophilum*, have the smallest genome size [15]. The biomass yields during the growth phase on glucose varied highly between species and there was a correlation between yields and morphology: species with low yields all showed pellet formation and/or wall growth, while species with high yields showed dispersed growth. In case of wall growth, this can be explained by the biomass determination method used in this study: broth was sampled from a port in the reactor, which in the case of wall growth resulted in a lower biomass value than the real biomass value in the reactor, and a lower yield. One hypothesis for the low yields observed in pelleted growth cultures is that dense hyphal packing may often result in diffusional limitation of both nutrients and oxygen, which can cause autolysis of the pellet core [22]. This can lead to a decrease of dry cell mass, as was previously observed in cultures of *P. chrysogenum* where hyphal elements showed cell wall degradation and clear signs of autolysis, even during the culture growth phase [23].

Similar to the growth profiles, the UHPLC–QTOFMS metabolite profiles showed to be reproducible in the biological triplicates of the batch fermentations. The base peak chromatograms revealed that all ten species produced secondary metabolites in submerged cultivations, with most of the metabolites being unique to one of the ten species. The media composition had an influence on secondary metabolite profiles, with eight out of ten species producing a higher number of compounds in CM than in DM. This is not surprising as the yeast extract containing CYA medium is regarded as a secondary metabolite inducing medium [24]. Some species such as *P. solitum* and *P. flavigenum* produced very few metabolites in both media; these species could be interesting for the development of secondary metabolite free production hosts, especially *P. flavigenum* since it showed to have a high reproducibility and a high growth rate. It has been shown in both *P. chrysogenum* [25] and in *Streptomyces* species [26, 27] that removal of highly producing secondary metabolite pathways can increase the production of other native, or heterologous secondary metabolites. However, the species producing very few secondary metabolites still contain many secondary metabolite biosynthetic gene clusters (BGCs). In a recent large scale genomic analysis of *Penicillium* genomes, it was found that 24 *Penicillium* genomes (including the ten used in this study) contained in total 1317 putative BGCs, corresponding to an average of 55 clusters per species [15]. This means that a great number of BGCs were inactive in the conditions used in this study, especially when taken into account that one BGC is often responsible for the biosynthesis of a whole family of related metabolites.

Dereplication of the extracts showed that the species produced a broad spectrum of secondary metabolites belonging to several different metabolite classes (polyketides, non-ribosomal peptides and terpenes). In total we identified 34 different known compounds, of which many for the first time in submerged fermentations. Here, it has to be taken into account that because of the chemical diversity of the secondary metabolites the extraction could not be optimal for all compounds, and it was only possible to obtain semi-quantitative data. Several of the identified compounds are known to have antibacterial, antifungal and anti-cancer activities, such as the antifungal griseofulvin, the antibiotic roquefortines [28] as well as the promising anti-cancer metabolites andrastin A [29, 30] and calbistrin A [31]. At the same time also several mycotoxins such as patulin and penicillic acid were produced of which elimination need to be considered in a production process. Additionally, we described three new species–metabolite relationships: to our knowledge, this is the first time that *P. decumbens* and *P. vulpinum* are described as andrastin A producers and *P. flavigenum*

as chrysogine producer. Furthermore, it was shown that the species produce many unknown compounds at levels above the used detection area threshold; these could be interesting targets for isolation, bioactivity testing and structural elucidation, especially because it is less laborious to isolate metabolites from submerged cultures than from agar plates.

In order to utilize metabolic reprogramming techniques to optimize production of a secondary metabolite, the corresponding BGC of the metabolite must first be identified. The BGCs responsible for production of the detected andrastin A, patulin, griseofulvin, pseurotin A, fungisporin and roquefortine/meleagrin pathways have all previously been identified and characterized [32–42]. Although the experimental characterization was done in different species, represents for each of the listed BGCs, except for andrastin A, were recently identified in at least one of the ten analyzed species used in this study [15]. Several other metabolites with predicted BGCs and previously detected in solid cultivations of the corresponding strains, could not be detected in the submerged cultivations of the same strain. This was the case for yanuthones previously detected in extracts of *P. flavigenum* grown on agar plates [15], and fumagillin, pyripyropenes, austalides, and tryptoquivalines, previously detected in extracts of *P. arizonense* grown on agar plates [14]. In this context it is interesting that a species like *P. arizonense* only secreted pseurotin A, despite the large potential for producing a vast number of secondary metabolites. This would be an advantage if pseurotin A has to be produced in an industrial fermentation, but a disadvantage if any of its other secondary metabolites should be produced [14]. Furthermore we detected small amounts of penicillin G in the CM extracts of the known penicillin G producer *P. nalgiovense*, but the intensity and area were lower than the threshold values used in this study. It should be considered that in the experiments conducted in this study the biomass was filtered out and only the secreted metabolites were studied, while in plate experiments the intracellular metabolites are also taken into account.

Conclusions

The results of this study show that the fermentation properties of the ten analysed *Penicillium* species and the highly reproducible performance in bioreactors should be considered as very encouraging for the application of native hosts for production via submerged fermentation. The *Penicillium* species (isolates) showed promising growth characteristics for use in large-scale industrial bioprocesses and produced a diverse range of interesting secondary metabolites. The production of specific secondary metabolites can subsequently be optimized by

process optimization, classical strain development and/or with metabolic engineering approaches. Furthermore, all ten species used in this study were recently genome sequenced, which makes the use of rational approaches for improving the product yields or eliminating the production of other compounds more accessible thanks to the ongoing development of CRISPR/Cas9 genome editing techniques, which has recently been successfully demonstrated in filamentous fungi [43, 44]. The work conducted in this study can assist in deciding the optimization strategy for specific secondary metabolites produced by one of the ten *Penicillium* species.

Additional files

Additional file 1. Physiological data. Physiological characteristics for triplicate fermentations of each species in DM and CM. Figures show the CO_2 exhaust values and log value of the accumulated CO_2. Furthermore a red line through the data points in exponential phase and highlighted in light grey the non-exponential phase; the time between start of non-exponential growth until the maximum CO_2 off gas value. Additionally the R-squared value for the exponential phase, CO_2 production rate, the number of data points in exponential phase, the end point (in hours) of the exponential phase and the end point of the non-exponential phase (in hours) are shown.

Additional file 2. Secondary metabolite data. In each separate tab one species with all triplicate samples taken in the DM and CM media. The time points of the samples taken (in hour after inoculation) and the metabolites detected with their maximum *m/z* value, retention time and intensity. The metabolites were first identified if they had an absolute area higher than 500,000 counts and intensity higher than 10,000 in the time last sample. The presence of these compounds was searched for in all the other samples. Data was generated by UHPLC–DAD–QTOFMS on a maXis HD orthogonal acceleration quadrupole time-of-flight mass spectrometer (Bruker Daltonics).

Authors' contributions
MW, KFN, and JN conceived the study. SG and JCN performed the cultivations and data analysis on physiological parameters. SG performed the chemical extractions and analysis of metabolites. SG, KFN, JCF and TOL evaluated the data on secondary metabolites. SG and JCN wrote the manuscript. MW and RF supervised the work and co-wrote the manuscript. All authors read and approved the final manuscript.

Author details
[1] Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark. [2] Department of Biology and Biological Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden. [3] Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

Competing interests
The authors declare that they have no competing interests.

References
1. Visagie CM, Houbraken J, Frisvad JC, Hong S-B, Klaassen CHW, Perrone G, et al. Identification and nomenclature of the genus *Penicillium*. Stud Mycol. 2014;78:343–71.
2. Frisvad JC, Smedsgaard J, Larsen TO, Samson RA. Mycotoxins, drugs and other extrolites produced by species in Penicillium subgenus Penicillium. Stud Mycol. 2004;49:201–41.
3. Fleming A. On the antibacterial action of cultures of a *penicillium*, with special reference to their use in the isolation of B. influenzae. Bull World Health Organ. 1929;79:780–90.
4. Oxford AE, Raistrick H, Simonart P. Studies in the biochemistry of micro-organisms: griseofulvin, C(17)H(17)O(6)Cl, a metabolic product of *Penicillium griseofulvum* Dierckx. Biochem J. 1939;33:240–8.
5. Gosio B. Sperimentate su culture pure di bacilli del carbonchio demonst-rarano notevole potere antisettica. GR Accad Med Torino. 1893;61:484.
6. Brown BAG, Srnale TC, Pharmaceuticals B, Park B, King TJ, Hasenkamp R, et al. Crystal and molecular structure of compactin, a new antifungal metabolite from *Penicillium brevicompactum*. J Chem Soc Perkin. 1976;1:1165–8.
7. Frisvad JC, Filtenborg O. Terverticillate penicillia: chemotaxonomy and mycotoxin production. Mycologia. 1989;81:837–61.
8. Florey HW, Jennings MA, Gilliver K, Sanders AG. Mycophenolic acid—an antibiotic from *Penicillium brevicompactum* Dierckx. Lancet. 1946;247:46–9.
9. United Nations [Internet]. [Cited 2017 July 9]. http://www.un.org/pga/71/2016/09/21/press-release-hl-meeting-on-antimicrobial-resistance/.
10. Aminov RI. A brief history of the antibiotic era: lessons learned and challenges for the future. Front Microbiol. 2010;1:1–7.
11. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, et al. A new antibiotic kills pathogens without detectable resistance. Nature. 2015;517:455–9.
12. O'Connor SE. Engineering of secondary metabolism. Annu Rev Genet. 2015;49:1–24
13. Gibbs PA, Seviour RJ, Schmid F. Growth of filamentous fungi in submerged culture: problems and possible solutions. Crit Rev Biotechnol. 2000;20:17–48.
14. Grijseels S, Nielsen JC, Randelovic M, Nielsen J, Nielsen KF, Workman M, et al. *Penicillium arizonense*, a new, genome sequenced fungal species, reveals a high chemical diversity in secreted metabolites. Sci Rep. 2016;6:35112.
15. Nielsen JC, Grijseels S, Prigent S, Ji B, Dainat J, Nielsen KF, et al. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. Nat Microbiol. 2017;2:17044.
16. Kildgaard S, Mansson M, Dosen I, Klitgaard A, Frisvad JC, Larsen TO, et al. Accurate dereplication of bioactive secondary metabolites from marine-derived fungi by UHPLC–DAD–QTOFMS and a MS/HRMS library. Mar Drugs. 2014;12:3681–705.
17. Boyle DT. A rapid method for measuring specific growth rate of microorganisms. Biotechnol Bioeng. 1977;19:297–300.
18. Krijgsheld P, Bleichrodt R, van Veluw GJ, Wang F, Müller WH, Dijksterhuis J, et al. Development in *Aspergillus*. Stud Mycol. 2013;74:1–29.
19. Gomez R, Schnabel I, Garrido J. Pellet growth and citric acid yield of *Aspergillus niger* 110. Enzyme Microb Technol. 1988;10:188–91.
20. Nielsen J, Johansen CL, Jacobsen M, Krabben P, Villadsen J. Pellet formation and fragmentation in submerged cultures of *Penicillium chrysogenum* and its relation to penicillin production. Biotechnol Prog. 1995;11:93–8.
21. Robin J, Jakobsen M, Beyer M, Noorman H, Nielsen J. Physiological characterisation of *Penicillium chrysogenum* strains expressing the expandase gene from *Streptomyces clavuligerus* during batch cultivations. Growth and adipoyl-7-aminodeacetoxycephalosporanic acid production. Appl Microbiol Biotechnol. 2001;57:357–62.
22. Prosser JI, Tough AJ. Growth mechanisms and growth kinetics of filamentous microorganisms. Crit Rev Biotechnol. 1991;10:253–74.

23. McNeil B, Berry DR, Harvey LM, Grant A, White S. Measurement of autolysis in submerged batch cultures of *Penicillium chrysogenum*. Biotechnol Bioeng. 1998;57:297–305.

24. Frisvad JC. Media and growth conditions for induction of secondary metabolite production. Fungal Second Metab Methods Protoc Methods Mol Biol. 2012;944:47–58.

25. Salo OV, Ries M, Medema MH, Lankhorst PP, Vreeken RJ, Bovenberg RAL, et al. Genomic mutational analysis of the impact of the classical strain improvement program on β-lactam producing *Penicillium chrysogenum*. BMC Genom. 2015;16:937.

26. Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. Proc Natl Acad Sci. 2010;107:2646–51.

27. Gomez-Escribano JP, Bibb MJ. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. Microb Biotechnol. 2011;4:207–15.

28. Kopp-Holtwiesche B, Rehm HJ. Antimicrobial action of roquefortine. Eur J Appl Microbiol Biotechnol. 1979;6:397–401.

29. Rho MC, Toyoshima M, Hayashi M, Uchida R, Shiomi K, Komiyama K, et al. Enhancement of drug accumulation by andrastin A produced by *Penicillium* sp. FO-3929 in vincristine-resistant KB cells. J Antibiot. 1998;51:68–72.

30. Shiomi K, Uchida R, Inokoshi J, Tanaka H, Iwai Y, Omura S. Andrastins A-C, new protein farnesyltransferase inhibitors, produced by *Penicillium* sp. FO-3929. Tetrahedron Lett. 1996;37:1265–8.

31. Bladt TT, Durr C, Knudsen PB, Kildgaard S, Frisvad JC, Gotfredsen CH, et al. Bio-activity and dereplication-based discovery of ophiobolins and other fungal secondary metabolites targeting leukemia cells. Molecules. 2013;18:14629–50.

32. Matsuda Y, Awakawa T, Abe I. Reconstituted biosynthesis of fungal meroterpenoid andrastin A. Tetrahedron. 2013;69:8199–204.

33. Rojas-Aedo JF, Gil-Durán C, Del-Cid A, Valdés N, Pamela Á, Vaca I, et al. The biosynthetic gene cluster for andrastin A in *Penicillium roqueforti*. Front Microbiol. 2017;8:1–11.

34. Tannous J, El Khoury R, Snini SP, Lippi Y, El Khoury A, Atoui A, et al. Sequencing, physical organization and kinetic expression of the patulin biosynthetic gene cluster from *Penicillium expansum*. Int J Food Microbiol. 2014;189:51–60.

35. Chooi YH, Cacho R, Tang Y. Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. Chem Biol. 2010;17:483–94.

36. Maiya S, Grundmann A, Li X, Li SM, Turner G. Identification of a hybrid PKS/NRPS required for pseurotin A biosynthesis in the human pathogen *Aspergillus fumigatus*. ChemBioChem. 2007;8:1736–43.

37. Wiemann P, Guo C-J, Palmer JM, Sekonyela R, Wang CCC, Keller NP. Prototype of an intertwined secondary-metabolite supercluster. Proc Natl Acad Sci USA. 2013;110:17065–70.

38. Tsunematsu Y, Fukutomi M, Saruwatari T, Noguchi H, Hotta K, Tang Y, et al. Elucidation of pseurotin biosynthetic pathway points to trans-acting C-methyltransferase: generation of chemical diversity. Angew Chem Int Ed. 2014;53:8475–9.

39. Ali H, Ries MI, Lankhorst PP, Van Der Hoeven RAM, Schouten OL, Noga M, et al. A non-canonical NRPS is involved in the synthesis of fungisporin and related hydrophobic cyclic tetrapeptides in *Penicillium chrysogenum*. PLoS ONE. 2014;9:e98212.

40. Ali H, Ries MI, Nijland JG, Lankhorst PP, Hankemeier T, Bovenberg RAL, et al. A branched biosynthetic pathway is involved in production of roquefortine and related compounds in *Penicillium chrysogenum*. PLoS ONE. 2013;8:1–12.

41. Ries MI, Ali H, Lankhorst PP, Hankemeier T, Bovenberg RAL, Driessen AJM, et al. Novel key metabolites reveal further branching of the roquefortine/meleagrin biosynthetic pathway. J Biol Chem. 2013;288:37289–95.

42. García-Estrada C, Ullán RV, Albillos SM, Fernández-Bodega MÁ, Durek P, Von Döhren H, et al. A single cluster of coregulated genes encodes the biosynthesis of the mycotoxins roquefortine C and meleagrin in *Penicillium chrysogenum*. Chem Biol. 2011;18:1499–512.

43. Pohl C, Kiel JAKW, Driessen AJM, Bovenberg RAL, Nygård Y. CRISPR/Cas9 based genome editing of *Penicillium chrysogenum*. ACS Synth Biol. 2016;5:754–64.

44. Nødvig CS, Nielsen JB, Kogle ME, Mortensen UH. A CRISPR-Cas9 system for genetic engineering of filamentous fungi. PLoS ONE. 2015;10:e0133085.

Filamentous ascomycetes fungi as a source of natural pigments

Rebecca Gmoser[1,2*], Jorge A. Ferreira[1], Patrik R. Lennartsson[1] and Mohammad J. Taherzadeh[1]

Abstract

Filamentous fungi, including the ascomycetes *Monascus*, *Fusarium, Penicillium* and *Neurospora*, are being explored as novel sources of natural pigments with biological functionality for food, feed and cosmetic applications. Such edible fungi can be used in biorefineries for the production of ethanol, animal feed and pigments from waste sources. The present review gathers insights on fungal pigment production covering biosynthetic pathways and stimulatory factors (oxidative stress, light, pH, nitrogen and carbon sources, temperature, co-factors, surfactants, oxygen, tricarboxylic acid intermediates and morphology) in addition to pigment extraction, analysis and identification methods. Pigmentation is commonly regarded as the output of secondary protective mechanisms against oxidative stress and light. Although several studies have examined pigmentation in *Monascus* spp., research gaps exist in the investigation of interactions among factors as well as process development on larger scales under submerged and solid-state fermentation. Currently, research on pigmentation in *Neurospora* spp. is at its infancy, but the increasing interest for biorefineries shows potential for booming research in this area.

Keywords: Pigments, *Neurospora*, Carotenoids, Edible filamentous fungi, Ascomycetes

Background

For a long time, filamentous fungi have been used for the industrial production of commercially relevant products, including enzymes, antibiotics, feed products, and many others [1]. The biorefinery concept, i.e., converting substrates to value-added products, is widely accepted within the research community. Therefore, research towards the diversification of established and future facilities for the production of numerous novel and valuable products as well as by-products through fermentation is presently a hot topic. First-generation ethanol plants are good examples where side stream products are utilized to supplement already existing products (e.g., ethanol, animal feed and CO_2) by producing substances such as organic acids, enzymes, ethanol, biomass for food and/or feed applications, and pigments [1]. In particular, the interest for fermentation-derived pigments in the food and feed industry has increased in recent years [2]. This interest in food-grade pigments is because of the pigments' ability to

enhance the products' natural color in order to indicate freshness, appearance, safety, and sometimes even to add a novel sensory aspect to attract consumers [3, 4].

Producing pigments from filamentous fungi has great potential [2, 4, 5], not only as an added-value product for biorefineries but also as an alternative to synthetic or other natural pigments that have limitations [6–8]. The increasing demand for pigments of natural origin, particularly in the food sector [9, 10], further increases the interest to investigate filamentous fungi as potential pigment producers. To improve the chances of the pigments and biomass being free from mycotoxins (toxins of fungal origin [11]), particular interest has been taken in edible filamentous fungi that have been used in traditional food products and that can naturally synthesize and secrete pigments. Figure 1 presents an overview of the sources of natural pigments, highlighting the main focus of this review.

A few strains of ascomycetes filamentous fungi being considered as potential pigment producers include, some strains of *Talaromyces* (e.g., *T. purpurogenus* and *T. atroroseus* producing red pigments), *Cordyceps unilateralis* (deep blood red pigment) [15], *Herpotrichia rhodosticta*

*Correspondence: rebecca.gmoser@hb.se
[2] University of Borås, Allégatan 1, 503 32 Borås, Sweden
Full list of author information is available at the end of the article
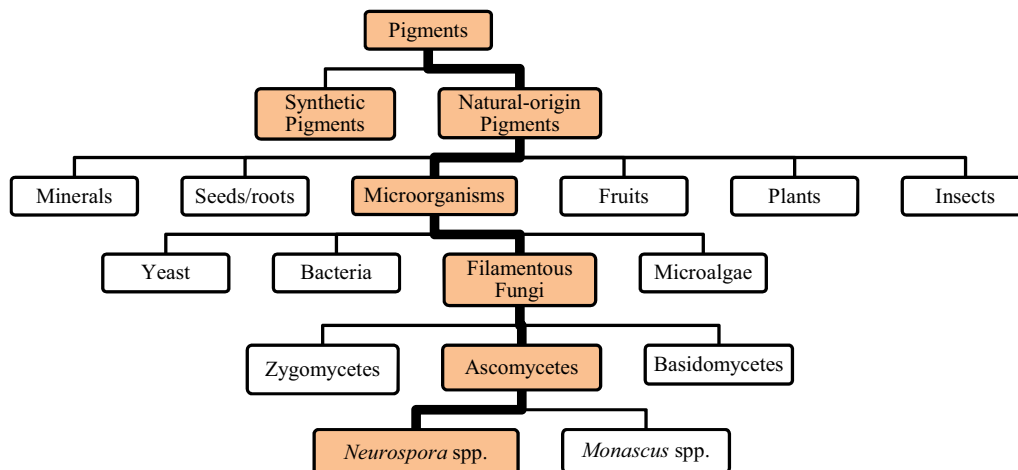
Fig. 1 Overview of some sources that can be used for extraction of synthetic or natural-origin pigments [12–14]

(orange), *Curvularia lunata* and several species of *Drechslera* (many different pigments). Strains of these species are promising because they are non-mycotoxigenic and non-pathogenic to humans. Nevertheless, the individual mycotoxin profiles of these strains remain to be explored [12]. Some other pigment producing fungi for their use in the production of potential food colorants are species of *Eurotium* and *Fusarium oxysporum* (yellow and red pigments, respectively) [13], *Fusarium fujikuroi* (red [16] and orange pigments [17]) and strains of *Penicillium* [13] such as *P. citrinum*, *P. islandicum* [5], *P. aculeatum* and *P. pinophilum* [5]. However, several species of *Penicillium* are able to produce known toxic metabolites [5] and *Eurotium* spp. and *F. oxysporum* have been shown to produce mycotoxins as well. The potential production of mycotoxins is a major problem which limits the commercial application of these strains of fungi [4]. This problem, together with the increasing demand for natural coloring alternatives from both customers and regulators [4], has triggered investigations and screens for other potential pigment-producing genera of fungi.

After more than two centuries of research, *Neurospora* spp. have been generally recognized as safe (GRAS) [18] with no record of mycotoxin production [1], and thus considered safe for animal and human consumption [19]. This fungus is able to grow rapidly on various types of substrates [20], such as industrial residuals and lignocellulose, to produce ethanol, biomass, and pigments. Particularly, the ascomycete *N. intermedia*, used for preparation of the Indonesian food oncom, has recently been reported as a potential biomass and ethanol producer from waste streams of the industrial process of ethanol production from agricultural grains [1, 21]. However, despite ongoing research and the availability of microorganism-derived

pigments, little is known about the production of pigments by *N. intermedia* [6, 9]. Since *Neurospora* species are able to accumulate orange pigments, other species of *Neurospora* such as *Neurospora crassa*, which is genetically and biochemically one of the most well-studied eukaryotic microorganisms [22], have been investigated more extensively regarding the biosynthesis of carotenoids and its regulation [17, 23–25]. *N. crassa* accumulates a mixture of carotenoid pigments such as γ-carotene and neurosporaxanthin [6, 7], the latter carotenoid acid has also been isolated from *Neurospora sitophila* [8].

Several factors have been reported to influence pigment production in ascomycetes [26], although studies on most, if not all, of the mentioned factors are still scarce and superficial. Moreover, even though a great deal of research is available on pigment production by *Monascus* spp. and *Fusarium* spp., information on the performance of the process using both submerged and solid-state fermentation on a larger scale is missing. A deeper understanding and better overview of the factors influencing pigment production, particularly in *Neurospora* as well as in *Monascus* and *Fusarium* spp., is thus desired in order to optimize the process.

The present review gathers available research on pigment production by *Neurospora* species. It identifies main research gaps and consequently provides future research avenues and main challenges towards the use of *Neurospora* spp. for the production of pigments.

Filamentous ascomycetes fungi as pigment producers

Filamentous ascomycetes fungi are known to produce an extraordinary range of colors. There are a wide selection of non-pathogenic strains of filamentous fungi that are

non-toxin producers and can be used as potential sources of natural food colorants with improved functionality [7]. The ability of these fungi to grow on residuals of different complexity (e.g., starch-based, lignocellulose-based residuals) is well-documented, showing versatility regarding different processes that can be built around the ascomycetes fungi [1]. Unlike the use of pigments from vegetables and fruits, the cultivation of ascomycetes does not compete with agricultural land for food production, and therefore, the synthesis of pigments is faster due to time-efficient and simple fermentation processes [7, 27]. The fermentation processes generate high yields of biomass together with value-added products such as pigments, organic acids and alcohols [1].

Regarding pigment production from ascomycetes, even though there is much research on the factors that influence pigmentation in *Monascus* spp. [2, 13, 26, 28–31] and to some extend for *Fusarium* spp. [32–35], the correlations between the factors are still not fully understood [26]. It is likely that a variety of factors with a complex interplay are involved and that they vary among species. Research using *N. intermedia* for pigment production is limited to only a few studies [6, 9, 36]. Moreover, similar to the research with *Monascus*, studies on large-scale production are nonexistent in the literature [26]. Thus, the specific culture conditions that induce pigment production and their properties and the link between different pigments and the level and activity of carotenoid biosynthetic enzymes are not well understood. The physicochemical properties of pigments are further discussed in the review by Priatni [37]. The available information on pigment-producing *Neurospora* spp. has been compiled in Table 1, including fermentation mode and extraction solvent for the produced pigments, along with the pigment concentration.

The diverse classes of pigments synthesized and secreted by ascomycetes are commonly reported as secondary metabolites [22] of both known and unknown functions [5]. Pigments are generally produced in the cell cytoplasm as a response to disadvantageous environmental conditions, such as nutrient limitation [7, 41], and this process is controlled by a complex regulatory network. Different pigments help to improve fungal survival; for example, carotenoids protect against harmful ultraviolet radiation and light (lethal photooxidation), melanins protect against environmental stress, and flavins serve as cofactors in enzyme catalysis [12]. Pigments produced by filamentous fungi include melanins (dihydroxynaphthalene melanin; a complex aggregate of polyketides), phenazines, flavins (riboflavin), quinones (anthraquinones, naphthaquinones and azaphilones) and carotenoids [2, 12]. Generally, these pigments are chemically classified as either carotenoids or polyketides [5] based on different

biosynthetic pathways. *Monascus* spp. pigments are generally produced through the polyketide pathway with some related routes from other pathways such as fatty acid biosynthesis [5, 42], whereas *N. intermedia* pigments are produced through the carotenoid biosynthetic pathway [6]. *Monascus* spp. are usually used as model fungi for polyketide biosynthesis. However, due to the complex pathways involved, there is only rudimentary knowledge about the polyketide pigment regulation. *Fusarium* spp. produces both polyketides (e.g., the polyketide-derived pigment bikaverins) and carotenoids (e.g., neurosporaxanthin) [32]. Polyketides do not seem to be produced in *Neurospora* spp. [43]. Figure 2 illustrates the different pathways involved in polyketide and carotenoid fungal pigment production in broad terms. Carotenoids are yellow to orange-red pigments that are widely used as food colorants. They are produced mainly by microbes belonging to *Myxococcus*, *Streptomyces*, *Mycobacterium*, *Agrobacterium* and *Sulfolobus* [3].

Carotenoids belong to the subfamily of isoprenoids, which include a large and diverse class of naturally occurring chemicals [46]. As members of isoprenoids [47], these molecules are derived from eight C_5 isoprene units and contain 40 carbon units. Based on their molecular structures, carotenoids are therefore further split into two divisions: carotenes and xanthophylls. Carotenes are hydrocarbons, while xanthophylls are derivatives containing oxygen in various functional groups in otherwise similar structures to that of carotenes [24, 48]. Examples of some common carotenoid structures are presented in Fig. 3 [14, 25, 37].

Most of the *Neurospora* spp. have been identified in tropical and subtropical areas in the world and, to some extent, in temperate areas of western North America and Europe as well [49, 50]. Five species of *Neurospora* were identified in Europe, namely, *N. crassa, N. discreta, N. intermedia, N. sitophila* and *N. tetrasperma*. These strains are similar in morphology, and the color of their conidia are orange or yellow-orange caused by the different types of carotenoids. *N. crassa* is the most well-known. Most attempts to increase its carotenoid production has been carried out through photoinduction [51, 52]. Since the carotenoid biosynthetic pathway of *N. intermedia* is comparable to that of *N. crassa* [53] and *F. fujikuroi* [14, 54], studies using *these fungi* are also of interest.

