FEBS Letters

FEBS PRESS science publishing by scientists

ARA1 regulates not only L-arabinose but also D-galactose catabolism in *Trichoderma reesei*

Tiziano Benocci[1], Maria Victoria Aguilar-Pontes[1], Roland Sándor Kun[1], Bernhard Seiboth[2], Ronald P. de Vries[1] and Paul Daly[1]

1 Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, The Netherlands
2 Research Area Biochemical Technology, Institute of Chemical, Environmental and Biological Engineering, TU Wien, Vienna, Austria

Correspondence
R. P. de Vries, Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
Fax: +31 302122601
Tel: +31 302122600
E-mail: r.devries@cbs.knaw.nl

(Received 5 October 2017, revised 18 November 2017, accepted 29 November 2017, available online 19 December 2017)

doi:10.1002/1873-3468.12932

Edited by Ulf-Ingo Flügge

***Trichoderma reesei* is used to produce saccharifying enzyme cocktails for biofuels. There is limited understanding of the transcription factors (TFs) that regulate genes involved in release and catabolism of L-arabinose and D-galactose, as the main TF XYR1 is only partially involved. Here, the *T. reesei* ortholog of ARA1 from *Pyricularia oryzae* that regulates L-arabinose releasing and catabolic genes was deleted and characterized by growth profiling and transcriptomics along with a *xyr1* mutant and *xyr1/ara1* double mutant. Our results show that in addition to the L-arabinose-related role, *T. reesei* ARA1 is essential for expression of D-galactose releasing and catabolic genes, while XYR1 is not involved in this process.**

Keywords: D-galactose; L-arabinose; lignocellulose degradation; metabolism; regulation; *Trichoderma reesei*

Trichoderma reesei is one of the most exploited fungi by the biotechnology industry, especially for production of plant biomass-degrading (PDB) enzymes [1]. Plant biomass is composed of three major polysaccharides (cellulose, hemicellulose, and pectin) [2] and while several regulators for cellulose degradation have been characterized, there is only a limited number of characterized *T. reesei* regulators of degradation of the latter two polysaccharides besides xylan [3]. A better understanding of the regulatory system for plant biomass degradation is required for future strain improvement [4]. In particular, there is limited understanding of the regulation of genes involved in L-arabinose and D-galactose release from plant biomass and subsequent catabolism.

L-arabinose is the third most abundant monosaccharide in lignocellulosic plant biomass [5,6], and it is catabolized by filamentous fungi through the pentose catabolic pathway (PCP; Fig. S1) [7,8]. D-galactose is also present in most plant biomass and this monosaccharide also forms with D-glucose the disaccharide lactose, which is the only cheap and soluble inducer of (hemi)-cellulolytic enzymes in *T. reesei* [4]. D-galactose, depending on the fungal species and growth conditions, is catabolized through the oxidoreductive pathway [9,10] and/or the Leloir pathway [11,12] (Fig. S1). These three metabolic pathways are conserved in filamentous fungi, but the enzymes that catabolize the reactions can differ in number and type of activity, depending on the species [13].

Only two L-arabinose-responsive regulators have been characterized in fungi: AraR in Aspergilli [14] and recently ARA1 [15] in the rice-blast fungus

Abbreviations
ABF, α-L-arabinofuranosidase; ADH, alcohol dehydrogenase; AEP, aldose 1-epimerase/mutarotase; AGD, α-glucosidase; AGL, α-galactosidase; AGU, α-glucuronidase; AXE, acetyl xylan esterase; AXL, α-xylosidase; BGA, β-galactosidase; BXL, β-xylosidase; GLN, endo-β-1, 6-galactanase; GUS, β-glucuronidase; LPMO, lytic polysaccharide monooxygenase; PCP, pentose catabolic pathway; PGA, endo-polygalacturonase; XEG, xyloglucanase; XYN, β-xylanase.

Pyricularia oryzae (*Magnaporthe oryzae*) showing differences and similarities in function [3,15–17]. AraR and ARA1 both control the arabinanolytic system and L-arabinose catabolism in their respective species, but in Aspergilli the full control of pentose utilization is shared with the (hemi-)cellulolytic regulator XlnR [16], while this is not the case in *P. oryzae* [15]. In *P. oryzae*, ARA1 fully controls the L-arabinose catabolism and its release from plant biomass, apparently without any requirement for XLR1 [15]. ARA1 and AraR are not sequence orthologs. ARA1, which is present in three classes (Sordariomycetes and Leotiomycetes), has low sequence identity and clusters separately in a phylogenetic analysis to AraR, which is present only in the Eurotiomycetes [3]. Although ARA1 is widely distributed, it has only been characterized in the plant pathogen *P. oryzae*, and has not been characterized in any saprobic fungus, such as *T. reesei*.

For D-galactose, the regulation of its releasing and catabolic enzymes is poorly understood [18], but differs from *Saccharomyces* spp. [19,20]. D-galactose-responsive regulators, so far, have been described only in Aspergilli (GalX [21] and GalR [3,17,22]) and *Saccharomyces* spp., with the Gal1/Gal3-Gal4-Gal80-dependent regulatory circuit that does not exist in filamentous fungi [19,20,23]. *Trichoderma reesei* XYR1 (ortholog of XlnR) is (partially) involved in D-galactose response, based on studies using lactose [18,24–26] instead of D-galactose. In this context XYR1 is mainly controlling (hemi)-cellulolytic gene expression [3,18,24,25,27,28]. There are no transcriptional studies in *T. reesei* of the role of XYR1 or any other regulator in response to D-galactose alone, whereas at an enzymatic level, D-galactose alone induces the production of cellulases at low growth rates [27].

In *T. reesei,* XYR1 is only partially involved in the regulation of L-arabinose utilization, regulating *abf2*, *bxl1* [a β-xylosidase (BXL) with a separate α-L-arabinofuranosidase (ABF) domain], *xyl1*, and slightly *abf1* and *abf3* [29]. Although no other TFs have been identified, the presence of a specific arabinanolytic regulator in *T. reesei* has been suggested [7,29]. Furthermore, three *T. reesei* enzymes involved in the PCP (XYL1, LAD1, and XDH1) are also involved in D-galactose catabolism [13] (Fig. S1). We hypothesize that ARA1 is the main regulator in L-arabinose utilization [3] and would respond to both L-arabinose and D-galactose. In this study, we tested this hypothesis by characterizing a deletion strain of ARA1 through growth profiling and transcriptomics on L-arabinose and D-galactose, and also demonstrate that XYR1 is not involved in the response to D-galactose.

