Title
A comparative systems analysis of polysaccharide-elicited responses in Neurospora crassa reveals carbon source-specific cellular adaptations

Permalink
https://escholarship.org/uc/item/6mj36935

Journal
Molecular Microbiology, 91(2)

ISSN
0950-382X

Authors
Benz, PJ
Chau, BH
Zheng, D
et al.

Publication Date
2014

DOI
10.1111/mmi.12459

Peer reviewed
A comparative systems analysis of polysaccharide-elicited responses in *Neurospora crassa* reveals carbon source-specific cellular adaptations

J. Philipp Benz,1* Bryant H. Chau,1† Diana Zheng,1 Stefan Bauer,1 N. Louise Glass1,2 and Chris R. Somerville1,2

1Energy Biosciences Institute and 2Department 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 coexpression 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.

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 fulfils a multitude of functions for the plant (Somerville *et al*., 2004; Popper *et al*., 2011), but can also be used as an alternative energy source such as for the fermentation to biofuels (Pauly and Keegstra, 2008; Somerville *et al*., 2010; Chundawat *et al*., 2011; Jordan *et al*., 2012; Youngs and Somerville, 2012). 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 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) (Harris *et al*., 2010; Phillips *et al*., 2011; Beeson *et al*., 2012; Levasseur *et al*., 2013). 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 and 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 and 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 and de Vries, 2011).

Among plant cell wall polysaccharides, pectin has the most structural and functional complexity (for review, see Willats *et al*., 2001; Vincken *et al*., 2003; Mohnen, 2008; Caffall and Mohnen, 2009; Harholt *et al*., 2010). 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 and Mohnen, 2009). It is made of $\alpha$-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 $\rightarrow\alpha$-D-GalA-$1,2\rightarrow\alpha$-L-Rha-$1,4\rightarrow$. 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 $\beta$-linked D-xylose (D-Xyl)-residues. RG-II has a backbone of about eight $\alpha$-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 and 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 acetyl esterases (CE12 and CE13) and rhamnogalacturonan acetyl esterases (CE12)]. The majority of the side-chain active enzymes that work on RG-I substitutions include arabinanases (GH43, GH93), arabinosidases/$\alpha$-L-arabinofuranosidases (GH3, GH43, GH51, GH54 and GH62), galactanases (GH5, GH16, GH30, GH35 and GH53), $\alpha$- and $\beta$-galactosidases (GH1, GH2, GH4, GH27, GH35, GH36, GH42, GH43, GH57, GH97 and GH110), $\beta$-glucuronidases (GH1, GH2 and GH79) and feruloyl esterases (CE1) (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. Lara-Marquez et al., 2011; Kubicek, 2013). 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. de Vries et al., 2002; Paper et al., 2007; Martens-Uzunova and Schaap, 2009; Martinez et al., 2009; Shah et al., 2009; Tian et al., 2009; Tsang et al., 2009; Braaksma et al., 2010; Wymelenberg et al., 2010; Berka et al., 2011; Adav et al.,...
2012; Couturier et al., 2012; Delmas et al., 2012; Hakkinen et al., 2012; Navarrete et al., 2012; Sun et al., 2012), 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 (Mohnen, 2008; Caffall and Mohnen, 2009; Harholt et al., 2010). 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 (Perkins et al., 1976; Perkins and Turner, 1988; Turner et al., 2001; Dunlap et al., 2007). In the laboratory, N. crassa has been developed to study a variety of genetic, molecular, biochemical and genomic phenomena since the 1920s (Davis, 2000; Davis and Perkins, 2002; Borkovich et al., 2004; Dunlap et al., 2007), and recently has been used as a model system for polysaccharide degradation by filamentous fungi (Tian et al., 2009; Sun et al., 2012). 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 (Tian et al., 2009; Coradetti et al., 2012; Sun et al., 2012), 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 secretemes 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 secreteme in more detail, WT culture supernatants were analysed by mass spectrometry (LC-MS/MS; see Experimental procedures) 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 secretemes 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 secretemes showed a significant overlap (e.g. 78% of the pectin secreteome 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 acetylated CE12-1 (NCU09976). The pectin methylsterase 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 endo-arabinanase (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/NCU08176) and a β-glucuronidase (GH79-1/NCU00937; An et al., 1994; Renard et al., 1999).