Carotenoids

Carotenoids are among the most common of all natural pigments on the market [46], and among them, β-carotene, lycopene, astaxanthin, canthaxanthin, lutein, and capxanthin have been exploited commercially (Fig. 3) [8]. Carotenoids are usually extracted from carrots, citrus peels, tomatoes, and algae [6]. Bacteria and

Table 1 Pigment production by *Neurospora* spp. using different experimental set-ups

Ascomycete	Substrate	Mode of operation[b]	Pigment(s)	λ_{max} (nm)[a]	Pigment extraction solvent	Production	References
Neurospora crassa	60% tapioca by product and 40% tofu waste	SSF	β-Carotene (yellow-orange)	429, 452, 478	THF	295 µg/g	[13, 38, 39]
Neurospora crassa 74-OR23-1A	Vogel's minimal broth	SmF, in 500-mL Erlenmeyer flask	Neurosporaxanthin	477	Acetone	–	[23, 40]
Neurospora crassa	Not defined medium with addition of 1% Bacto-Difco agar	Fernbach flasks (200 ml) for 15 days, followed by 1 day in plastic tent (oxygen atmosphere) (SSF)	Neurosporaxanthin	477.5, 504 (in petroleum ether), 471.5, 496 (in acetone)	Acetone, petroleum ether	0.33 mg/g dry weight	[22, 23]
Neurospora sitophila	Not defined medium with addition of 1 g/l Bacto-Difco yeast extract	Aerated flasks for 6 days, wet mycelial was then placed on petri dishes in a transparent plastic tent for 1 day. (SSF)	Neurosporaxanthin	477 (petroleum ether)	Methanol, petroleum ether	0.04 mg/g dry weight	[22, 23]
Neurospora intermedia N-1	Maltose, peptone, yeast extract, Mg^{2+} fermentationmedia	SmF* in 1 l Erlenmeyer flask	Yellow-orange	480	1 g of spores was extracted with 5 ml of acetone	24.31 µg/g spores	[9]
Neurospora intermedia (PTCC 5291)	Vogel's growth medium	SmF* in 250 ml Erlenmeyer flasks	Mixture of carotenoids	470	50 mg dried mycelia were extracted in 3 ml methanol, re-extracted with 3 ml of acetone	500 mg/l	[36]

[a] Total carotenoids content is determined by spectral absorption at specific wavelengths maxima depending on which solvent the carotenoids are extracted in. The absorbance value correlates to the amount of carotenoids present in the samples

[b] Solid-state fermentation (SSF), submerged fermentation (SmF)

a
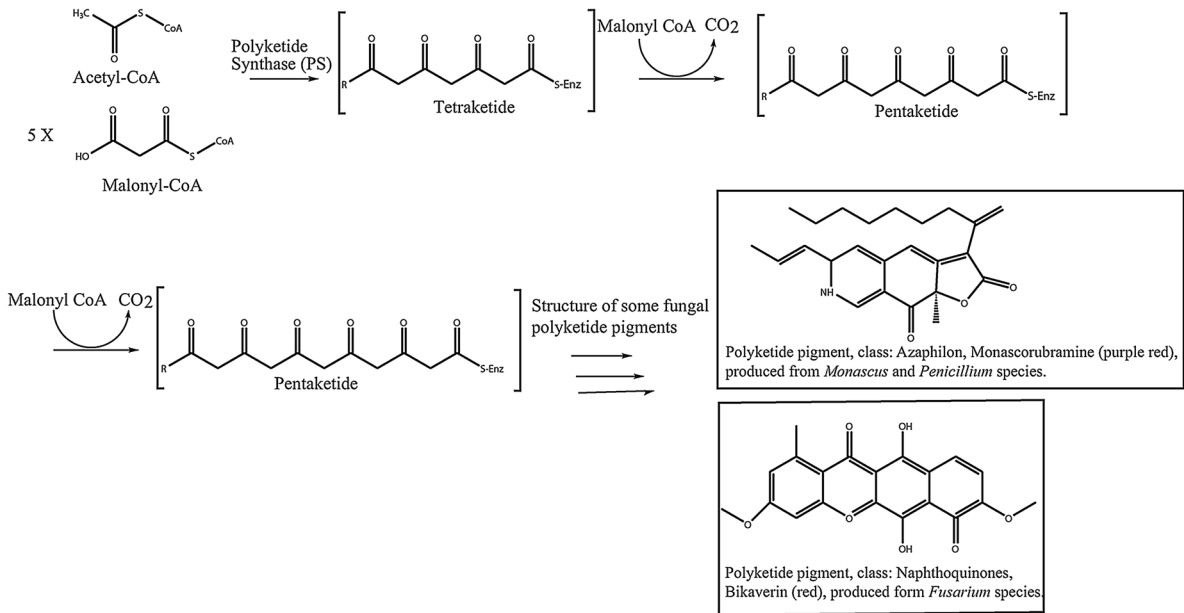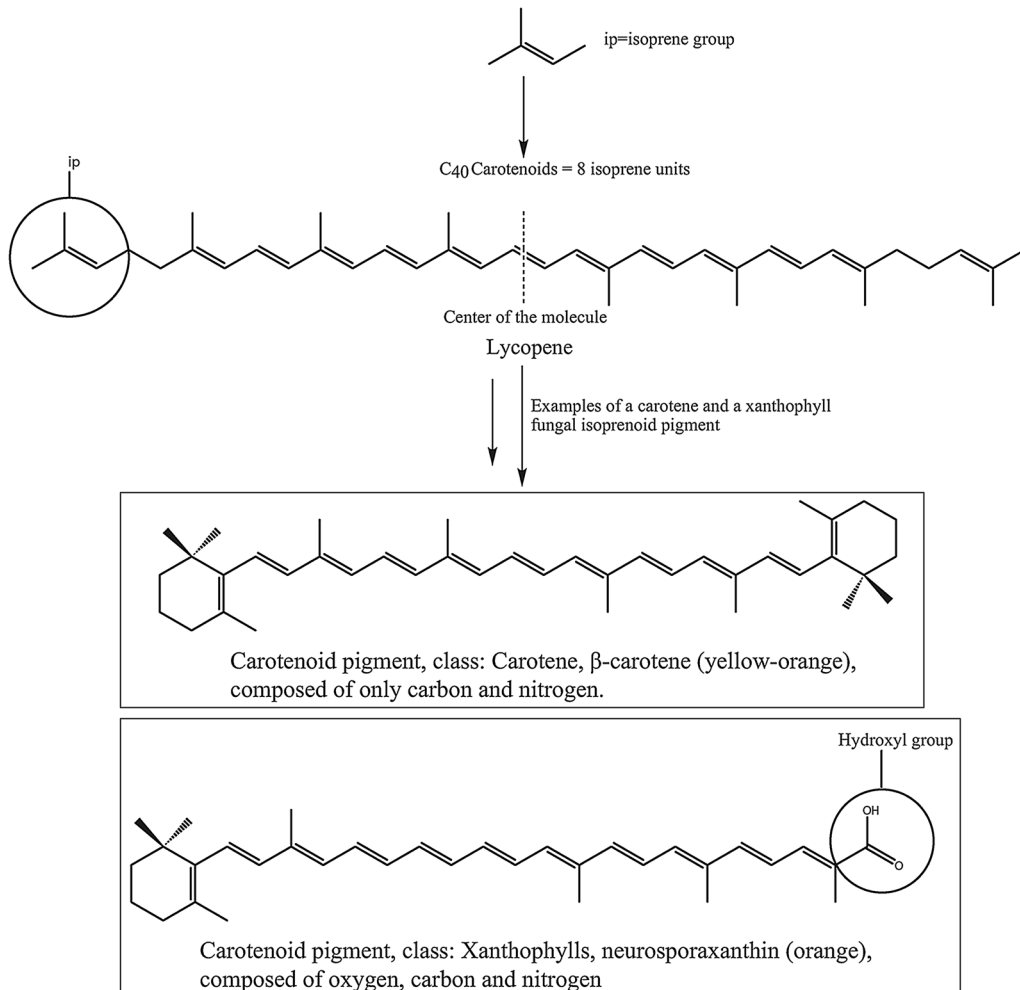
Polyketide pigment, class: Azaphilon, Monascorubramine (purple red), produced from *Monascus* and *Penicillium* species.

Polyketide pigment, class: Naphthoquinones, Bikaverin (red), produced form *Fusarium* species.

b

ip=isoprene group

C_{40} Carotenoids = 8 isoprene units

Center of the molecule

Lycopene

Examples of a carotene and a xanthophyll fungal isoprenoid pigment

Carotenoid pigment, class: Carotene, β-carotene (yellow-orange), composed of only carbon and nitrogen.

Hydroxyl group

Carotenoid pigment, class: Xanthophylls, neurosporaxanthin (orange), composed of oxygen, carbon and nitrogen

Fig. 3 Chemical structures of some common carotenoids found in microorganisms that are also of economic value [8] as well as the structure of the carotenoid acid neurosporaxanthin [22]

carotenogenic species of fungi can also produce them. It is the conjugated double bond system in the carotenoid structure that acts as the chromophore for wavelength-selective (light) absorption, giving these compounds an attractive bright yellow to red color. The absorbed energy leads to electron excitation and changes in orbital occupancy, local bonding, and charge distribution. The non-absorbed light is transmitted and/or reflected to be captured by the eye [44]. Each desaturated reaction shifts the absorption maxima towards longer wavelengths resulting in different yellow to red colors of carotenoids [14]. The double bonds in the carotenoid structure can exist in *trans* and *cis*, but in nature, they are generally in all-*trans* form [55]. The transformation from *trans* to *cis* can be done by acid, heat, oxygen and exposure to light

[56]. Nevertheless, the shift from an all-*trans* arrangement to *cis* only results in a minor loss in color strength and hue [55].

Similar to other metabolites, carotenoids have ecological functions and are of value to the fungi. For example, as mentioned before, they can protect against lethal photooxidation [12]. Sterols, dolichols, and ubiquinones fulfill essential cell functions, while secondary carotenoids, such as astaxanthin and canthaxanthin, are accumulated as a response to environmental stress [46]. For example, they can be integrated into the cell membrane to improve its fluidity under high or low temperatures, high light conditions, or when the lipids become more unsaturated. Furthermore, carotenoids serve as precursors of several physiologically important compounds in fungi, such as apocarotenoids (e.g., the fungal pheromone trisporic acid), which are synthesized through the oxidative cleavage of carotenoids. *Neurospora* is one of the carotenogenic genus of fungi, producing a mixture of carotenoid and apocarotenoid pigments. The major component of the carotenoids produced by *Neurospora* is the C_{35}-apocarotenic acid, neurosporaxanthin (see Fig. 3) [14, 47]. The closely related filamentous fungus *F. fujikuroi* is also known to synthesis neurosporaxanthin, and have contributed extensively to a better understanding of the neurosporaxanthin pathway and its regulation [35, 57].

Carotenoid biosynthesis

There is a limited amount of knowledge on the regulation of carotenoid biosynthesis, accumulation, and storage in filamentous fungi. Studies covering this area have only resulted in hypotheses and further research to explain the regulation on a cellular level is required. It has been proposed that genes encoding enzymes involved in isoprenoid and carotenoid biosynthesis are subjected to positive and negative feedback regulatory processes. Several of these molecules are able to mediate signaling processes in order to attain a balanced supply of precursors and assure the adjustment of biosynthesis in response to developmental and environmental cues [58]. It has been hypothesized that the genes are silent under optimal culture conditions and only activated under certain conditions [13]. Even though the molecular mechanisms are awaiting a more detailed understanding, the carotenoid biosynthesis process is undoubtedly influenced by culture conditions, and can therefore be optimized [13].

It is known that filamentous fungi are able to sense and respond to external signals, such as environmental stress. Heterotrimeric G proteins (G proteins) are key signaling elements that communicate external signals to the cells, triggering the upstream regulation of fungal secondary metabolites (e.g., carotenoids) as a response [59]. The levels and activities of carotenoid biosynthetic enzymes and the total carbon flux through the synthesizing system are also important stimuli [46]. These carotenoid molecules are hydrophobic, and they can generally be found in lipid globules and the endoplasmic reticulum membranes. Accordingly, the enzymes involved in carotenogenesis are membrane-bound. However, the regulatory mechanisms controlling the cell compartment where carotenoids are synthesized or stored are still unknown for the *Neurospora* spp. [14]. Structural properties and biosynthesis aspects of carotenoids focusing on *Neurospora* spp. are available in a review by Priatni [37].

The genes encoding the enzymes involved in carotenoid biosynthesis have been biochemically investigated [14]. *N. crassa* have been used, along with *F. fujikuroi* that has a similar carotenoid pathway as *N. crassa* [60], as a research model to investigate the carotenoid biosynthesis and researcher have made discoveries on the regulation of pigment biosynthesis by using *Fusarium* species [61]. Acetyl-CoA has been suggested to have an impact on the production of secondary metabolites, such as pigments, by being a primary precursor, depending on the available enzyme pool [59]. Figures 4, 5 and 6 show the carotenoid biosynthetic pathway of *Neurospora* spp. initiated via the mevalonate pathway that leads to the synthesis of short, five-carbon isoprenoid precursors. The isoprenoid biosynthesis pathway is common for all carotenoids. In addition, the last steps in Figs. 4, 5 and 6 present the carotenogenic pathway in *N. crassa*. It has been suggested that the biosynthesis pathway in *N. intermedia* is similar to that in *N. crassa* [37, 53, 62]. The carotenoid pathway is initiated by the condensation of two geranylgeranyl pyrophosphate (GGPP) molecules by the bifunctional enzyme with both phytoene synthase and lycopene cyclase activity, *al-2* of *N. crassa*, to produce the colorless carotene–phytoene, the precursor to different carotenoids. The corresponding enzyme in *F. fujikuroi* is named *CarRA*. Phytoene desaturase, encoded by phytoene dehydrogenases *al-1* in *N. crassa* (*CarB* in

Fig. 5 The possible carotenoid biosynthetic pathway of *Neurospora crassa*. The gene products/enzymes responsible for each enzymatic reaction are indicated. Site of chemical changes from precursor molecules are *shaded*. Molecular groups that distinguish xanthophylls from carotenes are marked with *red circles*. The pigments found in *N. intermedia* N-1 are presented in *boxes*

F. fujikuroi) (Figs. 4, 5 and 6), mediates the introduction of up to five conjugated double bonds into a phytoene backbone to produce dehydrogenated colored carotenoids [14]. The carotenoids are then subjected to one or two cyclization reactions by cyclases, which introduce α- or β-ionone rings at one or both ends of the polyene chain [63]. The action of *al-1* and *CarB* results in different colored intermediates, namely,

3,4-didehydrolycopene, ζ-carotene, neurosporene, lycopene and β-carotene. 3,4-Didehydrolycopene is synthesized by lycopene cyclase to yield the reddish carotene torulene. One cleavage reaction and two oxidation steps are then required to produce the final apocarotenoid neurosporaxanthin in *Neurospora* from torulene. First, torulene is converted into β-apo-4-carotenal by the torulene-cleaving oxygenase *Cao-2* in

Continued
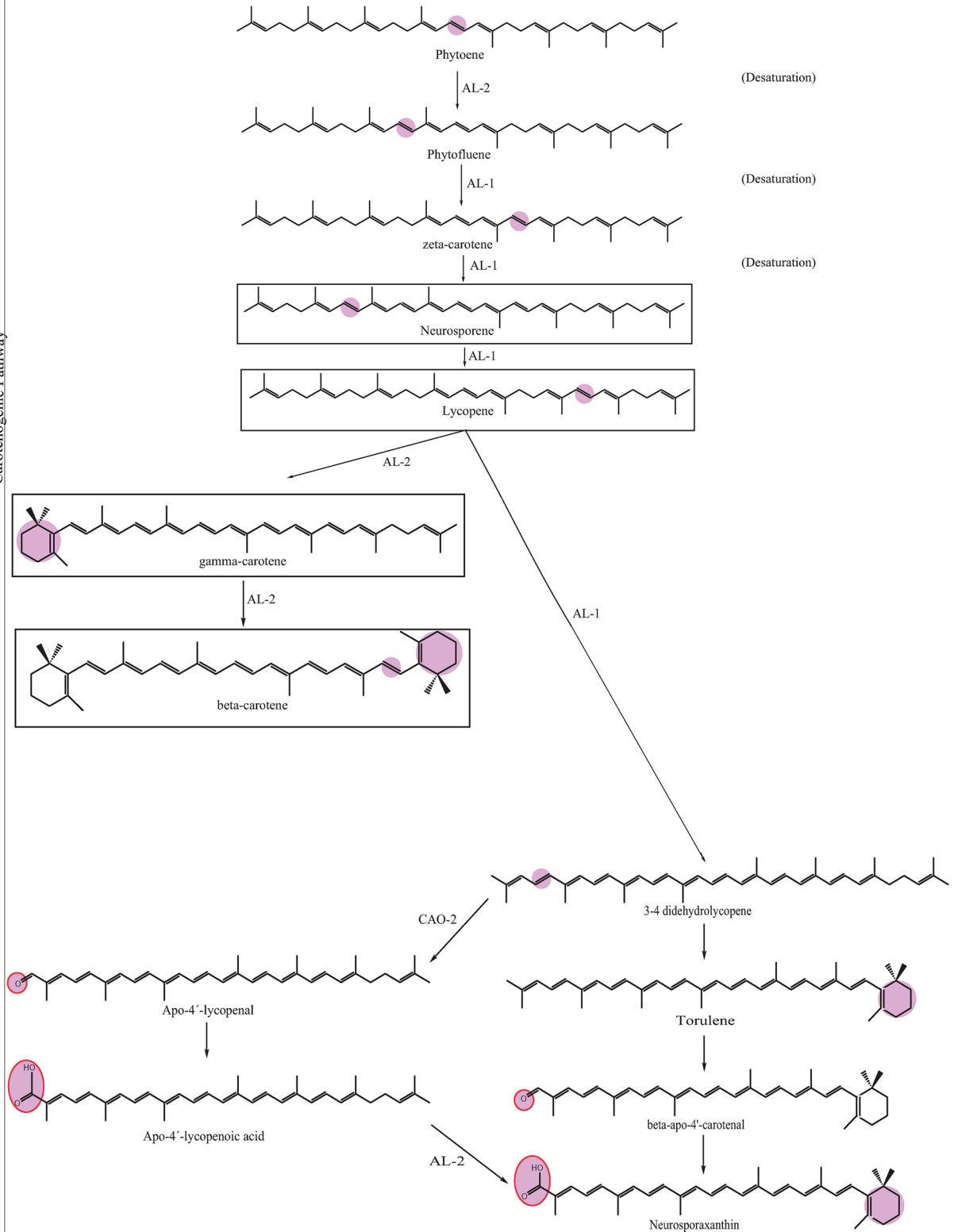
Fig. 6 The possible carotenoid biosynthetic pathway of *Neurospora crassa*. The gene products/enzymes responsible for each enzymatic reaction are indicated. Site of chemical changes from precursor molecules are *shaded*. Molecular groups that distinguish xanthophylls from carotenes are marked with *red circles*. The pigments found in *N. intermedia* N-1 are presented in *boxes*
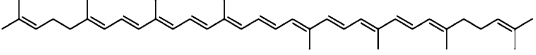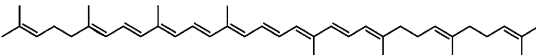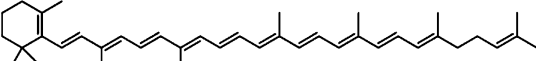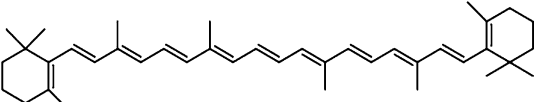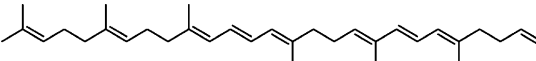
N. crassa or *CarT* in *F. fujikuroi*. It was the identification of *carT* that led to the identification of its *Neurospora* orthologue, *cao-2*. Next, β-apo-4′-carotenal is further oxidized to neurosporaxanthin by the aldehyde dehydrogenase *ylo-1* in *N. crassa* and *CarD* in *F. fujikuroi* [14, 54]. In summary, the five genes needed to produce neurosporaxanthin in *N. crassa* are *al-3, al-2, al-1, cao-2* and *ylo-1.* Thus, the order of reactions in the biosynthesis and, therefore, the type of carotenoids produced depends on the degree of oxidative stress [35] or temperature of growth [14, 37, 47], among other things. Singgih Marlia, et al. [9] evaluated the carotenogenesis of *N. intermedia* N-1 in a liquid fermentation system and were able to identify five carotenoid compounds in the spores, namely lycopene, neurosporene, γ-carotene, β-carotene and phytoene [9]. Table 2 includes the structures and colors of carotenoid compounds found in *N. intermedia* N-1.

Factors that have been previously reported to influence pigment production in other strains can thus be considered as potentially important in other filamentous fungi as well. Based on already existing processes and previous knowledge of carotenoid production pathway in filamentous fungi, these factors are individually discussed in further sections.

Factors influencing pigment production

Pigment biosynthesis is greatly influenced by fermentation conditions, such as medium composition and process parameters. Sexual interactions have also been reported to increase the biosynthesis of the carotenoid β-carotene in some fungal species (e.g., *B. trispora*), where trisporic acids (substances with hormonal activity formed upon mating) were suggested to mediate the stimulatory effect [70]. Therefore, the various factors influencing pigmentation are interesting to consider in order to optimize the process of pigment production. The factors stimulating carotenoid production for microorganisms are summarized by Bhosale [46]. Overall, conditions that typically stress the cells, thus threatening cell growth, trigger carotenoid biosynthesis [8]. These typically include nutrition depletion (N and P) and certain levels of oxidative stress. Carotenoid accumulation usually occurs during the later stages of cultivation which also indicates the association of nutrient depletion with carotenoid synthesis [8]. This section addresses the effects of light, pH, nitrogen and carbon sources, temperature, co-factors, surface active agents, oxygen level, tricarboxylic acid intermediates and morphology in regards to *Neurospora* spp. carotenoid production, complemented with interesting discoveries on the regulation

Table 2 Structures, application areas and color of the carotenoid compounds found in *Neurospora intermedia* N-1 namely lycopene, neurosporen, γ-caroten, β-carotene and phytoene

Carotenoids	Structure	Applications	Color
Lycopene		High antioxidative activity [64] with beneficial effects on health by fighting many diseases [65]. Used in nutraceuticals and related applications [66]. Used for food coloring (E160d) [67]	Dark red [7]
Neurosporen		Antioxidative properties [68]. Data not found on industrial applications	Yellow-orange [68]
γ-Carotene		Data not found on industrial applications	Orange red [7]
β-Carotene		Used for food coloring (E160a) [67]. Beneficial effects on health by fighting many diseases [65]	Yellow to Orange [7]
Phytoene		Key carotenoid intermediate as a precursor to other carotenoids	Colorless [69]

of pigment biosynthesis, particularly in *Fusarium* and *Monascus* species.

Oxidative stress

To survive and compete in the environment, fungi produce enzymes and secondary metabolites with various activities. One such factor influencing this processes is oxidative stress [71].

The degree of oxidative stress has been shown to influence pigment production in filamentous fungi, indicating that these pigments are involved in the defense mechanism of the fungi. Numerous studies have reported observations [23, 37, 46, 63, 72–77] that support secondary protective roles of carotenoids against oxidative damage [63, 77]. For example, carotenoids can act as antioxidants against reactive oxygen species (ROS) [62] or alleviate cell membrane damage [63]. They have been proposed to act as antioxidative agents to extend the survival time of the fungi by synergistic effects with other antioxidants [62]. The ability of carotenoids to function as antioxidants may be the reason why dietary carotenoids have been shown to inhibit the onset of many diseases, such as cancer, in which ROS are thought to play a role [46]. It is the conjugated polyene chain of carotenoids that is able to chemically react as the quencher of singlet molecular oxygen [63], with varying efficiencies among carotenoids [72] depending on the structures beyond the polyene chain [63]. Astaxanthin, among other xanthophylls, have higher antioxidant activities than hydrocarbon carotenes, which make it advantageous for the fungi to synthesize. For example, the fungi synthesize astaxanthin at the expense of β-carotene under enhanced oxidative stress. This was confirmed when an increased synthesis of astaxanthin and less β-carotene by *Neurospora* was observed at increased oxidative stress levels (by addition of H_2O_2 and $CuSO_4$) [63]. Other studies exposing carotenoid producing fungal strains to oxidants in order to increase synthesis of carotenoids have also been carried out [62, 78]. For example, in addition to the generation of reactive oxygen species (ROS) by respiration [8], exposure of filamentous fungi to paraquat (PQ) or hydrogen peroxide (H_2O_2) promote oxidative stress [79]. One study investigated the accumulation of the apocarotenoid neurosporaxanthin in *N. crassa*. They induced oxidative stress by exposing the fungi to high concentrations of oxygen and extracellular hydrogen peroxide (H_2O_2). In response to the elevated oxygen, the expression of genes encoding enzymes that are involved in the synthesis of carotenoids increased by a factor of five. On the other hand, H_2O_2 exposure resulted in a twofold increase in the accumulation of *al-1* mRNA [73]. Moreover, the addition of H_2O_2 to the fungi is suggested to work both as a pigmentation

trigger and as an antimicrobial agent, making it an interesting factor to consider for further research, provided that the fungi do not consume it too fast [62].

Stressing the cells by inducing the generation of active oxygen molecules in order to enhance carotenoid production [46] can be done in different ways, which are proposed as secondary factors. These secondary factors are discussed below together with factors that are able to inhibit cell growth.

Light

Carotenoids, in general, protect the fungi against UV-damage. The photoprotective function of *Neurospora* carotenoids are suggested to be due to their ability, as antioxidants, to quench damaging single molecular oxygen species generated by UV radiation [14]. Thus, light has been reported to influence carotenoid synthesis by inducing the enzymes involved in carotenoid synthesis [8]. The use of light as a factor in pigment production has been reported for *Monascus* spp. [27, 31], *Fusarium* spp. [80–82] and to some extent for *Neurospora* spp. [14, 23, 37, 63].

Carotenoids do not play a major physiological role in fungal cells, but they may have beneficial effects under certain adverse conditions, such as abnormal levels of UV light. This was corroborated by a study where albino mutants of carotenogenic fungi in *N. crassa* and others were compared with the counterparts of the same species with functional carotenoid synthesis. The lack of carotenoids showed no apparent phenotypic consequence on growth or morphology in laboratory cultures [37, 77]. *N. crassa* has been used as a model organism for photobiology research [83].

Stimulation of carotenoid synthesis by light is achieved at the transcriptional level, e.g., as seen by an increase in the mRNA levels of structural genes (*al-1, al-2* and *cao-2*) when neurosporaxanthin was produced by *N. crassa* upon irradiation [63]. The *al-3* gene in *N. crassa*, which is responsible for GGPP synthesis, has further been confirmed to be strongly regulated by light. The response was reported to be mediated by the photoreceptor and transcriptional factor called the white collar complex (WCC) comprised by the photoreceptors white collar (WC)-1 and its interacting partner WC-2 WC-1 is a blue light photoreceptor, responsible for the light-induced response. The protein contains a DNA-binding zinc-finger domain (that is able to bind specifically to the promoter of a blue light-regulated gene), and a PAS domain, called LOV (from "light, oxygen and voltage"). The WC-1 LOV domain binds to a light absorbing Flavin adenine dinucleotide (FAD) chromophore that convert light to mechanical energy. FAD displays a maximum absorption of light at 450 nm [84, 85], which explains the WC-1's

sensitivity to blue light. Thus, *Neurospora* perceives light through the WCC, which is activated by blue light and binds directly to light regulatory elements (LRE) in the promoters of their target genes to activate transcription [14, 86]. Consequently, light can be used to activate genes involved in the biosynthesis of pigments that have photoprotective functions [87]. Regarding *Fusarium* spp., carotenoid production is also stimulated by light, through the transcriptional induction of the structural genes *carRA, carB, carC and carT*, as the major regulatory. However, in *F. fujikuroi*, the main photoreceptor seemed to not be a White Collar protein, as found in e.g., *Neurospora* species [32, 35, 60]. Photocarotenogenesis in *F. fujikuroi* relies primarily on other putative photoreceptors, such as cryptochromes [60].