Materials and methods

Strains, media, and growth conditions

Trichoderma reesei QM9414 ATCC 26921 [30] was used as reference strain in all experiments. All *T. reesei* plate cultures were incubated at 28 °C on PDA (Difco, Boom BV, Meppel, the Netherlands) for sporulation, or minimal medium (MM) [31] with 18 $g \cdot L^{-1}$ Select agar (Invitrogen, Carlsbad, CA, USA) during the transformation or growth profiling (in this case Na-citrate was removed). The growth profile was performed on MM with 25 mM D-glucose, L-arabinose, D-xylose, D-galactose, 1% arabinogalactan, 1% arabinan, and 1% arabinoxylan in 9-mm Petri dishes. Duplicate plates were inoculated with 2 μL containing 1×10^3 spores, which were pregerminated overnight in MM with 1% fructose and 0.1% peptone, and incubated in the dark for at least 5 days at 28 °C. Pregermination facilitates reproducible growth on C sources where *T. reesei* spores germinate infrequently or do not germinate. D-Fructose was used for the pregermination as it did not appear to cause a strong CRE1 repression effect in contrast to D-glucose (data not shown). Independent deletion strains for the same gene were generated, confirmed by molecular methods, and tested for growth on a subset of C-sources to confirm the reliability of attributing the observed phenotypes to deleted gene (data not shown). We selected one strain for each deletion to use in further studies and deposited these at the Westerdijk Fungal Biodiversity Institute collection, with strain numbers CBS143330 (Δ*ara1*) and CBS143331 (Δ*xyr1*Δ*ara1*).

A transfer experiment was performed for transcriptomics. 250 mL of complete medium (CM) [32] containing 2% D-fructose in 1-L Erlenmeyer flasks was inoculated with 2.5×10^8 fresh spores, harvested from a PDA plate, and incubated in a rotatory shaker at 28 °C for 20 h at 250 r.p.m. The mycelium was harvested by filtration, washed with liquid Mandels Andreotti medium (MA) [33] (without carbon source) and 2.5 g mycelium (wet weight) was transferred to 250-mL Erlenmeyer flasks containing 50 mL MA with 25 mM D-galactose or L-arabinose, and incubated in a rotatory shaker at 28 °C and 250 r.p.m. After 2 h of incubation, the mycelium was harvested by vacuum filtration, dried between tissue paper, directly frozen in liquid nitrogen, and stored at -45 °C, as previously described [15]. All experiments were performed in duplicates.

Molecular biology methods

The *T. reesei ara1* gene was identified as previously described [3]. The hygromycin B^R cassette was amplified from the plasmid pLH1hph [19] and fused with 1 kb flanking regions up- and downstream of the *ara1* gene by fusion-PCR and purified as described by Klaubauf *et al.* [15]. This *ara1* deletion cassette was used to transform

spores of *T. reesei* QM9414 Δ*tku70* [34] (single mutant) or *T. reesei* QM9414 Δ*xyr1* [35] (double mutant) by electroporation as described by Schuster *et al.* [36], using a Bio-Rad Gene Pulser Electroporator System (Veenendaal, the Netherlands) set at 1.8 kV, 800 Ω and 25 μF. DNA from transformants was screened by PCR for the absence of *ara1* and the correct positioning of the insert as described by Klaubauf *et al.* [15]. The absence of ectopic integrations was confirmed by Southern blot (Fig. S2) using DIG Easy Hyb kit (Roche, Woerden, the Netherlands) and Anti-Digoxigenin-AP, Fab fragments (Roche) with a probe designed to hybridize to part of the hygromycin resistance gene sequence and amplified with the PCR DIG Probe Synthesis Kit (Roche), according to the manufacturer protocols. Primers used for PCR reactions are listed in Table S1. Total RNA was extracted from mycelium, purified, and checked for quality as described previously [15].

RNA sequencing and read mapping

RNA samples (5–41 μg DNase-treated total RNA) were processed by GenomeScan (Leiden, The Netherlands) according to the company's workflow pipeline. The sample preparation was performed according to the protocol 'NEBNext Ultra Directional RNA Library Prep Kit for Illumina'. The size of the library was between 300–500 bp, a concentration of 1.6 pM of DBA was used for sequencing via Illumina NextSeq 500. On average 75 bp reads were constituted, producing ~ 450 MB raw yields for each sample. Raw reads were produced from the original image data by base calling with the Illumina data analysis pipeline RTA v2.4.11 and BCL2DASTQ v2.17. After data filtering, the adaptor was trimmed using TRIMMOMATIC v0.30 [37]. Clean reads were mapped to the genome of *T. reesei* QM6A v.2 (http://genome.jgi.doe.gov/Trire2/Trire2.home.html) using Tophat [38] with a default mismatch rate of 2%. Reads were mapped to the genes using HTSeq-count [39] with option stranded as reverse for counts. In total, 90% of the reads mapped to the genome and 79% mapped to the genes. Additionally, FPKM values were calculated using Cufflinks [40] with default options. The RNA-seq data have been submitted to Gene Expression Omnibus (GEO) [41] with accession number: GSE104606.

RNA-seq data analysis

DESEQ2 v1.16.1 [42] was used to determine which genes were differentially expressed (DE) between pairs of conditions. The parameters used to call a gene DE between conditions were adjusted to P-value < 0.05 and a fold change > 2.5 (DESeq2 log2 fold change > 1.32). Also, genes with FPKM < 15 in the reference strain were considered too lowly expressed and then excluded from the sets of DE genes in the mutants on the same sugar. Transcriptomics analysis focused only on genes encoding CAZy, carbon catabolic enzymes and TFs (Table S2). The CAZy gene set was made by combining the annotations from Hakkinen *et al.* [43] with the annotations from the JGI database (http://genome.jgi.doe.gov/mycocosm/annotations/browser/cazy/summary;ffmgHL?p=Trire2), and then selecting only the PBD CAZymes. *Trichoderma reesei* carbon catabolic genes were identified by orthology with the carbon catabolic genes of *Aspergillus niger* CBS 513.88 using orthoMCL. The parameters for detecting orthoMCL clusters were according to Li *et al.* [44], using inflation factor 1, E value cutoff 1 E–3, percentage match cutoff 60% [45] and the final set was manually curated based on literature [13]. The *T. reesei* transcription factors (TF) were those identified previously by Benocci *et al.* [3]. These genes previously identified as DE were used for the Venn diagrams (Figs 2–4) and the heatmap (Fig. S4). Venn diagrams were generated using VENNDIAGRAM 1.6.17 and GRIDEXTRA 2.2.1 package, the heatmap was generated using HEATMAP.2 function from GPLOTS_3.0.1 package in R statistical language and ENVIRONMENT 3.4.0 [46].