The difference between the pectin and OPP secretemes 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 LPMOs (GH61-13/PMO-3/NCU07898 and GH61-5/NCU08760), two acetylxyylan 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 secretemes (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 gram (or 1%) of commercially supplied pectin (data not shown). Within the pectin and OPP secretemes, 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 secretemes (Tian et al., 2009; Sun et al., 2012), showed significant differences (Fig. 2C and 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 secreteme), 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 secreteme overlapped with the xylan secreteme, versus only 12 proteins that overlapped with the Avicel secreteme. Of the nine proteins in common between all three secretemes, six are associated with the N. crassa cell wall and/or implicated in cell wall remodelling: NEG-1/NCU04395, NCW-1/NCU05137, Mwg1/NCU05974, ACW-12/NCU08171, ACW-1/NCU08936, and EglC/NCU09175 (Oyama et al., 2006; Maddi et al., 2009). 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 it has very little redundancy within its genome (Galagan et al., 2003; Borkovich et al., 2004; Galagan and Selker, 2004). This is also true for the pectinase complemen, 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 analysed for their growth phenotype on pectin as a sole carbon-source (C-source) (Fig. 3). Three deletion strains (Δgh28-1, Δgh28-2 and the pectin methyltransferase Δce8-1) showed a strong growth phenotype, accumulating only 20–40% of WT biomass when grown on pectin (Fig. 3A and C). They also secreted less protein, and had a reduced rate of pectin consumption (Fig. 3A and 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 and Seymour, 2002) were not able to complement the loss-of-function of the polygalacturonases 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 Experimental procedures). 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 exopolygalacturonase gh28-2 was the most highly expressed pectinase, followed by the pectin methyltransferase 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 exopolygalacturonase 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

---

Table 1. *N. crassa* genes encoding pectinolytic enzymes that are strongly induced by pectin.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene symbol</th>
<th>Gene product; (putative) function</th>
<th>CAZy family</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Backbone-degrading pectinases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU02369</td>
<td>gh28-1</td>
<td>(endo-) polygalacturonase</td>
<td>GH28</td>
<td>yes</td>
</tr>
<tr>
<td>NCU08681</td>
<td>gh28-2</td>
<td>(exo-) polygalacturonase</td>
<td>GH28</td>
<td>yes</td>
</tr>
<tr>
<td>NCU10045</td>
<td>ce8-1</td>
<td>pectin methyltransferase</td>
<td>CE8</td>
<td>yes</td>
</tr>
<tr>
<td>NCU06326</td>
<td>ply-1</td>
<td>pectate lyase 1</td>
<td>PL1</td>
<td>yes</td>
</tr>
<tr>
<td>NCU08176</td>
<td>ply-2</td>
<td>pectate lyase A</td>
<td>PL3</td>
<td>yes</td>
</tr>
<tr>
<td>NCU05598</td>
<td>asd-1</td>
<td>rhamnogalacturonase A</td>
<td>PL4</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09976</td>
<td>ce12-1</td>
<td>rhamnogalacturonan acetyltransferase</td>
<td>CE12</td>
<td>yes</td>
</tr>
<tr>
<td>NCU02654</td>
<td>gh105-1</td>
<td>unsaturated rhamnogalacturonyl hydrolase</td>
<td>GH105</td>
<td>no</td>
</tr>
<tr>
<td>Side-chain active enzymes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU09170</td>
<td>gh43-4</td>
<td>α-L-arabinofuranosidase II</td>
<td>GH43</td>
<td>yes</td>
</tr>
<tr>
<td>NCU02343</td>
<td>gh51-1</td>
<td>α-L-arabinofuranosidase 2</td>
<td>GH51</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09775</td>
<td>gh54-1</td>
<td>α-L-arabinofuranosidase</td>
<td>GH54</td>
<td>yes</td>
</tr>
<tr>
<td>NCU05965</td>
<td>gh43-7</td>
<td>probable endo-arabinanase</td>
<td>GH43</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00972</td>
<td>gh53-1</td>
<td>rhamnogalactan endo-1,4-β-galactosidase</td>
<td>GH53</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00642</td>
<td>gh35-1</td>
<td>β-galactosidase</td>
<td>GH35</td>
<td>yes</td>
</tr>
<tr>
<td>NCU04623</td>
<td>gh35-2</td>
<td>β-galactosidase</td>
<td>GH35</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00810</td>
<td>gh2-3</td>
<td>β-galactosidase</td>
<td>GH2</td>
<td>no</td>
</tr>
<tr>
<td>NCU05882</td>
<td>gh5-5</td>
<td>endo-β-1,6-galactanase</td>
<td>GH5</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00937</td>
<td>gh79-1</td>
<td>β-glucuronidase</td>
<td>GH79</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09491</td>
<td>fae-1</td>
<td>feruloyl esterase B</td>
<td>CE1</td>
<td>yes</td>
</tr>
</tbody>
</table>