A study on carotenogenic fungi, including *N. crassa*, has shown that upon prolonged light exposure, there was a reduction in the level of transcription compared to that with exposure to light pulses [63]. Another report demonstrated that exposing *Neurospora* to light resulted in a rapid accumulation of colored carotenoids after 1 h, with increasing concentrations for up to 12 h. Aerobic conditions are required for these light responses to keep the photoreceptor in the right oxidation state [14]. The observation showing a cessation in the activation of gene transcription by light after a certain time is probably due the transient nature of light-induced transcription in *Neurospora*. Incubation in the dark is thus required before light responding transcription can be activated again. Regulation by light of pigment biosynthesis (photocarotenogenesis), is mediated, as other photoresponses in *Neurospora*, by the WCC. Upon extended illumination, the WCC-dependent transcript level decreases. The degree of 'photoadaptation' is by the fungus is modulated by the blue-light photoreceptor VVD. Strains with mutations in the *vvd* exhibited a sustained photoactivation of genes required for carotenogenesis [14, 87]. The amount and the intensity of light tolerated have been shown to vary with the strain, ranging 1000–5000 Lx [46]. Carotenoid photostimulation has also been reported to vary with different wavelengths for *Neurospora* spp. For instance, red light did not induce carotenoid biosynthesis in *Neurospora*, whereas wavelengths within 450–480 nm were shown to be effective [14]. Even though carotenogenesis in hyphal cells is induced by blue light and is lacking in mycelia grown in dark conditions, the synthesis of carotenoids has been found to be independent of light when coupled to conidiation and results in a pale pigmentation. [87]. Thus, *Neurospora* cultivated in the dark accumulate high amounts of the colorless precursor to carotenoids, phytoene, in their conidia. Illumination of these cultures grown in the dark induces transcription of enzymes responsible for the desaturation of phytoene,

which leads to the formation of colored carotenoids [87]. More detailed descriptions of the regulation of carotenoid biosynthesis by light in *Neurospora* are very well illustrated by Díaz-Sánchez Violeta et al. [35], Avalos Javier et al. [14], Olmedo Maria et al. [87] and Muñoz Victor et al. [88].

Further studies on light as a factor in pigmentation have found that *N. crassa* accumulates the orange carotenoid neurosporaxanthin after light exposure, with higher pigment content correlated to strains at lower latitudes. The increased carotenoid accumulation was suggested to serve as a second protection from harmful effects of UV radiation in *Neurospora* spp., since species that accumulate more carotenoids are more resistant to UV radiation [23, 37]. The ability of the fungi to accumulate different amounts of carotenoids [23] and the higher resistance to irradiation being connected to carotenoid production [72] strongly support light as a stimulating factor in pigment production. An attempt to increase carotenoid biosynthesis by *N. intermedia* in liquid substrate fermentation showed that fermentation with incubation for 3 days in the dark followed by 4 days under blue light was better than fermentation for 7 days in the dark. However, this difference in carotenoid production in different phases of life cycle has only been investigated in yeast (*Sporobolomyces ruberrimus* and *Cystofilobasidium*), and it showed a higher tolerance in the stationary phase compared to the exponential growth phase [89]. Direct information on such phenomenon for filamentous fungi is therefore still lacking.

Nitrogen source

Limiting the availability of nitrogen is suggested to increase the concentration of total pigments in most strains of filamentous fungi [8], and this has been investigated mainly for *M. purpureus* [90–93] and to some extend in *Fusarium* [17].

The pigment variation depending on the nitrogen source is suggested to be influenced by the rate of amino acid metabolism [8] rather than directly by the nitrogen compound itself when the source is in the form of amino acids [94]. In this regard, amino acids that are metabolized slowly are favored since they induce nitrogen limitation to a greater extent. The carbon to nitrogen (C/N) ratio in the culture medium has also been proposed to be an important factor in addition to the nitrogen source. Generally, a high C/N ratio has been reported to promote carotenogenesis in many fungi as this condition limits the cell's access to nitrogen [8]. The increased synthesis of pigments in limited nitrogen has been suggested to be related to a response mechanism to excess energy and carbon that cannot be used for protein synthesis or growth [8]. The balance between carbon and nitrogen

sources at C/N 9:1 was reported to increase β-carotene production by *N. crassa*. The optimal ratio of carbon and nitrogen for the growth of *Neurospora* spp. in general has been reported to be within 7:1 and 15:1 [95], but its link to pigment production has not been investigated. Similar stimulation of carotenogenesis by a limiting nitrogen content has been reported for *Fusarium* spp. Rodríguez-Ortiz Roberto et al. [32] discovered that nitrogen exhaustion increased carotenoid production in both wild type and the carotenoid-overproducing mutants (*carS*, containing high levels of mRNA for the *car* genes) which had a high synthesis of carotenoids irrespective of illumination. The authors suggest regulatory connection between carotenoid biosynthesis and nitrogen controlled biosynthetic pathways in *Fusarium* [22]. The results indicate similarities in the regulation of nitrogen on carotenoid synthesis with *N. crassa*. Incubation of *N. crassa* under nitrogen starvation also increased the levels of the corresponding *al-1* and *al-2* independent on light [9].

Using a complex nitrogen source or the addition of individual amino acids have been shown to influence the number, type and excretion rates of different pigments that are formed [31, 42, 75, 96]. It has been suggested that the stimulating effect of amino acids on the production or liberation of pigments is caused by an increase in solubility because the derivatives linked to amino acids are more soluble than the original pigments [97]. It may be possible to obtain a desired number of pigment components by using a defined nitrogen source which provides specific amino acid moieties that can be incorporated into water-soluble extracellular pigments [97]. However, both positive and negative effects of different amino acids and their concentration limitations on pigment production have been reported. Furthermore, the mechanism behind these findings are still scarce in the literature and limited to only a few strains based on studies with *Monascus* spp. Such a phenomenon has not, to the best of our knowledge, been investigated for *Neurospora* spp.

Carbon source

The carbon source is the most studied parameter regarding its effect on carotenogenesis in different fungi [77]. Singgih Marlia et al. [9] evaluated the carotenogenesis of *N. intermedia* in liquid fermentation. The highest production of carotenoids (24.31 µg/g spores) was achieved when 2% w/v maltose was used under aerobic conditions [9]. The other tested carbon sources included glucose, sucrose and maltose. Interestingly, a higher than 18 g/l glucose concentration was reported to reduced pigment production, perhaps due to respiro-fermentative metabolism [92]. These findings, however, oppose the previously mentioned report that suggested that an excess amount of energy and carbon limits the nitrogen source

and stimulates pigmentation. The initial amount of the carbon source added to the cells has been investigated by Hailei et al. [41]. Their findings showed that it was the accumulation of glucose metabolites (glycerol, ethanol, organic acids and other substances), not the exhaustion of glucose, that played a major role [13, 41]. Another potential explanation for the relationship between stress due to carbon starvation and pigment production in *Neurospora* spp. could be the suppression of central carbon metabolism, which results in an increase in the acetyl-CoA pool available for other pathways, such as the mevalonate pathway (shown in part one in Figs. 4, 5, 6) [51]. The addition of ethanol as a carbon source has been reported to increase carotenoid production by growth inhibition, activation of oxidative metabolism and induction of HMG-CoA reductase [8], resulting in the production of lycopene and neurosporaxanthin in *N. crassa*.

The specific time of addition of different carbon sources has also been shown to influence carotenogenesis in *N. intermedia*, with a positive effect seen by addition at the middle of the log phase than at the beginning of the stationary phase. From this the authors concluded that the carotenoid content likely increased in the middle of log phase and then remained constant, which is typical for secondary metabolites [36]. However, general information regarding the time of addition on pigment biosynthesis is scarce in the literature and needs to be investigated further.

Although studies regarding the effect of carbon source is limited for *Neurospora*, more extensive research can be found for *Monascus* spp. [31, 77, 90, 98–101].

pH

It has been previously reported that many kinds of fungi in submerged cultivation respond to more acidic pH conditions with the accumulation of pigments [102], probably due to the stress condition. Changes in pH during growth depend primarily on the nitrogen and carbon source in the medium [103]. Since research studies have supported the influence of pH on the pigment production by *Monascus* spp., it has been hypothesized as a factor to influence pigmentation in *N. intermedia* as well [31, 42, 44, 90, 92, 104]. The strong effect of pH on the biosynthesis of pigments in *Monascus* spp. has been proposed to be associated with changes in the activities of proteins. Changing the pH from neutral or slightly alkaline to more acidic has been shown to favor the cyclization of lycopene to β-carotene, and this has been applied in a patented fermentation process to improve lycopene yield [70]. It has also been noted that maximum pigment production is associated with the combined effect of pH and temperature in the culture medium. The effect was suggested to be associated with cellular growth, oxidation

processes and metabolic flows regulated by molecules such as adenosine triphosphate (ATP). A change in the pH affects oxidation and reduction processes of molecules in the cell, thereby affecting the redox flux and the oxidative state of important energy molecules such as ATP, which has important roles in the cell metabolism. Changing the pH can therefore cause different metabolic pathways, substrate oxidation, and regulation of metabolic and osmotic processes, resulting in different end-products [104]. Nevertheless, the influence of pH on pigment production by *Neurospora* spp. is currently unexplored in the literature.

Temperature

Changes in temperature affect cellular growth and metabolite production by enzymes, including those involved in carotenoid production in microorganisms [46, 77]. Thus, temperature is suggested to be a factor that may be utilized to regulate enzymatic processes connected to pigment production by the fungal cell [104]. The order of reactions in carotenoid biosynthesis can result in different carotenoids being produced depending on the temperature of growth [14, 47]. For example, at low temperature, the oxidative cleavage in the carotenoid pathway is favored over the cyclization reactions. This implies a competition between the enzymatic activities for the oxygenase and the cyclase for the same substrate (3,4-didehydro-lycopene, Fig. 4, 5, 6), and the conditions favoring one or the other depends on the growth temperature [14]. Cultivation at lower temperatures has also been reported to block the production of torulene and result in greater accumulation of β-carotene, probably due to changes in enzymatic activities [105]. Furthermore, illumination specifically at low temperatures increased the proportion of neurosporaxanthin in *Neurospora*. In general, the temperature influences not only the type of carotenoids produced but also the total carotenoid content in various carotenoid producing fungi [14, 17]. Studies have shown that exposing *Neurospora* mycelia to temperatures between 12 and 6 °C with illumination resulted in the highest response, although this phenomenon has not been described in other carotenoid-producing fungi [14]. This increase in carotenoid production below the optimal growth temperature is suggested to be an acclimating response compensating for the downregulation of metabolic processes and other changes, such as the fluidity of the cell membranes [46]. This strengthens the suggestion that carotenoids serve as a secondary protection for the cells.

Co-factors

Co-factors such as metal ions and salts greatly affect fungal metabolism [106] and have been demonstrated

to influence carotenoid synthesis [107]. Magnesium and calcium are considered macronutrients for filamentous fungi, whereas iron, manganese, zinc and copper are considered micronutrients [46]. It is probable that the effect of such co-factors on carotenogenesis occurs due to an activation or inhibition mechanism on specific carotenogenic enzymes [77], such as microbial desaturases [94]. Only a few attempts to address the biological roles of these cations in fungi have been reported [46]. One of the studies added up to 12 mM Magnesium ions which showed a stimulatory effect on pigmentation by *N. intermedia* [9, 31]. The magnesium ions were reported to stimulate the conversion of GGPP into phytoene (catalyzed by phytoene-synthase enzyme). Phytoene is desaturated to produce lycopene, which is a precursor of cyclic carotenoids in *N. intermedia* [9, 106]. Another study aimed to optimize red pigment biosynthesis by *M. purpureus* under solid-state fermentation and achieved higher concentrations of pigments by adding manganese rather than other macronutrients such as magnesium and calcium. This result was in line with findings from another study that cultivated *B. trispora* in trace amounts of manganese ions [46] or trace amounts of copper ions [108]. Manganese is a known cofactor for enzymes involved in carotenoid biosynthesis [109]. Manganese-dependent enzymes act on carotenoid production by influencing the concentration of cyclic AMP. cAMP has been shown to control a variety of functions in fungi [110, 111], which may include pigment production. One study investigated the role of exogenous cAMP on conidiation and carotenoid biosynthesis in *N. crassa* and found that it suppressed conidiation and lowered carotenoid synthesis [111]. However, the influence of abnormally low levels of cAMP is yet to be investigated. Nevertheless, another study reported that mutants of *N. crassa* with defects in the acyA gene coding for cAMP resulted in lower intracellular cAMP levels and contained more carotenoid pigments than wild-type cells [112]. García-Martínez et al. [113] also reported enhanced production of red pigments by *F. fujikuroi* with defects in the acyA gene [113]. A negative relationship between cAMP level and the accumulation of carotenoids in *N. crassa* [73, 114] has also been reported.

Another potential mechanism for how metal ions affect pigment production relates to the formation of active oxygen radicals ("Carotenoid biosynthesis" section) [94]. For example, pigment production was suggested to be induced by using ferrous ions to generate hydroxyl radicals ($H_2O_2 + Fe^{2+} \rightarrow HO^- + HO^*$). Additionally, copper, zinc, lanthanum and cerium have been shown to have stimulatory effects on carotenoid yield via generation of active oxygen radicals [46]. However, negative effects on pigment production have also been reported

for ferrous or cobalt ions, thereby indicating other effects of the metal ions in contributing to cell growth rather than promoting metabolic pathways [102]. Inhibition of carotenoid production was also shown in the case of adding manganese ions to a culture of *Xanthophyllomyces dendrorhous* containing oxygen radical generators. This effect was suggested to result from manganese acting as a scavenger or antioxidant by de-activating radicals, thereby resulting in lower stress to the fungi. The effect was thus suggested to be dependent on the concentration of manganese [46].

Surface active agents

It has been reported that surface active agents (also known as surfactants), such as corn oil [90], Triton X-100 [13, 115], Tween-20, Tween-80 [42, 90], Span 20 [116] and olive oil, show positive effects on the metabolism of both intra- and extracellular pigments [42] depending on the strain and surfactant used. The amount of pigments that can be produced extra- vs. intracellularly varies with different cultivation factors and choice of strain. Secretion of extracellular pigments is favored over intracellular production since it requires less work to extract the pigments [115]. It has been reported that pigments access the aqueous environment by association with proteins or other polar compounds [117]. Surfactants are amphipathic substances that are able to adsorb onto surfaces of interfaces in dispersions and alter the interfacial free energy. Nonionic surfactants are able to form micelles in aqueous solutions and can be used as permeabilization agents for the secretion of hydrophobic intracellular pigments [115]. Transporting intracellular pigments to extracellular micelles will prevent pigment degradation and lower the intracellular pigment concentration, which otherwise decreases yield by product inhibition [13]. Surfactants are also suggested to act on cell membranes, increasing their permeability to release both enzymes and pigments into the medium. This was observed when Tween 80 was added to a culture with *Aspergillus amylovorus*, and it increased the pigment concentration in the medium by a factor of six [118]. It was also suggested that Triton X-100, which is able to solubilize membrane proteins, can increase the access of carotenoids in *Neurospora* to the aqueous environment by associating the pigments with membrane-bound enzymes [111]. In addition, Span 20 was suggested to affect β-carotene production by altering the fungal morphology [116] ("Carotenoid biosynthesis" section). However, when *N. crassa* was cultivated with the addition of Tween 40 (0.8%), carotene production was increased, but the carotene remained inside the cells [119]. Therefore, the way in which surfactants act on pigments is not yet fully understood since their effects are not consistent [118].

Oxygen level

Because carotenogenesis is an aerobic process, oxygen supply is another important parameter of carotenogenesis [77, 120, 121]. Poor oxygen supply to the culture has been reported to decrease pigment production [104], and a critical dissolved oxygen tension (DOT) between 15 and 20% air saturation has been suggested for efficient carotenoid synthesis in fungi [8]. The crucial role of oxygen in pigment production is probably due to its ability to act as an electron acceptor during oxidative phosphorylation and as a substrate of monooxygenase. Oxidative phosphorylation and monooxygenation are involved in metabolite synthesis. In particular, monooxygenases are more active in secondary metabolism [31]. Thus, the relationship between oxygen supply and pigmentation is important to understand in order to achieve optimal production of pigments.

Tricarboxylic acid intermediates

Intermediates of the tricarboxylic acid (TCA) cycle, such as succinate, citrate and malate, have been reported to stimulate carotenoid biosynthesis under aerobic conditions [46]. These intermediates form a carbon skeleton that can be used in pigment synthesis by different carotenoid-synthesizing species. Two potential explanations for the effect on pigment production by TCA cycle intermediates have been put forward in the literature. One explanation is that these intermediates are able to act specifically on some of the key enzymes involved in the isoprenoid biosynthesis pathway such as acetyl CoA-carboxylase, the starting substance for isoprenoid synthesis. The other explanation is that the effect is caused by the increasing pool of oxaloacetate from the added intermediates, since oxaloacetate is further decarboxylated to pyruvate leading to an increase of acetyl-CoA [46]. Likewise, high respiration rates and tricarboxylic acid (TCA) cycle activity are also associated with the production of large quantities of ROS [46], which could be another explanation for the increased pigment production when these intermediates are added.

Addition of citrate in *X. dendrorhous* and malate in *Blakeslea* and supplementation of 28 mM citrate to *Blakeslea trispora* have been reported to increase carotenoid production [8]. The degree of stimulation by the intermediates of the TCA cycle has also been reported to depend on the time during the cultivation when they were added to the medium [46]. However, these observations have not been investigated in ascomycetes, as far as we know.

Morphology

Filamentous fungi can adopt diverse morphologies when cultivated in submerged cultures such as uniform

and long filaments or entangled filaments in pellets or clumps. Morphology is also considered as a factor in pigmentation since it has been reported to influence the secretion of metabolites [13, 116]. A high fungal concentration with entangled mycelia or filaments results in a highly viscous suspension with non-Newtonian properties, which reduces the homogeneity of nutrition, temperature, oxygen, and other parameters. Growth in the form of dense pellets generates a less viscous medium with Newtonian properties, but the internal mass transfer rate is limited by pellet size and compactness. These changes in gas–liquid-mass transfer can thus affect the formation and secretion of products/metabolites [13]. One specific morphology that promotes pigmentation could be aimed for by adjusting stimulatory factors such as the cAMP level or fermentation conditions such as pH or oxygen supply rate. The influence of pH on morphology is explained in the study by Méndez et al. [104]. The study suggests that chemical or structural changes in the cell membrane induce morphology changes in the cell wall, which leads to an overproduction of pigments as a defense mechanism to regulate damage at the membrane level [104]. However, clearly established links between fungal morphology and pigmentation levels are still missing in the literature [104].

Isolation, analysis and identification of carotenoids
Extraction of carotenoids
In general, the isolation of intracellular carotenoids from filamentous fungi commonly involves a pretreatment step such as drying and/or cell disruption, an extraction step and a saponification step. Cell disruption is often necessary for intracellular carotenoids in order to increase the recovery. Many different methods for cell disruption have been suggested in the literature, including mechanical disruption (e.g., sonication, high-pressure homogenization, grinding, and bead-milling), and non-mechanical disruption (e.g., microwave assisted extraction, enzymatic hydrolysis, and ionic liquids) [122]. The preferred extraction procedure is based on a quick process that efficiently releases all the pigments from the matrix into the solution without altering them [56], while using environmental-friendly solvent(s), if possible. There is no standard method for carotenoid extraction from fungi, but based on previous studies, the general process involves the mixing of dried or wet biomass with organic solvents (liquid–liquid system), followed by mechanical disruption of the cells and subsequent centrifugation or filtration (pigment particles are approximately 1–2 μm). After filtration, the solid residue is re-extracted, and the process is repeated until the residue becomes pale. Three extractions are usually sufficient [56]. A saponification step is sometimes required for fat-rich biomasses in order

to remove lipid contamination and to hydrolyze carotenoids found in ester or di-ester forms. The contaminating lipids may otherwise interfere with the chromatographic separation, identification and quantification of the carotenoids in later stages. When indispensable, saponification is most commonly carried out with 10% potassium hydroxide in methanol or ethanol, at temperatures below 60 °C to prevent carotenoid degradation [122, 123]. The carotenoid solution is then washed with water to remove the alkali. However, carotenoids with allylic hydroxyl and keto-groups, such as neurosporaxanthin, undergo oxidation in the presence of alkali and air. Saponification is not recommended for these carotenoids, or it must be carried out anaerobically. This procedure is further described by Schiedt Katharina et al. [124]. Finally, the total carotenoid content is commonly analyzed by measuring the maximum absorbance of the extracted pigments by spectral analysis using a spectrophotometer. The maximum absorbance for these pigments is in the range of 450–480 nm, depending on the type(s) of carotenoids present in the sample [14]. It should be noted that the position of the absorption maxima and the shape of the spectrum can vary by a few nanometers depending on the temperature and the presence of organic solvents [125].

There are some aspects to take into consideration when selecting the organic extraction solvent/solvents to use in order to optimize the solubility and stability of the carotenoids. Solvents with low boiling points are preferred since they are more easily removed through evaporation. If the fungal biomass contains large amounts of water, a water-miscible organic solvent (e.g., methanol or acetone) should be used for better solvent penetration [56]. Furthermore, the choice of solvent/solvents depends on the polarity of the carotenoids. Organic solvents commonly used to extract pigments include benzene, petroleum ether, hexane, acetone, chloroform, dimethyl ether, methanol, ethanol, [126], other alcohols, and ethyl lactate (described as a so-called green solvent) [127]. The organic solvents allowed in the EU for the extraction of natural food colorants are water, ethyl acetate, acetone, n-butanol, methanol, ethanol and hexane, while those allowed in the US are isopropanol, methanol, ethanol, hexane, and acetone [128]. Acetone is the most widely used solvent for carotenoid extraction since it is fairly harmless, inexpensive and readily available [56]. However, the polarity of the solvents and the pigments need to be considered. A better extraction is obtained when polar carotenoids, such as neurosporaxanthin, are dissolved in polar solvents, while less-polar carotenoids, such as lycopene, and γ- and β-carotene, dissolve better in non-polar solvents [129]. Therefore, if the fungal biomass contains a mixture of pigments, two- or three-stage extraction methods using organic solvents of different polarity may

be required. A combination of different solvents has also been suggested to disrupt the cell and release intracellular pigments to a greater extent. For example, a mixture of petroleum ether, dimethyl sulfoxide and acetone was found to improve the extent of carotenoid recovery from *Rhodotorula glutinis* compared to the individual solvents [126]. Moreover, novel alternative methods that are more environmentally friendly compared to the conventional organic solvent extraction approaches are being developed. These include super- or sub-critical solvent extraction, switchable solvent extraction, wet extraction, and, more recently, ionic liquid-assisted extraction [122] and CO_2 extraction [127]. These methods should be considered not only for environmental advantages but also for potential cost saving and increased efficiency [51, 122]. Research on "greener" extraction processes for fungal pigments should thus be further evaluated as an alternative to existing processes. The potential of emerging greener extraction systems for filamentous fungi pigments have been reviewed by Dufossé [15], Gong Mengyue et al. [122].

Only a few studies focusing on the isolation of carotenoids from *Neurospora* are available. One study analyzed the carotenoids produced by *N. intermedia* in liquid fermentation. Disruption and solubilization of the carotenoids present in the biomass were made using acetone as the extraction solvent and sonication for cell disruption. The total carotenoid content was determined to be 0.8 mg/g spores by measuring their absorbance by UV/Vis spectrophotometer at 480 nm [9]. Another study [36] utilized a two-stage extraction. Methanol was used as the first extraction solvent and acetone was applied as the second extraction solvent on dry mycelia, at 60 and 50 °C, respectively. The resulting methanol/acetone fraction containing the extracted carotenoids was collected, centrifuged and the spectral absorption at 470 nm was determined [36]. In another study, extraction of carotenoids from conidia of *N. crassa* was performed at an elevated temperature but with ethanol as the solvent, and the carotenoid content was estimated by measuring the absorbance at 475 nm [112]. The results obtained were expressed in terms of units of absorbance (U/g dry cell mass), a value proportional to pigment concentration [92]. The λ_{max} values and solvents used for extraction of various carotenoid pigments from *Neurospora* spp. are cited in Table 1.

Purification, characterization and identification of carotenoids

Since *N. intermedia* and most other filamentous fungi accumulate a complex mixture of pigments, the next step includes techniques to purify and quantify the pigments that are present in the mixture. A purification step

is needed if the extracted carotenoids are to be used, for example, in food and cosmetic applications, or for certain quantification methods [51]. Purification and quantification of the extracted pigments are based on chromatographic and spectroscopic properties, and chemical tests. Column chromatography, thin-layer chromatography (TLC), ultraviolet–visible spectrometry, and high-performance liquid chromatography (HPLC) with online photodiode array detection are commonly used to separate, identify and quantify carotenoids [130]. HPLC systems combined with photodiode array detector (PDAD) separates and identify carotenoids found in fungi, in view of their high sensitivity, reproducibility and short analysis time, while minimizing isomerization and oxidation of unsaturated carotenoids [131]. This technique relies on the characteristic differences in wavelength maxima for each carotenoid and the spectral fine structure. Larger numbers of conjugated double bonds in the carotenoid structure will shift the wavelength maxima (λ_{max}) towards longer wavelengths The long conjugated polyene system, which makes the *trans* isomers linear and rigid molecules, is an important property for the interaction of carotenoids with the stationary phase in HPLC and for their absorbance of light in the visible region at 400–550 nm (Figs. 4, 5, 6). Generally, carotenoids absorb light maximally at three wavelengths (three-peak spectrum) where most other substances do not absorb. This property can be utilized when carotenoids are to be identified in complex mixtures [56]. In a previous study, measurements of the amount of β-carotene produced from *N. intermedia* have been carried out by HPLC using a C_{18} column with acetonitrile: methanol: 2-propanol (85:10:5) as the mobile phase. Detection of the carotene was carried out with a UV/Vis detector at 450 nm and the compound was compared with β-carotene standards [9].

Characterization and identification of carotenoids are described in more detail by Gross [123].

Industrial applications and challenges
Potential application areas and present status of the market

The pigment industry, mainly the food industry, is looking for potential sources and uses of natural-origin pigments without harmful environmental and health-related side effects, in addition to new colorants with improved functionality and a constant supply of raw materials from cheap and reliable sources. Consequently, the application areas of naturally derived pigments from *N. intermedia* are broad and have a bright future on the market if successful at an industrial scale.

Potential applications of carotenoids include their use in animal feed to improve the nutritional profile and to enhance the appearance of poultry skin, salmon meat and

shade of egg yolks, the colors of which are determined by the animals' diet. For instance, the apocarotenoid astaxanthin, mostly produced from microalga and bacteria, has an orange-to-reddish color and is used currently as feed-additives in aquaculture due to its health-promoting properties. This compound provides the typical orange pigments to salmonids, lobsters and trout in addition to being used as a color enhancer in the diets of chickens to improve the color of egg yolks to meet consumers' expectations [132]. Some natural-origin pigments are also used as intermediates for dyestuff for textiles and biodegradable polymers [133]. Furthermore, they can replace synthetic pigments used in food, drugs, cosmetics and healthcare products. Extraction of pigments from plants has been the predominant source of natural pigments thus far, but the use of these pigments is limited by their available quantities [134], irregularity of harvests, land use and labor-intensive characteristics [8]. Microbial pigments, on the other hand, have shown greater stability against external stress such as light, pH and temperature and have high water solubility [6, 104] compared to pigments from plant sources [7]. Microbial pigment production is also an environmental friendly method compared to synthetic pigment production [8]. To date, more than 600 carotenoids have been found to be produced by carotenogenic microorganisms, but only astaxanthin and β-carotene are commercially produced by microbial fermentation. In 2010, synthetic colorants accounted for 40% of the colorants available in the market, whereas natural-derived colorants and nature-identical colorants (i.e., man-made pigments which are also found in nature [135]) accounted for 31 and 29% of the market, respectively. However, due to the advantageous associated with fermentation-derived natural pigments from microbial sources mentioned above [9] these may be a promising alternative that could tackle some of the current problems. According to Mapari et al. [7], products with natural colorants are expected to replace synthetic colorants in the future.

Microalgae are one alternative source of colorants, but their low productivity limits their use on a commercial scale. The carotenoids in fungi (include β-carotene, γ-carotene, torulene and their hydroxyl- and keto-derivatives [136]) grant industrial interest. These natural-origin pigments could be exploited for their antioxidant, provitamin A activity [137] and beneficial effects on the onset of many diseases [13, 65]. Strains of *Basidiomycetous* fungi have been used for coloring silk and wool but are limited by their difficulty to grow under laboratory conditions and are therefore not suitable for production on an industrial scale. On the other hand, ascomycetes have been traditionally used in different parts of the world for hundreds of years for food coloring, and they can be easily cultivated to give high yields [5]. The interest in these pigments is also growing because many ascomycetes are known to secrete pigments with improved functionality (e.g., with anticancer properties) [13]. Particular attention has been given on *Monascus* pigments which have been shown to possess heat and pH stability during food processing [12].

