

## Cell Wall Degrading Enzymes in *Trichoderma asperellum* Grown on Wheat Bran

Lasse Bech\*, Peter Kamp Busk and Lene Lange

Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Copenhagen, Denmark

### Abstract

*Trichoderma asperellum* is a filamentous fungus that is able to produce and secrete a wide range of extracellular hydrolytic enzymes used for plant cell wall degradation. The *Trichoderma* genus has attracted considerable attention from the biorefinery industry due to the production of cell wall degrading enzymes and strong secretion ability of this genus. Here we report extensive transcriptome analysis of plant cell wall degrading enzymes in *T. asperellum*. The production of cell wall degrading enzymes by *T. asperellum* was tested on a range of cellulosic materials under various conditions. When *T. asperellum* was grown on wheat bran, the greatest range of enzymes activity was detected and a total of 175 glycoside hydrolases from 48 glycoside hydrolase families were identified in the transcriptome. The glycoside hydrolases were identified on a functional level using the bioinformatical tool Peptide Pattern Recognition enabling an efficient enzyme discovery. This was furthermore used to re-annotate CAZymes present in five publically available *Trichoderma* species, hereby elucidating differences in CAZymes on a functional level in contrary to glycoside hydrolase family level. This comparison supports the theory that the glycoside hydrolases have evolved from a common ancestor, followed by a specialization in which saprotrophic fungi such as *T. reesei* and *T. longibrachiatum* lost a significant number of genes including several glycoside hydrolases.

**Keywords:** *Trichoderma asperellum*, transcriptome, CAZymes, Peptide Pattern Recognition, plant cell wall degrading enzymes (CWDE).

### Introduction

In recent years a significant effort has been put into efficient utilization of plant biomass, such as of agricultural crops, crop residues and industrial byproducts, for conversion into a range of value-added bio-products to maximize the value derived from the biomass feedstock (Bozell & Petersen, 2010; Ragauskas et al., 2006). Plants have evolved complex structural and chemical mechanisms for resisting degradation of their structural sugars by for example fungi (Popper, 2008). To overcome the problem of recalcitrance of the plant cell wall, industry uses a range of pretreatment approaches such as use of cell wall degrading enzymes to convert the plant cell wall polymers into smaller molecules (Chundawat, Beckham, Himmel, & Dale, 2011; Himmel et al., 2007). In nature, various microorganisms produce enzymes that act independently or in synergy to break down the plant cell wall. Although it is not fully known how many enzymes are involved in plant cell-wall degradation, three general categories of enzymes are considered necessary to hydrolyze native cell-wall materials: Cellulases, hemicellulases, and accessory enzymes. The recalcitrance of plant cell walls to enzymatic degradation and the high cost of the necessary hydrolytic enzymes are significant barriers to the global and large-scale production of biofuels and value-added bio-based products (Merino & Cherry, 2007). There is a need to develop more efficient and cost effective enzyme mixtures for the conversion of biomass to fermentable sugars to increase production of cellulose-derived value-added products, such as food, feed, chemicals and biofuels. To address this challenge, it is essential to gain a better understanding of the interactions between plant cell wall polysaccharides and cell wall degrading enzymes (CWDE). The complexity of the plant cell wall is mirrored by the diverse arsenal of CWDE produced by lignocellulose-degrading microbes. Enzymes acting on sugar structures are collectively called Carbohydrate-Active enZymes (CAZymes) (Cantarel et al., 2009). The diversity of CAZymes reflects the structural diversity of plant cell walls. Each type of CAZyme function is represented in multiple families determined by sequence

and structural similarities (Henrissat & Bairoch, 1993; Henrissat & Bairoch, 1996; Henrissat, Teeri, & Warren, 1998). There are currently 133 glycoside hydrolase (GH) families in the CAZy database, <http://www.cazy.org/>. As a consequence of more genomes being sequenced a large number of enzymes are either labelled as “hypothetical” or even annotated incorrectly (Schnoes et al. 2009; Gerlt et al. 2011). High throughput annotation of CAZymes is done largely based on GH family using Blast or HMMer programs, leaving many proteins without a functional annotation (Lombard, Golaconda Ramulu, Drula, Coutinho, & Henrissat, 2014). A more specific annotation is achieved by a following manual analysis as for example described for the re-annotation of *Trichoderma reesei* CAZymes (Häkkinen et al., 2012).

Many *Trichoderma* species are strong opportunistic invaders. They are fast growing and abundantly producers of spores, and contain highly active CAZymes and secondary metabolites (Monte, 2010). *T. asperellum* is a mycoparasitic species that are well known and widely used for their ability to inhibit the growth of plant pathogens. Their multi-enzymatic systems position them as some of the best anti-microbial agents (Borrero, Trillas, Delgado, & Aviles, 2012; Mbarga et al., 2012; Schuster & Schmoll, 2010). In addition to anti-microbial capability, *T. asperellum* has shown its potential in secreting other GHs under appropriate culture conditions (Aires et al., 2012; Marx et al., 2013), such as on cellulose (Raghuwanshi, Deswal, Karp, & Kuhad, 2014) and on sugar bagasse (Marx et al., 2013). On the latter substrate,

**\*Corresponding author:** Lasse Bech, Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Copenhagen, Denmark, Tel: (+45)9940 7512; E-mail: [lla@adm.aau.dk](mailto:lla@adm.aau.dk)

**Received** December 16, 2014; **Accepted** December 19, 2014; **Published** December 26, 2014

**Citation:** Bech L, Busk PK, Lange L (2014) Cell Wall Degrading Enzymes in *Trichoderma asperellum* Grown on Wheat Bran. Fungal Genom Biol 4: 116. doi:10.4172/2165-8056.1000116

**Copyright:** © 2014 Bech L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*T. asperellum* showed promising results as an enzyme producer with a higher diversity of hemi-cellulases and  $\beta$ -glucosidases compared to *T. reesei*.

The purpose of the present work was to analyze the GH enzymes in the transcriptome of *T. asperellum* when grown on wheat bran. This was done partly to understand the plant cell wall degrading system of *T. asperellum* and to identify enzymes important for the degradation of complex sugar structures, as well as to discover new enzymes. Several *Trichoderma* genomes are now available, which permitted a comparative study of GHs with the bioinformatic tool Peptide Pattern Recognition (PPR). With PPR it is possible to predict not only the GH family but also the enzyme function, so that the CAZymes may be compared on a functional level (Busk & Lange, 2013).