Results and Discussion

Growth of Δ*ara1* mutant reduced on L-arabinose and abolished on D-galactose

Three *T. reesei* deletion mutants (Δ*xyr1*, Δ*ara1*, Δ*xyr1*-Δ*ara1*) were phenotypically screened by growth profiling using various carbon sources including L-arabinose and D-galactose monomers as well as polymers that these monomers compose, and compared to the reference strain QM9414 (Fig. 1). Deletion of *ara1* affected the growth on D-galactose stronger than L-arabinose and arabinan. It reduced growth on L-arabinose and arabinan, resulting in a smaller and thinner colony, and abolished growth on D-galactose, but did not affect growth on D-xylose. The effects on growth on these carbon sources were also analyzed in the Δ*xyr1* mutant. Similar to the *ara1* deletion, deletion of *xyr1* reduced growth on L-arabinose and arabinan, as previously shown by Akel *et al.* [29]. In contrast to the deletion of *ara1*, deletion of *xyr1* severely reduced growth on D-xylose, but did not affect growth on D-galactose. This is similar to what was observed in *P. oryzae*, where deletion of *xlr1* severely reduced growth on D-xylose [47], while ARA1 is required for growth on L-arabinose and arabinan [15]. The effect of an *ara1* deletion on growth on D-galactose in *P. oryzae* cannot be studied because this fungus cannot grow on D-galactose as a single carbon source.

In *T. reesei*, the reduced growth of both Δ*xyr1* and Δ*ara1* on L-arabinose suggested combined regulation and this prompted us to delete *ara1* in the Δ*xyr1* background, especially as this was not done previously

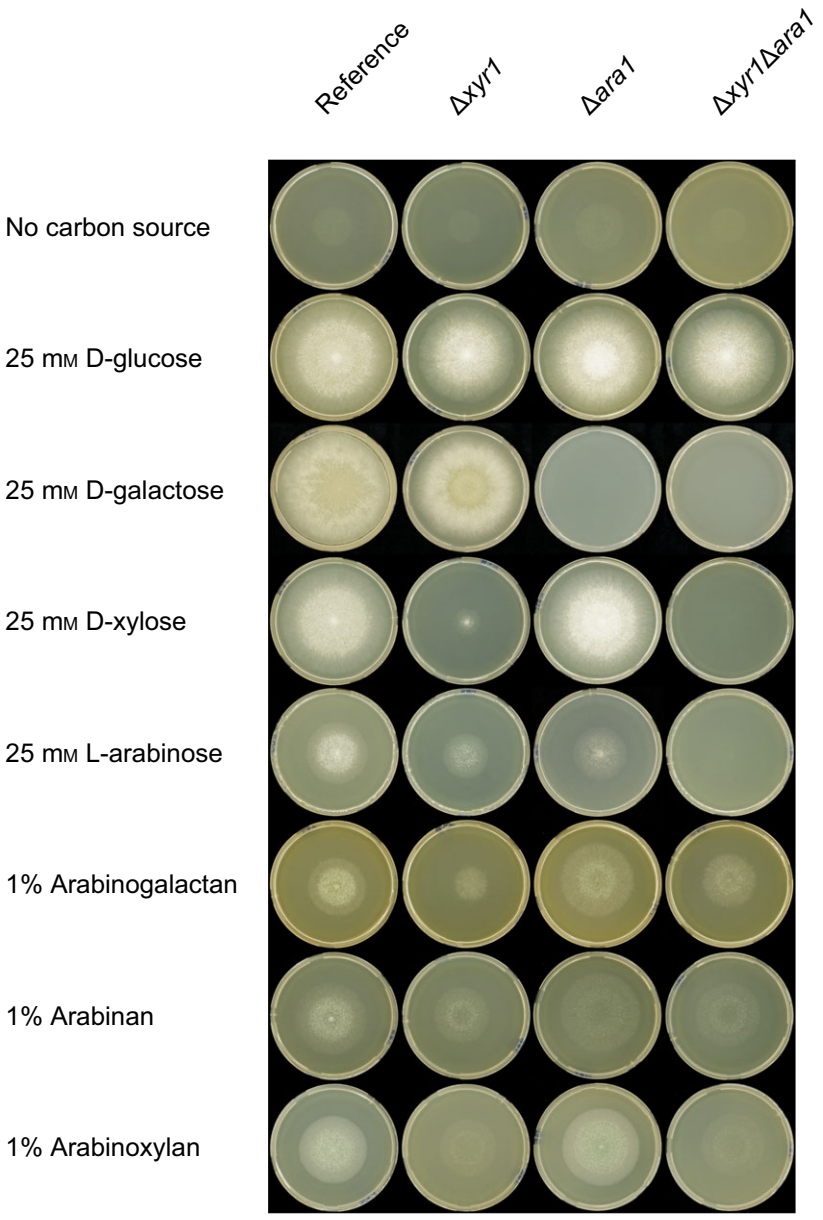

Fig. 1. Growth profile of *Trichoderma reesei* mutants. Plates were incubated for 5 days, except for D-galactose plates (10 days).

in *P. oryzae*. The double deletion (Δ*xyr1*Δ*ara1*) abolished growth on L-arabinose and D-xylose (Fig. 1). The pattern discerned from the *T. reesei* double deletion showed that the presence of either XYR1 or ARA1 is essential for growth on these pentoses and that there were no other compensating TFs able to support growth. This is the same as in *A. niger* [14,48], where deletion of both XlnR and AraR (which is a functional but not genetic ortholog of ARA1 [3]) abolished growth on L-arabinose and D-xylose. Surprisingly, the double deletion (Δ*xyr1*Δ*ara1*) did not abolish, but only reduced growth on the polymers arabinan, arabinogalactan and arabinoxylan. This could be explained by small amounts of sugars other than L-arabinose and D-galactose in these polymers, which may support limited growth. According to the manufacturer, arabinan contains 1.4% of L-rhamnose and 0.7% of galacturonic acid and arabinoxylan contains less than 1% of D-glucose and D-mannose. Alternatively, these saccharides or growth on polymeric substrates could induce other TFs that compensate for the loss of XYR1 and ARA1.

