Fungal Genetics and Biology 112 (2018) 40–46

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Regular Articles

Comparative analysis of basidiomycete transcriptomes reveals a core set of expressed genes encoding plant biomass degrading enzymes

Mao Peng[a], Maria Victoria Aguilar-Pontes[a], Matthieu Hainaut[b], Bernard Henrissat[b,c,d], Kristiina Hildén[e], Miia R. Mäkelä[a,e], Ronald P. de Vries[a,e,*]

[a] Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
[b] CNRS, Aix-Marseille Université, Marseille, France
[c] INRA, USC 1408 AFMB, 13288 Marseille, France
[d] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
[e] Department of Food and Environmental Sciences, University of Helsinki, Viikinkaari 9, Helsinki, Finland

ARTICLE INFO

Keywords:
Basidiomycete
Plant biomass degradation
Comparative transcriptomics
CAZymes

ABSTRACT

Basidiomycete fungi can degrade a wide range of plant biomass, including living and dead trees, forest litter, crops, and plant matter in soils. Understanding the process of plant biomass decay by basidiomycetes could facilitate their application in various industrial sectors such as food & feed, detergents and biofuels, and also provide new insights into their essential biological role in the global carbon cycle. The fast expansion of basidiomycete genomic and functional genomics data (e.g. transcriptomics, proteomics) has facilitated exploration of key genes and regulatory mechanisms of plant biomass degradation. In this study, we comparatively analyzed 22 transcriptome datasets from basidiomycetes related to plant biomass degradation, and identified 328 commonly induced genes and 318 repressed genes, and defined a core set of carbohydrate active enzymes (CAZymes), which was shared by most of the basidiomycete species. High conservation of these CAZymes in genomes and similar regulation pattern in transcriptomics data from lignocellulosic substrates indicate their key role in plant biomass degradation and need for their further biochemical investigation.

1. Introduction

Plant biomass is the most abundant renewable carbon source on earth and a substrate for the majority of fungal species. Basidiomycete fungi comprise 32% of the described fungi (Kirk et al., 2008). Many of them can efficiently use plant biomass as carbon source by degrading plant cell wall polysaccharides via the production of various carbohydrate active enzymes (CAZymes) (Lombard et al., 2014). Wood-decaying basidiomycetes have traditionally been classified as white rot and brown rot species, based on differences in their ability to degrade the major components of plant cell wall. White rot fungi are able to degrade both lignin and polysaccharides (cellulose and hemicellulose), while brown rot fungi efficiently degrade polysaccharides, but only modify lignin.

Recent advances in high-throughput “omics” technologies have greatly enhanced our understanding of the complex biomechanisms involved in lignocellulose degradation by basidiomycetes (Ohm et al., 2014). With the accumulation of an increasing amount of sequence data, comparative genomics of different fungi has become an effective way to reveal evolutionary patterns as well as the potential function of interesting genes. The comparison of 33 sequenced basidiomycete genomes revealed that the current dichotomous classification of white rot and brown rot fungi is not adequate to describe the whole spectrum of the wood-decaying basidiomycetes (Riley et al., 2014). For instance, *Botryobasidium botryosum* and *Jaapia argillacea* have lost ligninolytic class II peroxidases like brown rot fungi, but possess diverse enzymes acting on crystalline cellulose normally associated with white rot fungi, and thus represent so-called grey rot species. In addition, the fast expansion of basidiomycete functional genomics data (e.g. transcriptomics, proteomics) plays a key role in exploring and validating new functions of plant biomass degradation related genes and the regulatory mechanisms driving their expression (Gaskell et al., 2016; Vanden Wymelenberg et al., 2009).

Meta-analysis of transcriptomic datasets has been a powerful approach to discover new gene expression patterns that are difficult to detect in individual experiments. By combining multiple studies that share a similar experimental setup their statistical power can be increased to reveal a more precise and reliable set of differentially

* Corresponding author at: P.O. Box 85167, 3508 AD Utrecht, The Netherlands.
E-mail address: r.devries@westerdijkinstitute.nl (R.P. de Vries).

http://dx.doi.org/10.1016/j.fgb.2017.08.001
Received 17 February 2017; Received in revised form 25 July 2017; Accepted 7 August 2017
Available online 10 August 2017
1087-1845/ © 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

expressed genes (Ramasamy et al., 2008). Moreover, cross species integration and comparisons has already been used to identify evolutionarily conserved genes reflecting adaptation of organisms to specific environmental changes (Shaar-Moshe et al., 2015). In animal and plant studies, meta-analysis of transcriptome datasets has been extensively used for finding sets of key genes involved in specific biological processes (Mustroph et al., 2010) and disease mechanisms (Rhodes et al., 2004). However, very few meta-analysis of multiple transcriptome studies on basidiomycetes have been reported that go beyond two or three species. The increasing number of fungal plant biomass degradation related transcriptome sets provides options for integration and comparison of these transcriptomes of different basidiomycetes to promote detection of a core set of conserved genes involved in lignocellulose degradation.

In this study, we collected a large set of transcriptome datasets associated with plant biomass degradation, which consists of 22 experiments across 10 basidiomycete species. We aimed to evaluate whether it is possible to compare these heterogeneous studies and still obtain reliable conclusions or leads on fungal lignocellulosic biomass degradation, meaning the identification of a core set of consistently regulated genes/pathways with significant statistical p-values and important biochemical function that could be discovered via integrative analysis of previously published heterogeneous datasets. Through comparing gene expression profiles of these fungi growing on complex lignocellulose media and simple carbon sources, we identified a core set of up-regulated genes that are shared among these basidiomycetes in response to plant biomass degradation (Suppl. Fig. S1). In addition, our study revealed that many currently documented CAZy families contain more than one orthologous groups, and significantly different expression patterns among these orthologs were observed during lignocellulose degradation.

