Published in final edited form as:
Mol Microbiol. 2014 January ; 91(2): 275–299. doi:10.1111/mmi.12459.

A comparative systems analysis of polysaccharide-elicited responses in N*eurospora crassa* reveals carbon source-specific cellular adaptations

J. Philipp Benz[1,*], **Bryant H. Chau**[1,3], **Diana Zheng**[1], **Stefan Bauer**[1], **N. Louise Glass**[1,2], and **Chris R. Somerville**[1,2]

[1]Energy Biosciences Institute, University of California Berkeley, Berkeley, California, USA

[2]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California, USA

Summary

Filamentous fungi are powerful producers of hydrolytic enzymes for the deconstruction of plant cell wall polysaccharides. However, the central question of how these sugars are perceived in the context of the complex cell wall matrix remains largely elusive. To address this question in a systematic fashion we performed an extensive comparative systems analysis of how the model filamentous fungus *Neurospora crassa* responds to the three main cell wall polysaccharides: pectin, hemicellulose and cellulose. We found the pectic response to be largely independent of the cellulolytic one with some overlap to hemicellulose, and in its extent surprisingly high, suggesting advantages for the fungus beyond being a mere carbon source. Our approach furthermore allowed us to identify carbon source-specific adaptations, such as the induction of the unfolded protein response on cellulose, and a commonly induced set of 29 genes likely involved in carbon scouting. Moreover, by hierarchical clustering we generated a co-expression matrix useful for the discovery of new components involved in polysaccharide utilization. This is exemplified by the identification of *lat-1*, which we demonstrate to encode for the physiologically relevant arabinose transporter in *Neurospora*. The analyses presented here are an important step towards understanding fungal degradation processes of complex biomass.

Keywords

pectin; xylan; cellulose; systems analysis; *Neurospora crassa*; polysaccharide perception

Introduction

The cell walls of all higher plants are composed of four polymeric building blocks: the polyphenol lignin and the three polysaccharides cellulose, hemicellulose and pectin. Together with a number of enzymes, structural proteins and proteoglycans these components form an intricately linked network that fulfills a multitude of functions for the plant (Popper *et al.*, 2011; Somerville *et al.*, 2004), but can also be used as an alternative energy source such as for the fermentation to biofuels (Chundawat *et al.*, 2011; Jordan *et al.*, 2012; Pauly & Keegstra, 2008; Youngs & Somerville, 2012; Somerville *et al.*, 2010). Cellulose is the major structural and load-bearing polysaccharide of the plant cell wall and is composed of unbranched, linear chains of β-1,4-linked glucan. Four enzyme classes are known to be

[*]corresponding author ph.benz@berkeley.edu Phone: +1 (510) 666-2556.
[3]present address: Rinat Laboratories, Pfizer Inc., 230 East Grand Avenue, South San Francisco, CA 94080, USA

involved in cellulose deconstruction, which are classified according to their substrate preferences and mode of action (Carbohydrate Active Enzymes database; http://www.cazy.org/) (Cantarel *et al.*, 2009): endoglucanases (GH5, GH7, GH12, GH45), cellobiohydrolases (GH6, GH7), β-1,4-glucosidases (GH1, GH3) and polysaccharide monooxygenases (LPMOs; formerly GH61) (Beeson *et al.*, 2012; Harris *et al.*, 2010; Levasseur *et al.*, 2013; Phillips *et al.*, 2011).

The term hemicellulose comprises a diverse group of polysaccharides. Current definitions require hemicelluloses to have a β-1,4-linked backbone with an equatorial configuration and an ability to hydrogen-bond to cellulose (Scheller & Ulvskov, 2010; Albersheim *et al.*, 2011). Xylan is usually regarded as the major component hemicellulose, but it can also include arabinoxylan, xyloglucan, mannan, glucomannan, galactomannan, and mixed-linkage glucan. Arabinan, galactan, and arabinogalactan do not fit this classification and can be grouped with pectins (Scheller & Ulvskov, 2010). Consistent with the variability of the substrate, a number of enzyme classes are involved in hemicellulose degradation (for an overview, see van den Brink & de Vries (2011)).

Among plant cell wall polysaccharides, pectin has the most structural and functional complexity (for review, see Caffall & Mohnen, 2009; Harholt *et al.*, 2010; Mohnen, 2008; Vincken *et al.*, 2003; Willats *et al.*, 2001). Four structural domains are most commonly differentiated: homogalacturonan (HG) and rhamnogalacturonan I (RG-I) as well as the substituted galacturonans xylogalacturonan (XG) and rhamnogalacturonan II (RG-II). HG is the major pectic backbone polymer, typically accounting for more than 60% of the pectin in the plant cell walls (Caffall & Mohnen, 2009). It is made of α-1,4-linked D-galacturonic acid (D-GalA) residues, which can be methyl esterified at the *C*-6 carboxyl and acetylated at the *O*-2 and/or *O*-3 positions (Fig. 1). RG-I, the second most abundant pectic polysaccharide in most plant cell walls, has a unique backbone consisting of repeating units of $[\rightarrow\alpha\text{-D-GalA-1,2-}\alpha\text{-L-Rha-1,4}\rightarrow]_n$. About half of the L-rhamnose (L-Rha) subunits in RG-I are substituted at the *C*-4 position with arabinan, galactan, or arabinogalactan side-chains. XG is a HG substituted at *O*-3 with one or two β-linked D-xylose (D-Xyl)-residues. RG-II has a backbone of about eight α-1,4-linked D-GalA units substituted with four very distinct side branches composed of 12 different types of sugars connected by over 20 discrete linkages, which are strongly conserved across vascular plants (O'Neill *et al.*, 2004). Due to the high complexity of the pectic heteropolysaccharides, a number of enzymes with a range of activities are necessary for its efficient degradation (Jayani *et al.*, 2005; Martens-Uzunova & Schaap, 2009). Most generally, backbone-digesting enzymes can be differentiated from enzymes active on the side-chain sugars. The former group consists of hydrolases, such as polygalacturonases and rhamnogalacturonases (both GH28), polysaccharide lyases (pectin and pectate lyases (PL1, 2 and 3) and rhamnogalacturonan lyases (PL4)) and esterases (pectin methylesterases (CE8), pectin acetylesterases (CE12 and CE13) and rhamnogalacturonan acetylesterases (CE12)). The majority of the side-chain active enzymes that work on RG-I substitutions include arabinanases (GH43, GH93), arabinosidases/α-L-arabinofuranosidases (GH3, GH43, GH51, GH54 and GH62), galactanases (GH5, GH16, GH30, GH35 and GH53), α- and β-galactosidases (GH1, GH2, GH4, GH27, GH35, GH36, GH42, GH43, GH57, GH97 and GH110), β-glucuronidases (GH1, GH2 and GH79) and feruloyl esterases (CE1) (Fig. 1) (Lara-Marquez *et al.*, 2011).

Saprobic filamentous fungi that degrade plant biomass contribute extensively to global carbon recycling, and are also the main source of commercial hydrolases with biotechnological applications, including the production of lignocellulosic biofuels (see e.g. Kubicek, 2013; Lara-Marquez *et al.*, 2011). Although a number of studies have been published describing the fungal responses to a variety of carbohydrate substrates – from monosaccharides to complex (often pretreated) biomass (e.g. Adav *et al.*, 2012; Berka *et al.*,

2011; Braaksma *et al.*, 2010; Couturier *et al.*, 2012; de Vries *et al.*, 2002; Delmas *et al.*, 2012; Hakkinen *et al.*, 2012; Martens-Uzunova & Schaap, 2009; Martinez *et al.*, 2009; Navarrete *et al.*, 2012; Paper *et al.*, 2007; Shah *et al.*, 2009; Sun *et al.*, 2012; Tian *et al.*, 2009; Tsang *et al.*, 2009; Wymelenberg *et al.*, 2010), our knowledge of fungal strategies associated with deconstruction of intact plant cell walls is still incomplete. While these studies provide important information about the hydrolytic capabilities of filamentous fungi, the signals derived from the individual carbohydrates are most likely perceived in the context of the complex polysaccharide network of intact plant cell walls. For further understanding of the fungal response to biomass, it is therefore important to consider the internal structure of the cell wall and to perform a systematic, bottom-up analysis involving all the building blocks, such as cellulose, hemicellulose and pectin. Although the abundance of pectin is low in some plant lineages (in particular grasses of the order Poales), its unique localization and function in the middle lamella connecting the cells of all plants makes it a critical cell wall constituent (Caffall & Mohnen, 2009; Harholt *et al.*, 2010; Mohnen, 2008). Since saprophytic fungi have evolved closely with their hosts, we reasoned that pectin may also play a central role in the orientation of fungi within their biomass environment.

To perform a systematic analysis of fungal sensing of individual plant cell wall components, we chose to use the cellulolytic fungus *Neurospora crassa*. In nature, *N. crassa* degrades plant biomass killed by fire and is adapted to burned habitats (Dunlap *et al.*, 2007; Perkins & Turner, 1988; Perkins *et al.*, 1976; Turner *et al.*, 2001). In the laboratory, *N. crassa* has been developed to study a variety of genetic, molecular, biochemical and genomic phenomena since the 1920s (Borkovich *et al.*, 2004; Davis, 2000; Davis & Perkins, 2002; Dunlap *et al.*, 2007), and recently has been used as a model system for polysaccharide degradation by filamentous fungi (Sun *et al.*, 2012; Tian *et al.*, 2009). Our study presented here had three main goals. The first was to elucidate the proteomic and transcriptional response of *N. crassa* to pectin, thereby identifying the “toolbox” that *N. crassa* uses to degrade pectin and pectin-rich substrates. We subsequently integrated these data with previously obtained proteomic and transcriptomic data during cellulose and hemicellulose degradation (Coradetti *et al.*, 2012; Sun *et al.*, 2012; Tian *et al.*, 2009), an approach that allowed us to delineate common traits from unique cellular adaptations associated with utilization of a specific polysaccharide. Finally, by making use of these refined data sets, we aimed to identify and test the functionality of heretofore unknown factors involved in polysaccharide degradation.

Results

The size of the pectic secretome correlates to the complexity of the substrate

To assess secretome differences when *N. crassa* is degrading pectin, cellulose or hemicellulose, we compared the secreted proteins of wild type (WT) cultures on all these substrates by SDS-PAGE (Fig. 2A). Orange peel powder (OPP) (Fig. S1; Rivas *et al.*, 2008) was used as a representative, albeit pectin-rich, complex carbon source. The pectin and OPP secretomes exhibited a banding pattern of proteins quite unlike those secreted by *N. crassa* when degrading crystalline cellulose (Avicel) or xylan. The OPP secretome was similar to the pectin one, but included additional protein bands, presumably due to the presence of other polysaccharides (Fig. 2A).

To survey the *N. crassa* pectin secretome in more detail, WT culture supernatants were analyzed by mass spectrometry (LC-MS/MS; see Materials and Methods) after 4 to 5 days of growth on pectin or OPP, which allowed a maximum yield of secreted protein before substrate depletion. Sample preparation was performed in two ways: either by in-solution digests after fractionation by ion-exchange chromatography (Supporting data set 1) or by in-gel digestion after separation on SDS-PAGE and excision of the major bands (Fig. S2). The two approaches were complementary, with the in-solution digest resulting in the

identification of more proteins, while the band excision method allowed the assignment of proteins to the major bands visible on the gel.

The LC-MS/MS data showed that growth on pectin leads to the most complex secreted protein complement of all three major plant cell wall polysaccharides described so far in *N. crassa* (Fig. 2C). A total of 80 proteins were identified on pectin (combined from in-solution digest (74) and band excision (35)) and 90 from OPP (Supporting data set 1), which included carbohydrate-active enzymes from 28 and 32 different CAZy families, respectively. Not surprisingly, when ordered by their functional categories (Fig. S3) (Ruepp *et al.*, 2004), the pectin and OPP secretomes were dominated by the category "C-compound and carbohydrate degradation" (P-values of 1.59×10^{-13} and 3.54×10^{-27}, respectively). The pectin and OPP secretomes showed a significant overlap (e.g. 78% of the pectic secretome were also found on OPP; Fig. 2B), including six dedicated pectinases: the GH28 polygalacturonases GH28-1 (NCU02369) and GH28-2 (NCU06961), the pectate lyases PLY-1 (NCU06326) and PLY-2 (NCU08176), the rhamnogalacturonan lyase ASD-1 (NCU05598), and the rhamnogalacturonan acetylesterase CE12-1 (NCU09976). The pectin methylesterase CE8-1 (NCU10045) was detected on pectin, but not on OPP. In addition to the backbone-acting pectinases described above, a number of enzymes were present with the potential to degrade the side-chains of RG-I: a probable endoarabinanase (GH43-7/NCU05965), several α-L-arabinofuranosidases (GH51-1/NCU02343, GH43-4/NCU09170, GH54-1/NCU09775), two β-galactosidases (GH35-1/NCU00642 and GH35-2/NCU04623), a putative endo-β-1,6-galactanase (GH5-5/NCU05882) and a β-glucuronidase (GH79-1/NCU00937; An *et al.*, 1994; Renard *et al.*, 1999)

The difference between the pectin and OPP secretomes could be attributed to the presence of xylan and cellulose in the OPP, with the identification of five endo- and exoglucanases (CBH-1/NCU07340, GH5-1/NCU00762, GH6-2/NCU09680, GH7-1/NCU05057, and GH7-2/NCU04854), two PMOs (GH61-13/PMO-3/NCU07898 and GH61-5/NCU08760), two acetylxylan esterases (CE1-1/NCU04870 and CE5-3/NCU09664), three endo-xylanases (GH10-1/NCU05924, GH11-1/NCU02855, and GH11-2/NCU07225) and a xyloglucanase (GH74-1/NCU05955) being specific for OPP. In addition to predicted carbohydrate active enzymes, a number of proteases/peptidases (10 total) were identified in both the pectin and OPP secretomes (enrichment of "protein/peptide degradation" category; P=0.0239 and 0.0493, respectively). These data suggested the presence of proteoglycans in commercial pectin preparations. Indeed, a protein concentration determination assay (Bradford) revealed the presence of 10 mg protein per g (or 1%) of commercially supplied pectin (data not shown). Within the pectin and OPP secretomes, 25 proteins annotated as "hypothetical protein" were detected (http://www.broadinstitute.org/; 2 specific for pectin, 7 specific for OPP, and 16 for both substrates) (Supporting data set 1).

A comparison between the pectin, xylan and cellulose secretomes (Sun *et al.*, 2012; Tian *et al.*, 2009), showed significant differences (Fig. 2C,D; Supporting data set 1): 55 proteins were uniquely identified on pectin, including all but one pectinase (GH28-2 was also identified in the xylan secretome), one probable endo-arabinanase (GH43-7), an α-L-arabinofuranosidase (GH54-1), a β-galactosidase (GH35-2), a endo-β-1,6-galactanase (GH5-5), and the β-glucuronidase GH79-1. Overall, 22 proteins identified in the pectin secretome overlapped with the xylan secretome, versus only 12 proteins that overlapped with the Avicel secretome. Of the nine proteins in common between all three secretomes, six are associated with the *N. crassa* cell wall and/or implicated in cell wall remodeling: NEG-1/NCU04395, NCW-1/NCU05137, Mwg1/NCU05974, ACW-12/NCU08171, ACW-1/NCU08936, and EglC/NCU09175 (Maddi *et al.*, 2009; Oyama *et al.*, 2006). In addition, one endo-xylanase (GH10-2/NCU08189) and two proteins of unknown function were detected (NCU00798 and NCU09024).

The activity of endo- and exo-acting polygalacturonases is required for robust growth on pectin

An advantage of *N. crassa* as a model system is that is has very little redundancy within its genome (Borkovich *et al.*, 2004; Galagan *et al.*, 2003; Galagan & Selker, 2004). This is also true for the pectinase complement, which generally has only one annotated gene for each enzymatic function. The only exception is the presence of two GH28 polygalacturonase genes. Since a comprehensive single gene-deletion strain collection is available for *Neurospora* (Dunlap *et al.*, 2007), deletion strains for all eight annotated and dedicated pectinases were analyzed for their growth phenotype on pectin as a sole carbon-source (C-source) (Fig. 3). Three deletion strains (Δ*gh28-1*, Δ*gh28-2* and the pectin methylesterase Δ*ce8-1*) showed a strong growth phenotype, accumulating only 20% - 40% of WT biomass when grown on pectin (Fig. 3A,C). They also secreted less protein, and had a reduced rate of pectin consumption (Fig. 3A,B). These data indicate that growth on pectin relies primarily on the activity of the polygalacturonases to liberate D-GalA, while pectate lyases (many of which are most active at alkaline pH) (Tucker & Seymour, 2002) were not able to complement the loss-of-function of the polygalacturonase mutants under these conditions. Residual growth is probably due to utilization of the pectin side-chain sugars (such as L-Ara and D-Gal).

Transcriptional analyses of N. crassa exposed to pectin and OPP shows increased expression of genes encoding pectinolytic enzymes

As a complement to the characterization of the pectin/OPP secretome, we assessed the transcriptional response of *N. crassa* when exposed to these substrates. RNA-Seq data were obtained from 16 h sucrose-grown WT cultures transferred for 4 h to pectin or OPP. These data were compared via a pairwise analysis to transcriptional profiles previously obtained from 16 h sucrose-grown cultures transferred for 4 h to media with no carbon source (NoC) (Coradetti *et al.*, 2012). From these analyses, we identified 548 and 692 genes that showed statistically significant increased expression levels and 453 and 617 genes that showed statistically significant decreased expression levels on pectin or OPP, respectively (Fig. S4; Supporting data set 2; see Materials and Methods). Similar to the secretome data, the overlap of the responses to pectin and OPP was extensive with 469 genes showing increased expression (>60% total overlap between the two data sets; Fig. 4A), and 359 genes being repressed under both conditions. The Pearson's correlation coefficient of transcript expression between pectin and OPP was 0.94, indicating that when *N. crassa* is transferred to OPP it primarily reacts to the pectin in the plant cell walls (Fig. 4B).

All predicted pectinase genes as well as a number of genes encoding enzymes with the potential to degrade pectin side-chains were among the most strongly upregulated (Fig. 4C; Table 1; Supporting data set 2). The exo-polygalacturonase *gh28-2* was the most highly expressed pectinase, followed by the pectin methylesterase *ce8-1* and both pectate lyase genes *ply-1* and *ply-2*. Interestingly, the remaining pectinase genes were expressed at an order of magnitude less, including the endo-polygalacturonase *gh28-1*, which despite its much lower transcript level, is needed for efficient polygalacturonase breakdown (see above). Of the side chain-active enzymes, the three most highly expressed genes encode enzymes involved in the degradation of the arabinans (*gh51-1*, *gh43-4* and *gh43-7*), suggesting that the efficient removal of this oligomer is beneficial for accessibility to the backbone sugars.