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.
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 versus 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 with 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 with 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. 4. Transcriptional induction of pectinolytic genes by pectin and OPP in N. crassa.
A. Comparison of transcriptional response after 4 h exposure of a 16 h culture to pectin or to 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 log10 scale. Genes with expression values $< 1$ FPKM are not shown.
C. Transcript abundances in FPKM (log10 scale) for a selection of genes encoding pectinolytic enzymes found to be significantly upregulated 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 versus 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; $\cap$: intersection between sets.

© 2013 John Wiley & Sons Ltd, Molecular Microbiology
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 $5.04 \times 10^{-7}$; orange colouration 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 co-ordinately 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 homologue grp78 (NCU03982), the protein disulphide isomerase $pdiA$ and $prpA$ homologues NCU09223 and NCU08813, 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 LPMOs were uniquely induced on cellulose ($gh61$-2/NCU07760, $gh61$-4/ $pmo$-2/NCU01050, $gh61$-6/ NCU03328, $gh61$-12/NCU02344, $gh61$-13/ $pmo$-3), three LPMO 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 $= 5.17 \times 10^{-5}$), 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 $htg$-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 methyltransferase $ce8$-1). Additionally, enzymes that cleave oligosaccharides into monosaccharides were also present in this group, including a $\beta$-galactosidase ($gh53$-2), two $\beta$-xylanases ($gh3$-8 and $gh3$-7), one $\beta$-mannosidase ($gh2$-1), and two $\alpha$-L-arabinofuranosidases ($gh51$-1 and $gh54$-1). Intriguingly, $\beta$-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 celloolitanins (Znameroski et al., 2012) as opposed to hemicellulose and pectin signalling, 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 D-Xyl 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 and 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/