2. Materials and methods

2.1. Transcriptomic data collection and identification of differentially expressed genes

To collect a large transcriptome set related with plant biomass degradation of basidiomycetes, we searched the NCBI GEO database (Edgar et al., 2002) (https://www.ncbi.nlm.nih.gov/geo/) with key words basidiomycete and plant biomass degradation. The matched datasets where manually filtered to select the experiments which shared a similar research focus of comparing fungi grown on complex lignocellulose and simple carbon sources, with at least 2 replicates for each condition. In total, 22 transcriptome datasets from 10 basidiomycete species (including 14 datasets for 6 white rot species, 7 for 3 brown rot species and 1 for litter-degrading species) were collected (Table 1). The lignocellulose substrates varied from ball-milled aspen and pine sapwood to cellulose and lignin. The simple carbon sources included glucose- and maltose-amended Melin-Norkrans (MMN) agar medium, Highley's basal medium, and malt extract agar medium. The experimental platforms included Roche-NimbleGen microarray, Agilent and Illumina RNA-seq (Suppl. Table S1).

Due to the heterogeneity of the experimental platforms, the preprocessed gene expression profiling data was downloaded from the Series Matrix File of the GEO database. All the collected gene expression data had been processed with background correction, quality control and normalization in their original studies and was ready for further statistical analysis. For each dataset, samples from fungi grown on lignocellulose were compared with their control samples which contained glucose or maltose as carbon source. Student *t*-test (p-value ≤ 0.05) was applied to determine differentially expressed genes (DEGs) between lignocellulose medium and glucose- or maltose-amended medium. For the orthologous genes which have two or more copies in a specific species, only under the situation that all of the gene copies in a comparison experiment showed the same up- or down-regulated patterns, the corresponding ortholog was considered as DEG.

2.2. Identification of orthologous genes across species

To compare the transcriptomic response related to plant biomass degradation across different basidiomycetes, it is necessary to identify the evolutionarily conserved orthologous genes among them. To accomplish this, the predicted proteome of each basidiomycete was downloaded from the JGI MycoCosm database (Grigoriev et al., 2014) and used for an all-against-all BLAST comparison with subsequent clustering of orthologous genes. The OrthoMCL method was used to infer orthologous genes across species (Li et al., 2003) with parameters set as: E-value $1E^{-5}$, inflation level 1.5 and sequence coverage 60%. The OrthoMCL algorithm generates clusters of proteins where each cluster consists of orthologs or “recent” paralogs across multiple fungal species. The proteins belonging to the same cluster have putatively evolved from a common ancestor by speciation and share similar biological functions. Orthologous clusters containing at least one detected gene among the collected datasets were extracted. The enrichment of induced and repressed genes in each specific orthologous group were examined by hypergeometric distribution analysis as applied in a previous study (Mustroph et al., 2010). All the expressed genes were combined into a matrix according to their orthologous relationships. The relative expression values of each ortholog in each dataset were calculated as fold change between complex lignocellulose containing and glucose- or maltose-amended medium. To visualize gene expression from heterogeneous platforms and make diverse dataset comparable, gene expression features in each species were binned into the following four categories: repressed, induced, absent and no-regulation and labeled respectively in different colors in the heatmap (Figs. 2and 3 and

Table 1
Datasets included in the meta-analysis.

Species	Rot-type	Experimental design	Platform	Transcriptome size	GEO ID	References
Agaricus bisporus	Litter	Wheat straw VS MMN	Agilent microarray	10,413	GSE39569	Morin et al. (2012)
Phlebiopsis gigantea	White	Pine wood VS glucose	Illumina RNA-seq	11,892	GSE53112	Hori et al. (2014)
Phanerochaete chrysosporium	White	Cellulose, ball-milled aspen VS glucose	Roche-NimbleGen	9201	GSE14736	Vanden Wymelenberg et al. (2009)
Gelatoporia subvermispora	White	Cellulose, ball-milled aspen VS glucose	Roche-NimbleGen	12,096	GSE34636	Fernandez-Fueyo et al. (2012)
Heterobasidion annosum	White	Wood VS malt extract agar medium amended with glucose	Roche-NimbleGen	7248	GSE39805	Raffaello et al. (2014)
Heterobasidion irregulare	White	Lignin, wood shavings, cellulose VS MMN medium	Roche-NimbleGen	7248	GSE30230	Olson et al. (2012)
Pycnoporus coccineus	White	Aspen, pine VS maltose	Illumina RNA-seq	12,690	GSE74234	Couturier et al. (2015)
Serpula lacrymans	Brown	Wood VS MMN medium	Roche-NimbleGen	12,797	GSE27839	Eastwood et al. (2011)
Postia placenta	Brown	Cellulose, ball-milled aspen VS glucose	Roche-NimbleGen	12,438	GSE12540	Martinez et al. (2009)
Wolfiporia cocos	Brown	Aspen, cellulose, pine VS glucose	Illumina RNA-seq	12,748	GSE78007	Gaskell et al. (2016)

