Title

Enzymatic Adaptation of Podospora anserina to Different Plant Biomass Provides Leads to Optimized Commercial Enzyme Cocktails

Journal

Biotechnology Journal, 14(4)

ISSN

1860-6768

Authors

Benocci, Tiziano

Daly, Paul

Aguilar-Pontes, Maria Victoria

et al.

Publication Date

2019-04-01

DOI

10.1002/biot.201800185

Peer reviewed

Biotechnology Journal

www.biotechnology-journal.com

Enzymatic Adaptation of *Podospora anserina* to Different Plant Biomass Provides Leads to Optimized Commercial Enzyme Cocktails

*Tiziano Benocci, Paul Daly, Maria Victoria Aguilar-Pontes, Kathleen Lail, Mei Wang, Anna Lipzen, Vivian Ng, Igor V. Grigoriev, and Ronald P. de Vries**

As a late colonizer of herbivore dung, *Podospora anserina* has evolved an enzymatic machinery to degrade the more recalcitrant fraction of plant biomass, suggesting a great potential for biotechnology applications. The authors investigated its transcriptome during growth on two industrial feedstocks, soybean hulls (SBH) and corn stover (CS). Initially, CS and SBH results in the expression of hemicellulolytic and amylolytic genes, respectively, while at later time points a more diverse gene set is induced, especially for SBH. Substrate adaptation is also observed for carbon catabolism. Overall, SBH resulted in a larger diversity of expressed genes, confirming previous proteomics studies. The results not only provide an in depth view on the transcriptomic adaptation of *P. anserina* to substrate composition, but also point out strategies to improve saccharification of plant biomass at the industrial level.

1. Introduction

Podospora anserina is a late colonizer of herbivore dung, suggesting that this fungus evolved an enzymatic machinery to degrade the more recalcitrant fraction of plant biomass.[1] It has a genome rich in cellulases and hemicellulases, including many LPMOs (33 in AA9), as well as putative ligninolytic enzymes.[2–4] In contrast, it has a significantly reduced potential for pectin, sucrose, and inulin degradation.[3,5] This fits with the composition of its natural substrate, which is rich in (hemi-) cellulose and lignin, but poor in pectin.[6,7] These features make *P. anserina* a highly interesting species as a source of novel enzymes for biotechnology.[2]

A recent proteomics study revealed that the set of plant biomass degrading enzymes produced by *P. anserina* is strongly dependent on the composition of the substrate.[8]

In this study we investigated the expression of genes encoding PBD enzymes during growth on two industrial feedstocks, soybean hulls (SBH) (dicot) and corn stover (CS) (monocot), which differ strongly in their composition (Table S1, Supporting Information).[9–11] CS is richer in (hemi-) cellulose content, especially arabinoxylan, with cellobiose, D-xylose, and L-arabinose as main potential inducers. SBH is richer in pectin, xyloglucan, galactomannan, and starch, with a broader range of potential inducers than CS, such as cellobiose, D-xylose, L-arabinose, amylose, D-galactose, mannobiose, L-rhamnose, and D-galacturonic acid.

2. Experimental Section

2.1. Strains, Media, and Growth Conditions

Maintenance and cultivation of *P. anserina* S mat+ (CBS143333) was performed as described previously.[12] Briefly, 2.5 g mycelium from pre-culturing media was washed with M2 before transfer into 250 mL Erlenmeyer flasks containing 50 mL M2 with 1% corn stover or soybean hulls, and incubated in a rotatory shaker at 27 °C and 120 rpm. After 4, 24, and 48 h, the mycelium was harvested as described by Klaubauf et al.[13] All experiments were performed in triplicate.

2.2. Molecular Biology Methods, RNA Sequencing, and Read Mapping

Total RNA was extracted from mycelium, purified and checked for quality as described previously.[13] RNA sequencing was

T. Benocci, Dr. P. Daly, M. V. Aguilar-Pontes, Prof. R. P. de Vries
Fungal Physiology
Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology
Utrecht University
Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands
E-mail: r.devries@westerdijkinstitute.nl

K. Lail, M. Wang, A. Lipzen, V. Ng, Prof. I. V. Grigoriev
US Department of Energy Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA

Prof. I. V. Grigoriev
Department of Plant and Microbial Biology
University of California Berkeley
Berkeley, CA 94598, USA