Comparative transcriptome analysis highlights the specific cellular adaptations in response to different polysaccharides

The plant cell wall is a complex matrix of polysaccharides. To put the pectin transcriptional response into the context of the cell wall, we compared the pectin transcriptome to that of

cellulose and xylan (Coradetti *et al.*, 2012; Sun *et al.*, 2012). For this analysis, the data sets were normalized by identifying genes with increased expression on a given C-source *vs*. both NoC (starvation) and sucrose (carbon catabolite repressed) (Fig. S5A; Supporting data set 2). We performed RNA-Seq from a 16 h culture transferred for 4 h to xylan and calculated the correlation coefficient to a previously published microarray data set (Sun *et al.*, 2012) (Fig. S5B). A correlation factor of 0.768 was sufficiently robust for use of the RNA-Seq experiment as a valid replicate of the published microarray data. The identified cellulose, xylan, and pectin regulons contained similar numbers of genes: 189 for pectin (reduced from 548 when compared only to NoC (see above)), 117 for xylan and 212 for cellulose (Coradetti *et al.*, 2012) (Fig. S5A; Supporting data set 2).

The regulons as defined above contain genes that are significantly induced upon exposure to a particular C-source compared to the relevant controls (no carbon and sucrose). However, they do not provide information about the specificity of the response regarding other C-sources/inducers being present in the cell wall. Therefore, Venn diagrams were generated to visualize the overlap as well as the unique features of each response (Fig. 5A; Fig. S5C; Supporting data set 3). Unique gene sets included 115 genes for pectin, 120 for cellulose, but only 20 for xylan. Additionally, 21 genes were found to be in common between the pectin and cellulose responses, 25 genes between pectin and xylan, 43 between xylan and cellulose and 29 for all three responses.

The genes found in each intersection were manually grouped according to their annotated functions (Fig. 5B) and subjected to a functional category analysis (Fig. S6) (Ruepp *et al.*, 2004). With the exception of the group of genes that overlapped between pectin and xylan, a significant fraction of genes in each comparison included those coding for “hypothetical proteins”. In addition, “metabolism”-associated functions were prevalent in each group (predominantly carbohydrate-related functional categories 01.05 or 01.05.02; Fig. S6), such as CAZymes (in blue, red and green), as well as genes involved in sugar metabolism (purple) and metabolite transport (light blue; Fig. 5B). Genes coding for enzymes involved in fatty acid and protein metabolism (turquoise and orange, respectively), were also prevalent in several of the data sets.

In addition to these common observations, several specific features were identified. The gene set exclusively expressed on pectin contained a number of proteases, peptidases, and other factors involved in amino acid-metabolism (FunCat group 14.13 with a P-value of 5.04×10^{-7}; orange coloration in Fig. 5B). The expression of two genes encoding enzymes involved in phosphorus metabolism (*pho-2*, an alkaline phosphatase and *pho-5*, a high affinity phosphate permease) were specific for the xylan response. Similarly, the specific response for cellulose included at least five coordinately induced genes encoding predicted components of the Sec61 translocon (Zimmermann *et al.*, 2011): SEC61 subunits alpha, beta and gamma (NCU08897, NCU08379 and NCU04127), SEC62 (NCU06333), and a SEC63 Brl domain containing protein (NCU00169). These observations are indicative of a need to adapt the secretory machinery to increased production of newly synthesized hydrolases under cellulolytic conditions, and suggests that the protein flux upon growth on cellulose is substantially higher than that on hemicellulose or pectin. The additional presence of factors involved in ER protein folding in the cellulose-specific gene pool, such as the Kar2p/BiP homolog *grp78* (NCU03982), the protein disulfide isomerase *pdiA* and *prpA* homologs NCU09223 and NCU00813, respectively, as well as the DnaK chaperone encoding gene NCU09485 and the calnexin/calreticulin NCU09265 is a further indication of an ER-associated stress response (Guillemette *et al.*, 2007).

Genes coding for hydrolytic enzymes and sugar metabolite transporters display differential expression patterns

Our comparative analysis showed that only three pectinase genes were exclusively upregulated on pectin (*gh28-1*, *gh105-1*/NCU02654, and *ply-1*). In contrast, *gh28-2* was also induced on xylan, *asd-1* and *ce12-1* also on Avicel and *ply-2* as well as *ce8-1* on all three C-sources. Regarding cellulase genes, of the eight present in this analysis (from GH families 5, 6, 7, 45), five were specifically induced on Avicel (*cbh-1*, *gh6-2*, *gh7-2*, *gh7-4*/NCU05104, and *gh45-1*/NCU05121), while three were induced on all substrates (*gh5-1*, *gh6-3*/NCU07190, and *gh7-1*). Moreover, five PMOs were uniquely induced on cellulose (*gh61-2*/NCU07760, *gh61-4*/*pmo-2*/NCU01050, *gh61-6*/NCU03328, *gh61-12*/NCU02344, *gh61-13*/*pmo-3*), three PMO genes were induced on cellulose and xylan (*gh61-1*/NCU02240, *gh61-5*, and *gh61-7*/NCU00836), and one was induced on cellulose and pectin (*gh61-3*/NCU02916). Interestingly, none of the hemicellulase genes was exclusively induced only on xylan. The largest fraction (five out of 15) was induced on all three substrates (*gh10-1*, *gh10-2*, *gh11-1*, *gh51-1*, *gh54-1*) while three were induced by xylan and Avicel (*gh11-2*, *gh43-2*/NCU01900, *gh43-5*/NCU09652), three by pectin and Avicel (*gh10-3*/NCU04997, *gh43-6*/NCU07326, *gh53-1*/NCU00972), one by pectin and xylan (*gh43-7*) and two on pectin only (*gh43-4* and *gh67-1*/NCU07351).

A total of 160 genes with a Pfam description (Punta *et al.*, 2012) of “Major Facilitator Superfamily” (MFS), “MFS transporter”, or “Sugar (and other) transporter” are found in the *N. crassa* genome. As evident from the light blue bar in Fig. 5B and the small P-value of FunCat group 20.01.03 (C-compound and carbohydrate transport; P-value = 5.17×10^{-9}), metabolite transporters were enriched in the relatively small 25-gene overlap between the pectin and xylan transcriptome (group “Xy ∩ Pe”; NCU00809, NCU01132, NCU02188, NCU04537, NCU04963, NCU05627, and *hgt-1*). These data are consistent with the fact that the largest variety of sugars is found both in the pectic and hemicellulosic fractions of the plant cell wall.

Identification of a group of genes induced in presence of all three polysaccharides

Several observations make the group of 29 genes significantly induced under all three carbon sources unique (Fig. 5; Table 2; Supporting data set 3): more than 75% are carbohydrate-active enzymes (CAZymes), including a number of polysaccharide esterases (two acetylxylan esterases (*ce1-1* and *ce5-3)*, the feruloyl esterase *fea-1* and the pectin methylesterase *ce8-1*). Additionally, enzymes that cleave oligosaccharides into monosaccharides were also present in this group, including a β-galactosidase (*gh53-2*), two β-xylosidases (*gh3-8* and *gh3-7*), one β-mannosidase (*gh2-1*), and two α-L-arabinofuranosidases (*gh51-1* and *gh54-1*). Intriguingly, β-glucosidase genes were absent from this group, but instead two endo- and one exo-glucanase genes (*gh5-1*, *gh6-3* and *gh7-1*) were present. Considering that cellulolytic gene induction is thought to proceed via the perception of short cellodextrins (Znameroski *et al.*, 2012) as opposed to hemicellulose and pectin signaling, which largely act via monosaccharides (e.g. de Vries *et al.*, 2002; Foreman *et al.*, 2003), these data suggest that the cell uses this pool of enzymes under any carbon-induced condition to scout for (and produce) inducing molecules from an array of possible substrates usually present in complex plant biomass. In support of this hypothesis, 24 genes from this gene set (>82%) were also found to be part of the OPP regulon (Supporting data set 3), a C-source that contains all three polysaccharides.

Hierarchical clustering of whole-genome transcriptomes distinguishes clusters of genes co-regulated in response to specific plant cell wall-derived C-sources

Comparative analyses using Venn diagrams do not consider absolute expression values and therefore do not make use of the most powerful feature of RNA-Seq data. As a complementary analysis, we performed hierarchical clustering of whole-genome transcriptional responses to pectin, xylan and Avicel as well as the complex carbon source OPP, with the relevant control conditions NoC and sucrose. We arbitrarily set the threshold so that 17 clusters were differentiated, such that the transcriptional responses to the different C-sources were well separated (Fig. 6; Supporting data set 4). The most cellulose-specific genes are found in cluster 3 (425 genes), the most xylan-specific in cluster 4 (355 genes), and the most pectin-specific in cluster 10 (555 genes). Genes in cluster 1 seem to be predominantly involved in the starvation response (NoC), and clustered most closely with the Avicel-specific response, while the xylan response clustered closest to sucrose, reflective of the easy digestibility of this carbon source relative to Avicel.

Several genes of metabolic pathways clustered closely together. For example, genes involved in DXyl catabolism clustered next to each other in the xylan-specific cluster 4: xylose reductase (*xyr-1*/NCU08384), xylitol dehydrogenase (NCU00891) and D-xylulose kinase (*xyk-1*/NCU11353) were only interrupted by a MFS transporter gene (NCU06384), with L-xylulose reductase (NCU09041) being located slightly more distant. Intriguingly, L-xylulose reductase is the only one of these enzymes not exclusively involved in D-Xyl metabolism, and represents a link to the catabolic pathway of L-arabinose (L-Ara) (Seiboth & Metz, 2011). The L-Ara catabolic genes themselves are present in the xylan/pectin overlap cluster 6, which includes genes involved in both pectin- (D-GalA) and hemicellulose (L-Ara) metabolism, such as the α-L-arabinofuranosidase *gh51-1*, a probable endo-arabinanase (*gh43-7*), the arabinogalactan endo-1,4-β-galactosidase *gh53-1* (De Vries *et al.*, 2002), as well as the L-arabinitol 4-dehydrogenase *ard-1*/NCU00643. The first three D-GalA catabolism pathway genes: GalA reductase (NCU09533), L-galactonate dehydratase (NCU07064), and L-threo-3-deoxy-hexulosonate aldolase (NCU09532), also clustered closely together (within five positions).

Genes for the seven secreted pectinases plus the intracellular predicted unsaturated rhamnogalacturonyl hydrolase *gh105-1* were in the top 50 genes of the pectin-specific cluster 10. This cluster could be further subdivided into four sub-clusters (10.1-10.4), with 10.1 harboring the most pectin-specific gene set (181 genes), including all eight pectinase genes and a number of genes encoding side-chain active enzymes, such as two α-L-arabinofuranosidases (*gh54-1* and *gh43-4*), a putative endo-1,6-galactanase (*gh5-5*), an exo-1,3-galactanase (*gh43-3*/NCU06861), three β-galactosidases (*gh2-3*/NCU00810, *gh35-2* and *gh35-1*), a feruloyl esterase (*fae-1*/NCU09491) and a predicted α-glucuronidase (*gh67-1*), that could potentially degrade glucurono(arabino)xylans. Another subcluster, 10.4, included the most OPP-specific genes (192 genes), while the remaining two included genes that were also robustly induced under NoC conditions (Cluster 10.2; 142 genes) or had similar expression levels to those on sucrose (Cluster 10.3; 40 genes).

A similar sub-division was also created for the cellulose-specific cluster 3. While the clusters 3.2 (28 genes), 3.3 (49 genes) and 3.4 (114 genes) were less specific for cellulose and also displayed robust expression on other C-sources, the most strongly and specifically cellulose-induced genes were located in cluster 3.1 (such as *cbh-1*, *gh5-1*, *gh6-2*, *cdh-1*/NCU00206, *gh61-4/pmo-2* and *gh61-13*/*pmo-3;* 234 genes total), which also shows a substantial overlap (52%) with the 212-gene “Avicel regulon” (Coradetti *et al.*, 2012).

The identification of new factors associated with polysaccharide degradation by mutant analyses of genes encoding pectin-specific secreted proteins

Using the hierarchically clustered transcriptional response as well as our secretome data, we analyzed the phenotype of 21 pectin-grown strains carrying gene deletions for proteins that were secreted by *N. crassa* when grown on pectin and OPP, and that were also part of the most pectin-specific gene expression clusters (clusters 6 and 10; Fig. S7). Four of the 21 deletion strains showed reduced growth (and/or protein secretion/pectin consumption) including a β-galactosidase mutant (Δ*gh35-1*), a mutant of a probable endo-arabinanase (Δ*gh43-7*) and mutants in two genes encoding "hypothetical proteins": Δ*NCU07923* and Δ*NCU05788* (Fig. 7). Strains containing deletions of genes coding for three proteases/peptidases (SPR-7/NCU07159, NCU00263 and NCU06720), the β-glucuronidase GH79-1, the "hypothetical proteins" NCU08432 and (particularly) NCU09525, as well as the circadian rhythm controlled protein CCG-14/NCU07787 grew significantly better on pectin.

Identification of transcription factors in the clusters

The coordinate expression of genes is often mediated by transcription factors that co-regulate a number of genes involved in a particular metabolic response (e.g. Aro *et al.*, 2005). While these do not necessarily follow the same expression pattern as their target genes, in some cases a positive feedback regulation can place them in clusters related to that of their target genes in a hierarchical clustering approach such as performed here. When we looked for transcription factors with a known impact on cellulase expression, we found this to be the case for the major cellulase regulator *clr-1*/NCU07705, which resides in cluster 3.3 (Coradetti *et al*., 2012). CLR-1 directly induces expression of *clr-2*/NCU08042, explaining its clustering in very close proximity to the other main cellulases in the top part of cluster 3.1. The hemicellulase regulator *xlr-1*/NCU06971 is also found in cluster 3.1, which is in line with its modulating role for full induction of a subset of cellulase genes under cellulolytic conditions (Sun *et al*., 2012).

No major transcription factor mediating the pectin response in filamentous fungi has been described today. Within the pectin-specific cluster 10, we identified a total of 15 (mostly putative) transcription factors: three in cluster 10.1 (*pp-1*/NCU00340, NCU06068, and NCU02142), four in cluster 10.2 (NCU03643, NCU09033, NCU00808, NCU05022), three in cluster 10.3 (NCU03184, NCU07379, and NCU00233), and five in cluster 10.5 (*bek-2*/NCU07139, NCU09315, NCU03421, NCU08594, and NCU00329) (Colot *et al.*, 2006; Leeder *et al.*, 2013; Li *et al.*, 2005). However, preliminary data indicate that none of the available deletion strains for these transcription factors is essential for pectinase induction (data not shown).

The presence of cellulose is associated with processing of the transcription factor hac-1 and increased expression of genes predicted to be involved in the unfolded protein response

Similar to *clr-2* and *xlr-1*, the predicted transcription factor *hac*-1 (NCU01856) involved in the unfolded protein response (UPR) was found to localize to the most cellulose-specific cluster 3.1. Moreover, with the exception of *sec61 alpha* and *gamma* (cluster 12), all factors involved in the ER stress response (and present in the 120 gene cellulose-specific regulon; see above) are located in cluster 3.1, as well as a number of additional genes that likely take part in the adaptation of the secretory machinery to the trafficking of newly translated cellulases, including additional chaperones, translocon components and members of the signal peptidase complex, *N*- and *O*-linked glycosylation factors as well as genes encoding proteins essential for the anterograde ER-to-Golgi vesicle trafficking (Supporting data set 4; Geysens *et al.*, 2009; Guillemette et al., 2007; Kwon *et al.*, 2012; Szilagyi *et al.*, 2013). This observation raised the question whether growth on cellulose triggers the UPR, which

functions to increase the secretory capacity of the ER under stress conditions. The key event for the activation of UPR is the non-spliceosomal, endonucleolytic cleavage of a short, unconventional intron in the mRNA of the transcription factor HAC-1, which leads to a frame shift in translation and the addition of a putative activation domain to the C-terminus of the protein (reviewed by: Geysens et al., 2009; Moore & Hollien, 2012; Saloheimo & Pakula, 2012; Walter & Ron, 2011). We found that *hac-1* mRNA was efficiently spliced under artificial ER-stress conditions (treatment with dithiothreitol; DTT) in a conserved location, although the intron is slightly longer than reported in other filamentous fungi (23 nt instead of 20 nt; Saloheimo *et al.*, 2003) (Fig. 8A, Fig. S8A-C). With this information we tested the (relative) abundance of spliced (ON) *vs*. non-spliced (OFF) *hac-1* transcript in cultures growing on cellulose as compared to control cultures on sucrose or treated with DTT (Fig. 8B). The total *hac-1* transcript was found to be induced about 4-fold with a concomitant shift of the ON:OFF ratio towards more spliced transcript. We concluded that while a clear *hac-1* activation is apparent under cellulolytic conditions, the response is more graded as compared to the more stringent response to DTT. Thus, increased *hac-1* splicing on cellulose is consistent with the observed increased expression of genes associated with the UPR and indicates that *N. crassa* suffers ER stress when required to deconstruct crystalline cellulose.

Identification and characterization of an L-arabinose transporter in the pectin-/hemicellulose shared cluster

The clustering of poorly characterized genes or genes encoding proteins of unknown function with other, well-annotated genes can guide in the development of testable hypotheses (Eisen *et al.*, 1998). As an example, an uncharacterized MFS sugar transporter gene (NCU02188) was found to cluster directly next to the L-arabinitol 4-dehydrogenase gene (*ard-1*) involved in L-Ara catabolism (Fig. S9A; Supporting data set 4) and was designated *lat-1* (L-arabinose transporter-1). To determine whether the transporter is involved in the metabolism of L-Ara, we performed growth assays of sucrose pre-grown WT and Δ*lat-1* cultures on purified arabinan and compared the phenotype to growth on galactan, xylan, and pectin by measuring the mycelial dry weight over the course of four days (Fig. 9A; Fig. S9B-D). A clear growth retardation was only observable on arabinan, corroborating our hypothesis. To further support our hypothesis, we performed sugar uptake assays, in which the pectin-induced WT and Δ*lat-1* deletion strain were incubated in a mixture of the four most relevant pectin side-chain sugars: D-Xyl, D-galactose (D-Gal), L-Rha, and L-Ara. The concentrations of these sugars in the culture supernatants were monitored over time (Fig. 9B). Analysis of initial transport rates (first five minutes) revealed that uptake of L-Rha and D-Xyl was similar to WT in the Δ*lat-1* mutant while L-Ara uptake was almost completely abolished. Since D-Gal uptake was also reduced by about 40%, we determined whether this was an indirect effect resulting from the mis-expression of a potential L-Ara induced D-Gal transporter in the Δ*lat-1* background. To this end, we induced cultures with 1 mM D-Xyl instead of pectin (which contains L-Ara) before measuring sugar uptake. No reduction of D-Gal uptake as compared to the WT was observed under these conditions (Fig. S9E-G), suggesting that LAT-1 is specific for L-Ara.