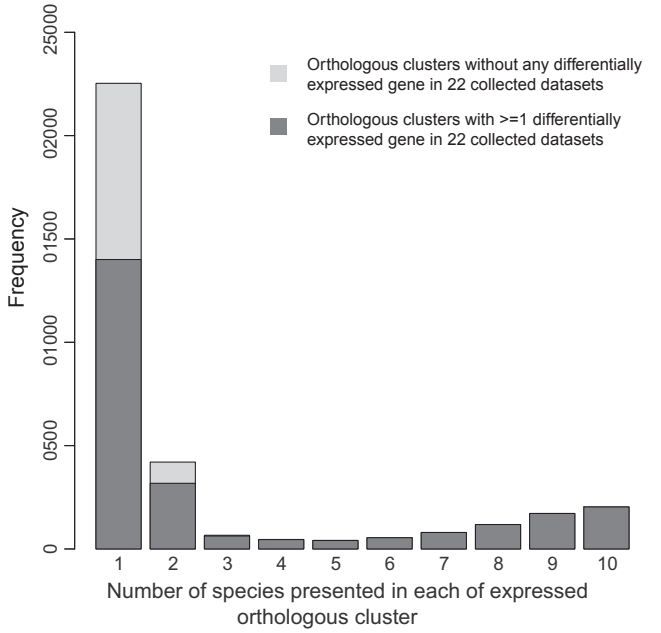

Fig. 1. The distribution of the expressed orthologous clusters among the 10 different basidiomycete species studied. In total 34,548 orthologous clusters were collected in this study. Among them, 24,950 orthologs were differentially expressed in at least one dataset (highlighted in dark grey), and 9598 orthologs were not significantly changed in any of dataset (highlighted in light grey).

Suppl. Fig. S2). Since many orthologous groups contained more than one regulated gene in specific species, we labeled them in a bright color to separate them from the orthologous groups with a single regulated gene.

2.3. Gene ontology and function catalogue annotation

The Gene ontology annotation of consistently induced and repressed orthologous gene clusters were analyzed by BLAST2GO (version 4.0) (Gotz et al., 2008) based on the best BLASTP hits (with default threshold of E value $\leq 1e-3$ and HSP length of 33) to basidiomycete proteins in the NCBI RefSeq non-redundant database. Since most orthologous clusters contain multiple genes from different species, we randomly picked one gene from each orthologous cluster and annotated its function. 88% of the induced genes (289 out of 328) and 87% of the repressed genes (277 out of 318 genes) have an assigned GO annotation (Suppl. Table S3A and S3B). The GO molecular function terms (level 4) with more than five genes in either induced or repressed gene list were shown on the bar-plot of Fig. 2.

In addition, the FunCat annotation available in the FunCat database (Ruepp et al., 2004) and Pedant3 database (Frishman et al., 2003) was also extracted to provide further gene functional information. Because of the limited number of basidiomycete species that have been well annotated by the FunCat team, only 38% (246 out 646) of consistently regulated genes have mapped with FunCat.

2.4. CAZymes identification

The predicted proteome of 10 basidiomycete species was analyzed by the CAZy annotation pipeline and assigned to specific carbohydrate-active enzyme families (Lombard et al., 2014). For the orthologous groups with significantly regulated CAZymes, their substrates were mapped and summarized according to the enzyme-substrate relation provided by the CAZy database (http://www.cazy.org/).

3. Results

3.1. Identification of orthologous genes across diverse basidiomycetes

OrthoMCL analysis resolved 34,548 orthologous clusters which contained at least one gene expression value in any of the 22 transcriptomics datasets across 10 studied basidiomycetes. Since our study focused on commonly regulated genes across diverse species, orthologous clusters which were not differentially expressed in any of the collected datasets were excluded from further analysis. In total, 24,950 orthologous clusters which had a differentially expressed gene in at least one dataset were used for statistical analysis. Of these, 56% (14,005) orthologs had expression values in only one basidiomycete species, and 8% (2034) orthologs had expression values in all studied basidiomycete species. Fig. 1 shows the number of organisms distributed in each orthologous cluster which indicates the high specificity of species, substrate and experimental setup, of fungal transcriptomes. Among the above 24,950 studied orthologs, only 2% (392 genes) of the transcriptome was annotated as CAZymes. The orthologous gene information and the related gene expression values are summarized in Suppl. Table S2.

3.2. Commonly regulated genes in response to plant biomass degradation

To detect commonly regulated genes in response to plant biomass degradation across diverse basidiomycetes, 44,251 induced and 39,770 reduced DEGs from all collected experimental datasets were organized to their corresponding orthologous clusters. The enrichment of regulated genes in specific orthologous groups compared to total regulated genes ratio in all detected genes was tested by hypergeometric distribution analysis (Suppl. Table S2). Using the enrichment p-value ≤ 0.01 and presence in more than half of all studied organisms and half of total datasets as criteria, we identified 328 induced genes (Suppl. Table S3A) and 318 repressed genes (Suppl. Table S3B), which are defined as the consistently differentially expressed genes in response to plant biomass substrates across the tested basidiomycetes (Fig. 2).

Gene Ontology (GO) analysis of 328 consistently up-regulated genes showed that the top enrichment molecular function categories (level 4) include genes encoding hydrolases acting on glycosidic bonds (53 genes), cation binding proteins (28 genes), substrate-specific transmembrane proteins (18 genes), coenzyme binding proteins (15 genes), polysaccharide binding proteins (12 genes), oxidoreductases acting on CH-OH groups (7 genes), oxidoreductases acting on CH-CH groups (5 genes), peptidases (5 genes) and phosphorus transferases (6 genes) (Fig. 2 and Table S3A). In contrast, the GO molecular function terms for consistently repressed genes contain much less genes encoding hydrolases active on glycosidic bonds (3 genes), oxidoreductases acting on CH-CH groups (0 genes), and coenzyme binding proteins (3 genes), but higher numbers of genes encoding peptidases (10 genes), nucleotide binding proteins (45 genes), and phosphorus transferases (16 genes) (Fig. 2 and Suppl. Table S3B). The commonly induced/repressed gene sets revealed in our study indicate that several core functional gene groups have been adopted by a broad range of basidiomycetes for degradation of plant biomass.