Another potential pigment producing fungi with biotechnological applications are species of *Fusarium* that synthesize e.g., bikaverins (red pigment) [16] and carotenoids [32]. Research projects using *Fusarium sporotrichioides* for production of β-carotene and Lycopene [15, 138] and *F. fujikuroi* for improved neurosporaxanthin [40] and β-carotene [139] production suggest high potential in the field. Bikaverin is also biotechnologically relevant due to its selective toxic effect against tumoral cells [140] and its antibiotic activity against some phytopathogenic fungi and protozoa [141, 142]. One drawback with *Fusarium* spp. is, depending on the growth conditions, its potential of co-production of mycotoxins such as trichothecenes, fusarins, and zearalenone [10, 22]. This can be overcome by carefully regulated cultivation conditions or genetically modifying the fungi. *Fusarium venenatum* is known for its myco-protein production used for human consumption [11]. The potential industrial interest may increase by genetically modify the fungus to produce carotenoids that would add health value to the myco-protein product [143].

Even though most fermentative food-grade pigments from filamentous fungi are at a development or research stage [13], there are a few that already exist in the market. These include Arpink red™ (now Natural Red™) from *Penicillium oxalicum* (manufactured from Ascolor Biotech) [5, 65, 83], which has received a 2-year approval by the EU to be used as a food additive in the Czech Republic from 2004 to 2006 (current status of approval unknown) [4], riboflavin from *Ashbya gossypii* [97, 144], lycopene and β-carotene from *B. trispora* (produced by Gist-Brocades, now DSM; approved in 2000 by the EU Scientific Committee on Food Safety) [46, 47, 97] and previously mentioned *Monascus* pigments [66, 97]. The industrial production and use of β-carotene and lycopene from *B. trispora* as food colorant have been approved by the European Commission in 2000 and 2006, respectively [51]. As an example, β-carotene has been developed to yield up to 30 mg/g dry mass or approximately 3 g/l of culture in a submerged fermentation process [2]. Production of β-carotene and lycopene at larger scales (25 m^3) under normal fermentation conditions are expected in the near future [46]. Regarding the industrial applications for *Monascus* pigments, the most common species used are *M. purpureus*, *M. pilosus* and *M. ruber* [13, 145] due to their production of orange, red and yellow pigments

[29, 30] that are used as natural colorants for making red rice, red soybean cheese, wine, marine and meat products [4, 31]. Their yellow and orange polyketide pigments have been commercially produced and legally used as food colorants in the Southeast Asia for more than a thousand years [2].

Until now, β-carotene has been the most economically important carotenoid with a market share of US$261 million in 2010, followed by astaxanthin and lutein [77] with global market shares of US$252 million (2010) and US$60 million (2011), respectively. However, only 2% β-carotene is derived from natural sources [122], mainly from palm oil and to some extent fermentatively produced by *B. trispora* [12]. β-Carotene is the most well-known dietary source of provitamin A and believed to be the most important carotenoid in human nutrition [8]. It has been used in health and food products and to enhance the color and appearance of fish, birds, processed meats and tomato ketchup among other applications. Lycopene, produced mainly from tomatoes [12], has many applications within the food industry due to its red color and strong antioxidant activity (strongest among the carotenoids) [8], and it is used mainly as a nutritional supplement or active ingredient in cosmetic products [146]. It has also been suggested to have anti-carcinogenic properties [147]. Additional carotenoid compounds found in *N. intermedia* and their industrial applications can be found in Table 2.

Challenges for large-scale pigment production and future prospective

Pigments used for food colorants need to be approved by a regulatory agency and the most important factor in their consideration is safety. Consequently, there are many limitations and legislations covering this issue [13]. The approval (2000) by the EU for the use of filamentous fungal carotenoids as food colorants has renewed the interest on the use of fungi as carotenoid producers. The United States regulation on the use of pigments in food and conditions for their use is outlined in Code of Federal Regulations, Title 21 (21CRF), which is also followed by several other countries, while the Australian and Japanese legislations are also used in a number of other Asian countries. The legal use of filamentous fungi in different parts of the world varies with local and traditional usage of colorants [2]. For example, *Monascus* pigments have been used for more than thousands of years in China, Japan, and other Southeast Asian countries but are not permitted as colorants in the European Union and in the United States [97, 145]. A legislation by the European Parliament reinforced the need for alternative "green" carotenoids by requesting that foods containing synthetic colorants, including quinolone yellow and tartrazine, require a label stating "may have an adverse effect on activity and attention in children" [5].

Other important desired features include a high pigment yield, ability to dissolve in water and the production of stable pigments [7]. Furthermore, the potential production of other secondary metabolites such as mycotoxins [30, 120], the kinds of carotenoids produced, [77] and whether the pigments are produced extra- or intracellularly are important factors to consider [97]. The production of pigments will also depend on consumer approval and the production costs required to bring the product to the market [7].

Regarding the pigment yield, there are generally two natural ways in which the amount of pigments can be increased by improving the fungal growth, or by increasing the cellular accumulation of pigments [148, 149]. The problem, however, lies with the difficulty in increasing both the biomass and the pigment production, which would be optimal for industrial production. Biomass and pigment yields are suggested to be negatively correlated [104]. An increase in the biomass yield is connected to the abundance of nutrients in the medium whereas pigment production seems to be increased under nutrient-poor conditions and external stresses. Therefore, the relationship between biomass and pigment production needs to be fully elucidated in order to optimize and control pigment production. One way to control pigmentation in fungi is to use recombinant DNA techniques. Genetic modifications have been previously applied to alter the activity of enzymes involved in carotenoid biosynthesis [46, 148–151]. All carotenoids share a common precursor, which can be utilized to manipulate the carotenoid biosynthetic pathway by combining biosynthetic genes and produce a much wider range of carotenoids [8]. However, genetic modification lowers the acceptance for the use of the produced pigments in food and feed industries [42].

Another challenge involves the differences in bioavailability and absorption rates between different types of pigments. Bioavailability refers to the amount of pigments absorbed in the body (to become available for physiological functions or storage) and is thus desired if the pigments are to be used in the feed industry, for instance. A high absorption rate, in this context, relates to the release of pigments to the product matrix, for example, in coloring a food product. Naturally, the food matrix itself will affect the absorption and the release of pigments [137]. Hence, the choice of pigment is important, as different pigments vary in bioavailability and absorption to the animal, human and food product matrices [127]. There are still limited data about differences in bioavailability and absorption between various natural-origin pigments, but, in general, it seems that more polar carotenoids

(xanthophylls) are absorbed more efficiently than hydrocarbon carotenoids (carotenes) [152, 153].

The few strains of filamentous fungi that already exist in the market still have some challenges that need to be addressed to be able to compete with other pigment sources on the market. For example, the carotenoids produced from *B. trispora* exhibit good pH stability in most foods but are easily oxidized [12], although this can be countered by addition of antioxidants [154]. Challenges for the commercial production of pigments produced from *Monascus* include the pigments' low water solubility, sensitivity to heat, instability at pH 2–10, and fading color intensity with light. To make the pigments more water-soluble, methods have been developed to substitute the replaceable oxygen in the pigment structure with nitrogen from the amino group of various compounds such as proteins, peptides and amino acids, and some patents have addressed these challenges [144]. Furthermore, when using *Monascus* for the production of pigments for feed and food applications, another problem is the co-production of citrinin, an azaphilone with nephrotoxic and hepatotoxic properties. Safety concerns regarding citrinin, a compound classified as a potential human carcinogen [30, 120], limits the use of *Monascus* pigments although some edible *Monascus* spp., such as *M. purpureus*, have been used for the production of red fermented rice for over a thousand years in Asian countries, [143]. This has prevented the approval of *Monascus* pigments as food colorants in the European Union (EU) and the United States (US) [5]. Similarly, *P. oxalicum* also produces the yellow toxic pigment, secalonic acid D [5]. On the other hand, *N. intermedia*, isolated in 1842, has neither been observed to cause diseases in plants or animals nor to produce dangerous secondary metabolites (e.g., mycotoxins), and thus, it has been extensively used in the food and beverage industry [19]. This edible fungus has traditionally been used for the production of the Indonesian fermented food oncom [6], contributing to the characteristic orange color of the dish. Hence, *N. intermedia* is generally recognized as safe (GRAS) [19]. Moreover, it is closely related to *N. crassa*, a very well-evaluated model organism [155]. Nevertheless, among the *Neurospora* spp., only *N. crassa* has so far been evaluated for industrial production of pigments in a research project [130].

Since most fungi produce a mixture of pigments, another challenge is to be able to direct the pigment production towards one specific colored dye in the future. The amount of different pigments can be adjusted by altering the substrates, the operational conditions (pH, dissolved oxygen, temperature) and fermentation mode [solid state fermentation (SSF) or submerged fermentation (SmF)].

The choice between solid state fermentation (SSF) and submerged fermentation (SmF) has been discussed extensively in previous reviews regarding pigment production by filamentous fungi, especially for fermentation with *Monascus* spp. [26]. SSF is generally the process for the production of *Monascus* spp. pigments and is known to provide more pigments than SmF. However, research on the use of SmF has demonstrated easier handling, lower production costs, shorter cultivation times and higher product quality [31, 156]. Furthermore, there are several SmF techniques that can be used, such as batch or continuous mode, in order to achieve optimal pigment yields [13]. To date, most of the industrial production of carotenoids by filamentous fungi use SmF, while the use of SSF is still in the exploratory stage [107]. SmF may generate a higher pigmentation because oxygen and light are required for maximum pigmentation. To the best of our knowledge, studies focusing on pigmentation by *N. intermedia* have not addressed different fermentation strategies. Therefore, it remains to be seen if SSF will compete with SmF in the future. Furthermore, if pigment production is to be incorporated in an already existing biorefinery, the impeller design and configuration or the aeration rate in the stirred-tank and airlift bioreactors should be optimized for optimal gas exchange in order to increase pigmentation.

Pigments as a value-added product within biorefineries

To produce cost-effective pigments with less environmental impacts, the choice of substrate is of great importance. Using residues from industrial processes to create new products is certainly something worth aspiring to.

Substrates commonly used for solid state fermentation of *Neurospora* spp. have been based on waste products of plants or cereals that are rich in amino-acids and carbohydrates [106]. β-Carotene production by *N. crassa* has been investigated in a study by Thomson ISI [39] where the fungi were grown on various residues with the aim to produce carotenoid rich feed. A mixture of tapioca by-product (60%) and tofu waste (40%) resulted in the highest β-carotene content (295 μg/g) [95]. *N. intermedia* was also suggested to produce high concentrations of carotenoids when the solid waste from tofu production was used as substrate [37]. Similarly, when *Neurospora* spp. were grown on 80% sago waste and 20% tofu waste, 246 μg/g of β-carotene was produced [95]. Studies on different waste sources as substrates for fermentation-derived pigment production by *Neurospora* spp. are interesting from an economical and environmental point of view and thus need to be investigated further.

Neurospora intermedia is under evaluation [1] as a second biocatalyst in the industrial production of bioethanol for yielding ethanol and fungal biomass

from fermentation left-overs [120]. The production of pigment-containing biomass would open the possibility for their use in salmon or chicken feed, which would be in line with the studies aimed to improve the nutrient content in poultry feed [95]. The pigment-containing biomass could potentially be used for human consumption as well. The process has been scaled up to 80 m^3 and further expansion is expected in the near future. Therefore, the successful control of pigment production by *N. intermedia* could reinforce the biorefinery character of the 1st-generation ethanol plants via the production of value-added products. Figure 7 shows an example of how pigment production can be integrated into an already existing biorefinery plant by using the residues as substrates.

Future prospects

Due to increased health and environmental concerns along with tougher regulations regarding the use of synthetic pigments, intensive research is being carried out to find sources of natural-origin pigments. The use of filamentous fungi as an industrial source of biomass and value-added products such as organic acids, enzymes and pigments is already a reality.

Research focusing on *N. intermedia* is also expected to increase due to its potential ability to grow on more economically and environmentally friendly substrates from process leftovers. This also entails the production of value-added products, such as pigments, by process diversification. The potential use of the fungus for this purpose is further strengthened by its pigment production without co-production of any mycotoxins. Therefore, the food-grade ability of the fungus will tentatively aid in public acceptance regarding the use of its pigments for food and feed applications.

Over the years, several studies have focused on factors that stimulate pigment production in filamentous fungi. However, these studies have been superficial and have only briefly considered the optimization of pigment production at a larger scale and considered the circumstances that lead to pigment production as well as the potential regulation towards different color dyes. Moreover, the correlations between the factors are not entirely clear either. Hence, the substrate, cultivation conditions and bioreactor design need to be developed and optimized in order to control the process towards pigments and other desired products. From an industrial point of view, it would be interesting to investigate if pigment optimization is correlated with other valuable features, such as favorable morphological structures (pellets) and/or production of other metabolites. To date, the use of *N. intermedia* for pigment production on an industrial scale has not yet been explored.

As the research goes on, knowledge gaps are filled, hypotheses are proven and production processes are verified, it remains to be seen if pigment production using

Fig. 7 Schematic process scheme of the main starch-based bioethanol process stages leading to the production of ethanol, DDGS and CO$_2$. Adapted from Ferreira et al. [120], and suggested production of pigments represented in *dashed lines* using filamentous fungi
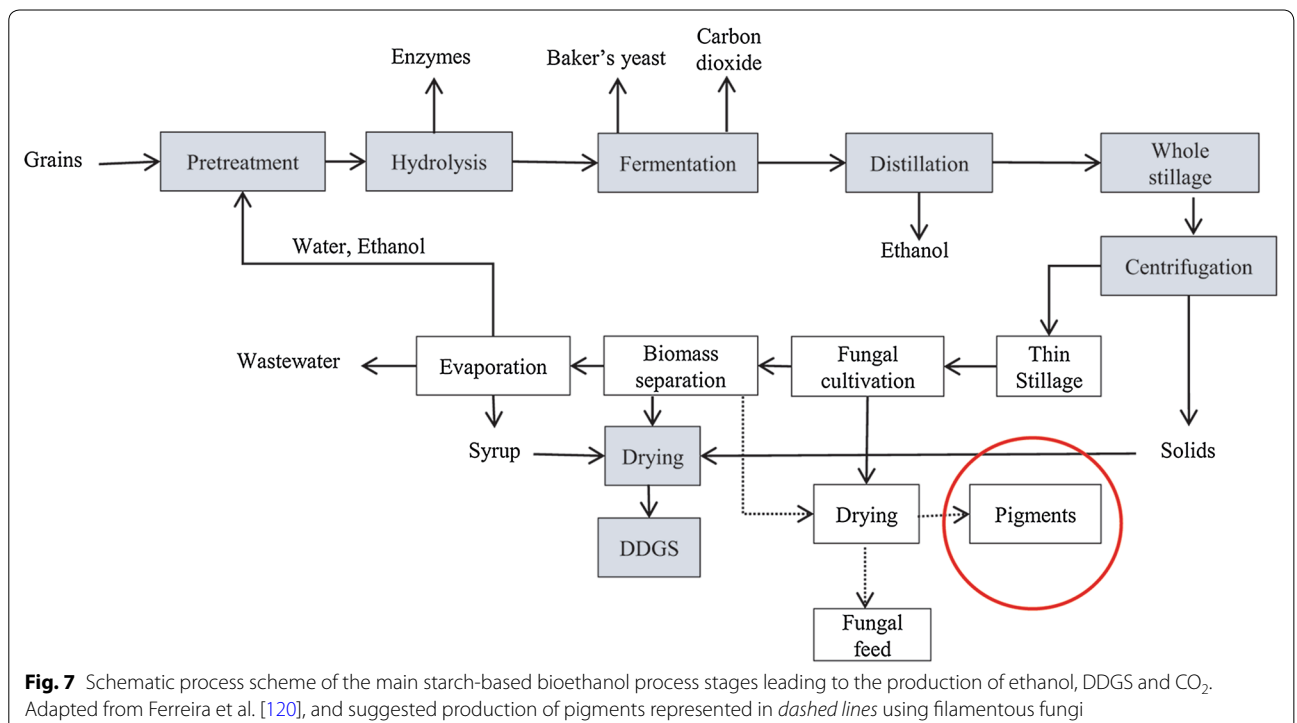
this fungus will be able to compete with other processes where plants or other organisms, such as algae, are used. Therefore, optimization but also comprehensive characterization of pigments produced by *N. intermedia* will naturally play a crucial role in this regard.

Conclusions

Natural-origin pigments, such as carotenoids produced from filamentous fungi, are valuable bioactive compounds with an increasing market demand to replace chemically synthesized pigments. However, there is a need to improve fermentation strategies, control pigment production and decrease production costs in order to compete with synthetic pigments. Carotenoid production in fungi seem to be triggered by environmental stresses, and these conditions can be achieved by regulation of the fermentation process. *N. intermedia* has been foreseen to be a very promising candidate for pigment production due to its color, absence of mycotoxin production, and versatility regarding substrates that it can grow on. The production of pigments adds great potential for biorefineries in which filamentous fungi can be core biocatalysts to convert by-products into several value-added products. By filling the gap in knowledge regarding the production of pigments by *N. intermedia*, the process of pigment-production might be accomplished by research efforts to include the fungus in 1st-generation ethanol plants.

Author details
[1] Swedish Centre for Resource Recovery, University of Borås, 501 90 Borås, Sweden. [2] University of Borås, Allégatan 1, 503 32 Borås, Sweden.

References

1. Ferreira JA, Mahboubi A, Lennartsson PR, Taherzadeh MJ. Waste biorefineries using filamentous ascomycetes fungi: present status and future prospects. Bioresour Technol. 2016;215:334–45.
2. Dufosse L, Fouillaud M, Caro Y, Mapari Sameer AS, Sutthiwong N. Filamentous fungi are large-scale producers of pigments and colorants for the food industry. Curr Opin Biotechnol. 2014;26:56–61.
3. Malik K, Tokkas J, Goyal S. Microbial pigments: a review. Int J Microbial Res Technol. 2012;1:361–5.
4. Caro Y, Anamale L, Fouillaud M, Laurent P, Petit T, Dufosse L. Natural hydroxyanthraquinoid pigments as potent food grade colorants: an overview. Nat Prod Bioprospect. 2012;2:174–93.
5. Mapari Sameer AS, Thrane U, Meyer AS. Fungal polyketide azaphilone pigments as future natural food colorants? Trends Biotechnol. 2010;28:300–7.
6. Gusdinar T, Singgih M, Priatni S, Sukmawati AE, Suciati T. *Neurospora intermedia* N-1 (Encapsulation and the stability of carotenoids from *Neurospora intermedia* N-1). J Manusia dan Lingkungan. 2014;18:206–11.
7. Pagano MC, Dhar PP. Fungal pigments: an overview. In: Gupta VK, Mach RL, Sreenivasaprasad S, editors. Fungal bio-molecules: sources, applications and recent developments. 1st ed. London: Wiley; 2015.
8. Sanchez S, Rutz B, Rodríguez-Sanoja R. Microbial production of carotenoids. In: Brian M, David A, Ioannis G, Linda H, editors. Microbial production of food ingredients, enzymes and nutraceuticals. Burlington: Elsevier Science; 2013. p. 194–223.
9. Singgih M, Andriatna W, Damayanti S, Priatni S. Carotenogenesis study of *Neurospora intermedia* N-1 in liquid. J Chem Pharm Res. 2015;7:842–7.
10. Umeno D, Tobias AV, Arnold FH. Diversifying carotenoid biosynthetic pathways by directed evolution. Microbiol Mol Biol Rev. 2005;69:51–78.
11. Blumenthal CZ. Production of toxic metabolites in *Aspergillus niger*, *Aspergillus oryzae*, and *Trichoderma reesei*: justification of mycotoxin testing in food grade enzyme preparations derived from the three fungi. Regul Toxicol Pharmacol. 2004;39:214–28.
12. Mapari SAS, Nielsen KF, Larsen TO, Frisvad JC, Meyer AS, Thrane U. Exploring fungal biodiversity for the production of water-soluble pigments as potential natural food colorants. Curr Opin Biotechnol. 2005;16:231–8.
13. Torres FAE, Zaccarim BR, de Lencastre Novaes LC, Jozala AF, Dos Santos CA, Teixeira MFS, Santos-Ebinuma VC. Natural colorants from filamentous fungi. Appl Microbiol Biotechnol. 2016;100:2511–21.
14. Avalos J, Corrochano LM. Carotenoid biosynthesis in *Neurospora*. In: Kasbekar DP, McCluskey K, editors. *Neurospora*: genomics and molecular biology. Great Britain: Horizon Scientific Press; 2013. p. 227–41.
15. Dufossé L. Current and potential natural pigments from microorganisms (bacteria, yeasts, fungi, microalgae). In: Carle R, Schweiggert R, editors. Handbook on natural pigments in food and beverages: industrial applications for improving food color. Cambridge: Woodhead Publishing; 2016. p. 337–52.
16. Wiemann P, Willmann A, Straeten M, Kleigrewe K, Beyer M, Humpf H-U, Tudzynski B. Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. Mol Microbiol. 2009;72:931–46.
17. Avalos J, Cerdà-Olmedo E. Carotenoid mutants of *Gibberella fujikuroi*. Curr Genet. 1987;11:505–11.
18. Mahboubi A, Ferreira JA, Taherzadeh MJ, Lennartsson PR. Value-added products from dairy waste using edible fungi. Waste Manag. 2017;59:518–25.
19. Perkins DD, Davis RH. Evidence for safety of *Neurospora* species for academic and commercial uses. Appl Environ Microbiol. 2000;66:5107–9.
20. Pandit A, Maheshwari R. Life-history of *Neurospora intermedia* in a sugar cane field. J Biosci. 1996;21:57–79.
21. Ferreira JA, Lennartsson PR, Taherzadeh MJ. Production of ethanol and biomass from thin stillage using food-grade zygomycetes and ascomycetes filamentous fungi. Energies. 2014;7:3872–85.
22. Aasen Arne J, Liaaen-Jensen SL. Fungal carotenoids Π.* The structure of the carotenoid acid neurosporaxanthin. Acta Chem Scand. 1965;19:1843–53.
23. Luque EM, Gutiérrez G, Navarro-Sampedro L, Olmedo M, Rodríguez-Romero J, Ruger-Herreros C, Tagua VG, Corrochano LM. A relationship between carotenoid accumulation and the distribution of species of the fungus *Neurospora* in Spain. PLoS ONE. 2012;7:e33658.
24. Armstrong GA, Hearst HE. Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. FASEB J. 1996;10:228–37.
25. Arrach N, Schmidhauser T, Avalos J. Mutants of the carotene cyclase

domain of al-2 from *Neurospora crassa*. Mol Genet Genomics. 2002;266:914–21.

26. Vendruscolo F, Bühler RMM, de Carvalho JC, de Oliveira D, Moritz DE, Schmidell W, Ninow JL. *Monascus*: a reality on the production and application of microbial pigments. Appl Biochem Biotechnol. 2016;178:211–23.

27. Velmurugan P, Lee YH, Venil CK, Lakshmanaperumalsamy P, Chae J-C, Oh B-T. Effect of light on growth, intracellular and extracellular pigment production by five pigment-producing filamentous fungi in synthetic medium. J Biosci Bioeng. 2010;109:346–50.

28. Lin Y-L, Wang T-H, Lee M-H, Su N-W. Biologically active components and nutraceuticals in the *Monascus*-fermented rice: a review. Appl Microbiol Biotechnol. 2008;77:965–73.

29. Babitha S, Soccol CR, Pandey A. Effect of stress on growth, pigment production and morphology of *Monascus* sp. in solid cultures. J Basic Microbiol. 2007;47:118–26.

30. Dikshit R, Tallapragada P. *Monascus purpureus*: a potential source for natural pigment production. J Microbiol Biotechnol Res. 2011;1:164–74.

31. Feng Y, Shao Y, Chen F. *Monascus* pigments. Appl Microbiol Biotechnol. 2012;96:1421–40.

32. Rodríguez-Ortiz R, Limón MC, Avalos J. Regulation of carotenogenesis and secondary metabolism by nitrogen in wild-type *Fusarium fujikuroi* and carotenoid-overproducing mutants. Appl Environ Microbiol. 2009;75:405–13.

33. Giordano W, Domenech CE. Aeration affects acetate destination in *Gibberella fujikuroi*. FEMS Microbiol Lett. 1999;180:111–6.

34. Giordano W, Avalos J, Cerdá-Olmedo E, Domenech CE. Nitrogen availability and production of bikaverin and gibberellins in *Gibberella fujikuroi*. FEMS Microbiol Lett. 1999;173:389–93.

35. Díaz-Sánchez V, Estrada AF, Trautmann D, Al-Babili S, Avalos J. The gene carD encodes the aldehyde dehydrogenase responsible for neurosporaxanthin biosynthesis in *Fusarium fujikuroi*. FEBS J. 2011;278:3164–76.

36. Khiabani MS, Esfahani ZH, Azizi M-H, Sahari MA. Effective factors on stimulate and stability of synthesised carotenoid by *Neurospora intermedia*. Nutr Food Sci. 2011;41:89–95.

37. Priatni S. Review: Potential production of carotenoids from *Neurospora*. Nusantara Biocenter. 2014;6:63–8.

38. Rivera Vélez SM. Guide for carotenoid identification in biological samples. J Nat Prod. 2016;79:1473–84.

39. Thomson ISI, Nuraini S, Suslina AL. Improving the quality of tapioca by product through fermentation by *Neurospora crassa* to produce β carotene rich feed. Pak J Nutr. 2009;8:487–90.

40. Avalos J, Prado-Cabrero A, Estrada AF. Neurosporaxanthin production by *Neurospora* and *Fusarium*. In: José-Luis Barredo, editor. Microbial carotenoids from fungi: methods and protocols, vol. 898. New York: Springer; 2012. p. 263–74.

41. Hailei W, Ping L, Yufeng L, Zhifang R, Gang W. Overproduction of a potential red pigment by a specific self-immobilization biomembrane-surface liquid culture of *Penicillium* novae-zeelandiae. Bioprocess Biosyst Eng. 2012;35:1407–16.

42. Chen G, Wu Z. Production and biological activities of yellow pigments from *Monascus*. World J Microbiol Biotechnol. 2016;32:1–8.

43. Hajjaj H, Klaébé A, Loret MO, Goma G, Blanc PJ, François J. Biosynthetic pathway of citrinin in the filamentous fungus *Monascus ruber* as revealed by ^{13}C nuclear magnetic resonance. Appl Environ Microbiol. 1999;65:311–4.

44. Delgado-Vargas F, Jiménez AR, Paredes-López O. Natural pigments: carotenoids, anthocyanins, and betalains—characteristics, biosynthesis, processing, and stability. Crit Rev Food Sci Nutr. 2000;40:173–289.

45. Schuemann J, Hertweck C. Biosynthesis of fungal polyketides. In: Anke T, Weber D, editors. Physiology and genetics, vol. 15. 1st ed. Berlin: Springer; 2009. p. 331–51.

46. Bhosale P. Environmental and cultural stimulants in the production of carotenoids from microorganisms. Appl Microbiol Biotechnol. 2004;63:351–61.

47. Estrada AF, Maier D, Scherzinger D, Avalos J, Al-Babili S. Novel apocarotenoid intermediates in *Neurospora crassa* mutants imply a new biosynthetic reaction sequence leading to neurosporaxanthin formation. Fungal Genet Biol. 2008;45:1497–505.

48. Kaczor A, Barańska M, Czamara K. Carotenoids: overview of nomenclature, structures, occurrence and functions. In: Kaczor A, Baranska M, editors. Carotenoids: nutrition, analysis and technology. 1st ed. London: Wiley; 2016. p. 1–13.

49. Jacobson DJ, Powell AJ, Dettman JR, Saenz GS, Barton MM, Hiltz MD, Dvorachek WH, Glass NL, Taylor JW, Natvig DO. *Neurospora* in temperate forests of western North America. Mycologia. 2004;96:66–74.

50. Jacobson DJ, Dettman JR, Adams RI, Boesl C, Sultana S, Roenneberg T, Merrow M, Duarte M, Marques I, Ushakova A. New findings of *Neurospora* in Europe and comparisons of diversity in temperate climates on continental scales. Mycologia. 2006;98:550–9.

51. Caro Y, Venkatachalam M, Lebeau J, Fouillaud M, Dufossé L. Pigments and colorants from filamentous fungi. In: Mérillon J-M, Ramawat KG, editors. Fungal metabolites. Berlin: Springer; 2015. p. 499–568.

52. Harding RW, Turner RV. Photoregulation of the carotenoid biosynthetic pathway in albino and white collar mutants of *Neurospora crassa*. Plant Physiol. 1981;68:745–9.

53. Priatni S, Singgih M, Kardono LBS, Gusdinar T. Molecular identification of *Neurospora* sp. N-1 isolated from Indonesian red fermented cake. J Pure Appl Microbiol. 2010;4:527–31.

54. Jin JM, Lee J, Lee Y-W. Characterization of carotenoid biosynthetic genes in the ascomycete *Gibberella zeae*. FEMS Microbiol Lett. 2010;302:197–202.