To determine whether LAT-1 functions as a symporter or a facilitator-type of transporter, WT *N. crassa* was induced with pectin and L-Ara uptake subsequently analyzed in presence or absence of low concentrations (5 μM) of the uncoupler CCCP (Fig. S9H). In presence of CCCP, the uptake of L-Ara was strongly inhibited, suggesting that LAT-1 operates as H^+ symporter. Uptake of D-Glc was much less affected, presumably since D-Glc is partly taken up by facilitator-type transporters, which are insensitive to uncouplers.

Since our uptake assays only followed the disappearance of sugar from the extracellular medium, we confirmed the Δ*lat-1* phenotype by measuring the intracellular L-Ara concentration. Pectin-induced *N. crassa* WT accumulated 1.3 +/− 0.2 mM L-Ara after 20 min (plus some converted to L-arabitol; not quantified), while L-Ara (and L-arabitol) was barely detectable intracellularly in the Δ*lat-1* mutant (Fig. 9C).

We hypothesized that a drastically reduced L-Ara uptake as observed in Δ*lat-1* should also affect downstream signaling cascades that regulate L-Ara catabolism. In filamentous fungi only the L-arabitol dehydrogenase ARD-1 is dedicated to L-Ara metabolism while all other enzymes are shared with D-Xyl metabolism (Seiboth & Metz 2011). Consequently, we found *ard-1* to be specifically induced by L-Ara, while the general pentose reductase encoding gene *xyr-1* responded to both D-Xyl and L-Ara (Fig. 9D,E). A specific response to L-Ara was also found for *lat-1* itself (Fig. 9D), further corroborating its co-expression with *ard-1* as observed in the clustering data (Fig. S9A). To investigate the role of LAT-1 in signaling, we analyzed whether the regulation of *ard-1* is affected in Δ*lat-1*. For greatest specificity, very low concentrations of sugar (2 μM) were chosen for the induction experiment. The L-Ara dependent induction of *ard-1* was specifically abolished in Δ*lat-1* while the D-Xyl dependent induction of *xyr-1* and the D-GalA dependent induction of the main exo-polygalacturonase encoding *gh28-2* were WT-like (Fig. 9E). Based on these results, *lat-1* encodes the physiologically relevant L-Ara transporter in *N. crassa* with a role in both pentose metabolism as well as L-Ara signaling.

Discussion

N. crassa displays a robust and specific response to pectin comparable in scale to the cellulolytic and xylanolytic responses

Our systems analysis provides evidence that *N. crassa* employs a “toolbox” for the specific degradation of pectin, which is by no means insignificant in comparison to its cellulolytic or hemicellulolytic counterparts (e.g. Fig. 2 and 5). This is particularly remarkable when taking into account the generally low amount of pectin in mature plant cell walls. In part, this may be due to the complexity of the pectic heteropolysaccharide, but also suggests that the decomposition of pectin has additional advantages over being a mere C-source. The importance of pectin is most likely the result of its strategic localization in the middle lamella and primary cell walls, where it acts like the “glue” between the individual cells. Targeted digestion of the pectic fraction would therefore allow dissolution (maceration) of individual plant cells and thus a better penetration of plant tissue.

From our comparative secretome and transcriptome analyses of *N. crassa*, we identified the majority of genes and proteins needed for the deconstruction of the HG and RG-I components of pectin in the plant cell wall (Fig. 10). No dedicated enzymes for the degradation of either XG or RG-II were identified. However, the substrate specificities of GH28-1 and GH28-2 have not been extensively studied and therefore may have XG hydrolase activity, which has been reported for members of this glycoside hydrolase family (Kester *et al.*, 1999; Zandleven *et al.*, 2005). The lack of RG-II degrading factors, on the other hand, is not surprising, since no organism has so far been reported to have this ability.

We were furthermore able to refine the predicted functions for a number of secreted enzymes by comparing the transcriptional responses of *N. crass*a to individual plant cell wall components in a systematic fashion. For example, the exo-α-L-1,5-arabinanase (GH93-1/NCU09924) could act on the RG-I side chain arabinan (Benoit *et al.*, 2012), but was most highly expressed on xylan (Cluster 4), suggesting a role in the hydrolysis of hemicellulose-associated arabinose. Conversely, the pectin methylesterase *ce8-1* and the feruloyl esterase *fae-1* display their strongest induction on pectin. However, moderate

induction was also detected on xylan and cellulose (part of the central 29-gene set), suggesting an additional role in generating signaling molecules for induction (see below).

Compared to *A. niger*, which currently represents the "gold standard" for industrial pectinase production, the number of genes encoding pectinolytic enzymes in *N. crassa* is small, reminiscent of what has been observed in the ascomycete plant pathogen *Magnaporthe oryzae* (Benoit *et al*., 2012; Pel *et al.*, 2007). For example, *N. crassa* has two GH28 genes, while *A. niger* has 22 (Martens-Uzunova & Schaap, 2009; van den Brink & de Vries, 2011). However, most enzyme classes necessary for pectin breakdown are present in the *N. crassa* genome, in contrast to other industrial fungi such as *Trichoderma reesei*, which does not encode any lyases or pectin methyl-/acetylesterases (van den Brink & de Vries, 2011). The only enzyme classes lacking in *N. crassa* are an α-rhamnosidase (GH78) (Mutter *et al.*, 1994; Yadav *et al.*, 2010), an unsaturated glucuronyl hydrolase (GH88), and a dedicated pectin acetylesterase (CE12). A more detailed characterization of GH105-1 (a rhamnogalacturonan hydrolase) as well as CE12-1 (a rhamnogalacturonan acetylesterase) may reveal that their respective substrate specificities are broad enough to include the HG parts of the pectic polysaccharide, and thus compensate for the absence of any dedicated enzyme. The limited cohort of genes in *N. crassa* facilitates genetic analyses, since deletion of single genes often translates directly into an observable phenotype. The more elaborate enzyme systems present, for example, in the Aspergilli likely allow for higher efficiency of degradation and represent adaptation to specialized habitats (van den Brink & de Vries, 2011). While these much more complex enzyme systems cannot be modeled in *N. crassa*, we suggest that *N. crassa*'s pectinolytic complement might prove to be a good, genetically tractable, example of a minimal "toolbox" necessary for effective pectin degradation.

The analysis of deletion strains in pectin-regulated genes reveals their impact on the utilization process

Our transcriptome and secretome data analysis identified predicted genes/proteins associated with pectin degradation, as well as a large number of genes/proteins with either a general biochemical function (MFS transporter, for example) or with no annotated function. The near full genome deletion set available for *N. crassa* (Dunlap *et al.*, 2007) provides a powerful tool to assess their roles in plant biomass deconstruction. By analyzing deletion mutants for genes encoding each of the backbone-acting pectinases, we found that the pectin methylesterase *ce8-1*, as well as both GH28 polygalacturonases *gh28-1* and *gh28-2*, were important for efficient pectin utilization. These three mutants display very similar phenotypes indicating a functional interdependence. Interestingly, the two GH28 enzymes were not able to compensate for the loss of the other, suggesting a different and probably synergistic mechanism of cleavage. Although both polygalacturonase mutants showed little pectin consumption after four days, the Δ*gh28-2* cultures had cleared while the Δ*gh28-1* cultures were still visibly opaque and more viscous (data not shown), indicative of an intact pectin matrix. These observations, in conjunction with a phylogenetic comparison to the *Aspergillus niger* endo- and exo-polygalacturonases (Fig. S10) confirm that GH28-2 is an exo-acting polygalacturonase related to PgxB, while GH28-1 is endo-acting, clustering with enzymes such as PgaI, PgaC and PgaII (Benen *et al.*, 1999; Delourdes *et al.*, 1991; Martens-Uzunova *et al.*, 2006). One or both GH28 enzymes further relies on the pectin methylesterase CE8-1 to modify the substrate, since the absence of this non-hydrolytic enzyme also leads to a similar phenotype. To our knowledge, this is the first report of an *in vivo* synergism between these enzymes in saprobic fungi, which has previously been observed only *in vitro* (Benen *et al.*, 1999; Christgau *et al.*, 1996; Dahodwala *et al.*, 1974; Jansen *et al.*, 1945; Massiot *et al.*, 1997; Pressey & Avants, 1982; Wakabayashi *et al.*, 2003). In plant pathogenic fungi, several polygalacturonase or pectin methylesterase enzymes have been demonstrated to be important for growth on pectin and (often) essential

for virulence (Isshiki *et al.*, 2001; Moran-Diez *et al.*, 2009; Schell *et al.*, 1988; Scott-Craig *et al.*, 1998; Valette-Collet *et al.*, 2003). Therefore, efficient pectin degradation is equally vital for the decomposition of dead plant biomass as it is for the colonization of living hosts.

In addition to genes encoding enzymes with predicted function, we also identified genes encoding “hypothetical proteins” that were significantly induced on pectin and identified proteins in the pectin and OPP secretomes that have unknown biochemical function. At least 13 such proteins were specifically secreted on pectin (part of the 55-protein pectin-only group) and seven were also part of the most pectin-specific expression cluster 10. When the respective deletion mutants were grown on pectin, five showed either a defect or increased fitness on pectin, with one (Δ*NCU09525*) showing markedly increased fitness and one (Δ*NCU07923*) showing markedly decreased fitness. Both the phenotypic data and the expression profiling data strongly suggest that these proteins play a role in pectin degradation. Thus, our data has provided information on these proteins with no available biochemical information that can guide further research to address their specific function in biomass degradation.

Identification of an MFS-type transporter for L-arabinose through polysaccharide-dependent co-expression clustering

As a proof-of-concept of whether the resolution of the hierarchical clustering matrix was sufficient for a directed experimental approach aimed at the elucidation of the specific biochemical function of proteins, we analyzed the phenotype of the deletion strain for *lat-1* (NCU02188). *lat-1* is annotated merely as an MFS-type “sugar transporter”, but our data demonstrate that it indeed encodes the physiologically relevant L-Ara transporter in *N. crassa*, showing the applicability of this approach to assess new protein functions associated with plant biomass deconstruction.

The closest relatives of LAT-1 that have been biochemically characterized are from ascomycete yeasts: LAT1 from *Ambrosiozyma monospora* and AraT from *Pichia stipitis* (46% and 39% amino acid identity with LAT-1, respectively) (Subtil & Boles, 2011; Verho *et al.*, 2011). These recently identified transport proteins also facilitate the uptake of L-Ara, and were expressed in *Saccharomyces cerevisiae* to improve its pentose fermentation capability. LAT-1 is conserved among ascomycete fungi (BLASTp (Altschul *et al.*, 1990); Fig. S11), but more distant homologs within basidiomycete species and even in bacteria are present (data not shown). A paralog of *lat-1* is also present in the *N. crassa* genome: NCU08152 (39% amino acid identity). It is conceivable that NCU08152 can also transport L-Ara, however, in *N. crassa* its expression level is extremely low and therefore most likely not of physiological relevance. Interestingly, the LAT-1-related clade in filamentous ascomycete fungi is smaller than the NCU08152 clade due to the absence of close homologs in some species (Fig. S11), suggesting that these rely more on the NCU08152-like transporters for L-Ara transport.

Complexities in the plant cell wall structure may explain a differential expression behavior of hydrolytic genes

Our data demonstrate that the regulation of the pectinolytic machinery is a highly coordinated process that is largely independent of the cellulose “regulon” as well as sufficiently different from the xylan response to generate a distinct expression cluster (Cluster 10). However, we also observed that the pectin response significantly overlaps with the hemicellulosic one, both on the level of the transcriptome (Cluster 6) as well as the secretome (Fig. 2D). Pectin and hemicellulose polysaccharides have a number of sugar moieties in common (L-Ara, D-Gal, D-Xyl and D-GlcA). Some of the expression complexity therefore likely results from the fact that several of these monosaccharides can

have signaling function (as shown e.g. in *Aspergillus*; Martens-Uzunova & Schaap, 2009)). Moreover, it is feasible that these (and other) signals can act either individually or in an additive fashion to affect gene expression. Cluster 6 is an example for several independent regulatory circuits (such as D-GalA and L-Ara or D-Xyl) likely acting in parallel. Pectinolytic genes under the control of such multiple regulatory systems have been reported in *A. niger* (de Vries *et al.*, 2002). The OPP-specific cluster 10.4, on the other hand, displays the characteristics of a group of genes that are most strongly induced on a complex substrate (similar to observations by Hakkinen *et al.*, 2012), probably due to the additive effects of several induction pathways. Analogously, also in plant pathogenic fungi certain proteins can only be detected when cultivated *in planta* (being exposed to all native inducers at the same time) and cannot be found in the secretome under more controlled *in vitro* conditions (Paper *et al*., 2007).

An alternative (non mutually-exclusive) explanation involves the need for a spatiotemporal regulation of the plant cell wall matrix decomposition. The outermost layers of the wall, the middle lamella as well as the primary cell wall, must be degraded first and are mostly made of pectin and hemicellulose. A parallel induction of gene-sets required for their utilization might therefore be beneficial. Cellulose, on the other hand, is predominantly found in the secondary cell wall, and thus physically already well separated from pectin-containing fractions. The cellulose microfibrils are, however, embedded in a hemicellulose matrix (Somerville *et al*., 2004). The fact that genes encoding hemicellulases were expressed with either pectinases or cellulases in *N. crassa* is therefore consistent with the overall organization of polysaccharides within intact plant cell walls.

In addition to spatial constraints, many hydrolases clearly display differential expression over time, and the complete hydrolysis of plant biomass likely requires a cascade-like action of different enzymes (Hakkinen *et al*., 2012). Our approach to focus on the four hour time point after carbon source switch allows a very comparable view at gene induction close to its early maximum for the majority of genes, but it is limited by the fact that late-responsive events will not be detected (compare de Souza *et al.*, 2011; Hakkinen *et al.*, 2012; Martens-Uzunova & Schaap, 2009). Ultimately, a combination of abundance and accessibility will govern gene induction. Pectin, for example, is a highly accessible matrix polysaccharide as opposed to cellulose, which (although usually abundantly present) is covered by hemicelluloses (see above). This discrepancy becomes obvious also in our data sets, in which we observed the pectic response to be dominant when *N.crassa* was growing on OPP, while the main cellulases (such as *cbh-1* and others present in the top of cluster 3.1) do not display any induction at all after four hours. Later time points would probably reveal an induction of these genes on OPP (containing > 15% cellulose) as well, and indeed, CBH-1 is detected in the OPP secretome after 4-5 days (Fig. 2D; Supporting data set 1). Similarly, even within a polysaccharide, main chain degrading enzymes might only be induced after the blocking side chains have been taken off by other enzymes. An example of this can be seen in Martens-Uzunova & Schaap (2009) where rhamnogalacturonan lyases, which are rapidly induced by D-Rha, show a clear delay in induction on sugar beet pectin (Fig. 2C in Martens-Uzunova & Schaap, 2009).

The identification of a subset of genes induced by all carbon-sources may allow N. crassa to “taste” the environment and adjust its metabolism to its “flavors”.

In the presence of a preferred carbon source, the expression of lignocellulolytic genes in filamentous fungi is subject to carbon catabolite repression (CCR) (Delmas *et al*., 2012; Ebbole, 1998; Flipphi & Felenbok, 2004; Ruijter & Visser, 1997; Strauss *et al.*, 1995; Sun & Glass, 2011). Under starvation conditions (“no carbon” in this study), CCR is relieved and many lignocellulolytic genes (>100 CAZymes in *N. crassa*) display an elevated level of

expression (Coradetti *et al.*, 2012; Delmas *et al.*, 2012; Nitsche *et al.*, 2012; Tian *et al.*, 2009). In this situation, small amounts of enzyme are secreted into the environment to "scout" for new C-sources. If found, hydrolysis will lead to the production of metabolites that elicit a strong induction of genes encoding enzymes involved in degradation and metabolism of the respective C-source. Our data enrich this model in two ways: First, by identifying C-source-specific cellular adaptations, such as the activation of UPR on cellulose as well as an apparent need for higher phosphate metabolism in presence of xylan. And second, by the discovery of a (relatively small) set of 29 genes that are not only strongly induced in presence of a suitable C-source (for example, pectin), but that also display a significant, albeit lower, induction on non-dedicated substrates (for example, both hemicellulose and cellulose) or complex substrate mixtures (such as OPP). The gene set contains a number of genes coding for esterases, endo- and exo-acting hydrolases, as well as enzymatic activities for the release of monosaccharides, with the notable exception of β-glucosidases. We hypothesize that the composition of enzymes produced by this set of genes would be perfectly suited to produce signaling molecules from a variety of possible C-sources, allowing fine-tuning of cellular metabolism and the hydrolytic enzyme repertoire. Most of the genes in this set also display a small to moderate upregulation on "no carbon" *vs*. sucrose, and might therefore be already involved in the starvation scouting (Delmas *et al*., 2012). However, a further probing of the immediate environment in presence of C-source (or "tasting") could conceivably facilitate the efficient degradation of the plant cell wall in the context of its three-dimensional ultra-structure.

Outlook

In the future, the incorporation of additional conditions, such as a variety of complex C-sources with varying polysaccharide compositions, as well as exposure of *N. crassa* to monosaccharides and monomeric compounds (such as L-Ara, L-Rha, ferulic acid, lignin) will increase the resolution of the clustered comparative analyses and provide concrete functional hypotheses that can be tested. It is an intriguing thought that an extended study of the occurrence and activity of the hydrolyzing enzymes in plant cell wall-deconstructing fungi might also provide insight into polysaccharide structures and linkages present in plant cell walls. Moreover, understanding the expression, induction and regulation characteristics of the corresponding genes in detail might allow to deduce the proximities of certain polymers in the plant cell wall, which are difficult to visualize by most other, albeit more direct, experimental approaches.

