Samal *et al. Biotechnol Biofuels (2017) 10:225*
DOI 10.1186/s13068-017-0901-2

Biotechnology for Biofuels

RESEARCH

Open Access

Network reconstruction and systems analysis of plant cell wall deconstruction by *Neurospora crassa*

Areejit Samal[1,2,3,4], James P. Craig[2,5], Samuel T. Coradetti[2,5†], J. Philipp Benz[2,6†], James A. Eddy[1], Nathan D. Price[1*] and N. Louise Glass[2,5*]

Abstract

Background: Plant biomass degradation by fungal-derived enzymes is rapidly expanding in economic importance as a clean and efficient source for biofuels. The ability to rationally engineer filamentous fungi would facilitate biotechnological applications for degradation of plant cell wall polysaccharides. However, incomplete knowledge of biomolecular networks responsible for plant cell wall deconstruction impedes experimental efforts in this direction.

Results: To expand this knowledge base, a detailed network of reactions important for deconstruction of plant cell wall polysaccharides into simple sugars was constructed for the filamentous fungus *Neurospora crassa*. To reconstruct this network, information was integrated from five heterogeneous data types: functional genomics, transcriptomics, proteomics, genetics, and biochemical characterizations. The combined information was encapsulated into a feature matrix and the evidence weighted to assign annotation confidence scores for each gene within the network. Comparative analyses of RNA-seq and ChIP-seq data shed light on the regulation of the plant cell wall degradation network, leading to a novel hypothesis for degradation of the hemicellulose mannan. The transcription factor CLR-2 was subsequently experimentally shown to play a key role in the mannan degradation pathway of *N. crassa*.

Conclusions: Here we built a network that serves as a scaffold for integration of diverse experimental datasets. This approach led to the elucidation of regulatory design principles for plant cell wall deconstruction by filamentous fungi and a novel function for the transcription factor CLR-2. This expanding network will aid in efforts to rationally engineer industrially relevant hyper-production strains.

Keywords: Biofuels, Systems biology, *Neurospora crassa*, Network reconstruction, Plant cell wall degradation network, Transcriptional regulatory networks, CLR-2, Mannan

Background

Plant biomass, primarily composed of lignocellulose, is a renewable and environmentally clean energy source, and a promising feedstock for the production of next-generation biofuels and specialty chemicals [1–3]. A principal barrier to economical production of biofuels is the high production cost of biomass depolymerization enzymes [4]. Filamentous fungi are among the most efficient degraders of lignocellulosic biomass in nature and play a key role in carbon recycling [5, 6]. Industrially relevant strains, such as *Trichoderma reesei*, were constructed through multiple rounds of random mutagenesis and can secrete over 100 g/L of protein [7, 8]. However, rationally engineering strains of filamentous fungi to further enhance the secretion of enzymes is a major challenge in bioenergy research [9]. To meet this challenge and aid future experimental efforts, a system-level understanding of plant cell wall deconstruction by filamentous fungi is necessary [5, 10].

*Correspondence: nprice@systemsbiology.org; Lglass@berkeley.edu
†Samuel T. Coradetti and J. Philipp Benz contributed equally to this work
[1] Institute for Systems Biology, Seattle, WA 98109, USA
[2] Energy Biosciences Institute, University of California Berkeley, Berkeley, CA 94704, USA
Full list of author information is available at the end of the article

BioMed Central

© The Author(s) 2017. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

The model filamentous fungus *Neurospora crassa* has well-developed genetics, biochemistry, molecular biology, and a well-annotated genome [11–14]. In nature, *N. crassa* colonizes freshly burnt plant biomass and shows robust growth on lignocellulose [5, 15–19]. The suite of experimental resources available for *N. crassa* makes it an ideal model system for bioenergy-related research, particularly for the elucidation of plant cell wall deconstruction mechanisms and nutrient sensing [5, 16–20]. Research on *N. crassa* contributed to the discovery of a new class of enzymes called lytic polysaccharide monooxygenases (LPMOs), which greatly increase synergy in cellulose degradation. In addition, novel cellodextrin transporters from *N. crassa* were utilized to engineer improved yeast strains for sugar fermentation [21–24]. A network reconstruction encompassing the present knowledge of metabolic reactions, enzymes, and associated genes in *N. crassa* dedicated to the deconstruction of plant cell wall polysaccharides into simple fermentable sugars will further expedite experimental efforts.

The availability of fully sequenced genomes and accumulated wealth of biochemical evidence led to the reconstruction of genome-scale and manually curated metabolic networks for more than 50 organisms across the three domains of life [25, 26]. These genome-scale metabolic networks have been widely analyzed using constraint-based modeling methods to predict the response to environmental and genetic perturbations [27, 28]. Notably, only a few curated genome-scale metabolic reconstructions have been built for filamentous fungi [29–34]. While a manually curated genome-scale metabolic network for *N. crassa* exists [34], this reconstruction and those built for other ascomycete fungi [29–34] are limited by significant knowledge gaps, specifically pathways for the degradation and utilization of plant cell wall polysaccharides.

To overcome this limitation, we built a detailed network of biochemical reactions important for the degradation of plant cell wall polysaccharides into simple fermentable sugars in *N. crassa* (Fig. 1; Additional file 1). Plant cell walls are largely composed of complex polysaccharides that include cellulose, hemicellulose, and pectin [5, 35–40]. Cellulose is the most abundant plant cell wall polysaccharide and is an unbranched structure composed of linear chains of β-1,4-linked D-glucose residues. The second most abundant is the heterogeneous group of hemicelluloses, composed of several branched polymers, including xylan, xyloglucan, mannan, and mixed-linkage glucan. Pectin is a minor constituent of mature plant cell walls but the most complex heteropolysaccharide. Its main constituents are homogalacturonan, xylogalacturonan, and rhamnogalacturonan I.

The plant cell wall degradation network (PCWDN) reconstruction and annotation pipeline described here involved the integration of five heterogeneous data types: functional genomics, transcriptomics, proteomics, genetics, and biochemical information, along with extensive manual curation based on more than 130 research articles (Fig. 2; Additional file 1). The combined annotation information was encapsulated in a feature matrix, which was used to assign annotation confidence scores to PCWDN genes. Comparative analysis of RNA sequencing (RNA-seq)-based global transcriptome profiles underlined the importance of PCWDN genes for adaptations to different plant cell wall polysaccharides. Subsequent analyses of RNA-seq and ChIP-seq data within the context of the *N. crassa* PCWDN provided novel insights into the roles of key transcription factors (TFs) in the deconstruction of plant biomass, which were tested here. The *N. crassa* PCWDN will serve as a scaffold for the integration and systems analyses of diverse experimental data, helping elucidate the regulatory design principles underlying plant cell wall deconstruction by filamentous fungi.

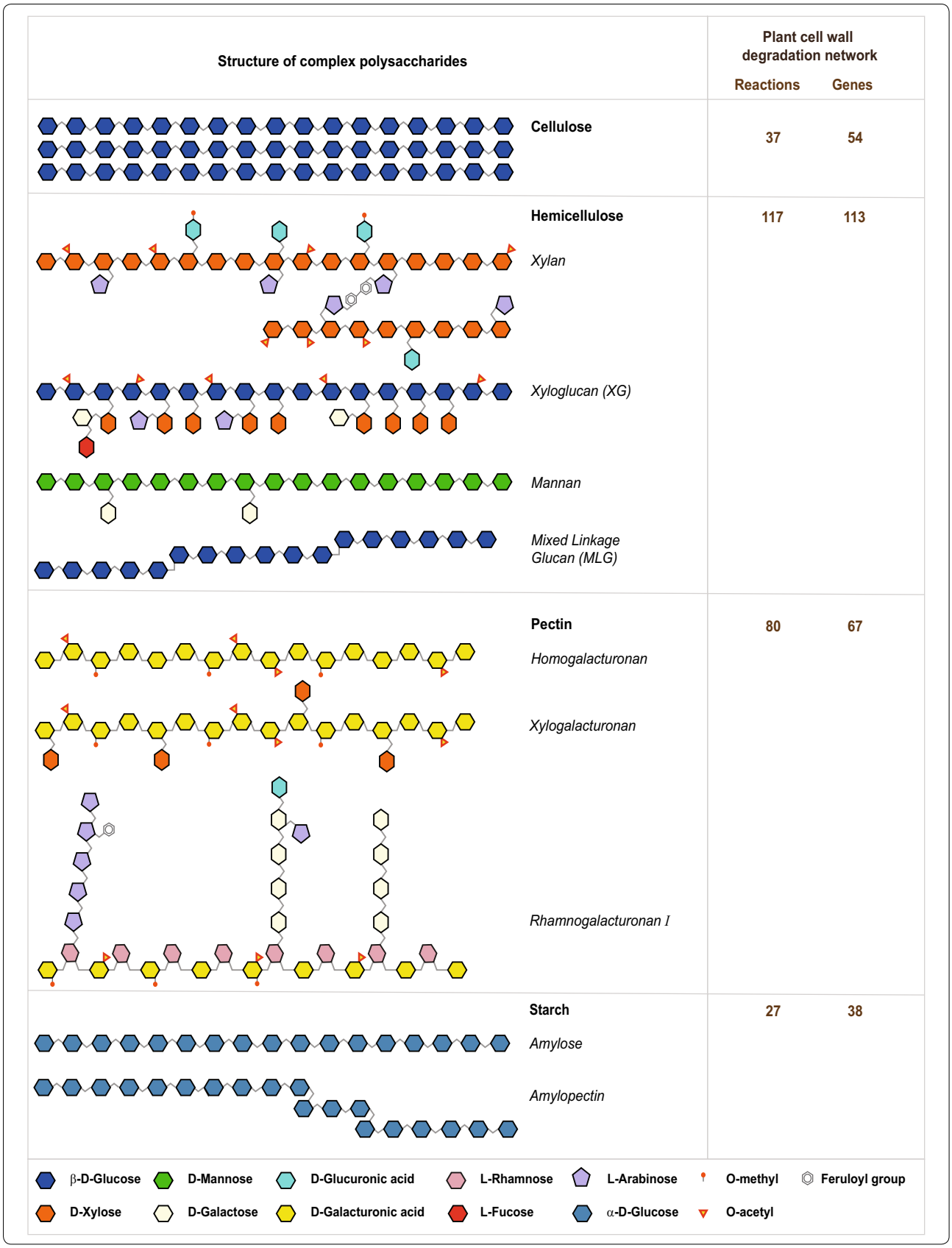
Results and discussion

Network reconstruction and annotation pipeline

We assembled an initial list of biochemical reactions and associated genes in the PCWDN v0.0 of *N. crassa* by combining information on predicted enzymes and transporters involved in the degradation of plant cell wall polysaccharides from the following sources: the Carbohydrate-Active enZYmes database (CAZY) [41], the *N. crassa* e-Compendium [42], the genome annotation for *N. crassa* OR74A [12, 13], and TransportDB [43] ("Methods"; Fig. 2). Specifically, 110 out of 231 CAZY genes predicted to encode carbohydrate-active enzymes in the genome were included in the PCWDN v0.0 (Additional file 2). The remaining 121 CAZY genes mainly belong to families of enzymes active on chitin and chitosan and thus are not likely to be involved in the PCWDN,

(See figure on next page.)

Fig. 1 Schematic illustration of the structure of different plant cell wall polysaccharides along with the associated reactions and genes in the PCWDN of *N. crassa*. Cellulose has an unbranched structure composed of linear chains of β-1,4-linked D-glucose residues. Hemicellulose comprises several branched polymers including xylan, xyloglucan (XG), mannan, and mixed-linkage glucan (MLG). Pectin is a family of several polymers including homogalacturonan, xylogalacturonan, and rhamnogalacturonan I. Starch is a polymer composed of amylose and amylopectin. On the right, the number of reactions and genes involved in the degradation of cellulose, hemicelluloses, pectin, and starch are indicated that are captured in our PCWDN


Structure of complex polysaccharides
Plant cell wall degradation network
Reactions
Genes
Cellulose
37
54
Hemicellulose
117
113
Xylan
Xyloglucan (XG)
Mannan
Mixed Linkage Glucan (MLG)
Pectin
80
67
Homogalacturonan
Xylogalacturonan
Rhamnogalacturonan I
Starch
27
38
Amylose
Amylopectin
β-D-Glucose
D-Mannose
D-Glucuronic acid
L-Rhamnose
L-Arabinose
O-methyl
Feruloyl group
D-Xylose
D-Galactose
D-Galacturonic acid
L-Fucose
α-D-Glucose
O-acetyl