The difference in how ARA1 functions in L-arabinose and arabinan utilization in *T. reesei* and *P. oryzae*, whereby only *P. oryzae* ARA1 was essential for growth on those carbon sources, could partly

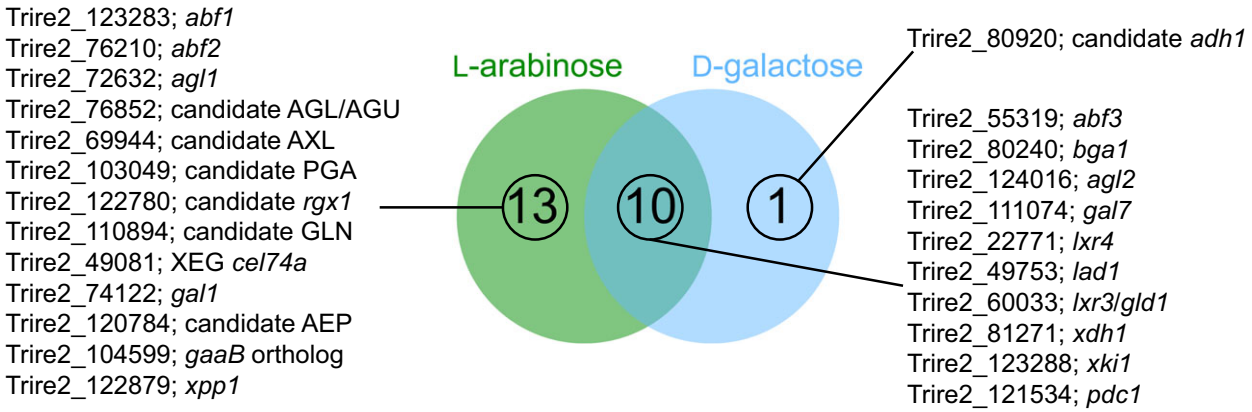

Fig. 2. Comparison of genes regulated by ARA1 in L-arabinose and D-galactose. Genes regulated (> 2.5-fold change, > 15 FPKM in reference strain, *P*-value < 0.05) by ARA1 in both conditions are shown. Genes are listed in the order of PBD, catabolic and TF gene categories, whereby listed first in each category are those genes involved in L-arabinose or D-galactose release or catabolism.

be due to the different lifestyles of these fungi. A saprobe (*T. reesei*) may have more redundancy in regulators to adapt to different environments than a plant pathogen (*P. oryzae*) that is specialized to its hosts with only one regulator. This has already been speculated for the orthologs of XlnR/XYR1/XLR1 from fungi with different lifestyles [3,49]. A technical factor of note is that the *T. reesei* spores were pregerminated, whereas *A. niger* and *P. oryzae* [15] spores were inoculated on carbon sources without any pregermination. After confirming our initial hypothesis that ARA1 is involved in regulating the response to both L-arabinose and D-galactose by finding reduced or abolished growth of the mutants on these sugars, we then analyzed for which genes the expression was affected by these mutants during growth on these sugars.

Transcriptomics showed that ARA1 regulates genes involved in D-galactose catabolism/conversion in addition to L-arabinose utilization

Mycelia from precultures were transferred to medium containing either L-arabinose or D-galactose and sampled after 2 h for transcriptomic analysis. Duplicate samples were highly reproducable (Figure S3). The transcriptomic analysis focused only on genes encoding CAZymes, carbon catabolic enzymes and TFs involved in plant biomass utilization (Table S2), which were downregulated compared to QM9414 in either of the sugars (Table S2). In total, 10 genes were downregulated on both monomers, 13 genes only on L-arabinose, and 1 gene only on D-galactose (Fig. 2, Table S2). The shared set of 10 genes consisted of seven catabolic genes (*pdc1*, *gal7*, *lxr4*, *lad1*, *lxr3*/*gld1*, *xdh1* and *xki1*; involved in PCP, D-galactose oxidoreductive pathway or the Leloir pathway), one β-galactosidase (*bga1*), one α-galactosidase (*agl2*), and one ABF (*abf3*). The 13 genes downregulated in Δ*ara1* only on L-arabinose consisted of three catabolic genes (*gal1, putative* glycolytic aldose 1-epimerase/mutarotase (AEP) Trire2_120784 and putative *gaaB* orthologs Trire2_104599), two ABFs (*abf1* and *abf2*), one AGL (*agl1*), one putative AGL/α-glucuronidase (Trire2_76852), one putative endo-β-galactanase (Trire2_110894), two putative pectinases (Trire2_103049 and *rgx1*), one putative α-xylosidase (Trire2_69944), one xyloglucanase (XEG) (*cel74a*), and the secondary metabolism repressor *xpp1* [50]. One gene [Trire2_80920, putative alcohol dehydrogenase (ADH) *adh1*] was downregulated only on D-galactose, possibly due to an indirect effect because the downregulation was small at less than threefold (Table S2).

When the genes regulated by ARA1 on one or both sugars are analyzed from the perspective of the sugars they utilize, this shows that ARA1 regulates genes involved in D-galactose in addition to L-arabinose release and catabolism along with genes involved in pectin backbone utilization. Of the 24 genes regulated by ARA1, 10 genes are involved in D-galactose release and catabolism, seven genes are involved in L-arabinose release and catabolism and four genes are involved in pectin backbone release and catabolism. This is the first report of ARA1 regulating enzymes involved in D-galactose release and catabolism. Regulation of genes involved in pectin backbone release and catabolism was found also for *A. niger* AraR [17,51,52], matching with the composition of pectin where D-galacturonic acid, L-arabinose, and D-galactose are the predominant saccharides [53–55].