## Materials and methods

### Strains

*Trichoderma asperellum*, strain number CBS 433.97 from the Centraalbureau, Schimmelcultures, CBS, The Netherlands, was used in this study. For genome comparison the following genome sequences were used: *Trichoderma reesei* (Genbank accession number GCA\_000167675.2); *Trichoderma atroviride* (GCA\_000171015.2); *Trichoderma virens* (GCA\_000170995.2); *Trichoderma hamatum* (GCA\_000331835.1); *Trichoderma longibrachiatum* (GCA\_000332775.1)

### Biomass

*T. asperellum* was cultivated on 1.5% (w/v) agar plates (Sigma Aldrich, UK) with 2.5% (w/v) different cellulosic material for 3, 5 and 7 days each at 20°C and 30°C, where additionally each combination of incubation duration and temperature was conducted at three levels of pH (pH 4, pH 6.5 and pH 8). Cellulosic material included: wheat bran (Finax, Denmark), oatmeal (Kornkammeret, Denmark), *Spirodella polyrrhiza* (BD-D2013-7, CIB, Chengdu, China), *Lemna minor* (BD-L2013-7, CIB, Chengdu, China), wheat straw, *Brassica oleracea* var. *medullosa*, *Sinapis alba*, *Cannabis sativa* (Molinuevo-Salces, Fernández-Varela, & Uellendahl, 2014).

### Activity

Agar plate cultures of *T. asperellum* on each of the above cellulosic biomass substrates were washed with 10ml H<sub>2</sub>O containing 1% Tween 80 to extract secreted enzymes. Enzyme activity was measured on AZCL-plates prepared using 1% agarose in buffer, pH 5.8, containing 0.1% azurine cross-linked (AZCL) substrates (Megazymes, Bray, Ireland) as described in the protocol. Inoculation was done by placing 15  $\mu$ L of enzyme blend from the growth plates in holes punched in AZCL-agarose plates. Enzyme activity was indicated by the area (cm<sup>2</sup>) of blue color zones resulting from hydrolysis of the substrate. The extent of the blue zones, indicating the enzyme activity, was measured after 24h at 30°C.

### RNA isolation

Mycelia were scraped from the surface of two agar plates on which *T. asperellum* had grown on wheat bran for seven days at 30°C, frozen in liquid nitrogen and ground to a powder. RNA was extracted from the mycelia with fenzol and the RNA Total Maxi kit protocol (A&A Biotechnology, Gdynia, Poland).

### Transcriptome sequencing

RNA was sequenced by Beijing Genomics Institute (BGI) using

the Illumina HiSeq 2000 platform. Reads were assembled using Trinity (Grabherr et al., 2011) by first combining clean reads with a certain length of overlap to form longer fragments without N into contigs. The reads were then mapped back to contigs with paired-end reads in order to detect contigs from the same transcript as well as to determine the distances between those contigs. Trinity then connected the contigs, using N to represent unknown sequences between every two contigs. From this, scaffolds were made. Paired-end reads were used again for gap-filling of scaffolds to obtain sequences with the fewest Ns which could not be extended on either end. Such sequences were defined as unigenes. TGICL (Pertea et al., 2003) at default parameters was used to cluster these unigenes to acquire non-redundant unigenes of the greatest length possible. In the final step, a Blast X alignment was performed between the unigenes and the protein databases of NR, Swiss-Prot, KEGG, and COG. The best aligning results were used to determine the sequence directions of those unigenes. If the results of different databases conflicted with each other, we followed a priority order of NR–Swiss-Prot–KEGG–COG. Orientation and CDS of sequences that have no hits in blast search were predicted using ESTScan (Iseli, Jongeneel, & Bucher, 1999). Original transcript sequences (5'→3') were chosen if their orientations were unable to be determined by above mentioned approaches. Unigene expression levels were calculated using the RPKM method (Reads Per kb per Million reads) (Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008). The formula is  $RPKM = (1000000 * C) / (N * L * 1000)$ . This formula assigns RPKM(A) to be the expression of gene A, C to be number of reads that uniquely aligned to gene A, N to be total number of reads that uniquely aligned to all genes, and L to be number of bases on gene A. The RPKM method is able to eliminate the influence of different gene length and sequencing discrepancy on the calculation of gene expression. Therefore the calculated gene expression can be directly used for comparing the difference of gene expression among samples.

### Data processing

Further isolation of GHs was done by searching the obtained unigene annotation for the keywords: "glycoside, hydrolase, GH, EC: 3.2.1". These genes were afterwards checked for duplications by blasting gene fragments in NCBI and reverse blasting these results against a local database of obtained unigenes. Blast results lead to the construction of full unigenes without duplications. CLC Main workbench was used for local blast database and for collecting partial genes into full genes.

### Peptide Pattern Recognition

Peptide Pattern Recognition (PPR) was used for the identification of GHs on a functional level in the *Trichoderma* genomes (Busk & Lange, 2013).

### Phylogeny

Sequences were aligned using ClustalX 2.1 (<http://clustalx.software.informer.com/2.1/>) with default alignment parameters. Phylogenetic analysis was performed using the neighbor-joining algorithm of ClustalX 2.1 with default parameters (gapped regions were included). Bootstrap analysis (1000 trials) provided a measure of confidence for the detected relationships as described above. The resulting trees were visualized by the program Figtree (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Results

The choice of substrate is important for the successful production of CAZymes, because different biomasses induce distinct fungal enzyme

responses depending on the biomass structure and composition. An initial screening of ten carbohydrate-containing growth substrates was performed to induce the widest and largest production of CAZymes (Supplementary information). *T. asperellum* produced a wide range of enzymes after only three days incubation. The highest activity was recorded for endo-xylanases and  $\beta$ -glucanases on arabinoxylan, xylan and  $\beta$ -glucan. After three days wheat bran induced the widest range of enzyme activity as well as the largest response in terms of halo size, closely followed by duckweed, *Spirodella polyrhiza*. At five days the overall AZCL response increased for all growth substrates except PDA. The widest activity range with the largest halos was observed with wheat bran substrate at seven days (Table 1). Enzymes induced by growth on wheat bran resulted in activity on 8 out of the 13 tested AZCL-substrates, with the highest activity being recorded for xylanases with halo sizes of 4.9cm<sup>2</sup> and 4.5cm<sup>2</sup> on AZCL-arabinoxylan and AZCL-xylan, respectively. Second, the duckweed *Spirodella polyrhiza* induced enzymes that were detected on 7 out of 13 tested AZCL-substrates, but to a lesser degree than wheat bran. This was indicated by significantly less by cellulase and 1,4- $\beta$ -D-mannanase activity. Enzyme activities when grown on wheat bran were further investigated under different conditions. These included growth durations for five and seven days, temperatures from 20°C to 30°C and pH 4, pH 6.5 and pH 8 respectively. Neither the temperature nor the pH of the growth medium had any significant effect on the range of activities measured.