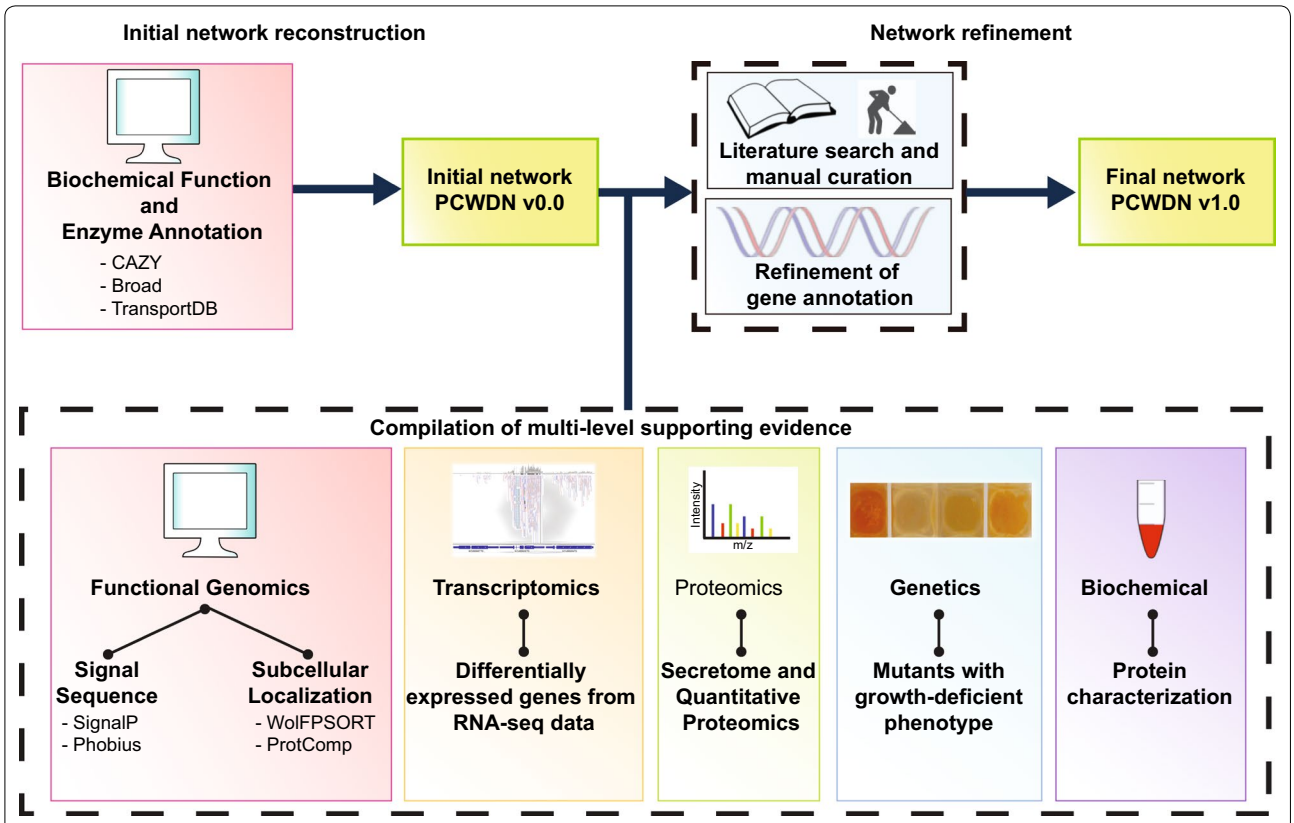


Fig. 2 Schematic illustration of the pipeline for reconstruction and annotation of the PCWDN of *N. crassa*. An initial scaffold network, PCWDN v0.0, was assembled based on annotation information in several databases. Extensive literature-based manual curation was then performed to fill knowledge gaps in the initial PCWDN v0.0. Gene annotations in the final PCWDN v1.0 were refined based on multi-level supporting evidence from five heterogeneous data types

but rather in remodeling the fungal cell wall (Additional file 2).

The annotation of PCWDN genes in the above databases has not been updated with data on plant cell wall deconstruction by *N. crassa*. For example, the current OR74A genome annotation is unable to differentiate between cellulolytic LPMOs, hemicellulose-active LPMOs, and starch-active LPMOs [44–50]. Thus, we performed extensive literature-based manual curation involving more than 130 research articles (Additional file 1) to fill the knowledge gaps in the initial PCWDN v0.0 and compiled multi-level supporting evidence as described below from five heterogeneous data types: functional genomics, transcriptomics, proteomics, genetics, and biochemical characterizations, to annotate genes in the final PCWDN v1.0 of 202 reactions and 168 genes (Fig. 2; Additional file 1). The 202 reactions in the final PCWDN of *N. crassa* were further subdivided into 101 extracellular reactions, 35 transport reactions, and 66 intracellular reactions (Additional file 1).

Functional genomics-based annotation

An important annotation feature of PCWDN enzymes is their predicted subcellular localization. For example, the hydrolysis of cellodextrins into D-glucose by β-glucosidases can occur in the extracellular space or in the intracellular space (Additional file 1). Of the 168 PCWDN genes, products of 103, 19, and 46 genes are associated with extracellular, transport, and intracellular reactions, respectively. We used SignalP [51] and Phobius [52] to predict the presence of signal peptides in PCWDN proteins to determine if they were destined toward the secretory pathway ("Methods"). We found that 89 out of the 103 gene products (~86%) associated with extracellular reactions were predicted to have a signal peptide by at least one of the two tools, while no gene products associated with transport or intracellular reactions were predicted to have a signal peptide by either of the two tools (Additional file 3). WoLF PSORT [53] and ProtComp were also used to predict subcellular localization of proteins ("Methods"). Predictions from at least

one of the two tools matched the assigned localization for 90 out of the 103 gene products (~87%) associated with extracellular reactions, while the predictions from at least one of the two tools matched the assigned localization for all gene products associated with transport or intracellular reactions in the PCWDN (Additional file 3).

Using the compiled functional genomics-based information (Additional file 3), we assessed the annotation support for the 168 PCWDN members. Specifically, functional genomics information was considered to support the annotation of a PCWDN enzyme if the following three conditions were satisfied (Fig. 3): (i) gene annotation in CAZY database [41] or Broad OR74A genome [12, 13] or TransportDB [43] matches the assigned biochemical function in the network; (ii) SignalP or Phobius predicts the presence of a signal peptide in extracellular (secreted) enzymes and the absence of a signal peptide in intracellular enzymes; (iii) subcellular localization predictions from WoLF PSORT or ProtComp match the assigned localization in the network. Based on this definition, we obtained functional genomics-based annotation support for 145 of the 168 PCWDN genes (Additional file 3).

Transcriptomics-based annotation

To augment the annotation of PCWDN genes, we used RNA-seq data and compared the transcriptional profiles of *N. crassa* WT strain (FGSC 2489 [12]; "Methods") grown under different carbon source regimes corresponding to the different carbohydrates that make up the plant cell wall. Previous studies [16, 19] generated RNA-seq data from shift experiments, in which a 16-h-old culture of *N. crassa* WT was shifted for 4 h to minimal media with no carbon (NoC) source or one of five carbon sources: sucrose, cellobiose (CB), Avicel (microcrystalline cellulose), xylan, pectin, or orange peel powder (OPP, a pectin-rich substrate) ("Methods"; Additional file 4). We replicated this experimental design, generating RNA-seq data for four additional carbon sources: xyloglucan (XG), mannan, mixed-linkage glucan (MLG), and starch ("Methods"; Additional file 4). This approach ensured that a comparative analysis of transcriptional profiles could be performed between all tested plant cell wall polysaccharides. A pipeline consisting of standard software was used to analyze the RNA-seq data and identify differentially expressed genes between the nine different conditions corresponding to different polysaccharides and the two controls corresponding to NoC and sucrose conditions ("Methods"; Additional file 4).

Based on the compiled RNA-seq data, we generated transcriptomics-based annotation support for the PCWDN genes (Fig. 3; Additional file 4). We considered RNA-seq data to support the annotation of a gene involved in the degradation of a specific polysaccharide if its transcript was significantly differentially expressed and upregulated in the relevant carbon source as compared to the NoC control. For example, RNA-seq data supported the annotation of the gene *gh6-3* (NCU07190), predicted to encode an exo-β-1,4-glucanase (cellobiohydrolase) involved in cellulose degradation, as this gene is upregulated on Avicel in comparison to the NoC control (Additional file 4). We obtained transcriptomics-based annotation support for 106 of the 168 PCWDN genes (Additional file 4).

Proteomics-based annotation

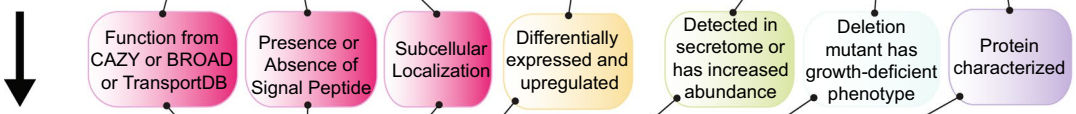
To further enrich the PCWDN and provide an added layer of confidence in the annotation, we compiled proteomics data from previous *N. crassa* studies [15, 17–19, 54–57] (Additional file 5). In these studies, the secretomes from *N. crassa* grown on sucrose, Avicel, xylan, pectin, OPP, and NoC had been characterized using a shotgun proteomics approach [15, 17–19, 54, 55] or by quantitative proteomics [57] (Additional file 5). The compiled data were used to generate proteomics-based support for the annotation of PCWDN genes. Proteomics data supported the annotation of an enzyme if the protein was detected in the secretome or displayed an increased abundance in a carbon source as compared to the NoC control (Fig. 3). For example, proteomics data further supported the annotation of *gh6-3*, as the encoded protein was detected in the secretome and also increased in abundance on Avicel compared to the NoC control (Additional file 5). In this way, we obtained proteomics-based annotation support for 68 of the 168 PCWDN genes (Additional file 5). In cases where an enzyme in the PCWDN was detected in a secretome, but was not predicted to be secreted using functional

(See figure on next page.)

Fig. 3 Feature matrix and annotation confidence scores for genes encoding AA9 LPMOs in *N. crassa*. Combined annotation information from the five different data types was captured in a feature matrix, and a method was devised to assign annotation confidence scores to PCWDN genes. A differential weighting system was used to account for the different levels of confidence associated with the information from each data type. The majority of genes encoding AA9 LPMOs of class 1 (3 out of 5 genes) and class 2 (2 out of 3 genes) are well characterized, while only 1 out of 6 genes encoding AA9 LPMOs of class 3 is well characterized. *NoC* no carbon, *CB* cellobiose, *XG* xyloglucan, *MLG* mixed-linkage glucan, *OPP* orange peel powder

Feature Matrix

Gene	Functional Genomics							Transcriptomics (Differential Expression relative to NoC condition)									Proteomics		Genetics		Biochemical	
	CAZY	BROAD Annotation version 12	TransportDB	SignalP	Phobius	WoLF PSORT Subcellular Localization	ProtComp Subcellular Localization	Sucrose	CB	Avicel	Xylan	XG	Mannan	MLG	Pectin	OPP	Starch	Detected in Secretome	Increased abundance in Avicel versus NoC	Condition	References	Details	References
NCU00836	AA9	glycosylhydrolase family 61-7		Y	Y	Extracellular	Extracellular	0	1	1	1	1	0	1	1	1	0						
NCU01867	AA9	glycosylhydrolase family 61-10		Y	Y	Extracellular	Extracellular	-1	-1	0	-1	0	0	0	-1	-1	-1					Neurospora crassa	23102010
NCU02344	AA9	fungal cellulose binding domain-containing protein		Y	Y	Extracellular	Mitochondria	1	1	1	1	0	0	1	1	1	1						
NCU03328	AA9	glycosylhydrolase family 61-6		Y	Y	Extracellular	Extracellular	-1	-1	1	-1	0	0	1	-1	-1	-1					Neurospora crassa	23102010
NCU08760	AA9	glycosylhydrolase family 61-5		Y	Y	Extracellular	Extracellular	0	0	1	1	0	0	1	0	1	0	Y	Y			Neurospora crassa	22004347 22188218 23102010
NCU03000	AA9	glycosylhydrolase family 61-8		Y	Y	Extracellular	Extracellular	-1	-1	0	0	0	0	0	0	0	0						
NCU05969	AA9	glycosylhydrolase family 61-9		Y	Y	Extracellular	Extracellular	1	1	0	0	0	0	1	1	1	0						
NCU07520	AA9	glycosylhydrolase family 61-11		Y	Y	Extracellular	Mitochondria	-1	-1	0	0	0	0	0	0	-1	0						
NCU07760	AA9	glycosylhydrolase family 61-2		Y	Y	Extracellular	Extracellular	1	1	1	1	0	1	1	1	1	0						
NCU07898	AA9	endoglucanase IV		Y	Y	Extracellular	Extracellular	-1	0	1	1	1	0	1	1	1	0	Y				Neurospora crassa	22004347 22188218
NCU07974	AA9	endoglucanase B		Y	Y	Extracellular	Extracellular	1	1	0	0	0	0	0	1	1	0						
NCU01050	AA9	glycosylhydrolase family 61-4		Y	Y	Extracellular	Extracellular	0	1	1	1	1	0	1	0	0	0	Y	Y			Neurospora crassa	22004347 22188218
NCU02240	AA9	glycosylhydrolase family 61-1		Y	Y	Extracellular	Extracellular	-1	-1	1	1	0	1	1	0	0	-1	Y					
NCU02916	AA9	glycosylhydrolase family 61-3		Y	Y	Extracellular	Extracellular	0	1	1	1	0	0	1	1	1	0					Neurospora crassa	22729546 23102010 24733907



Annotation support and confidence score

Gene	Annotation Support: Functional Genomics X 1	Transcriptomics X 2	Proteomics X 4	Genetics X 8	Biochemical X 16	Confidence Score /31	Our Annotation
NCU00836	0	1	0	0	0	0.06	lytic polysaccharide monooxygenase 1 (LPMO1)
NCU01867	1	0	0	0	1	0.55	lytic polysaccharide monooxygenase 1 (LPMO1)
NCU02344	1	1	0	0	0	0.10	lytic polysaccharide monooxygenase 1 (LPMO1)
NCU03328	1	1	0	0	1	0.61	lytic polysaccharide monooxygenase 1 (LPMO1)
NCU08760	1	1	1	0	1	0.74	lytic polysaccharide monooxygenase 1 (LPMO1)
NCU03000	1	0	0	0	0	0.03	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU05969	1	1	0	0	0	0.10	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU07520	1	0	0	0	0	0.03	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU07760	1	1	0	0	0	0.10	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU07898	1	1	1	0	1	0.74	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU07974	1	0	0	0	0	0.03	lytic polysaccharide monooxygenase 3 (LPMO3)
NCU01050	1	1	1	0	1	0.74	lytic polysaccharide monooxygenase 2 (LPMO2)
NCU02240	1	1	1	0	0	0.23	lytic polysaccharide monooxygenase 2 (LPMO2)
NCU02916	1	1	0	0	1	0.61	lytic polysaccharide monooxygenase 2 (LPMO2)

genomics tools, the proteomics-based evidence was considered more reliable and given priority. This approach led to the re-annotation of 4 PCWDN genes: *gla-1*(NCU01517), *gh6-3* (NCU07190), *gh5-7* (NCU08412), and *ce5-2* (NCU09663).