55. Bauernfeind JC. Carotenoids as colorants and vitamin A precursors. Technological and nutritional applications. London: Academic Press; 1981.

56. Rodriguez-Amaya DB, Kimura M. HarvestPlus handbook for carotenoid analysis. Princeton: Citeseer; 2004.

57. Prado-Cabrero A, Scherzinger D, Avalos J, Al-Babili S. Retinal biosynthesis in fungi: characterization of the carotenoid oxygenase CarX from *Fusarium fujikuroi*. Eukaryot Cell. 2007;6:650–7.

58. Baranskia R, Cazzonellib CI. Carotenoid biosynthesis and regulation in plants. In: Kaczor A, Baranska M, editors. Carotenoids: nutrition, analysis and technology. 1st ed. New York: Wiley; 2016. p. 162–89.

59. Yu J-H, Keller N. Regulation of secondary metabolism in filamentous fungi. Annu Rev Phytopathol. 2005;43:437–58.

60. Estrada AF, Avalos J. The white collar protein WcoA of *Fusarium fujikuroi* is not essential for photocarotenogenesis, but is involved in the regulation of secondary metabolism and conidiation. Fungal Genet Biol. 2008;45:705–18.

61. Avalos J, Estrada AF. Regulation by light in *Fusarium*. Fungal Genet Biol. 2010;47:930–8.

62. Iigusa H, Yoshida Y, Hasunuma K. Oxygen and hydrogen peroxide enhance light-induced carotenoid synthesis in *Neurospora crassa*. FEBS Lett. 2005;579:4012–6.

63. Avalos J, Limón MC. Biological roles of fungal carotenoids. Curr Genet. 2015;61:309–24.

64. Feofilova EP, Tereshina VM, Memorskaya AS, Dul'kin LM, Goncharov NG. Fungal lycopene: the biotechnology of its production and prospects for its application in medicine. Microbiology. 2006;75:629–33.

65. López-Nieto MJ, Costa J, Peiro E, Méndez E, Rodríguez-Sáiz M, de la Fuente JL, Cabri W, Barredo JL. Biotechnological lycopene production by mated fermentation of *Blakeslea trispora*. Appl Microbiol Biotechnol. 2004;66:153–9.

66. Huang W, Li Z, Niu H, Li D, Zhang J. Optimization of operating parameters for supercritical carbon dioxide extraction of lycopene by response surface methodology. J Food Eng. 2008;89:298–302.

67. Currrent EU approved additives and their E Numbers. [https://www.food.gov.uk/science/additives/enumberlist].

68. Ramaprasad EVV, Sasikala Ch, Ramana ChV. Neurosporene is the major carotenoid accumulated by *Rhodobacter viridis* JA737. Biotechnol Lett. 2013;35:1093–7.

69. Meléndez-Martínez AJ, Mapelli-Brahm P, Benítez-González A, Stinco CM. A comprehensive review on the colorless carotenoids phytoene and phytofluene. Arch Biochem Biophys. 2015;572:188–200.

70. Nelis HJ, De Leenheer AP. Microbial sources of carotenoid pigments used in foods and feeds. J Appl Bacteriol. 1991;70:181–91.

71. Schuster A, Schmoll M. Biology and biotechnology of *Trichoderma*. Appl Microbiol Biotechnol. 2010;87:787–99.

72. Michiko S, Takeshi E, Umeo T. Inactivation of *Neurospora crassa* conidia by singlet molecular oxygen generated by a photosensitized reaction. J Bacteriol. 1979;138:293–6.

73. Yang Q, Borkovich KA. Mutational activation of a Gαi causes uncontrolled proliferation of aerial hyphae and increased sensitivity to heat and oxidative stress in *Neurospora crassa*. Genetics. 1999;151:107–17.

74. Michán S, Lledías F, Hansberg W. Asexual development is increased in *Neurospora crassa* cat-3-null mutant strains. Eukaryot Cell. 2003;2:798–808.

75. Linden H. A white collar protein senses blue light. Science. 2002;297:777–8.

76. Yen H-W, Zhang Z. Enhancement of cell growth rate by light irradiation in the cultivation of *Rhodotorula glutinis*. Bioresour Technol. 2011;102:9279–81.

77. Mata-Gómez LC, Montañez JC, Méndez-Zavala A, Aguilar CN. Biotechnological production of carotenoids by yeasts: an overview. Microbial Cell Fact. 2014;13:1.

78. Schroeder WA, Johnson EA. Singlet oxygen and peroxyl radicals regulate carotenoid biosynthesis in *Phaffia rhodozyma*. J Biol Chem. 1995;270:18374–9.

79. Angelova MB, Pashova SB, Spasova BK, Vassilev SV, Slokoska LS. Oxidative stress response of filamentous fungi induced by hydrogen peroxide and paraquat. Mycol Res. 2005;109:150–8.

80. Avalos J, Schrott EL. Photoinduction of carotenoid biosynthesis in *Gibberella fujikuroi*. FEMS Microbiol Lett. 1990;66:295–8.

81. Rau W, Rau-Hund A. Light-dependent carotenoid synthesis. Planta. 1977;136:49–52.

82. Garbayo I, Vilchez C, Nava-Saucedo JE, Barbotin JN. Nitrogen, carbon and light-mediated regulation studies of carotenoid biosynthesis in immobilized mycelia of *Gibberella* fujikuroi. Enzyme Microbial Technol. 2003;33:629–34.

83. Linden H, Ballario P, Giuseppe Macino. Blue light regulation in *Neurospora crassa*. Fungal Genet Biol. 1997;22:141–50.

84. Lewis JA, Escalante-Semerena JC. The FAD-dependent tricarballylate dehydrogenase (TcuA) enzyme of *Salmonella enterica* converts tricarballylate into cis-aconitate. J Bacteriol. 2006;188:5479–86.

85. Kritskiĭ MS, Belozerskaia TA, Sokolovskiĭ VY, Filippovich SY. Photoreceptor apparatus of a fungus *Neurospora crassa*. Mol Biol. 2004;39:602–17.

86. Talora C, Franchi L, Linden H, Ballario P, Giuseppe Macino. Role of a white collar-1–white collar-2 complex in blue-light signal transduction. EMBO J. 1999;18:4961–8.

87. Olmedo M, Ruger-Herreros C, Corrochano LM. Regulation of gene transcription by light in *Neurospora*. In: Kasbekar DP, McCluskey K, editors. *Neurospora*: Genomics and molecular biology. 1st ed. Great Britain: Horizon Scientific Press; 2013. p. 155–66.

88. Muñoz V, Butler WL. Photoreceptor pigment for blue light in *Neurospora crassa*. Plant Physiol. 1975;55:421–6.

89. Moliné M, Libkind D, del Carmen DM, van Broock M. Photoprotective role of carotenoids in yeasts: response to UV-B of pigmented and naturally-occurring albino strains. J Photochem Photobiol B Biol. 2009;95:156–61.

90. Carvalho JC, Pandey A, Babitha S, Soccol CR. Production of *Monascus* biopigments: an overview. Agro Food Ind Hi Technol. 2003;14:37–43.

91. Gunasekaran S, Poorniammal R. Optimization of fermentation conditions for red pigment production from *Penicillium* sp. under submerged cultivation. Afr J Biotechnol. 2008;7:1894–8.

92. Mukherjee G, Singh SK. Purification and characterization of a new red pigment from *Monascus purpureus* in submerged fermentation. Process Biochem. 2011;46:188–92.

93. Babitha S, Soccol CR, Pandey A. Jackfruit seed—a novel substrate for the production of *Monascus* pigments through solid-state fermentation. Food Technol Biotechnol. 2006;44:465–71.

94. Milan C, Ivana M, Vladimíra H, Rapta P, Breierová E. In: Hou CT, Shaw J-F, editors. Biocatalysis and agricultural biotechnology. London: CRC Press; 2009. p. 355–75.

95. Nuraini S, Latif SA. Improving the quality of tapioca by product through fermentation by *Neurospora crassa* to produce $ carotene rich feed. Pak J Nutr. 2009;8:487–90.

96. Jung H, Kim C, Kim K, Shin CS. Color characteristics of *Monascus* pigments derived by fermentation with various amino acids. J Agric Food Chem. 2003;51:1302–6.

97. Mapari Sameer AS, Meyer Anne S, Thrane U, Frisvad JC. Identification of potentially safe promising fungal cell factories for the production of polyketide natural food colorants using chemotaxonomic rationale. Microb Cell Fact. 2009;8:1.

98. Kim SJ, Kim G-J, Park D-H, Ryu Y-W. High-level production of astaxanthin by fed-batch culture of mutant strain *Phaffia rhodozyma* AJ-6-1. J Microbiol Biotechnol. 2003;13:175–81.

99. Gu WL, An GH, Johnson EA. Ethanol increases carotenoid production in *Phaffia rhodozyma*. J Ind Microbiol Biotechnol. 1997;19:114–7.

100. Meinicke RM, Vendruscolo F, Esteves MD, de Oliveira D, Schmidell W, Samohyl RW, Ninow JL. Potential use of glycerol as substrate for the production of red pigments by *Monascus ruber* in submerged fermentation. Biocatal Agric Biotechnol. 2012;1:238–42.

101. Sharmila G, Nidhi B, Muthukumaran C. Sequential statistical optimization of red pigment production by *Monascus purpureus* (MTCC 369) using potato powder. Ind Crops Prod. 2013;44:158–64.

102. Cho YJ, Park JP, Hwang HJ, Kim SW, Choi JW, Yun JW. Production of red pigment by submerged culture of *Paecilomyces sinclairii*. Lett Appl Microbiol. 2002;35:195–202.

103. Patakova P. *Monascus* secondary metabolites: production and biological activity. J Ind Microbiol Biotechnol. 2013;40:169–81.

104. Méndez A, Pérez C, Montañéz JC, Martínez G, Aguilar CN. Red pigment production by *Penicillium purpurogenum* GH2 is influenced by pH and temperature. J Zhejiang Univ Sci B. 2011;12:961–8.

105. Hayman EP, Yokoyama H, Chichester CO, Simpson KL. Carotenoid biosynthesis in *Rhodotorula glutinis*. J Bacteriol. 1974;120:1339–43.

106. Singgih M, Julianti E. Food colorant from microorganisms. In: Liong M-T, editor. Beneficial microorganisms in food and nutraceuticals, vol. 27. Switzerland: Springer; 2015. p. 265–84.

107. Li XL, Cui XH, Han JR. Sclerotial biomass and carotenoid yield of *Penicillium* sp. PT95 under oxidative growth conditions and in the presence of antioxidant ascorbic acid. J Appl Microbiol. 2006;101:725–31.

108. Govind NS, Amin AR, Modi VV. Stimulation of carotenogenesis in *Blakeslea trispora* by cupric ions. Phytochemistry. 1982;21:1043–4.

109. Goodwin TW. Biosynthesis of carotenoids. In: Goodwin TW, editor. The biochemistry of the carotenoids, vol. I. London: Chapman and Hall; 1980. p. 33–76.

110. Cohen RJ. Cyclic AMP levels in *Phycomyces* during a response to light. Nature. 1974;251:144–6.

111. Pall ML. Adenosine 3′,5′-phosphate in fungi. Microbiol Rev. 1981;45:462–80.

112. Murayama T, Uno I, Hamamoto K, Ishikawa T. A cyclic adenosine 3′,5′-monophosphate-dependent protein kinase mutant of *Neurospora crassa*. Arch Microbiol. 1985;142:109–12.

113. García-Martínez J, Ádám AL, Avalos J. Adenylyl cyclase plays a regulatory role in development, stress resistance and secondary metabolism in *Fusarium fujikuroi*. PLoS ONE. 2012;7:e28849.

114. Kritsky MS, Sokolovsky VY, Belozerskaya TA, Chernysheva EK. Relationship between cyclic AMP level and accumulation of carotenoid pigments in *Neurospora crassa*. Arch Microbiol. 1982;133:206–8.

115. Hu Z, Zhang X, Wu Z, Qi H, Wang Z. Perstraction of intracellular pigments by submerged cultivation of *Monascus* in nonionic surfactant micelle aqueous solution. Appl Microbiol Biotechnol. 2012;94:81–9.

116. Choudhari SM, Ananthanarayan L, Singhal RS. Use of metabolic stimulators and inhibitors for enhanced production of β-carotene and lycopene by *Blakeslea trispora* NRRL 2895 and 2896. Bioresour Technol. 2008;99:3166–73.

117. Britton G. Structure and properties of carotenoids in relation to function. FASEB J. 1995;9:1551–8.

118. Reese ET, Anne Maguire. Surfactants as stimulants of enzyme production by microorganisms. Appl Microbiol. 1969;17:242–5.

119. Krzeminski Leo F, Quackenbush FW. Stimulation of carotene synthesis in submerged cultures of *Neurospora crassa* by surface-active agents and ammonium nitrate. Arch Biochem Biophys. 1960;88:64–7.

120. Ferreira JA. Integration of filamentous fungi in ethanol dry-mill biorefinery. Doctorial thesis, University of Borås, Faculty of Textiles, Engineering and Business, Swedish Center for Resource Recovery; 2015.

121. Zalokar M. Variations in the production of carotenoids in *Neurospora*. Arch Biochem Biophys. 1957;70:561–7.

122. Gong M, Bassi A. Carotenoids from microalgae: a review of recent developments. Biotechnol Adv. 2016;34:1396–412.

123. Gross J. Carotenoids. In: Gross J, editor. Pigments in vegetables: chlorophylls and carotenoids. New York: Springer; 2012. p. 75–254.

124. Schiedt K, Bischof S, Glinz E. Metabolism of carotenoids and in vivo racemization of (3S, 3′S)-astaxanthin in the crustacean *Penaeus*. Methods Enzymol. 1993;214:148–68.

125. Socaciu C. Food colorants: chemical and functional properties. Cliy-Napoca: CRC Press; 2007.

126. Park PK, Kim EY, Chu KH. Chemical disruption of yeast cells for the isolation of carotenoid pigments. Sep Purif Technol. 2007;53:148–52.

127. Schulz H. Carotenoid bioavailability from the food matrix: toward efficient extraction procedures. In: Kaczor A, Baranska M, editors. Carotenoids: nutrition, analysis and technology. 1st ed. London: Wiley; 2016. p. 191–211.

128. Lauro GJ, Francis J. Natural food colorants: science and technology. London: Marcel Dekker, CRC Press; 2000.

129. Rivera S, Canela R. Influence of sample processing on the analysis of carotenoids in maize. Molecules. 2012;17:11255–68.

130. Fraser PD, Bramley PM. Methodologies for the analysis of fungal carotenoids. In: Barredo JL, editor. Microbial processes and products, vol. 18. Totowa: Humana Press; 2005. p. 273–82.

131. Khachik F. Analysis of carotenoids in nutritional studies. In: Britton G, Pfander H, Synnøve Liaaen-Jensen HC, editors. Carotenoids, vol. 5. Basel: Birkhäuser Verlag; 2009. p. 7–44.

132. Shah Md MR, Liang Y, Cheng JJ, Daroch M. Astaxanthin-producing green microalga *Haematococcus pluvialis*: from single cell to high value commercial products. Front Plant ci. 2016;7:1–28.

133. Mohamed MS, Mohamad R, Manan MA, Ariff AB. Enhancement of red pigment production by *Monascus purpureus* FTC 5391 through retrofitting of helical ribbon impeller in stirred-tank fermenter. Food Bioprocess Technol. 2012;5:80–91.

134. Velmurugan P, Lee YH, Nanthakumar K, Kamala-Kannan S, Dufossé L, Mapari SAS, Oh B-T. Water-soluble red pigments from Isaria farinosa and structural characterization of the main colored component. J Basic Microbiol. 2010;50:581–90.

135. Mortensen A. Carotenoids and other pigments as natural colorants. Pure Appl Chem. 2006;78:1477–91.

136. Sandmann G, Takaichi S, Fraser PD. C 35-apocarotenoids in the yellow mutant *Neurospora crassa* YLO. Phytochemistry. 2008;69:2886–90.

137. Saini RK, Nile SH, Park SW. Carotenoids from fruits and vegetables: chemistry, analysis, occurrence, bioavailability and biological activities. Food Res Int. 2015;76:735–50.

138. Jones JD, Hohn TM, Leathers TD. Genetically modified strains of *Fusarium sporotrichioides* for production of lycopene and β-carotene. San Diego: Society of Industrial Microbiology Annual meeting; 2004. p. 91.

139. Prado-Cabrero A, Schaub P, Díaz-Sánchez V, Estrada AF, Al-Babili S, Avalos J. Deviation of the neurosporaxanthin pathway towards β-carotene biosynthesis in Fusarium fujikuroi by a point mutation in the phytoene desaturase gene. FEBS J. 2009;276:4582–97.

140. Fuska J, Proksa B, Fuskova A. New potential cytotoxic and antitumor substances I. In vitro effect of bikaverin and its derivatives on cells of certain tumors. Neoplasma. 1974;22:335–8.

141. Balan J, Fuska J, Kuhr I, Kuhrova V. Bikaverin, an antibiotic from *Gibberella fujikuroi*, effective against *Leishmania brasiliensis*. Folia Microbiol. 1970;15:479–84.

142. Son SW, Kim HY, Choi GJ, Lim HK, Jang KS, Lee SO, Lee S, Sung ND, Kim J-C. Bikaverin and fusaric acid from *Fusarium oxysporum* show antioomycete activity against *Phytophthora infestans*. J Appl Microbiol. 2008;104:692–8.

143. Shi Y-C, Pan T-M. Beneficial effects of *Monascus purpureus* NTU 568-fermented products: a review. Appl Microbiol Biotechnol. 2011;90:1207–17.

144. Dufossé L. Microbial production of food grade pigments. Food Technol Biotechnol. 2006;44:313–23.

145. Chen W, He Y, Zhou Y, Shao Y, Feng Y, Mu L, Chen F. Edible filamentous fungi from the species *Monascus*: early traditional fermentations, modern molecular biology, and future genomics. Compr Rev Food Sci Food Saf. 2015;14:555–67.

146. Ciriminna R, Fidalgo A, Meneguzzo F, Meneguzzo F, Ilharco LM, Pagliaro M. Lycopene: emerging production methods and applications of a valued carotenoid. ACS Sustain Chem Eng. 2016;4:643–50.

147. Naviglio D, Pizzolongo F, Ferrara L, Naviglio B, Santini A. Extraction of pure lycopene from industrial tomato waste in water using the extractor Naviglio. Afr J Food Sci. 2008;2:037–44.

148. Loto I, Gutiérrez MS, Barahona S, Sepúlveda D, Martínez-Moya P, Baeza M, Cifuentes V, Alcaíno J. Enhancement of carotenoid production by disrupting the C22-sterol desaturase gene (CYP61) in *Xanthophyllomyces dendrorhous*. BMC Microbiol. 2012;12:235.

149. Martín JF, Gudiña E, Barredo JL. Conversion of beta-carotene into astaxanthin: two separate enzymes or a biofunctional hydroxylase-ketolase protein. Microb Cell Fact. 2008;7:1–10.

150. Dannert CS, Umeno D, Arnold FH. Molecular breeding of carotenoid biosynthetic pathway. Nat Biotechnol. 2000;18:750–3.

151. Duffosé L. Microbial and microalgal carotenoids as colorants and supplements carotenoids. In: Britton G, Liaaen-Jensen S, Pfander H, editors. Carotenoids, vol. 5. Basel: Birkhäuser Verlag; 2009. p. 83–98.

152. Furr HC, Clark RM. Intestinal absorption and tissue distribution of carotenoids. J Nutr Biochem. 1997;8:364–77.

153. Salter A, Wiseman H, Tucker G. Phytonutrients. 1st ed. London: Blackwell; 2012.

154. Ciegler A, Lagoda AA, Sohns VE, Hall HH, Jackson RW. Beta-carotene production in 20-liter fermentors. Biotechnol Bioeng. 1963;5:109–21.

155. Powell AJ, Jacobson DJ, Salter L, Natvig DO. Variation among natural isolates of *Neurospora* on small spatial scales. Mycologia. 2003;95:809–19.

156. Evans PJ, Wang HY. Pigment production from immobilized *Monascus* sp. utilizing polymeric resin adsorption. Appl Environ Microbiol. 1984;47:1323–6.

Truncation of the transcriptional repressor protein Cre1 in *Trichoderma reesei* Rut-C30 turns it into an activator

Alice Rassinger[1], Agnieszka Gacek-Matthews[2,3], Joseph Strauss[2], Robert L. Mach[1] and Astrid R. Mach-Aigner[1*]

Abstract

Background: The filamentous fungus *Trichoderma reesei* (*T. reesei*) is a natural producer of cellulolytic and xylanolytic enzymes and is therefore industrially used. Many industries require high amounts of enzymes, in particular cellulases. Strain improvement strategies by random mutagenesis yielded the industrial ancestor strain Rut-C30. A key property of Rut-C30 is the partial release from carbon catabolite repression caused by a truncation of the repressor Cre1 (Cre1-96). In the *T. reesei* wild-type strain a full *cre1* deletion leads to pleiotropic effects and strong growth impairment, while the truncated *cre1-96* enhances cellulolytic activity without the effect of growth deficiencies. However, it is still unclear which function Cre1-96 has in Rut-C30.

Results: In this study, we deleted and constitutively expressed *cre1-96* in Rut-C30. We found that the presence of Cre1-96 in Rut-C30 is crucial for its cellulolytic and xylanolytic performance under inducing conditions. In the case of the constitutively expressed Cre1-96, the cellulase activity could further be improved approximately twofold. The deletion of *cre1-96* led to growth deficiencies and morphological abnormalities. An in silico domain prediction revealed that Cre1-96 has all necessary properties that a classic transactivator needs. Consequently, we investigated the cellular localization of Cre1-96 by fluorescence microscopy using an eYFP-tag. Cre1-96 is localized in the fungal nuclei under both, inducing and repressing conditions. Furthermore, chromatin immunoprecipitation revealed an enrichment of Cre1-96 in the upstream regulatory region of the main transactivator of cellulases and xylanases, Xyr1. Interestingly, transcript levels of *cre1-96* show the same patterns as the ones of *xyr1* under inducing conditions.

Conclusions: The findings suggest that the truncation turns Cre1 into an activating regulator, which primarily exerts its role by approaching the upstream regulatory region of *xyr1*. The conversion of repressor proteins to potential activators in other biotechnologically used filamentous fungi can be applied to increase their enzyme production capacities.

Keywords: Carbon catabolite repression, *Trichoderma reesei*, Cre1, Gene regulation, Transcription factor, Cellulases, Xylanases, Chromatin

Background

Cellulose and hemicellulose are the most abundant biopolymers in plants. After the industrial processing of trees, crops and other plants, which are grown for food and other purposes, a lot of cellulosic and hemicellulosic waste accumulates [1]. The quality and composition of this waste can be quite versatile, depending on the branch of industry they originate from. However, they all share a significant, unused carbohydrate content that can be utilized for the production of valuable products [1]. The main challenge for an economic utilization of these waste products is the efficient conversion of cellulose-rich biomass to products such as (ligno)cellulosic ethanol [2]. One main limitation is the extraction of monomeric and dimeric sugars such as cellobiose, D-glucose and D-xylose from cellulose and hemicellulose [3]. The rigidity of the

---
*Correspondence: astrid.mach-aigner@tuwien.ac.at
[1] Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Str. 1a, 1060 Vienna, Austria
Full list of author information is available at the end of the article

structure of cellulose and hemicellulose requires first mechanical and chemical treatment, which demand high temperatures, harsh chemicals and create an ecologically difficult disposable waste stream. Secondly, hydrolysis of cellulose and hemicellulose is done enzymatically. For the hydrolysis in industrial scale, the main bottleneck is an affordable price on bulk amounts of cellulose and hemicellulose degrading enzymes [3]. The filamentous fungus *Trichoderma reesei* (*T. reesei)* is one of the top producers of such enzymes (e.g. cellobiohydrolases (EC 3.2.1.91), endoglucanases (EC 3.2.1.4), endo-β-1,4-xylanases (EC 3.2.1.8), β-xylosidases (EC 3.2.1.37) (reviewed in [4])) in industry. Those enzymes are moderately expressed in the presence of cellulose and the hemicellulose xylan and stronger by the respective degradation products. Surprisingly, lactose also triggers the expression of these enzymes even though it is not present in the natural environment of the fungus. Although the exact induction mechanism is not fully understood, the uptake of lactose by a permease is necessary for the activation of cellulase gene expression [5].

Anyhow, the enzyme formation is limited by carbon catabolite repression (CCR) in the presence of high concentrations of easily metabolizable monomeric carbohydrates, such as D-glucose or D-xylose [6]. The uptake of D-glucose enables the fungus to rapidly gain energy; hence, the degradation of complex biopolymers by the cellulolytic and xylanolytic enzymes is shut down. The CCR mechanism is well conserved amongst various organisms ranging from bacteria to humans. Based on the sequence homologies to CreA from *Aspergillus* species, the Carbon catabolite repressor protein Cre1 (encoded by *cre1*) was described as the regulator of CCR in *T. reesei* during the 1990ies [7]. Cre1 is a C_2H_2 zinc finger protein and binds to a 5′-SYGGRG-3′ motif within upstream regulatory regions (URR) of cellulase and xylanase encoding genes (e.g. *cbh1* [8], *xyn1* [9]). Its regulon also comprises sugar transporters, developmental processes, and parts of the chromatin remodelling machinery such as nucleosome positioning [10, 11]. Most notably, Cre1 acts negatively on the transcription of the main and essential transactivator of cellulolytic and xylanolytic enzyme expression, Xyr1 [12]. Thus, Xyr1 is also a subject to CCR mediated by Cre1 [13]. With regards to the industry-scale production of hydrolytic enzymes, top producing *T. reesei* strains became a necessity. Random mutagenesis yielded the mutant strain Rut-C30, which achieves enzyme yields of 20 g/L [14]. The nowadays used industrial *T. reesei* strains (yielding up to 100 g/L [15]) are based on Rut-C30 and thus share a similar genetic background. Predominately, this includes a truncation of Cre1, which led to partial de-repression from CCR on D-glucose [16]. Nevertheless, with regards

to the wild-type system, we refer in this manuscript to D-glucose as a repressing condition. In 2014, Mello-de-Sousa and colleagues used the *T. reesei* wild-type strain to demonstrate that this truncated Cre1 (Cre1-96) positively influences cellulase expression, while the full deletion of *cre1* leads to strong pleiotropic effects and growth impairment [17]. The enhancement of cellulase expression by Cre1-96 was attributed to a chromatin opening in the URR of cellulase-encoding genes and also of the *xyr1* gene. However, the impact of Cre1-96 was never studied directly in Rut-C30. The exact regulatory mechanism of Cre1-96 and its role as a putative new transcription factors in industrial strains still remain to be elucidated. In this study, we investigated the effects of a *cre1-96* deletion in Rut-C30 on its growth behaviour, the enzymatic activities and the transcriptional profiles of cellulase- and xylanase-encoding genes (*cbh1*, *xyn1*) and of *xyr1*. To determine the subcellular localization of the putative transcription factor, the nuclear import was examined under cellulase inducing and repressing conditions. Moreover, we performed chromatin immunoprecipitation and nuclease digestion to learn which genes are targeted by Cre1-96 and what is its impact on the DNA accessibility within the URR of its target genes. Finally, we constitutively expressed *cre1-96* in Rut-C30 and examined the impact on the cellulolytic activities.

Results

Deletion and constitutive expression of *cre1-96* in *T. reesei* Rut-C30

To identify the function of Cre1-96 in Rut-C30, the encoding gene was deleted from the genome. Therefore, a deletion cassette was integrated by homologous recombination at the *cre1-96* locus, resulting in a gene replacement of *cre1-96* in Rut-C30. Two *cre1-96* deletion strains were identified by diagnostic PCR (Additional file 1: Figure S1). Both deletion strains were used throughout this study and are in the following termed Rut-C30Δ*cre1-96* (1) and (2) in the figures. In the parent strain Rut-C30, the structural gene of *cre1-96* was put under the control of the *tef1* promoter. The homologous integration of this expression cassette at the *cre1-96* locus was again verified by diagnostic PCR and the resulting strain is in the following termed Rut-C30OE*cre1-96* (Additional file 2: Figure S2).