3.3. Consistently up-regulated CAZymes

With the special interest in the up-regulated CAZyme encoding genes, we further focused on the functional analysis of these plant biomass degradation related genes. To recognize commonly expressed CAZymes across basidiomycetes, the orthologous groups that were significantly enriched with up-regulated CAZymes were separated from the above identified commonly regulated gene list. In total, 23% (75 out of 328 genes) consistently up-regulated genes encoded CAZymes (Suppl. Table S3A). Compared to 2% (392) CAZymes in all collected 24,950 orthologous clusters (Suppl. Table S2), there was a 12-fold

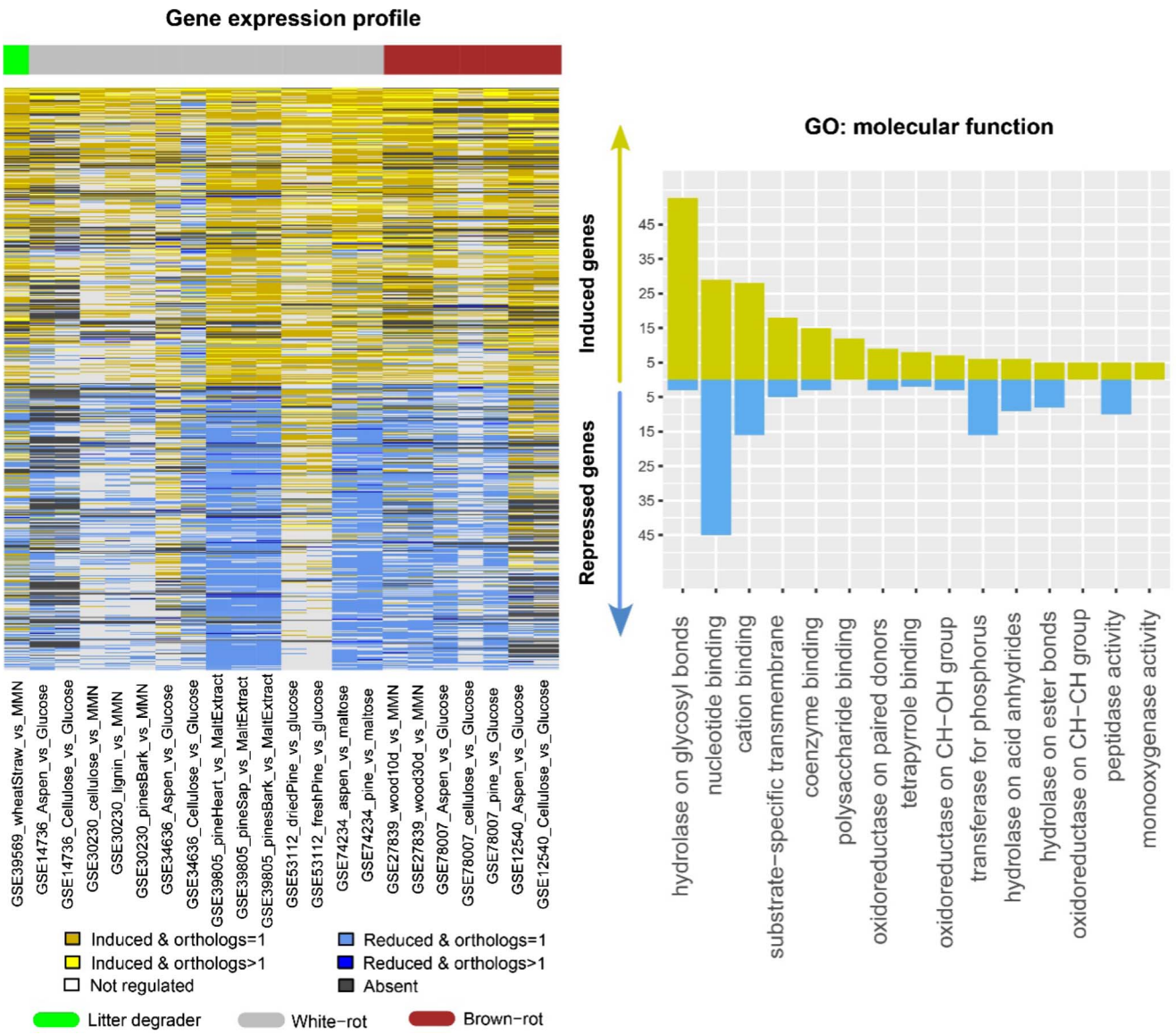

Fig. 2. Commonly up/down-regulated genes and their enrichment of GO molecular function.

higher enrichment of CAZymes in the consistently up-regulated gene set, which indicates the high specificity of our method to detect key genes from large heterogeneous datasets.

By including enzyme-substrate information provided by CAZyme database (http://www.cazy.org/), we defined a core set of 50 CAZymes from the above 328 consistently induced CAZymes which are related to degradation of plant biomass. Their plant biomass related substrate, and expression patterns in the studied basidiomycetes are summarized in Fig. 3 and Suppl. Table S3C. Most of these conserved induced genes encode enzymes involved in degradation of polymeric substrates, such as cellulose (16 genes), xylan (8 genes) and pectin (11 genes). The high genomic conservation and similar up-regulated pattern in lignocellulose transcriptomics of this core set of genes indicates their key role in plant biomass degradation.