Experimental Procedures

Strains, media, and growth conditions

N. crassa WT (FGSC #2489) and gene deletion strains used in this study were obtained from the Fungal Genetics Stock Center (FGSC; www.fgsc.net). Due to a former mis-annotation, Δ10966 was found to actually be a deletion of two neighboring genes: the pectate lyase *ply-1* (NCU06326) and NCU06327, encoding a cytochrome P450 protein. The Δ*gh28-1* strain (ΔNCU02369; FGSC #16219) was only available as a heterokaryon and was therefore backcrossed to WT to obtain homokaryotic strains and to eliminate the Δ*mus-51*::bar^{+} background (strains were confirmed to be hygromycin-resistant but sensitive to 200μg/ml glufosinate-ammonium). The strain for Δ*ply-2* (ΔNCU08176; FGSC #21118 and FGSC #21119) could not be successfully confirmed by PCR genotyping; new gene deletion strains were therefore constructed. The DNA cassette used to delete NCU08176 was provided by the *Neurospora* functional genomics project (http://www.dartmouth.edu/~neurosporagenome/protocols.html) and the knockout procedure was performed as described in Phillips *et al.* (2011). Strains were deposited to the FGSC (ΔNCU02369 (homokaryon): FGSC #10763/10764; ΔNCU08176: FGSC #10765/10766). *N. crassa* was

grown on 1x Vogel's salts (Vogel, 1956) with either 2% (w/v) sucrose (Fisher S3-12), or 1% pectin (Sigma P9135), 2% xylan (beechwood, Sigma X4252), 2% orange peel powder (OPP; see below), or 2% Avicel (PH 101, Fluka 11365) at 25 °C and 200 rpm with constant light unless stated otherwise.

Preparation of orange peel powder (OPP) and compositional analysis

Squeezed-out oranges (peels and residual pulp) from a local organic food market were de-pulped, cut into small pieces, and soaked in water for overnight to remove water-extractable sugars. The peels were subsequently extracted 3× with 70% ethanol over the course of 2 days. Between extractions the peels were filtered through two layers of cheese-cloth. After the third extraction the peels were filtered and squeezed-out again and transferred to a 40 °C drying oven for several days. The dry peels were finally ground to a fine powder in a household electric coffee grinder (Black & Decker). Final water content of the OPP was 6%.

The compositional analysis was essentially done as described in Haffner *et al*. (2013). OPP was ground in a canister ball mill (Kleco, USA), dried overnight at 45 °C in a vacuum oven and extracted three times each with ethanol and 1:1 chloroform:methanol. Dried and extracted OPP was then analyzed for structural carbohydrates, lignin and ash following essentially NREL standard laboratory protocols (Sluiter *et al.*, 2011; Sluiter *et al.*, 2005; Sluiter *et al.*, 2008). Released monosaccharides and acetate were analyzed by high-performance liquid chromatography (HPLC) as well as by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD). The fraction of crystalline cellulose of total glucan was calculated by subtracting the amount of glucan detected after a 4% sulfuric acid hydrolysis from that detected after 72%/4% hydrolysis. The remainder was considered (non-crystalline) glucan.

The HPLC analysis as well as the determination of ash and Klason lignin content was done exactly as described in Haffner *et al*. (2013). For HPAEC-PAD an ICS-3000 instrument (Thermo Fisher, USA) was used. Samples were injected onto a 3×150 mm CarboPac PA20 column (Thermo Fisher) equipped with a 3x30 mm guard column of the same material and eluted at 30 °C using an isocratic mobile phase of 2 mM KOH at 0.4 mL/min for 23 min.

The amount of starch was determined using the sample protocol (a) of the ‘Total starch assay procedure’ kit (Megazyme, Ireland) (one replicate only).

Growth, protein secretion and pectin consumption assays

The analyses of deletion strains were usually performed in 24 deep-well plates in a volume of 3 mL from directly inoculated conidia (10^6/mL) over the course of four days (Fig. 3 & Fig. 7). Dry weight was determined after an overnight incubation of the mycelial mass in aluminum pans in a 105 °C oven. For characterization of the Δ*lat-1* phenotype on arabinan, WT and Δ*lat-1* cultures were pre-grown for 16 hours on 2% sucrose and then transferred to 0.5% arabinan for another four days. Each day the dry weight of triplicate cultures was determined and the remaining cultures transferred to fresh carbon source.

The amount of secreted protein was measured using the Bio-Rad Protein assay (Bio-Rad). Pectin consumption over time was followed using the phenol-sulfuric acid assay (PSA) (Dubois *et al.*, 1956). For this, pectin-grown *N. crassa* cultures were initially cleared by centrifugation at 20,000 × g for 5 min. 2.5 μL of the supernatant was diluted into 150 μL Milli-Q water and mixed with 150 μL of a 5% phenol solution. Finally, 750 μL of concentrated sulfuric acid was quickly added, and the suspension vortexed for 5-10 seconds. After an incubation of about seven minutes at room temperature (RT), 200 μL of the

reactions were analyzed in a plate-reader spectrophotometer (Paradigm, Beckman-Coulter) at 487 nm.

Sample preparation for proteomics and mass spectrometry

For in-solution digests, *N. crassa* was inoculated from conidia into 1× Vogel's, 1% pectin (or 2% OPP) solution at an OD_{600} of 0.04 (total volume 200 mL), and then grown for 4-5 days as described above. The mycelial biomass was removed by passing the culture over two layers of cheese-cloth; the flow-through was then filtered through a 0.22 μm filtration device (Corning). In the following, the culture supernatant was fractionated by cation exchange chromatography using an ÄKTAexplorer system (GE Healthcare). For this, 1/10 volume of a 10x buffer A (final concentration: 50 mM citric acid/Na_2HPO_4, pH 3.0) was added, and the sample loaded at onto a 5 mL HiTrap-SP column (GE Healthcare). The column was washed with 20 column volumes (CV) of buffer A, and the elution proceeded using a step-gradient of buffer B (buffer A + 1.5 M NaCl): 0-33% B in 10 CV, 33-66% B in 10 CV, and 66-100% B in 10 CV. Fractions of 1.5 mL each were collected. After control of fractionation by SDS-PAGE, ten representative fractions were prepared for mass-spectrometry. To this end, the samples were precipitated with four volumes of ice-cold acetone, and the resulting pellet resuspended in 100 μL of 50 mM Tris, pH 8.5, 5 mM dithiothreitol (DTT), 6 M urea. After incubation at 65 °C for 60 minutes the samples were diluted with 700 μL 25 mM ammonium bicarbonate (AmBic) and 140 μL methanol. 50 μL of 100 μg/mL trypsin (Trypsin Gold, Promega) in 50 mM sodium acetate, pH 5.0 was then added and samples incubated overnight at 37 °C on a rocking platform. Following digestion, the samples were dried in a speed-vac and the residue resuspended in 250 μL MS-grade water. The resuspension and drying step was repeated two more times after which the pellet was resuspended in 100 μL 0.1% TFA/water. This solution was desalted with C18 OMIX microextraction zip tips according to the manufacturer's instructions (Varian, Inc.; now Bond Elut OMIX, Agilent Technologies). Finally, the eluted peptide samples were dried in a speed-vac and resuspended in 50 μL water/acetonitrile (97:3, v/v, 0.1% formic acid, v/v). All reagents used were MS grade.

For in-gel trypsin-digestion, cleared *N. crassa* pectin culture supernatant (as described above) was concentrated 20× in 10 kDa MWCO centrifugal filtration devices (Vivaspin) and applied to SDS-PAGE (12.5% acrylamide). After staining with GelCode Blue (Thermo Fisher Scientific Inc.) the major visible bands were excised and prepared for LC-MS/MS. For this, the bands were cut into small pieces and washed 4× with 25 mM AmBic in 1:1 acetonitrile (ACN)/water by 10 min of vortexing at each step. The pieces were subsequently vacuum centrifuged to complete dryness and submitted to a reducing treatment (10 mM DTT in 25 mM AmBic with 10% ACN for 1 h at 56 °C) followed by alkylation (55 mM iodoacetamide in 25 mM AmBic for 45 min at RT in the dark). Next, the gel pieces were washed 4x by two rounds of 25 mM AmBic followed by 25 mM AmBic in 1:1 ACN/water, and then vacuum centrifuged to complete dryness. Residues were then rehydrated with 12.5 ng/μL trypsin (Trypsin Gold, Promega) in 25 mM AmBic and digestion allowed to proceed at 37 °C for overnight. The peptides were eluted 1x with 2-3 gel-piece volumes of water and 3× with water/ACN/formic acid (45:50:5, v/v) by vortexing for 10 min followed by 5 min of sonication at each step. Finally, the solvent was removed in a lyophilizer and the peptides resuspended in 10 μL of water/acetonitrile (97:3, v/v, 0.1 % formic acid, v/v).

Samples were analyzed using a QTOF mass spectrometer (Agilent Technologies, USA) equipped with a ChipCube interface and a 1200 Series liquid chromatograph (Agilent) with degasser, capillary pump, nano pump and micro WPS thermostatted autosampler. The column used was a ProtID-Chip-43 (Agilent) consisting of an analytical column with 43 mm length and 75 μm inner diameter with Zorbax 300SB-C18 stationary phase of 5 μm particle

size, and of an enrichment column of 4 mm length of the same material. Samples (1 μL in-gel digest, 0.5 μL in-solution digest) were transferred onto the enrichment column by a capillary pump flow of 4 μL/min of water/acetonitrile (97:3, v/v, 0.1 % formic acid, v/v). The flow from the enrichment column was automatically switched to the analytical column after one sample volume had passed through the enrichment column. Compounds were eluted by a gradient of solvent A (water/acetonitrile 98:2, v/v, 0.1% formic acid, v/v) and solvent B (acetonitrile/water 95:5, v/v, 0.1% formic acid, v/v) at a nano pump flow rate of 0.6 μL/min with a program from 3% B to 95% B in 7 min, then in 0.1 min to 3% B, 1.9 min isocratic (for in-gel digest); or 3% B to 50% B in 38 min, then 50% B to 95% B in 5 min, and in 0.1 min to 3% B, 1.9 min isocratic (for in-solution digest). The mass spectrometer was operated in positive ESI mode, fragmentor voltage 175 V, skimmer 65 V, OCT 1 RF Vpp 750 V, 4 L/min drying gas, 300 °C gas temperature with the following auto MS/MS settings: MS scan range 300-2400 m/z (1 spectrum/s), MS/MS scan range 59-3000 m/z (1 spectrum/s), precursor selection: charge state 2 or higher, 5000 counts min absorbance, 4 m/z isolation width, max 3 precursors per cycle, excluded after 1 spectrum for 0.5 min, collision energy formula: 3.7 × (m/z) / 100 + 2.5.

Database mining was performed using the Spectrum Mill MS Proteomics Workbench (Agilent), or the ProteinProspector batch-tag web service (prospector.ucsf.edu/). MS/MS search against a custom *N. crassa* database was limited to a maximum of 2 missed trypsin cleavages, a mass tolerance of +/− 20 ppm for peptide precursors and +/− 50 ppm for product ions (ProteinProspector: precursor mass tolerance: 100 ppm, product ions: 0.1 Da). Carbamidomethylation at Cys was set as fixed modification, while variable modifications included were: Cys carboxymethylation, carbamylated Lys, oxidized Met, pyroglutamic acid (at N-term), Ser/Thr/Tyr phosphorylation, and His acetylation.

Media shift assays for transcriptional studies (RNAseq and qPCR)

To ensure optimal comparability of RNAseq data sets, transfer assays for RNAseq were performed following the procedures described in Coradetti *et al.* (2012) and Znameroski *et al*. (2012). In short, *N. crassa* cultures were pre-grown from 10 day-old conidia for 16 h in 100 mL of 1× Vogel's salts plus 2% (w/v) sucrose using 250 mL shake-flasks. The mycelia were then washed three times in 1x Vogel's salts without added carbon (NoC) and transferred to 1% (w/v) pectin/OPP or 2% xylan for induction. After an additional 4 h, the mycelial mass was harvested over a Whatman glass microfiber filter (GF/F) on a Buchner funnel and subsequently flash-frozen in liquid nitrogen to be stored at −80 °C. RNA was then prepared using the TRIzol reagent (Invitrogen) as described by Tian *et al*. (2009). For transfer of NoC, sucrose, and Avicel cultures see Coradetti *et al*. (2012).

Gene expression by qPCR was performed essentially as described by Znameroski *et al*. (2012) with some modifications. Briefly, cultures were pre-grown from 9-10 day-old conidia for 16 h in 3 mL of 1× Vogel's salts plus 2% (w/v) sucrose in 24 deep-well plates. The mycelia were then washed three times in 1× Vogel's salts without added carbon (NoC) and transferred to 1x Vogel's salts plus the respective new carbon source (as indicated). The mycelial mass was harvested after an additional 4 h by a quick blotting dry on Whatman tissues and subsequent flash-freezing in liquid nitrogen to be stored at −80 °C. Total RNA from frozen samples was isolated using Zirconia/Silica beads (0.5 mm diameter; Biospec) and a Mini-Beadbeater-96 (Biospec) with 1 mL TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The total RNA was further digested with TURBO DNA-free (Ambion) and purified using an RNeasy kit (Qiagen). RNA concentration and integrity was checked by Nanodrop and agarose gel electrophoresis.

RNA-Seq

Library preparation for RNA sequencing (RNA-Seq) was performed essentially as described by Coradetti *et al*. (2012). These were sequenced in SR50 mode on an Illumina HiSeq2000 at the UC Davis Genome Center, and the files analyzed using the Illumina RTA 1.12 software. Mapping of the reads was done against the current version at the time of the *N. crassa* OR74A genome (v10) (Galagan *et al.*, 2003) using Tophat v1.2.0 (http://tophat.cbcb.umd.edu/) (Langmead *et al.*, 2009). Transcript abundance was estimated with Cufflinks v0.9.3 in FPKMs (fragments per kilobase of transcript per million mapped reads) using upper quartile normalization and mapping against reference isoforms from the Broad Institute (http://cufflinks.cbcb.umd.edu/) (Roberts *et al.*, 2011a; Roberts *et al.*, 2011b; Trapnell *et al.*, 2010). Profiling data are available at the Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/; accession no. GSE42692). Genes with a multiple-hypothesis adjusted P-value of <0.05 using Cuffdiff where called as significantly differentially expressed between conditions (see Supporting data set 2). Independent triplicate cultures were harvested and analyzed for *N. crassa* WT on pectin and duplicate cultures for orange peel powder. The RNAseq data were validated by qPCR analysis of the expression values of the major pectinases in independently prepared samples (Fig. S13).

Hierarchical Clustering of RNA-Seq data

Averaged FPKMs of all RNA-Seq library replicates were hierarchically clustered with the 'Hierarchical Clustering Explorer' v3.0 software (http://www.cs.umd.edu/hcil/multi-cluster/hce3.html). Genes with a consistently low expression (<10 FPKMs under all conditions) were not included in the analysis. Data (8664 genes total) were initially normalized row-by-row by standardization (i.e. the deviation from the mean was divided by the standard deviation). The average linkage method (unweighted pair group method with arithmetic mean) was used for cluster generation, with centered Pearson's correlation as distance/similarity measure. Nodes were arranged with the smaller subtrees kept to the right.

qPCR

Quantitative RT-PCR (qPCR) was performed using the "EXPRESS One-Step SYBR GreenER with Premixed ROX" kit (Life Technologies/Invitrogen) and the StepOnePlus Real-Time PCR System (Applied Biosystems). Reactions were performed in triplicate (each from three biological replicates) with a total reaction volume of 10 μL including 300 nM each of forward and reverse primers (Fig. S12) and 75 ng template RNA. Data Analysis was performed by the StepOne Software (Applied Biosystems) using the Relative Quantitation/Comparative CT (ΔΔCT) setting. Data were normalized to the endogenous control actin (NCU04173) with expression on sucrose or no carbon as the reference sample (as indicated). For quantitation of the *hac-1* splice versions, the primers used (Fig. S12) had been previously verified to be specific for their targets in RT-PCR and qPCR trials (not shown) using cDNA clones obtained from DTT treatment ("ON") or mock treatment ("OFF") (as shown in Fig. 8A).

Monosaccharide uptake assays

Assays following the sugar concentration in the medium were performed according to Galazka *et al.* (2010). Briefly, *N. crassa* cultures were pre-grown from 10 day-old conidia for 16 h in 3 mL of 1× Vogel's salts plus 2% (w/v) sucrose using 24 deep-well plates. The mycelia were then washed three times in 1× Vogel's salts without added carbon (NoC) and transferred to 0.5% (w/v) pectin or 1 mM D-Xyl for induction. After an additional 4 h, the mycelia were washed again as above and transferred into the uptake buffer (1x Vogel's salts plus 90 μM of each monosaccharide) for pre-equilibration and to reduce the dilution of the final uptake reaction. Two mycelia of the same genotype were combined into one well to

increase the biomass at this stage. In case of the use of the uncoupler carbonyl cyanide *m*-chlorophenyl hydrazone (CCCP), this was included starting at the pre-equilibration and added from a 10 mM stock solution in ethanol; control cultures received the equivalent amounts of ethanol without CCCP. After 5 min, the mycelia were transferred into fresh uptake buffer (with or without CCCP) to start the reaction. Time points of the supernatants were usually taken at 0 min, 5 min, and up to 40 min. The samples were cleared by centrifugation (1 min at 20,000 × g) and 50 μL diluted into 450 μL of H_2O. The monosaccharide concentrations were then quantified by HPAEC-PAD: a sample size of 25 μL was injected onto a Dionex CarboPac PA20 column (3×30 mm guard and 3×150 mm analytical) and eluted using an isocratic mobile phase of 18 mM KOH at 0.4 mL/min and 30 °C over 11 min. After the uptake assay, the fungal biomass was blotted dry and completely dried over night at 105 °C to determine the dry weight (generally between 2.5 and 3.0 mg) for data normalization.

Intracellular L-arabinose quantification

For analysis of intracellular L-Ara, mycelia pre-grown for 16 h in 2% sucrose were first switched to 0.5% pectin for another 4 h and then incubated for 20 min with 90 μM of L-Ara in 1x Vogel's salts in a standard uptake assay (see above). The uptake was rapidly stopped by a transfer of the biomass to 3 mL of −20 °C cold 50% methanol. The mycelia were washed by two additional transfers to fresh −20 °C cold 50% methanol, and finally transferred to pre-weighed microcentrifuge tubes and frozen in liquid nitrogen. After lyophilization over night the dry weight was determined for data normalization. Tissue was extracted by bead-beating for 2 min in presence of about 0.3 g of 0.5 mm silica/Zirkonia beads (Biospec) and 1 mL of 2:2:1 chloroform-methanol-water. After 10 min on ice and centrifugation for 10 min at 20.000 × g and 4 °C, the aqueous phase (400 μL) was transferred to a new tube and the solvent removed in a speed-vac. The pellet was resuspended in 50 μL of ddH_2O, centrifuged again to pellet any potentially insoluble debris, and the clear supernatant diluted 10-fold with ddH_2O for analysis by HPAECPAD (see above). Data were calculated with the assumption of an intracellular volume of 2.4 mL per g of dry biomass (Slayman & Tatum, 1964).