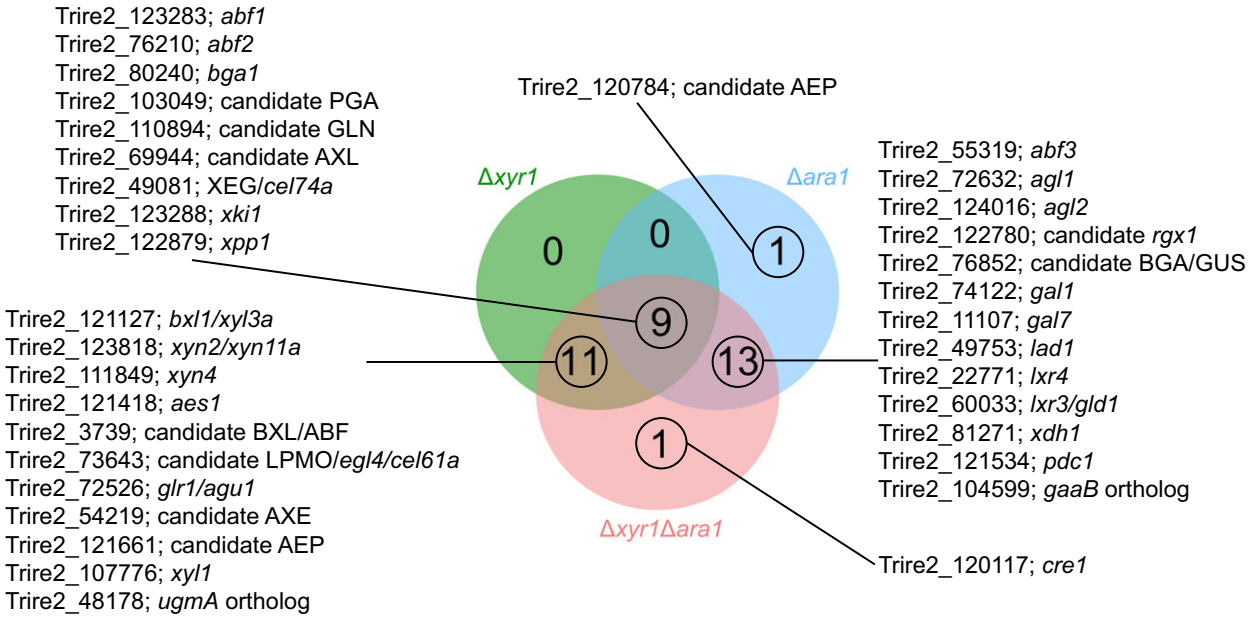

Fig. 3. Comparison of genes downregulated by the three mutants (Δ*xyr1*, Δ*ara1*, Δ*xyr1*Δ*ara1*) on L-arabinose. Genes downregulated (> 2.5 fold change, > 15 FPKM in reference strain, *P*-value < 0.05) in each mutant in L-arabinose are shown. Genes are listed in the order of PBD, catabolic and TF gene categories, whereby listed first in each category are those genes involved in L-arabinose or D-galactose release or catabolism.

ARA1 and XYR1 coregulate L-arabinose response, but ARA1 has a larger impact in L-arabinose utilization

The growth profile suggested that both ARA1 and XYR1 coregulated the response to L-arabinose, whereby growth was reduced in each single mutant on L-arabinose and arabinan and abolished in the double mutant on L-arabinose. In order to investigate the relative regulatory contribution of ARA1 and XYR1 in response to L-arabinose, the transcriptomes from the three mutants (Δ*xyr1*, Δ*ara1*, Δ*xyr1*Δ*ara1*) on L-arabinose after 2 h were analyzed. The transcriptomics showed that both XYR1 and ARA1 regulate genes in response to L-arabinose (Fig. 3). The genes that are involved in L-arabinose utilization are either regulated by only ARA1 or both TFs, with one exception (*xyl1*) regulated by only XYR1. Of the 13 genes that were regulated only by ARA1, four are involved in L-arabinose release and catabolism (three PCP genes, *lad1*, *lxr3*, *xdh1,* and one ABF *abf3*), along with genes involved in D-galactose release and catabolism (*gal7, lxr4, gal1, lad1, xdh1*, putative BGA Trire2_76852, *agl1*, and *agl2*). Of the nine genes regulated by both TFs, three genes are involved in L-arabinose release and catabolism (one PCP gene *xki1* and two ABFs *abf1* and *abf2*). This was previously reported for XYR1, with respect to *abf2* and to a lesser extent *abf1*, but not for *xki1* [7,29]. The genes regulated by both also included two D-galactose-releasing enzymes (*bga1* and Trire2_110894), two D-xylose-releasing enzymes (*cel74a* and Trire2_69944), one pectinase (Trire2_103049) and the TF *xpp1* (secondary metabolism repressor. The 11 genes that were regulated only by XYR1 encode mainly xylanases along with *bxl1* (described as BXL with a separated ABF activity) and D-xylose reductase (*xyl1*), which were shown previously to be regulated by XYR1 in L-arabinose [29]. The *cre1* gene was the only one (of the gene categories analyzed) that was downregulated more than 2.5-fold in the Δ*xyr1*Δ*ara1* strain, possibly due to an indirect effect.

Direct comparison of the expression level between the two single mutants of the genes regulated by both ARA1 and XYR1 (Table S2, Table 1, Fig. S4) showed that the expression was significantly lower in Δ*ara1* for two of the three genes involved in L-arabinose utilization (*abf1* and *abf2*). The third gene involved was the xylulokinase *xki1*, where both regulators contributed similarly. The *xki1* expression in the Δ*xyr1*Δ*ara1* mutant was significantly lower compared to each single mutant indicating an additive effect of the TFs on the regulation of *xki1* (Table S2). Although the literature indicated that XYR1 was partially involved in regulating L-arabinose utilization during growth on L-arabinose

Table 1. Summary table with DE genes in L-arabinose. DE genes downregulated in L-arabinose and their main regulator(s) are shown. In bold is marked the main regulator, while in brackets and smaller font the regulator which has a minor impact. Impact of each regulator is according to log2 fold change by direct comparison of the two single mutants (Δ*ara1* Vs Δ*xyr1*; Table S2). / = both ARA1 and XYR1 have the same impact in regulation, + = both ARA1 and XYR1 are required for transcription of target gene (gene is DE only in Δ*xyr1*Δ*ara1*), AR = gene involved in L-arabinose release; AC = gene involved in L-arabinose catabolism, GR = gene involved in D-galactose release, GC = gene involved in D-galactose catabolism, AGC = gene involved in L-arabinose and D-galactose catabolism.