Interestingly, the data also revealed significant differences between the species, as not all CAZymes conserved in the basidiomycete genomes have a similar regulation pattern related to plant biomass degradation, which can be seen for example in families GH45, GH125, and AA1_2 (Suppl. Table S2). Also, the orthologous clusters from the same CAZy family may present clearly different expression patterns across species. For instance, orthologs of GH5, GH3 and AA9 were revealed to have different expression patterns (Suppl. Fig. S2).

4. Discussion

Recent advance in genome sequencing and comparative genomics studies have revealed that different basidiomycetes contains a different set of potential CAZymes for adaptation to their unique natural niches (Martinez et al., 2009; Ohm et al., 2010; Rytioja et al., 2014; Shah et al., 2016). However, the comparative studies of large-scale functional genomics data are still in their infancy. By comparing a large set of transcriptomic data, our study provided new insights into plant biomass degradation by basidiomycetes. Clear differences in the expression patterns of orthologous genes in different species and plant biomass substrates were revealed. This demonstrates that the integration of transcriptome data can provide a more comprehensive and dynamic picture of how genes are regulated in specific conditions compared to static gene copy comparisons in comparative genomics studies. Subfamilies for GH5 have been previously defined with specific enzyme activities and annotated structure features (Aspeborg et al., 2012). Subfamily classifications have also been described for CAZy families GH13 (Stam et al., 2006) and GH43 (Mewis et al., 2016), but not yet for other CAZy families that also contain multiple activities, such as GH3 and AA9 (Suppl. Fig. S2). The diverse expression patterns of the members of these families supports the need for further sub-classification and