Table 2. The 29-gene set upregulated on pectin, xylan and cellulose.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene symbol</th>
<th>Gene product; (putative) function</th>
<th>CAZy family</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbohydrate esterases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU04870</td>
<td>ce1-1</td>
<td>acetyl xylan esterase</td>
<td>CE1</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09664</td>
<td>ce5-3</td>
<td>acetyl xylan esterase</td>
<td>CE5</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09491</td>
<td>fea-1</td>
<td>feruloyl esterase B</td>
<td>CE1</td>
<td>yes</td>
</tr>
<tr>
<td>NCU10045</td>
<td>ce8-1</td>
<td>pectinesterase</td>
<td>CE8</td>
<td>yes</td>
</tr>
<tr>
<td>NCU05751</td>
<td>ce3-2</td>
<td>cellulose-binding protein</td>
<td>CE3</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09416</td>
<td></td>
<td>cellulose-binding GDSL lipase/acylhydrolase</td>
<td>CE16; CBM1</td>
<td>yes</td>
</tr>
<tr>
<td>Endo-/exo-acting glycoside hydrolases and polysaccharide lyases</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU02855</td>
<td>gh11-1</td>
<td>endo-1,4-β-xylanase A</td>
<td>GH11</td>
<td>yes</td>
</tr>
<tr>
<td>NCU05924</td>
<td>gh10-1</td>
<td>endo-1,4-β-xylanase</td>
<td>GH10</td>
<td>yes</td>
</tr>
<tr>
<td>NCU08189</td>
<td>gh10-2</td>
<td>endo-1,4-β-xylanase</td>
<td>GH10</td>
<td>yes</td>
</tr>
<tr>
<td>NCU05057</td>
<td>gh7-1</td>
<td>endoglucanase EG-1</td>
<td>GH7</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00762</td>
<td>gh5-1</td>
<td>endoglucanase 3</td>
<td>GH5; CBM1</td>
<td>yes</td>
</tr>
<tr>
<td>NCU07190</td>
<td>gh6-3</td>
<td>exoglucanase 3</td>
<td>GH6</td>
<td>yes</td>
</tr>
<tr>
<td>NCU08176</td>
<td>ply-2</td>
<td>pectate lyase A</td>
<td>PL3</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09924</td>
<td>gh93-1</td>
<td>exo-α-L-1,5-arabinanase</td>
<td>GH93</td>
<td>yes</td>
</tr>
<tr>
<td>NCU06143</td>
<td>gh115-1</td>
<td>candidate glucuronidase</td>
<td>GH115</td>
<td>yes</td>
</tr>
<tr>
<td>Glycoside hydrolases releasing mono- from oligosaccharides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU00709</td>
<td>gh3-8</td>
<td>β-xylosidase</td>
<td>GH3</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09923</td>
<td>gh3-7</td>
<td>β-xylosidase</td>
<td>GH3</td>
<td>yes</td>
</tr>
<tr>
<td>NCU02343</td>
<td>gh51-1</td>
<td>α-L-arabinofuranosidase 2</td>
<td>GH51</td>
<td>yes</td>
</tr>
<tr>
<td>NCU09775</td>
<td>gh54-1</td>
<td>α-L-arabinofuranosidase</td>
<td>GH54</td>
<td>yes</td>
</tr>
<tr>
<td>NCU04623</td>
<td>gh35-2</td>
<td>β-galactosidase</td>
<td>GH35</td>
<td>yes</td>
</tr>
<tr>
<td>NCU00890</td>
<td>gh2-1</td>
<td>β-mannosidase</td>
<td>GH2</td>
<td>no</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCU03322</td>
<td></td>
<td>GDSL family lipase</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>NCU04475</td>
<td></td>
<td>lipase B</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>NCU07035</td>
<td></td>
<td>monooxygenase</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>NCU08384</td>
<td>xyr-1</td>
<td>xylene reductase</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>NCU08746</td>
<td></td>
<td>starch/chitin binding domain-containing protein</td>
<td>CBM20</td>
<td>yes</td>
</tr>
<tr>
<td>NCU01430</td>
<td></td>
<td>hypothetical protein</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>NCU09415</td>
<td></td>
<td>hypothetical protein</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>NCU09926</td>
<td></td>
<td>hypothetical protein</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

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.

© 2013 John Wiley & Sons Ltd, Molecular Microbiology
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 harbouring 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(arabinono)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 analysed the pheno-
type 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 (Δgh35-1), a mutant of a probable endo-arabinanase (Δgh43-7) and mutants in two genes encoding ‘hypothetical proteins’: ΔNCU07923 and ΔNCU05788 (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 co-ordinate 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 celluloytic 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, 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) (Li et al., 2005; Colot et al., 2006; Leeder et al., 2013). 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; Guillemette et al., 2007; Geysens et al., 2009; 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 frameshift in translation and the addition of a putative activation domain to the C-terminus of the protein (reviewed by: Geysens et al., 2009; Walter and Ron, 2011; Moore and Hollien, 2012; Saloheimo and Pakula, 2012). 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) versus non-spliced (OFF) hac-1 transcript in cultures growing on cellulose as compared with control cultures on sucrose or treated with DTT (Fig. 8B). The total hac-1 transcript was found to be induced about fourfold 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 cellulosytic conditions, the response is more graded as compared with 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-arabinobitol 4-dehydrogenase gene (ard-1) involved in L-Ara catabolism (Fig. S9A; Supporting data set 4; cluster 6) 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 Δlat-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 4 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 Δlat-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