### CAZymes in the transcriptome of *T. asperellum* grown on wheat bran:

Duplicates of mycelium from *T. asperellum* grown on wheat bran at 30°C for seven days were harvested and their RNA content was isolated and sequenced according to the description in materials and methods. The sequencing resulted in 12,901,784 reads that were assembled into 32,709 unigenes with an average length of 654nt. The transcriptome sequences were deposited in Transcriptome Shotgun Assembly (TSA) database in Genbank at the NCBI with the accession number SRR1575447. A BLASTp and BLASTx search against the National Center for Biotechnology Information (NCBI)'s protein non-redundant (NR) database showed that 66% of unigenes had significant hits (1e-05). The other 34% of the unigenes had no significant similarity to known proteins or domains in the NCBI database. The latter may

contain sequences that mostly covered long untranslated regions of the mRNAs. Among the 32,709 unigenes, growth on wheat bran induced 175 GHs from 48 GH families in *T. asperellum* (Figure 1 and detailed list in the supplementary information Table S4). The majority of the GH families were represented by 1-2 genes, but some families were represented with significantly more genes. The most transcribed family was GH18 with 20 genes, all of which were predicted to be chitinases. 14 genes were transcribed from the diverse GH16 family, of which most were predicted to be endo-1,3- $\beta$ -glucanases. There were 11 genes each from the GH5 family and GH3 family of which the majority predicted to be endo- $\beta$ -1,4-glucanases and  $\beta$ -glucosidases, respectively. To determine which genes were transcribed the most, the genes were listed in terms of sequenced raw reads. The most transcribed CWDE was shown to be an amylase represented with 18148 raw reads, which was three times more than the next enzymes. A 1,3- $\beta$ -glucanase and a chitinase were the next most transcribed CAZyme with 7743 and 6831 raw reads, respectively.

### Transcriptome derived secretome

To analyze which secreted GHs were transcribed in *T. asperellum*, the sequenced GHs were analyzed for the presence of a signal peptide. Genes without sequenced 5' end were evaluated based on whether their homologue from *Trichoderma atroviride* contains a signal peptide. The transcriptome derived secretome contained 111 enzymes from 37 GH families, which corresponded to 62% of the transcribed GH enzymes covering 75% of the transcribed GH families (supplementary information). The majority of the signal peptide-containing sequences belonged to GH family GH18, GH16 and to some degree GH3 (Figure 1). By contrast, GH76 was represented to a higher degree in the secretome compared with the transcriptome. Listing genes with signal peptide according to number of raw reads showed that the 34 most transcribed genes corresponded to 90% of the raw reads (Table 2).

Different substrates induced different CAZyme responses in *T. asperellum*, based on the initial screening of substrates. To analyze the correlation between growth substrate composition and the secreted CAZymes, the enzymes found in the transcriptome derived secretome were divided into groups of starch degrading, cellulose degrading, hemicellulases, chitinases and unknown based on their predicted

7 days																				
AZCL/Medium	PDA		WBA		<i>Spirodella polyrhiza</i>		<i>Lemna minor</i>		Oakmea		wheat bran wheatstraw		<i>Brassica oleracea</i>		<i>Sinapis alba</i>		Cannabies			
Amylose	0	±0	0	±0	0,0	±0	0	±0	0	±0	0,0	±0	0,0	±0	0,0	±0	0,0	±0		
Arabinan	0	±0	0	±0	0,0	±0	0	±0	0	±0	0,6	±0	0,0	±0	0,0	±0	0,0	±0		
Arabinoxylan	0	±0	38	±0	3,8	±0	35	±0	35	±0	4,9	±0,2	0,0	±0	3,8	±0	4,2	±0,1	3,8	±0,1
Beta-glucan	0	±0	8	±0	11	±0	6	±0	9	±0	31	±0	15	±0	20	±0	15	±0	20	±0
Casein	0	±0	6	±0	8	±0	8	±0	6	±0	11	±0	8	±0	13	±0	11	±0	15	±0
Cellulose	0	±0	8	±0	8	±0	0	±0	8	±0	18	±0	0	±0	8	±0	8	±0	8	±0
Curdlan	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0
Galactan	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0
Galactomannan	0	±0	8	±0	8	±0	0	±0	0	±0	28	±0	0	±0	8	±0	8	±0	8	±0
Pullulan	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0
Rhamnogalactorunan	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0	0	±0
Xylan	0	±0	35	±0,1	45	±0,1	31	±0	25	±0	45	±0,2	0	±0	35	±0	38	±0	31	±0
Xyloglucan	0	±0	8	±0	15	±0	8	±0	11	±0	18	±0,1	0	±0	8	±0	15	±0	11	±0
Sum diameters	0		110		133		88		95		207		35		139		145		132	

**Table 1:** Enzyme activities on different AZCL-substrates after seven days growth on different media, represented by the size of blue AZCL-halos in cm<sup>2</sup>.



function. The transcriptional level based on raw reads of these enzyme groups was compared with the macromolecular composition of wheat bran (Figure 2). A major part of the GHs predicted to be in the secretome were starch degrading enzymes, especially the highly transcribed amylase. The second most produced enzymes were hemicellulases, corresponding to the relatively high degree of arabinoxylan in wheat bran. The least secreted enzymes were cellulases, where the majority of the transcripts were two endo-1,3- $\beta$ -glucanases from GH16, a family, where similar enzymes has shown also to contain 1,4- $\beta$ -glucanases activity. Though, endo-1,3- $\beta$ -glucanases are found to be more involved in anti-microbial actions than cellulose degradation (Vazquez-Garciduenas, Leal-Morales, & Herrera-Estrella, 1998). To a lesser transcription degree, two  $\beta$ -glucosidases (GH3) and a cellobiosidase (GH7) were present in the transcriptome derived secretome. Moreover, a relatively large part of the transcriptome derived secretome was chitinases, probably induced to rearrange the cell wall of fungus itself or as a constant anti-microbial response.

### Composition of CAZymes across six *Trichoderma* species

Using a multigene phylogenetic approach, the *Trichoderma* genus can be divided into four big clades (Kullnig-Gradinger, Szakacs, & Kubicek, 2002). Clade A “section *Trichoderma*” consists of among others, *T. atroviride*, *T. hamatum* and *T. asperellum*. Clade B “section *Pachybasium*” consists of among others *T. virens*. Clade C “section *Longibrachiatum*” contains *T. longibrachiatum* and *T. reesei*. A small clade D consisting of *H. aureoviridis*. Peptide Pattern Recognition was used to compare the GHs from the genome of five *Trichoderma* species and the transcriptome from *T. asperellum* at a functional level (Table 3). Based on PPR analysis significantly more GHs were identified in the represented species from section *Trichoderma* and *Pachybasium* compared to the