Genetics-based annotation

Next, we mined published literature to compile an experimentally verified dataset of *N. crassa* deletion strains for PCWDN genes with a growth-deficient phenotype as compared to the parental WT strain (FGSC 2489) (Additional file 6). This dataset was used to assign genetics-based annotation support to PCWDN enzymes. For example, the genetics-based dataset supported the annotation of the gene *gh10-2* (NCU08189) as an endo-β-1,4-xylanase involved in xylan degradation, since the deletion strain for this gene exhibited a growth-deficient phenotype on xylan [17] (Additional file 6). Overall, genetics-based information supported the annotation of 19 out of the 168 PCWDN genes (Additional file 6).

Biochemical characterization of enzymes

Lastly, we mined the literature for biochemical data to support the annotation of PCWDN enzymes (Fig. 2). Gene products for 33 out of the 168 PCWDN genes have been biochemically characterized in *N. crassa* [20–24, 46–50, 58–68] (Additional file 7). An additional extensive literature search was performed to determine if ortholog/paralogs of *N. crassa* PCWDN genes had been biochemically characterized in other filamentous fungi [39, 69, 70] (Additional file 7). We used OrthoMCL [71, 72] to determine the orthology/paralogy of PCWDN genes in other filamentous fungi ("Methods"). In this way, biochemical-based annotation support was obtained for 113 out of the 168 PCWDN genes (Additional file 7).

Feature matrix and annotation confidence score

The combined annotation information for PCWDN genes from the five different data types was captured in a feature matrix (Fig. 3; Additional file 1). We next devised a simple method based on the feature matrix to assign annotation confidence scores to PCWDN genes. A differential weighting system was used to account for the different levels of confidence associated with the annotation information for each of the five heterogeneous data types with each annotation level superseded by the next (Fig. 3). Annotation support from biochemical characterizations was given the highest level of confidence (factor of 16), followed by published mutant phenotypes (factor of 8), proteomic data (factor of 4), transcriptomic data (factor of 2), and functional genomics-based predictions (factor of 1). To obtain an overall annotation confidence score in the range of 0–1, the weighted sum of evidence support from the five data types was normalized by dividing by the maximal possible score of 31 (Fig. 3). Note that the chosen factors in the differential weighting system are the simplest possible that enable mathematical mapping from diverse evidence support values to a unique annotation confidence score. Moreover, the chosen factors are such that the contribution from a given data type to the confidence score is always more than the combined contribution from all other data types with less confidence. That is, a gene with annotation support from only biochemical characterization has a higher confidence score than a gene with combined annotation support from genetics, proteomics, transcriptomics, and functional genomics-based information. Additional file 1 lists the annotation confidence scores for the 168 PCWDN genes, and Fig. 3 presents the confidence scores for the genes belonging to CAZY class AA9 (Auxiliary Activity Family 9) encoding LPMOs involved in cellulose degradation [44, 47, 50]. For this particular case, the majority of genes encoding AA9 LPMOs within class 1 (3 out of 5 genes) and class 2 (2 out of 3 genes) are well characterized, while this is true for only 1 out of 6 genes encoding class 3 AA9 LPMOs.

Comparative global transcriptome analysis

To better define the genome-wide response of *N. crassa* to plant cell wall substrates, we analyzed combined RNA-seq data for possible correlations between the response patterns for various carbon sources. Figure 4a shows the level of correlation across the whole transcriptome for each pair of growth conditions. Hierarchical clustering of the pairwise correlation matrix across carbon sources revealed four main clusters. Avicel (cellulose) and MLG (a hemicellulose component) were in the first cluster with highly correlated gene expression patterns, likely due to the fact that both polysaccharides have a backbone that is rich in β-1,4-linked D-glucose residues. Similarly, sucrose and starch conditions fell into a second cluster, while xylan and pectin conditions were in a third cluster. This analysis also identified a fourth cluster where the gene expression in two hemicellulose components, mannan and XG, were very similar to the NoC starvation condition. In support of these observations, the *N. crassa* WT strain was observed to grow poorly on both mannan and XG. Principal component analysis of expression data under the different conditions gave a similar result (Fig. 4b), where the first two principal components together explained more than 79% of the total variance.

For a more detailed comparison of the *N. crassa* transcriptional response upon exposure to different plant cell wall polysaccharides, we identified the regulons for each polysaccharide using the NoC and sucrose conditions as controls. Following Benz et al. [19], the regulon (or

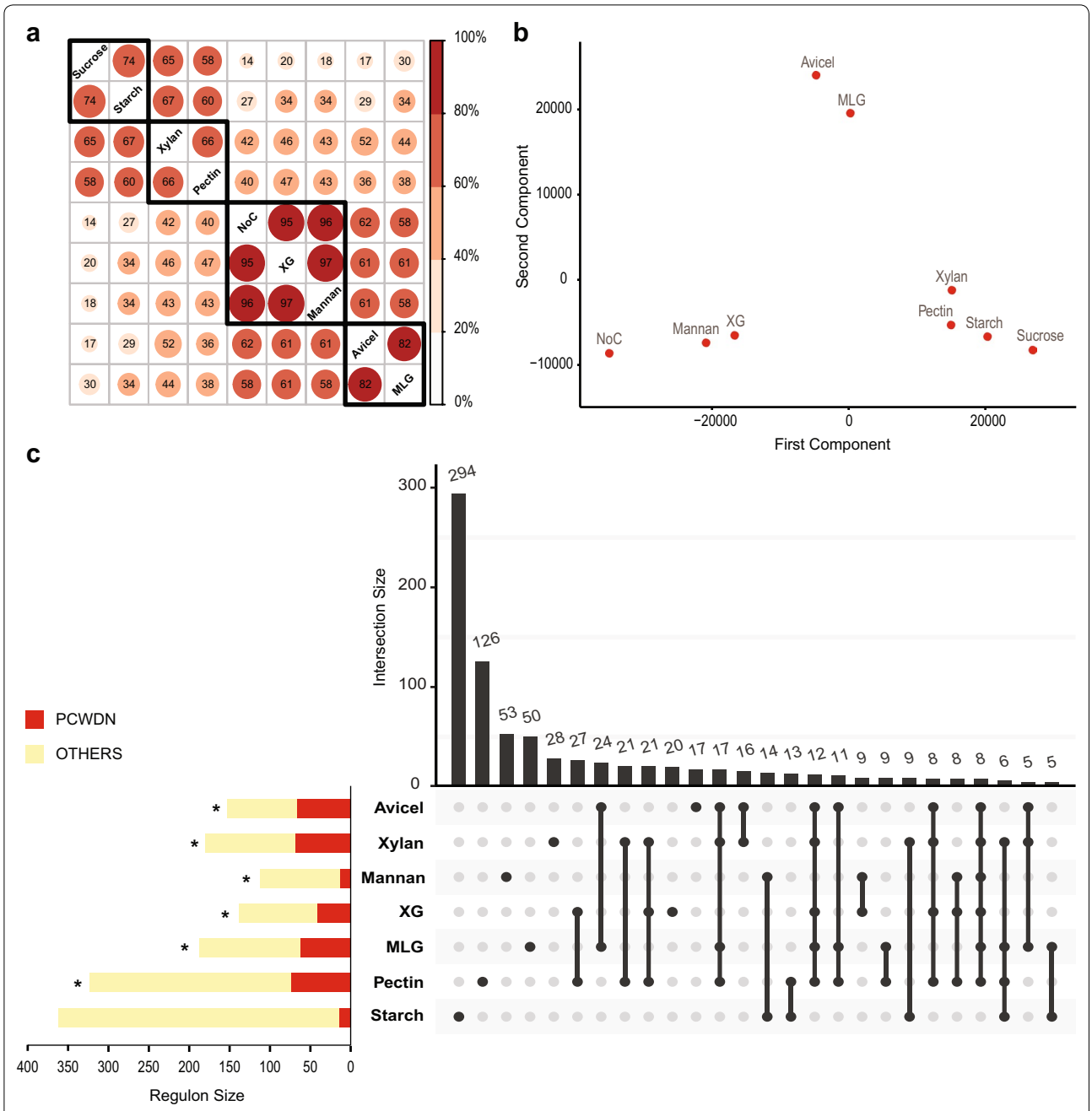


Fig. 4 Comparative analysis of global transcriptome profiles of *N. crassa* in nine conditions. **a** Correlation of the whole transcriptomes across each pair of conditions. Hierarchical clustering of the pairwise correlation matrix led to the identification of four clusters of highly correlated conditions. **b** Principal component analysis of the whole transcriptomes in all nine conditions. The first and second principal components explain >58% and 20%, respectively, of the total variance. Thus, the first two components together explain more than 79% of the total variance. **c** Horizontal bar plot showing the sizes of regulons for seven conditions and the overlap of each regulon with the 168 PCWDN genes. Statistically significant overlaps ($p < 10^{-6}$) were marked with an asterisk. The vertical bar plot shows the 26 intersection sets among the seven regulons with 5 or more genes and was generated using UpSetR [108]

up-regulon) for a given growth condition was defined as genes that were upregulated and differentially expressed in relation to both controls. Based on this definition, the regulons for Avicel, xylan, XG, mannan, MLG, pectin, and starch were determined to contain 153, 180, 138, 112, 188, 323, and 363 genes, respectively (horizontal bar plot in Fig. 4c; Additional file 4; for the corresponding down-regulons, see Additional files 4, 8). Although starch was

determined to have the largest regulon with 363 genes, it was fairly distinct among the regulons with 294 unique genes (81%). By comparison, mannan was 47% unique, followed by pectin (39%), MLG (28%), xylan (16%), XG (15%), and Avicel (11%). The overlap between each of the seven regulons was determined, and the vertical plot in Fig. 4c shows all regulon comparisons with an overlap of 5 or more genes. Interestingly, XG and pectin regulons contained the highest overlap with 27 genes in common, followed by the Avicel regulon that overlapped most highly with MLG (24 genes), xylan with pectin (21 genes), and mannan with starch (14 genes). All 26 intersection sets from this analysis were subjected to functional category analysis based on gene annotations in FunCatDB [73] (Additional file 9). Most of the intersection sets involving either Avicel, xylan, XG, mannan, MLG, or pectin regulons were enriched in metabolic genes ("Metabolism and Energy").