DOI: 10.1002/biot.201800185

performed using Illumina HiSeq2500 (yield 1TB of 1×101 bp). Raw fastq file reads were filtered and trimmed using the JGI QC pipeline. Using BBDuk [BBDuk: https://sourceforge.net/projects/bbmap/] raw reads were evaluated for artifact sequence by kmer matching (kmer = 25), allowing 1 mismatch and detected artifact was trimmed from the 3′ end of the reads. RNA spike-in reads, PhiX reads, and reads containing any Ns were removed. Quality trimming was performed using the phred trimming method set at Q6. Reads under the length threshold were removed. Filtered reads from each library were aligned to the reference genome (https://genome.jgi.doe.gov/Podan2/download/Podan2_AssemblyScaffolds_Repeatmasked.fasta.gz) using HISAT version 0.1.4-beta.[14] featureCounts[15] was used to generate the raw gene counts using gff3 annotations. On average 94% of the reads mapped to the genome. The RNA-seq data have been deposited at the Sequence Read Archive at NCBI with individual sample BioProject Accession numbers (PRJNA442509 to PRJNA442527).

2.3. Data Analysis

DESeq2 (version 1.10.0)[16] 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 p-value ≤ 0.05, foldchange > 2.5 (log2foldchange > 1.32) and FPKM ≥ 17 in at least one condition. Genes with FPKM values <17 in both conditions were considered lowly expressed and ignored in the analysis.

Transcriptome analysis focused on genes encoding CAZymes, carbon catabolic enzymes, and transcription factors (TFs). The CAZy gene set was made by combining previous annotations[8] with the JGI database (https://genome.jgi.doe.gov/mycocosm/annotations/browser/cazy/summary;mYVY-F?p=Podan2), and then selecting only the plant biomass degrading (PBD) CAZymes. We re-annotated the AA9 LPMOs family (Figure S1, Supporting Information) based on phylogeny. *P. anserina* carbon catabolic genes were identified by orthology with the genes of *Aspergillus niger* CBS 513.88 using orthoMCL. The parameters for detecting orthoMCL clusters were according to Li et al.,[17] using inflation factor 1, *E* value cutoff 1 E-3, percentage match cutoff 60%[18] and the final set was manually curated based on literature.[19] The *P. anserina* TFs were those identified previously.[20]

PBD genes were clustered using the heatmap.2 function from the gplots_3.0.1 package in R statistical language and environment 3.4.0. Raw gene counts were used to evaluate the level of correlation between biological replicates using Pearson's correlation matrix (Figure S3, Supporting Information).

3. Results and Discussion

3.1. *Podospora anserina* Adapted PBD CAZy Gene Expression to Different Substrates

Mycelia from pre-cultures were transferred to CS or SBH and sampled after 4, 24, and 48 h for transcriptome analysis, focused on genes encoding CAZymes involved in plant biomass degradation (PBD), carbon catabolic enzymes, and related TFs (Table S2, Supporting Information). In order to investigate the adaptation to each substrate, we initially compared the PBD CAZyme transcriptome profiles by clustering all six conditions (Figure S2, Supporting Information). The 4 h samples of the substrates clustered together, while for 24 and 48 h the samples clustered by substrate, indicating that at later time points a larger difference in the response to the substrates became apparent.

A deeper analysis using the fold change in the PBD CAZy gene expression (**Figure 1**A) showed that there were more substrate-adapted induced CAZymes than commonly induced CAZymes, when the number of CS-adapted and SBH-adapted genes was combined. The total number of expressed PBD CAZy genes increased from 4 to 24 h and then reached a plateau for both substrates. However, the number of substrate adapted PBD CAZy genes increased over time for SBH, while it decreased for CS. Initially, CS induced mostly (hemi-)cellulolytic genes, while SBH induced mainly amylolytic genes, matching with the easily metabolized compounds of each substrate. The commonly induced genes (Figure 1A) at each time point encode enzymes acting on starch, (hemi-)cellulose, lignin and pectin, suggesting that these may be the core gene set responding to plant biomass in this fungus. Core gene sets encoding plant biomass degrading enzymes were detected for basidiomycete fungi[21] and 60% were shared with the *P. anserina* commonly induced genes, resulting in 18 shared enzyme activities. This high proportion of shared enzymes between basidiomycetes and *P. anserina* suggests that, besides their different ecological niches, these fungi live in a natural habitat similar in terms of composition, evolving a similar enzymatic machinery to degrade it. In addition, these common activities could be used to develop a general commercial enzyme cocktail, which could be tweaked depending on the final use instead of developing many substrate-specific enzyme cocktails at industrial level.