Regulator	Trire2_ID	Monomer release and catabolism	Annotation (gene name)	Δ*ara1* vs Δ*xyr1* log2 fold change
ARA1 (XYR1)	123283	AR	ABF I (*abf1*)	−5.61*
ARA1 (XYR1)	76210	AR	Candidate ABF (*abf2*)	−4.16*
ARA1	55319	AR	Candidate ABF (*abf3)*	
ARA1 (XYR1)	80240	GR	BGA (*bga1*)	−6.28*
ARA1	76852	GR	Candidate BGA/GUS	
ARA1	72632	GR	AGL (*agl1*)	
ARA1	124016	GR	AGL (*agl2*)	
ARA1 (XYR1)	110894	GR	Candidate endo-β-1, 6-galactanase	−3.87*
ARA1 (XYR1)	103049	GR	Candidate PGA	−6.15*
ARA1	122780	GR	Candidate exo-rhamnogalacturonase (*rgx1*)	−3.00*
XYR1	123818		Endo-β-1,4-xylanase (*xyn2*)	
XYR1	111849		Endo-β-1,4-xylanase (*xyn4*)	
XYR1	121127	AR	BXL (*bxl1*)	
XYR1	3739		Candidate BXL/ABF	
ARA1 (XYR1)	69944		Candidate AXL/α-glucosidase	−2.42*
XYR1	54219		Candidate AXE	
XYR1	121418		Acetyl esterase (*aes1*)	
XYR1	72526		AGU (*glr1/agu1*)	
ARA1/XYR1	49081		XEG (*cel74a*)	0.06
XYR1	73643		Candidate LPMO (*egl4/cel61a*)	7.26*
XYR1	107776	AGC	*xyl1*	3.17*
ARA1	49753	AGC	*lad1*	
ARA1	60033	AC	*lxr3/gld1*	
ARA1	81271	AGC	*xdh1*	
ARA1/XYR1	123288	AC	*xki1*	0.58*
ARA1	22771	GC	*lxr4*	
ARA1	74122	GC	*gal1*	
ARA1	111074	GC	*gal7*	
XYR1	48178		*ugmA* ortholog	
ARA1	104599		*gaaB* ortholog	
ARA1	120784		AEP	
ARA1 (XYR1)	121534		Pyruvate decarboxylase (*pdc1*)	−1.93*
XYR1	121661		AEP	
ARA1/XYR1	122879		*xpp1*	−0.45
ARA1 + XYR1	120117		*cre1*	0.045

*P-value < 0.05.

or arabinan [7,29], the Δ*ara1* mutant we made allowed us to clearly show that ARA1 in response to L-arabinose is the main regulator. ARA1 regulates most of the genes for L-arabinose utilization, with the exception of the first and last step of L-arabinose catabolism, controlled respectively by XYR1 and both TFs (Table S2, Fig. S4, Table 1). In addition, our study confirms that the arabinanolytic enzyme BXL1 is regulated only by XYR1 [29] (Table S2, Fig. S4, Table 1). With regard to regulation of TFs in the single mutants, it is of note that *xyr1* was not DE in Δ*ara1* or *ara1* DE in Δ*xyr1* in response to either L-arabinose or D-galactose (Table S2). Therefore, the genes DE in each mutant are not due to an indirect effect from altered expression of *xyr1* or *ara1*.

ARA1, but not XYR1, regulates PBD CAZymes and metabolic genes in response to D-galactose

Whether there was coregulation by ARA1 and XYR1 of the response to D-galactose (in a similar manner to the coregulated response to L-arabinose)

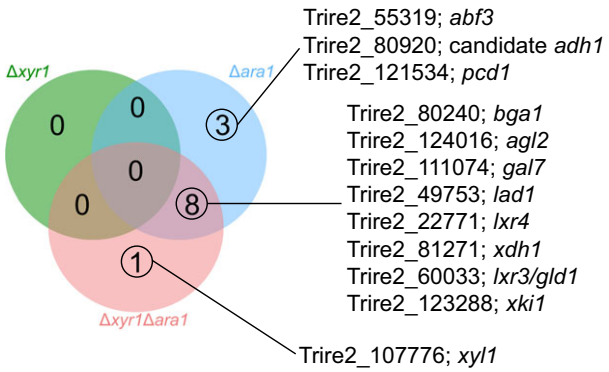

Fig. 4. Comparison of genes downregulated by the three mutants (Δ*xyr1*, Δ*ara1*, Δ*xyr1*Δ*ara1*) on D-galactose. Genes downregulated (> 2.5-fold change, > 15 FPKM in reference strain, *P*-value < 0.05) in each mutant in D-galactose are shown. Genes are listed in the order of PBD, catabolic and TF gene categories, whereby listed first in each category are those genes involved in L-arabinose or D-galactose release or catabolism.

was investigated by comparing the genes downregulated in the single and double mutants in response to D-galactose. Of 218 genes analyzed, none were downregulated in Δ*xyr1* in response to D-galactose, while 11 genes were regulated by ARA1, including four D-galactose catabolic genes (*lxr4*, *lad1*, *xdh1*, *gal7*) and two genes encoding D-galactose-releasing enzymes (*bga1*, *agl2*; Table S2, Fig. 4). The D-xylose reductase *xyl1* was the only gene downregulated in the double mutant, but not in Δ*ara1* or Δ*xyr1*, in response to D-galactose, which indicates that a contribution from both TFs may be required to regulate *xyl1*.

Previous studies indicated that XYR1 was partially involved in D-galactose release and catabolism [3,25,28], but our analysis (Fig. 4, Table S2) clearly showed that XYR1 did not regulate any PBD CAZyme or metabolic genes in response to D-galactose, and that instead ARA1 was the regulator responding to D-galactose. In particular, transcriptomic analysis showed that *ara1* deletion caused the blockage of D-galactose catabolism, downregulating both the Leloir and D-galactose oxidoreductive pathways. The only exception is the D-xylose reductase *xyl1*, which is regulated by only XYR1 (Table S2, Fig. 4), which is consistent with previous studies [7,28,29]. This blockage may explain why the Δ*ara1* strain could not grow on D-galactose (Fig. 1). Considering that most of the studies on D-galactose regulation were performed on lactose (where XYR1 was partially involved), this suggests that lactose and D-galactose responses use (partially) different mechanisms. In addition, our results clearly showed that the main BGA *bga1* (Trire2_ 80240) was regulated by ARA1 on D-galactose (Fig. 4, Table S2). A previous study of the *T. reesei* Δ*xyr1* mutant on lactose showed that XYR1 mediates lactose induction and regulates *bga1* [24]. This suggests that the regulator of *bga1* transcription can differ depending on whether the inducing sugar is lactose or D-galactose.

