Title

Glucose-Mediated Repression of Plant Biomass Utilization in the White-Rot Fungus Dichomitus squalens.

Journal

Applied and Environmental Microbiology, 85(23)

ISSN

0099-2240

Authors

Daly, Paul
Peng, Mao
Di Falco, Marcos
et al.

Publication Date

2019-12-01

DOI

10.1128/aem.01828-19

Peer reviewed

Glucose-mediated repression of plant biomass utilization in the white-rot fungus *Dichomitus squalens*

Paul Daly[a], Mao Peng[a], Marcos Di Falco[b], Anna Lipzen[c], Mei Wang[c], Vivian Ng[c], Igor V. Grigoriev[c], Adrian Tsang[b], Miia R. Mäkelä[d], Ronald P. de Vries[a, d]#

Affiliations:

[a]Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Utrecht, The Netherlands.
[b]Center for Structural and Functional Genomics, Concordia University, Montreal, Quebec, Canada.
[c]US Department of Energy Joint Genome Institute, Walnut Creek, CA, USA.
[d]Department of Microbiology, University of Helsinki, Helsinki, Finland.

Running title: Glucose-mediated repression in *Dichomitus squalens*

#Address correspondence to Ronald P. de Vries, Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands. tel: + 31 (0)30 2122600, fax: + 31 (0)30 2512097. Email: r.devries@wi.knaw.nl

Keywords: *Dichomitus*, carbon catabolite repression, CAZymes, regulation

Abstract

The extent of carbon catabolite repression (CCR) at a global level is unknown in wood-rotting fungi, which are critical to the carbon cycle and are a source of biotechnological enzymes. CCR occurs in the presence of sufficient concentrations of easily metabolizable carbon sources (*e.g.* glucose), down-regulating the expression of genes encoding enzymes involved in the breakdown of complex carbon sources. We investigated this phenomenon in the white-rot fungus *Dichomitus squalens* using transcriptomics and exo-proteomics. In *D. squalens* cultures, approximately 7% of genes were repressed in the presence of glucose compared to Avicel or xylan alone. The glucose-repressed genes included the essential components for utilization of plant biomass – Carbohydrate Active enZyme (CAZy) and carbon catabolic genes. The majority of polysaccharide degrading CAZy genes were repressed and included activities towards all major carbohydrate polymers present in plant cell walls, while also repression of ligninolytic genes occurred. The transcriptome-level repression of the CAZy genes observed on the Avicel cultures was strongly supported by exo-proteomics. Protease encoding genes were generally not glucose-repressed indicating their likely dominant role in scavenging for nitrogen rather than carbon. The extent of CCR is surprising given that *D. squalens* rarely experiences high free sugar concentrations in its woody environment and indicates that biotechnological use of *D. squalens* for modification of plant biomass would benefit from de-repressed or constitutively CAZymes-expressing strains.

Importance

White-rot fungi are critical to the carbon cycle because they can mineralise all wood components using enzymes that also have biotechnological potential. The occurrence of carbon catabolite repression (CCR) in white-rot fungi is poorly understood. Previously, CCR in wood-rotting fungi has only been demonstrated for a small number of genes. We demonstrated widespread glucose-mediated CCR of plant biomass utilisation in the white-rot fungus *D. squalens*. This indicates that the CCR mechanism has been largely retained even though wood-rotting fungi rarely experience commonly considered CCR conditions in their woody environment. The general lack of repression of genes encoding proteases along with the reduction in secreted CAZymes during CCR suggested that the retention of CCR may be connected with the need to conserve nitrogen use while growing on nitrogen-scarce wood. The widespread repression indicates that de-repressed strains could be beneficial for enzyme production.

Introduction

White-rot fungi break down all wood components, are critical to the carbon cycle and are a source of biotechnologically relevant enzymes (1). However, the extent of carbon catabolite repression (CCR) on plant biomass degradation in wood-rotting fungi is currently unknown. CCR occurs in the presence of sufficient concentrations of easily metabolizable carbon sources (*e.g.* glucose), down-regulating the expression of genes encoding enzymes involved in the breakdown of complex carbon sources. CCR has been studied mainly in ascomycete fungi (2) and has been examined at a

global level in various major groups of fungi, such as recently in anaerobic gut fungi (3), but not in wood-rotting fungi, such as *Dichomitus squalens*.

The analysis of glucose-mediated CCR in wood-rotting fungi is limited to small sets of transcripts and enzymatic activities. In the white-rot fungus *Phanerochaete chrysosporium*, glucose-mediated repression of two cellulase-encoding genes has been reported (4), while in the brown-rot fungus *Postia placenta*, glucose-mediated repression was observed for four hemicellulase, but not for four cellulase-encoding genes (5). In selected *Polyporales* white-rot fungi, cellulase and xylanase activities were significantly lowered when the cultures were supplemented with glucose (6).

Plant biomass utilisation by fungi requires the secretion of extracellular enzymes to degrade the polymeric components, transporters for the sugars released from the plant biomass and intracellular catabolic enzymes. There is evidence for catabolite repression of each of these functional categories in various ascomycete fungi (2) but not in wood-rotting basidiomycete fungi. Genes encoding other enzymes, such as proteases, that can scavenge for the scarce nitrogen in wood by hydrolyzing proteins (7), have been shown to be carbon catabolite repressed in ascomycete fungi (8-10), but there are reports indicating that protease activities in basidiomycete leaf-litter degrading fungi lack CCR (11). A transcriptional regulator of CCR, called CRE1/CreA, has been extensively studied in several ascomycete fungi (12), while other proteins, such as kinases, can mediate CCR independent of CRE1/CreA (13). Recently, *cre1* was deleted in the white-rot fungus *Pleurotus ostreatus* and exo-proteomics detected increased abundance of a subset of Carbohydrate Active enZymes (CAZymes) under apparently non-repressing

conditions (14). Analysis of the extent of glucose-mediated CCR can demonstrate if there is a requirement to delete transcriptional or post-transcriptional regulators of CCR in wood-rotting fungi for improved enzyme yields and/or biomass degradation.

In *D. squalens*, sugar inducers of polysaccharide degrading CAZyme encoding genes have been identified (15) demonstrating an inducible, *i.e.* not constitutive, system for these enzymes, which provides it with an ability to partially tailor its molecular responses to wood composition (16). *D. squalens* has biotechnological potential as a source of enzymes and for bioremediation and can be genetically modified (17), making it a highly suitable species to analyse the effects of CCR. The global effect of high glucose concentrations on the utilisation of plant biomass components is not known in any wood rotting fungus. Occurrence of glucose-mediated repression could be a tool to indicate whether *e.g.* particular CAZyme families or expansin-like domain-containing proteins in wood-rotting fungi, are involved in plant biomass degradation. In this study, we investigated the occurrence and extent of glucose-mediated repression in the white-rot basidiomycete *D. squalens* using two CAZyme-inducing polysaccharide substrates and time-points.

Results

Identification of carbon catabolite repressing conditions in *D. squalens*

Whether excess glucose could repress the secreted protein pattern from *D. squalens* cultures growing on ring-plates containing either Avicel or xylan was examined (Fig. 1A). The secreted proteins after 2 d, visualized using SDS-PAGE, showed a clear repression pattern on each substrate, albeit stronger on Avicel, and were analysed using exo-proteomics (Fig. 1B). Hierarchical clustering of the exo-proteins showed that the biological replicates from a condition generally clustered together (Fig. S1A). To investigate whether this repression extended beyond secreted proteins, the transcriptomes after 2 d as well as 5 h were analysed. Hierarchical clustering of all expressed genes showed two major time-dependent clusters (Fig. S1B). At the earlier time-point, the genes clustered by polysaccharide substrate, and presence and absence of glucose. In contrast at 2 d, transcripts from the Avicel-only cultures clustered separately compared to the other cultures from 2 d. Time, and not the type of polysaccharide or the presence of glucose, appeared to be the most dominant factor in the clustering, suggesting that gene induction and/or repression changed substantially over time.

Extensive glucose-mediated repression was observed at a global level

Glucose-mediated repression of *D. squalens* genes was extensive with approximately 7% of genes (1,042/15,295) repressed in one or more conditions (Table S1). A gene was considered glucose-repressed if the expression was > two-fold lower in the presence of glucose, the $P_{adj} < 0.05$ and the FPKM > 10 in the absence of glucose. A

large majority of the glucose-repressed genes were only repressed in one condition (Fig. 2A). Not all genes were expressed in all these conditions, and a lack of induction on one of the polysaccharides, preventing detection of repression, could explain why approximately a quarter of the genes were found repressed on only one polysaccharide (Fig. 2B). The remainder of the genes repressed on only one polysaccharide were expressed on both, suggesting differences in how glucose-mediated repression was functioning on each polysaccharide. On all four of the repressed conditions, there was enrichment of GO terms representing processes directly related to plant biomass utilisation, but also other processes (Fig. 2C and Table S2). The GO term "hydrolase activity, hydrolyzing O-glycosyl compounds" was enriched in glucose-repressed genes on both polysaccharides supplemented with glucose compared to the polysaccharides without glucose. In addition, "sugar transmembrane transporter activity" was enriched in the genes repressed on Avicel supplemented with glucose compared to Avicel alone. Transport of the sugars released from the polysaccharides is critical to growth on plant biomass and enrichment in the glucose-repressed genes of GO terms for degradation as well as transport demonstrates that repression of these two processes is coordinated. The GO terms enriched among the genes repressed on Avicel supplemented with glucose at 2 d suggested repression of protein production from the GO terms "cellular amino acid biosynthetic process" and "tRNA aminoacylation for protein translation". Other GO terms were enriched in repressed genes from a subset of the conditions and displayed repression of biological functions beyond those directly related to plant biomass degradation (Table S2). These included GO terms for lipase activity repressed on

xylan at 5 h, alcohol metabolic process on xylan at 2 d and organic acid biosynthetic process on Avicel at 2 d.