Cre1-96 is required for cellulolytic and xylanolytic performance of Rut-C30

To investigate a possible impact of Cre1-96 on cellulase and xylanase gene expression, the *cre1-96* deletion strain and its parent strain Rut-C30 were grown on plates containing lactose or carboxymethylcellulose (CMC) to resemble cellulase-inducing conditions (Fig. 1). Further,
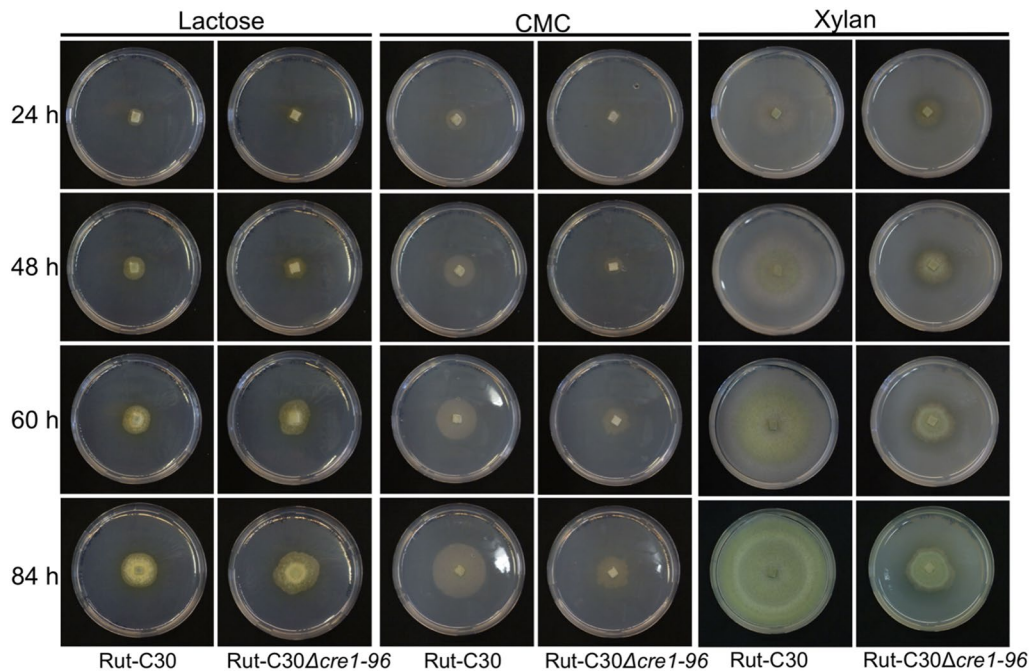
Fig. 1 Growth behaviour of Rut-C30Δ*cre1-96* under cellulase inducing conditions. The *T. reesei* strains Rut-C30 and Rut-C30Δ*cre1-96* were pre-grown on MEX plates and were then transferred in biological duplicates to MA medium plates supplemented with 1% (w/v) lactose, CMC or xylan. Plates were incubated at 30 °C and pictures were taken after 24, 48, 60 and 84 h
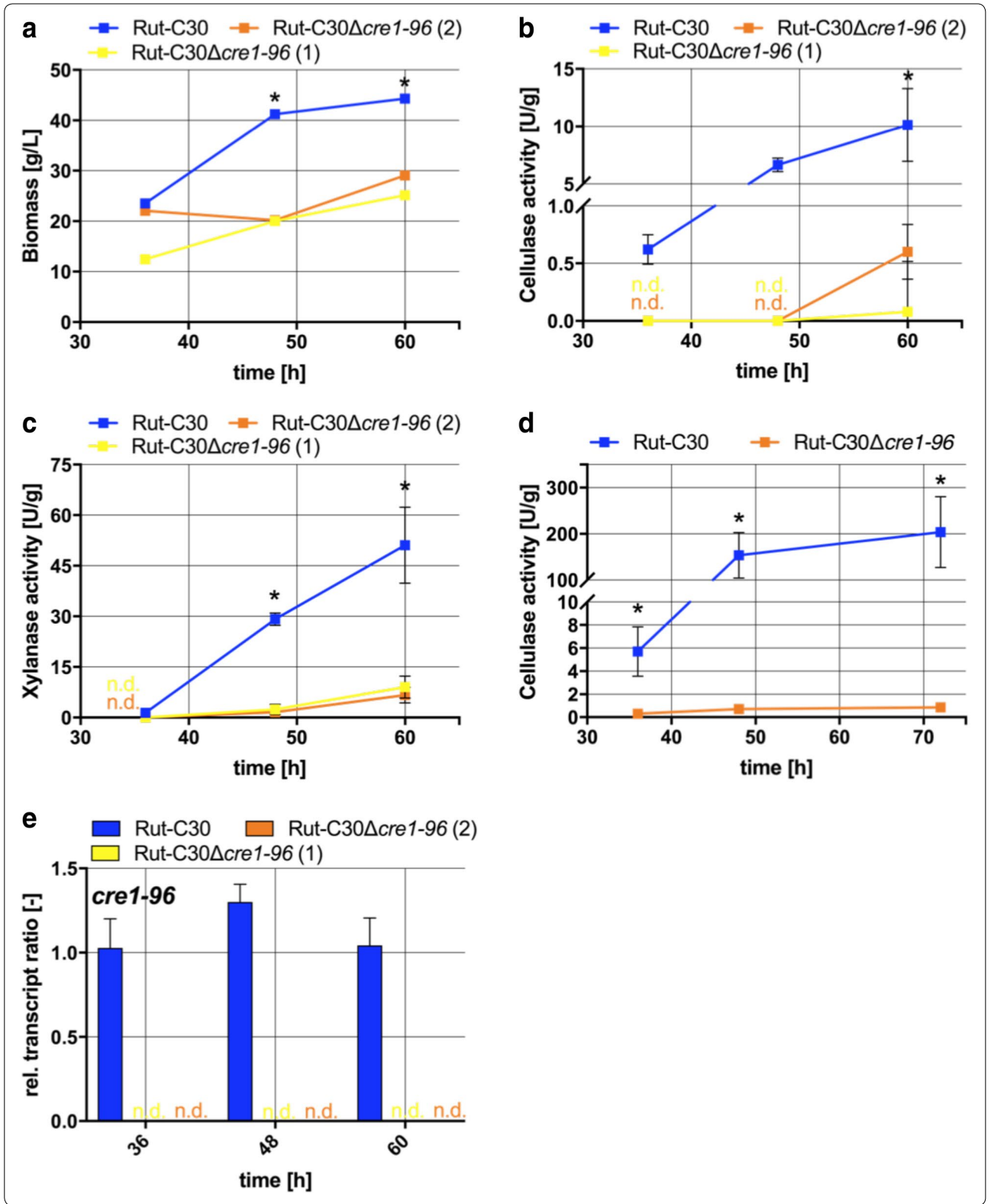
they were grown on xylan for induction of xylanase expression (Fig. 1), on a non-inducing condition (glycerol) and on a repressing condition (D-glucose) (Additional file 3: Figure S3). Photos of the plates were taken after 24, 48, 60 and 84 h of growth. On lactose, no clear differences in growth were obtained between the two tested strains at any time point (Fig. 1). However, the radial colony formation seemed to be abnormal after 60 and 84 h when *cre1-96* was absent (Fig. 1). On CMC and xylan, growth deficiencies were observed in the *cre1-96* deletion strain at all times points in comparison to the parent strain. The colony was clearly reduced in size, while no influence on sporulation was visible (Fig. 1). On glycerol and D-glucose no obvious growth reduction was visible at any time point. However, the spore pigmentation changed in colour intensity (from yellow to light yellow or white) on glycerol and in shade (from green or yellow to brownish) on D-glucose comparing the *cre1-96* deletion strain to Rut-C30 after 60 and 84 h (Additional file 3: Figure S3).

To learn whether the slower growth of the strains carrying the *cre1-96* deletion results from less cellulase and xylanase activity, we tested supernatants from cultivations under inducing conditions (lactose) but also under repressing conditions by enzymatic assays. Supplementary, the abundance of *cre1-96* transcript was determined under inducing conditions. In contrast to the growth

experiments on plates, the biomass formation in the liquid cultures was now also reduced on lactose in the Δ*cre1-96* strains (Fig. 2a). For this reason, the obtained cellulolytic and xylanolytic activities (Fig. 2b, c) were normalized to the biomass. Normalized to the biomass, the *cre1-96* deletion caused a complete loss of cellulolytic and of xylanolytic activity at earlier time points (36 and 48 h) and a strong reduction is observed after 60 h (Fig. 2b, c). Expression of *cre1-96* itself was equally high at all time points under inducing conditions and necessary for the enzyme production (Fig. 2e). Obviously, the presence of *cre1-96* is needed for a good performance in cellulase and xylanase production. Importantly, cellulolytic activities were also lost when D-glucose is used as the carbon source, which is not the case in the parent strain Rut-C30 (Fig. 2d). This reflects that in Rut-C30 the production of cellulases and xylanases is positively influenced by the presence of Cre1-96 regardless if inducing or repressing conditions are prevailing.

Cre1-96 influences the transcript formation of *cbh1*, *xyn1* and *xyr1*

The findings on the reduced enzyme activities prompted us to examine whether Cre1-96 regulates Cre1-target genes on the transcriptional level under inducing conditions. Therefore, we measured the transcript levels of *cbh1*, *xyn1* and *xyr1* on lactose in Rut-C30 and both

Truncation of the transcriptional repressor protein Cre1 in Trichoderma reesei Rut-C30 turns...

203

cre1-96 deletion strains. In the case of *cbh1*, the transcript levels were significantly reduced in the deletion strains compared to the parent strain at all time points (Fig. 3a). In the case of *xyn1*, the transcript levels were also significantly reduced in the deletion strains compared to its parent strain (Fig. 3b). Generally spoken, the *cbh1* and *xyn1* transcriptional profiles matched the measured enzymatic activities in Rut-C30 and both deletion strains. Interestingly, the transcript levels of *xyr1* were also reduced in the deletion strains compared to its parent strains after 36 and 48 h (Fig. 3c), but not any more at the later time point (60 h). To summarize, Cre1-96 has an impact on the formation of *cbh1* and *xyn1* transcript levels, and also on those of the main activator Xyr1 under inducing conditions.

Cre1-96 only indirectly regulates genes involved in the lactose metabolism

The observed differences in the biomass formation in liquid culture on lactose (compare Fig. 2a) gave rise to the possibility that the lactose metabolism could be altered in the *cre1-96* deletion strains. Several genes are necessary for the conversion of lactose to D-galactose and D-glucose. The lactose hydrolysis depends on the extracellular β-galactosidase Bga1 and on the D-xylose reductase Xyl1. A deletion of *xyl1* results in reduced growth on lactose, which is explained by low transcript levels of *bga1* [18]. Here, we investigated the genes coding for the D-xylose reductase (*xyl1*), the β-galactosidase (*bga1*), the galactokinase (*gal1*) and a lactose specific permease (*Tre3405*) [5]. A previous study demonstrated that Xyr1 is involved in the regulation of some lactose metabolism genes by activating *xyl1* and *bga1*, but not *gal1* transcription [19]. As it seems that Cre1-96 has an influence on the *xyr1* transcript formation (compare Fig. 3c), it is very likely that also genes involved in the lactose metabolism are affected by the *cre1-96* deletion. Significantly reduced transcripts of the *xyl1*, *bga1* and *Tre3405* genes were detected in the *cre1-96* deletion strain (Fig. 4a–c), whereas *gal1* transcripts accumulated to similar levels (Fig. 4d). Altogether, this suggests that Cre1-96 acts directly on *xyr1* transcript formation and thereby indirectly influences the transcript levels of *xyl1*, *bga1* and *Tre3405*.

Cre1-96 fulfils the requirements for a transcription factor

Enzymatic measurements and transcript analysis suggested that Cre1-96 exerts a positive effect on cellulase and xylanase gene expression, on the transcript formation of cellulase and xylanase-encoding genes, and most importantly on *xyr1*. To be considered as an activator, some properties need to be fulfilled. First of all, Cre1-96 needs to bind the DNA of its target genes, which is supported by previously reported in vivo footprinting experiments and in vitro protein-DNA binding studies for Cre1-96 and Cre1 [7, 17, 20]. A second essential prerequisite is its localization in the nucleus, at least transiently. In silico domain analysis revealed that Cre1-96 has a putative bipartite nuclear localization signal (NLS) (TVIK – linker – RPYK) located at amino acids (aa) positions 33–63 (Fig. 5a). This bipartite NLS was found with a score of 5 in the case of both proteins, Cre1 and Cre1-96. Scores ranging from 3 to 5 suggest that the protein can be localized both in the nucleus and the cytoplasm. Besides this, alignment of Cre1-96 and Cre1 to homologues from other filamentous fungi revealed further conserved domains or amino acids (Additional file 4: Figure S4). In Cre1-96 a part of a zinc finger domain was identified at 59–79 aa, and a putative transactivation domain (25–37 aa) as well (Fig. 5b). The conserved sequence parts, which are missing in Cre1-96 compared to the full-length Cre1, are a part of the full zinc finger binding domain (87–109 aa), stretches of acidic amino acids (121–129 aa, 243–246 aa, 359–374 aa), two other conserved domains (256–289 aa and 317–325 aa), the nuclear export signal (NES, 304–312 aa), a C-terminal repression domain (317–343 aa) and the phosphorylation site at Ser241 [20] (Fig. 5a). To summarize the in silico analysis, the truncated protein Cre1-96 has lost many potentially important domains but still contains all domains that are essential for a transcription factor, i.e. a DNA-binding domain, NLS
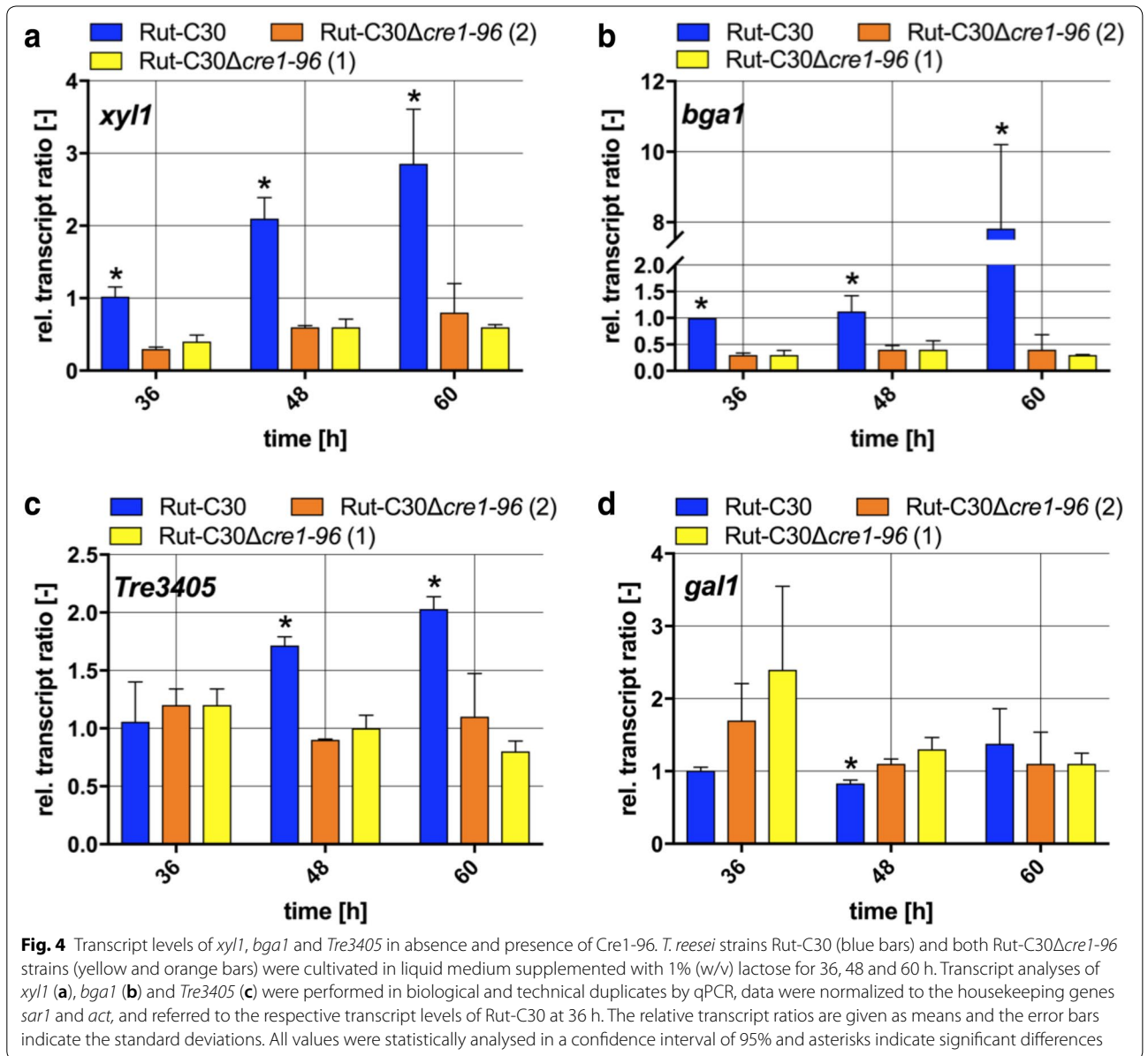
Fig. 3 Transcript levels of *cbh1*, *xyn1* and *xyr1* in absence and presence of Cre1-96. *T. reesei* strains Rut-C30 (blue bars) and both Rut-C30Δ*cre1-96* strains (yellow and orange bars) were cultivated in liquid medium supplemented with 1% (w/v) lactose for 36, 48 and 60 h. Transcript analyses of *cbh1* (**a**), *xyn1* (**b**) and *xyr1* (**c**) were performed in biological and technical duplicates by qPCR, data were normalized to the housekeeping genes *sar1* and *act,* and referred to the respective transcript levels of Rut-C30 at 36 h. The relative transcript ratios are given as means and the error bars indicate the standard deviations. All values were statistically analysed in a confidence interval of 95% and asterisks indicate significant differences

and one N-terminal acidic region potentially functioning as activator domain.

To monitor the localization of Cre1-96 in the fungal hyphae, a strain expressing an eYFP-tagged Cre1-96 was generated and cultivated in liquid medium containing ᴅ-glucose or lactose. It should be noted that we used QM6a for analysis of the nuclear transport of Cre1-96 to exclude any cross-genetic effects resulting from the other mutations present in Rut-C30. Confocal fluorescence microscopy was performed using a droplet of the liquid culture and the visualization of the fungal nuclei was achieved with Hoechst staining. The localization of Cre1-96 was determined by the detection of fluorescence emission of eYFP. Merging of

the eYFP signal and the nuclei fluorescence emissions revealed the presence of Cre1-96 in the nuclei of *T. reesei*. Nuclear localization of Cre1-96 was observed under both, repressing (Fig. 6a) and inducing conditions (Fig. 6b), similar to the full length Cre1 [21].

Cre1-96 targets Cre1-binding sites within the URR of *xyr1*

To learn where the transcription factor Cre1-96 is targeted to, we performed chromatin immunoprecipitation (ChIP) followed by qPCR analyses. For this purpose, the strain expressing eYFP-tagged Cre1-96 was used. As an initial control we tested cellulase activities and biomass formation in tagged and untagged strains to exclude any impact of the eYFP-tag.

Fig. 4 Transcript levels of *xyl1*, *bga1* and *Tre3405* in absence and presence of Cre1-96. *T. reesei* strains Rut-C30 (blue bars) and both Rut-C30Δ*cre1-96* strains (yellow and orange bars) were cultivated in liquid medium supplemented with 1% (w/v) lactose for 36, 48 and 60 h. Transcript analyses of *xyl1* (**a**), *bga1* (**b**) and *Tre3405* (**c**) were performed in biological and technical duplicates by qPCR, data were normalized to the housekeeping genes *sar1* and *act,* and referred to the respective transcript levels of Rut-C30 at 36 h. The relative transcript ratios are given as means and the error bars indicate the standard deviations. All values were statistically analysed in a confidence interval of 95% and asterisks indicate significant differences

The results of these preliminary experiments show no impact of the tag on Cre1-96 function (Additional file 5: Figure S5). As the nuclear localization of Cre1-96 was observed after 16 h cultivation in liquid malt extract (MEX) medium supplemented with ᴅ-glucose, this condition was chosen for the ChIP experiment. An enrichment of Cre1-96 was identified with anti-GFP antibodies (please note that they are able to bind eYFP) and qPCR. Since we had already indications that *xyr1* is a target of Cre1-96, specific primers were chosen for the analysis of Cre1-96 associated DNA within the URR of *xyr1*. The relative amount of Cre1-96 targeted DNA is almost threefold enriched in this target region

compared to the non-target housekeeping gene *sar1* indicating that indeed the truncated Cre1-96 protein might directly activate *xyr1* transcription (Fig. 7).

Chromatin accessibility is only moderately affected by a *cre1-96* deletion

Previous reports demonstrated a role of Cre1-96 in promoting chromatin accessibility [17]. Hence, we analysed the chromatin accessibility in the URR of Cre1-96 target genes (i.e. *xyr1*, *xyn1* and *cbh1)* in the *cre1-96* deletion strain and its parent strain under inducing conditions. Both strains were cultivated in liquid medium on lactose. The fungal mycelium was harvested after 36, 48

Fig. 5 *In silico* domain prediction of Cre1 (**a**) and Cre1-96 (**b**). The putative domains of Cre1 (A) and Cre1-96 (B) were predicted by a number of in silico prediction tools and alignment algorithms as described in the Methods section. Numbers indicate the amino acid (aa) positions and coloured boxes indicate identified domains: blue, C_2H_2 zinc finger; yellow, linker; pink, nuclear localization signal (NLS); black-white striped, transactivation domain (TAD); grey, other conserved region; turquoise, Q (glutamine), DE (aspartic (D) and glutamic acid (E)); green, Q-X_7 Q-X_7-Q; dark blue, nuclear export signal (NES) and violet, repression domain. Phosphorylation site Ser241 according to [20] is marked. The predicted zinc finger domain superfamily are presented as amino acid sequence and the zinc finger binding sites (cysteine (C) or histidine (H)) are coloured in blue. The e-value is given below the amino acid sequence

Fig. 6 Confocal microscopy analysis of the localization of Cre1-96. The *T. reesei* strain QM6a*cre1-96::eyfp* was cultivated in liquid medium supplemented with 1% (w/v) D-glucose (**a**) or 1% (w/v) lactose (**b**). For the visualization of fungal nuclei a Hoechst staining was performed. The following pictures were imaged to localize Cre1-96::eYFP in the fungal hyphae: detection of eYFP (EYFP), detection of the nuclei within the fungal cells (Hoechst), overlay of eYFP and Hoechst emissions (Merge), and brightfield image (DIC). Scales are given in the DIC pictures

and 60 h, followed by chromatin accessibility real-time PCR (CHART-PCR). In the case of *xyr1*, significant differences in the chromatin accessibility were just found after 60 h (Fig. 8a). However, there is no relation between the chromatin status and the transcript level suggesting that chromatin accessibility as measured by our assays is not changing with transcriptional activity. In the case of *cbh1*, significant opening of chromatin in Rut-C30 went along with higher transcript level compared to the deletion strain (Fig. 8b). However, this could be only observed for one time point (i.e. 48 h). Finally, the chromatin accessibility in the *xyn1* URR did differ between Rut-C30

and the *cre1-96* deleted strain at two time points of investigation (Fig. 8c). However, again a transcription-related change in accessibility could not be observed.

Constitutively expressed *cre1-96* enhances cellulase activity

Based on above findings, we have solid indication that Cre1-96 is a necessary activating regulator for cellulase gene expression in Rut-C30. For benefits towards biotechnological applications, we constructed a *T. reesei* strain having a constitutively expressed *cre1-96* under the control of the *tef1* promoter (in the following termed
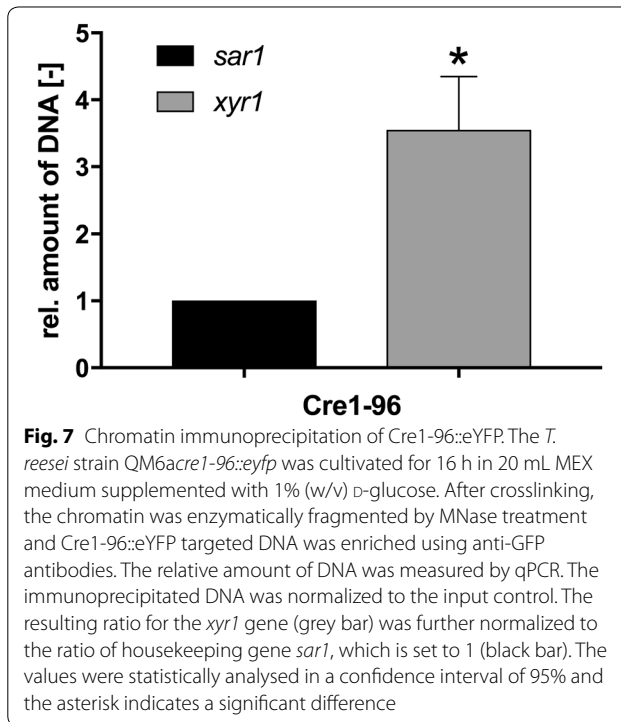
Fig. 7 Chromatin immunoprecipitation of Cre1-96::eYFP. The *T. reesei* strain QM6a*cre1-96::eyfp* was cultivated for 16 h in 20 mL MEX medium supplemented with 1% (w/v) ᴅ-glucose. After crosslinking, the chromatin was enzymatically fragmented by MNase treatment and Cre1-96::eYFP targeted DNA was enriched using anti-GFP antibodies. The relative amount of DNA was measured by qPCR. The immunoprecipitated DNA was normalized to the input control. The resulting ratio for the *xyr1* gene (grey bar) was further normalized to the ratio of housekeeping gene *sar1*, which is set to 1 (black bar). The values were statistically analysed in a confidence interval of 95% and the asterisk indicates a significant difference

Rut-C30OE*cre1-96)*, which was cultivated in parallel with the deletion strain and the parent strain in liquid medium on lactose. Cellulase activities were subsequently measured in the culture supernatants. Rut-C30OE*cre1-96* had a constant increase in the cellulase activity over time, and it was significantly higher than in the other two strains from 48 h of incubation time on (Fig. 9a). After 60 h of incubation, Rut-C30OE*cre1-96* outcompeted Rut-C30 in cellulolytic performance almost twofold. With regard to the growth on cellulase inducing substrates, we observed a similar (24 and 48 h) or a slightly faster growth (60 and 84 h) of Rut-C30OE*cre1-96* compared to Rut-C30 on CMC plates (Fig. 9b). On lactose, no visible differences in colony size were observed amongst all three strains (Fig. 9b). Similar growth was observed under non-inducing (glycerol) and repressing conditions (ᴅ-glucose) (Additional file 6: Figure S6). Importantly, Rut-C30OE*cre1-96* did not show the non-radial growth that was observed for Rut-C30Δ*cre1-96* (Additional files 7 and 8: Tables S1 and S2).

Discussion

The carbon catabolite repressor Cre1 represses transcription of its targets genes by binding their URR. Surprisingly, if truncated, Cre1-96 is still able to bind DNA but converts to a putative activator. Like Cre1/Cre1-96, the transcription factor PacC, which is involved in pH-regulation in *A. nidulans* [22], is a zinc finger protein. PacC is processed by the Pal signalling pathway under alkaline pH and subsequently, moves into the nucleus. In its truncated form it acts as repressor of acidic-expressed genes. Even if its molecular action, namely competition for DNA binding, is different from Cre1-96, in both cases a truncated zinc finger protein acts as transcription factor. The detailed findings of the Cre1 repressor truncation in *T. reesei* are discussed below.

When *cre1* is exchanged for *cre1-96* in the wild-type strain QM6a, higher *cbh1*, *cbh2* and *xyr1* transcript levels were obtained compared to a full deletion of *cre1* [17]. However, Rut-C30 that natively carries Cre1-96 had even higher transcript levels of those genes than the QM6a-CREI$_{96}$ strain. Therefore, we deleted *cre1-96* in Rut-C30 to study the effects on transcript and corresponding enzyme levels.