In conclusion, our results confirm the initial hypothesis, where ARA1 is involved in both L-arabinose and D-galactose utilization (Table 1). XYR1 is still necessary for the full control of L-arabinose catabolism and releasing enzymes, in agreement with previous reports of XYR1 partially regulating L-arabinose utilization [7,29]. However, we showed that ARA1 had the larger regulatory impact on this pathway (Table 1, Table S2 and Fig. S4). In addition, ARA1 was essential for D-galactose utilization in *T. reesei*, regulating both catabolism and release. Previously, Kowalczyk *et al.* [17] showed that the arabinanolytic regulator AraR in *Aspergillus nidulans* was partially involved in D-galactose catabolism, but not in D-galactose release. No regulators that primarily regulate D-galactose release and catabolism have been shown to be involved in L-arabinose release or catabolism [17,21,22]. This is the first report of an L-arabinose- or D-galactose-responsive regulator that is involved to similar extents in both the release and catabolism of both L-arabinose and D-galactose. This dual role of ARA1 in *T. reesei* matches with the composition of lignocellulose in the natural environment of *T. reesei* where both monomers are generally present. *Trichoderma reesei* has evolved a less complex regulation of the release and catabolism of these monomers, in terms of the number of TFs required, than *A. niger* where this utilization is regulated by two TFs GalX [21] and AraR [16]. Additional investigations are needed in other species to verify if this dual role of ARA1 (regulating D-galactose utilization in addition to L-arabinose utilization) is unique to *T. reesei* or is a common feature in all the fungal classes where ARA1 is present (Sordariomycetes and Letiomycetes).

Acknowledgements

This study was supported by a grant of the Dutch Technology Foundation STW, Applied Science Division of NWO, and the Technology Program of the Ministry of Economic Affairs 016.130.609 to RPdV. PD was supported by a grant from the ALW division of NWO (824.15.023) to RPdV. BS was supported by the Austrian Science Fund project P24219. We acknowledge support of GenomeScan, project number AGS17086/103062, for the transcriptome sequencing.

Author contributions

TB carried out the experiments, analyzed the data, and wrote the manuscript. MVAP performed part of the bioinformatics analysis. RSK generated the *ara1* deletion strains. RPdV and PD supervised the experiments and the manuscript. BS reviewed the manuscript and provided *T. reesei* QM9414 strains ATCC 26921 (reference), Δ*tku70*, Δ*xyr1*. All authors read and approved the final manuscript.

References

1 Bischof RH, Ramoni J and Seiboth B (2016) Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb Cell Fact* **15**, 106.

2 de Vries RP and Visser J (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microb Mol Biol Rev* **65**, 497–522.

3 Benocci T, Aguilar-Pontes MV, Zhou M, Seiboth B and de Vries RP (2017) Regulators of plant biomass degradation in ascomycetous fungi. *Biotechnol Biofuels* **10**, 152.

4 Gupta VK, Steindorff AS, de Paula RG, Silva-Rocha R, Mach-Aigner AR, Mach RL and Silva RN (2016) The post-genomic era of *Trichoderma reesei*: what's next? *Trends Biotechnol* **34**, 970–982.

5 Harris PJ, Kelderman MR, Kendon MF and Mckenzie RJ (1997) Monosaccharide compositions of unlignified cell walls of monocotyledons in relation to the occurrence of wall-bound ferulic acid. *Biochem System Ecol* **25**, 167–179.

6 Makela MR, Donofrio N and de Vries RP (2014) Plant biomass degradation by fungi. *Fungal Genet Biol* **72**, 2–9.

7 Seiboth B and Metz B (2011) Fungal arabinan and L-arabinose metabolism. *Appl Microbiol Biotechnol* **89**, 1665–1673.

8 Witteveen CFB, Busink R, Vandevondervoort P, Dijkema C, Swart K and Visser J (1989) L-Arabinose and D-Xylose Catabolism in *Aspergillus niger*. *J Gen Microbiol* **135**, 2163–2171.

9 Fekete E, Karaffa L, Sándor E, Bányai I, Seiboth B, Gyémánt G, Sepsi A, Szentirmai A and Kubicek CP (2004) The alternative D-galactose degrading pathway of *Aspergillus nidulans* proceeds via L-sorbose. *Arch Microbiol* **181**, 35–44.

10 Mojzita D, Herold S, Metz B, Seiboth B and Richard P (2012) L-xylo-3-hexulose reductase is the missing link in the oxidoreductive pathway for D-galactose catabolism in filamentous fungi. *J Biol Chem* **287**, 26010–26018.

11 Holden HM, Rayment I and Thoden JB (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J Biol Chem* **278**, 43885–43888.

12 Frey PA (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J* **10**, 461–470.

13 Khosravi C, Benocci T, Battaglia E, Benoit I and de Vries RP (2015) Sugar catabolism in *Aspergillus* and other fungi related to the utilization of plant biomass. *Adv Appl Microbiol* **90**, 1–28.

14 Battaglia E, Hansen SF, Leendertse A, Madrid S, Mulder H, Nikolaev I and de Vries RP (2011) Regulation of pentose utilisation by AraR, but not XlnR, differs in *Aspergillus nidulans* and *Aspergillus niger*. *Appl Microbiol Biotechnol* **91**, 387–397.

15 Klaubauf S, Zhou M, Lebrun MH, de Vries RP and Battaglia E (2016) A novel L-arabinose-responsive regulator discovered in the rice-blast fungus *Pyricularia oryzae* (*Magnaporthe oryzae*). *FEBS Lett* **590**, 550–554.

16 Battaglia E, Zhou M and de Vries RP (2014) The transcriptional activators AraR and XlnR from *Aspergillus niger* regulate expression of pentose catabolic and pentose phosphate pathway genes. *Res Microbiol* **165**, 531–540.

17 Kowalczyk JE, Gruben BS, Battaglia E, Wiebenga A, Majoor E and de Vries RP (2015) Genetic Interaction of *Aspergillus nidulans galR*, *xlnR* and *araR* in Regulating D-Galactose and L-Arabinose Release and Catabolism Gene Expression. *PLoS ONE* **10**, e0143200.

18 Ivanova C, Baath JA, Seiboth B and Kubicek CP (2013) Systems analysis of lactose metabolism in *Trichoderma reesei* identifies a lactose permease that is essential for cellulase induction. *PLoS ONE* **8**, e62631.

19 Hartl L, Kubicek CP and Seiboth B (2007) Induction of the gal pathway and cellulase genes involves no transcriptional inducer function of the galactokinase in *Hypocrea jecorina*. *J Biol Chem* **282**, 18654–18659.

20 Escalante-Chong R, Savir Y, Carroll SM, Ingraham JB, Wang J, Marx CJ and Springer M (2015) Galactose metabolic genes in yeast respond to a ratio of galactose and glucose. *Proc Natl Acad Sci USA* **112**, 1636–1641.

21 Gruben BS, Zhou M and de Vries RP (2012) GalX regulates the D-galactose oxido-reductive pathway in *Aspergillus niger*. *FEBS Lett* **586**, 3980–3985.