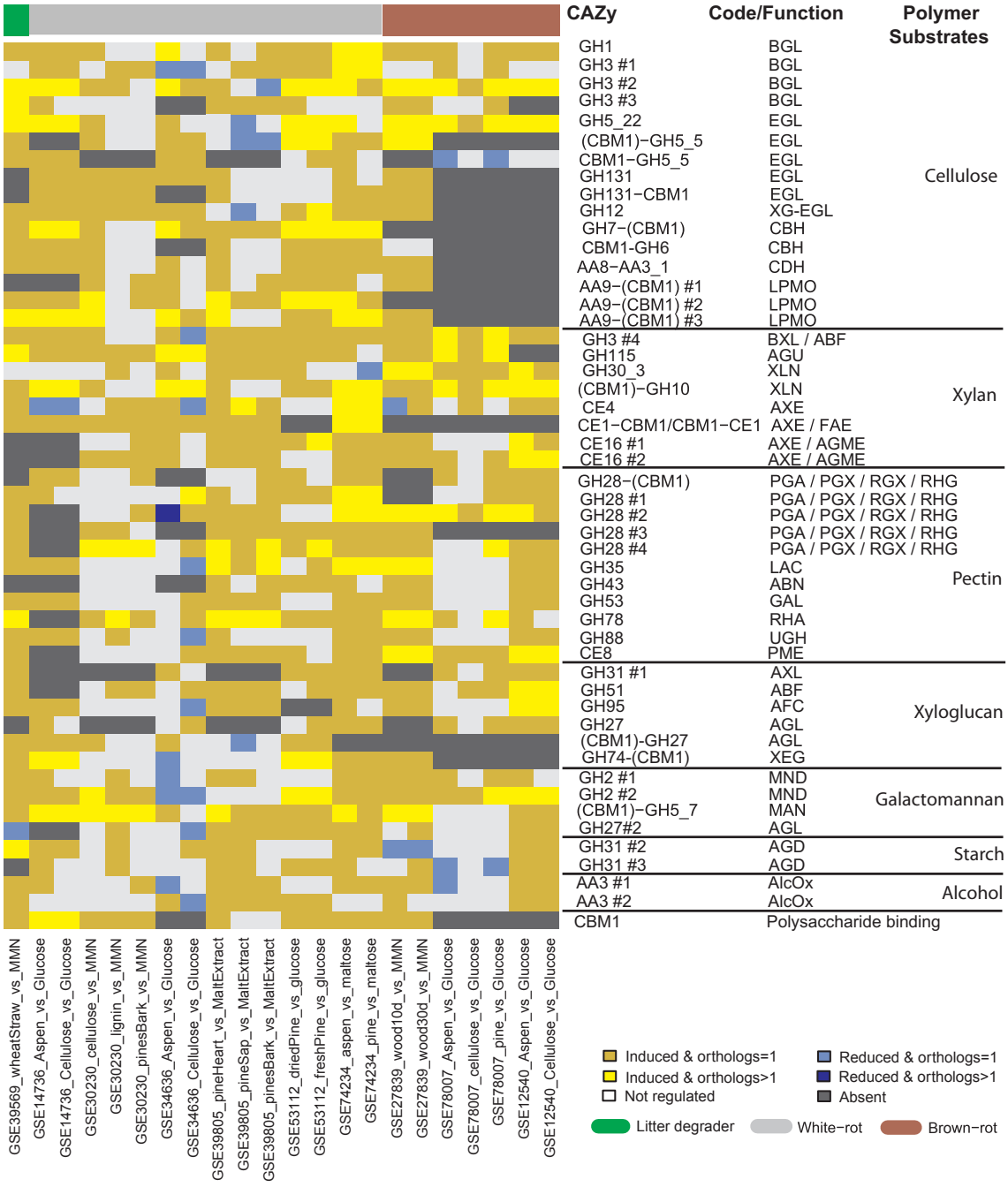

Fig. 3. Core set of plant biomass degradation related genes identified in this study. Abbreviations: ABF, α-arabinofuranosidase; ABN, endoarabinanase; AXE, acetyl xylan esterase; AGME, acetyl glucomannan esterase; AFC, α-fucosidase; AGD, α-1,4-glucosidase; AGL, α-1,4-D-galactosidase; AGU, α-glucuronidase; AXL, α-xylosidase; BGL, β-1,4-glucosidase; BXL, β-1,4-xylosidase; CBH, cellobiohydrolase; CDH, cellobiose dehydrogenase; AlcOx, alcohol oxidase; EGL, β-1,4-endoglucanase; FAE, feruloyl esterase; GAL, β-1,4-endogalactanase; LAC, β-1,4-galactosidase; LPMO, lytic polysaccharide mono-oxygenase; MAN, β-1,4-endomannanase; MND, β-1,4-D-mannosidase; PGA, endo-polygalacturonase; PGX, exo-polygalacturonase; PME, pectin methyl esterase; RGX, rhamnogalacturonan galacturonohydrolase/exorhamnogalacturonase; RHA, α-rhamnosidase; RHG, rhamnogalacturonan hydrolase/endorhamnogalacturonase; XEG, xyloglucan β-1,4-endoglucanase; XG-EGL, xyloglucanase; XLN, β-1,4-endoxylanase; UGH, d-4,5-unsaturated glucuronyl hydrolase. The CBM1 in the brackets indicate that some members in the orthologous groups may not contain the CBM1 domain. The '#' was added to several CAZymes for indicating that they have been matched by multiple orthologous clusters.

functional characterization of CAZy families.

Besides meta-analysis of data from one single "omics" platform, the integration of heterogeneous and large "omics" data from different experimental platforms and molecular levels (genomics, transcriptomics, proteomics and metabolomics) is the next big challenge of today's biological studies (Chen et al., 2012; Gomez-Cabrero et al., 2014; Miyauchi et al., 2016; Zhang et al., 2010). Studies have been performed to identify key genes involved in plant cell wall degradation in e.g. *Phanerochaete chrysosporium* (Vanden Wymelenberg et al., 2009), *Neurospora crassa* (Tian et al., 2009), *Dichomitus squalens* (Rytioja et al., 2016) and *Agaricus bisporus* (Patyshakuliyeva et al., 2015), by combining evidence from proteomics and transcriptomics data. Now that more fungal functional genomics data associated with plant biomass degradation is accumulating, a deeper understanding of lignocellulose degradation could be achieved via an integrated analysis of "omics" data from these different species and platforms.

The 50 CAZymes involved in plant biomass degradation for which the corresponding genes were considered to be consistently upregulated

in the presence of plant biomass, cover a range of activities acting on various plant polymers. However, when these are ranked based on the number of datasets they are upregulated in, an interesting pattern emerges. Only three CAZyme orthologous groups are upregulated in 19 of the 22 datasets and two in 18 of the 22 datasets (Suppl. Table S3C). The first three encode a putative β-glucosidase, β-xylosidase, and endoxylanase, while the last two encode a putative endomannanase and α-glucuronidase. The first two are crucial enzymes for the complete hydrolysis of cellulose and xylan, while the third and fourth are highly important enzymes for degradation of xylan and mannan, respectively, by hydrolyzing the backbone of these polysaccharides into oligosaccharides (Rytioja et al., 2014). Surprisingly, the latter, α-glucuronidase, is an enzyme that removes the 4-O-methyl-glucuronic acid side groups from xylan thus providing access for endoxylanases to the backbone of xylan (Chong et al., 2011). The consistent upregulation of this gene suggests that the tested basidiomycetes all commonly encounter glucuronoxylans and that 4-O-methyl-glucuronic acid side groups significantly hinder xylan hydrolysis by these fungi.

The next set of most prevalently upregulated genes in the presence of plant biomass are mainly a set of cellulases, followed by a diverse number of (hemi-)cellulases and pectinases (Fig. 3). The presence of the cellulases is expected as a high amount of cellulose is present in all plant biomass substrates. However, the relative amount of pectin in plant biomass is highly variable, from highly abundant in some dicots (e.g. soy beans) to very low (e.g. in wood) (Edwards and Doran-Peterson, 2012). The common upregulation of a basic set of pectinases in most of the datasets (even on substrates with low pectin content), suggests that the presence of pectin is a significant barrier for enzymes targeting other plant polymers. Pectin has been reported to act as a glue in the cell wall, linking hemicellulose, lignin and cellulose polymers into a tight matrix and determining cell growth (Anderson, 2016), which may explain why the (partial) degradation of pectin is a core component of plant biomass degrading strategies in fungi.

The observation that not all orthologous genes targeting a certain polysaccharide (e.g. cellulose) are consistently upregulated in different fungal species or on different substrates indicates the diversity of plant biomass degradation approaches of fungi. While the substrate-related differences could be caused by structural differences in the plant polymers, the species-specific differences more likely relate to evolutionary adaptations of the species. This could be due to differences in their biotope, which would also involve different plant biomass compositions, but could also hint at the specific components of the biomass these species primarily target.

5. Conclusions

In this study, transcriptomic profiling of plant biomass degradation in 10 basidiomycetes was systematically evaluated to identify consistently regulated genes. Despite the variations in growth conditions and species, a core set of commonly up-regulated plant biomass degradation related genes was discovered. Further investigation of this core gene-set could help to determine their specific biological function and regulatory mechanisms during the process of plant biomass degradation. Also, a more extensive study including a larger dataset of more fungal species would likely result in an even more robust analysis and be able to distinguish between basic core genes and genes that are consistently upregulated on specific plant biomass substrates.

While our study focused on consistent patterns between fungi and datasets, this approach could also be used to identify significantly different patterns in either different datasets from the same species, or in different species grown under similar conditions.