Phylogenetic analysis

Protein sequences were retrieved from the NCBI database. Sequences were aligned and phylogenetic trees constructed using the program Phylogeny.fr (Dereeper *et al.*, 2008) (alignment by MUSCLE excluding alignment curation by Gblocks and tree generation by PhyML). For Fig. S11 the Newick output file was subsequently visualized using iTOL (Letunic & Bork, 2007). PelA from *A. niger* (Fig. S10) and Gal2p from *Saccharomyces cerevisiae* (Fig. S11) were used as outgroup, respectively.

Localization of the unconventional intron in hac-1

Sucrose pre-grown (16h) WT *N. crassa* were shifted for 4h to either fresh 2% sucrose or 2% sucrose + 10 mM DTT to induce the UPR. RNA was extracted, cDNA prepared and the region surrounding the putative intron amplified (primers: hac1_intron_fwd & hac1_intron_rev; Fig. S12). The PCR products were separated into a higher MW and a lower MW band (not shown), which were cloned into *E.coli* and several clones sequenced. The majority of high MW band clones had the “OFF” sequence while all lower MW band clones had the “ON” sequence (compare Fig. 8A).

A clear assignment of the intron is made difficult by a “CCTG” repeat at both borders, but the location depicted in Fig. 8A is homologous with those in *A. nidulans* and *T. reesei* (compare Fig. S8A). Also, the secondary structure prediction (ensemble centroid structure) using the program *S*fold 2.2 (default parameters; http://sfold.wadsworth.org; Ding *et al.*,

2005; Ding & Lawrence, 2003) and the sequence 5'-CGATGCGACACAACATCCTGCTGTGGTGTTGTGTGATGACCTGCAGTGTCGG-3' indicates these positions are exposed for IRE-1-dependent cleavage.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We want to thank James Craig, Samuel Coradetti, Mark Toews, Ana B. Ibáñez, Christina Kuchelmeister, Adam Cohen, Mara Bryan, and Monique Benz for excellent technical assistance and Junior Prof. Kai Heimel (Uni Goettingen, GER) for help with *hac-1*. We thank William Beeson, Christopher Phillips and Elizabeth Znameroski for helpful advice, and Morgann Reilly additionally for her comments on the manuscript. We acknowledge the FGSC and the *Neurospora* Genome Project for continuous support, and we are pleased to acknowledge use of materials generated by the NIH grant P01 GM068087 'Functional analysis of a model filamentous fungus'. This work was supported by grants from the Energy Biosciences Institute to CRS and NLG. JPB was supported by a Feodor-Lynen fellowship from the Humboldt-Foundation (GER).

References

Adav SS, Chao LT, Sze SK. Quantitative secretomic analysis of Trichoderma reesei strains reveals enzymatic composition for lignocellulosic biomass degradation. Mol Cell Proteomics. 2012; 11

Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A. Plant Cell Walls - From Chemistry to Biology. Garland Science. 2011

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. J Mol Biol. 1990; 215:403–410. [PubMed: 2231712]

An J, O'Neill MA, Albersheim P, Darvill AG. Isolation and structural characterization of beta-D-glucosyluronic acid and 4-O-methyl beta-D-glucosyluronic acid-containing oligosaccharides from the cell-wall pectic polysaccharide, rhamnogalacturonan I. Carbohydr Res. 1994; 252:235–243. [PubMed: 8137363]

Aro N, Pakula T, Penttila M. Transcriptional regulation of plant cell wall degradation by filamentous fungi. Fems Microbiol Rev. 2005; 29:719–739. [PubMed: 16102600]

Beeson WT, Phillips CM, Cate JHD, Marletta MA. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J Am Chem Soc. 2012; 134:890–892. [PubMed: 22188218]

Benen JAE, Kester HCM, Visser J. Kinetic characterization of Aspergillus niger N400 endopolygalacturonases I, II and C. Eur J Biochem. 1999; 259:577–585. [PubMed: 10092840]

Benoit I, Coutinho PM, Schols HA, Gerlach JP, Henrissat B, de Vries RP. Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. BMC Genomics. 2012; 13

Berka RM, Grigoriev IV, Otillar R, Salamov A, Grimwood J, Reid I, Ishmael N, John T, Darmond C, Moisan MC, Henrissat B, Coutinho PM, Lombard V, Natvig DO, Lindquist E, Schmutz J, Lucas S, Harris P, Powlowski J, Bellemare A, Taylor D, Butler G, de Vries RP, Allijn IE, van den Brink J, Ushinsky S, Storms R, Powell AJ, Paulsen IT, Elbourne LDH, Baker SE, Magnuson J, LaBoissiere S, Clutterbuck AJ, Martinez D, Wogulis M, de Leon AL, Rey MW, Tsang A. Comparative genomic analysis of the thermophilic biomass-degrading fungi Myceliophthora thermophila and Thielavia terrestris. Nat Biotechnol. 2011; 29:922–U222. [PubMed: 21964414]

Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, Plamann M, Goodrich-Tanrikulu M, Schulte U, Mannhaupt G, Nargang FE, Radford A, Selitrennikoff C, Galagan JE, Dunlap JC, Loros JJ, Catcheside D, Inoue H, Aramayo R, Polymenis M, Selker EU, Sachs MS, Marzluf GA, Paulsen I, Davis R, Ebbole DJ, Zelter A, Kalkman ER, O'Rourke R, Bowring F, Yeadon J, Ishii C, Suzuki K, Sakai W, Pratt R. Lessons from the genome sequence of Neurospora crassa: Tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol R. 2004; 68:1–108.

Braaksma M, Martens-Uzunova ES, Punt PJ, Schaap PJ. An inventory of the Aspergillus niger secretome by combining in silico predictions with shotgun proteomics data. BMC Genomics. 2010; 11:584. [PubMed: 20959013]

Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohyd Res. 2009; 344:1879–1900.

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–D238. [PubMed: 18838391]

Christgau S, Kofod LV, Halkier T, Andersen LN, Hockauf M, Dorreich K, Dalboge H, Kauppinen S. Pectin methyl esterase from Aspergillus aculeatus: Expression cloning in yeast and characterization of the recombinant enzyme. Biochem J. 1996; 319:705–712. [PubMed: 8920970]

Chundawat SPS, Beckham GT, Himmel ME, Dale BE. Deconstruction of lignocellulosic biomass to fuels and chemicals. Annu Rev Chem Biomol. 2011; 2:121–145.

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 Nat Acad Sci USA. 2006; 103:10352–10357. [PubMed: 16801547]

Conner AH, Anderson L. Tautomerization and mutarotation of beta-L-arabinopyranose -participation of both furanose anomers. Carbohyd Res. 1972; 25:107–116.

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 Nat Acad Sci USA. 2012; 109:7397–7402. [PubMed: 22532664]

Couturier M, Navarro D, Olive C, Chevret D, Haon M, Favel A, Lesage-Meessen L, Henrissat B, Coutinho PM, Berrin JG. Post-genomic analyses of fungal lignocellulosic biomass degradation reveal the unexpected potential of the plant pathogen Ustilago maydis. BMC Genomics. 2012; 13:57. [PubMed: 22300648]

Dahodwala S, Humphrey A, Weibel M. Pectic Enzymes - Individual and Concerted Kinetic-Behavior of Pectinesterase and Pectinase. J Food Sci. 1974; 39:920–926.

Davis, RH. Neurospora : contributions of a model organism. Oxford University Press; Oxford: 2000.

Davis RH, Perkins DD. Neurospora: a model of model microbes. Nature Reviews Genetics. 2002; 3:397–403.

de Souza WR, de Gouvea PF, Savoldi M, Malavazi I, de Souza Bernardes LA, Goldman MH, de Vries RP, de Castro Oliveira JV, Goldman GH. Transcriptome analysis of Aspergillus niger grown on sugarcane bagasse. Biotechnol Biofuels. 2011; 4:40. [PubMed: 22008461]

de Vries RP, Jansen J, Aguilar G, Parenicova L, Joosten V, Wulfert F, Benen JA, Visser J. Expression profiling of pectinolytic genes from Aspergillus niger. FEBS Lett. 2002; 530:41–47. [PubMed: 12387863]

De Vries RP, Parenicova L, Hinz SW, Kester HC, Beldman G, Benen JA, Visser J. The beta-1,4-endogalactanase A gene from Aspergillus niger is specifically induced on arabinose and galacturonic acid and plays an important role in the degradation of pectic hairy regions. Eur J Biochem. 2002; 269:4985–4993. [PubMed: 12383257]

Delmas S, Pullan ST, Gaddipati S, Kokolski M, Malla S, Blythe MJ, Ibbett R, Campbell M, Liddell S, Aboobaker A, Tucker GA, Archer DB. Uncovering the genome-wide transcriptional responses of the filamentous fungus Aspergillus niger to lignocellulose using RNA sequencing. PLoS Genet. 2012; 8

Delourdes M, Polizeli TM, Jorge JA, Terenzi HF. Pectinase production by Neurospora crassa - Purification and biochemical characterization of extracellular polygalacturonase activity. J Gen Microbiol. 1991; 137:1815–1823. [PubMed: 1835496]

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008; 36:W465–W469. [PubMed: 18424797]

Ding Y, Chan CY, Lawrence CE. RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble. RNA. 2005; 11:1157–1166. [PubMed: 16043502]

Ding Y, Lawrence CE. A statistical sampling algorithm for RNA secondary structure prediction. Nucleic Acids Res. 2003; 31:7280–7301. [PubMed: 14654704]

Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F. Colorimetric method for determination of sugars and related substances. Anal Chem. 1956; 28:350–356.

Dunlap, JC.; Borkovich, KA.; Henn, MR.; Turner, GE.; Sachs, MS.; Glass, NL.; McCluskey, K.; Plamann, M.; Galagan, JE.; Birren, BW.; Weiss, RL.; Townsend, JP.; Loros, JJ.; Nelson, MA.; Lambreghts, R.; Colot, HV.; Park, G.; Collopy, P.; Ringelberg, C.; Crew, C.; Litvinkova, L.; DeCaprio, D.; Hood, HM.; Curilla, S.; Shi, M.; Crawford, M.; Koerhsen, M.; Montgomery, P.; Larson, L.; Pearson, M.; Kasuga, T.; Tian, C.; Basturkmen, M.; Altamirano, L.; Xu, J. Enabling a community to dissect an organism: overview of the Neurospora Functional Genomics Project.. In: Hall, JC.; Dunlap, JC.; Friedmann, T.; van Heyningen, V., editors. Advances in Genetics. Academic Press; 2007. p. 49-96.

Ebbole DJ. Carbon catabolite repression of gene expression and conidiation in Neurospora crassa. Fungal Genet Biol. 1998; 25:15–21. [PubMed: 9806802]

Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA. 1998; 95:14863–14868. [PubMed: 9843981]

Flipphi, M.; Felenbok, B. The onset of carbon catabolite repression and interplay between specific induction and carbon catabolite repression in Aspergillus nidulans.. In: Brambl, R.; Marzluf, GA., editors. Biochemistry and Molecular Biology. Vol. III. Springer; Berlin-Heidelberg: 2004. p. 403-420.

Foreman PK, Brown D, Dankmeyer L, Dean R, Diener S, Dunn-Coleman NS, Goedegebuur F, Houfek TD, England GJ, Kelley AS, Meerman HJ, Mitchell T, Mitchinson C, Olivares HA, Teunissen PJ, Yao J, Ward M. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus Trichoderma reesei. J Biol Chem. 2003; 278:31988–31997. [PubMed: 12788920]

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang S, Nielsen CB, Butler J, Endrizzi M, Qui D, Ianakiev P, Bell-Pedersen D, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stange-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li W, Pratt RJ, Osmani SA, DeSouza CP, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B. The genome sequence of the filamentous fungus Neurospora crassa. Nature. 2003; 422:859–868. [PubMed: 12712197]

Galagan JE, Selker EU. RIP: the evolutionary cost of genome defense. Trends Genet. 2004; 20:417–423. [PubMed: 15313550]

Galazka JM, Tian C, Beeson WT, Martinez B, Glass NL, Cate JH. Cellodextrin transport in yeast for improved biofuel production. Science. 2010; 330:84–86. [PubMed: 20829451]

Geysens S, Whyteside G, Archer DB. Genomics of protein folding in the endoplasmic reticulum, secretion stress and glycosylation in the aspergilli. Fungal Genet Biol. 2009; 46:S121–S140.

Guillemette T, van Peij NNME, Goosen T, Lanthaler K, Robson GD, van den Hondel CAMJJ, Stam H, Archer DB. Genomic analysis of the secretion stress response in the enzyme-producing cell factory Aspergillus niger. BMC Genomics. 2007; 8

Haffner FB, Mitchell VD, Arundale RA, Bauer S. Compositional analysis of Miscanthus giganteus by near infrared spectroscopy. Cellulose. 2013; 20:1629–1637.

Hakkinen M, Arvas M, Oja M, Aro N, Penttila M, Saloheimo M, Pakula TM. Re-annotation of the CAZy genes of Trichoderma reesei and transcription in the presence of lignocellulosic substrates. Microb Cell Fact. 2012; 11:134. [PubMed: 23035824]

Harholt J, Suttangkakul A, Scheller HV. Biosynthesis of pectin. Plant Physiol. 2010; 153:384–395. [PubMed: 20427466]

Harris PV, Welner D, McFarland KC, Re E, Poulsen JCN, Brown K, Salbo R, Ding HS, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry. 2010; 49:3305–3316. [PubMed: 20230050]

Isshiki A, Akimitsu K, Yamamoto M, Yamamoto H. Endopolygalacturonase is essential for citrus black rot caused by Alternaria citri but not brown spot caused by *Alternaria alternata*. Mol Plant Microbe In. 2001; 14:749–757.

Jansen EF, MacDonnell LR, Jang R. Simultaneous actions of polygalacturonase and pectinesterase on pectin. Arch Biochem Biophys. 1945; 8:113–118.

Jayani RS, Saxena S, Gupta R. Microbial pectinolytic enzymes: a review. Process Biochem. 2005; 40:2931–2944.

Jordan DB, Bowman MJ, Braker JD, Dien BS, Hector RE, Lee CC, Mertens JA, Wagschal K. Plant cell walls to ethanol. Biochem J. 2012; 442:241–252. [PubMed: 22329798]

Kester HCM, Benen JAE, Visser J. The exopolygalacturonase from *Aspergillus tubingensis* is also active on xylogalacturonan. Biotechnol Appl Bioc. 1999; 30:53–57.

Kou SC, Christen. & Cirillo VP. Galactose transport in *Saccharomyces cerevisiae* 2 - characteristics of galactose uptake and exchange in galacktokinaseless cells. J Bacteriol. 1970; 103:671–&. [PubMed: 5474882]

Kubicek CP. Systems biological approaches towards understanding cellulase production by *Trichoderma reesei*. J Biotechnol. 2013; 163:133–142. [PubMed: 22750088]

Kwon MJ, Jorgensen TR, Nitsche BM, Arentshorst M, Park J, Ram AFJ, Meyer V. The transcriptomic fingerprint of glucoamylase over-expression in Aspergillus niger. BMC Genomics. 2012; 13

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. [PubMed: 19261174]

Lara-Marquez A, Zavala-Paramo MG, Lopez-Romero E, Camacho HC. Biotechnological potential of pectinolytic complexes of fungi. Biotechnol Lett. 2011; 33:859–868. [PubMed: 21246254]

Leeder AC, Jonkers W, Li J, Glass NL. Germination and Early Colony Establishment in *Neurospora crassa* Requires a MAP Kinase Regulatory Network. Genetics. 2013 doi: 10.1534/genetics.113.156984.

Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007; 23:127–128. [PubMed: 17050570]

Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels. 2013; 6:41. [PubMed: 23514094]

Li D, Bobrowicz P, Wilkinson HH, Ebbole DJ. A mitogen-activated protein kinase pathway essential for mating and contributing to vegetative growth in Neurospora crassa. Genetics. 2005; 170:1091–1104. [PubMed: 15802524]

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; 46:768–781. [PubMed: 19555771]

Martens-Uzunova ES, Zandleven JS, Benen JAE, Awad H, Kools HJ, Beldman G, Voragen AGJ, Van den Berg JA, Schaap PJ. A new group of exo-acting family 28 glycoside hydrolases of Aspergillus niger that are involved in pectin degradation. Biochem J. 2006; 400:43–52. [PubMed: 16822232]

Martens-Uzunova ES, Schaap PJ. Assessment of the pectin degrading enzyme network of Aspergillus niger by functional genomics. Fungal Genet Biol. 2009; 46:S170–S179. [PubMed: 19618506]

Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, Hammel KE, Wymelenberg AV, Gaskell J, Lindquist E, Sabat G, BonDurant SS, Larrondo LF, Canessa P, Vicuna R, Yadav J, Doddapaneni H, Subramanian V, Pisabarro AG, Lavin JL, Oguiza JA, Master E, Henrissat B, Coutinho PM, Harris P, Magnuson JK, Baker SE, Bruno K, Kenealy W, Hoegger PJ, Kues U, Ramaiya P, Lucash S, Salamov A, Shapiro H, Tu H, Chee CL, Misra M, Xie G, Teter S, Yaver D, James T, Mokrejs M, Pospisek M, Grigoriev IV, Brettin T, Rokhsar D, Berka R, Cullen D. Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA. 2009; 106:1954–1959. [PubMed: 19193860]

Massiot P, Perron V, Baron A, Drilleau JF. Release of methanol and depolymerization of highly methyl esterified apple pectin with an endopolygalacturonase from Aspergillus niger and pectin methylesterases from A-niger or from orange. Food Sci Technol-Leb. 1997; 30:697–702.

Mohnen D. Pectin structure and biosynthesis. Curr Opin Plant Biol. 2008; 11:266–277. [PubMed: 18486536]

Moore KA, Hollien J. The unfolded protein response in secretory cell function. Annu Rev Genet. 2012; 46:165–183. [PubMed: 22934644]

Moran-Diez E, Hermosa R, Ambrosino P, Cardoza RE, Gutierrez S, Lorito M, Monte E. The ThPG1 endopolygalacturonase is required for the *Trichoderma harzianum*-plant beneficial interaction. Mol Plant Microbe In. 2009; 22:1021–1031.