Glucose repressed a broad range of lignocellulose degrading activities

The *D. squalens* genome encodes 233 putative plant biomass degrading (PBD) CAZy genes and almost half (108) of these were glucose-repressed in at least one condition. The majority of the repressed PBD CAZy genes were repressed on only one of the two polysaccharides (Fig. 2B). This was due to either a lack of induction by released sugars on the other polysaccharide or differences in how the repression was being regulated, as more than half of these PBD CAZy genes repressed on only one polysaccharide were expressed substantially on both substrates without glucose. The glucose-repressed genes on Avicel were less affected by the time-point compared to xylan, where 2.5 times more CAZy genes were repressed on xylan at 2 d compared to 5 h. In addition to genes encoding CAZymes acting on polysaccharides, half (5/10) of the manganese peroxidase (MnP) encoding genes in *D. squalens* were repressed on xylan at 2 d. The total CAZy gene expression was five-fold higher on Avicel than on xylan at 2 d, and this likely contributed to approximately six-fold repression on Avicel compared to less than two-fold repression detected on xylan (Fig. S2). Hierarchical clustering also illustrated the higher CAZy gene expression on Avicel (Fig. 3 and Fig. S3). The most striking example of glucose-mediated repression was in Cluster 3, which contained mainly genes encoding cellulose-acting CAZymes that were highly expressed on Avicel at 2 d and strongly repressed with average 100-fold repression by glucose (Fig. 3). On xylan, the repressed genes mainly encoded xylan-acting enzymes that were side-chain rather than backbone acting, although in the exo-proteome, there

was one endo-xylanase protein lower produced in the presence of glucose at 2 d (Table S3). From the transcriptome analysis, it was clear that genes encoding a broad-range of activities were repressed including hemicellulolytic and pectinolytic activities.

Exo-proteomics identified 57 PBD CAZymes in at least one of the 2 d conditions (Table S3). Strikingly, the five most abundant proteins secreted into the Avicel ring-plates made up almost 80% of the total protein abundance, whereas in the ring-plates where Avicel was supplemented with glucose, these same proteins only made up approximately 10% of the total protein abundance. A protein was considered significantly lower produced in the presence of glucose if the abundance was > two-fold lower and the $P < 0.05$ or alternatively if only identified in the absence of glucose. There was a better positive correlation for Avicel than xylan when the proteins that were significantly lower produced in the presence of glucose were compared to the repressed genes. On Avicel at 2 d, almost all of the CAZymes that were lower produced in the presence of glucose also had the encoding gene repressed (18/19). In contrast, less than half of the CAZymes that were lower produced in the presence of glucose on xylan at 2 d had the encoding gene repressed (11/24).

Repression of catabolism of plant-derived sugars

Sugars are catabolised via several carbon catabolic pathways and intermediate metabolites can be inducers of genes encoding catabolic enzymes and CAZymes (18). Putative carbon catabolic enzymes in *D. squalens* were identified by homology to characterized enzymes from ascomycete fungi. Genes encoding 21 putative catabolic enzymes were glucose-repressed and half of these are involved in catabolism of

198 hemicellulose or pectin sugars, such as pentoses, mannose or galacturonic acid (Table
199 S1). Only a subset of the expressed genes from catabolic pathways related to
200 hemicellulose or pectin derived sugars (pentose catabolic pathway, Leloir pathway,
201 D-mannose pathway, D-galacturonic acid pathway and L-rhamnose pathway) were
202 glucose repressed at a time-point on Avicel or xylan. Repression of the entry point to
203 these pathways could repress the flux through the pathway obviating the need to
204 repress genes encoding subsequent enzymes, but the putative entry point to these
205 pathways was not preferentially repressed. Some carbon catabolic genes were higher
206 expressed in the presence of glucose, but these were not shared across the repressed
207 conditions. Glycolysis and the tricarboxylic acid (TCA) cycle catabolise glucose or its
208 derivatives in the main energy generating reactions of the cell. Surprisingly, few
209 glycolytic or TCA cycle genes were higher expressed in the presence of high glucose
210 concentrations and none at 5 h. In *Aspergillus oryzae*, many of the glycolytic and
211 TCA cycle genes are induced in glucose-rich compared to glucose-depleted
212 conditions (19). The general trend for glycolytic or TCA cycle genes was to be neither
213 induced nor repressed in the presence of excess glucose, but instead to remain
214 constitutively expressed. The same trend was observed for genes from the pentose
215 phosphate pathway (PPP), that catabolises glucose in parallel to glycolysis. This was
216 surprising because there could be a greater requirement for the ribose 5-phosphate
217 generated by the PPP for use in synthesising nucleotides in DNA replication if
218 replication was occurring more frequently at higher growth rates such as in the
219 presence of excess glucose. In contrast to the repression of *D. squalens* CAZy genes,
220 the magnitude of the repression was generally not greater than two-fold (Fig. S4) and

the proportion of the carbon catabolic enzymes encoding genes that were repressed (27%) was smaller than for CAZy genes (46%).

Limited glucose-mediated repression of proteases

Proteases function in scavenging for nitrogen during fungal wood degradation as the nitrogen content of wood is very low (7). The GO term for "serine-type peptidase activity" was enriched in the genes that were glucose-repressed on either of the polysaccharides at 2 d (Fig. 2C). However, only a small proportion of protease encoding genes (23/215) were repressed in at least one condition, mainly at the later time-point, while the expression of three-quarters of protease encoding genes was not significantly different (Table S1). This was reflected in the total abundance of transcripts encoding for proteases where the levels were not lower in the presence of excess glucose (Fig. 4). Exo-proteomics identified 26 proteases in at least one of the 2 d conditions (Table 1). There were only three proteases lower produced in the presence of glucose in the cultures growing on each polysaccharide whereas the majority of the proteases were not significantly lower produced in the presence of glucose. Indeed, the proteases were the other major functional category of secreted proteins, besides those degrading polysaccharides, secreted into the ring-plates in all conditions.

Candidate transcriptional regulator of glucose repression

Direct transcriptional repression via transcription factors binding to the promoters of the glucose-repressed genes likely plays a role in *D. squalens*. The putative *D. squalens cre1* ortholog, identified as the reciprocal BLASTP match of *P. ostreatus cre1* (14), was glucose-repressed on Avicel at 2 d, but was not induced by the high

glucose concentrations in the studied cultures. An alignment that included characterised CRE1/CreA proteins and putative CRE1 orthologs from three basidiomycetes, including *D. squalens*, showed conservation of the zinc binuclear cluster (Fig. S5). Surprisingly, there was little conservation between the ascomycete and basidiomycete proteins for the remainder of the sequence suggesting that how the transcription factor functions could differ substantially between the phyla. However, there was substantial conservation within each phyla for the proteins for the reminder of the sequence suggesting conservation of function within the phyla. The conservation of the zinc binuclear cluster supported analysis of the presence of the binding motif of CRE1/CreA that has been identified in ascomycetes (20, 21) in the promoters of the *D. squalens* glucose-repressed genes. There was a trend for enrichment (~ 30%) of motifs in glucose-repressed *D. squalens* genes, as measured by the mean number of motifs per gene promoter (Table 2 and Table S4).

Repression of candidate proteins with potential roles in biomass degradation

A gene being subject to glucose-mediated carbon catabolite repression when cultured with Avicel or xylan makes the encoded protein a stronger candidate for involvement in plant biomass degradation. The previous section showed how genes encoding proteins with well-established involvement in plant biomass degradation, such as cellulases, were strongly glucose-repressed thus allowing inferences to be made about other repressed genes.

β-glucuronidases from CAZy family GH79 cleave β-linked glucuronic acid residues in arabinogalactan-proteins (22), but these proteoglycans are a minor component of most lignocellulosic residues and thus a role for β-glucuronidases in

degradation of lignocellulose-rich residues is controversial. Recently, it was shown that β-glucuronidases can have a role in cellulose degradation, as polysaccharide monoxygenases catalyse oxidation of cellulose to glucuronic acid-containing oligosaccharides (23). Three *D. squalens* GH79 genes, two of which clustered closely with characterised β-glucuronidases from *Aspergillus niger* and *Neurospora crassa* (22) in phylogenetic analysis (Fig. S6), were repressed by glucose, further supporting a possible role for these proteins in lignocellulose degradation. The proteins from these three GH79s were detected with two of the three proteins lower produced in the presence of glucose (Table S3).