Funding

The comparative genomics study was carried out on the Dutch national e-infrastructure with the support of SURF Foundation (e-infra140010). MVAP 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. BH and MH were funded by ANR grant N°21000602 (Funtune).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fgb.2017.08.001.

References

Anderson, C.T., 2016. We be jammin': an update on pectin biosynthesis, trafficking and dynamics. J. Exp. Bot. 67, 495–502.
Aspeborg, H., Coutinho, P.M., Wang, Y., Brumer 3rd, H., Henrissat, B., 2012. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol. Biol. 12, 186.
Chen, R., Mias, G.I., Li-Pook-Than, J., Jiang, L., Lam, H.Y., Chen, R., Miriami, E., Karczewski, K.J., Hariharan, M., Dewey, F.E., Cheng, Y., Clark, M.J., Im, H., Habegger, L., Balasubramanian, S., O'Huallachain, M., Dudley, J.T., Hillenmeyer, S., Haraksingh, R., Sharon, D., et al., 2012. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 148, 1293–1307.
Chong, S.L., Battaglia, E., Coutinho, P.M., Henrissat, B., Tenkanen, M., de Vries, R.P., 2011. The alpha-glucuronidase Agu1 from *Schizophyllum commune* is a member of a novel glycoside hydrolase family (GH115). Appl. Microbiol. Biotechnol. 90, 1323–1332.
Couturier, M., Navarro, D., Chevret, D., Henrissat, B., Piumi, F., Ruiz-Duenas, F.J., Martinez, A.T., Grigoriev, I.V., Riley, R., Lipzen, A., Berrin, J.G., Master, E.R., Rosso, M.N., 2015. Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. Biotechnol. Biofuels 8, 216.
Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., Asiegbu, F.O., Baker, S.E., Barry, K., Bendiksby, M., Blumentritt, M., Coutinho, P.M., Cullen, D., de Vries, R.P., Gathman, A., Goodell, B., Henrissat, B., Ihrmark, K., Kauserud, H., Kohler, A., et al., 2011. The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science 333, 762–765.
Edgar, R., Domrachev, M., Lash, A.E., 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30, 207–210.
Edwards, M.C., Doran-Peterson, J., 2012. Pectin-rich biomass as feedstock for fuel ethanol production. Appl. Microbiol. Biotechnol. 95, 565–575.
Fernandez-Fueyo, E., Ruiz-Duenas, F.J., Ferreira, P., Floudas, D., Hibbett, D.S., Canessa, P., Larrondo, L.F., James, T.Y., Seelenfreund, D., Lobos, S., Polanco, R., Tello, M., Honda, Y., Watanabe, T., Watanabe, T., Ryu, J.S., Kubicek, C.P., Schmoll, M., Gaskell, J., Hammel, K.E., et al., 2012. Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc. Natl. Acad. Sci. U.S.A. 109, 5458–5463.
Frishman, D., Mokrejs, M., Kosykh, D., Kastenmuller, G., Kolesov, G., Zubrzycki, I., Gruber, C., Geier, B., Kaps, A., Albermann, K., Volz, A., Wagner, C., Fellenberg, M., Heumann, K., Mewes, H.W., 2003. The PEDANT genome database. Nucleic Acids Res. 31, 207–211.
Gaskell, J., Blanchette, R.A., Stewart, P.E., BonDurant, S.S., Adams, M., Sabat, G., Kersten, P., Cullen, D., 2016. Transcriptome and secretome analyses of the wood decay fungus *Wolfiporia cocos* support alternative mechanisms of lignocellulose conversion. Appl. Environ. Microbiol. 82, 3979–3987.
Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkenschlager, M., Gisel, A., Ballestar, E., Bongcam-Rudloff, E., Conesa, A., Tegner, J., 2014. Data integration in the era of omics: current and future challenges. BMC Syst. Biol. 8 (Suppl. 2), I1.
Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36, 3420–3435.
Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res. 42, D699–D704.
Hori, C., Ishida, T., Igarashi, K., Samejima, M., Suzuki, H., Master, E., Ferreira, P., Ruiz-Dueñas, F.J., Held, B., Canessa, P., Larrondo, L.F., Schmoll, M., Druzhinina, I.S., Kubicek, C.P., Gaskell, J.A., Kersten, P., et al., 2014. Analysis of the *Phlebiopsis gigantea* genome, transcriptome and secretome provides insight into its pioneer colonization strategies of wood. PLoS Genet. 10, e1004759.
Kirk, P.M., Cannon, P.F., Minter, D.W., Stalpers, J.A., 2008. Dictionary of the Fungi, 10th ed. CAB International, Wallingford, UK.
Li, L., Stoeckert Jr., C.J., Roos, D.S., 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13, 2178–2189.
Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490–D495.
Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C.P., Ferreira, P., Ruiz-Duenas, F.J., Martinez, A.T., Kersten, P., Hammel, K.E., Vanden Wymelenberg, A., Gaskell, J., Lindquist, E., Sabat, G., Bondurant, S.S., Larrondo, L.F., Canessa, P., Vicuna, R., Yadav, J., et al., 2009. Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc. Natl. Acad. Sci. U.S.A. 106, 1954–1959.