*Longibrachiatum* species. *T. virens* contained a GH profile more like *T. asperellum*, *T. atroviride* and *T. hamatum* rather than *T. reesei* and *T. longibrachiatum*. The difference was mostly due to a higher content of hemicellulases and chitinases in the species from *Trichoderma* and *Pachybasium* section, but they also contained a higher number of genes within all major enzyme classes (starch degrading, cellulose degrading, hemicellulases, chitinases). PPR analysis identified relatively few starch degrading enzymes, both in terms of functions but also in gene copies, though these were the most transcribed enzyme group based on raw reads. Twice as many hemicellulose degrading as cellulose functions were found across all species, though more were present in *T. virens*, *T. atroviride* and *T. hamatum* compared to *T. reesei* and *T. longibrachiatum*. In terms of gene copies, *T. reesei* and *T. longibrachiatum* had roughly the same number cellulose acting enzymes as hemicellulose, while *T. virens*, *T. atroviride* and *T. hamatum* expressed around 20% more hemicellulose acting enzymes compared to cellulose acting enzymes. The PPR analysis of GHs at a functional level makes it possible to pinpoint differences in the enzymatic potential of the fungi in terms of specific functions (Table 3). This has revealed for example fewer genes for the *Longibrachiatum* section within functions with high copy number like chitinase (EC. 3.2.1.14) and  $\beta$ -glucosidase (EC. 3.2.1.21). Further, *T. reesei* and *T. longibrachiatum* have lost 9-10 functions, for which the species in section *Trichoderma* have a low copy number; these include a  $\beta$ -fructofuranosidase, galactan 1,3- $\beta$ -galactosidase, and  $\alpha$ -L-rhamnosidase. The PPR analysis also showed a fructan  $\beta$ -fructosidase to be present only in *T. hamatum* and an exo- $\alpha$ -sialidase present only in *T. asperellum*. Bacterial exo- $\alpha$ -sialidase has been shown to cleave sialic acids from sialyloligosaccharides, gangliosides and glycoproteins. These glycoconjugates are mostly found on surface exposed locations and thought to be among other things involved in cell-to-cell interactions and adhesion (Kiyohara et al., 2011).

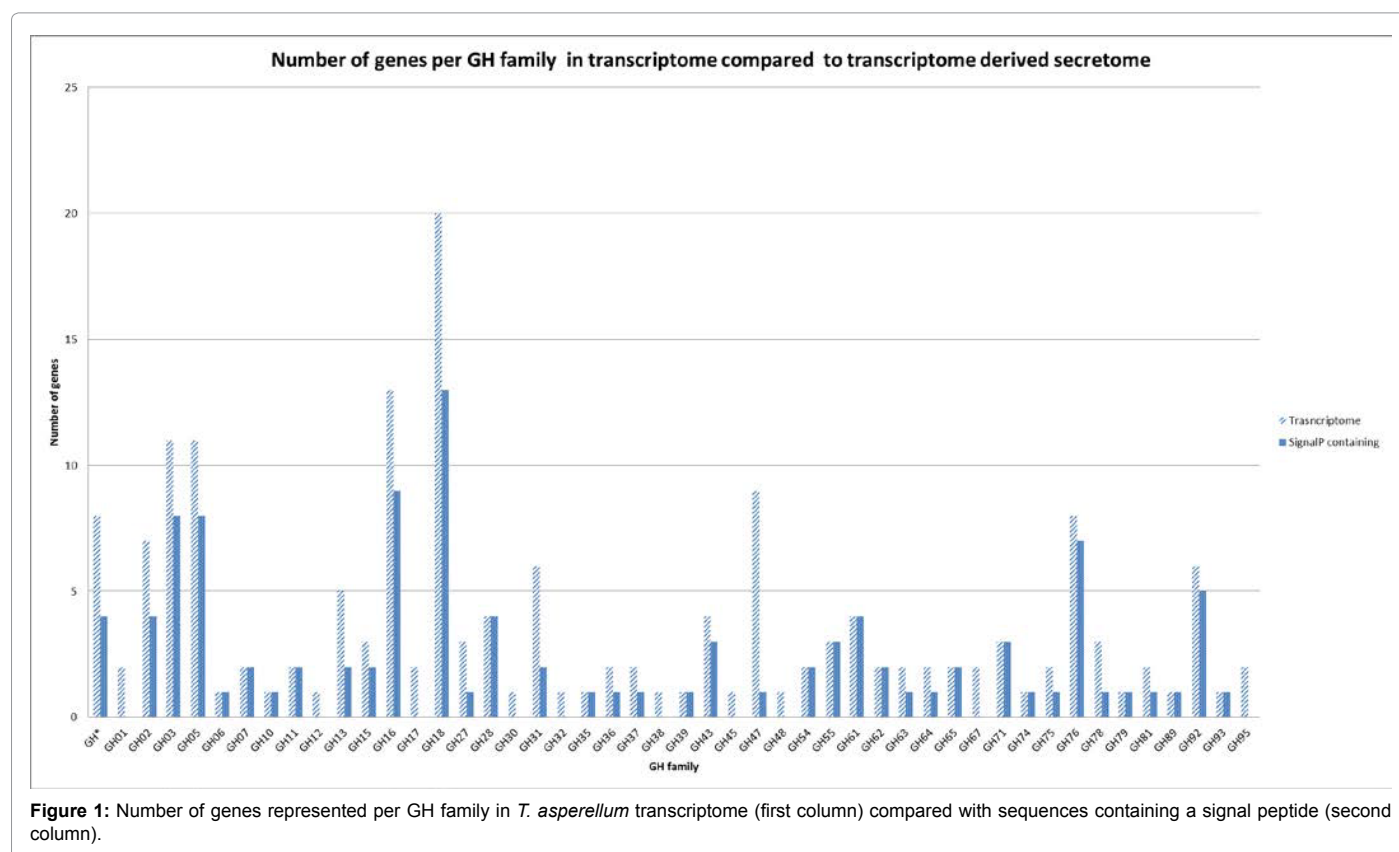
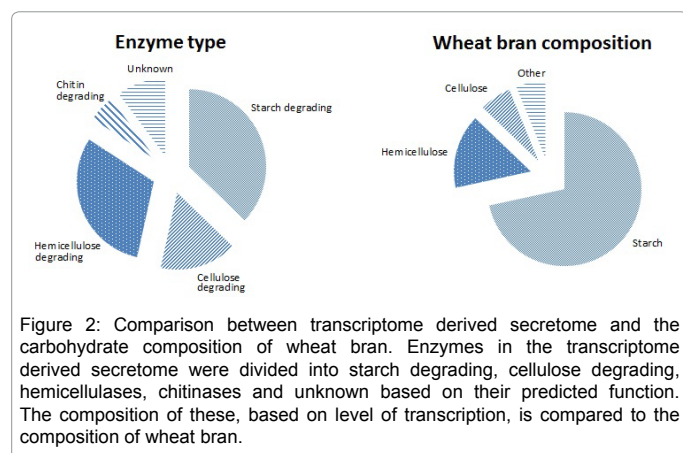


Figure 1: Number of genes represented per GH family in *T. asperellum* transcriptome (first column) compared with sequences containing a signal peptide (second column).