Mutter M, Beldman G, Schols HA, Voragen AGJ. Rhamnogalacturonan alpha-L-rhamnopyranohydrolase - a novel enzyme specific for the terminal nonreducing rhamnosyl unit in rhamnogalacturonan regions of pectin. Plant Physiol. 1994; 106:241–250. [PubMed: 7972516]

Navarrete M, Callegari E, Eyzaguirre J. The effect of acetylated xylan and sugar beet pulp on the expression and secretion of enzymes by *Penicillium purpurogenum*. Appl Microbiol Biot. 2012; 93:723–741.

Nitsche BM, Jorgensen TR, Akeroyd M, Meyer V, Ram AFJ. The carbon starvation response of *Aspergillus niger* during submerged cultivation: Insights from the transcriptome and secretome. BMC Genomics. 2012; 13

O'Neill MA, Ishii T, Albersheim P, Darvill AG. Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annual Review of Plant Biology. 2004; 55:109–139.

Oyama S, Inoue H, Yamagata Y, Nakajima T, Abe K. Functional analysis of an endo-1,6-beta-D-glucanase gene (neg-1) from *Neurospora crassa*. Biosci Biotech Bioch. 2006; 70:1773–1775.

Paper JM, Scott-Craig JS, Adhikari ND, Cuom CA, Walton JD. Comparative proteomics of extracellular proteins in vitro and in planta from the pathogenic fungus *Fusarium graminearum*. Proteomics. 2007; 7:3171–3183. [PubMed: 17676664]

Pauly M, Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J. 2008; 54:559–568. [PubMed: 18476863]

Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, Schaap PJ, Turner G, de Vries RP, Albang R, Albermann K, Andersen MR, Bendtsen JD, Benen JA, van den Berg M, Breestraat S, Caddick MX, Contreras R, Cornell M, Coutinho PM, Danchin EG, Debets AJ, Dekker P, van Dijck PW, van Dijk A, Dijkhuizen L, Driessen AJ, d'Enfert C, Geysens S, Goosen C, Groot GS, de Groot PW, Guillemette T, Henrissat B, Herweijer M, van den Hombergh JP, van den Hondel CA, van der Heijden RT, van der Kaaij RM, Klis FM, Kools HJ, Kubicek CP, van Kuyk PA, Lauber J, Lu X, van der Maarel MJ, Meulenberg R, Menke H, Mortimer MA, Nielsen J, Oliver SG, Olsthoorn M, Pal K, van Peij NN, Ram AF, Rinas U, Roubos JA, Sagt CM, Schmoll M, Sun J, Ussery D, Varga J, Vervecken W, van de Vondervoort PJ, Wedler H, Wosten HA, Zeng AP, van Ooyen AJ, Visser J, Stam H. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol. 2007; 25:221–231. [PubMed: 17259976]

Perkins DD, Turner BC. *Neurospora* from natural populations: toward the population biology of a haploid eukaryote. Exp Mycol. 1988; 12:91–131.

Perkins DD, Turner BC, Barry EG. Strains of *Neurospora* collected from nature. Evolution. 1976; 30:281–313.

Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011; 8:785–786. [PubMed: 21959131]

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; 6:1399–1406. [PubMed: 22004347]

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; 62:567–590. [PubMed: 21351878]

Pressey R, Avants JK. Solubilization of Cell-Walls by Tomato Polygalacturonases - Effects of Pectin-Esterases. J Food Biochem. 1982; 6:57–74.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD. The Pfam protein families database. Nucleic Acids Res. 2012; 40:D290–D301. [PubMed: 22127870]

Renard CMGC, Crepeau MJ, Thibault JF. Glucuronic acid directly linked to galacturonic acid in the rhamnogalacturonan backbone of beet pectins. Eur J Biochem. 1999; 266:566–574. [PubMed: 10561599]

Rivas B, Torrado A, Torre P, Converti A, Dominguez JM. Submerged citric acid fermentation on orange peel autohydrolysate. J Agric Food Chem. 2008; 56:2380–2387. [PubMed: 18321055]

Roberts A, Pimentel H, Trapnell C, Pachter L. Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics. 2011a; 27:2325–2329. [PubMed: 21697122]

Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol. 2011b; 12

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; 32:5539–5545. [PubMed: 15486203]

Ruijter GJ, Visser J. Carbon repression in *Aspergilli*. FEMS Microbiol Lett. 1997; 151:103–114. [PubMed: 9228741]

Saloheimo M, Pakula TM. The cargo and the transport system: secreted proteins and protein secretion in Trichoderma reesei (Hypocrea jecorina). Microbiology. 2012; 158:46–57. [PubMed: 22053009]

Saloheimo M, Valkonen M, Penttila M. Activation mechanisms of the HAC1-mediated unfolded protein response in filamentous fungi. Mol Microbiol. 2003; 47:1149–1161. [PubMed: 12581366]

Schell MA, Roberts DP, Denny TP. Analysis of the *Pseudomonas solanacearum* polygalacturonase encoded by Pgla and its involvement in phytopathogenicity. J Bacteriol. 1988; 170:4501–4508. [PubMed: 3049534]

Scheller HV, Ulvskov P. Hemicelluloses. Annual Review of Plant Biology. 2010; 61:263–289. 61.

Scott-Craig JS, Cheng YQ, Cervone F, De Lorenzo G, Pitkin JW, Walton JD. Targeted mutants of *Cochliobolus carbonum* lacking the two major extracellular polygalacturonases. Appl Environ Microb. 1998; 64:1497–1503.

Seiboth B, Metz B. Fungal arabinan and L-arabinose metabolism. Appl Microbiol Biot. 2011; 89:1665–1673.

Shah P, Gutierrez-Sanchez G, Orlando R, Bergmann C. A proteomic study of pectin-degrading enzymes secreted by *Botrytis cinerea* grown in liquid culture. Proteomics. 2009; 9:3126–3135. [PubMed: 19526562]

Slayman CW, Tatum EL. Potassium Transport in *Neurospora* .I. Intracellular Sodium + Potassium Concentrations + Cation Requirements for Growth. Biochim Biophys Acta. 1964; 88:578–&. [PubMed: 14249101]

Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J. Laboratory Analytical Procedures (LAP). National Renewable Energy Laboratory (NREL); Golden, CO: 2011. Determination of Structural Carbohydrates and Lignin in Biomass..

Sluiter, A.; Hames, B.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D. Laboratory Analytical Procedures (LAP). National Renewable Energy Laboratory (NREL); Golden, CO: 2005. Determination of Ash in Biomass..

Sluiter, A.; Ruiz, R.; Scarlata, C.; Sluiter, J.; Templeton, D. Laboratory Analytical Procedures (LAP). National Renewable Energy Laboratory (NREL); Golden, CO: 2008. Determination of Extractives in Biomass..

Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, Vorwerk S, Youngs H. Toward a systems approach to understanding plant-cell walls. Science. 2004; 306:2206–2211. [PubMed: 15618507]

Somerville C, Youngs H, Taylor C, Davis SC, Long SP. Feedstocks for lignocellulosic biofuels. Science. 2010; 329:790–792. [PubMed: 20705851]

Strauss J, Mach RL, Zeilinger S, Hartler G, Stoffler G, Wolschek M, Kubicek CP. Cre1, the carbon catabolite repressor protein from *Trichoderma reesei*. FEBS Lett. 1995; 376:103–107. [PubMed: 8521952]

Subtil T, Boles E. Improving L-arabinose utilization of pentose fermenting Saccharomyces cerevisiae cells by heterologous expression of L-arabinose transporting sugar transporters. Biotechnol Biofuels. 2011; 4:38. [PubMed: 21992610]

Sun J, Glass NL. Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. PLoS ONE. 2011; 6:e25654. [PubMed: 21980519]

Sun J, Tian C, Diamond S, Glass NL. Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. Eukaryot Cell. 2012; 11:482–493. [PubMed: 22345350]

Szilagyi M, Miskei M, Karanyi Z, Lenkey B, Pocsi I, Emri T. Transcriptome changes initiated by carbon starvation in *Aspergillus nidulans*. Microbiol-SGM. 2013; 159:176–190.

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 Nat Acad Sci USA. 2009; 106:22157–22162. [PubMed: 20018766]

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010; 28:511–U174. [PubMed: 20436464]

Tsang A, Butler G, Powlowski J, Panisko EA, Baker SE. Analytical and computational approaches to define the *Aspergillus niger* secretome. Fungal Genet Biol. 2009; 46:S153–S160. [PubMed: 19618504]

Tucker, GA.; Seymour, GB. Modification and degradation of pectins. In Pectins and their manipulation. Seymour, GB.; Knox, JP., editors. Blackwell Publishing Ltd; Oxford: 2002. p. 150-173.

Turner BC, Perkins DD, Fairfield A. *Neurospora* from natural populations: A global study. Fungal Genet Biol. 2001; 32:67–92. [PubMed: 11352529]

Valette-Collet O, Cimerman A, Reignault P, Levis C, Boccara M. Disruption of *Botrytis cinerea* pectin methylesterase gene Bcpme1 reduces virulence on several host plants. Mol Plant Microbe In. 2003; 16:360–367.

van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biot. 2011; 91:1477–1492.

Verho R, Penttila M, Richard P. Cloning of two genes (LAT1,2) encoding specific L-arabinose transporters of the L-arabinose fermenting yeast *Ambrosiozyma monospora*. Appl Biochem Biotech. 2011; 164:604–611.

Vincken JP, Schols HA, Oomen RJFJ, McCann MC, Ulvskov P, Voragen AGJ, Visser RGF. If homogalacturonan were a side chain of rhamnogalacturonan I. Implications for cell wall architecture. Plant Physiol. 2003; 132:1781–1789. [PubMed: 12913136]

Vogel HJ. A convenient growth medium for *Neurospora* (medium N). Microbial Genetics Bulletin. 1956; 13:42.

Wakabayashi K, Hoson T, Huber DJ. Methyl de-esterification as a major factor regulating the extent of pectin depolymerization during fruit ripening: a comparison of the action of avocado (Persea americana) and tomato (Lycopersicon esculentum) polygalacturonases. J Plant Physiol. 2003; 160:667–673. [PubMed: 12872489]

Walter P, Ron D. The unfolded protein response: from stress pathway to homeostatic regulation. Science. 2011; 334:1081–1086. [PubMed: 22116877]

Willats WGT, McCartney L, Mackie W, Knox JP. Pectin: cell biology and prospects for functional analysis. Plant Mol Biol. 2001; 47:9–27. [PubMed: 11554482]

Wymelenberg AV, Gaskell J, Mozuch M, Sabat G, Ralph J, Skyba O, Mansfield SD, Blanchette RA, Martinez D, Grigoriev I, Kersten PJ, Cullen D. Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. Appl Environ Microb. 2010; 76:3599–3610.

Yadav V, Yadav PK, Yadav S, Yadav KDS. alpha-L-Rhamnosidase: a review. Process Biochem. 2010; 45:1226–1235.

Youngs H, Somerville C. Development of feedstocks for cellulosic biofuels. F1000 Biology Reports 2012. 2012; 4:11.

Zandleven J, Beldman G, Bosveld M, Benen J, Voragen A. Mode of action of xylogalacturonan hydrolase towards xylogalacturonan and xylogalacturonan oligosaccharides. Biochem J. 2005; 387:719–725. [PubMed: 15560751]

Zimmermann R, Eyrisch S, Ahmad M, Helms V. Protein translocation across the ER membrane. BBA-Biomembranes. 2011; 1808:912–924. [PubMed: 20599535]

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 Nat Acad Sci USA. 2012; 109:6012–6017. [PubMed: 22474347]

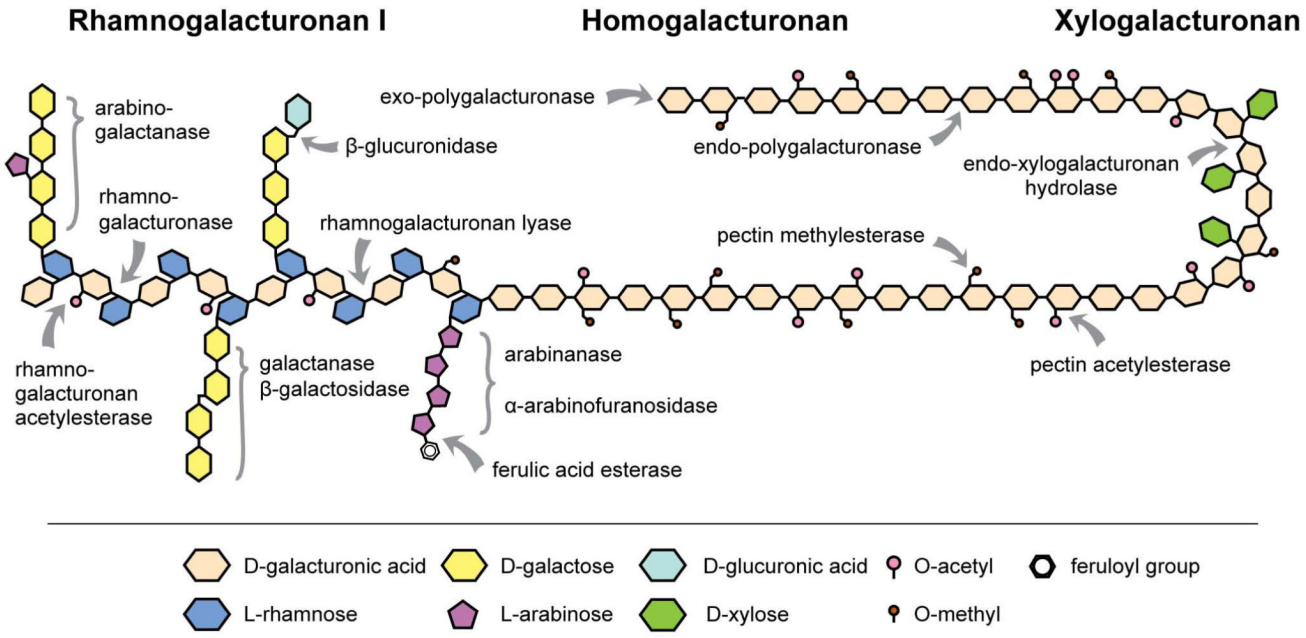

Fig. 1.

Schematic structure of pectin and the sites of action of pectinolytic enzymes. Pectin is a family of at least four different types of polysaccharides: homogalacturonan (HG), rhamnogalacturonan I (RG-I), xylogalacturonan (XG), and rhamnogalacturonan II (which is not shown here since it is not known to be a substrate for microbial degradation). The figure shows representative structures and is not to scale. HG and RG-I are generally much more abundant in plant cell walls than the other components. Please note that the enzyme classes indicated here are not necessarily found in all fungi.

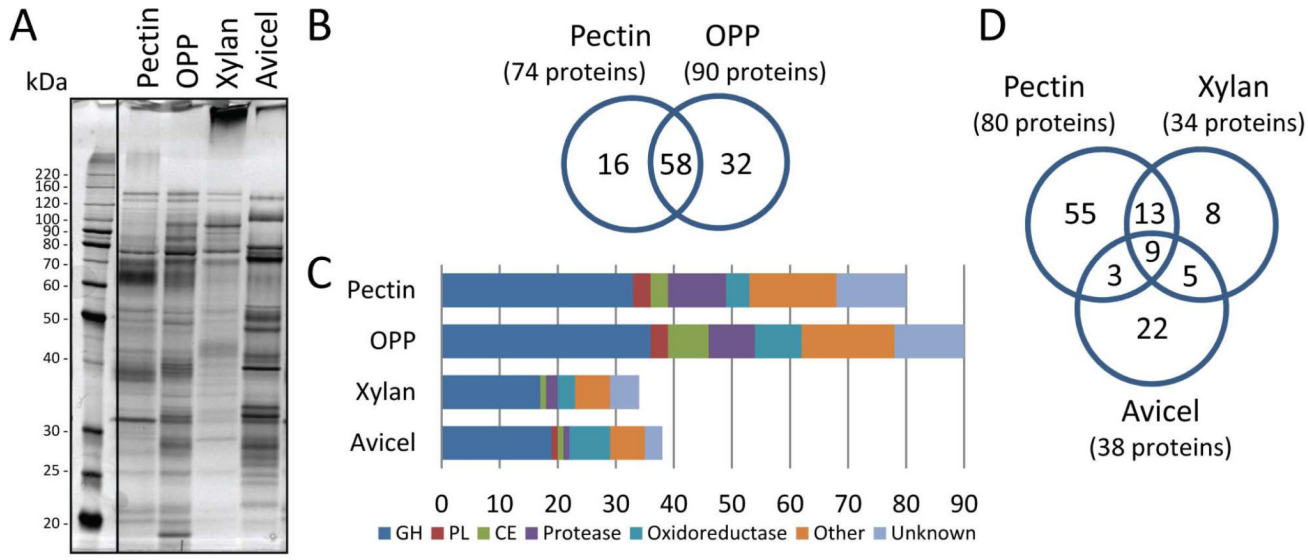

Fig. 2.

Comparative proteomic analysis of the *N. crassa* secretome on pectin and orange peel powder. (A) Silver stained SDS-PAGE of *N. crassa* secretomes from culture supernatants of pectin, orange peel powder (OPP), xylan, or Avicel as a sole carbon source. (B) Comparison of the secretomes from pectin or OPP-grown cultures as determined by LC-MS/MS after ion-exchange fractionation and in-solution digest. In the pectin and OPP secretomes, 74 and 90 proteins were identified with confidence, respectively. (C) Functional categorization of the pectin secretome (in-solution digest and gel-excision combined) and OPP secretome and comparison to those described for xylan (2% for four days; Sun *et al.*, 2012) and Avicel (2% for seven days; Tian *et al.*, 2009). The high number of CAZymes (dark blue, red, and green bars) indicates the complexity of pectin as substrate. Polysaccharide monooxygenases (PMOs; formerly GH61) are here counted as oxidoreductases (turquoise). (D) Three-way comparison of pectin, xylan and Avicel secretomes. The majority of the proteins (9) common to all three conditions may be involved in fungal cell wall remodeling and integrity.