22 Christensen U, Gruben BS, Madrid S, Mulder H, Nikolaev I and de Vries RP (2011) Unique regulatory mechanism for D-galactose utilization in *Aspergillus nidulans*. *Appl Environ Microbiol* **77**, 7084–7087.

23 Sellick CA, Campbell RN and Reece RJ (2008) Galactose metabolism in yeast-structure and regulation of the leloir pathway enzymes and the genes encoding them. *Int Rev Cell Mol Biol* **269**, 111–150.

24 Stricker AR, Steiger MG and Mach RL (2007) Xyr1 receives the lactose induction signal and regulates lactose metabolism in *Hypocrea jecorina*. *FEBS Lett* **581**, 3915–3920.

25 Bischof R, Fourtis L, Limbeck A, Gamauf C, Seiboth B and Kubicek CP (2013) Comparative analysis of the *Trichoderma reesei* transcriptome during growth on the cellulase inducing substrates wheat straw and lactose. *Biotechnol Biofuels* **6**, 127.
26 Portnoy T, Margeot A, Seidl-Seiboth V, Le Crom S, Ben Chaabane F, Linke R, Seiboth B and Kubicek CP (2011) Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. *Eukaryot Cell* **10**, 262–271.
27 Karaffa L, Fekete E, Gamauf C, Szentirmai A, Kubicek CP and Seiboth B (2006) D-Galactose induces cellulase gene expression in *Hypocrea jecorina* at low growth rates. *Microbiology* **152**, 1507–1514.
28 Seiboth B, Gamauf C, Pail M, Hartl L and Kubicek CP (2007) 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* **66**, 890–900.
29 Akel E, Metz B, Seiboth B and Kubicek CP (2009) Molecular regulation of arabinan and L-arabinose metabolism in *Hypocrea jecorina* (*Trichoderma reesei*). *Eukaryot Cell* **8**, 1837–1844.
30 Montenecourt BS and Eveleigh DE (1977) Preparation of mutants of *Trichoderma reesei* with enhanced cellulase production. *Appl Environ Microbiol* **34**, 777–782.
31 Penttila M, Nevalainen H, Ratto M, Salminen E and Knowles J (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene* **61**, 155–164.
32 de Vries RP, Burgers K, van de Vondervoort PJ, Frisvad JC, Samson RA and Visser J (2004) A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. *Appl Environ Microbiol* **70**, 3954–3959.
33 Mandels M and Andreotti RE (1978) Problems and challenges in cellulose to cellulase fermentation. *Process Biochem* **13**, 6–15.
34 Guangtao Z, Hartl L, Schuster A, Polak S, Schmoll M, Wang T, Seidl V and Seiboth B (2009) Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. *J Biotechnol* **139**, 146–151.
35 Stricker AR, Grosstessner-Hain K, Wurleitner E and Mach RL (2006) Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. *Eukaryot Cell* **5**, 2128–2137.
36 Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP and Schmoll M (2012) A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. *Biotechnol Biofuels* **5**, 1.
37 Bolger AM, Lohse M and Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.
38 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R and Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**, R36.
39 Kim D, Langmead B and Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357–360.
40 Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL and Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**, 562–578.
41 Edgar R, Domrachev M and Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**, 207–210.
42 Love MI, Huber W and Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.
43 Hakkinen M, Arvas M, Oja M, Aro N, Penttila M, Saloheimo M and Pakula TM (2012) Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. *Microb Cell Fact* **11**, 134.
44 Li L, Stoeckert CJ Jr and Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13**, 2178–2189.
45 Boekhorst J and Snel B (2007) Identification of homologs in insignificant blast hits by exploiting extrinsic gene properties. *BMC Bioinformatics* **8**, 356.
46 R Core Team (2017) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
47 Battaglia E, Klaubauf S, Vallet J, Ribot C, Lebrun M-H and de Vries RP (2013) Xlr1 is involved in the transcriptional control of the pentose catabolic pathway, but not hemi-cellulolytic enzymes in *Magnaporthe oryzae*. *Fungal Genet Biol* **57**, 76–84.
48 Battaglia E, Visser L, Nijssen A, van Veluw GJ, Wosten HA and de Vries RP (2011) Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in Eurotiales. *Stud Mycol* **69**, 31–38.
49 Klaubauf S, Narang HM, Post H, Zhou M, Brunner K, Mach-Aigner AR, Mach RL, Heck AJR, Altelaar AFM and de Vries RP (2014) Similar is not the same: differences in the function of the (hemi-)cellulolytic regulator XlnR (Xlr1/Xyr1) in filamentous fungi. *Fungal Genet Biol* **72**, 73–81.
50 Derntl C, Kluger B, Bueschl C, Schuhmacher R, Mach RL and Mach-Aigner AR (2017) Transcription factor

Xpp1 is a switch between primary and secondary fungal metabolism. *Proc Natl Acad Sci USA* **114**, E560–E569.
51 de Souza WR, Maitan-Alfenas GP, de Gouvêa PF, Brown NA, Savoldi M, Battaglia E, Goldman MH, de Vries RP and Goldman GH (2013) The influence of *Aspergillus niger* transcription factors AraR and XlnR in the gene expression during growth in D-xylose, L-arabinose and steam-exploded sugarcane bagasse. *Fungal Genet Biol* **60**, 29–45.
52 Kowalczyk JE, Lubbers RJM, Peng M, Battaglia E, Visser J and de Vries RP (2017) Combinatorial control of gene expression in *Aspergillus niger* grown on sugar beet pectin. *Sci Rep* **7**, 12356.
53 Ridley BL, O'Neill MA and Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **57**, 929–967.
54 Willats WG, McCartney L, Mackie W and Knox JP (2001) Pectin: cell biology and prospects for functional analysis. *Plant Mol Biol* **47**, 9–27.
55 Muller-Maatsch J, Bencivenni M, Caligiani A, Tedeschi T, Bruggeman G, Bosch M, Petrusan J, Van Droogenbroeck B, Elst K and Sforza S (2016) Pectin content and composition from different food waste streams. *Food Chem* **201**, 37–45.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. L-arabinose and D-galactose catabolism in *Trichoderma reesei*.

Fig. S2. Southern blot of Δ*ara1* and Δ*xyr1*Δ*ara1*.

Fig. S3. Pearson correlation matrix of *Trichoderma reesei* transcriptomes.

Fig. S4. Heatmap with expression values of all DE genes.

Table S1. PCR primers used in this study.

Table S2. *Trichoderma reesei* selected genes and their transcriptome.