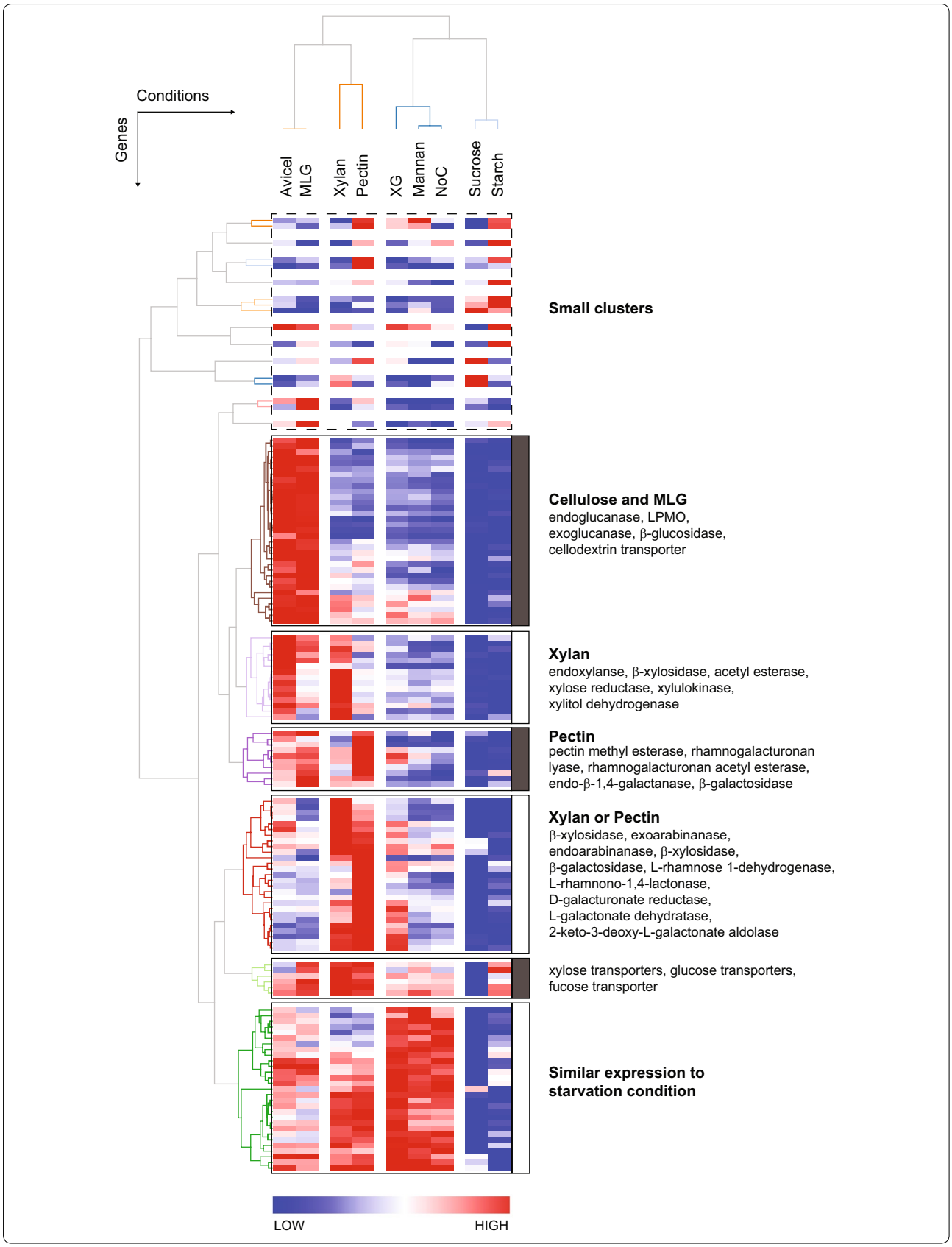
We next computed the relative abundance of the 168 PCWDN genes within each of the seven polysaccharide regulons and the probability of their enrichment. For Avicel, 66 PCWDN genes were represented in the regulon of 153 genes, a 43% enrichment ($p < 10^{-78}$). The xylan regulon contained 68 PCWDN genes (37% enrichment; $p < 10^{-76}$), the mannan regulon contained 13 PCWDN genes (29%; $p < 10^{-7}$), the XG regulon contained 41 PCWDN genes (43%; $p < 10^{-39}$), the MLG regulon contained 62 PCWDN genes (32%; $p < 10^{-64}$), the pectin regulon contained 74 PCWDN genes (22%; $p < 10^{-65}$), and the starch regulon contained 14 PCWDN genes (3.8%; $p < 10^{-2}$) (Fig. 4c). The observed overlap between PCWDN genes and regulons was statistically highly significant, since the PCWDN genes account for only 1.7% of genes (168 out of 9758) within the *N. crassa* genome. Thus, our PCWDN genes capture a substantial (and relevant) part of the transcriptional response of *N. crassa* upon exposure to Avicel (cellulose), xylan, mannan, XG, MLG, and pectin.

Clustering of transcriptome profiles within the context of the PCWDN

Comparative global transcriptional analyses underlined the importance of PCWDN genes for adaptation to different plant cell wall polysaccharides. We next performed hierarchical clustering [74] of the RNA-seq data for the 168 PCWDN genes in all nine conditions (Avicel, xylan, XG, mannan, MLG, pectin, starch, sucrose, and NoC). The conditions (*x*-axis in Fig. 5) clustered similarly to what was obtained from correlation and principal component analysis of the genome-wide transcriptome profiles presented in Fig. 4a, b. The genes (*y*-axis in Fig. 5) were grouped into six major clusters ("Methods"; Additional file 10). The 33 genes in the largest cluster were highly expressed under Avicel and MLG as compared to other conditions. Notably, this cluster contains most of the endo-β-1,4-glucanases, exo-β-1,4-glucanases, β-glucosidases, AA9 LPMOs, and the identified cellodextrin transporters. The second cluster contained 15 genes, showed the highest expression on xylan, and consisted of genes encoding endo-xylanases, β-xylosidases, acetyl xylan esterases, xylose reductase, xylulokinase, and xylitol dehydrogenase. The 10 genes in the third cluster were most strongly expressed on pectin and included genes encoding pectin methyl esterases, a rhamnogalacturonan lyase, a rhamnogalacturonan acetyl esterase, an endo-β-1,4-galactanase, and a β-galactosidase. The fourth cluster composed of 27 genes showed transcriptional activity on both xylan and pectin and consisted mainly of genes encoding β-xylosidases, exo- and endo-arabinanases, α-arabinosidases, β-galactosidases, and genes involved in L-rhamnose and D-galacturonic acid metabolism. The fifth cluster composed of 6 genes with broader expression pattern contained four sugar transporters, and the final cluster composed of 29 genes contained those factors most strongly induced during starvation (excluding mannan and XG).

The AA9 LPMOs, previously annotated as CAZY class GH61, degrade cellulose by oxidative cleavage [44, 46, 47, 49]. Most AA9 LMPOs were highly expressed on Avicel and MLG and their expression was correlated across the two conditions (Additional file 11). In comparison to Avicel and MLG, the expression of 10 of the 14 LPMOs was much lower on XG and negligible on starch. Based on these observations, we concluded that these genes are specifically induced by non-substituted β-D-glucans and hypothesize that LPMOs in AA9 CAZY class also act on MLG, which is now included in the PCWDN of *N. crassa* (Additional file 1).

A recent study [49] suggested that the product of AA9 LPMO gene NCU02916 can also act on XG. However, NCU02916 was not induced when grown on XG, and the *N. crassa* WT strain (FGSC 2489) grew poorly on XG, with the transcriptome profile highly correlated with that of the NoC control (Fig. 4a). Thus, it is difficult to rule out the possibility that the AA9 LPMOs also act on XG, and we therefore included such reactions in the PCWDN of *N. crassa* (Additional file 1). However, another recent study [50] characterized a starch-specific LPMO, NCU08746, and this gene is classified in the new CAZY class AA13 [41]. Consistent with this finding, NCU08746 was differentially upregulated under starch conditions (Additional file 4), and thus NCU08746 was assigned to starch-specific LPMO reaction in the PCWDN (Additional file 1).


Conditions
Genes
Avicel
MLG
Xylan
Pectin
XG
Mannan
NoC
Sucrose
Starch
Small clusters
Cellulose and MLG
endoglucanase, LPMO,
exoglucanase, β-glucosidase,
cellodextrin transporter
Xylan
endoxylanse, β-xylosidase, acetyl esterase,
xylose reductase, xylulokinase,
xylitol dehydrogenase
Pectin
pectin methyl esterase, rhamnogalacturonan
lyase, rhamnogalacturonan acetyl esterase,
endo-β-1,4-galactanase, β-galactosidase
Xylan or Pectin
β-xylosidase, exoarabinanase,
endoarabinanase, β-xylosidase,
β-galactosidase, L-rhamnose 1-dehydrogenase,
L-rhamnono-1,4-lactonase,
D-galacturonate reductase,
L-galactonate dehydratase,
2-keto-3-deoxy-L-galactonate aldolase
xylose transporters, glucose transporters,
fucose transporter
Similar expression to
starvation condition
LOW
HIGH

(See figure on previous page.)
Fig. 5 Hierarchical clustering of transcriptome profiles within the context of the PCWDN of *N. crassa*. The heatmap shows the result of two-dimensional clustering of the RNA-seq data for all 168 PCWDN genes in nine conditions corresponding to Avicel, xylan, XG, mannan, MLG, pectin, starch, sucrose, and NoC, which led to the identification of six major clusters. Note that the clustering of conditions was similar to that obtained from correlation and principal component analysis of the global transcriptome profiles (Fig. 4a, b)

Regulation of the PCWDN by key transcription factors

Previous research [16, 17] on plant cell wall deconstruction by *N. crassa* led to the identification of two essential transcription factors (TFs), CLR-1 (NCU07705) and CLR-2 (NCU08042), for cellulose utilization, and one essential TF, XLR-1 (NCU06971), for utilization of xylan. All three TFs are conserved across ascomycete fungi [16, 75–80]. Using next-generation sequencing of chromatin-immunoprecipitated DNA (ChIP-seq), a recent study [81] identified binding regions for CLR-1, CLR-2, and XLR-1 across the *N. crassa* genome under sucrose, Avicel, xylan, and NoC conditions. CLR-1, CLR-2, and XLR-1 bound to the regulatory regions of 293, 164, and 84 genes, respectively, in the *N. crassa* genome [81]. Integrating this information, we determined that CLR-1, CLR-2, and XLR-1 bound the promoter regions of 27 ($p < 10^{-12}$), 37 ($p < 10^{-30}$), and 20 ($p < 10^{-17}$) genes, respectively, within the 168 PCWDN genes (Fig. 6). Of the 27 PCWDN genes regulated by CLR-1, 21 are involved in cellulose utilization, including endo-β-1,4-glucanases, AA9 LPMOs, exo-β-1,4-glucanases, β-glucosidases, or cellodextrin transporters, while 17 out of 20 PCWDN genes regulated by XLR-1 are involved in xylan utilization including endo-β-1,4-xylanases, β-xylosidases, β-arabinofuranosidases, or xylodextrin transporters [81] (Fig. 6). However, in the case of CLR-2, only 23 out of 37 PCWDN genes directly regulated by the TF are involved in cellulose utilization, while 12 other PCWDN genes are involved in xylan or mannan utilization [81] (Fig. 6). These observations suggest that CLR-2, unlike CLR-1 or XLR-1, has a broader regulatory role in plant cell wall deconstruction, and which is not limited to cellulose utilization.

CLR-1 is known to directly regulate *clr-2* [81]. Intriguingly, we found that 19 out of 27 PCWDN genes directly regulated by CLR-1 are also directly regulated by CLR-2 (Fig. 6). These observations indicate that CLR-1 and CLR-2 predominantly employ feed-forward loops (FFLs) [82–86] to regulate the cellulose utilization pathway within the PCWDN of *N. crassa*.

clr-2, *clr-1*, and *gh5-7* are important for mannan degradation

Our comparative transcriptome analysis revealed that the pattern of gene expression of *N. crassa* WT exposed to mannan was highly correlated to the NoC starvation condition. In contrast to the cellulose and xylan utilization pathways, the *N. crassa* genome has a mostly non-redundant mannan utilization pathway, with just one predicted endomannanase, *gh5-7* (NCU08412), and one intracellular β-mannosidase, *gh2-1* (NCU00890) (Fig. 7a; Additional file 1). Note that a novel β-1,4-mannanase belonging to the CAZY family GH134 recently found in *Aspergillus nidulans* [87] has no close homolog in *N. crassa*. Based on ChIP-seq data [81], CLR-2 directly regulates both genes, while CLR-1 directly regulates only *gh5-7* and XLR-1 does not directly regulate any genes in the mannan pathway (Fig. 6). Additionally, *gh5-7* and *gh2-1* were induced on mannan, albeit to a much lower extent than on Avicel or MLG, a context where these enzymes should not be directly required (Fig. 7b). By contrast, the *clr-1* and *clr-2* genes were highly expressed in Avicel and MLG conditions, but not induced by mannan (Fig. 7b). These observations led us to hypothesize that CLR-1 and CLR-2 may be important for the induction of *gh5-7*, constituting the first step in the mannan utilization pathway.

To test this hypothesis, we first designed experiments to confirm that GH5-7, as the only predicted endomannanase in *N. crassa*, is critical for mannan utilization. As the deletion strain for *gh5-7* was not available, we constructed a Δ*gh5-7* strain (ΔNCU08412) following standard procedures ([88]; "Methods"). Since pure mannan acts as a poor inducer (Fig. 8d), we used Konjac glucomannan instead of pure mannan to test our hypothesis. As predicted, the Δ*gh5-7* mutant was found to have a strong growth phenotype, accumulating only 50–60% of WT biomass when grown on glucomannan as a sole carbon source (Fig. 8a, b). Additionally, both the Δ*clr-1* and Δ*clr-2* strains also grew poorly on glucomannan, accumulating only 10–20% of WT biomass, while the Δ*xlr-1* strain showed no significant growth phenotype (Fig. 8a, b).

The results presented above showed that *clr-1*, *clr-2*, and the endomannanase gene *gh5-7* are important for mannan utilization by *N. crassa*. Together with the observation that CLR-2 directly regulates *gh5-7* and *gh2-1*, and that the expression of *clr-2* was negligible in the WT strain under pure mannan conditions (Fig. 7b), we hypothesized that an engineered strain mis-expressing *clr-2* would show enhanced growth on mannan. A *clr-2* mis-expression strain (Pc2) constructed in a previous study [76] results in a strain that constitutively expresses *clr-2* and cellulases.