Gene ID	Raw reads	BGI-annotation	PPR-annotation	EC Number	Family
Unigene396_Lyx4	18148	Amylase	Glu DaY $\epsilon$ , $\kappa$ - $\alpha$ -glu Dosidase	3.2.1.3	GH15
Unigene562_Lyx4	6831	Chitinase	Chitinase	3.2.1.14	GH18
Unigene1796_Lyx4	5134	$\beta$ - $\epsilon$ , $\uparrow$ -gluDaY	gluDaY eYdo- $\epsilon$ , $\uparrow$ - $\beta$ -D-gluDosidase	3.2.1.39	GH16
Unigene5138_Lyx4	2506	$\beta$ -glLJDosidase		*	GH76
Unigene2441_Lyx4	2498	Xylanase	eYdo- $\epsilon$ , $\kappa$ - $\beta$ -dzLJlaYase	3.2.1.8	GH11
Unigene2199_Lyx4	1702	$\beta$ - $\epsilon$ , $\uparrow$ -edzogludaYase	gluDaY $\epsilon$ , $\uparrow$ - $\beta$ -gluDosidase	3.2.1.58	GH55
Unigene5138_Lyx4	1507	$\alpha$ -glLJDosidase		*	GH76
Unigene3977_Lyx4	1494	$\alpha$ -glLJDosidase	*mannan endo-1,6-alpha-mannosidase	*	GH76
Unigene5438_Lyx4	1482	$\alpha$ -glLJDosidase		*	GH76
Unigene4139_Lyx4	1376	$\alpha$ -L-aradiYofuraYosidase	gluDaY $\epsilon$ , $\uparrow$ - $\beta$ -gluDosidase	3.2.1.58	GH54
Unigene4884_Lyx4	1359	Chitinase	chitinase	3.2.1.14	GH18
Unigene5992_Lyx4	1358	$\alpha$ -glLJDosidase	*mannan endo-1,6-alpha-mannosidase	*	GH76
Unigene4496_Lyx4	1351	Mannosyl-oligosaccharide glucosidase	mannosyl-oligosaccharide glucosidase	3.2.1.106	GH63
Unigene4405_Lyx4	1284	$\beta$ -gluDosidase	$\beta$ -gluDosidase	3.2.1.21	GH03
Unigene2969_Lyx4	1158	$\alpha$ -gluDosidase	$\alpha$ -gluDosidase	3.2.1.20	GH31
Unigene10314_Lyx4	1046	$\beta$ -dzLJllosidase/ $\alpha$ -L-aradiYofuraYosidase	dzLJlaY $\epsilon$ , $\kappa$ - $\beta$ -dzLJllosidase	3.2.1.37	GH43
Unigene4419_Lyx4	1041	eYdo- $\epsilon$ , $\kappa$ - $\beta$ -gluDaYase		*	GH05
Unigene6388_Lyx4	1041	$\beta$ - $\epsilon$ , $\uparrow$ - $\epsilon$ , $\kappa$ -gluDaYase		*	GH16
Unigene8823_Lyx4	939	$\alpha$ -glLJDosidase		*	GH76
Unigene5396_Lyx4	812	$\alpha$ -gluDosidase	$\alpha$ -gluDosidase	3.2.1.20	GH31
Unigene9301_Lyx4	743	EYdo- $\epsilon$ , $\uparrow$ ; $\kappa$ - $\beta$ -gluDaYase	gluDaY eYdo- $\epsilon$ , $\uparrow$ - $\beta$ -D-gluDosidase	3.2.1.39	GH16
Unigene9622_Lyx4	639	Cellobiohydrolase	AA9	AA9	GH06
Unigene3741_Lyx4	605	Chitinase	chitinase	3.2.1.14	GH18
Unigene4031_Lyx4	604	$\beta$ -galaDtosidase	galaDturaY $\epsilon$ , $\kappa$ - $\alpha$ -galaDturoYidase	3.2.1.67	GH28
Unigene9036_Lyx4	595	$\beta$ -dzLJllosidase	dzLJlaY $\epsilon$ , $\kappa$ - $\beta$ -dzLJllosidase	3.2.1.37	GH03
Unigene7861_Lyx4	589	$\beta$ -gluDosidase	$\beta$ -gluDosidase	3.2.1.21	GH03
Unigene8089_Lyx4	588	Chitinase	chitinase	3.2.1.14	GH18
Unigene14133_Lyx4	581	$\alpha$ - $\epsilon$ , $\uparrow$ -waYosidase	waYosLJl-oligosadDharide $\epsilon$ , $\uparrow$ - $\alpha$ -waYosidase	3.2.1.113	GH47
Unigene7664_Lyx4	526	$\beta$ -galaDtosidase	$\beta$ -galaDtosidase	3.2.1.23	GH02
Unigene6759_Lyx4	513	$\beta$ - $\epsilon$ , $\uparrow$ ; $\kappa$ -eYdogluDaYase	gluDaY eYdo- $\epsilon$ , $\uparrow$ - $\beta$ -D-gluDosidase	3.2.1.39	GH16
Unigene6619_Lyx4	494	Cellobiohydrolase	Delulose $\epsilon$ , $\kappa$ - $\beta$ -Delodfosidase ;reduDiYg eYd	3.2.1.176	GH07
Unigene6963_Lyx4	467	Chitinase	chitinase	3.2.1.14	GH18
Unigene8432_Lyx4	466	$\alpha$ -L-aradiYofuraYosidase	$\alpha$ -N-aradiYofuraYosidase	3.2.1.55	GH62
Unigene14373_Lyx4	421	$\alpha$ - $\epsilon$ , $\uparrow$ -waYosidase		*	GH92

**Table 2:** Transcriptome derived secretome represented by the 34th most transcribed genes containing signal peptide, corresponding to 90% of total secretome in terms of raw reads. Enzymes are listed by number of raw reads. The columns contain BGI-annotation as described in materials and methods. PPR annotation is a result of the predicted function by EC number.



evolution on a single-gene level as a further examination of the genetic relationship between the *Trichoderma* species featured in the study. These genes were chosen based on their high transcription level and presence in the transcriptome derived secretome or because of their significance as the only gene with a specific function. The closest blast hits in NCBI were genes from *T. atroviride* or *T. hamatum* and the closest 20-30 hits were sequences from other ascomycete genes. The chosen protein sequences were aligned with all other enzymes from the corresponding GH family or function across all the *Trichoderma* species using ClustalX. When protein sequences were illustrated in a phylogenetic tree, clades form with related sequences from each *Trichoderma* and often according to their greater phylogenetic relationship on a species level (supplementary information). This generally form clades with one related sequence from each *Trichoderma* species, where sequences from the section *Trichoderma* (*T. asperellum*, *T. hamatum* and *T. atroviride*) exhibit closer relationship relative to sequences from section *Longibrachiatum* (*T. longibrachiatum* and *T. reesei*) and to sequences from the outlier *T. virens* from section

### Phylogenetic relationship on a single gene level

Five CWDE-encoding genes were chosen to investigate the

*Pachybasium*. Moreover the genes formed clades with one enzyme from each *Trichoderma*, although on some clades a gene was missing from one or more *Trichoderma*. The same trend was seen for the

other selected enzymes. To illustrate this point, the phylogenetic tree of endo-1,3- $\beta$ -D-glucosidase (EC. 3.2.1.39) is shown (Figure 3).