Previously, genes encoding expansin-like proteins were expressed when *D. squalens* was cultured with plant biomass substrates suggesting a role in the degradation (24). Expansin-like proteins such as the *Trichoderma reesei* SWO1 can disrupt the cellulose structure (25) and can boost the sugar release in saccharification of a non-pretreated grass substrate (26). In *D. squalens*, there was limited repression of genes encoding expansin-like proteins and their total expression was similar in absence or presence of glucose (Table S1). The only expansin-like protein detected in the exo-proteins was not lower produced in the presence of glucose (Table S3), but that does not exclude their role in plant biomass degradation, just indicates a different regulatory control of these genes.

Apart from plant biomass degradation, the gene encoding a *D. squalens* protein, FIP-dsq2 (Dicsqu464_1_PID_921554), which was reported to have immunomodulatory and anti-cancer properties (27), was glucose-repressed on Avicel at 5 h and 2 d as well as xylan at 2 d.

Discussion

We have demonstrated using transcriptomics and exo-proteomics widespread glucose-mediated repression in a wood-rotting basidiomycete fungus. Although high concentrations of free sugars are naturally rare for a wood-rotting fungus, they were insightful in suggesting that the existence of CCR in *D. squalens* was partly to conserve use of the scarce nitrogen in its woody biotope. There was widespread glucose-mediated repression of CAZy genes but not genes encoding nitrogen-scavenging proteases. Also, the pattern of repression provided support for direct and indirect mechanisms for how *D. squalens* reduced the expression of its CAZy genes.

The strong repression of secreted CAZy genes, but not genes encoding proteases, by glucose supplementation, suggests that CCR has been maintained by *D. squalens* to strongly conserve nitrogen use in its nitrogen-scarce woody biotope. It is possible that the sensing of increasing sugar concentrations signals successful breakdown of plant biomass and less of a need for high CAZyme secretion levels, which could be a drain on the limited nitrogen supply. Instead the scarce nitrogen could be diverted to support fungal growth and achieve a lower mycelial C:N balance than would be possible if a higher protein secretion level was maintained. It did appear that there was more growth in the presence of excess glucose as the mycelium was substantially thicker than when cultured on either polysaccharide in the absence of excess glucose. Previously, in three leaf-litter degrading basidiomycetes, the total protease activity was not lowered by glucose supplementation (11). The lack of repression in *D. squalens* is partly in contrast to the CCR of proteases that can be found in ascomycete fungi (10, 28, 29) possibly due to differences in the biotope or

313 lifestyle of the fungi. Regulation of proteases is a complex phenomenon and in
314 particular the low nitrogen concentration in the culture medium, intended to replicate
315 environmental conditions, may be a competing regulatory factor with any repressing
316 effect of the glucose due to a physiological requirement of *D. squalens* to scavenge
317 for nitrogen from proteins.

318 There were interesting temporal patterns to the induction and repression on
319 both Avicel and xylan. On Avicel there was higher CAZyme encoding transcript
320 levels at 2 d compared to 5 h and a three-time larger average fold-change of
321 repression at the later time-point. The higher transcript levels at 2 d could be
322 explained by the induction taking longer than 5 h to peak but the possibility cannot be
323 excluded that the low concentrations of glucose released from Avicel at 5 h were
324 sufficient to repress. On xylan at 2 d, the correlation between the CAZymes lower
325 produced in the presence of glucose and repressed genes was ~ 50%, whereas on
326 Avicel the correlation was ~ 95%. The lack of repression of the genes encoding for
327 half of the CAZymes lower produced on xylan in the presence of glucose suggested
328 that the repression of the genes occurred earlier than two days and had now ceased.
329 Secondary (indirect) effects could also be contributing to the repression of these
330 genes.

331 Whether the repression effect is primary (direct) or secondary (indirect) is
332 difficult to ascertain. The presence of a *D. squalens* CRE1 ortholog and an enrichment
333 of the CRE1/CreA binding motif in the promoters of the repressed genes suggest that
334 there was a direct repression effect at the transcript level. In addition, at least two
335 indirect repression effects were also suggested by the datasets. The enriched GO
336 terms for “tRNA aminoacylation for protein translation” and “cellular amino acid

biosynthetic process" in the Avicel cultures supplemented with glucose suggest an indirect effect leading to less exo-proteins due to less protein synthesis and translation in the presence of glucose. The genes encoding β-glucosidases were glucose-repressed mainly at 5 h on Avicel and the three β-glucosidase proteins detected were not found to be lower produced in the presence of glucose. These non-repressed β-glucosidases may indirectly contribute to an apparent repressing effect by reducing the induction by means of hydrolysis of cellobiose, which was previously found to be a major inducer of *D. squalens* CAZymes (15).

Analysis of glucose repression in orthologs of *D. squalens* glucose-repressed genes of the white-rot fungus *P. chrysosporium* (4) and the brown-rot fungus *P. placenta* (5) revealed similar and contrasting repressing trends. The *D. squalens* orthologs of the cellobiohydrolase *Pc_cel6A* and cellobiose dehydrogenase *Pc_cdh* were also glucose-repressed on Avicel at both time points, while the β-glucosidase *Pc_bgl3A* was not repressed in *P. chrysosporium*, but the *D. squalens* ortholog was repressed on Avicel at both time-points (see Table S1 for the IDs of the *D. squalens* orthologs). The 30-fold lower concentration of glucose used by H. Suzuki et al. (4) may be a factor in the lack of repression compared to *D. squalens*, as well as the lack of an inducing carbon source in their study. In *P. placenta*, four cellulase genes analysed by quantitative PCR were not found to be repressed by glucose compared to a control without a carbon source (5), while in our study *D. squalens* orthologs for three of these genes (β-glucosidase *Pp_bgl1*, and two endoglucanases *Pp_cel5B* and *Pp_cel12A*) were repressed by glucose on Avicel at both time-points. Four *P. placenta* hemicellulose encoding genes were repressed by glucose (5) and the *D. squalens* orthologs of these were also repressed (β-xylosidase *Pp_bxl1*, β-mannanase

Pp_man5A and endo-xylanases *Pp_xyn10A-1* and *Pp_xyn10A-2*). Both of these studies (4, 5) consider glucose repression compared to a control without a carbon source whereas in our study, the control is an inducing polysaccharide.

It is somewhat surprising that high glucose concentrations had a widespread physiological effect in a wood-rotting fungus that would rarely experience such high concentrations in its woody environment, as recalcitrant lignocellulose degradation is a slow process occurring over several years (30). Glucose could potentially be sensed from catabolic intermediates that are also intermediates from catabolism of other sugars. *D. squalens* is predominantly found on softwoods (31) where galactoglucomannan, containing a backbone of glucose and mannose, is the main hemicellulose. In ascomycetes, mannose can be catabolised in two steps to D-fructose-6-phosphate that is the same intermediate that glucose is catabolised to (18). Thus D-fructose-6-phosphate accumulation could be the signal for excess mannose release more readily than other plant biomass derived sugars such as xylose which are catabolised by more complex pathways. There is evidence for cross-talk between cellulose and mannan sensing pathways in *N. crassa* (32), although the levels of other sugars besides glucose, such as mannose, are also likely to be released slowly from the wood.

Testing lower glucose concentrations and investigating the repressing effect of other sugars was beyond the scope of this study. Also, investigating the repressing effect of high nitrogen concentrations on protease expression and production would be of interest as our study used media with a low nitrogen concentration. The experimental conditions developed here facilitate screening based on protein secretion pattern or candidate CAZymes before more detailed transcriptome analysis. The

conditions also facilitate analysis of deletions of candidate regulatory proteins, such as the *D. squalens cre1* ortholog. Moreover, other *D. squalens* strains could be tested, as differences in CAZy gene induction between *D. squalens* monokaryotic and dikaryotic strains was reported previously (15).

In conclusion, for applications, de-repressed strains would be of benefit for a higher yield of CAZymes from *D. squalens*, particularly for the strongly repressed genes encoding cellobiohydrolases. Wood-rotting fungi seem to maintain the CCR physiological response found in other major fungal lineages despite the rarity of high free sugar concentrations in their natural biotope. The scarcity of nitrogen in wood may be a contributing factor to the maintenance of CCR.

Materials and Methods

Culturing conditions and biomolecule harvesting and extraction

D. squalens monokaryon CBS464.89 was maintained on malt extract agar (MEA) at 4°C and at -80°C in 30% glycerol. For pre-cultures, 0.5 cm plugs from a culture of *D. squalens* growing on MEA, was inoculated onto a 7.6 cm diameter perforated (0.1 µM) polycarbonate membrane (Maine Manufacturing) on top of low-nitrogen asparagine-succinate (LN-AS) agar (1.5% w/v) adjusted to pH 4.5 (33) containing either 1% w/v Avicel cellulose (Fluka) or beech wood xylan (Sigma) and incubated at 28°C for 3.5 d, after which the colony diameter was ~4.5 cm. The polycarbonate membrane containing the colony was then transferred to 9 cm ring-plates (34) containing LN-AS liquid medium containing the same polysaccharide (1% w/v) as the pre-culture plate or that polysaccharide supplemented with 3% w/v D-glucose. The rings in ring-plate are numbered from one to five from the centre to the outermost ring (see (34) for image with numbering of the rings). The well of each ring has a depth and width of 0.5 cm and is separated from adjacent rings by a barrier of 0.1 cm. After 5 h or 2 d from the transfer, the mycelia growing above rings two to five were sampled for RNA extraction. Also, 2 d after transfer, the liquid in rings two to five were pooled (for SDS-PAGE and exo-proteomics analysis), flash frozen in liquid nitrogen and stored at -20°C. LN-AS medium or agar containing Avicel or beech wood xylan was autoclaved at 121°C for 15 min, and the filter-sterilised vitamin solution was added after autoclaving. A 20% w/v glucose stock solution was sterilised using a membrane filter.