The total expression of PBD encoding genes (Figure 1B) was analyzed, resulting in a similar pattern to the number of substrate-adapted CAZy genes: decreased expression in CS over time, but increased expression in SBH.

3.2. *P. anserina* Partially Adapted Sugar Catabolism But Not Its Regulatory System to the Substrate Composition

Most of the sugar catabolic genes were induced similarly by both substrates (**Figure 2**A). Genes of the pentose catabolic pathway (PCP) and pentose phosphate pathway (PPP) were more induced at 4 h on CS than on SBH, while genes D-galacturonic acid catabolism were more induced at 24 and 48 h on SBH, suggesting a partial adaptation to the substrate also at the sugar catabolic level. Total sugar catabolic gene expression was highest at 4 h, mainly due to glycolytic genes, and decreased over time for both substrates with a larger decrease for CS (Figure 2B).

The induction of regulatory system related to plant biomass degradation was apparent on both substrates (Figure S4, Supporting Information), suggesting limited adaptation to the substrates at the level of regulator gene expression. This is not surprising, because these TFs are often post-transcriptionally regulated, such as by phosphorylation[22] or nucleo-cytoplasmic shuttling.[23] Only at 4 h a significantly higher induction on CS than SBH was observed for the orthologs of cellulolytic

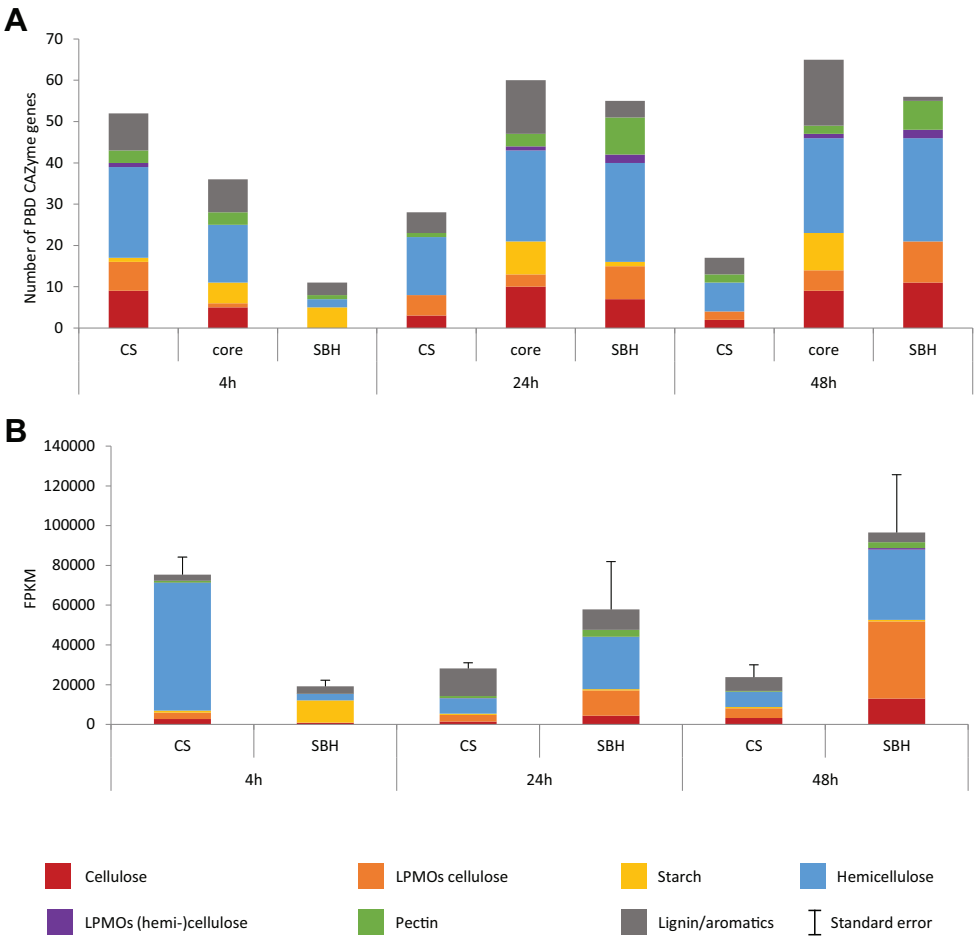

Figure 1. Comparison of the expression of PBD encoding genes in CS and SBH. Fold-change analysis (A) was performed comparing differentially expressed gene (DE) between CS and SBH at the same time point. DE gene has p-value ≤ 0.05, foldchange > 2.5 (log2foldchange > 1.32), and FPKM ≥ 17 in at least one condition. Expression analysis (B) was performed summing the average of the three replicates. Error bar represents standard error on the total PBD encoding genes expression.