We observed growth deficiencies, i.e. a slower growth, reduced biomass formation and growth abnormalities, in the Rut-C30 strain lacking *cre1-96*. Under cellulase inducing conditions, we found reduced growth in the *cre1-96* deletion strain compared to its parental strain on CMC plates and lactose liquid cultures (compare Figs. 1, 2a) and significant differences in the transcript ratios of genes involved in the lactose metabolism (compare Fig. 4). This indicates a change either in uptake of degradation products into the cell by transporters or in the enzymatic activity required for the conversion of CMC or lactose into an inducing substance (e.g. transglycosylation by BGLI). We did not observe differences in growth on lactose between the parent and the *cre1-96* deletion strain in the case of cultivation on plates while we did observe differences in the case of cultivation in liquid medium (i.e. determination of the mycelial biomass weight). Interestingly, Cánovas and colleagues found that the biomass accumulation from plates does usually not correlate with the radial growth diameter [23] so these abnormalities cannot be explained by the current model. Particularly, a non-radial growth of fungal hyphae was observed on lactose in the *cre1-96* deletion strain (compare Fig. 1). At this point it has to be mentioned that, a *cre1* deletion in the *T. reesei* strain QM6a leads to strongly impaired growth and morphological changes [8]. Portnoy and colleagues identified several genes involved in hyphal development (e.g. RAS1, PhiA, MedA), which are regulated by Cre1 on ᴅ-glucose [10]. Also in other filamentous fungi, like *Neurospora crassa* (*N. crassa*), CRE-1 seems to influence the hyphal growth and polarity because the enzyme activity of an involved cAMP protein kinase A is dependent on CRE-1 [24]. Altogether, this implies that Cre1-96 might exert additional functions (similar to Cre1), besides its role in cellulase and hemicellulase gene expression. Another aspect worth considering was reported by dos Santos Castro and colleagues. RNA

Fig. 8 Chromatin accessibility under cellulase inducing conditions. *T. reesei* strains Rut-C30 (blue symbols) and Rut-C30Δ*cre1-96* (orange symbols) were cultivated in liquid medium supplemented with 1% (w/v) lactose for 36 (squares), 48 (triangles) and 60 h (dots). CHART-PCR and transcript analysis of *cbh1* (**b**), *xyn1* (**c**) and *xyr1* (**a**) were performed in biological and technical replicates. Both sets of data were normalized to the housekeeping genes *sar1* and *act*. The relative transcript ratios are given as means and plotted on the x-axis and the chromatin accessibility indices (CAI) are plotted on the y-axis. Error bars indicate the standard deviations. All CAI values were statistically analysed in a confidence interval of 95% and asterisks indicate significant differences amongst the chromatin-related data

sequencing analysis of *T. reesei* QM9414 under repressing (D-glucose) and inducing (cellulose, α-sophorose) conditions [25] indicated that several MFS permeases are differentially expressed on D-glucose. Notably, amongst the strongest down-regulated genes in the absence of Cre1 are proteins involved in cellular transport, such as MFS permeases [10]. This indicates that Cre1 and most probably Cre1-96 might also play a role in the sugar uptake in the cell.

With regard to the cellulase and xylanase activity, we observed either a loss or strong reduction in enzymatic activities in the *cre1-96* deletion strain compared to its

parental strain (compare Fig. 2b, c). In Rut-C30, the transcript profile of *xyr1* relates to the profile of *cbh1* (compare Fig. 3), which is in full agreement with previously published results [26]. Most interestingly, the transcription profile of *cre1-96* relates to the profile of *xyr1* (compare Figs. 2d, 3c). Thus, Cre1-96 might have an effect on the regulation of *xyr1* transcription.

The fluorescence microscopy revealed that Cre1-96 is under repressing and inducing conditions present in the nucleus. A carbon source-dependent shuttling of Cre1 between cytosol and nucleus was proposed by Lichius and colleagues [21]. In silico analysis suggests that Cre1
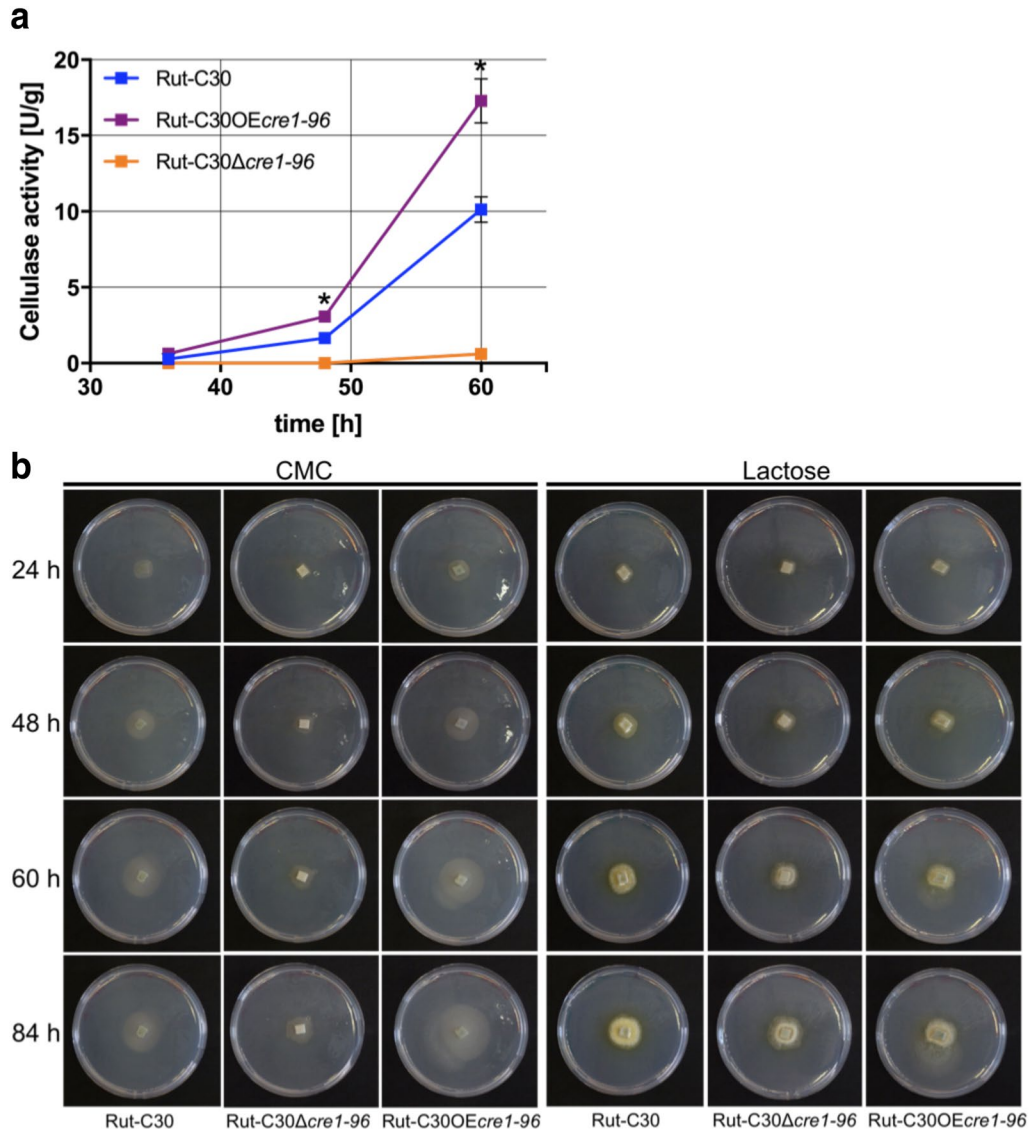
Fig. 9 Cellulase activity in presence of a constitutively expressed Cre1-96. **a** *T. reesei* strains Rut-C30 (blue squares), Rut-C30OE*cre1-96* (purple squares) and Rut-C30Δ*cre1-96* (orange squares) were cultivated in liquid medium supplemented with 1% (w/v) lactose for 36, 48 and 60 h. The endo-cellulolytic activities in the culture supernatants were measured in biological and technical duplicates and normalized to the biomass. The enzymatic activities are given as means and the error bars indicate the standard deviations. Error bars are not shown for standard deviations ≤ 3.5%. The values were statistically analysed by an ordinary one-way ANOVA and a Tukey's posthoc test in a confidence interval of 95%, and asterisks indicate significant differences. **b** For the growth assays, the *T. reesei* strains Rut-C30, Rut-C30OE*cre1-96* and Rut-C30Δ*cre1-96* were pre-grown on MEX plates and were then transferred to MA medium plates supplemented with 1% (w/v) CMC or 1% (w/v) lactose. Plates were incubated at 30 ℃ and pictures were taken after 24, 48, 60 and 84 h

has a nuclear export signal (NES) at amino acids positions 304–312 (LPSLRNLSL, predicted by using [27]). Cre1-96 lacks this putative NES due to its truncation and therefore, could remain inside the nucleus regardless the used carbon source. Besides this, Cre1-96 possesses a putative N-terminal transactivation domain as Cre1 does, but Cre1-96 importantly lacks the C-terminus of Cre1 (compare Fig. 5) that highly likely mediates repression

as it was described for CreA in *Aspergillus nidulans* [28]. Taken together the extended residence in the nucleus, the presence of a putative transactivating domain and the lack of the repression domain, would explain the positive impact of Cre1-96 on the cellulase activity in Rut-C30 compared to the *cre1-96* deletion strain under cellulase inducing conditions.

Notably, Cre1-96 also lacks the previously identified phosphorylation site Ser241 [20]. In the case of Cre1, Ser241 needs to be phosphorylated for an efficient DNA binding under repressing conditions. However, Czifersky and colleagues reported that GST fusion proteins of Cre1 fragments without Ser241 bind in vitro regardless the tested condition. Besides this, previously published in vivo footprinting results also supported the capability of Cre1-96 to bind DNA [17].

With regard to the targeting of Cre1-96, we found that Cre1-96 is enriched on its DNA binding sites in the *xyr1* URR. However, the chromatin accessibility in the *xyr1* URR is not significantly different. Neither for *cbh1* or *xyn1* any cohesive trend could be observed. Noteworthy, an earlier reported nucleosomal mapping of *cbh1* and *cbh2* promoter regions showed no positioned nucleosomes under repressing and inducing conditions in Rut-C30 [29, 30]. This lack of positioned nucleosomes is the likely explanation of the similarity in DNA accessibility observed in our experiments between conditions of expressed and non-expressed genes.

Anyhow, we propose here that a truncation of Cre1 positively influences the transactivator Xyr1 and thus enhances cellulolytic performance and phenotypically converts the carbon catabolite repressor into an activator.

Conclusions

Due to a truncation the Cre1 repressor can turn into an activator as seen in Cre1-96, which now functions to activate cellulase and xylanase expression. Cre1-96 meets all requirements for a transcription factor. It localizes to the nucleus and directly binds to the URR of target genes, in particular of the main transactivator of the mentioned enzymes, Xyr1, and most probably exerts thereby its activation role. Our findings encourage testing this strategy to increase enzymatic performance in other filamentous fungi, which contain functional Cre1 homologues.

Methods
Fungal strains

The *T. reesei* strains QM6aΔ*tmus53* [31], the QM6aΔ*tmus53* bearing an eYFP-tagged Cre1-96 (referred to in the text as QM6a*cre1-96::eyfp*, this study), Rut-C30Δ*tmus53* (referred to in the text as Rut-C30, VTT Finland), two *cre1-96* deletion strains Rut-C30Δ*tmus53*Δ*cre1-96* (referred to in the text as Rut-C30Δ*cre1-96* (1) and (2), this study) and the *cre1-96* constitutively expressing strain Rut-C30Δ*tmus53*OE*cre1-96* (referred to in the text as Rut-C30OE*cre1-96*, this study) were maintained on malt extract (MEX) agar plates containing 0.1% (w/v) peptone from casein at 30 °C. Uridine was added to a final concentration of 5 mM for all Rut-C30 strains. For strain selection hygromycin B was added

to a final concentration of 113 U/mL for the QM6a-related strains and 56.5 U/mL for the Rut-C30-related strains. Homokaryon selection was carried out on MEX/peptone/hygromycin B plates, with uridine if applicable, and 0.1% (w/v) Igepal C-60.

Growth conditions

If not indicated otherwise in a Methods section, for cultivation experiments 10^6 conidia spores per mL were incubated in 100 mL Erlenmeyer flasks on a rotary shaker (180 rpm) at 30 °C for 60 h in 30 mL of MA medium supplemented with 0.1% (w/v) peptone and 1% (w/v) lactose or 1% (w/v) D-glucose as sole carbon source. If not stated otherwise, all strains were cultivated in triplicates and were harvested after 36, 48 and 60 h of cultivation. Fungal mycelia were separated from the supernatant by filtering with Miracloth (EMD Millipore, part of Merck KGaA, Darmstadt, Germany). Mycelia grown on lactose were weighted immediately before shock freezing and wet weight was used as the biomass reference for the enzymatic assays. Frozen mycelia were used for genomic DNA extraction, RNA extraction and for chromatin digestion. Culture supernatants were used for the measurement of enzymatic activities in technical duplicates.

For the comparison of the growth behaviour on plates, the strains Rut-C30, Rut-C30Δ*cre1-96* (1) and (2) and Rut-C30OE*cre1-96* were pre-grown on solid MEX media with 0.1% (w/v) peptone and were transferred in biological duplicates onto MA agar plates supplemented with 1% (w/v) lactose, CMC, xylan, glycerol or D-glucose for 84 h. Due to the same growth behaviour of the duplicates, only one replicate is displayed in the figures. In the case of the deletion strains, Rut-C30Δ*cre1-96* (2) is shown.

Plasmid construction

Escherichia coli strain Top10 (Invitrogen, part of Life Technologies, Paisley, UK) was used for all cloning purposes throughout this study and grown on LB medium at 37 °C. Generation of competent *E. coli* cells and subsequent transformation was performed according to standard protocols using CaCl$_2$. If applicable, ampicillin and hygromycin B were added to final concentrations of 100 mg/mL and 113 U/mL, respectively.

PCRs for all cloning purposes were performed with Pwo DNA Polymerase (peqlab VWR, Radnor, Pennsylvania, USA) or Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Waltham, Massachusetts, USA) according to the manufacturer's instructions. All used primers were purchased from Sigma Aldrich and are listed in Table 1.

For the construction of the *cre1-96* deletion cassette the 5′-flank of *cre1-96* was amplified by PCR using chromosomal DNA of *T. reesei* QM6aΔ*tmus53* (identical

Table 1 Primers used for strain construction in this study

Primer name	Sequence 5′–3′	References
5′cre1_NotI fwd	GCGGCCGCTGGAGGTGACGAGAAGAAAAATTCAGG	This study
5′cre1_XmaI rev	CCCGGGAGTCAAAAAGCAAGTACGCGACGTTG	This study
hph_XmaI_BamHI	CCCGGGTTGGATCCAGGGAGACGAGGTTGTGATGAATAC	This study
hph_SpeI rev	ACTAGTAAGTAGCACCGCTGTCGTCTG	This study
5Pcre1_NotI fwd	GCGGCCGCAGCCAAGACTCAGCATAAAGAGGTTG	This study
5Pcre1_XmaI rev	CCCGGGAGGTACCAAACAAAGCGAGCAAGTAC	This study
ptef_BspEI fwd	TCCGGATGTGTGACAGCTCGCGCAG	This study
ptef_NdeI rev	CATATGTGACGGTTTGTGTGATGTAGCGTG	This study
cre1-96_NdeI fwd	CATATGATGCAACGAGCACAGTCTGCC	This study
Cre1-96_BamHI rev	GGATCCTTAGAAAAAAAAGCAGGTAATGGAGGTGC	This study
cre1-96_BspEI fwd	TCCGGAATGCAACGAGCACAGTCTGCC	This study
cre1-96-TAA_NdeI rev	CATATGGAAAAAAAAGCAGGTAATGGAGGTGC	This study
linker_NheI rev	GCTAGCGCGGGGGGCGCAC	This study
linker_NdeI fwd	CATATGCACAACATGGTCAAGCAGAAGC	This study
YFP_NheI fwd	GCTAGCATGGTCAGCAAGGGCGAGG	This study
YFP_BamHI rev	GGATCCCTTGTACAGCTCGTCCATGCCG	This study

sequence to Rut-C30) as template with the primers 5′cre1_NotI fwd and 5′cre1_XmaI rev. The PCR product was subcloned into pJET1.2 (Thermo Scientific) by blunt end ligation using T4 DNA ligase (Thermo Scientific) yielding pJET1.2-5′-cre1. The 3′-flank of *cre1* and a hygromycin B resistance cassette were amplified by PCR using chromosomal DNA of *T. reesei* QM6a-Cre1$_{96}$ [17] as template with the primers hph_XmaI_BamHI and hph_SpeI rev and was inserted into pJET1.2 by blunt end ligation yielding pJET1.2-hph. The hygromycin B resistance cassette bears the constitutive promoter of the *pki* gene, the hygromycin B structural gene and the terminator of *cbh2* [32]. The subcloned 5′-flank of *cre1* was recovered by NotI/XmaI digestion and was inserted into the NotI/XmaI-digested vector pJET1.2-hph. The resulting plasmid was termed pJET1.2-5hph3cre1. Subsequently, the vector pJET1.2-5hph3cre1 was cut by NotI and SpeI and the cassette was inserted into a NotI/SpeI-digested derivative pMS plasmid yielding pMS*-5hph3cre1. The orientation of the hygromycin B resistance gene and the *pki* promoter were in the opposite orientation as the 5′-flank and the 3′-flank of *cre1*. This was determined by plasmid sequencing (Microsynth, Balgach, Switzerland).

For the constitutive expression of *cre1-96* the promoter of the *tef1* gene was used. For this purpose, the promoter region (1500 bp upstream of ATG) of *tef1* (p*tef1*) was amplified by PCR using chromosomal DNA of *T. reesei* QM6aΔ*tmus53* as template with the primers ptef_BspEI fwd and ptef_NdeI rev. The structural gene *cre1-96* was amplified using the primers cre1-96_NdeI fwd and cre1-96_BamHI rev. Both PCR fragments were subcloned into pJET1.2 (Thermo Scientific), yielding pJET1.2-Ptef and

pJET1-2-cre1-96. Both plasmids were digested by BspEI/NdeI, the p*tef1* fragment was isolated and ligated into the BspEI/NdeI-digested pJET1.2-cre1-96 to yield pJET1.2-Ptefcre1-96. The 5′-flank of *cre1* started at -1500 bp until 2400 bp to avoid a residual background of the native *cre1* promoter. This 5′-flanking region was amplified using the primers 5Pcre1_NotI fwd and 5Pcre_XmaI rev and was subcloned into pJET1.2 by blunt end ligation using T4 DNA ligase (Thermo Scientific) yielding pJET1.2-5Pcre1. The 3′-flank of *cre1* and a hygromycin B resistance cassette was constructed as described for the *cre1-96* deletion cassette. The subcloned 5′-flank of *cre1* was recovered by NotI/XmaI digestion of pJET1.2-5Pcre1 and was inserted into the NotI/XmaI-digested vector pJET1.2-hph. The resulting plasmid was termed pJET1.2-5′cre1-hph. The plasmid pJET1.2-Ptefcre1-96 was BspEI/BamHI-digested, the fragment Ptefcre1-96 isolated and ligated by cohesive ends with a XmaI/BamHI-digested pJET1.2-5′cre1-hph to yield pJET1.2-3Ptefcre1-96. Subsequently, this plasmid was cut by NotI and SpeI and the cassette was inserted into a NotI/SpeI-digested derivative pMS plasmid yielding the final plasmid pMS*-Ptefcre1-96. The correct orientation and sequence of the plasmid were confirmed by sequencing (Microsynth).

For construction of pMS*-*cre1-96::eyfp* the coding sequence of *cre1-96*, a linker and *eyfp* were amplified by PCR using chromosomal DNA from *T. reesei* QM6a-Cre1$_{96}$ and the plasmid pCD-EYFP [33] as templates and the following primers: cre1-96_BspEI fwd and cre1-96-TAA_NdeI rev to amplify *cre1-96* from QM6a-Cre1$_{96}$; linker_NdeI fwd and linker_NheI rev to amplify the linker sequence from the pCD-EYFP

plasmid; YFP_NheI fwd and YFP_BamHI to amplify the coding sequence of *eyfp* from the pCD-EYFP plasmid. Importantly, the fluorescent tag was fused to the C-terminus of the Cre1-96 because this was reported to be necessary for proper recruitment and import in the case of the full length Cre1 [21]. The PCR products were subcloned into pJET1.2, yielding pJET1.2-cre1-96(-TAA), pJET1.2-linker and pJET1.2-YFP. The first two plasmids were digested with BspEI/NdeI, the cre1-96 fragment was isolated and ligated into the BspEI/NdeI-digested recipient vector pJET1.2-linker to generate the plasmid pJET1.2-cre1-96-linker. The insert cre1-96-linker was recovered by BspEI/NheI digestion and cloned into the BspEI/NheI-digested pJET1.2-YFP to yield pJET1.2-cre1-96::eyfp. A BspEI/BamHI double digest of pJET1.2-cre1-96::eyfp recovered the cre1-96::eyfp insert, which was cloned into the XmaI/BamH-digested vector pJET1.2-5hph3cre1 yielding pJET1.2-5cre1-96::eyfp. Finally, the 5cre1-96::eyfp was recovered by NotI/SpeI digestion of pJET1.2-5cre1-96::eyfp and was cloned into a NotI/SpeI-digested derivative pMS plasmid yielding the final plasmid.

Fungal protoplast transformation

Protoplast transformation of *T. reesei* was performed as described by Gruber et al. [34]. For the gene replacement of *cre1,* the plasmid pMS*-*cre1-96::eyfp* was linearized by NotI digestion and transformed into *T. reesei* QM6aΔ*tmus53*. For the deletion of *cre1-96*, the plasmid pMS*-5hph3cre1 was linearized by NotI digestion and transformed into *T. reesei* Rut-C30Δ*tmus53*. For the constitutive expression of *cre1-96* under the control of the promoter of *tef1,* the plasmid pMS*-*ptef*::cre1-96* was transformed into *T. reesei* Rut-C30Δ*tmus53*. Each transformation reaction was added to 40 mL melted, 50 °C warm MEX agar containing 1.2 M D-sorbitol. This mixture was poured into 4 sterile petri dishes, which were incubated at 30 °C for at least 2 h for protoplast regeneration. Appropriate amount of hygromycin B was added to 40 mL melted, 50 °C warm MEX agar containing 1.2 M D-sorbitol and was poured as a 10 mL-overlay on all 4 plates. Transformation plates were further incubated at 30 °C for 2–4 days until colonies were visible. The resulting candidates were subjected to 3 rounds of homokaryon selection by streaking.

Isolation of genomic DNA

Genomic DNA was isolated from approximately 50 mg mycelium in 1 mL CTAB buffer (1.4 M NaCl, 100 mM Tris–HCl pH 8.0, 10 mM EDTA, 2% (w/v) CTAB) by homogenization using a FastPrep(R)-24 cell disrupter

(MP Biomedicals, Santa Ana, California, USA) followed by a phenol/chloroform extraction. RNA was degraded using RNaseA (Thermo Scientific). DNA was precipitated with isopropanol, washed with 70% (w/v) ethanol, and dissolved in distilled H$_2$O.

Diagnostic PCR analysis

100 ng of chromosomal DNA was used as template in a 25-µL-PCR using GoTaq® G2 polymerase (Promega, Madison, Wisconsin, USA) according to manufacturer's instructions. Primer sequences are provided in Additional file 7: Table S1 and Additional file 8: Table S2. For subsequent agarose gel electrophoresis of DNA fragments a GeneRuler 1 kb DNA Ladder (Thermo Scientific) was applied for estimation of fragment size. DNA sequencing was performed at Microsynth.

RNA extraction and reverse transcription

Fungal mycelia were homogenized in 1 mL of peq-GOLDTriFast DNA/RNA/protein purification system reagent (peqlab VWR, Radnor, Pennsylvania, USA) using a FastPrep(R)-24 cell disrupter (MP Biomedicals). RNA was isolated according to the manufacturer's instructions, and the concentration was measured using the NanoDrop 1000 (Thermo Scientific). Reverse transcription of the isolated mRNA was carried out using the RevertAidTM H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) according to the manufacturer's instructions.

Transcript analysis

Quantitative PCR (qPCR) was performed in a Rotor-Gene Q system (Qiagen, Hilden, Germany). Reactions were performed in technical duplicates or triplicates. The amplification mixture (final volume 15 µL) contained 7.5 µL 2×iQ SYBR Green Mix (Bio-Rad, Hercules, California, USA), 100 nM forward and reverse primer, and 2.5 µL cDNA (diluted 1:20). Primer sequences and cycling conditions are provided in Table 2. Data normalization using *sar1* and *act* as reference genes and calculations were performed as previously published [35].

In silico prediction of protein domains

The protein sequences of Cre1 (JGI *Trichoderma reesei* QM6a v2.0 Database, Protein ID 120117) and Cre1-96 (JGI *Trichoderma reesei* Rut C-30 v1.0 Database, Protein ID 23706) were obtained from the respective genome databases [36, 37]. The identification of the DNA binding domain (i.e. C$_2$H$_2$ zinc finger and linker sequence) was achieved by using the NCBI conserved domain search [38]. Multiple sequence alignment of Cre1-96, Cre1 and its homologues of

Table 2 Primer used for qPCR

Primer name	Sequence 5′–3′	References
actfw	TGAGAGCGGTGGTATCCACG	[35]
actrev	GGTACCACCAGACATGACAATGTTG	[35]
sar1fw	TGGATCGTCAACTGGTTCTACGA	[35]
sar1rev	GCATGTGTAGCAACGTGGTCTTT	[35]
xyr1f	CCCATTCGGCGGAGGATCAG	[35]
xyr1r	CGAATTCTATACAATGGGCACATGGG	[35]
taqxyn1 f	CAGCTATTCGCCTTCCAACAC	[13]
taqxyn1 r	CCAAAGTTGATGGGAGCAGAA	[13]
cbh1f	GATGATGACTACGCCAACATGCTG	[12]
cbh1r	ACGGCACCGGGTGTGG	[12]
cre1_a_f	ACCTCCTGAATCCAACGTCGG	[17]
cre1_a_r	TGGGTGCGAATGTGCCTGG	[17]
bga1f	CGTTTGATCCTTTCGGCGGCT	[19]
bga1r	CCAAAGGTCATGTATATGTTGAAGATGGTC	[19]
gal1f	GGAGGCATGGACCAGGC	[19]
gal1r	GACATGCTTGTTGGAGGTGACG	[19]
xorf	CTGTGACTATGGCAACGAAAAGGAG	[19]
xorr	CACAGCTTGGACACGTGAAGAG	[19]
st RT 1	CCGTCTACCGTCTGTTGTGC	[5]
st RT 2	GAAGTAGGAAAGAACCGCATTG	[5]

Aspergillus nidulans (*A. nidulans*, NCBI Accession ID: XP_663799.1), *Aspergillus niger* (*A. niger*, NCBI Accession ID: XP_001399519.1), *Neurospora crassa* (*N. crassa*, NCBI Accession ID: XP_961994.1), *Trichoderma atroviride* (*T. atroviride*, NCBI Accession ID: XP_013941427.1), *Trichoderma virens* (*T. virens*, NCBI Accession ID: XP_013956509.1) and *Saccharomyces cerevisiae* (*S. cerevisiae*, NCBI Accession ID: NP_011480.1) was conducted using Clustal Omega [39] and identified conserved amino acids and protein domains. Prediction of the NLS was achieved by applying the NLS Mapper [40] on Cre1 and Cre1-96. For the in silico identification of the transactivation domain, the Nine Amino Acids Transactivation Domain (9aaTAD) Prediction Tool [41] was used [42]. As search specification, the less stringent pattern was chosen as the most adequate pattern for both proteins.

Confocal microscopy

The localization of the eYFP-labelled Cre1-96 was determined by confocal microscopy and image processing using Fiji [43]. Samples were prepared from liquid cultures. Therefore, 10^6 spores per mL of QM6a*cre1-96::eyfp* were used to inoculate 20 mL of MA medium supplemented with 1% (w/v) D-glucose and 1% (w/v) lactose and incubated at 30 °C and 180 rpm for 16 h. A 10-μL sample was taken and embedded between two glass coverslips (24×60, 24×24). For the nuclear staining, 4 μL

of a 1:10-diluted (distilled water) Hoechst 34580 stain (Thermo Scientific, 5 mg/mL in DMSO) was added before putting the glass coverslip on top of the sample and incubated for 10 min in darkness. Live-cell imaging was performed using a Nikon C1 confocal laser scanning unit sitting on top of a Nikon Eclipse TE2000-E inverted microscope base (Nikon Inc., Melville, New York, USA). An argon ion laser emitting a wavelength of 488 nm excited fluorescent proteins and Hoechst stained nuclei. The emission wavelength was detected with a photomultiplier in a range of 500–530 nm. Laser intensity and illumination time were kept the same for all samples. Pictures were taken as a single picture configuration at a resolution of 1024×1024 pixels.

Enzyme assays

Endo-xylanolytic and endo-cellulolytic activities of cultivation supernatants were measured with Xylazyme AX tablet assay and Azo-CMC-Cellulose assay (both Megazyme International Ireland, Wicklow, Ireland) according to the manufacturer's instructions. For the comparison of cellulolytic activities of Rut-C30OE*cre1-96* and the *cre1-96* deletion, only Rut-C30Δ*cre1-96* (2) was used due to similar results of previous experiments of this study (e.g. transcript analysis).