Total RNA was extracted from mycelium, purified and checked for quality as described previously (35). The mycelium was ground and mixed with Trizol reagent (Invitrogen) after harvesting and stored at -80°C before proceeding with remaining steps of RNA extraction.

Transcriptomic sequencing and analysis

Plate-based RNA sample preparation was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using the Illumina TruSeq Stranded mRNA HT sample prep kit utilizing poly-A selection of mRNA following the protocol outlined by Illumina: https://support.illumina.com/sequencing/sequencing_kits/truseq-stranded-mrna.html, and with the following conditions: total RNA starting material was 1 µg per sample and eight cycles of PCR was used for library amplification. The prepared libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then multiplexed with other libraries, and the pool of libraries was then prepared for sequencing on the Illumina NovaSeq sequencer using NovaSeq XP V1 reagent kits, S4 flow cell, and following a 2x150 indexed run recipe.

Using BBDuk (https://sourceforge.net/projects/bbmap), raw reads were evaluated for artifact sequence by kmer matching (kmer = 25), allowing one mismatch and detected artifact was trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads and reads containing any Ns were removed. Quality trimming was performed using the phred trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length 25 bases

or one third of the original read length – whichever was longer). Filtered reads from each library were aligned to the Dicsqu464_1 genome assembly (36) using HISAT2 version 2.1.0 (37). FeatureCounts (38) was used to generate the raw gene counts using gff3 annotations. Only primary hits assigned to the reverse strand were included in the raw gene counts (-s 2 -p --primary options).

Statistical analysis was performed using DESeq2 (39). Hierarchical clustering on the expressed PBD CAZy genes using the log2 FPKM values (+1) was performed in the R statistical environment using the gplots package with the Euclidian distance and complete linkage options selected. Transcripts were considered differentially expressed if the DESeq2 fold change was > 2 or < 0.5 and $P_{adj} < 0.05$ as well as the FPKM > 10 in at least one of the two conditions being compared. Transcripts with FPKM ≤ 10 were considered lowly (*i.e.* not substantially) expressed. The plant biomass degrading CAZy annotations (40) were the same as used previously (16). For protease annotations, the Merops (41) annotations from the Dicsqu464_1 JGI portal were used. The annotations for carbon catabolic enzymes in *D. squalens* were obtained by reciprocal BLASTP using characterized carbon catabolic enzymes from *Aspergillus* species. Gene ontology (GO) enrichment analysis was performed using BiNGO (42) using the default setting with "GO_full" selected for the ontology.

SDS-PAGE and protein LC-MS/MS sample preparation and analysis

The liquid from rings two to five from each ring-plate from the 2 d cultures was pooled before centrifuging at 3,200 x *g* for 1 h at 4°C to pellet any solid particles from the polysaccharide substrates. From Avicel-containing cultures, 4 mL was concentrated ~20-fold using Vivaspin 500 columns (5000 kDa molecular weight cut-

463 off, GE Life Sciences) by centrifuging at 15,000 x *g*, 4°C for 3 h. All of the liquid
464 from the concentrated supernatants were then precipitated. For the xylan-containing
465 cultures, 500 µL of the non-concentrated supernatants was precipitated.

466 Proteins were precipitated on ice for 1 h using twice the sample volume of a
467 solution of 20% trichloroacetic acid, 20 mM DTT and 80% acetone, then centrifuged
468 at 3,200 x *g* for 30 min at 4°C. The pellet was then mixed with a solution of 20 mM
469 DTT and 80% acetone and incubated overnight at -20°C, then centrifuged at 3,200 x g
470 for 30 min at 4°C. The air-dried pellet was re-suspended in 150 µL 0.25% w/v anionic
471 acid labile surfactant (AALS I) (Protea Biosciences) solution (prepared in 200 mM
472 ammonium bicarbonate pH 7.8).

473 From the protein samples re-suspended in the AALS solution, 7.5 µL was
474 analysed by SDS-PAGE. A 4X loading buffer (0.1 M Tris-HCl, pH 6.8, 42% glycerol,
475 4% w/v SDS, 0.02% w/v bromophenol blue and 0.6 M β-mercaptoethanol) was used
476 where samples were boiled for 2 min to denature the proteins, cooled on ice for 2 min
477 and centrifuged at ~10,000 × *g* for 2 min to remove insoluble material. The proteins
478 were separated using a 12% w/v acrylamide SDS-containing running gel along with a
479 PageRuler Plus Prestained Protein Ladder with a 10 to 250 kDa size range (Thermo
480 Fisher). The gels were silver-stained based on standard methods.

481 The objective of the proteomic analysis was to compare equivalent volumes of
482 the culture supernatant. For the six samples from xylan-containing cultures, the same
483 volume of each sample with digested with trypsin followed by analysing the same
484 volume of cleaned-up digests by LC-MS/MS. For the samples from Avicel-containing
485 cultures, a three times larger volume was digested from the Avicel cultures

supplemented with glucose followed by analysing the same volume of cleaned-up digests by LC-MS/MS. Protein samples were digested with trypsin for proteomic analysis as previously described (43). Dried peptide digest samples were solubilized in a solution of 5% acetonitrile, 0.1% formic acid and 4 fmol/µL of trypsin-digested bovine serum albumin (BSA) (Michrom Bioresources) used as internal standard. Five µL from all twelve cleaned-up digest samples was analyzed by LC-MS/MS using an Easy-LC II Nano-HPLC system connected in-line with a Velos LTQ-Orbitrap mass spectrometer (Thermo Fisher). LC-MS/MS data peptide and protein identification was done using the *D. squalens* protein sequence databases obtained from the Joint Genome Institute Dicsqu464_1 generated gene models (36). Protein identification and quantification was performed using the Proteome Discoverer 2.2 (Thermo Fisher) precursor ion quantitation workflow. Normalized individual protein area values were expressed as a fold value of the protein area value determined for the BSA internal standard.

The values for the samples from Avicel cultures supplemented with glucose, where three times larger volume was used for the trypsin digests, were divided by three. This correction for dilution was suitable as the amounts of proteins analysed were within the linear range of detection as validated by comparing by LC-MS/MS three-fold dilutions of two of the samples. A minimum of two peptides matched to a protein with at least one peptide of unique sequence was considered sufficient for identification in the dataset. For a protein to be considered present in a condition, the requirements were an abundance measurement in at least two of the replicates for that condition. For an identified protein to be considered significantly higher or lower produced, the requirements were > 2-fold difference in mean abundance values and P

510 < 0.05 from a two-tailed heteroscedastic (assuming unequal variances) t-test of the
511 log2 transformed abundance values. Where it was not possible to calculate a P-value,
512 a protein was considered as only present in one of the two conditions being compared
513 if there was abundance measurements in ≥ 2 of the replicates for the condition where
514 the protein was present and an abundance measurement in ≤ 1 replicate for the other
515 condition. The remaining proteins in the dataset were categorised as not significantly
516 different or not present in that comparison.

517 **Identification and alignment of CRE1/CreA protein sequences**

518 The *P. ostreatus* CRE1 (14) sequence was used to identify putative CRE1/CreA
519 orthologs by reciprocal BLASTp at MycoCosm in *D. squalens* (Dicsqu464_1), *P.*
520 *chrysosporium* (Phchr1) and *P. placenta* (PosplRSB12_1). The protein sequences of
521 the putative CRE1/CreA orthologs and the characterized CRE1/CreA from *P.*
522 *ostreatus* (14), *Fusarium oxysporum* (44), *Trichoderma reesei* (45), *N. crassa* (46)
523 and *Aspergillus nidulans* (47) were aligned using Clustal Omega (48) with default
524 parameters and visualized with Jalview (49). The conserved domain database (50)
525 was used to identify the zinc binuclear cluster domains which were annotated on the
526 alignment (Fig. S5).

527 **Promoter analysis for CRE1/CreA binding motifs**

528 Up to 1,000 bp length of promoter sequence upstream of the coding region of the
529 genes were obtained from the Dicsqu464_1 genome annotation (36). The reported
530 CreA binding motif 5'-SYGGRG-3', and its sub-motifs 5'-[GC][CT]GGGG-3' and
531 5'-[GC][CT]GGAG-3' were searched to both strands of promoter sequences as a
532 previous study (45) using an in-house Perl script.

Accession numbers

The reads from each of the RNAseq samples were deposited with the Sequence Read Archive at NCBI with individual sample accession numbers (SRP215076, 215080-81, 215085, 215089, 215091-92, 215095, 215099, 215112, 215118-19, 215122-23, 215127, 215138, 215142, 215150-55 and 215157). The mass spectrometry data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (51) partner repository with the dataset identifier PXD014774 and 10.6019/PXD014774.

Declarations

Competing interests

None declared.