regulators from *Neurospora crassa*, *clr-1* and *clr-2* or the (hemi-) cellulolytic regulator from *A. niger xlnR*,[20] which may explain the higher induction of (hemi-)cellulolytic genes on CS at 4 h. The regulatory system in *P. anserina* is poorly characterized, but our dataset suggests similarities to *N. crassa*, where *clr-1* and *clr-2* regulate cellulolytic genes and *xlr-1* hemicellulolytic genes.[24] To confirm this, a proper TF characterization is required, such as by deletion strains.

3.3. Transcriptomics Provides Candidate Genes and Leads for Improving Feedstock Saccharification

AA9 LPMOs are highly abundant in the *P. anserina* genome and can improve the efficiency of industrial enzyme cocktails.[25] Most characterized AA9 LPMOs act on cellulose, but activity toward other polymers has also been demonstrated.[26,27] So far, only six out of 33 *P. anserina* AA9 LPMOs have been characterized.[28] Co-expression analysis (Figure S2, Supporting Information) showed that the starch acting AA13 LPMO (Pa_3_10650|5411) clustered with amylases. Two LPMOs clustered with pectinases (Figure S2, Supporting Information, cluster 9), suggesting that these LPMOs could be active on pectin as well as on (hemi-)cellulose (Figure S1, Supporting Information). Two LPMOs were specifically induced only on SBH (Figure S2, Supporting Information, cluster 11). Two LPMOs (Figure S2, Supporting Information, cluster 7) were highly expressed at later time points in both substrates, suggesting that they could be particularly active on recalcitrant cellulose. Another LPMO (Figure S2, Supporting Information, cluster 4) was highly expressed in all conditions with the exception of 4 h on SBH. These enzymes were also found in a proteomics study,[8] indicating that they could significantly improve saccharification, as described in **Table 1**. Recently it was shown that LPMO can use other electron donors than CDHs[29] and that *P. anserina* AA3_2 enzymes can act as electron donors for LPMOs.[30] Our analysis identified an AA3_3 enzyme as a putative LPMO electron donor, as it was found in a high expression cluster (Figure S2, Supporting Information, cluster 7) with three LPMOs (Table 1). All these candidate genes for biotech applications (Table 1) will require further analysis to fully analyze their relevance (e.g., by construction of deletion strains and/or biochemical characterization of the enzymes). However, recent studies already demonstrated the relevance of the right