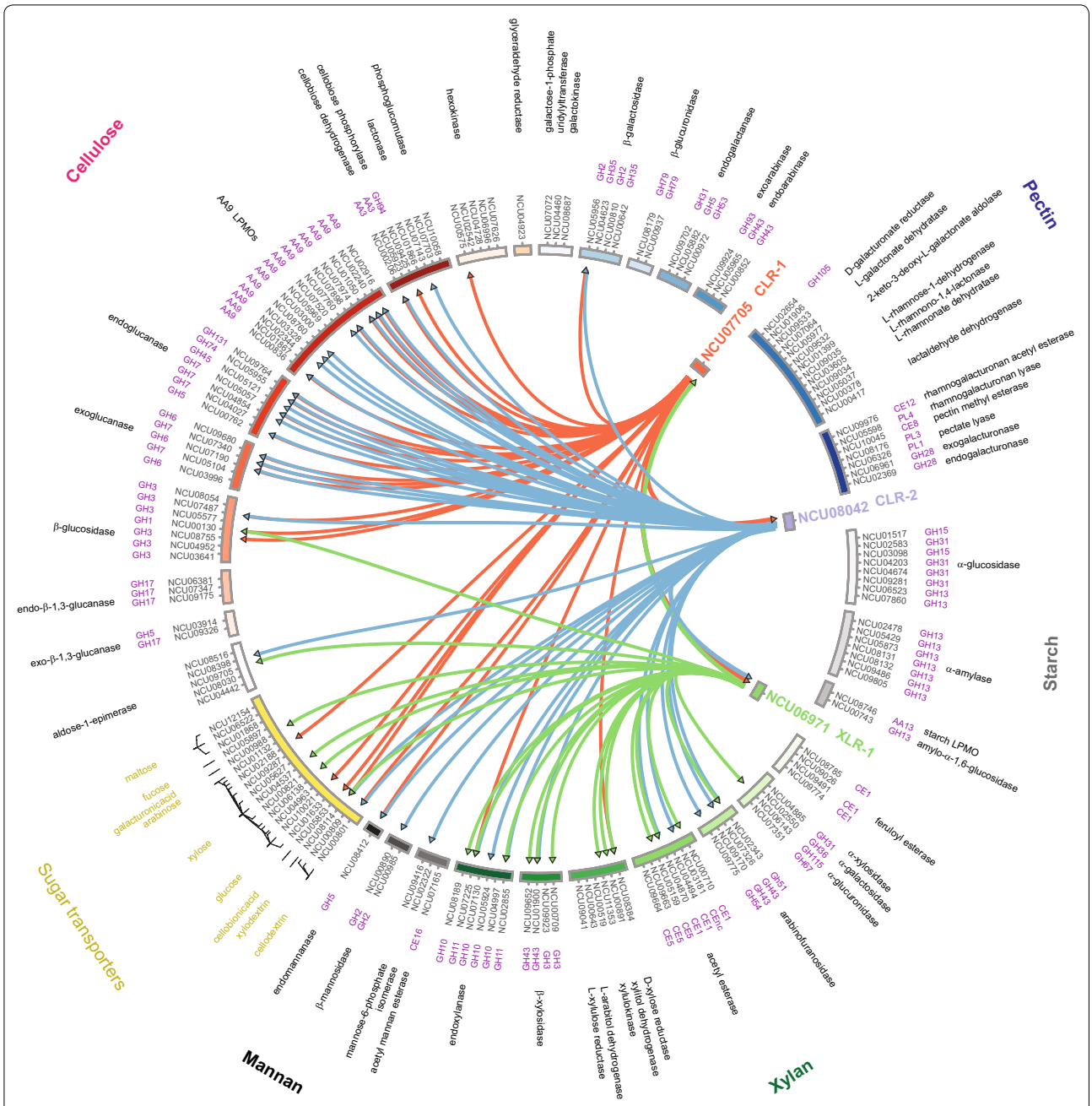
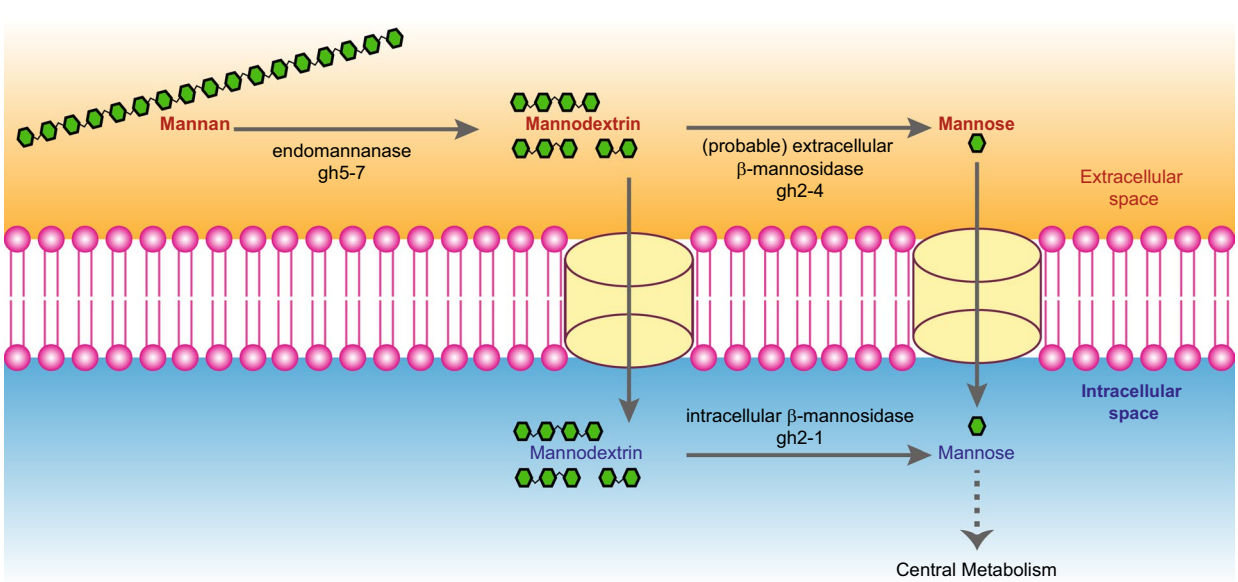


Fig. 6 Direct regulation of the PCWDN genes by key transcription factors for cellulose and xylan utilization. The PCWDN genes have been grouped based on their biochemical function and participation in degradation pathways of different plant cell wall polysaccharides. It is observed that CLR-1 directly regulates mostly genes involved in cellulose utilization, XLR-1 directly regulates mostly genes involved in xylan utilization, while CLR-2 has a much broader role in the regulation of genes involved in plant cell wall deconstruction

We evaluated the growth of the Pc2 strain on glucomannan as compared to the WT strain and determined that the Pc2 strain accumulated 90–100% more fungal biomass than WT on glucomannan, while secreting 7 times as much protein (Fig. 8a–c). Similarly, a *clr-1* mis-expression strain (Pc1) accumulated 90–100% more fungal biomass than WT on glucomannan, albeit exhibiting a similar protein secretion level (Fig. 8a–c). In contrast, a *xlr-1* mis-expression strain (Px1) accumulated fungal biomass at a similar level to WT on glucomannan (Fig. 8a–c).



b

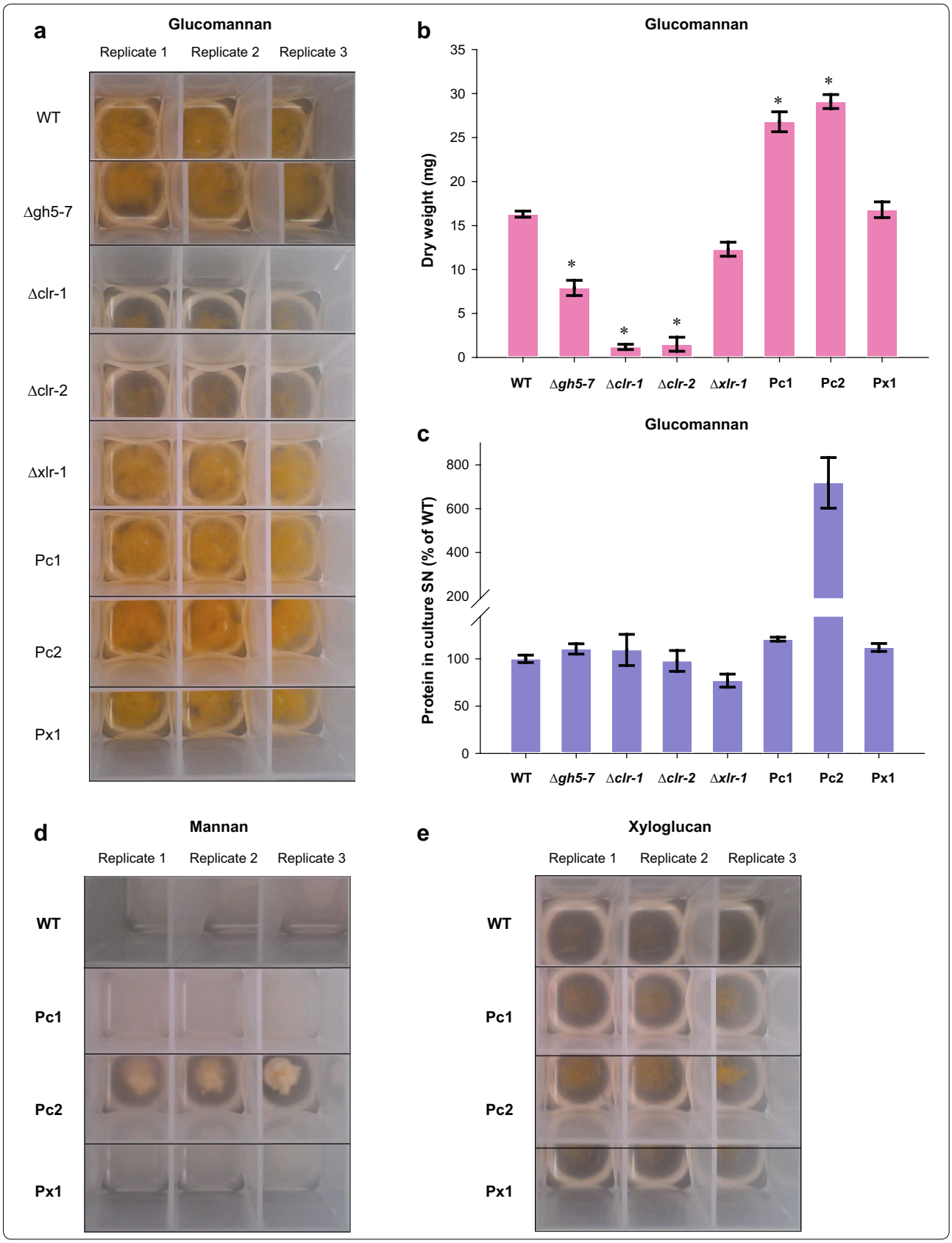
Locus	Name	Change in expression relative to NoC condition								Mean expression								
		Sucrose	Avicel	Xylan	XG	Mannan	MLG	Pectin	Starch	NoC	Sucrose	Avicel	Xylan	XG	Mannan	MLG	Pectin	Starch
NCU08412	*gh5-7*	0.02	47.88	13.26	2.96	20.82	76.47	6.05	0.60	25.8	0.6	1233.3	341.6	76.3	536.4	1969.8	155.8	15.4
NCU00985	*gh2-4*	0.00	1.05	0.03	1.00	5.74	0.73	0.09	0.05	38.7	0.1	40.6	1.3	38.8	221.9	28.3	3.7	2.1
NCU00890	*gh2-1*	0.22	19.58	7.01	1.46	6.16	22.47	4.09	0.17	2.5	0.5	48.5	17.4	3.6	15.3	55.6	10.1	0.4
NCU06971	*xlr-1*	0.02	0.43	0.73	1.27	0.84	3.48	0.92	0.06	56.7	0.9	159.3	41.4	72.2	47.8	197.5	52.0	3.3
NCU07705	*clr-1*	0.16	3.39	1.94	1.18	1.23	3.01	0.90	0.24	31.8	5.0	107.8	61.6	37.4	39.2	95.7	28.6	7.7
NCU08042	*clr-2*	0.12	87.00	1.16	1.33	1.35	70.96	2.12	0.13	2.0	0.2	172.5	2.3	2.6	2.7	140.7	4.2	0.3

Fig. 7 **a** Schematic diagram of the mannan degradation pathway. **b** Expression of genes encoding enzymes in the mannan degradation pathway and key transcription factors for cellulose and hemicellulose utilization across different conditions. Expression values for genes with more than two fold up-regulation relative to the starvation condition (NoC) are shaded in pink, while those with more than two fold down-regulation are shaded in blue. *XG* xyloglucan, *MLG* mixed-linkage glucan

We next used Pc2, Pc1, and Px1 strains to test their growth on pure mannan as a sole carbon source. Remarkably, the Pc2 strain exhibited a robust growth phenotype on pure mannan (Fig. 8d) and secreted 14 times more protein than the WT strain. However, both Pc1 and Px1 strains showed a poor growth phenotype similar to WT in pure mannan (Fig. 8d). Thus, these experimental results validated our hypothesis that mis-expression of *clr-2*, and thus induction of *gh5-7* and *gh2-1*, would restore growth on mannan. Furthermore, the Pc2 strain also exhibited robust growth on XG, while the Pc1 and Px1 strains showed growth characteristics that were similar to WT on XG as a sole carbon source (Fig. 8e). These novel insights into the regulation of the mannan and XG

(See figure on next page.)

Fig. 8 *clr-2* plays a major role in mannan and xyloglucan (XG) degradation. **a**–**c** Growth phenotypes of WT, Δ*gh5-7*, Δ*clr-1*, Δ*clr-2*, Δ*xlr-1*, Pc1 (P*ccg-1-clr-1*), Pc2 (P*ccg-1-clr-2*), and Px1 (P*ccg-1-xlr-*1) strains of *N. crassa* in the medium containing glucomannan as the sole carbon source after growth for 4 days. **a** Photograph of 3 mL cultures with replicates in 24-well deep-well plates. **b** Fungal dry weights after 4 days. Bars represent standard deviations. The asterisk indicates a significant difference from WT with an unadjusted *p* value of <0.003 using one-way ANOVA. **c** Secreted protein in culture supernatants (SN) relative to WT. Bars represent standard deviations. The concentration of secreted protein is shown relative to WT, which was set to 100%. **d** Growth phenotypes of WT, Pc1, Pc2, and Px1 strains of *N. crassa* in the medium containing pure mannan as the sole carbon source after growth for 4 days. **e** Growth phenotypes of WT, Pc1, Pc2, and Px1 strains of *N. crassa* in the medium containing XG as the sole carbon source after growth for 4 days


a
Glucomannan
Replicate 1
Replicate 2
Replicate 3
WT
Δgh5-7
Δclr-1
Δclr-2
Δxlr-1
Pc1
Pc2
Px1
b
Glucomannan
Dry weight (mg)
35
30
25
20
15
10
5
0
WT
Δgh5-7
Δclr-1
Δclr-2
Δxlr-1
Pc1
Pc2
Px1
c
Glucomannan
Protein in culture SN (% of WT)
800
600
400
200
100
50
0
WT
Δgh5-7
Δclr-1
Δclr-2
Δxlr-1
Pc1
Pc2
Px1
d
Mannan
Replicate 1
Replicate 2
Replicate 3
WT
Pc1
Pc2
Px1
e
Xyloglucan
Replicate 1
Replicate 2
Replicate 3
WT
Pc1
Pc2
Px1

degradation pathway by CLR-2 will aid future efforts to engineer improved strains for degradation of lignocellulosic biomass.