Enzyme group Annotation, <i>T.longibrochiatum</i> , EC number <i>T.reesei</i> , <i>T.virens</i> , <i>T.atroviride</i> , <i>T.hamatum</i> <i>T. asperellum</i> trans									
Starch degrading	$\alpha$ -amylase	3.2.1.1	1	1	2	2	1		2
	Glucoamylase	3.2.1.3	0	1	1	2	2		2
	Dextranase	3.2.1.11	0	0	1	0	1		0
	$\alpha$ -glucosidase	3.2.1.20	3	4	5	3	6		5
Cellulose degrading	Cellulase	3.2.1.4	5	4	7	4	5		4
	Endo- $\beta$ -1,4-glucanase	3.2.1.6	1	1	1	1	1		1
	Oligo-1,6-glucosidase	3.2.1.10	0	1	2	0	3		1
	$\beta$ -glucosidase	3.2.1.21	11	11	13	14	15		11
	Cellulose $\beta$ -D-glucanase ; reducing end	3.2.1.91	1	1	1	1	1		1
	Cellulose $\beta$ -D-glucanase ; non-reducing end	3.2.1.176	1	1	1	1	1		1
$\beta$ -glucanases	Glucoamylase	3.2.1.39	9	9	10	10	12		9
	Glucoamylase	3.2.1.58	4	4	9	6	7		4
	Glucoamylase	3.2.1.75	2	2	2	2	2		1
Hemicellulose degrading	Endo- $\beta$ -1,4-xylanase	3.2.1.8	3	4	6	5	7		7
	Polygalacturonase	3.2.1.15	0	1	3	2	2		3
	$\alpha$ -galactosidase	3.2.1.22	4	3	5	6	5		4
	$\beta$ -galactosidase	3.2.1.23	1	1	2	2	2		1
	$\alpha$ -mannosidase	3.2.1.24	1	1	2	1	1		1
	$\beta$ -mannosidase	3.2.1.25	5	4	6	6	4		3
	Xylotriase	3.2.1.37	3	3	4	3	3		3
	$\alpha$ -N-acetylglucosaminidase	3.2.1.55	3	1	5	5	6		4
	Galacturonylase	3.2.1.67	1	1	3	3	2		1
	Mannosylase	3.2.1.78	1	8	1	2	2		1
	Mannosyl-oligosaccharide glucosidase	3.2.1.106	1	2	1	1	1		1
	Mannosyl-oligosaccharide 6,7- $\alpha$ -mannosidase	3.2.1.113	8	0	8	8	8		7
	Xylotriase	3.2.1.131	2	2	3	3	3		2
	Galactosylase	3.2.1.145	0	0	0	1	1		1
	Xyloglucanase	3.2.1.151	1	1	2	2	3		1
	Galactosylase	3.2.1.164	1	1	1	1	1		1
	$\alpha$ -D-xylofuranosylase	3.2.1.177	1	1	2	2	2		2
Chitin degrading	Chitinase	3.2.1.14	14	17	28	23	25		18
	Chitosanase	3.2.1.132	3	3	4	6	5		1
	Endo- $\beta$ -1,4-N-acetylglucosaminidase	3.2.1.165	1	1	1	1	1		1
Others	AA9		2	3	3	2	2		4
	Endo- $\alpha$ -sialidase	3.2.1.18	0	0	0	0	0		1
	$\beta$ -fructofuranosylase	3.2.1.26	0	0	0	1	3		1
	$\alpha$ -trehalase		4	4	4	4	4		3
	$\beta$ -glucuronidase	3.2.1.31	1	2	2	2	2		0
	$\alpha$ -L-rhamnosidase	3.2.1.40	0	0	1	2	2		1
	Glucosylceramidase	3.2.1.45	1	1	1	1	1		1
	$\alpha$ -N-acetylglucosaminidase	3.2.1.50	1	1	0	0	0		0
	$\beta$ -N-acetylglucosaminidase	3.2.1.52	3	3	3	3	3		3
	Glucoamylase	3.2.1.59	3	3	4	2	3		3
	6,7- $\alpha$ -L-fucosidase	3.2.1.63	3	4	4	4	4		1
	Fructofuranosylase	3.2.1.80	0	0	0	0	1		0
	Total		110	118	164	150	166		123

**Table 3:** Comparison of GHs found in publically available genomes of *Trichoderma* species based on functions predicted by Peptide Pattern Recognition. In the left column are the functions predicted by PPR and below the species are the number of genes present in the individual species with this specific function.

## Discussion

The production of CAZymes by *Trichoderma* grown on different substrates illustrated that the choice of substrate is important for the successful production of CAZymes, because biomasses induce a different fungal enzyme response depending on structure and composition. Several *Trichoderma* species have been successfully cultivated on various lignocellulosic substrates, and their CAZymes analyzed (Hölker, Höfer, & Lenz, 2004; Singhania, Patel, Soccol, & Pandey, 2009). *T. asperellum* has shown its potential to secrete a range of GHs under appropriate culture conditions (Aires et al., 2012; Marx et al., 2013) and when grown on substrates such as cellulose (Raghuwanshi et al., 2014) and sugar bagasse (Marx et al., 2013). On the latter substrate *T. asperellum* showed promising results as an enzyme producer exhibiting a higher diversity of hemicellulases and  $\beta$ -glucosidases compared with *T. reesei*. In addition *T. viride* exhibited enhanced cellulase production when grown on wheat bran (Lan et al., 2013) as has *T. reesei* grown on substrates mixed with wheat bran (Brijwani, Oberoi, & Vadlani, 2010; Smits, Rinzema, Tramper, VanSonsbeek, & Knol, 1996). These studies indicate that several experimental parameters influence enzyme yields, including incubation time, extraction methods, and substrate loading. Other factors that improved cellulase production by *T. reesei* during solid state fermentation included relative humidity and temperature, continuous light exposure, and aeration (Mekala, Singhania, Sukumaran, & Pandey, 2008). In our study a range of enzyme activities were assayed by AZCL to evaluate the effect of specific growth substrate

on the production of CAZymes by *T. asperellum*. The analysis revealed that substrates induced different enzyme responses from *T. asperellum* and that more complex sugars such as wheat bran induced a higher and wider activity range than the simple sugars found in e.g. PDA medium, which induced no activity. Beside the sugar complexity, enzyme production also seemed to be dependent on the degree of pretreatment and hence the accessibility to the sugar structure. This may for example be the case with the enzyme response for wheat straw compared to wheat bran; a relatively mild pretreatment of the wheat straw may not have opened up its structure and allowed access to more complex sugars, and this therefore resulted in the production of a relatively simple enzyme cocktail by *T. asperellum*. Processed wheat bran on the other hand seemed to have the right composition for allowing a broad range of CAZymes to be produced at the same pretreatment level as wheat straw.