Chromatin immunoprecipitation (ChIP) and quantitative PCR analysis

T. reesei strain QM6a*cre1-96::eyfp* was grown for 16 h in MEX supplemented with 1% (w/v) D-glucose at 30 °C at 180 rpm. Crosslinking was performed with 1% (w/v) formaldehyde for 15 min at room temperature and gentle shaking every 2–3 min. Quenching was performed by the addition of 125 mM glycine at room temperature for 5 min and gently shaking. Mycelia were filtered by Miracloth, washed with distilled water, dry-pressed between sheets of Whatman paper and frozen in liquid nitrogen. The chromatin shearing and the ChIP protocol were performed according to [44] with the following adaptions. An amount of 100–200 mg of fungal mycelia was grinded in liquid nitrogen and suspended in MNase digestion buffer (50 mM Hepes–KOH pH 7.5, 50 mM NaCl, 1 mM $CaCl_2$, 5 mM $MgCl_2$, 1 mM PMSF, 1× fungal protease inhibitors (Sigma, St. Louis, Missouri, USA)). Chromatin shearing was enzymatically performed by using 0.4 U MNaseI (Sigma,) on 200 μL mycelia aliquots at 37 °C for 13 min. The reaction was stopped by adding 100 μL Lysis Buffer v2 (50 mM Hepes–KOH pH 7.5, 255 mM NaCl, 12 mM EDTA, 2% (w/v) Triton-X100, 0.2% (w/v) Na-deoxcholate, 1 mM PMSF, 1× fungal protease inhibitors (Sigma)). For the precipitation of the protein-antibody complex, an Anti-GFP antibody (ChIP grade; Abcam, Cambridge, UK) and Dynabeads® Protein A magnetic

Table 3 Primer used for ChIP-qPCR

Primer name	Sequence 5′–3′	Reference
sar1 3UTR f	TGACGGGGAGAACATGTGCTC	This study
sar1 3UTR r	ATGCGACTCCCACAAGTGGTG	This study
ChIP_xyr1 upstream f	TACACAAGAGCAATGGCCCTAGC	This study
ChIP_xyr1 upstream r	TGGATGGATGGAGAACGGGATG	This study

Table 4 Primer used for CHART-PCR

Primer name	Sequence 5′–3′	References
epiactinTr_f	CTTCCCTCCTTTCCTCCCCCTCCAC	[17]
epiactinTr_r	GCGACAGGTGCACGTACCCTCCATT	[17]
episar1Tr_f	GTCAGGAAATGCCGCACAAGCAAGA	[17]
episar1Tr_r	TGTGTTTTACCGCCTTGGCCTTTGG	[17]
epixyr1_1Tr_f	CCTTTGGCCATCTACACAAGAGCAA	[45]
epixyr1_1Tr_r	CGCAATTTTTATTGCTGTTCGCTTC	[45]
epicbh1_1Tr_f	AAGGGAAACCACCGATAGCAGTGTC	[46]
epicbh1_1Tr_r	TTTCACTTCACCGGAACAAACAAGC	[46]
epixn1_1Tr_f	GCACTCCAAGGCCTTCTCCTGTACT	[46]
epixyn1_1Tr_r	TAGATTGAACGCCACCCGCAATATC	[46]

beads (Thermo Scientific) were used. The obtained conjugate was washed 3 times with a low salt buffer (0.1% (w/v) SDS, 1% (w/v) Triton X-100, 2 mM EDTA pH 8.0, 20 mM Tris–HCl pH 8.0, 150 mM NaCl), once with a final wash buffer (0.1% (w/v) SDS, 1% (w/v) Triton X-100, 2 mM EDTA pH 8.0, 20 mM Tris–HCl pH 8.0, 500 mM NaCl) and once with TE buffer. Then, samples were eluted in TES buffer (10 mM Tris–HCl pH 8.0, 1 mM EDTA, 1% (w/v) SDS). Protein-bound DNA was treated with Proteinase K (Thermo Scientific) and DNA samples were purified using the MiniElute PCR Purification Kit (Qiagen) according to the manufacturer's protocol. The precipitated DNA was quantified by qPCR performed in iCycler Thermal Cycler (Bio-Rad) and the use of a standard curve. A reaction volume of 25 µL including the following compounds: 2× iQ SYBR® Green Supermix (Bio-Rad), 10 µM primers and 5 µL of immunoprecipitated and input DNA (1:5 diluted in EB) or genomic DNA for the standard curve. The annealing temperature was 60 °C and the primer sequences are provided in Table 3. The qPCR cycling protocol and the adequate amounts of reagents were chosen as recommend in the manufacturer's instructions. All experiments were performed in biological and technical duplicates.

Chromatin accessibility real-time PCR (CHART-PCR) assays
DNaseI digestions of chromatin and subsequent qPCR analyses were carried out as described before [17]. To be noted, only one of both deletion strains (Rut-C30Δ*cre1-96* (2)) was used for this analysis due to similar results from previous experiments of this study (e.g. transcript analysis). The qPCR analyses of the DNaseI-treated samples were performed to measure the relative abundance of DNA of the target regions. PCRs were performed in triplicates in a Rotor-Gene Q system (Qiagen) using the reaction mixture (final volume 20 µL) and the cycling conditions as described before [17]. Primer sequences are provided in Table 4. The amount of intact input DNA of each sample was calculated by comparing the threshold values of the PCR amplification plots with a standard curve generated for each primer set using serial dilutions of genomic, undigested DNA. The chromatin accessibility index (CAI) was defined as: $CAI = (Dc1 + Dc2)/2Ds$, where Ds is the amount of intact

DNA detected for each target region, and Dc1 and Dc2 are the amounts of intact DNA detected for the promoter regions of *sar1* and *act*, respectively, which were used as reference genes for normalization.

Additional files

Additional file 1: Figure S1. Deletion of *cre1-96* in *T. reesei* Rut-C30. (A) Rut-C30 was transformed with the plasmid pMS*-5hph3cre1 that bears the deletion cassette consisting of the hygromycin resistance gene under the *pki* promoter and the terminator of *cbh2* (dark grey arrow, *hph*) to replace the native *cre1-96* gene (light grey arrow, *cre1-96*). (B) Agarose gel electrophoresis of diagnostic PCR was performed. Primer pairs added to the respective PCR are indicated on top of the gel, the strain of which the genomic DNA was used as template is indicated below each lane. Candidate strains (Δ*cre1-96* (1) and (2)) yielded expected fragments with the primer pair 1F and 1R or 3F and 3R, and no fragment in case of primer pair 2F and 2R. Rut-C30 was applied as negative control in the case of the PCR using primer pair 1F and 1R and as positive control in the PCR using primer pair 2F and 2R. Water added to the respective PCR in a no template control PCR (NTC). A DNA ladder (L) was included for estimation of fragment size.

Additional file 2: Figure S2. Constitutive expression of *cre1-96* in *T. reesei* Rut-C30. (A) Rut-C30 was transformed with the plasmid pMS*-ptef::cre1-96 that bears the *tef1* promoter (white bar, p*tef1*), the *cre1-96* gene (light grey arrow, *cre1-96*), and the marker cassette (dark grey bar, *hph*). The latter consists of the hygromycin resistance gene under the *pki* promoter and the terminator of *cbh2*. (B) Agarose gel electrophoresis of diagnostic PCR was performed. Primer pairs added to the respective PCR are indicated on top of the gel, the strain of which the genomic DNA was used as template is indicated below each lane. A candidate strain (OE*cre1-96*) yielded expected fragments with all three primer pairs. Rut-C30 was applied as negative control in case of the PCR using primer pair 1F and 1R as well as 2F and 2R and as a positive control in the PCR using primer pair 3F and 3R. A DNA ladder (L) was included for estimation of fragment size.

Additional file 3: Figure S3. Growth behaviour of Rut-C30Δ*cre1-96* on glycerol and ᴅ-glucose. The *T. reesei* strains Rut-C30 and Rut-C30Δ*cre1-96* (2) were pre-grown on MEX plates and were then transferred to MA medium plates supplemented with 1 % (w/v) glycerol or ᴅ-glucose. Plates were incubated at 30 °C and pictures were taken after 24, 48, 60 and 84 hours.

Additional file 4: Figure S4. Multiple sequence alignment of Cre1 homologues. Multiple sequence alignment of *T. reesei* Cre1, Cre1-96 and Cre1 homologues of *A. nidulans*, *A. niger*, *N. crassa*, *T. atroviride*, *T. virens* and *S. cerevisiae* was conducted using Clustal Omega (http://www.ebi.ac.uk/

Tools/msa/clustalo/). Protein sequences were retrieved from respective genome databases. The alignment revealed conserved amino acids and protein domains based on sequence similarities.

Additional file 5: Figure S5. Cellulase activity and biomass formation of QM6a*cre1-96* and QM6a*cre1-96::eyfp* on ᴅ-glucose. The *T. reesei* strains QM6a*cre1-96* (blue bar, Cre1-96) and QM6a*cre1-96::eyfp* (purple bar, Cre1-96::eYFP) were cultivated in triplicates for 45 hours in MA medium supplemented with 1 % (w/v) ᴅ-glucose. Cellulase activities (A) of the culture supernatants were measured in technical duplicates and the biomass (B) was collected by filtration with miracloth and is depicted as dry weight on the y-axis.

Additional file 6: Figure S6. Growth behaviour of Rut-C30OE*cre1-96* on glycerol and ᴅ-glucose. The *T. reesei* strains Rut-C30, Rut-C30Δ*cre1-96* (2) and Rut-C30OE*cre1-96* were pre-grown on MEX plates and were then transferred to MA medium plates supplemented with 1 % (w/v) glycerol and ᴅ-glucose. Plates were incubated at 30 °C and pictures were taken after 24, 48, 60 and 84 hours.

Additional file 7: Table S1. Primers used for the diagnostic PCR of Rut-C30Δ*cre1-96* (1) and (2).

Additional file 8: Table S2. Primers used for the diagnostic PCR of Rut-C30OE*cre1-96*.

Abbreviations
aa: amino acid; CAI: chromatin accessibility index; CCR: carbon catabolite repression; CHART-PCR: chromatin accessibility real-time PCR; ChIP: chromatin immunoprecipitation; CMC: carboxymethylcellulose; Cre1: carbon catabolite repressor 1; CreA: carbon catabolite repressor A; EB: elution buffer; eYFP: enhanced yellow fluorescent protein; LB: lysogeny broth; MA: Mandels-Andreotti; MEX: malt extract; NES: nuclear export signal; NLS: nuclear localization signal; qPCR: quantitative PCR; TAD: transactivation domain; URR: upstream regulatory regions; Xyr1: xylanase regulator 1.

Authors' contributions
AR constructed the plasmids and the fungal strains, performed the growth experiments, the chromatin analyses (ChIP and CHART-PCR), the confocal microscopy, the in silico domain analysis, the transcript analyses, the enzymatic assays, and drafted this manuscript. AGM contributed to the chromatin analyses (ChIP). JS participated in the conception of the study, analysis of data and revision of the manuscript. RLM participated in the conception of the study. ARMA participated in the conception of the study, supervised the experiments, and revised the manuscript. All authors read and approved the final manuscript.

Author details
[1] Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Str. 1a, 1060 Vienna, Austria. [2] Fungal Genetics and Genomics Lab, Department of Applied Genetics and Cell Biology, BOKU-University of Natural Resources and Life Sciences, Konrad Lorenz Str. 24, 3430 Tulln/Donau, Austria. [3] Institute of Microbiology, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria.

Competing interests
The authors declare that they have no competing interests.

References
1. Singh A, Mishra P. Overview of problems and potential. In: Microbial pentose utilization—current applications in biotechnology. vol. 33; 1995. p. 1–3.
2. Kumar R, Singh S, Singh OV. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. J Ind Microbiol Biotechnol. 2008;35(5):377–91.
3. Singh A, Mishra P. Extraction of pentosans from lignocellulosic materials. In: Microbial pentose utilization—current applications in biotechnology. vol. 33; 1995. p. 71–98.
4. Aro NPT, Penttilä M. Transcriptional regulation of plant cell wall degradation by filamentous fungi. FEMS Microbiol Rev. 2005;29:719–39.
5. Ivanova C, Bââth JA, Seiboth B, Kubicek CP. Systems analysis of lactose metabolism in *Trichoderma reesei* identifies a lactose permease that is essential for cellulase induction. PLoS ONE. 2013;8(5):e62631.
6. Mach-Aigner AR, Pucher ME, Mach RL. D-Xylose as a repressor or inducer of xylanase expression in *Hypocrea jecorina* (*Trichoderma reesei*). Appl Environ Microbiol. 2010;76(6):1770–6.
7. Strauss J, Mach RL, Zeilinger S, Hartler G, Stoffler G, Wolschek M, et al. Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*. FEBS Lett. 1995;376(1–2):103–7.
8. Nakari-Setälä T, Paloheimo M, Kallio J, Vehmaanperä J, Penttilä M, Saloheimo M. Genetic modification of carbon catabolite repression in *Trichoderma reesei* for improved protein production. Appl Environ Microbiol. 2009;75(14):4853–60.
9. Mach RL, Strauss J, Zeilinger S, Schindler M, Kubicek CP. Carbon catabolite repression of xylanase I (*xyn1*) gene expression in *Trichoderma reesei*. Mol Microbiol. 1996;21(6):1273–81.
10. Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, et al. The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. BMC Genomics. 2011;12:269.
11. Ries L, Belshaw NJ, Ilmén M, Penttilä ME, Alapuranen M, Archer DB. The role of CRE1 in nucleosome positioning within the *cbh1* promoter and coding regions of *Trichoderma reesei*. Appl Microbiol Biotechnol. 2014;98(2):749–62.
12. Stricker AR, Grosstessner-Hain K, Würleitner E, Mach RL. Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. Eukaryot Cell. 2006;5(12):2128–37.
13. Mach-Aigner AR, Pucher ME, Steiger MG, Bauer GE, Preis SJ, Mach RL. Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. Appl Environ Microbiol. 2008;74(21):6554–62.
14. Montenecourt BS, Eveleigh DE. Production and characteriation of high yielding cellulase mutants of *Trichoderma reesei*. TAPPI J. 1979;28:101–8.
15. Ward M. Improving secreted enzyme production by *Trichoderma reesei*. In: 9th International workshop on *Trichoderma* and *Gliocladium*: 2006; Vienna.
16. Ilmén M, Thrane C, Penttilä M. The glucose repressor gene *cre1* of *Trichoderma*: isolation and expression of a full-length and a truncated mutant form. Mol Gen Genetics. 1996;251(4):451–60.
17. Mello-de-Sousa TM, Gorsche R, Rassinger A, Pocas-Fonseca MJ, Mach RL, Mach-Aigner AR. A truncated form of the Carbon catabolite repressor 1 increases cellulase production in *Trichoderma reesei*. Biotechnol Biofuels. 2014;7(1):129.
18. Seiboth B, Gamauf C, Pail M, Hartl L, Kubicek CP. The D-xylose reductase of *Hypocrea jecorina* is the major aldose reductase in pentose and D-galactose catabolism and necessary for beta-galactosidase and cellulase induction by lactose. Mol Microbiol. 2007;66(4):890–900.
19. Stricker AR, Steiger MG, Mach RL. Xyr1 receives the lactose induction signal and regulates lactose metabolism in *Hypocrea jecorina*. FEBS Lett. 2007;581(21):3915–20.

20. Cziferszky A, Mach RL, Kubicek CP. Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). J Biol Chem. 2002;277(17):14688–94.

21. Lichius A, Seidl-Seiboth V, Seiboth B, Kubicek CP. Nucleo-cytoplasmic shuttling dynamics of the transcriptional regulators XYR1 and CRE1 under conditions of cellulase and xylanase gene expression in *Trichoderma reesei*. Mol Microbiol. 2014;94(5):1162–78.

22. Tilburn J, Sarkar S, Widdick DA, Espeso EA, Orejas M, Mungroo J, et al. The Aspergillus PacC zinc finger transcription factor mediates regulation of both acid- and alkaline-expressed genes by ambient pH. EMBO J. 1995;14(4):779–90.

23. Cánovas D, Studt L, Marcos AT, Strauss J. High-throughput format for the phenotyping of fungi on solid substrates. Sci Rep. 2017;7:4289.

24. Sun J, Glass LN. Identification of the CRE-1 cellulotytic regulon in *Neurospora crassa*. PLoS ONE. 2011;6(9):e25654.

25. dos Santos Castro L, Pedersoli WR, Antonieto AC, Steindorff AS, Silva-Rocha R, Martinez-Rossi NM, et al. Comparative metabolism of cellulose, sophorose and glucose in *Trichoderma reesei* using high-throughput genomic and proteomic analyses. Biotechnol Biofuels. 2014;7(1):41.

26. Derntl C, Gudynaite-Savitch L, Calixte S, White T, Mach RL, Mach-Aigner AR. Mutation of the Xylanase regulator 1 causes a glucose blind hydrolase expressing phenotype in industrially used *Trichoderma* strains. Biotechnol Biofuels. 2013;6(1):62.

27. NetNES 1.1 Server. http://www.cbs.dtu.dk/services/NetNES/. Accessed 31 May 2017.

28. Ries L, Beattie SR, Espeso EA, Cramer RA, Goldman GH. Diverse Regulation of CreA Carbon Catabolite Repressor in *Aspergillus nidulans*. Genetics. 2016;203(1):335–52.

29. Ries L, Belshaw NJ, Ilmén M, Penttilä ME, Alapuranen M, Archer DB. The role of CRE1 in nucleosome positioning within the *cbh1* promoter and coding regions of *Trichoderma reesei*. Appl Microbiol Biotechnol. 2014;98(2):749–62.

30. Zeilinger S, Schmoll M, Pail M, Mach RL, Kubicek CP. Nucleosome transactions on the *Hypocrea jecorina* (*Trichoderma reesei*) cellulase promoter *cbh2* associated with cellulase induction. Mol Genet Genomics. 2003;270(1):46–55.

31. Steiger MG, Vitikainen M, Uskonen P, Brunner K, Adam G, Pakula T, et al. Transformation system for *Hypocrea jecorina* (*Trichoderma reesei*) that favors homologous integration and employs reusable bidirectionally selectable markers. Appl Environ Microbiol. 2011;77(1):114–21.

32. Mach RL, Schindler M, Kubicek CP. Transformation of *Trichoderma reesei* based on hygromycin B resistance using homologous expression signals. Curr Genet. 1994;25(6):567–70.

33. Derntl C, Kiesenhofer DP, Mach RL, Mach-Aigner AR. Novel strategies for genomic manipulation of *Trichoderma reesei* with the purpose of strain engineering. Appl Environ Microbiol. 2015;81(18):6314–23.

34. Gruber F, Visser J, Kubicek CP, de Graaff LH. Cloning of the *Trichoderma reesei pyrG* gene and its use as a homologous marker for a high-frequency transformation system. Curr Genet. 1990;18(5):447–51.

35. Steiger MG, Mach RL, Mach-Aigner AR. An accurate normalization strategy for RT-qPCR in *Hypocrea jecorina* (*Trichoderma reesei*). J Biotechnol. 2010;145(1):30–7.

36. JGI *Trichoderma reesei* v2.0 Genome Database. http://genome.jgi.doe.gov/Trire2/Trire2.home.html. Accessed 28 April 2014.

37. JGI *Trichoderma reesei* Rut C-30 v1.0 Genome Database. http://genome.jgi.doe.gov/TrireRUTC30_1/TrireRUTC30_1.home.html. Accessed 28 April 2014.

38. NCBI Conserved Domain Search. http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi. Accessed 01 June 2014.

39. Clustal Omega. http://www.ebi.ac.uk/Tools/msa/clustalo/. Accessed 13 June 2017.

40. cNLS Mapper. http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi. Accessed 08 September 2014.

41. Nine Amino Acids Transactivation Domain (9aaTAD) Prediction Tool. http://www.med.muni.cz/9aaTAD/. Accessed 25 May 2017.

42. Piskacek S, Gregor M, Nemethova M, Grabner M, Kovarik P, Piskacek M. Nine-amino-acid transactivation domain: establishment and prediction utilities. Genomics. 2007;89(6):756–68.

43. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676–82.

44. Lando D, Endesfelder U, Berger H, Subramanian L, Dunne PD, McColl J, et al. Quantitative single-molecule microscopy reveals that CENP-ACnp1 deposition occurs during G2 in fission yeast. Open Biol. 2012;2(7):120078.

45. Mello-de-Sousa TM, Rassinger A, Derntl C, Poças-Fonseca MJ, Mach-Aigner AR, Mach RL. The relation between chromatin status, Xyr1 and cellulase expression in *Trichoderma reesei*. Curr Genomics. 2016;17:1–8.

46. Mello-de-Sousa TM, Rassinger A, Pucher ME, dos Santos Castro L, Persinoti GF, Silva-Rocha R, et al. The impact of chromatin remodelling on cellulase expression in *Trichoderma reesei*. BMC Genomics. 2015;16:588.

Permissions

List of Contributors

Elke J. van Nieuwenhuijzen, Jos A. M. P. Houbraken and Robert A. Samson
Applied and Industrial Mycology, CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands

Peter J. Punt
TNO, Microbiology and Systems Biology, Zeist, The Netherlands
Dutch DNA Biotech BV, Zeist, The Netherlands

Guus Roeselers
TNO, Microbiology and Systems Biology, Zeist, The Netherlands
Danone Nutricia Research, Utrecht, The Netherlands

Olaf C. G. Adan
Department of Applied Physics, Section Transport in Permeable Media, University of Technology Eindhoven, Eindhoven, The Netherlands

Lara Hassan, Manfred J. Reppke, Nils Thieme and J. Philipp Benz
HFM, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

Steffen A. Schweizer and Carsten W. Mueller
Chair of Soil Science, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany

Elena Geib and Matthias Brock
Fungal Genetics and Biology, School of Life Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK

Julien Hoarau, Marie Watson, Alain Shum Cheong Sing and Isabelle Grondin
Antenne sud du laboratoire de chimie des Substances Naturelles et des Sciences des Aliments (LCSNSA), EA 2212, Université de la Réunion, UFR des Sciences et Technologies, 15 Avenue René Cassin, CS 92003, 97744 Saint-DenisCedex 9, France

Graziella Chuppa-Tostain
Antenne sud du laboratoire de chimie des Substances Naturelles et des Sciences des Aliments (LCSNSA), EA 2212, Université de la Réunion, UFR des Sciences et Technologies, 15 Avenue René Cassin, CS 92003, 97744 Saint-DenisCedex 9, France

Laboratoire de Physique et Ingénierie Mathématique pour l'Energie et l'Environnement (PIMENT), EA 4518, Université de la Réunion, UFR Sciences de l'Homme et de l'Environnement, 117 rue Général Ailleret, 97430 Le Tampon, France

Yanis Caro and Thomas Petit
Antenne sud du laboratoire de chimie des Substances Naturelles et des Sciences des Aliments (LCSNSA), EA 2212, Université de la Réunion, UFR des Sciences et Technologies, 15 Avenue René Cassin, CS 92003, 97744 Saint-DenisCedex 9, France
Département Hygiène Sécurité Environnement (HSE), Institut Universitaire de Technologie, Université de La Réunion, 40 Avenue de Soweto, 97410 Saint-Pierre, France

Laetitia Adelard
Laboratoire de Physique et Ingénierie Mathématique pour l'Energie et l'Environnement (PIMENT), EA 4518, Université de la Réunion, UFR Sciences de l'Homme et de l'Environnement, 117 rue Général Ailleret, 97430 Le Tampon, France

Isabelle Bourven
Groupement de Recherche Eau Sol Environnement (GRESE), EA 4330, Université de Limoges, Faculté des Sciences et Techniques, 123 Avenue A. Thomas, 87060 Limoges Cedex, France

Jean-Marie Francois
LISBP, UMR INSA-CNRS &/INRA 792, 135 Avenue de Rangueil, 31077 Toulouse Cedex 4, France

Elisabeth Girbal-Neuhauser
Laboratoire de Biotechnologies Agroalimentaire et Environnementale (LBAE), EA 4565, Université de Toulouse III, Institut Universitaire de Technologie, 24 Rue d'Embaquès, 32000 Auch, France

Yanhe Ma
Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Xiqidao 32, Tianjin Airport Economic Area, Tianjin 300308, China

Andreas Noack
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA

Xiaomei Zheng, Ping Zheng and Jibin Sun
Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Xiqidao 32, Tianjin Airport Economic Area, Tianjin 300308, China
Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

Zhang Kun
Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Xiqidao 32, Tianjin Airport Economic Area, Tianjin 300308, China
Key Laboratory of Systems Microbial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China
University of Chinese Academy of Sciences, Beijing 100049, China

Stefan Grätz, Dennis Kerwat, Lutz Adam, David Schirmer, Lennart Richter, Daniel Petras and Roderich D. Süssmuth
Department Biological Chemistry, Institute of Chemistry, Technische Universität Berlin, Straße des 17. Juni 124, 10623 Berlin, Germany

Simon Boecker
Department Biological Chemistry, Institute of Chemistry, Technische Universität Berlin, Straße des 17. Juni 124, 10623 Berlin, Germany
Department Applied and Molecular Microbiology, Institute of Biotechnology, Technische Universität Berlin, Gustav-Meyer-Allee 25, 13355 Berlin, Germany

Tabea Schütze and Vera Meyer
Department Applied and Molecular Microbiology, Institute of Biotechnology, Technische Universität Berlin, Gustav-Meyer-Allee 25, 13355 Berlin, Germany

Jens Christian Nielsen
Department of Biology and Biological Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden

Nancy P. Keller
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Department of Bacteriology, University of Wisconsin, Madison, WI 53706, USA

Philipp Wiemann
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Hexagon Bio, Menlo Park, CA 94025, USA

Alexandra A. Soukup
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Department of Cell and Regenerative Biology, University of Wisconsin, Madison, WI 53705, USA

Jacob S. Folz
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Davis Genome Center – Metabolomics, University of California, 451 Health Science Drive, Davis, CA 95616, USA

Pin-Mei Wang
Department of Medical Microbiology and Immunology, University of Wisconsin, Madison, WI 53706, USA
Ocean College, Zhejiang University, Hangzhou 310058, Zhejiang Province, People's Republic of China

Monika Schmoll
Center for Health and Bioresources, AIT Austrian Institute of Technology GmbH, Konrad Lorenz Straße 24, 3430 Tulln, Austria

Marco Oldiges
Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany
Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Alexander Eck and Matthias Schmidt
Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Wolfgang Wiechert
Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
Computational Systems Biotechnology (AVT.CSB), RWTH Aachen University, 52074 Aachen, Germany

Stefanie Hamer, Anna Joelle Ruff and Ulrich Schwaneberg
Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany
Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Jan Förster and Lars M. Blank
iAMB - Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, 52074 Aachen, Germany
Bioeconomy Science Center (BioSC), c/o Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Christoph Dattenböck, Alberto Alonso Monroy, Wolfgang Hinterdobler and Monika Schmoll
Center for Health and Bioresources, AIT Austrian Institute of Technology GmbH, Konrad Lorenz Straße 24, 3430 Tulln, Austria

Doris Tisch and Andre Schuster
Institute of Chemical Engineering, Research Area Molecular Biotechnology, TU Wien, 1060 Vienna, Austria

Hamed Hosseinpour Tehrani, Svenja Meyer, Lars M. Blank and Nick Wierckx
iAMB – Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany

Elena Geiser
iAMB – Institute of Applied Microbiology, ABBt – Aachen Biology and Biotechnology, RWTH Aachen University, Worringerweg 1, 52074 Aachen, Germany
BioSC, c/o Forschungszentrum Jülich, 52425 Jülich, Germany

Tadashi Takahashi, Masahiro Ogawa, Atsushi Sato and Yasuji Koyama
Noda Institute for Scientific Research, 399 Noda, Noda City, Chiba Pref 278-0037, Japan

Sietske Grijseels, Thomas Ostenfeld Larsen, Jens Christian Frisvad, Kristian Fog Nielsen, Rasmus John Normand Frandsen and Mhairi Workman
Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Jens Nielsen
Department of Biology and Biological Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden
Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

Jorge A. Ferreira, Patrik R. Lennartsson and Mohammad J. Taherzadeh
Swedish Centre for Resource Recovery, University of Borås, 501 90 Borås, Sweden

Rebecca Gmoser
Swedish Centre for Resource Recovery, University of Borås, 501 90 Borås, Sweden
University of Borås, Allégatan 1, 503 32 Borås, Sweden

Alice Rassinger, Robert L. Mach and Astrid R. Mach-Aigner
Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, Gumpendorfer Str. 1a, 1060 Vienna, Austria

Joseph Strauss
Fungal Genetics and Genomics Lab, Department of Applied Genetics and Cell Biology, BOKU-University of Natural Resources and Life Sciences, Konrad Lorenz Str. 24, 3430 Tulln/Donau, Austria

Agnieszka Gacek-Matthews
Fungal Genetics and Genomics Lab, Department of Applied Genetics and Cell Biology, BOKU-University of Natural Resources and Life Sciences, Konrad Lorenz Str. 24, 3430 Tulln/Donau, Austria
Institute of Microbiology, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria

Index