---

**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 frameshift 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 quantification, 16 h 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.
time (Fig. 9B). Analysis of initial transport rates (first 5 min) revealed that uptake of L-Rha and D-Xyl was similar to WT in the Δlat-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 Δlat-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 with 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 analysed 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 a 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 con-
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. Figs 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. crassa 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 methylsterase ceb-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 signalling molecules for induction (see below).

Compared with 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 (Pel et al., 2007; Benoit et al., 2012). For example, N. crassa has two GH28 genes, while A. niger has 22 (Martens-Uzunova and Schaap, 2009; van den Brink and 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-/acyltransferases (van den Brink and de Vries, 2011). The only enzyme classes lacking in N. crassa are an α-hamnosidase (GH78) (Mutter et al., 1994; Yadav et al., 2010), an unsaturated glucuronidic hydrolase (GH88), and a dedicated pectin acetyl esterase (CE12). A more detailed characterization of GH105-1 (a rhamnogalacturonan hydrolase) as well as CE12-1 (a rhamnogalacturonan acetyl esterase) 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.

© 2013 John Wiley & Sons Ltd, Molecular Microbiology
of degradation and represent adaptation to specialized habitats (van den Brink and de Vries, 2011). While these much more complex enzyme systems cannot be modelled 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 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 acetylemesterase; RGL: rhamnogalacturonan lyase.

**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 acetylemesterase; RGL: rhamnogalacturonan lyase.

© 2013 John Wiley & Sons Ltd, *Molecular Microbiology*
et al., 1999; Martens-Uzunova et al., 2006). One or both GH28 enzymes further relies on the pectin methyl-esterase 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 saprobiotic fungi, which has previously been observed only in vitro (Jansen et al., 1945; Dahodwala et al., 1974; Pressey and Avants, 1982; Christgau et al., 1996; Massiot et al., 1997; Benen et al., 1999; Wakabayashi et al., 2003). In plant pathogenic fungi, several polygalacturonase or pectin methyl-esterase enzymes have been demonstrated to be important for growth on pectin and (often) essential for virulence (Schell et al., 1988; Scott-Craig et al., 1998; Isshiki et al., 2001; Valette-Collet et al., 2003; Moran-Diez et al., 2009). 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 (ΔNCU09525) showing markedly increased fitness and one (ΔNCU07923) 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 coexpression 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 analysed 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 and 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 homologues within basidiomycete species and even in bacteria are present (data not shown). A parologue 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 homologues 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 behaviour of hydrolytic genes

Our data demonstrate that the regulation of the pectinolytic machinery is a highly co-ordinated 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 signalling function (as shown, e.g. in Aspergillus; Martens-Uzunova and 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 4 h 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 Martens-Uzunova and Schaap, 2009; de Souza et al., 2011; Hakkinen et al., 2012). 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 4 h. 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 and 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 and 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 ‘flavours’. In the presence of a preferred carbon source, the expression of lignocellulolytic genes in filamentous fungi is subject to carbon catabolite repression (CCR) (Strauss et al., 1995; Ruijter and Visser, 1997; Ebbole, 1998; Flipphi and Felenbok, 2004; Sun and Glass, 2011; Delmas et al., 2012). 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 (Tian et al., 2009; Coradetti et al., 2012; Delmas et al., 2012; Nitsche et al., 2012). 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 signalling 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’ versus 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 hydrolysing 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*

**Neurospora crassa** WT (FGSC #2489) and gene deletion strains used in this study were obtained from the Fungal Genetics Stock Center (FGSC; http://www.fgsc.net). Due to a former mis-annotation, Δ10966 was found to actually be a deletion of two neighbouring genes: the pectate lyase *ply-1* (NCU06326) and NCU06327, encoding a cytochrome P450 protein. The Δgh28-1 strain (ΔNCU02369; FGSC #16219) was only available as a heterokaryon and was therefore backcrossed to WT to obtain homokaryotic strains and to eliminate the background (strains were confirmed to be hygromycin-resistant but sensitive to 200 μg ml⁻¹ glufosinate-ammonium). The strain for *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).