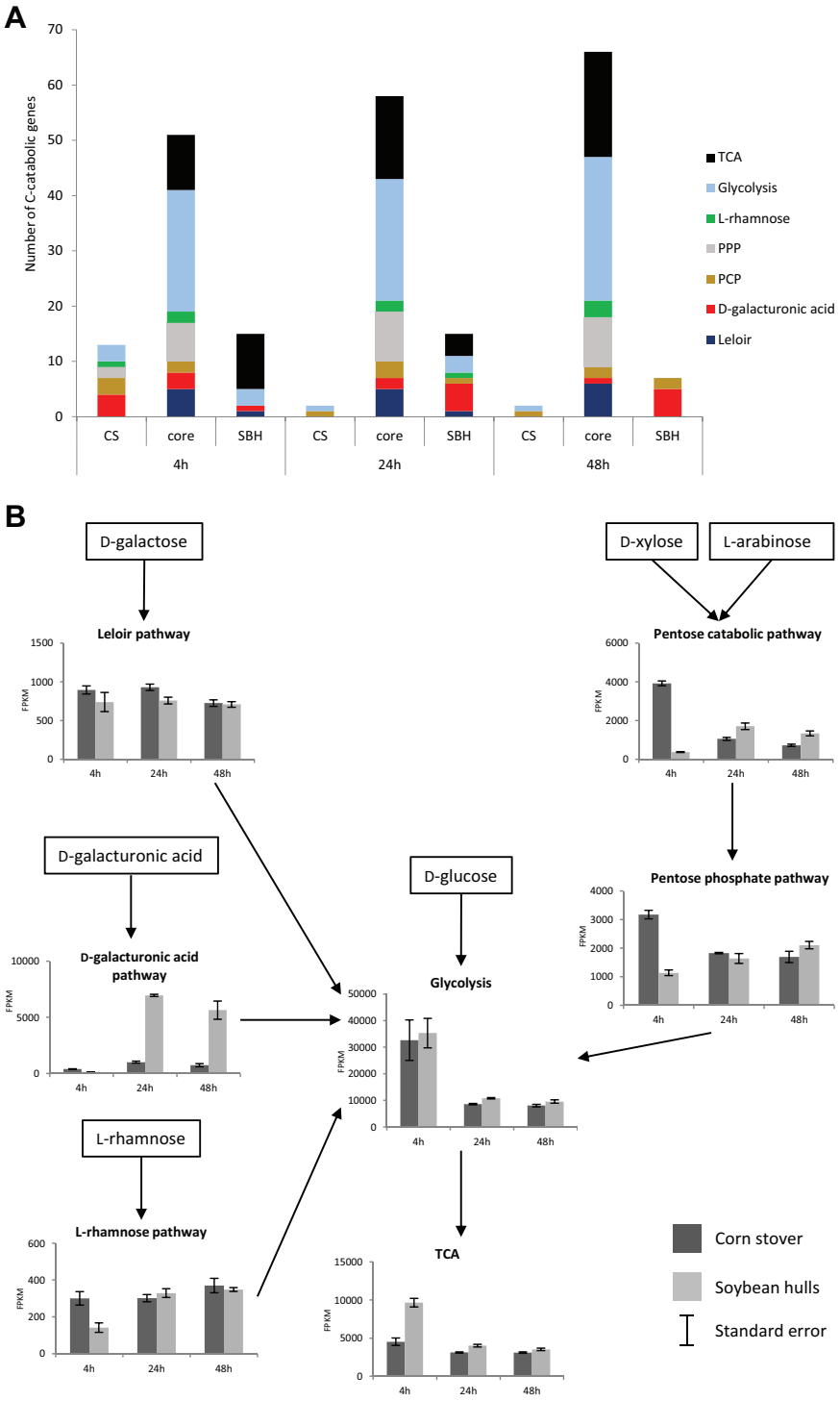

Figure 2. Comparison of the expression of carbon-catabolic genes in CS and SBH. Fold-change analysis (A) was performed comparing differentially expressed gene (DE) between CS and SBH at the same time point. DE gene has p-value ≤ 0.05, foldchange > 2.5 (log2foldchange > 1.32), and FPKM ≥ 17 in at least one condition. Expression analysis (B) was performed summing the average of the three replicates. Error bar represents standard error on the total carbon-catabolic encoding genes expression.

combination of enzymes,[31] suggesting a similar effect for the *P. anserina* enzymes.

The PBD CAZyme transcriptomic responses (Figure 1B) suggested other strategies for improving the saccharification reaction efficiency, by suggesting which enzyme activities to add. In addition our data provides a deep view on the induction system of the saccharifying enzymes, which could improve the production of these enzymes at industrial level, such as a more

Table 1. Selected genes for heterologous expression to improve saccharification efficiency.

Podan2_ ProteinID	CAZy domains and annotation	Cluster[a]	Putative degradation target(s)	Reason(s) for selection and potential improvement
1711	AA9|LPMO	4	Cellulose	Highly expressed in almost all conditions (except 4 h on SBH). Could improve cellulose degradation in many substrates.
3187	AA9|LPMO	7	Cellulose	Among the most highly expressed PBD CAZy genes at 24 and 48 h on SBH and CS.
9365	AA9|LPMO	7	Cellulose	Could improve degradation of recalcitrant cellulose that is enriched at later time points in cultures or reactions
2309	AA3_3|GMC	7	Cellulose	Found in highly expressed LPMO cluster. Putative alternative electron donor for the co-clustering LPMOs.
7738	AA9-CBM1|LPMO	9	Cellulose, pectin	Co-expressed in cluster with pectinases. Highly expressed at 48 h on SBH.
5547	AA9-CBM1|LPMO	9	(hemi-)cellulose, pectin	Potential for oxidative pectin degradation.
3438	AA9|LPMO	11	Cellulose	Highly expressed at 24–48 h on SBH but not induced on CS.
8951	AA9|LPMO	11	Cellulose	Could be induced by D-mannose/mannan, improving the degradation of substrate rich in mannan.