Acknowledgements

PD was supported by a grant of the Netherlands Scientific Organization NWO 824.15.023 to RPdV. The Academy of Finland grant no. 308284 to MRM is acknowledged. The work conducted by the U.S. Department of Energy Joint Genome Institute (JGI), was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The proteomics work was supported by Genome Canada and Genome Quebec.

References

1. Mäkelä M, Hildén K, Vries RP. 2014. Degradation and modification of plant biomass by fungi, p 175-208. *In* Nowrousian M (ed), Fungal Genomics. Springer, Cham, Switzerland.
2. Adnan M, Zheng W, Islam W, Arif M, Abubakar Y, Wang Z, Lu G. 2018. Carbon catabolite repression in filamentous fungi. Int J Mol Sci 19:48.
3. Henske JK, Gilmore SP, Haitjema CH, Solomon KV, O'Malley MA. 2018. Biomass-degrading enzymes are catabolite repressed in anaerobic gut fungi. AICHE J 64:4263-4270.
4. Suzuki H, Igarashi K, Samejima M. 2008. Real-time quantitative analysis of carbon catabolite derepression of cellulolytic genes expressed in the basidiomycete *Phanerochaete chrysosporium*. Appl Microbiol Biotechnol 80:99-106.
5. Zhang J, Schilling JS. 2017. Role of carbon source in the shift from oxidative to hydrolytic wood decomposition by *Postia placenta*. Fungal Genet Biol 106:1-8.
6. Kobakhidze A, Asatiani M, Kachlishvili E, Elisashvili V. 2016. Induction and catabolite repression of cellulase and xylanase synthesis in the selected white-rot basidiomycetes. Ann Agrar Sci 14:169-176.
7. Cowling EB, Merrill W. 1966. Nitrogen in wood and its role in wood deterioration. Can J Bot 44:1539-1554.
8. Katz ME, Bernardo SM, Cheetham BF. 2008. The interaction of induction, repression and starvation in the regulation of extracellular proteases in *Aspergillus nidulans*: evidence for a role for CreA in the response to carbon starvation. Curr Genet 54:47-55.
9. van den Hombergh JP, van de Vondervoort PJ, Fraissinet-Tachet L, Visser J. 1997. *Aspergillus* as a host for heterologous protein production: the problem of proteases. Trends Biotechnol 15:256-63.
10. Snyman C, Theron LW, Divol B. 2019. Understanding the regulation of extracellular protease gene expression in fungi: a key step towards their biotechnological applications. Appl Microbiol Biotechnol 103:5517-5532.
11. Kalisz HM, Wood DA, Moore D. 1987. Production, regulation and release of extracellular proteinase activity in basidiomycete fungi. Trans Br Mycol Soc 88:221-227.
12. Benocci T, Aguilar-Pontes MV, Zhou M, Seiboth B, de Vries RP. 2017. Regulators of plant biomass degradation in ascomycetous fungi. Biotechnol Biofuels 10:152.
13. Kunitake E, Li Y, Uchida R, Nohara T, Asano K, Hattori A, Kimura T, Kanamaru K, Kimura M, Kobayashi T. 2019. CreA-independent carbon catabolite repression of cellulase genes by trimeric G-protein and protein kinase A in *Aspergillus nidulans*. Curr Genet 65:941-952.
14. Yoav S, Salame TM, Feldman D, Levinson D, Ioelovich M, Morag E, Yarden O, Bayer EA, Hadar Y. 2018. Effects of *cre1* modification in the white-rot

fungus *Pleurotus ostreatus PC9*: altering substrate preference during biological pretreatment. Biotechnol Biofuels 11:212.

15. Casado López S, Peng M, Issak TY, Daly P, de Vries RP, Mäkelä M. 2018. Induction of genes encoding plant cell wall-degrading carbohydrate-active enzymes by lignocellulose-derived monosaccharides and cellobiose in the white-rot fungus *Dichomitus squalens*. Appl Environ Microbiol 84:e00403-18.
16. Daly P, Casado López S, Peng M, Lancefield CS, Purvine SO, Kim Y-M, Zink EM, Dohnalkova A, Singan VR, Lipzen A, Dilworth D, Wang M, Ng V, Robinson E, Orr G, Baker SE, Bruijnincx PCA, Hilden KS, Grigoriev IV, Mäkelä MR, de Vries RP. 2018. *Dichomitus squalens* partially tailors its molecular responses to the composition of solid wood. Environ Microbiol 20:4141-4156.
17. Daly P, Slaghek GG, Casado López S, Wiebenga A, Hilden KS, de Vries RP, Mäkelä MR. 2017. Genetic transformation of the white-rot fungus *Dichomitus squalens* using a new commercial protoplasting cocktail. J Microbiol Methods 143:38-43.
18. Khosravi C, Benocci T, Battaglia E, Benoit I, de Vries RP. 2015. Sugar catabolism in *Aspergillus* and other fungi related to the utilization of plant biomass, p 1-28. *In* Sariaslani S, Gadd GM (ed), Adv Appl Microbiol, vol 90. Academic Press.
19. Maeda H, Sano M, Maruyama Y, Tanno T, Akao T, Totsuka Y, Endo M, Sakurada R, Yamagata Y, Machida M, Akita O, Hasegawa F, Abe K, Gomi K, Nakajima T, Iguchi Y. 2004. Transcriptional analysis of genes for energy catabolism and hydrolytic enzymes in the filamentous fungus *Aspergillus oryzae* using cDNA microarrays and expressed sequence tags. Appl Microbiol Biotechnol 65:74-83.
20. Cubero B, Scazzocchio C. 1994. Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. EMBO J 13:407 - 415.
21. Mach R, Strauss J, Zeilinger S, Schindler M, Kubicek C. 1996. Carbon catabolite repression of *xyn1* (xylanase I-encoding) gene expression in *Trichoderma reesei*. Mol Microbiol 21:1273 - 1281.
22. Konishi T, Kotake T, Soraya D, Matsuoka K, Koyama T, Kaneko S, Igarashi K, Samejima M, Tsumuraya Y. 2008. Properties of family 79 beta-glucuronidases that hydrolyze beta-glucuronosyl and 4-O-methyl-beta-glucuronosyl residues of arabinogalactan-protein. Carbohydr Res 343:1191-201.
23. Chen J, Guo X, Zhu M, Chen C, Li D. 2019. Polysaccharide monooxygenase-catalyzed oxidation of cellulose to glucuronic acid-containing cello-oligosaccharides. Biotechnol Biofuels 12:42.
24. Rytioja J, Hildén K, Di Falco M, Zhou M, Aguilar-Pontes MV, Sietiö O-M, Tsang A, de Vries RP, Mäkelä MR. 2017. The molecular response of the white-rot fungus *Dichomitus squalens* to wood and non-woody biomass as examined by transcriptome and exoproteome analyses. Environ Microbiol 19:1237-1250.

25. Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyyssönen E, Bhatia A, Ward M, Penttilä M. 2002. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. Eur J Biochem 269:4202-4211.
26. Eibinger M, Sigl K, Sattelkow J, Ganner T, Ramoni J, Seiboth B, Plank H, Nidetzky B. 2016. Functional characterization of the native swollenin from *Trichoderma reesei*: study of its possible role as C1 factor of enzymatic lignocellulose conversion. Biotechnol Biofuels 9:178.
27. Li S, Jiang Z, Sun L, Liu X, Huang Y, Wang F, Xin F. 2017. Characterization of a new fungal immunomodulatory protein, FIP-dsq2 from *Dichomitus squalens*. J Biotechnol 246:45-51.
28. Jarai G, Buxton F. 1994. Nitrogen, carbon, and pH regulation of extracellular acidic proteases of *Aspergillus niger*. Curr Genet 26:238-244.
29. Cohen BL. 1973. Regulation of intracellular and extracellular neutral and alkaline proteases in *Aspergillus nidulans*. Microbiology 79:311-320.
30. Makinen H, Hynynen J, Siitonen J, Sievanen R. 2006. Predicting the decomposition of Scots pine, Norway spruce, and birch stems in Finland. Ecol Appl 16:1865-79.
31. Krah FS, Bassler C, Heibl C, Soghigian J, Schaefer H, Hibbett DS. 2018. Evolutionary dynamics of host specialization in wood-decay fungi. BMC Evol Biol 18:119.
32. Hassan L, Lin L, Sorek H, Sperl LE, Goudoulas T, Hagn F, Germann N, Tian C, Benz JP. 2019. Crosstalk of cellulose and mannan perception pathways leads to inhibition of cellulase production in several filamentous fungi. MBio 10:e00277-19.
33. Hatakka AI, Uusi-Rauva AK. 1983. Degradation of ^{14}C-labelled poplar wood lignin by selected white-rot fungi. Appl Microbiol Biotechnol 17:235-242.
34. Levin AM, de Vries RP, Wösten HAB. 2007. Localization of protein secretion in fungal colonies using a novel culturing technique; the ring-plate system. J Microbiol Methods 69:399-401.
35. Klaubauf S, Zhou M, Lebrun MH, de Vries RP, Battaglia E. 2016. A novel L-arabinose-responsive regulator discovered in the rice-blast fungus *Pyricularia oryzae* (*Magnaporthe oryzae*). FEBS Lett 590:550-8.
36. Casado López S, Peng M, Daly P, Andreopoulos B, Pangilinan J, Lipzen A, Riley R, Ahrendt S, Ng V, Barry K, Daum C, Grigoriev IV, Hilden KS, Mäkelä MR, de Vries RP. 2019. Draft genome sequences of three monokaryotic isolates of the white-rot Basidiomycete fungus *Dichomitus squalens*. Microbiol Resour Announc 8.
37. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357.
38. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30.
39. Love M, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.
40. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490-D495.

41. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res 46:D624-d632.
42. Maere S, Heymans K, Kuiper M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21:3448-9.
43. Budak SO, Zhou M, Brouwer C, Wiebenga A, Benoit I, Di Falco M, Tsang A, de Vries RP. 2014. A genomic survey of proteases in Aspergilli. BMC Genomics 15:523-523.
44. Jonkers W, Rep M. 2009. Mutation of CRE1 in *Fusarium oxysporum* reverts the pathogenicity defects of the FRP1 deletion mutant. Mol Microbiol 74:1100-13.
45. Portnoy T, Margeot A, Linke R, Atanasova L, Fekete E, Sandor E, Hartl L, Karaffa L, Druzhinina I, Seiboth B, Le Crom S, Kubicek C. 2011. The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei:* a master regulator of carbon assimilation. BMC Genomics 12:269.
46. Sun J, Glass NL. 2011. Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. PLoS ONE 6:e25654-e25654.
47. Drysdale MR, Kolze SE, Kelly JM. 1993. The *Aspergillus niger* carbon catabolite repressor encoding gene, *creA*. Gene 130:241-245.
48. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res 47:W636-W641.
49. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189-91.
50. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222-6.
51. Vizcaino JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44:D447-56.

Tables

Table 1. Summary of proteases identified in the exo-proteomics and their differential production at 2 d after transfer of *D. squalens* onto ring-plates containing either Avicel or xylan with or without supplementation with 3% w/v D-glucose (G).

Dicsqu464_1 protein ID	Protease functional group (from Merops annotations)	Status Avicel vs Avicel+G	Status xylan vs xylan+G
153959	Aspartyl protease	No sig. diff.	No sig. diff.
542799		Higher in Avicel	Higher in xylan+G
812714		Only in Avicel+G	Higher in xylan
813173		No sig. diff.	No sig. diff.
941015		Only in Avicel+G	No sig. diff.
976858		Not present	No sig. diff.
1001160		Higher in Avicel	No sig. diff.
367116	Metalloprotease	Higher in Avicel+G	Higher in xylan
830406		No sig. diff.	No sig. diff.
907349		Higher in Avicel+G	No sig. diff.
910760		No sig. diff.	No sig. diff.
636882	Serine protease	No sig. diff.	No sig. diff.
806666		No sig. diff.	No sig. diff.
917108		No sig. diff.	No sig. diff.
917629		No sig. diff.	Not present
921007		Higher in Avicel+G	No sig. diff.
926265		No sig. diff.	No sig. diff.
928513		Only in Avicel+G	No sig. diff.
929001		Higher in Avicel+G	Higher in xylan+G
977823		Higher in Avicel	Higher in xylan
1004022		No sig. diff.	No sig. diff.
917216	Other	No sig. diff.	No sig. diff.
944074		No sig. diff.	No sig. diff.
953225		No sig. diff.	No sig. diff.
969286		No sig. diff.	No sig. diff.

733 Table 2. Summary of enrichment of CRE1/CreA binding motifs in the promoters of
734 *D. squalens* genes which were repressed in the presence of glucose on either Avicel or
735 xylan.

Type of motif	Motif sequence	Mean motif number per gene (all genes)	% enrichment of motifs in repressed genes			
			Avicel 5 h (325)[a]	Avicel 2 d (476)[a]	xylan 5 h (94)[a]	xylan 2 d (466)[a]
single direct	[GC][CT]GG[AG]G	1.48	+12.2	+14.3	+0.4	+16.1
single reverse	C[CT]CC[AG][GC]	1.78	+32.4	+26.2	+18.7	+38.4
singles all	both of above motifs	3.21	+21.4	+20.2	+8.7	+28.3
pair 1	[GC][CT]GG[AG]G[AGCT](1,100)[GC][CT]GG[AG]G	0.26	+19	+18.8	-9.5	+19.5
pair 2	C[CT]CC[AG][GC][AGCT](1,100)C[CT]CC[AG][GC]	0.34	+43.9	+42	+8.8	+49.1
pair 3	[GC][CT]GG[AG]G[AGCT](1,100)C[CT]CC[AG][GC]	0.25	+48.1	+24.9	+59	+44.1
pair 4	C[CT]CC[AG][GC][AGCT](1,100)[GC][CT]GG[AG]G	0.21	+38.2	+29.8	-0.1	+47.2
pairs all	all of above paired motifs	0.8	+29.5	+26.6	+13.9	+37.2

736 [a]In parenthesis are the total number of genes that were repressed in that condition.

Figure legends

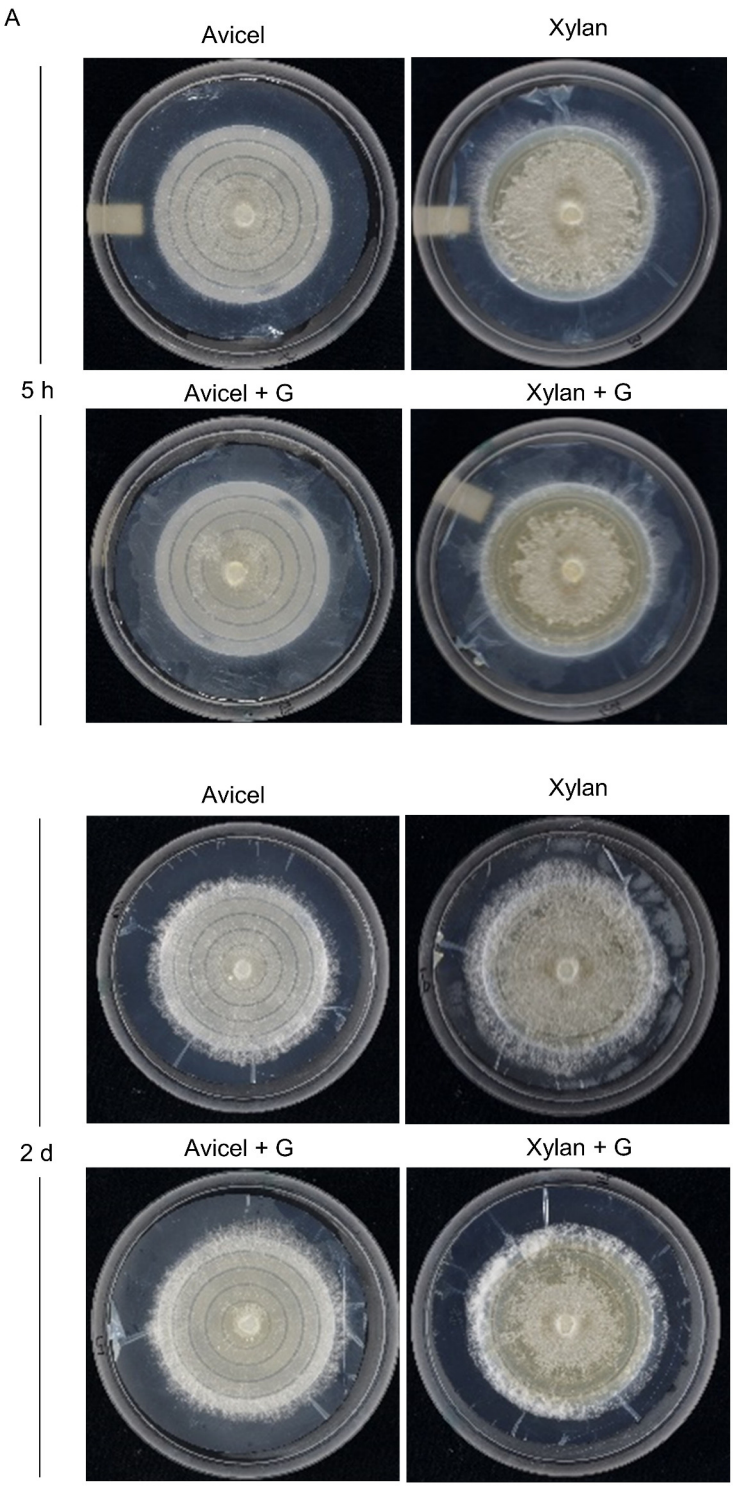
Fig. 1. Carbon catabolite repression phenotypes of *D. squalens* during growth on Avicel or xylan in the absence and presence of glucose. (A) Representative images of the *D. squalens* cultures on ring-plates containing either Avicel cellulose or beech wood xylan with or without supplementation with 3% w/v D-glucose (G) from either 5 h or 2 d after transfer of the pre-culture and (B) SDS-PAGE showing the banding pattern two days after transfer of pre-culture from protein samples used for proteomic analysis.