**Neurospora crassa** was grown on 1× 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% OPP (see below), or 2% Avicel (PH 101, Fluka 11365) at 25°C and 200 r.p.m. with constant light unless stated otherwise.

**Preparation of orange peel powder 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 through a 0.22 μm filtering device; 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 ÄKTA-explorer system (GE Healthcare). For this, 1/10 volume of a

© 2013 John Wiley & Sons Ltd, Molecular Microbiology
10× buffer A (final concentration: 50 mM citric acid/NaHPO₄, 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, 10 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 min the samples were diluted with 700 μl 25 mM ammonium bicarbonate (AmBic) and 140 μl methanol. Fifty microlitres of 100 μg ml⁻¹ 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; now Bond Elut OMIX, Agilent Technologies). Finally, the eluted peptide samples were dried in a speed-vac and resuspended in 50 μl water/acetone (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 (Vivascin) and applied to SDS-PAGE (12.5% acrylamide). After staining with GelCode Blue (Thermo Fisher Scientific) 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 acetone/triton (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 4× 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⁻¹ trypsin (Trypsin Gold, Promega) in 25 mM AmBic and digestion allowed to proceed at 37°C for overnight. The peptides were eluted 1× 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/acetone (97:3, v/v, 0.1% formic acid, v/v).

Samples were analysed 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⁻¹ 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 (acetoneitrile/water 95:5, v/v, 0.1% formic acid, v/v) at a nano pump flow rate of 0.6 μl min⁻¹ with a programme 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⁻¹ drying gas, 300°C gas temperature with the following auto MS/MS settings: MS scan range 300–2400 m/z (1 spectrum s⁻¹), MS/MS scan range 59–3000 m/z (1 spectrum s⁻¹), 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 (http://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 1× 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 microfibre 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 Avice 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- to 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 1× Vogel’s salts plus the respective new carbon source (as indicated). The mycelial mass was

© 2013 John Wiley & Sons Ltd, *Molecular Microbiology*
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 a RNA-prep 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 analysed 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/) (Trapnell et al., 2010; Roberts et al., 2011a,b). 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 analysed for N. crassa WT on pectin and duplicate cultures for OPP. 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/hec3.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 centred 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 Quantiﬁcation/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 quantification 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 (1× 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 H2O. The monosaccharide concentrations were then quantified by HPAEC-PAD: a sample size of 25 μl was injected onto a Dionex CarboPac PA20 column (3 x 30 mm guard and 3 x 150 mm analytical) and eluted using an isocratic mobile phase of 18 mM KOH at 0.4 ml min⁻¹ 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 1× 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 ddH2O, centrifuged again to pelatny potentially insoluble debris, and the clear supernatant diluted 10-fold with ddH2O for analysis by HPAEC-PAD (see above). Data were calculated with the assumption of an intracellular volume of 2.4 ml per gram of dry biomass (Slayman and 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 and Bork, 2007). PelA from A. niger (Fig. S10) and GalZp from Saccharomyces cerevisiae (Fig. S11) were used as outgroup respectively.

**Localization of the unconventional intron in hac-1**

Sucrose pre-grown (16 h) WT N. crassa were shifted for 4 h 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 hac1 intron_fwd & hac1 intron_rev; Fig. S12). The PCR localization of the unconventional intron in hac-1

**References**


© 2013 John Wiley & Sons Ltd, *Molecular Microbiology*


Phillips, C.M., Beeson, W.T., Cate, J.H., and Marietta, M.A. (2011) Cellulobiose dehydrogenase and a copper-dependent


**Supporting information**

Additional supporting information may be found in the online version of this article at the publisher’s web-site.