Comparison of the PCWDN with genome-scale metabolic networks of other fungi

A genome-scale metabolic model (iJDZ836) containing 836 metabolic genes that encode 1027 unique enzymatic activities was previously published for *N. crassa* [34] and captures biochemical reactions for catabolism of simple nutrients, central and energy metabolism, and biosynthesis of biomass precursors. By comparing reactions and genes in our PCWDN with those in iJDZ836, we found that 167 out of 202 PCWDN reactions (>82%) and 105 out of 168 PCWDN genes (>62%) were not contained in the model (Additional file 12). Additionally, among the 63 common genes, 23 PCWDN genes (>36%) had incorrect or outdated annotations (Additional file 12).

Apart from *N. crassa*, genome-scale metabolic models have also been reconstructed for a few other ascomycetes species, including *A. nidulans* [29], *Aspergillus niger* [30], *Aspergillus oryzae* [31], *Aspergillus terreus* [32], and *Penicillium chrysogenum* [33]. Using OrthoMCL [71, 72], we searched the genomes of these species for ortholog/paralogs of the 168 *N. crassa* PCWDN genes, and found 178 (*A. nidulans*), 160 (*A. niger*), 197 (*A. oryzae*), 192 (*A. terreus*), and 164 (*P. chrysogenum*) orthologous/paralogous genes ("Methods"; Additional file 12). To assess the coverage of polysaccharide degradation pathways in the reconstructed metabolic models for these species [29–33], the overlap between genes in other fungal metabolic models and ortholog/paralogs of the PCWDN genes was determined. Of the PCWDN ortholog/paralogs, 40–70% were not accounted for in the other filamentous fungal metabolic models (Additional file 12). These analyses highlight the significant knowledge gaps specific to pathways for degradation and utilization of plant cell wall polysaccharides in published genome-scale metabolic models for other ascomycete species.

Conclusions

Here, we have taken a comprehensive approach, using diverse datasets, to define genes in filamentous fungi involved in the deconstruction of plant biomass, using

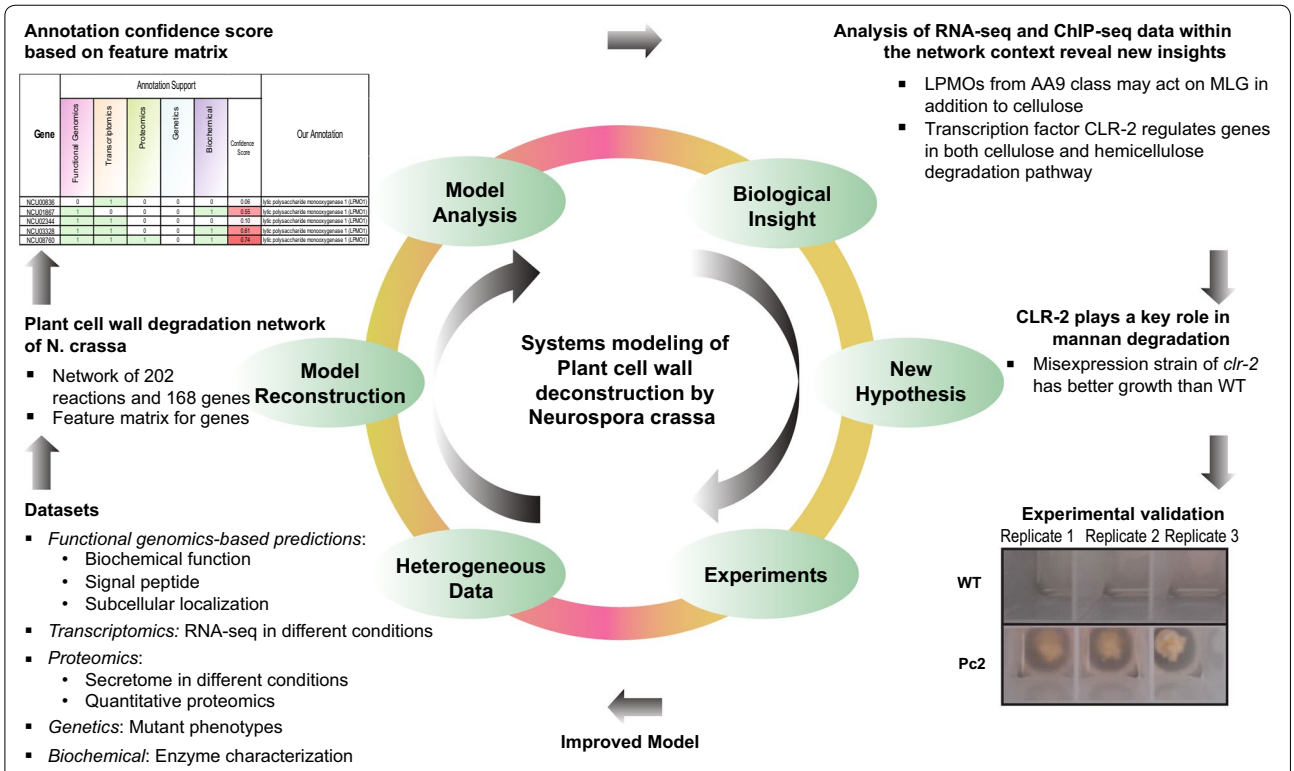


Fig. 9 Schematic diagram summarizing the systems approach undertaken here to reconstruct and analyze the plant cell wall degradation network (PCWDN) of *N. crassa*. We compiled information from diverse sources to build the network of biochemical reactions responsible for degrading plant cell wall polysaccharides into simple sugars. The combined annotation information was encapsulated in the form of a feature matrix and a simple method was devised to assign an annotation confidence score to each PCWDN gene. To demonstrate the utility of our PCWDN, we performed comparative transcriptomics analysis using RNA-seq data and integrated genome-wide binding data from ChIP-seq experiments for three key transcription factors regulating the plant cell wall degradation response. Integration of next-generation sequencing data within the PCWDN led to the hypothesis that CLR-2 is a key TF for deconstruction of mannan and XG, a hypothesis that was subsequently validated through experimentation

N. crassa as a model (Fig. 9). From these analyses, we developed hypotheses regarding the regulation of mannan degradation in *N. crassa* and experimentally tested the hypothesis that the transcription factor CLR-2, previously characterized to regulate cellulose utilization [16, 17], also plays a role in mannan degradation. Interestingly, *clr-2* is the ortholog of the transcriptional regulator ManR, which regulates mannan utilization in *A. oryzae* [89]. These data support the view that the role of *clr-2* orthologs in the regulation of genes involved in mannan utilization is conserved among filamentous fungi.

Current metabolic models for *N. crassa* [34] and other filamentous fungi [29–33] have significant knowledge gaps in the degradation pathways for most plant cell wall polysaccharides. For example, the *N. crassa* metabolic model iJDZ836 [34] has significant knowledge gaps regarding the pathways for degradation and utilization of plant cell wall polysaccharides. The iJDZ836 model contains neither the pectin degradation nor the D-galacturonic acid utilization pathway (Additional file 12) and lacks degradation pathways for most hemicellulosic polysaccharides including mannan, XG, and MLG (Additional file 12). Additionally, degradation pathways for cellulose and xylan were not captured in detail. For example, the AA9 LPMOs were incorrectly annotated as endo-β-1,4-glucanases and therefore do not contain reactions describing the oxidative cleavage of cellulose by LPMOs (Additional file 12). Likewise, the metabolic models for other filamentous fungi do not capture the LPMO-associated reactions and pectin degradation pathway. These limitations render the available genome-scale metabolic models for filamentous fungi not well suited for the investigation of plant cell wall deconstruction, while our PCWDN will serve as a valuable resource to fill this gap. In the future, our reconstructed network is expected to play a central role in the systems analysis of complex experimental datasets and will yield many more novel insights into plant cell wall deconstruction by filamentous fungi.

Methods

Databases and functional genomics tools

A list of predicted carbohydrate-active enzymes in the *N. crassa* genome was compiled from two databases, the CAZY [41, 90] and the *N. crassa* e-Compendium [42, 91]. Following manual curation, we generated an updated list of predicted carbohydrate-active enzymes in *N. crassa* (Additional file 2). TransportDB [43, 92] was used to obtain a list of predicted transporters in the *N. crassa* genome. Proteins destined for the secretory pathway are likely to have a signal peptide sequence in their N terminus, and we used two prediction tools, SignalP [51, 93] and Phobius [52, 94], to predict the presence of signal peptides in amino acid sequences of PCWDN enzymes (Additional file 3). In order to assign the subcellular localization of enzymes, we used WoLF PSORT [53, 95] and ProtComp [96] (Additional file 3). The mycoCLAP database [70, 97] was used extensively to compile the list of biochemically characterized lignocellulose-active proteins of fungal origin. OrthoMCL [71, 72, 98] is a tool to identify orthologous gene pairs across eukaryotic genomes. We have used OrthoMCL to determine the ortholog/paralogs of *N. crassa* PCWDN genes in more than 40 fungal genomes.

Strains and culture conditions

The *N. crassa* WT reference strain was OR74A (FGSC 2489), which was the background for all mutant strains [12, 88]. The deletion strains for *clr-1* (FGSC 11029), *clr-2* (FGSC 15835), and *xlr-1* (FGSC 11066 and 11067) were obtained from the Fungal Genetics Stock Center (FGSC) [99]. The mis-expression strains for *clr-1* (Pc1), *clr-2* (Pc2), and *xlr-1* (Px1) were obtained from a previous study [76]. The deletion strain for *gh5-7* (NCU08412) was not available in the FGSC collection and was constructed following standard procedures (Δ*gh5-7*; ΔNCU08412) [12, 88]. Briefly, the 5′ upstream and 3′ downstream genomic regions surrounding NCU08412 were PCR amplified from WT genomic DNA and joined through fusion PCR with the hygromycin phosphotransferase (hph) knockout cassette [88]. The resulting amplicon was transformed into FGSC 9718 (Δ*mus-51*) and selected on hygromycin slants. A homokaryotic strain was obtained through microconidia selection on water agar plates yielding the strain (Δ*gh5-7*::*hyg*^{R}; Δ*mus-51*).

All *N. crassa* strains were pre-grown for 24 h on 3 mL agar slants of Vogel's minimal media (VMM) [100] with 2% sucrose at 30 °C under dark conditions. The slants were placed under constant light at 25 °C to stimulate conidia production. For flask cultures, conidia were collected and inoculated (10^6 conidia/mL) into 100 mL liquid VMM (2% sucrose) at 25 °C under constant light and shaking (200 rpm).

Media shift experiments

Media shift experiments were performed in triplicate and followed the procedure described earlier in Coradetti et al. [16] and Znameroski et al. [18] to ensure optimal comparability with the previously published RNA-seq datasets. First, using shake flasks (200 rpm), *N. crassa* cultures were pre-grown from conidia for 16 h in 100 mL of VMM [100] with 2% sucrose. Next, the mycelia were passed over a Whatman glass microfiber filter and washed three times with VMM without a carbon source (NoC). The mycelial mass was then transferred to new shake flasks with 100 mL of VMM containing a specific

carbon source [2% XG (P-XYGLN, Megazyme) or 2% mannan (P-MANCB, Megazyme) or 2% MLG (P-BGBM, Megazyme) or 2% Starch]. After 4 h in the new carbon source, the mycelia were harvested over a filter, flash frozen in liquid nitrogen, and stored at −80 °C. Total RNA was extracted for library generation using the standard procedures as described in Tian et al. [15].

RNA sequencing and data analyses

Single-end libraries were prepared for RNA sequencing (RNA-seq) using an Illumina kit (RS-100-0801) following standard protocols as described in Coradetti et al. [16]. The cDNA libraries were sequenced on the Illumina HiSeq 2000 platform at the Vincent J. Coates Genomic Sequencing Laboratory, University of California, Berkeley. Three independent biological replicates of media shift cultures were sampled for *N. crassa* WT strains grown on VMM [100] with one of four carbon sources: XG, mannan, MLG, or starch. Profiling data for each of these carbon sources are contained in Additional file 4 and have also been deposited to the Gene Expression Omnibus (GEO) database [101] (Accession No. GSE90611). Previous studies [16, 19] generated profiling data for *N. crassa* WT strain grown in VMM with one of six carbon sources [sucrose, cellobiose (CB), Avicel, xylan, pectin, and OPP] or No Carbon (NoC) source, and these libraries were downloaded from the GEO database (Additional file 4).