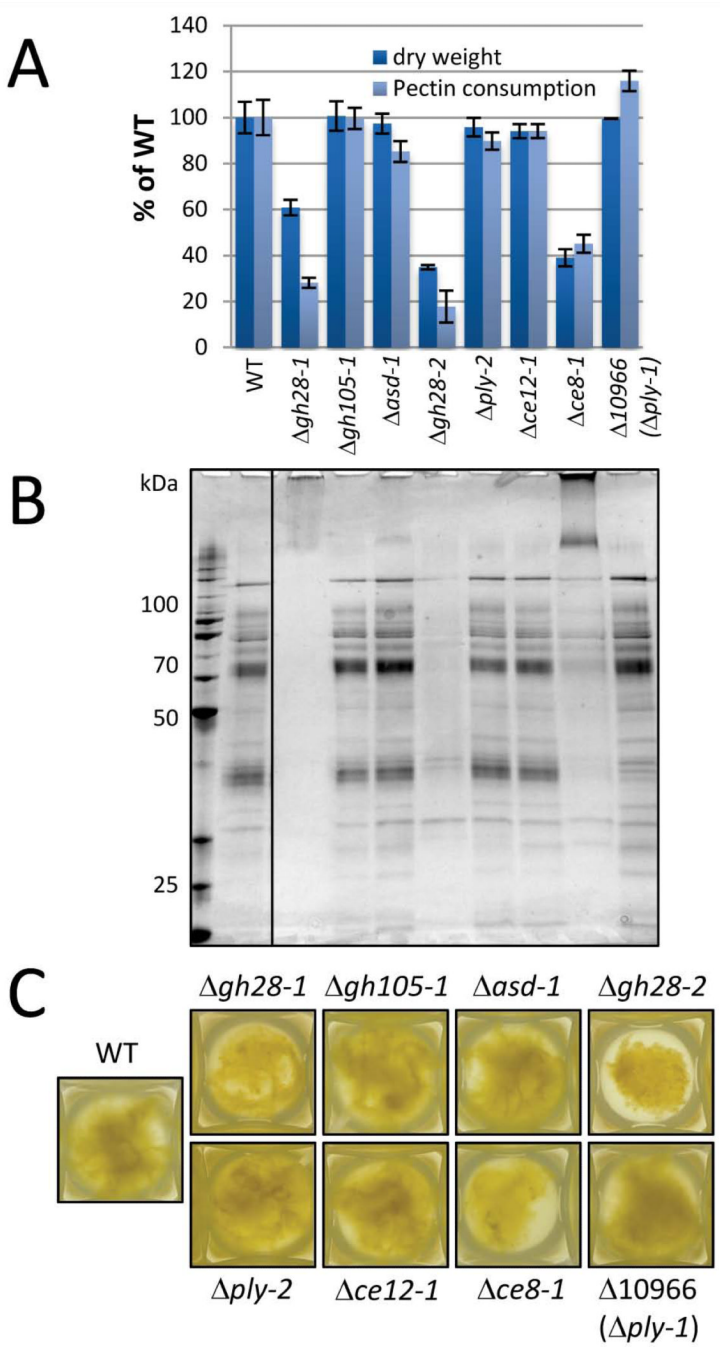

Fig. 3.

Pectin methylesterase and polygalacturonases are required for robust growth on pectin. (A) *N. crassa* strains containing deletions of genes involved in pectic backbone degradation (Table 1) were assayed for their phenotype when grown on pectin as sole C-source for four days. The fungal dry weight and the consumed pectin (the difference between day 0 and day 4 in total reducing ends in the culture supernatants) were measured and compared to WT (FGSC #2489). Bars represent standard deviations (n=3). (B) Silver stained SDS-PAGE of secreted proteins from the mutants in (A). The same volume (30 μL) of supernatant was loaded in each lane. (C) Visible growth phenotype of the same cultures as in (A) and (B)

after four days of growth on 1% pectin in 24-deep-well plates. Representative pictures are shown. Due to a former mis-annotation, δ10966 is actually a deletion of the pectate lyase *ply-1* (NCU06326) and NCU06327, encoding a cytochrome P450 protein. However, no phenotype was observed under these conditions.

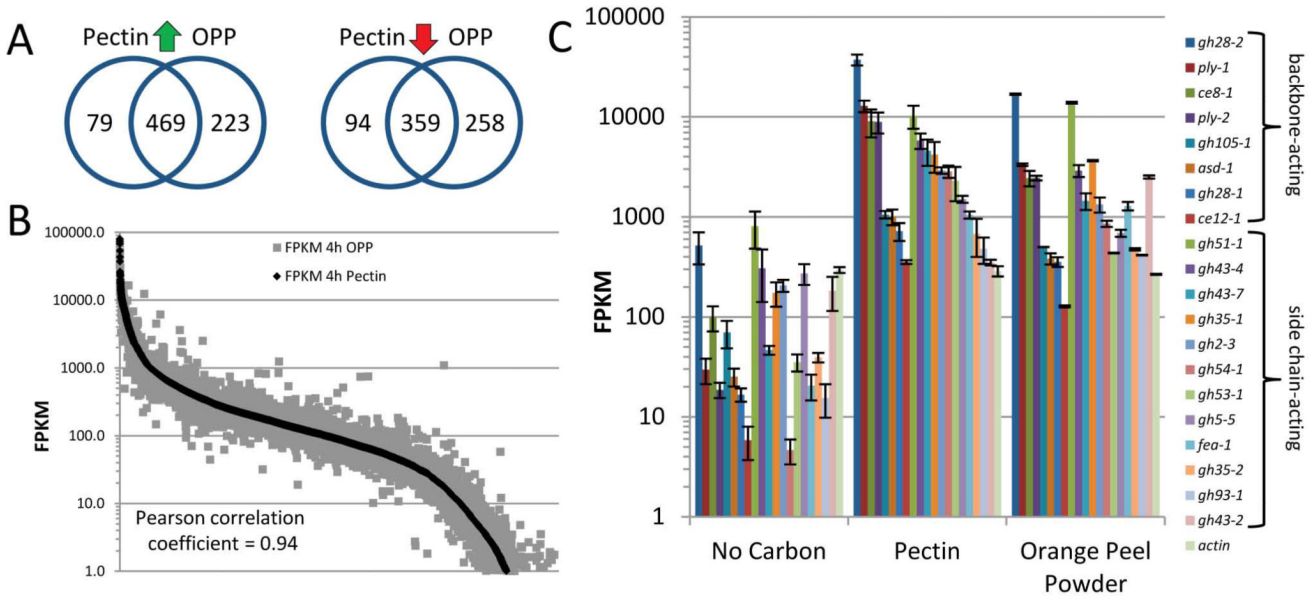

Fig. 4.

Transcriptional induction of pectinolytic genes by pectin and OPP in *N. crassa*. (A) Comparison of transcriptional response after 4 hr exposure of a 16 hr culture to pectin or to orange peel powder (OPP). Significantly up- (green arrow) or down- (red arrow) regulated genes versus the “no carbon” control condition were identified and compared. (B) Correlation plot of the full genome-expression profiles for pectin and OPP. All *N. crassa* genes are plotted along the x-axis and the corresponding FPKMs (fragments per kilobase of transcript per million mapped reads) along the y-axis using a $\log_{10}$ scale. Genes with expression values <1 FPKM are not shown. (C) Transcript abundances in FPKM ($\log_{10}$ scale) for a selection of genes encoding pectinolytic enzymes found to be significantly up-regulated on pectin or OPP versus both control conditions: “no carbon” (shown) and 2% sucrose (not shown). All eight genes encoding backbone-degrading enzymes as well as 12 genes coding for side-chain active enzymes (Table 1) are depicted and were ordered by expression strength on pectin within their category. Note the similarity of expression profiles between pectin and OPP for these genes. Actin was included as a non-induced (“house-keeping”) control. Error bars show standard deviation (NoC and pectin: n=3; OPP: n=2).

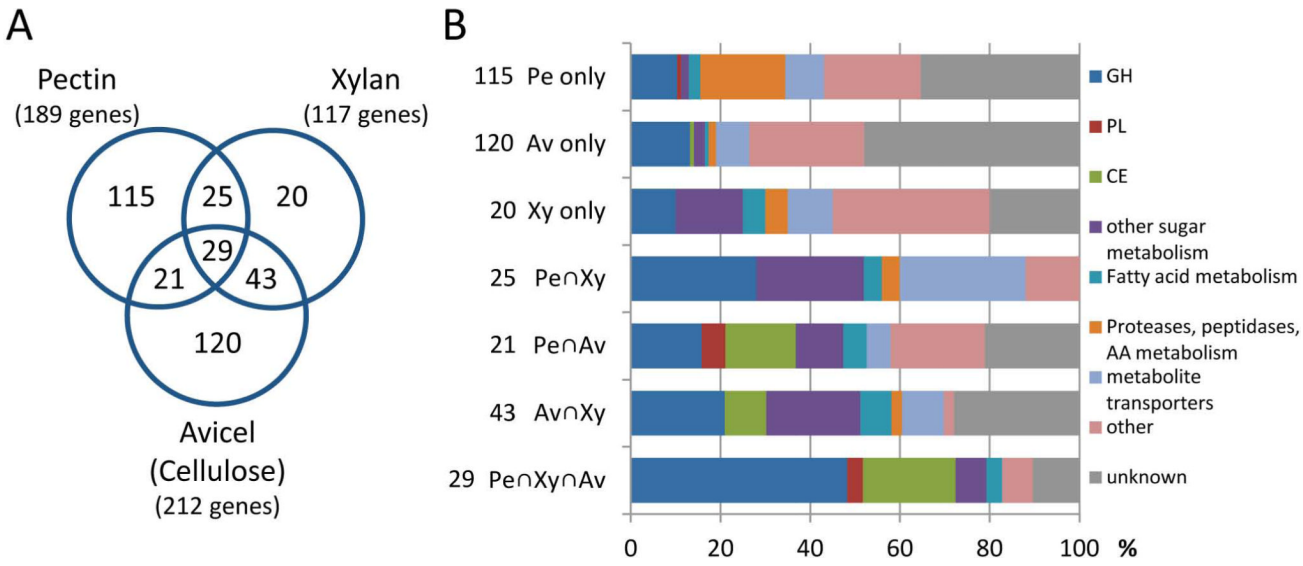

Fig. 5.

Comparative transcriptomics reveals the importance of polysaccharide esterases and oligosaccharide hydrolases. (A) Three-way comparison of the regulons for pectin, xylan and cellulose (Coradetti *et al.*, 2012). A regulon is comprised of all genes with elevated expression levels on a given C-source *vs*. both control conditions: NoC (starved) and sucrose (glucose-repressed) (see also Fig. S5 and Supporting data set 3). (B) Functional categorization of the genes in all seven Venn-diagram fractions resulting from the three-way comparison. The relative contribution of a functional group to each pool is depicted (total number of genes in each pool = 100%). Carbohydrate-active properties (glycoside hydrolases (GH), dark blue; polysaccharide lyases (PL), red; and carbohydrate esterases (CE), green) are clearly enriched in the four intersections. In particular the central section (29 genes induced under all conditions) is enriched for CEs and GHs acting as oligosaccharide hydrolases. Pe: pectin; Av: Avicel; Xy: xylan; ∩: intersection between sets.

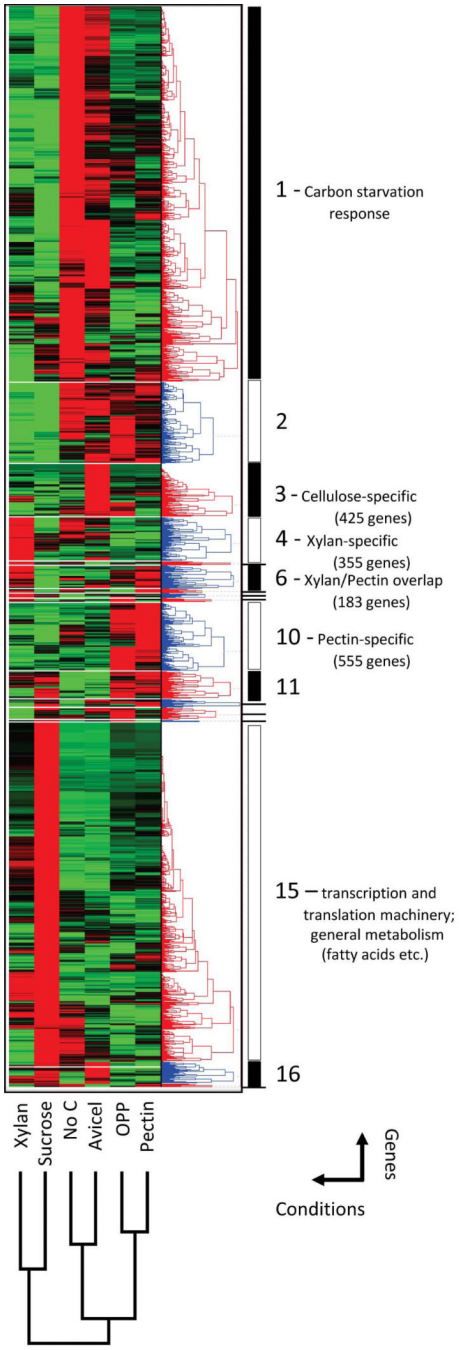

Fig. 6.

Hierarchical clustering of expression profiles for pectin, OPP, xylan, Avicel, sucrose and no carbon response. The averaged transcriptional expression profiles for the full-genome response to all three major plant cell wall polysaccharides (cellulose, xylan and pectin) as well as orange peel powder (OPP) and control conditions (no carbon and sucrose) were normalized and clustered in two dimensions: six conditions (horizontal) and ~8500 genes (vertical), using the software Hierarchical Clustering Explorer v3.0 (see also Supporting data set 4). The resulting heat map (normalized from bright green = low expression to bright red = strong expression) was divided into 17 sub-clusters. Large clusters are marked by black

and white bars, while smaller clusters are indicated by short dashes. Genes with consistently low expression (<10 FPKMs in all conditions) were not included in the clustering.

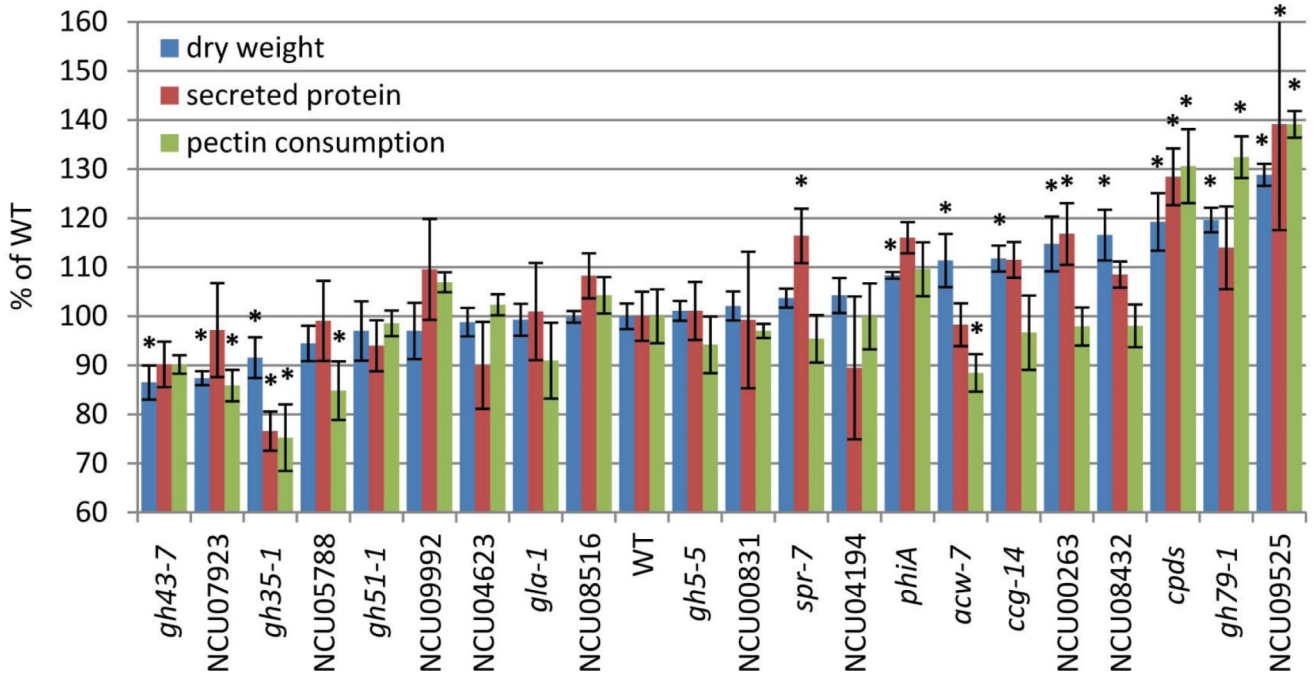

Fig. 7.

Screening of strains containing gene deletions for secreted proteins with pectin-specific expression patterns. *N. crassa* strains containing deletions of genes present in pectin-specific clusters 6 and 10 (Fig. 6) and that were also identified in the pectin and OPP secretomes (see Fig. S7 for annotation) were assayed for their phenotype when grown on pectin as sole C-source for four days. The fungal dry weight, the concentration of secreted protein, and the consumed pectin (the difference between day 0 and day 4 in total reducing ends in the culture supernatants) were measured and compared to WT (FGSC# 2489), which was set to 100%. The strains were ordered according to their recorded dry weight from the least (left) to the most (right). Bars represent standard deviations. (*) indicates a significant difference from WT with an unadjusted P-value of < 0.003 using a one-way Anova.