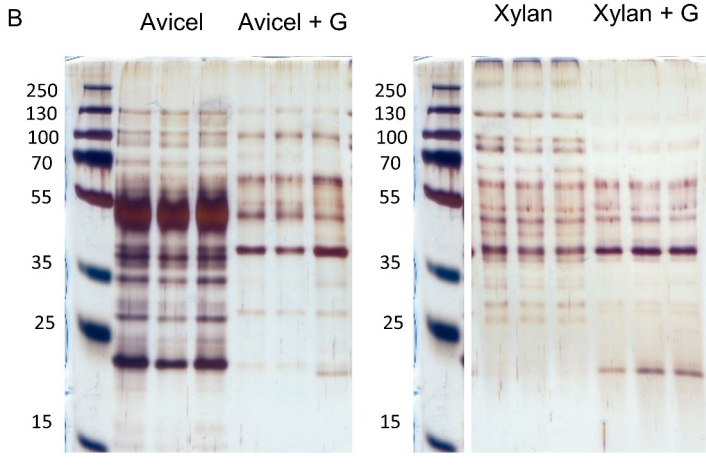
Fig. 2. Widespread glucose-mediated repression affecting a broad range of biological functions *in D. squalens*. (A) Comparison of the number of transcripts repressed by glucose (G) on Avicel or xylan after 5 h or 2 d of transfer from pre-culture. (B) Number of genes that were glucose-repressed on only one polysaccharide at a certain time-point. Each bar is divided into the proportions of these genes that were substantially expressed on both or only on the polysaccharide where glucose repression was detected. (C) Selection of enriched gene ontology (GO) terms showing broad effects of glucose-mediated repression. The size of the coloured bar indicates the percentage (for 100%, the bar would fill the entire cell) of the total genes annotated with an enriched GO term that were present in the repressed genes. (D) Comparison of plant biomass degrading CAZy genes repressed by glucose on Avicel or xylan after 5 h or 2 d of transfer of pre-culture.

Fig. 3. Repression of CAZymes affects a broad range of activities with strong support for the repressing pattern in the exo-proteomics. Hierarchical clustering, using

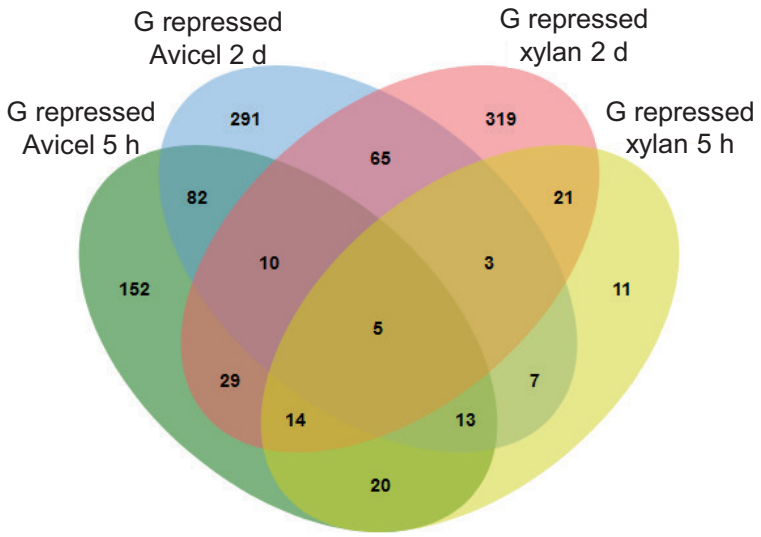
759 Euclidian distance, of transcript levels of plant biomass degrading CAZy encoding
760 genes from *D. squalens* mycelia grown on either Avicel or xylan with or without
761 supplementation with 3% w/v D-glucose (G). The CAZymes are colour-coded
762 according to substrate groups that they putatively act on (enzymes that act on multiple
763 polymers are coloured white). The protein abundance from exo-proteome is shown
764 adjacent to the transcript. See Fig. S3 for an image of the heatmap displaying
765 information about the gene each row corresponds to

766 Fig. 4. Total abundances of transcripts encoding proteases at 5 h and 2 d after transfer
767 of *D. squalens* onto ring-plates containing either Avicel or xylan with or without
768 supplementation with 3% w/v D-glucose (G). Error bars represent standard errors (n =
769 3).

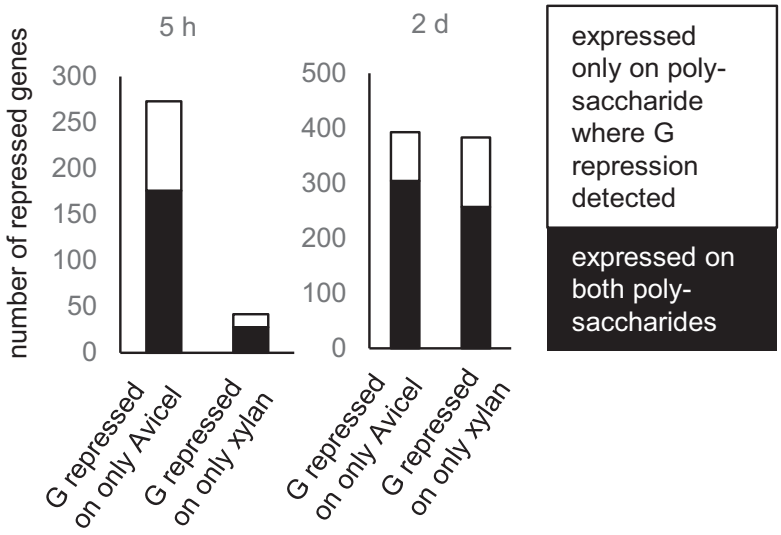

A
Avicel
Xylan
5 h
Avicel + G
Xylan + G
Avicel
Xylan
2 d
Avicel + G
Xylan + G


B
Avicel
Avicel + G
Xylan
Xylan + G
250
130
100
70
55
35
25
15
250
130
100
70
55
35
25
15

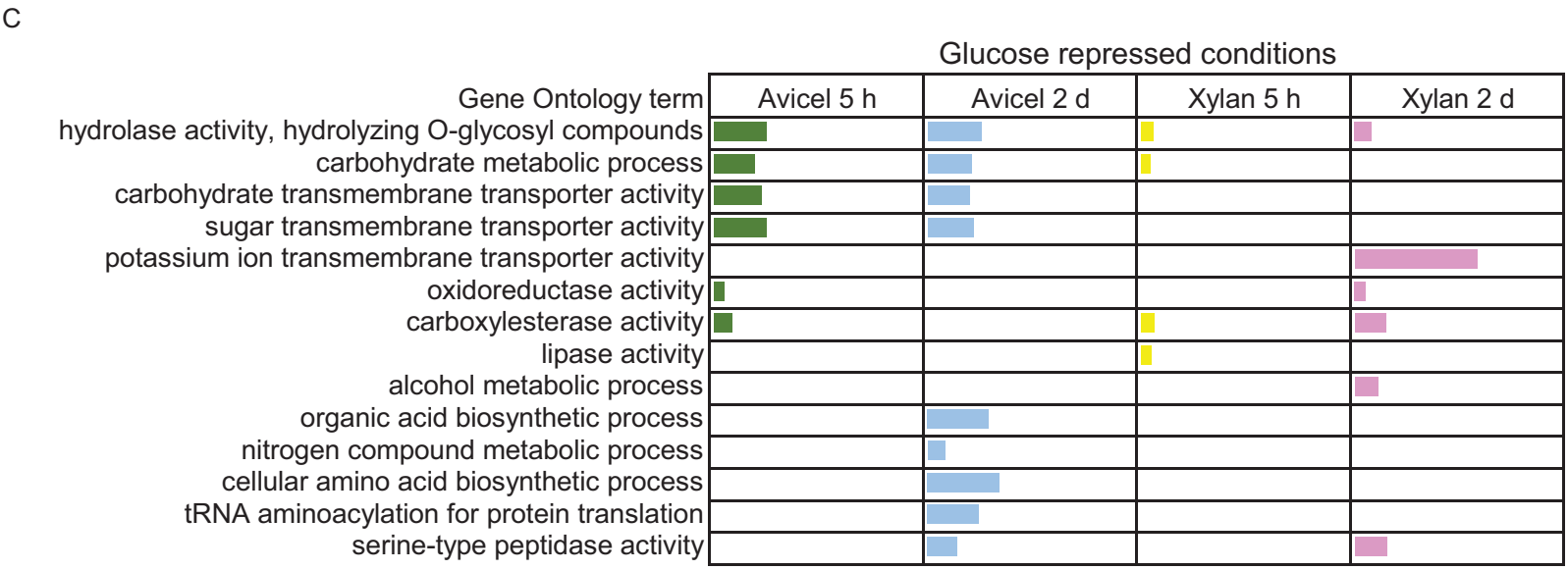
A


G repressed Avicel 2 d
G repressed xylan 2 d
G repressed Avicel 5 h
G repressed xylan 5 h
291
319
65
82
21
10
3
152
11
5
29
7
14
13
20