## CAZymes in the transcriptome of *T. asperellum* grown on wheat bran

To elucidate the enzymatic response on a transcript level, the transcriptome of *T. asperellum* was sequenced and glycoside hydrolases were identified the results. When *T. asperellum* was grown on wheat bran, 175 GHs from 48 GH families were induced. The genomes of phylogenetically similar fungi like *T. atroviride* encode 213 GHs (Xie et al., 2014), meaning that around 80% of the GHs present in the genome can be induced by growing *T. asperellum* on wheat bran. The most

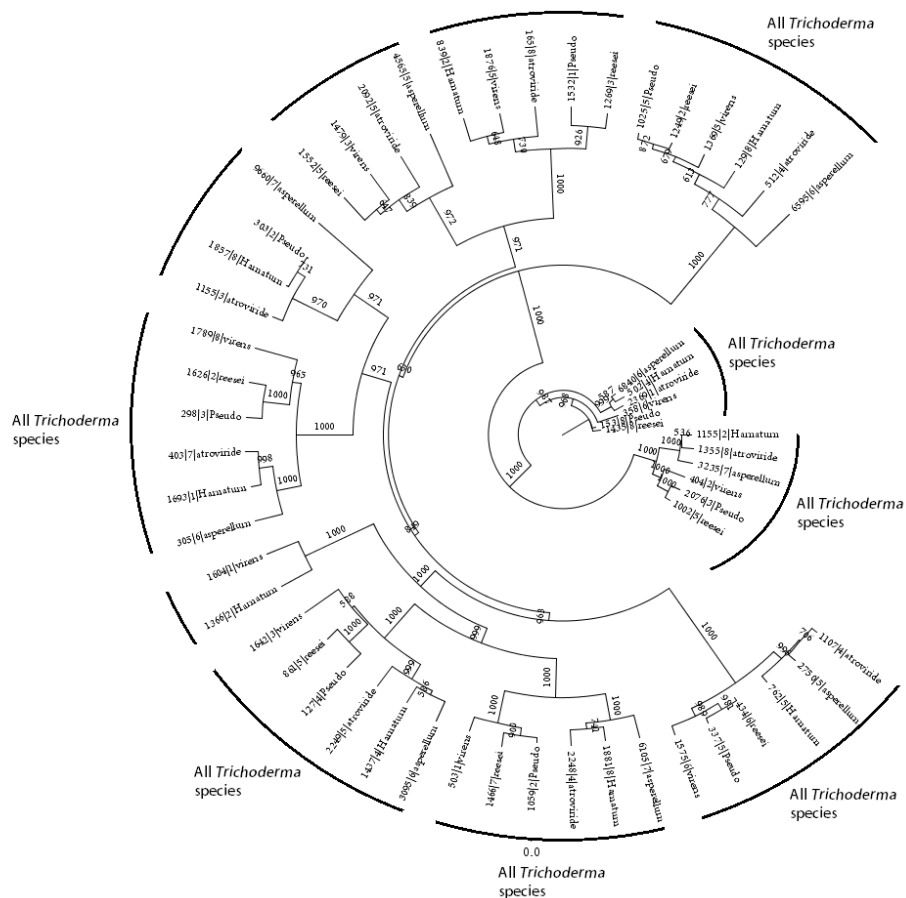


Figure 3: Phylogenetic tree of predicted glucan endo-1,3- $\beta$ -D-glucosidases (E.C. 3.2.1.39) present in the featured *Trichoderma* species.



represented family was GH18, which was predicted to be chitinases. This response corresponds to other experiments using *Trichoderma* species that exhibit a high production of chitinases (Marx et al., 2013; Tseng, Liu, Yang, Lo, & Peng, 2008). Moreover, several GH16-encoding genes were transcribed. This family contains enzymes with different functions, but all of the expressed GH16s were predicted to be  $\beta$ -glucanases which, like chitinases, are reported to be involved in mycoparasitic activities (Tseng et al., 2008). Compared with a similar experiment with *Trichoderma harzianum* (Uversky et al., 2014), GH16 genes were upregulated when growing on sugar bagasse compared with when growing on cellulose and lactose, while a GH16 glucan endo-1,3(4)- $\beta$ -glucosidase from *T. reesei* has been reported to be highly induced on cellulose (dos Santos Castro et al., 2014). A large transcription of GH5 genes was also observed. The genes from GH5 were predicted to contain several functions, including endo- $\beta$ -1,6-galactanase, endo-1,4-glucanase,  $\beta$ -1,6-glucanase,  $\beta$ -mannanase and xylanase. By contrast, the *T. harzianum* paper (Uversky et al., 2014) showed that each growth substrate induced its specific GH5 enzyme. This was possibly due to the simpler structure of substrates in that study compared to wheat bran, and hence fewer enzyme types were needed to bring about degradation. A more in-depth analysis of the enzymes is needed to identify which function is related to which growth substrate. Interestingly, the *Trichoderma* species contain only a few  $\alpha$ -glucosidases, making them the smallest group, but  $\alpha$ -glucosidases was seen to be the most transcribed enzyme group by *T. asperellum*, when grown on wheat bran. This may be due to the relatively simple structure of starch, hence needing fewer different enzymes compared to hemicellulose, which due to its more complex structure needs a more complex enzyme cocktail for degradation.

### Transcriptome derived secretome

To analyze which secreted GHs were transcribed in *T. asperellum*, the sequenced GHs were analyzed for the presence of a signal peptide. **The GH families present in the transcriptome derived secretome were mostly represented by one enzyme per family, but four families exhibited a higher representation: there were five enzymes from GH16, primarily  $\beta$ -glucanases; GH18, primarily chitinases; GH76,  $\alpha$ -glucosidases; and GH3,  $\beta$ -glucosidases and xylosidases. Besides being represented by more enzymes than the other families, these were also among the most transcribed genes.** Typically, enzyme blends from *Trichoderma* are low in  $\beta$ -glucosidase activity and are supplemented with  $\beta$ -glucosidases by co-cultivation with other high  $\beta$ -glucosidase-producing fungi to efficiently hydrolyze sugar structure (Singhania, Patel, Sukumaran, Larroche, & Pandey, 2013). The enhanced  $\beta$ -glucosidase activity of *T. asperellum* corresponds to other research where this fungus has been grown on sugar bagasse (Marx et al., 2013). But a relatively large part of the transcriptome derived secretome by number of genes was chitinases and  $\beta$ -glucanases. The *Trichoderma* genus has been reported to produce several  $\beta$ -glucanases and chitinases, which may either be involved in rearranging of the fungus' own cell wall or as a constant anti-microbial response (Ike et al., 2006; Limón & Codón, 2004; Tseng et al., 2008; Uversky et al., 2014). This indicates that wheat bran induces a significant number of anti-microbial enzymes and not just a CAZyme response corresponding with its biomass composition. This might provide the fungi with an advantage in nature, but could represent an energy waste for industrial biomass degradation.