[a] Corresponds to cluster in Figure S2, Supporting Information.

informed choice of the inducing substrate. CS appeared initially to require predominantly hemicellulases, while SBH required amylases and later degradation required a more diverse enzyme set, especially for SBH. These CAZyme inducing patterns matched well with the substrate composition of the two feedstocks, showing clear adaptation of the fungus to its substrate (Table S1, Supporting Information). For induction, where the choice of substrate and time affects the resulting enzyme cocktail, SBH resulted in higher PBD CAZy gene expression and diversity (24 and 48 h), which confirmed previous proteomics studies.[8,32] These strategies deserve deeper investigation, especially by comparing other crude substrates, in order to improve the efficiency of the enzyme cocktails.

4. Conclusion

This study shows that *P. anserina* adapts to different feedstocks, based on the substrate composition. From the results we can devise strategies to improve the efficiency of the enzyme cocktails, such as by adding novel activities (e.g., LPMOs or electron donors).

Abbreviations

ABF, α-L-arabinofuranosidase; AGD, α-1,4-glucosidase; AXE, acetyl xylan esterase; BGL, β-1,4-glucosidase; BXL, β-1,4-xylosidase; CBH, cellobiohydrolase; CDH, cellobiose dehydrogenase; CS, corn stover, EGL, β-1,4-endoglucanase; FAE, feruloyl esterase; LAC, β-1,4-galactosidase; LPMO, lytic polysaccharide monooxygenases; MAN, β-1,4-mannanase; MND, β-1,4-mannosidase; PBD, plant biomass degrading enzymes; SBH, soybean hulls; RHA, α-rhamnosidase; XG-EGL, xyloglucan β-1,4-endoglucanase; XLN, β-1,4-endoxylanase.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This study was supported by a grant of the 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. The work conducted by the U.S. Department of Energy Joint Genome Institute (JGI), was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. TB carried out the experiments, analyzed the data, and wrote the manuscript. MVAP performed part of the bioinformatics analysis. RPdV designed the study. RPdV and PD supervized the experiments and the manuscript writing. KL, MW, AL, VN, and IVG performed RNA sequencing. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no commercial or financial conflict of interest.

Keywords

CAZy, coprophilous fungi, lignocellulose degradation, lytic polysaccharide monooxygenases (LPMOs)

Received: April 10, 2018
Revised: September 10, 2018
Published online: October 8, 2018

[1] M. J. Richardson, *Fungal Divers.* **2002**, *10*, 101.
[2] M. Couturier, N. Tangthirasunun, X. Ning, S. Brun, V. Gautier, C. Bennati-Granier, P. Silar, J. G. Berrin, *Biotechnol. Adv.* **2016**, *34*, 976.
[3] E. Espagne, O. Lespinet, F. Malagnac, C. Da Silva, O. Jaillon, B. M. Porcel, A. Couloux, J. M. Aury, B. Segurens, J. Poulain, V. Anthouard, S. Grossetete, H. Khalili, E. Coppin, M. Dequard-Chablat, M. Picard, V. Contamine, S. Arnaise, A. Bourdais, V. Berteaux-Lecellier, D. Gautheret, R. P. de Vries, E. Battaglia,