Mewis, K., Lenfant, N., Lombard, V., Henrissat, B., 2016. Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. Appl. Environ. Microbiol. 82, 1686–1692.
Miyauchi, S., Navarro, D., Grigoriev, I.V., Lipzen, A., Riley, R., Chevret, D., Grisel, S., Berrin, J.G., Henrissat, B., Rosso, M.N., 2016. Visual comparative omics of fungi for plant biomass deconstruction. Front. Microbiol. 7, 1335.
Morin, E., Kohler, A., Baker, A.R., Foulongne-Oriol, M., Lombard, V., Nagy, L.G., Ohm, R.A., Patyshakuliyeva, A., Brun, A., Aerts, A.L., Bailey, A.M., Billette, C., Coutinho, P.M., Deakin, G., Doddapaneni, H., Floudas, D., Grimwood, J., Hilden, K., Kues, U., Labutti, K.M., et al., 2012. Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. Proc. Natl. Acad. Sci. U.S.A. 109, 17501–17506.
Mustroph, A., Lee, S.C., Oosumi, T., Zanetti, M.E., Yang, H., Ma, K., Yaghoubi-Masihi, A., Fukao, T., Bailey-Serres, J., 2010. Cross-kingdom comparison of transcriptomic adjustments to low-oxygen stress highlights conserved and plant-specific responses. Plant Physiol. 152, 1484–1500.
Ohm, R.A., de Jong, J.F., Lugones, L.G., Aerts, A., Kothe, E., Stajich, J.E., de Vries, R.P., Record, E., Levasseur, A., Baker, S.E., Bartholomew, K.A., Coutinho, P.M., Erdmann, S., Fowler, T.J., Gathman, A.C., Lombard, V., Henrissat, B., Knabe, N., Kues, U., Lilly, W.W., et al., 2010. Genome sequence of the model mushroom *Schizophyllum commune*. Nat. Biotechnol. 28, 957–963.
Ohm, R.A., Riley, R., Salamov, A., Min, B., Choi, I.G., Grigoriev, I.V., 2014. Genomics of wood-degrading fungi. Fungal Genet. Biol. 72, 82–90.
Olson, A., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., Canback, B., Coutinho, P.M., Cullen, D., Dalman, K., Deflorio, G., van Diepen, L.T., Dunand, C., Duplessis, S., Durling, M., Gonthier, P., Grimwood, J., Fossdal, C.G., Hansson, D., Henrissat, B., et al., 2012. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol. 194, 1001–1013.
Patyshakuliyeva, A., Post, H., Zhou, M., Jurak, E., Heck, A.J., Hildén, K.S., Kabel, M.A., Mäkelä, M.R., Altelaar, M.A., de Vries, R.P., 2015. Uncovering the abilities of *Agaricus bisporus* to degrade plant biomass throughout its life cycle. Environ. Microbiol. 17, 3098–3109.
Raffaello, T., Chen, H., Kohler, A., Asiegbu, F.O., 2014. Transcriptomic profiles of *Heterobasidion annosum* under abiotic stresses and during saprotrophic growth in bark, sapwood and heartwood. Environ. Microbiol. 16, 1654–1667.
Ramasamy, A., Mondry, A., Holmes, C.C., Altman, D.G., 2008. Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Med. 5, e184.
Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., Chinnaiyan, A.M., 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proc. Natl. Acad. Sci. U.S.A. 101, 9309–9314.
Riley, R., Salamov, A.A., Brown, D.W., Nagy, L.G., Floudas, D., Held, B.W., Levasseur, A., Lombard, V., Morin, E., Otillar, R., Lindquist, E.A., Sun, H., LaButti, K.M., Schmutz, J., Jabbour, D., Luo, H., Baker, S.E., Pisabarro, A.G., Walton, J.D., Blanchette, R.A., et al., 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc. Natl. Acad. Sci. U.S.A. 111, 9923–9928.
Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M., Mewes, H.W., 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 32, 5539–5545.
Rytioja, J., Hilden, K., Di Falco, M., Zhou, M., Aguilar-Pontes, M.V., Sietio, O.M., Tsang, A., de Vries, R.P., Makela, M.R., 2016. The molecular response of the white-rot fungus *Dichomitus squalens* to wood and non-woody biomass as examined by transcriptome and exoproteome analyses. Environ. Microbiol. 19, 1237–1250.
Rytioja, J., Hilden, K., Yuzon, J., Hatakka, A., de Vries, R.P., Makela, M.R., 2014. Plant-polysaccharide-degrading enzymes from Basidiomycetes. Microbiol. Mol. Biol. Rev. 78, 614–649.
Shaar-Moshe, L., Hubner, S., Peleg, Z., 2015. Identification of conserved drought-adaptive genes using a cross-species meta-analysis approach. BMC Plant Biol. 15, 111.
Shah, F., Nicolas, C., Bentzer, J., Ellstrom, M., Smits, M., Rineau, F., Canback, B., Floudas, D., Carleer, R., Lackner, G., Braesel, J., Hoffmeister, D., Henrissat, B., Ahren, D., Johansson, T., Hibbett, D.S., Martin, F., Persson, P., Tunlid, A., 2016. Ectomycorrhizal fungi decompose soil organic matter using oxidative mechanisms adapted from saprotrophic ancestors. New Phytol. 209, 1705–1719.
Stam, M.R., Danchin, E.G., Rancurel, C., Coutinho, P.M., Henrissat, B., 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. Protein Eng. Des. Sel. 19, 555–562.
Tian, C., Beeson, W.T., Iavarone, A.T., Sun, J., Marletta, M.A., Cate, J.H., Glass, N.L., 2009. Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. Proc. Natl. Acad. Sci. U.S.A. 106, 22157–22162.
Vanden Wymelenberg, A., Gaskell, J., Mozuch, M., Kersten, P., Sabat, G., Martinez, D., Cullen, D., 2009. Transcriptome and secretome analyses of *Phanerochaete chrysosporium* reveal complex patterns of gene expression. Appl. Environ. Microbiol. 75, 4058–4068.
Zhang, W., Li, F., Nie, L., 2010. Integrating multiple 'omics' analysis for microbial biology: application and methodologies. Microbiology 156, 287–301.