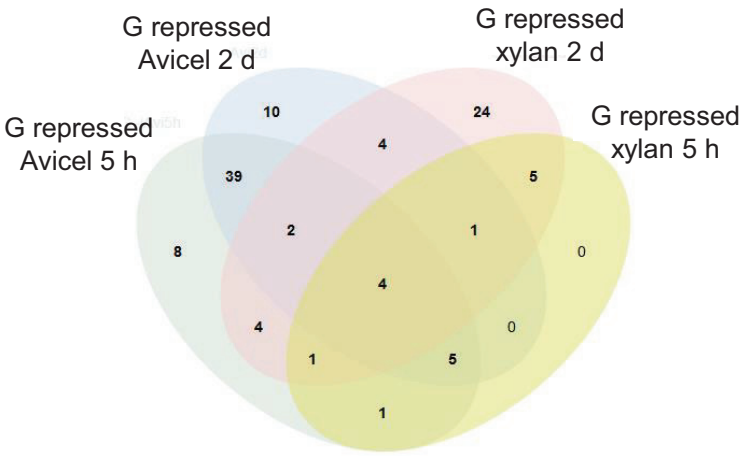
B

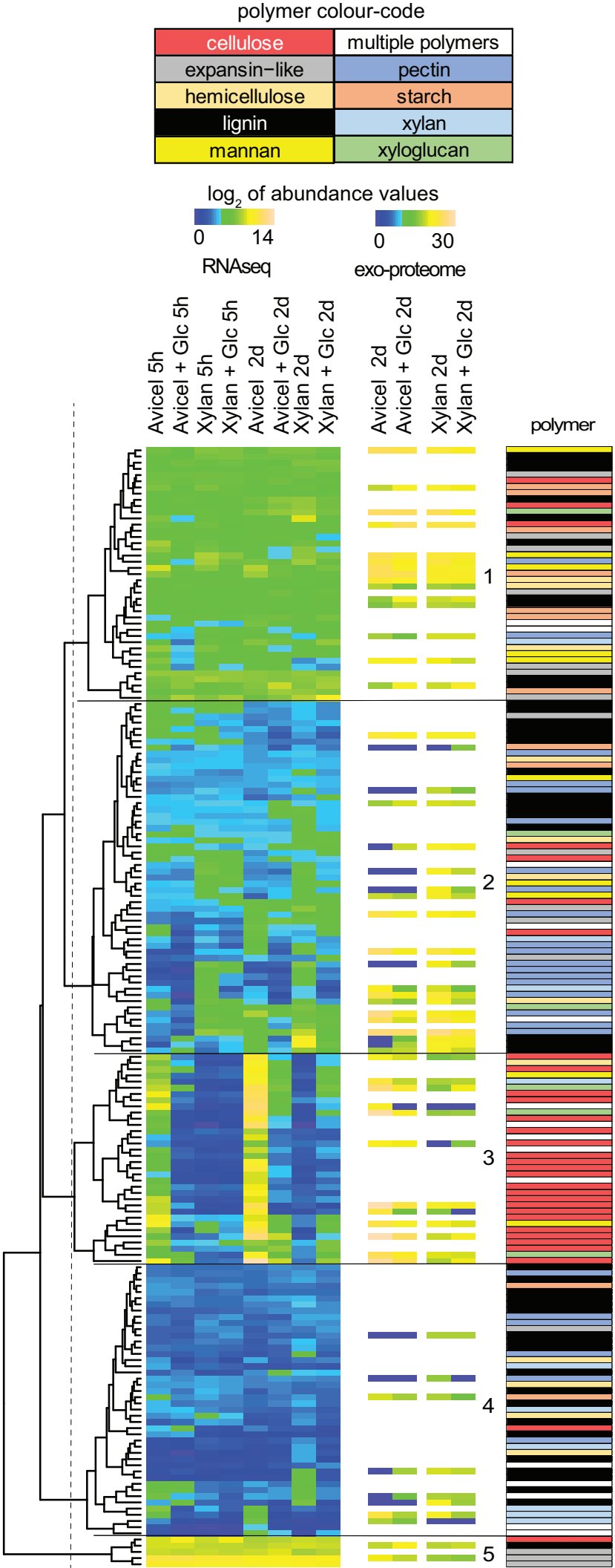

5 h
2 d
number of repressed genes
300
250
200
150
100
50
0
500
400
300
200
100
0
G repressed on only Avicel
G repressed on only xylan
G repressed on only Avicel
G repressed on only xylan
expressed only on poly-saccharide where G repression detected
expressed on both poly-saccharides

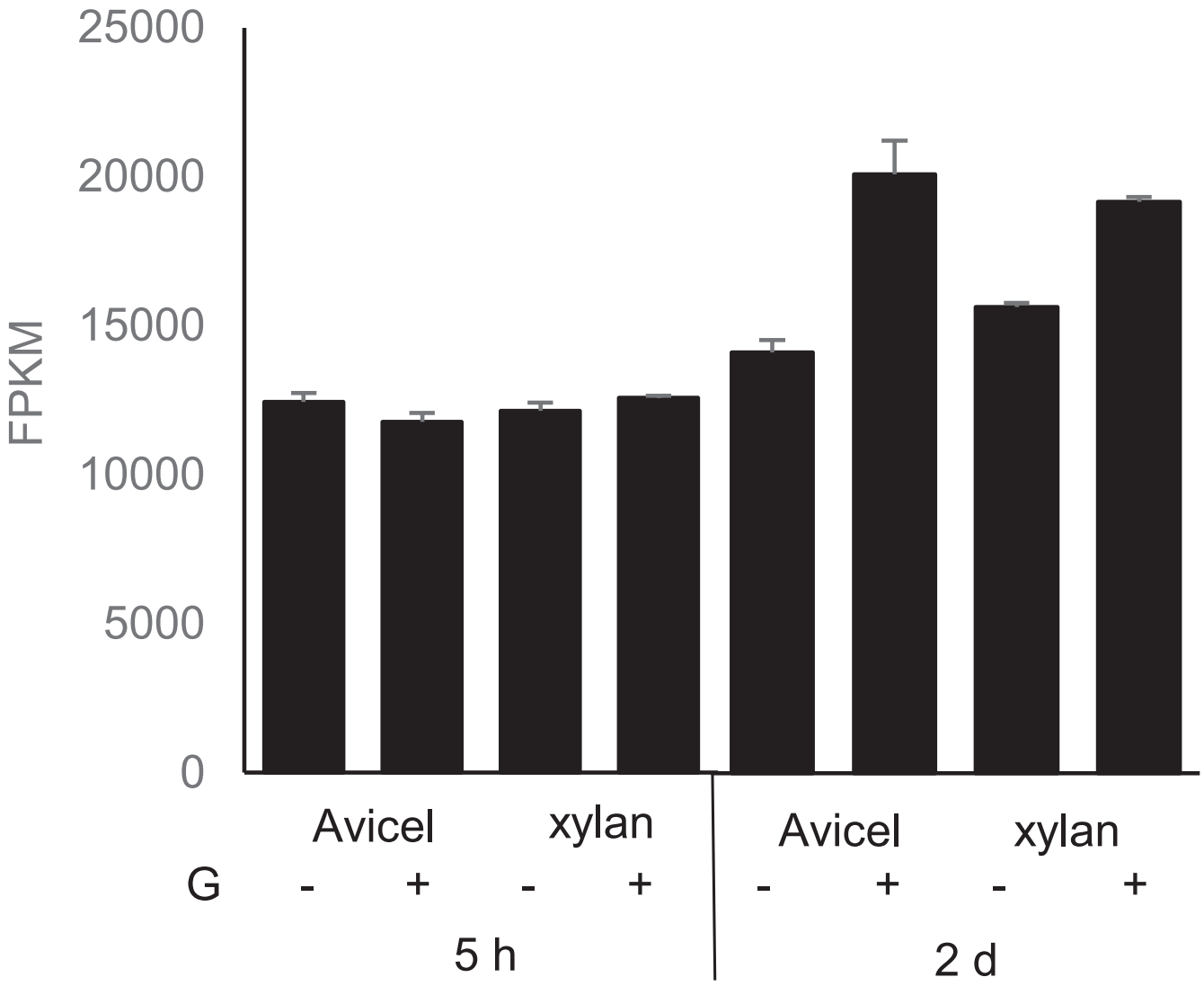
C


Glucose repressed conditions
Gene Ontology term
Avicel 5 h
Avicel 2 d
Xylan 5 h
Xylan 2 d
hydrolase activity, hydrolyzing O-glycosyl compounds
carbohydrate metabolic process
carbohydrate transmembrane transporter activity
sugar transmembrane transporter activity
potassium ion transmembrane transporter activity
oxidoreductase activity
carboxylesterase activity
lipase activity
alcohol metabolic process
organic acid biosynthetic process
nitrogen compound metabolic process
cellular amino acid biosynthetic process
tRNA aminoacylation for protein translation
serine-type peptidase activity

D


G repressed Avicel 2 d
G repressed xylan 2 d
G repressed Avicel 5 h
G repressed xylan 5 h
10
24
4
39
5
2
1
8
0
4
4
0
1
5
1


polymer colour-code
cellulose
multiple polymers
expansin−like
pectin
hemicellulose
starch
lignin
xylan
mannan
xyloglucan
log2 of abundance values
0
14
RNAseq
0
30
exo-proteome
Avicel 5h
Avicel + Glc 5h
Xylan 5h
Xylan + Glc 5h
Avicel 2d
Avicel + Glc 2d
Xylan 2d
Xylan + Glc 2d
Avicel 2d
Avicel + Glc 2d
Xylan 2d
Xylan + Glc 2d
polymer
1
2
3
4
5


25000
20000
15000
10000
5000
0
FPKM
Avicel
xylan
Avicel
xylan
G
-
+
-
+
-
+
-
+
5 h
2 d