P. M. Coutinho, E. G. Danchin, B. Henrissat, R. E. Khoury, A. Sainsard-Chanet, A. Boivin, B. Pinan-Lucarre, C. H. Sellem, R. Debuchy, P. Wincker, J. Weissenbach, P. Silar, *Genome Biol*. **2008**, *9*, R77.
[4] N. Xie, F. Chapeland-Leclerc, P. Silar, G. Ruprich-Robert, *Environ. Microbiol*. **2014**, *16*, 141.
[5] M. Paoletti, S. J. Saupe, *Genome Biol*. **2008**, *9*, 223.
[6] M. J. Richardson, *Mycol. Res*. **2001**, *105*, 387.
[7] P. Holter, *Ecol. Entomol*. **2016**, *41*, 367.
[8] M. R. Makela, O. Bouzid, D. Robl, H. Post, M. Peng, A. Heck, M. Altelaar, R. P. de Vries, *N. Biotechnol*. **2017**, *37*, 162.
[9] J. S. Yuan, K. H. Tiller, H. Al-Ahmad, N. R. Stewart, C. N. Stewart, Jr., *Trends Plant Sci*. **2008**, *13*, 421.
[10] A. U. Buranov, G. Mazza, *Ind. Crop Prod*. **2008**, *28*, 237.
[11] J. R. Mielenz, J. S. Bardsley, C. E. Wyman, *Bioresour. Technol*. **2009**, *100*, 3532.
[12] T. Benocci, R. P. de Vries, P. Daly, *J. Microbiol. Methods* **2018**, *146*, 33.
[13] S. Klaubauf, M. Zhou, M. H. Lebrun, R. P. de Vries, E. Battaglia, *FEBS Lett*. **2016**, *590*, 550.
[14] D. Kim, B. Langmead, S. L. Salzberg, *Nat. Methods* **2015**, *12*, 357.
[15] Y. Liao, G. K. Smyth, W. Shi, *Bioinformatics* **2014**, *30*, 923.
[16] M. I. Love, W. Huber, S. Anders, *Genome Biol*. **2014**, *15*, 550.
[17] L. Li, C. J. Stoeckert, Jr., D. S. Roos, *Genome Res*. **2003**, *13*, 2178.
[18] J. Boekhorst, B. Snel, *BMC Bioinformatics* **2007**, *8*, 356.
[19] C. Khosravi, T. Benocci, E. Battaglia, I. Benoit, R. P. de Vries, *Adv. Appl. Microbiol*. **2015**, *90*, 1.
[20] T. Benocci, M. V. Aguilar-Pontes, M. Zhou, B. Seiboth, R. P. de Vries, *Biotechnol. Biofuels* **2017**, *10*, 152.
[21] M. Peng, M. V. Aguilar-Pontes, M. Hainaut, B. Henrissat, K. Hilden, M. R. Makela, R. P. de Vries, *Fungal Genet. Biol*. **2017**, *112*, 40.
[22] A. A. Hasper, L. M. Trindade, D. van der Veen, A. J. van Ooyen, L. H. de Graaff, *Microbiology* **2004**, *150*, 1367.
[23] A. Lichius, V. Seidl-Seiboth, B. Seiboth, C. P. Kubicek, *Mol. Microbiol*. **2014**, *94*, 1162.
[24] A. Samal, J. P. Craig, S. T. Coradetti, J. P. Benz, J. A. Eddy, N. D. Price, N. L. Glass, *Biotechnol. Biofuels* **2017**, *10*, 225.
[25] K. S. Johansen, *Biochem. Soc. Trans*. **2016**, *44*, 143.
[26] G. Vaaje-Kolstad, Z. Forsberg, J. S. Loose, B. Bissaro, V. G. Eijsink, *Curr. Opin. Struct. Biol*. **2017**, *44*, 67.
[27] R. Berlemont, *Sci. Rep*. **2017**, *7*, 222.
[28] M. Fanuel, S. Garajova, D. Ropartz, N. McGregor, H. Brumer, H. Rogniaux, J. G. Berrin, *Biotechnol. Biofuels* **2017**, *10*, 63.
[29] N. Tangthirasunun, D. Navarro, S. Garajova, D. Chevret, L. C. Tong, V. Gautier, K. D. Hyde, P. Silar, J. G. Berrin, *Appl. Environ. Microbiol*. **2017**, *83*, e02716-16.
[30] S. Garajova, Y. Mathieu, M. R. Beccia, C. Bennati-Granier, F. Biaso, M. Fanuel, D. Ropartz, B. Guigliarelli, E. Record, H. Rogniaux, B. Henrissat, J. G. Berrin, *Sci. Rep*. **2016**, *6*, 28276.
[31] R. Tramontina, D. Robl, G. P. Maitan-Alfenas, *Biotechnol. J*. **2016**, *11*, 988.
[32] L. Poidevin, J. G. Berrin, C. Bennati-Granier, A. Levasseur, I. Herpoel-Gimbert, D. Chevret, P. M. Coutinho, B. Henrissat, S. Heiss-Blanquet, E. Record, *Appl. Microbiol. Biotechnol*. **2014**, *98*, 7457.