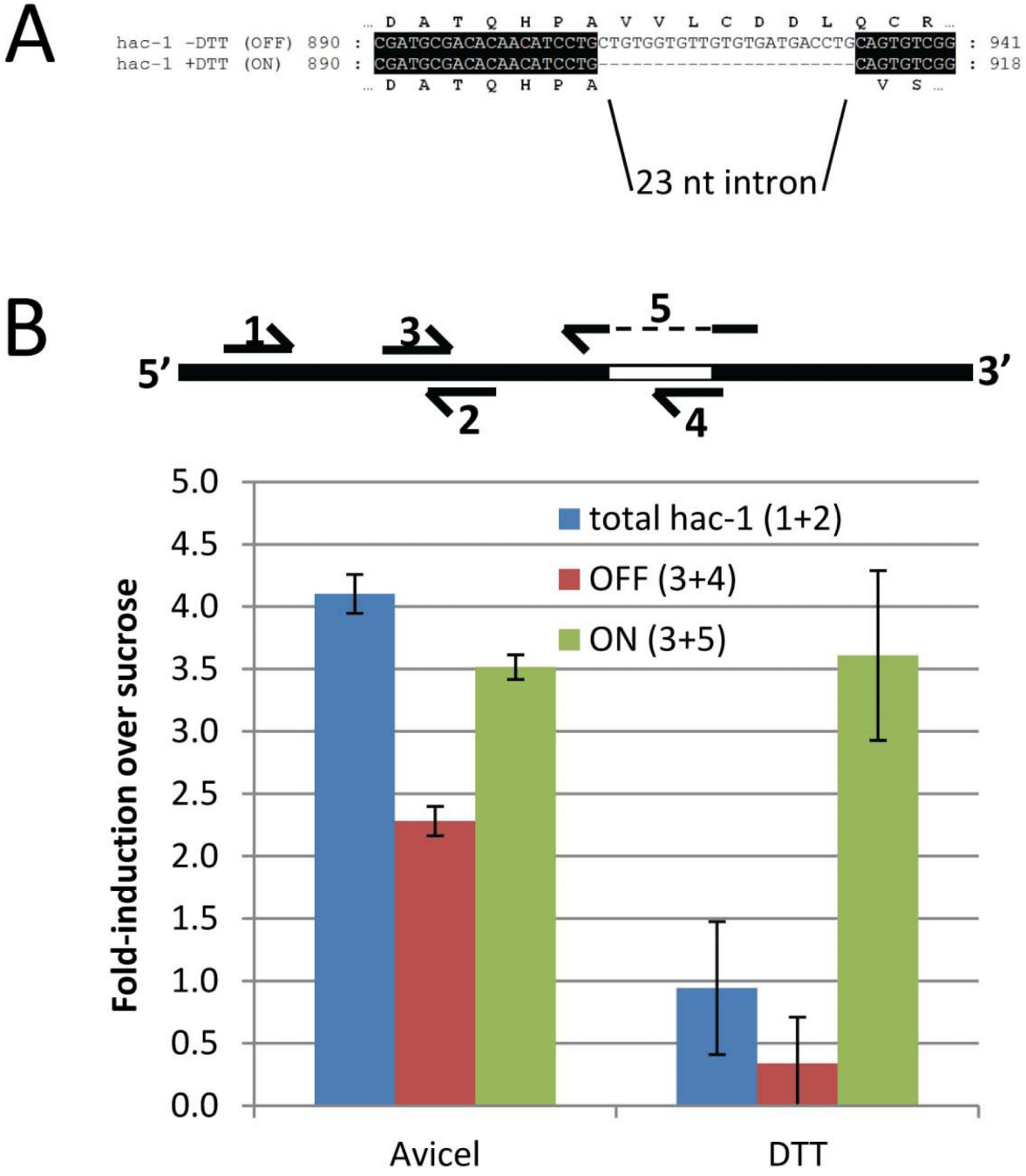

Fig. 8.

Growth on Avicel triggers the unfolded protein response (UPR). (A) Localization of the unconventional IRE-1-dependent intron in the *hac-1* mRNA. Shown is only the sequence in the immediate vicinity of the 23 nt intron including its translation, thereby showing the frame shift resulting from the splicing. (B) Relative quantification of the amount of *hac-1* mRNA splice variants by qPCR. Primers were chosen such that "total" *hac-1* transcript could be quantified (primers #1 and #2 in the above diagram) as well as the spliced ("ON"; primers #3 and #5), or non-spliced ("OFF"; primers #3 and #4) versions individually. For quantitation, 16h sucrose grown cultures were transferred to either fresh 2% sucrose, 2% sucrose + 10 mM DTT, or 0.5% Avicel. After 4h RNA was harvested, purified and

subjected to qPCR. The results are given as fold-induction over the expression state on 2% sucrose at 4h for each target.

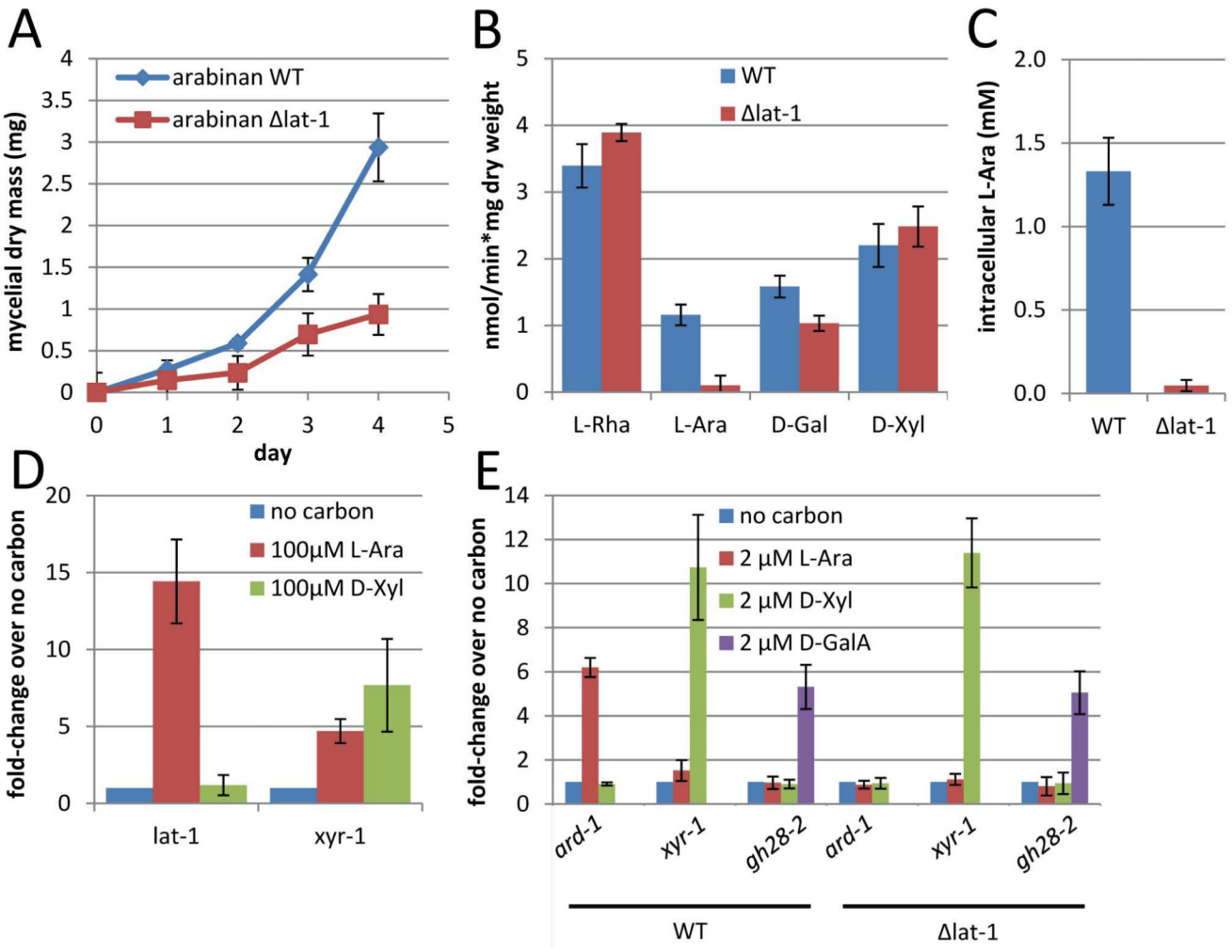

Fig. 9.

Characterization of the newly identified L-arabinose transporter LAT-1. (A) The Δ*lat-1* deletion strain displays a major growth phenotype on arabinan. For this experiment, sucrose pre-grown WT and Δ*lat-1* cultures were transferred to fresh 0.5% arabinan each day for 4 days. Shown is the average dry weight +/- SD from triplicate cultures at each transfer. (B) The Δ*lat-1* deletion strain is unable to take up L-Ara. *N. crassa* WT and the Δ*lat-1* strain were incubated with a 90 μM mixture (each) of the most abundant neutral sugars present in pectin: L-rhamnose (L-Rha), L-arabinose (L-Ara), D-galactose (D-Gal), and D-xylose (D-Xyl). Uptake rate was calculated from the monosaccharide consumption over the initial five minutes by analyzing aliquots of the supernatants by HPAEC-PAD. (C) Intracellular L-Ara accumulation. *N. crassa* WT and Δ*lat-1* were incubated for 20 min in 90 μM L-Ara, the uptake stopped in −20 °C cold 50% methanol, and the lyophilized biomass extracted by chloroform:methanol:water. The amount of L-Ara in the aqueous phase was determined by HPAEC-PAD. (D) *lat-1* is specifically induced by L-Ara. The expression strength of *lat-1* (NCU02188) and *xyr-1* (NCU08384) in *N. crassa* WT was assessed by qPCR after a 4h transfer to either no carbon or 100 μM L-Ara or D-Xyl. The data are normalized to expression on no carbon (=1) and are the average +/- SD from triplicate cultures. (E) The absence of LAT-1 specifically affects L-Ara signaling. The expression strength of *ard-1* (NCU00643), *xyr-1* (NCU08384) and *gh28-2* (NCU06961) in WT *vs*. Δ*lat-1* was assessed by qPCR after a 4h transfer to either no carbon or 2 μM of the indicated monosaccharides.

The data are normalized to expression on no carbon (=1) and are the average +/− SD from triplicate cultures. Specifically the induction of *ard-1* by L-Ara is absent in the Δ*lat-1* background.

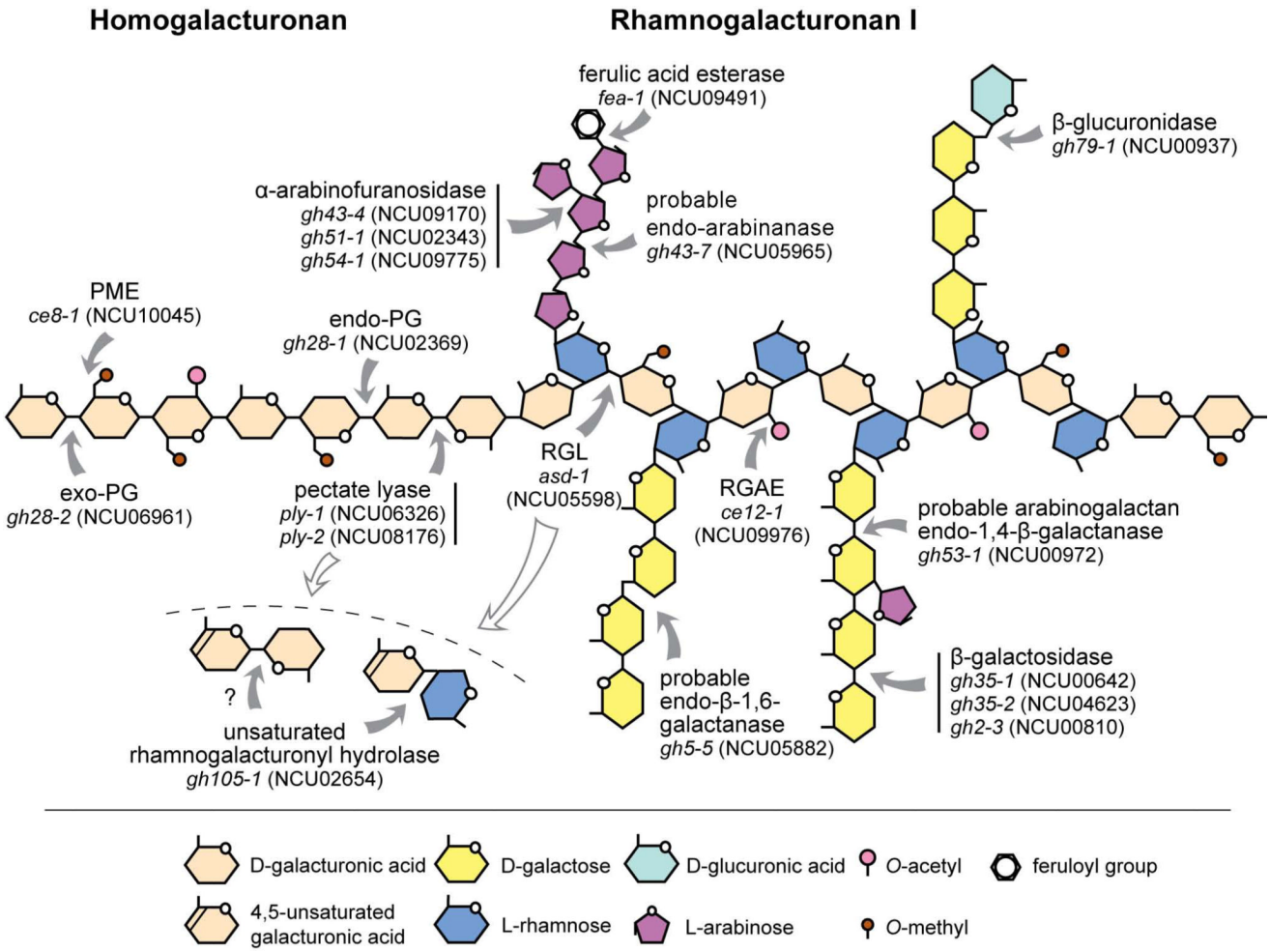

Fig. 10.

Model view of the pectin-degrading machinery of *N. crassa*. Depicted is an inventory of the *N. crassa* genes encoding enzymes with pectinolytic activities, that were found to be specifically upregulated on pectin and were also identified from the pectin and/or OPP secretomes (exception: *gh53-1* was not detectable in the secretome but significantly induced on pectin). Grey solid arrows indicate the points of attack on the pectin backbone or RG-I side chains. The open arrows indicate that the (unsaturated) D-GalA oligomers liberated by the action of the rhamnogalacturonan lyase ASD-1 (but possibly also the pectate lyases PLY-1 and PLY-2) are most likely imported for further hydrolysis inside the cell. Note that the pectin model is not to scale and that the linkages shown are of reported structures, but (in particular for RG-I) are not limited to these. RG-II and XG are not shown, since no dedicated enzymes were found in the *N. crassa* secretome and transcriptome. PG: polygalacturonase; PME: pectin methylesterase; RGAE: rhamnogalacturonan acetylesterase; RGL: rhamnogalacturonan lyase.

Table 1

N. crassa genes encoding pectinolytic enzymes that are strongly induced by pectin.

Locus	gene symbol	gene product; (putative) function	CAZy family	SP
Backbone-degrading pectinases				
NCU02369	*gh28-1*	(endo-) polygalacturonase	GH28	yes
NCU06961	*gh28-2*	(exo-) polygalacturonase	GH28	yes
NCU10045	*ce8-1*	pectin methylesterase	CE8	yes
NCU06326	*ply-1*	pectate lyase 1	PL1	yes
NCU08176	*ply-2*	pectate lyase A	PL3	yes
NCU05598	*asd-1*	rhamnogalacturonase B	PL4	yes
NCU09976	*ce12-1*	rhamnogalacturonan acetylesterase	CE12	yes
NCU02654	*gh105-1*	unsaturated rhamnogalacturonyl hydrolase	GH105	no
Side-chain active enzymes				
NCU09170	*gh43-4*	α-L-arabinofuranosidase II	GH43	yes
NCU02343	*gh51-1*	α-L-arabinofuranosidase 2	GH51	yes
NCU09775	*gh54-1*	α-L-arabinofuranosidase	GH54	yes
NCU05965	*gh43-7*	probable endo-arabinanase	GH43	yes
NCU00972	*gh53-1*	arabinogalactan endo-1,4-β-galactosidase	GH53	yes
NCU00642	*gh35-1*	β-galactosidase	GH35	yes
NCU04623	*gh35-2*	β-galactosidase	GH35	yes
NCU00810	*gh2-3*	β-galactosidase	GH2	no
NCU05882	*gh5-5*	endo-β-1,6-galactanase	GH5	yes
NCU00937	*gh79-1*	β-glucuronidase	GH79	yes
NCU09491	*fae-1*	feruloyl esterase B	CE1	yes

The corresponding gene loci and names, annotated protein function, respective CAZy family (Cantarel *et al.*, 2009), and the SignalP prediction (SP) (Petersen *et al.*, 2011) are indicated.

Table 2

The 29-gene set up-regulated on pectin, xylan and cellulose.

Locus	gene symbol	gene product; (putative) function	CAZy family	SP
Carbohydrate esterases				
NCU04870	*ce1-1*	acetyl xylan esterase	CE1	yes
NCU09664	*ce5-3*	acetyl xylan esterase	CE5	yes
NCU09491	*fea-1*	feruloyl esterase B	CE1	yes
NCU10045	*ce8-1*	pectinesterase	CE8	yes
NCU05751	*ce3-2*	cellulose-binding protein	CE3	yes
NCU09416		cellulose-binding GDSL lipase/acylhydrolase	CE16; CBM1	yes
Endo-/exo-acting glycoside hydrolases and polysaccharide lyases				
NCU02855	*gh11-1*	endo-1,4-beta-xylanase A	GH11	yes
NCU05924	*gh10-1*	endo-1,4-beta-xylanase	GH10	yes
NCU08189	*gh10-2*	endo-1,4-beta-xylanase	GH10	yes
NCU05057	*gh7-1*	endoglucanase EG-1	GH7	yes
NCU00762	*gh5-1*	endoglucanase 3	GH5; CBM1	yes
NCU07190	*gh6-3*	exoglucanase 3	GH6	yes
NCU08176	*ply-2*	pectate lyase A	PL3	yes
NCU09924	*gh93-1*	exo-αt-L-1,5-arabinanase	GH93	yes
NCU06143	*gh115-1*	candidate glucuronidase	GH115	yes
Glycoside hydrolases releasing mono- from oligosaccharides				
NCU00709	*gh3-8*	beta-xylosidase	GH3	yes
NCU09923	*gh3-7*	beta-xylosidase	GH3	yes
NCU02343	*gh51-1*	alpha-L-arabinofuranosidase 2	GH51	yes
NCU09775	*gh54-1*	alpha-L-arabinofuranosidase	GH54	yes
NCU04623	*gh35-2*	beta-galactosidase	GH35	yes
NCU00890	*gh2-1*	beta-mannosidase	GH2	no
Others				
NCU03322		GDSL family lipase		no
NCU04475		lipase B		yes
NCU07055		monooxygenase		yes
NCU08384	*xyr-1*	xylose reductase		no
NCU08746		starch/chitin binding domain-containing protein	CBM20	yes
NCU01430		hypothetical protein		no
NCU09415		hypothetical protein		no
NCU09926		hypothetical protein		yes

The corresponding gene loci and names, annotated protein function, respective CAZy family (Cantarel *et al.*, 2009), and the SignalP prediction (SP) (Petersen *et al.*, 2011) are indicated.