Sequenced libraries were mapped to the current version of the *N. crassa* OR74A genome (v12) using Tophat v2.0.5 [102–104]. Transcript abundance was estimated in FPKM (fragments per kilobase of transcript per million mapped reads) using Cufflinks v2.0.2 [102, 103, 105] with options of upper quartile normalization and mapping against reference isoforms. Differential expression analysis was conducted using Cuffdiff v2.0.2 [102, 103]. Genes with a multiple-hypothesis adjusted p value below 0.05 and at least twofold induction were determined to be significantly differentially expressed between conditions.

Starting from average FPKM of genes across RNA-seq library replicates for a condition, hierarchical clustering was performed using Cluster 3.0 software suite [74, 106]. Before clustering, genes were filtered out that displayed consistently low expression (<10 FPKM) in all conditions. FPKM were log-transformed, normalized across conditions, and centered on the geometric mean across conditions on a per-gene basis. The average linkage method was used for cluster generation, with Pearson's correlation as a similarity measure. Visualization of clusters was performed using GENE-E software suite [107].

Growth assays

Growth assays on cell wall substrates were performed in 3 mL liquid cultures in 24-well plate format (GE Healthcare Life Sciences 7701-5102) with breathable sealing tape (Fisher Scientific 1256705). 10^6 conidia/mL were inoculated into VMM with 0.5% wt/vol 1,4-β-D-mannan from carob (Megazyme P-MANCB), Konjac glucomannan (Megazyme P-GLCML), or xyloglucan from tamarind (Megazyme P-XYGLN) as the carbon source. Cultures were grown for 48 h at 25 °C with constant light and shaking at 250 rpm. At the end of the incubation, mycelia were concentrated by centrifugation at 3000 RCF for 10 min. Culture supernatants were then assayed for soluble protein with Bio-Rad protein assay dye reagent (Bio-Rad 500-0006), using bovine serum albumin (NEB 9001S) as the protein standard. Mycelia were washed twice in water and lyophilized before weighing for biomass determination.

Additional files

Additional file 1. List of reactions, compounds and genes in the plant cell wall degradation network (PCWDN) of *N. crassa*. The first, second and third sheets contain the list of reactions, genes and compounds, respectively, in the PCWDN. The fourth sheet gives the participation of the PCWDN genes in the cellulose, hemicellulose, pectin and starch degradation pathways. The fifth and sixth sheets contain the feature matrix and annotation confidence scores, respectively, for genes in the PCWDN of *N. crassa* based on functional genomics, transcriptomics, proteomics and genetics data as well as biochemical characterizations. The seventh sheet contains a list of research articles utilized to reconstruct the PCWDN. The eighth sheet contains information on the structural units comprising the backbone and side chains of different plant cell wall polysaccharides such as cellulose, xylan, xyloglucan, mannan, galactomannan, glucomannan, galactoglucomannan, mixed-linkage glucan, homogalacturonan, xylogalacturonan, rhamnogalacturonan I, amylose and amylopectin. *NoC* no carbon, *CB* cellobiose, *XG* xyloglucan, *MLG* mixed-linkage glucan, *OPP* orange peel powder.

Additional file 2. List of predicted genes coding for carbohydrate-active enzymes (CAZY) in *N. crassa* and the plant cell wall degradation network (PCWDN). List of CAZY genes in *N. crassa* was obtained from two sources: CAZY database [41] and the *N. crassa* e-Compendium [42]. The first sheet contains the consensus list of predicted CAZY genes in *N. crassa*. The second sheet contains the list of predicted CAZY genes in *N. crassa* from CAZY database. The third sheet contains the list of predicted glycoside hydrolase (GH) genes in *N. crassa* from e-Compendium [42]. The fourth sheet contains the list of CAZY classes known to be involved in native cell wall remodelling. *NoC* no carbon, *CB* cellobiose, *XG* xyloglucan, *MLG* mixed-linkage glucan, *OPP* orange peel powder.

Additional file 3. Functional genomics-based annotation support for genes in the plant cell wall degradation network (PCWDN) of *N. crassa*. The table contains information from following sources: CAZY, BROAD, TransportDB, SignalP, Phobius, WoLF PSORT and ProtComp.

Additional file 4. Transcriptomics-based annotation support for genes in the plant cell wall degradation network (PCWDN) of *N. crassa*. The first sheet contains information on the transcriptomics-based annotation support for each gene in the PCWDN. The second sheet contains a table with the list of RNA-seq libraries for *N. crassa* WT strain grown in different conditions along with their GEO [101] accession numbers and references for the profiling data. The third sheet contains a table with the expression of genes in different conditions, separately, for replicate RNA-seq libraries. The fourth sheet contains a table with the average expression (Mean FPKM) of genes in each condition across replicate RNA-seq libraries. The fifth sheet contains a table with information on differential expression of genes in each condition compared to the No Carbon control. The sixth sheet contains a table with information on differential expression

of genes in each condition compared to the sucrose control. The set of differentially expressed genes was computed using the Cuffdiff package [103]. The seventh and eighth sheets contain genes in the up-regulons and down-regulons, respectively, of *N. crassa* in seven conditions: Avicel, Xylan, Xyloglucan, Mannan, Mixed-Linkage Glucan, Pectin and Starch. *NoC* no carbon, *CB* cellobiose, *XG* xyloglucan, *MLG* mixed-linkage glucan, *OPP* orange peel powder.

Additional file 5. Proteomics-based annotation support for genes in the plant cell wall degradation network (PCWDN) of *N. crassa*. The first sheet contains information on the proteomics-based annotation support for each gene in the PCWDN. The second sheet contains the compiled dataset of secretome for *N. crassa* strains obtained under different conditions from published literature. The third sheet contains the list of references used to compile this dataset.

Additional file 6. Compiled dataset of experimentally validated deletion strains with growth-deficient phenotypes compared to WT for genes in the plant cell wall degradation network (PCWDN) of *N. crassa*.

Additional file 7. Biochemical characterization of genes in the plant cell wall degradation network (PCWDN) of *N. crassa*.

Additional file 8. Comparative analysis of the down-regulons for the seven conditions. Analogous to up-regulon (for brevity referred to as 'regulon' in main text), the down-regulon for a growth condition can be defined as genes that were downregulated and differentially expressed in comparison to both NoC and sucrose controls. **a** Horizontal bar plot shows the size of the down-regulons for seven conditions. The down-regulons for Avicel, xylan, XG, mannan, MLG, pectin and starch, were determined to contain 3, 45, 41, 30, 5, 35 and 236 genes, respectively (Additional file 4). Vertical bar plot shows the 11 intersection sets among the seven down-regulons with 4 or more genes and was generated using UpSetR [108]. It is seen that 220 out of 236 genes in the starch down-regulon, 26 out of 45 genes in the xylan down-regulon, 20 out of 41 genes in the XG down-regulon, 17 out of 35 genes in the pectin down-regulon, and 8 out of 30 genes in the mannan down-regulon, have no overlap with other down-regulons. It is also seen that 10 genes are common between XG and mannan down-regulons, 6 genes between xylan and starch down-regulons, 5 genes between xylan and pectin down-regulons, 4 genes between xylan and XG down-regulons, 4 genes between mannan and pectin down-regulons, and 4 genes between pectin and starch down-regulons. **b** Functional category analysis [73] of the 11 intersection sets among the down-regulons for seven conditions with four or more genes. Information on the functional category of *N. crassa* genes was obtained from Munich Information Center for Protein Sequence (MIPS) database [73, 109]. The 'Others' category includes genes with functional categorization different from the 10 categories listed in the legend. The 'Unknown function' category includes genes with unclassified or unknown function. The relative contribution of a functional category to each set of genes is depicted with the total number of genes in each pool equal to 100%.

Additional file 9. Functional category analysis of the 26 intersection sets among the up-regulons for seven conditions with 5 or more genes (shown in Fig. 4c). Information on the functional category of *N. crassa* genes was obtained from Munich Information Center for Protein Sequence (MIPS) database [73, 109]. The 'Others' category includes genes with functional categorization different from the 10 categories listed in the legend. The 'Unknown function' category includes genes with unclassified or unknown function. The relative contribution of a functional category to each set of genes is depicted with the total number of genes in each pool equal to 100%.

Additional file 10. Hierarchical clustering of genes in the plant cell wall degradation network (PCWDN) of *N. crassa* based on RNA-seq data obtained in nine different conditions. The first sheet lists the different clusters in the same order as shown in Fig. 5.

Additional file 11. Expression of characterized and predicted AA9 LPMOs in *N. crassa* grown on four plant cell wall polysaccharides with D-glucose backbone: Avicel, xyloglucan (XG), mixed-linkage glucan (MLG) and starch. In comparison to Avicel and MLG, the expression of LPMOs was much lower on XG and negligible on starch.

Additional file 12. Comparative analysis of the plant cell wall degradation network (PCWDN) with the genome-scale metabolic models of *N. crassa* and other filamentous fungi. The first and second sheets compare the list of reactions and genes, respectively, in the PCWDN and genome-scale metabolic model iJDZ836 of *N. crassa*. The third sheet gives the occurrence of orthologs or paralogs of *N. crassa* PCWDN genes in genome-scale metabolic models of other filamentous fungi.

Authors' contributions
AS, JPC, STC, JPB, NDP, and NLG designed the research; AS, JPC, STC, and JPB performed the research; AS, JPC, STC, JPB, and NLG analyzed the data; and AS, JPC, JPB, JAE, NDP, and NLG wrote the paper. All authors read and approved the final manuscript.

Author details
[1] Institute for Systems Biology, Seattle, WA 98109, USA. [2] Energy Biosciences Institute, University of California Berkeley, Berkeley, CA 94704, USA. [3] The Institute of Mathematical Sciences, Homi Bhabha National Institute, Chennai 600113, India. [4] The Abdus Salam International Centre for Theoretical Physics, 34151 Trieste, Italy. [5] Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA. [6] Holzforschung München, TUM School of Life Sciences Weihenstephan, Technische Universität München, 85354 Freising, Germany.

Acknowledgements
The authors thank Vincent Wu, Carly Grant, and Hillary Tunggal for technical assistance, and Xin Li and William Beeson for insightful discussions. The authors also thank the Fungal Genetics Stock Center and the *Neurospora* Genome Project for continuous support and acknowledge the use of materials generated by NIH Grant P01 GM068087 "Functional analysis of a model filamentous fungus."

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
The datasets supporting the conclusions of this article are included within the article and its additional files. RNA sequencing dataset generated for this article has been deposited to the Gene Expression Omnibus (GEO) database (Accession No. GSE90611).

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Funding
This work was supported by Grants from the Energy Biosciences Institute to NLG and NDP, and National Institutes of Health NRSA Trainee Grant 2 TR32 GM 7127-36 A1 to STC.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 June 2017 Accepted: 5 September 2017
Published online: 21 September 2017

References
1. Rubin EM. Genomics of cellulosic biofuels. Nature. 2008;454:841–5.
2. Pauly M, Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J. 2008;54:559–68.

3. Carroll A, Somerville C. Cellulosic biofuels. Annu Rev Plant Biol. 2009;60:165–82.
4. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 2007;315:804–7.
5. Glass NL, Schmoll M, Cate JH, Coradetti S. Plant cell wall deconstruction by ascomycete fungi. Annu Rev Microbiol. 2013;67:477–98.
6. Boddy L, Watkinson SC. Wood decomposition, higher fungi, and their role in nutrient redistribution. Can J Bot. 1995;73:1377–83.
7. Kubicek CP, Mikus M, Schuster A, Schmoll M, Seiboth B. Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. Biotechnol Biofuels. 2009;2:19.
8. Peterson R, Nevalainen H. Trichoderma reesei RUT-C30—thirty years of strain improvement. Microbiology. 2012;158:58–68.
9. Stephanopoulos G. Challenges in engineering microbes for biofuels production. Science. 2007;315:801–4.
10. Gilbert HJ. The biochemistry and structural biology of plant cell wall deconstruction. Plant Physiol. 2010;153:444–55.
11. Davis RH, Perkins DD. Neurospora: a model of model microbes. Nat Rev Genet. 2002;3:397–403.
12. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, et al. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature. 2003;422:859–68.
13. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, et al. Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev. 2004;68:1–108.
14. Dunlap JC, Borkovich KA, Henn MR, Turner GE, Sachs MS, Glass NL, McCluskey K, Plamann M, Galagan JE, Birren BW, et al. Enabling a community to dissect an organism: overview of the Neurospora functional genomics project. Adv Genet. 2007;57:49–96.
15. Tian C, Beeson WT, Iavarone AT, Sun J, Marletta MA, Cate JH, Glass NL. Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. Proc Natl Acad Sci USA. 2009;106:22157–62.
16. Coradetti ST, Craig JP, Xiong Y, Shock T, Tian C, Glass NL. Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. Proc Natl Acad Sci USA. 2012;109:7397–402.
17. Sun J, Tian C, Diamond S, Glass NL. Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. Eukaryot Cell. 2012;11:482–93.
18. Znameroski EA, Coradetti ST, Roche CM, Tsai JC, Iavarone AT, Cate JH, Glass NL. Induction of lignocellulose-degrading enzymes in *Neurospora crassa* by cellodextrins. Proc Natl Acad Sci USA. 2012;109:6012–7.
19. Benz JP, Chau BH, Zheng D, Bauer S, Glass NL, Somerville CR. A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptations. Mol Microbiol. 2014;91:275–99.
20. Cai P, Gu R, Wang B, Li J, Wan L, Tian C, Ma Y. Evidence of a critical role for cellodextrin transporter 2 (CDT-2) in both cellulose and hemicellulose degradation and utilization in *Neurospora crassa*. PLoS ONE. 2014;9:e89330.
21. Galazka JM, Tian C, Beeson WT, Martinez B, Glass NL, Cate JH. Cellodextrin transport in yeast for improved biofuel production. Science. 2010;330:84–6.
22. Li S, Du J, Sun J, Galazka JM, Glass NL, Cate JH, Yang X, Zhao H. Overcoming glucose repression in mixed sugar fermentation by co-expressing a cellobiose transporter and a beta-glucosidase in *Saccharomyces cerevisiae*. Mol BioSyst. 2010;6:2129–32.
23. Ha SJ, Galazka JM, Kim SR, Choi JH, Yang X, Seo JH, Glass NL, Cate JH, Jin YS. Engineered Saccharomyces cerevisiae capable of simultaneous cellobiose and xylose fermentation. Proc Natl Acad Sci USA. 2011;108:504–9.
24. Li X, Yu VY, Lin Y, Chomvong K, Estrela R, Park A, Liang JM, Znameroski EA, Feehan J, Kim SR, et al. Expanding xylose metabolism in yeast for plant cell wall conversion to biofuels. eLife. 2015;4:e05896.
25. Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO. Reconstruction of biochemical networks in microorganisms. Nat Rev Microbiol. 2009;7:129–43.
26. Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc. 2010;5:93–121.
27. Price ND, Papin JA, Schilling CH, Palsson BO. Genome-scale microbial in silico models: the constraints-based approach. Trends Biotechnol. 2003;21:162–9.
28. Price ND, Reed JL, Palsson BO. Genome-scale models of microbial cells: evaluating the consequences of constraints. Nat Rev Microbiol. 2004;2:886–97.
29. David H, Ozcelik IS, Hofmann G, Nielsen J. Analysis of *Aspergillus nidulans* metabolism at the genome-scale. BMC Genom. 2008;9:163.
30. Andersen MR, Nielsen ML, Nielsen J. Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol Syst Biol. 2008;4:178.
31. Vongsangnak W, Olsen P, Hansen K, Krogsgaard S, Nielsen J. Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. BMC Genom. 2008;9:245.
32. Liu J, Gao Q, Xu N, Liu L. Genome-scale reconstruction and in silico analysis of *Aspergillus terreus* metabolism. Mol BioSyst. 2013;9:1939–48.
33. Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. PLoS Comput Biol. 2013;9:e1002980.
34. Dreyfuss JM, Zucker JD, Hood HM, Ocasio LR, Sachs MS, Galagan JE. Reconstruction and validation of a genome-scale metabolic model for the filamentous fungus *Neurospora crassa* using FARM. PLoS Comput Biol. 2013;9:e1003126.
35. Mohnen D. Pectin structure and biosynthesis. Curr Opin Plant Biol. 2008;11:266–77.
36. Scheller HV, Ulvskov P. Hemicelluloses. Annu Rev Plant Biol. 2010;61:263–89.
37. Damager I, Engelsen SB, Blennow A, Moller BL, Motawia MS. First principles insight into the alpha-glucan structures of starch: their synthesis, conformation, and hydration. Chem Rev. 2010;110:2049–80.
38. Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, Tuohy MG, Kloareg B, Stengel DB. Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol. 2011;62:567–90.
39. van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol. 2011;91:1477–92.
40. Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G. Hemicellulose biosynthesis. Planta. 2013;238:627–42.
41. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 2009;37:D233–8.
42. Perkins DD, Radford A, Sachs MS. The Neurospora compendium: chromosomal loci. San Diego: Academic; 2001.
43. Ren Q, Chen K, Paulsen IT. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res. 2007;35:D274–9.
44. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen JC, Johansen KS, Krogh KB, Jorgensen CI, Tovborg M, Anthonsen A, et al. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proc Natl Acad Sci USA. 2011;108:15079–84.
45. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. Appl Environ Microbiol. 2011;77:7007–15.
46. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by *Neurospora crassa*. ACS Chem Biol. 2011;6:1399–406.
47. Beeson WT, Phillips CM, Cate JH, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J Am Chem Soc. 2012;134:890–2.
48. Kittl R, Kracher D, Burgstaller D, Haltrich D, Ludwig R. Production of four *Neurospora crassa* lytic polysaccharide monooxygenases in *Pichia pastoris* monitored by a fluorimetric assay. Biotechnol Biofuels. 2012;5:79.
49. Agger JW, Isaksen T, Varnai A, Vidal-Melgosa S, Willats WG, Ludwig R, Horn SJ, Eijsink VG, Westereng B. Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. Proc Natl Acad Sci USA. 2014;111:6287–92.

50. Vu VV, Beeson WT, Span EA, Farquhar ER, Marletta MA. A family of starch-active polysaccharide monooxygenases. Proc Natl Acad Sci USA. 2014;111:13822–7.
51. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8:785–6.
52. Kall L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. 2007;35:W429–32.
53. Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007;35:W585–7.
54. Sun J, Glass NL. Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. PLoS ONE. 2011;6:e25654.
55. Maddi A, Bowman SM, Free SJ. Trifluoromethanesulfonic acid-based proteomic analysis of cell wall and secreted proteins of the ascomycetous fungi *Neurospora crassa* and *Candida albicans*. Fungal Genet Biol. 2009;46:768–81.
56. Phillips CM, Iavarone AT, Marletta MA. Quantitative proteomic approach for cellulose degradation by *Neurospora crassa*. J Proteome Res. 2011;10:4177–85.
57. Xiong Y, Coradetti ST, Li X, Gritsenko MA, Clauss T, Petyuk V, Camp D, Smith R, Cate JH, Yang F, Glass NL. The proteome and phosphoproteome of *Neurospora crassa* in response to cellulose, sucrose and carbon starvation. Fungal Genet Biol. 2014;72:21–33.
58. Sun J, Phillips CM, Anderson CT, Beeson WT, Marletta MA, Glass NL. Expression and characterization of the *Neurospora crassa* endoglucanase GH5-1. Protein Expr Purif. 2011;75:147–54.
59. Sygmund C, Kracher D, Scheiblbrandner S, Zahma K, Felice AK, Harreither W, Kittl R, Ludwig R. Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. Appl Environ Microbiol. 2012;78:6161–71.
60. Yazdi MT, Radford A, Keen JN, Woodward JR. Cellulase production by *Neurospora crassa*: purification and characterization of cellulolytic enzymes. Enzym Microb Technol. 1990;12:120–3.
61. Zhang R, Fan Z, Kasuga T. Expression of cellobiose dehydrogenase from *Neurospora crassa* in *Pichia pastoris* and its purification and characterization. Protein Expr Purif. 2011;75:63–9.
62. Crepin VF, Faulds CB, Connerton IF. A non-modular type B feruloyl esterase from *Neurospora crassa* exhibits concentration-dependent substrate inhibition. Biochem J. 2003;370:417–27.
63. Konishi T, Kotake T, Soraya D, Matsuoka K, Koyama T, Kaneko S, Igarashi K, Samejima M, Tsumuraya Y. Properties of family 79 beta-glucuronidases that hydrolyze beta-glucuronosyl and 4-*O*-methyl-beta-glucuronosyl residues of arabinogalactan-protein. Carbohydr Res. 2008;343:1191–201.
64. Du J, Li S, Zhao H. Discovery and characterization of novel D-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. Mol BioSyst. 2010;6:2150–6.
65. Nair N, Zhao H. Biochemical characterization of an L-xylulose reductase from *Neurospora crassa*. Appl Environ Microbiol. 2007;73:2001–4.
66. Chomvong K, Kordic V, Li X, Bauer S, Gillespie AE, Ha SJ, Oh EJ, Galazka JM, Jin YS, Cate JH. Overcoming inefficient cellobiose fermentation by cellobiose phosphorylase in the presence of xylose. Biotechnol Biofuels. 2014;7:85.
67. Nihira T, Saito Y, Nishimoto M, Kitaoka M, Igarashi K, Ohtsubo K, Nakai H. Discovery of cellobionic acid phosphorylase in cellulolytic bacteria and fungi. FEBS Lett. 2013;587:3556–61.
68. Benz JP, Protzko RJ, Andrich JM, Bauer S, Dueber JE, Somerville CR. Identification and characterization of a galacturonic acid transporter from *Neurospora crassa* and its application for *Saccharomyces cerevisiae* fermentation processes. Biotechnol Biofuels. 2014;7:20.
69. Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR. Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc Natl Acad Sci USA. 2006;103:11417–22.
70. Strasser K, McDonnell E, Nyaga C, Wu M, Wu S, Almeida H, Meurs MJ, Kosseim L, Powlowski J, Butler G, Tsang A. mycoCLAP, the database for characterized lignocellulose-active proteins of fungal origin: resource and text mining curation support. Database (Oxford). 2015;2015:bav008.
71. Chen F, Mackey AJ, Stoeckert CJ Jr, Roos DS. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. Nucleic Acids Res. 2006;34:D363–8.
72. Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Iodice JB, Shanmugam D, Roos DS, Stoeckert CJ Jr. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. Curr Protoc Bioinform. 2011;35(6):12.
73. Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, Tetko I, Guldener U, Mannhaupt G, Munsterkotter M, Mewes HW. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. Nucleic Acids Res. 2004;32:5539–45.
74. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA. 1998;95:14863–8.
75. Calero-Nieto F, Di Pietro A, Roncero MI, Hera C. Role of the transcriptional activator xlnR of *Fusarium oxysporum* in regulation of xylanase genes and virulence. Mol Plant Microbe Interact. 2007;20:977–85.
76. Coradetti ST, Xiong Y, Glass NL. Analysis of a conserved cellulase transcriptional regulator reveals inducer-independent production of cellulolytic enzymes in *Neurospora crassa*. Microbiologyopen. 2013;2:595–609.
77. Marui J, Kitamoto N, Kato M, Kobayashi T, Tsukagoshi N. Transcriptional activator, AoXlnR, mediates cellulose-inductive expression of the xylanolytic and cellulolytic genes in *Aspergillus oryzae*. FEBS Lett. 2002;528:279–82.
78. Stricker AR, Grosstessner-Hain K, Wurleitner E, Mach RL. Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. Eukaryot Cell. 2006;5:2128–37.
79. van Peij NN, Gielkens MM, de Vries RP, Visser J, de Graaff LH. The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol. 1998;64:3615–9.
80. Yao G, Li Z, Gao L, Wu R, Kan Q, Liu G, Qu Y. Redesigning the regulatory pathway to enhance cellulase production in *Penicillium oxalicum*. Biotechnol Biofuels. 2015;8:71.
81. Craig JP, Coradetti ST, Starr TL, Glass NL. Direct target network of the *Neurospora crassa* plant cell wall deconstruction regulators CLR-1, CLR-2, and XLR-1. MBio. 2015;6:e01452–5.
82. Alon U. Network motifs: theory and experimental approaches. Nat Rev Genet. 2007;8:450–61.
83. Mangan S, Alon U. Structure and function of the feed-forward loop network motif. Proc Natl Acad Sci USA. 2003;100:11980–5.
84. Mangan S, Zaslaver A, Alon U. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J Mol Biol. 2003;334:197–204.
85. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. Network motifs: simple building blocks of complex networks. Science. 2002;298:824–7.
86. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat Genet. 2002;31:64–8.
87. Shimizu M, Kaneko Y, Ishihara S, Mochizuki M, Sakai K, Yamada M, Murata S, Itoh E, Yamamoto T, Sugimura Y, et al. Novel beta-1,4-mannanase belonging to a new glycoside hydrolase family in *Aspergillus nidulans*. J Biol Chem. 2015;290:27914–27.
88. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci USA. 2006;103:10352–7.
89. Ogawa M, Kobayashi T, Koyama Y. ManR, a novel Zn(II)2Cys6 transcriptional activator, controls the beta-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol. 2012;49:987–95.
90. CAZY database. http://www.cazy.org/. Accessed July 2013.
91. *N. crassa* e-compendium. http://www.bioinf.leeds.ac.uk/~gen6ar/newgenelist/genes/gene_list.htm. Accessed July 2013.
92. TransportDB. http://www.membranetransport.org/. Accessed July 2013.
93. SignalP. http://www.cbs.dtu.dk/services/SignalP/. Accessed July 2013.
94. Phobius. http://phobius.sbc.su.se/. Accessed July 2013.
95. WoLF PSORT. http://psort.hgc.jp/. Accessed July 2013.
96. ProtComp. http://linux1.softberry.com/berry.phtml?topic=protcomppl&group=programs&subgroup=proloc. Accessed July 2013.

97. mycoCLAP database http://mycoclap.fungalgenomics.ca/. Accessed July 2013.
98. OrthoMCL. http://orthomcl.org/orthomcl/.
99. Fungal Genetics Stock Center. http://www.fgsc.net/. Accessed July 2013.
100. Vogels H. A convenient growth medium for Neurospora. Microbiol Genet Bull. 1956;13:42–3.
101. Gene Expression Omnibus Database. https://www.ncbi.nlm.nih.gov/geo/. Accessed July 2013.
102. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105–11.
103. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–78.
104. Tophat. http://tophat.cbcb.umd.edu/.
105. Cufflinks. http://cufflinks.cbcb.umd.edu.
106. Cluster 3.0. http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm.
107. GENE-E. http://www.broadinstitute.org/cancer/software/GENE-E/.
108. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. IEEE Trans Vis Comput Graph. 2014;20:1983–92.
109. FuncatDB. http://mips.helmholtz-muenchen.de/funcatDB/. Accessed July 2013.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central