### Composition of CAZymes across six *Trichoderma* species

A drawback of the current classification of GHs in GH families is that each family consists of members with different functions. When

a number of enzymes in a GH family have been characterized, the PPR program is able to predict the function of other family members identified only by their sequence. PPR is a non-alignment based method for identifying conserved sequence motifs in biological sequences (for example peptides in proteins) (Busk & Lange, 2013). This provides an opportunity to analyze the fungi in more depth based on GHs and corresponding functions present in the genomes. In the case of family GH76, only one enzyme is characterized, thus PPR is unable to make a reliable prediction. Due to incomplete information about enzymatic properties of the GHs, PPR was only able to predict the function of about 80% of the GHs. The non-identified GHs were subsequently used in Blast search and their conserved domains analyzed to predict their function. Checking the unidentified GHs limits the possibility of them containing functions described in table 3. Based on the PPR analysis *T. atroviride*, *T. hamatum* and *T. virens* contained significantly more GHs than *T. reesei* and *T. longibrachiatum*. In the case of *T. hamatum* PPR identified almost 50% more GHs than in *T. reesei*. This difference was mostly due to a higher number of hemicellulases and chitinases, but *T. reesei* and *T. longibrachiatum* also contained fewer enzymes across all major enzyme groups such as starch degrading cellulases and hemicellulases compared to its mycoparasitic relatives. This was not only the case for enzymes with many copies; based on PPR *T. reesei* and *T. longibrachiatum* have lost 9-10 activities. These enzymes belong to a wide range of GH families and are involved in starch, cellulose and hemicellulose degradation. The consequence of this function loss is unclear because *T. reesei* is regarded as a relatively good lignocellulosic degrader. The overall picture of the PPR analysis corresponds to the current theory of mycoparasitic species, which suggests that the mycotroph-related genes arose in the common ancestor of *Trichoderma* that had the ancestral mycotrophic life style, and that some of these genes were subsequently lost in saprotrophic *T. reesei* (Kubicek et al., 2011). This supposedly happened as *T. reesei* became an efficient saprotroph on dead wood by following wood-degrading fungi into their habitat (Rossman, Samuels, Rogerson, & Lowen, 1999; Xie et al., 2014). These findings are also reflected in the phylogenetic comparison that shows a close relation between *T. reesei* and *T. longibrachiatum* and a more distant relation to the other species.

### Phylogenetic relationship on a single gene level

Phylogenetic trees of the enzymes illustrate clades with similar genes across the species, placing the individual enzymes according to their greater species phylogenetic relationship (Kullnig-Gradinger et al., 2002). This supports the theory that CAZymes in the *Trichoderma* species have evolved from a common ancestor, followed by an evolutionary process, which resulted in especially the saprotrophic species losing genes (Kubicek et al., 2011). This created smaller clades, where only homologues from some of the *Trichoderma* species were present. All *Trichoderma* species contain 6-7 enzymes predicted to be from GH76, and are thus relatively unchanged by the loss of GHs in saprotrophic species. This indicates that these enzymes may have a key role in the degradation of complex biomass, which is also reflected in the high representation of GH76 family in the secretome of *T. asperellum* when grown on wheat bran. To predict the function of the GH76s with greater accuracy, more members of the family needs to be characterized. One characterized gene from GH76 is a  $\alpha$ -1,6-mannanase (Maruyama & Nakajima, 2000), which shows close similarity to one of the predicted GH76 from *T. asperellum*. With respect to enzyme discovery, the transcriptome sequencing has elucidated the most transcribed GHs on a specific substrate, which is an indication of which enzymes are important and potentially effective for degradation. PPR on the other hand elucidated the differences between GHs represented



in the genomes and revealed specific enzymes and functions for single species within genus *Trichoderma*.

## Conclusion

When *T. asperellum* was grown on wheat bran, 175 GHs from 48 GH families were transcribed, corresponding to more than 80% of the GHs compared with the closely related *T. atroviride*. Based on number of raw reads, 90% of enzymes bound for secretion, which are used by the fungi use to degrade biomass, consisted of 35 different enzymes divided on 20 different families, indicating that 40% of the glycoside hydrolase families are represented in the secretome. Beside the main carbohydrate degrading enzymes, *T. asperellum* also transcribed several chitinases and  $\beta$ -glucanases for secretion. This result is suggested to be either for rearrangement of chitin of the fungus itself or a component in anti-microbial action. More detailed insight was gained into the evolutionary differences among *Trichoderma spp.* by comparing the glycoside hydrolases present in different species. It was shown that the genetically similar *T. atroviride*, *T. virens*, *T. hamatum* and *T. asperellum* contain a larger and broader repertoire of glycoside hydrolases compared to the genetically similar *T. reesei* and *T. longibrachiatum*. This is additionally supported by the phylogenetic relationship of individual enzymes that form clades of homologues with distances corresponding to the overall phylogenetic relationship between the species. Analysis of enzymes on a functional level instead of on a broader family level allowed the identification of specific functions lost in *T. reesei* and *T. longibrachiatum* compared to *T. atroviride*, *T. virens*, *T. hamatum* and *T. asperellum*, as well as unique functions present in certain species. The latter can be of importance for the discovery of novel glycoside hydrolases involved in plant cell wall degradation.

## Acknowledgements

This work was supported by the Sino-Danish center and Novozymes A/S.

## References

- Aires, R. d. S., Steindorff, A. S., Soller Ramada, M. H., Linhares de Siqueira, S. J., & Ulhoa, C. J. (2012). Biochemical characterization of a 27 kDa 1,3- $\beta$ -glucanase from *Trichoderma asperellum* induced by cell wall of *Rhizoctonia solani*. *Carbohydrate Polymers*, 87(2), 1219-1223.
- Borrero, C., Trillas, M. I., Delgado, A., & Aviles, M. (2012). Effect of ammonium/nitrate ratio in nutrient solution on control of fusarium wilt of tomato by *Trichoderma asperellum* T34. *Plant Pathology*, 61(1), 132-139.
- Bozell, J. J., & Petersen, G. R. (2010). Technology development for the production of biobased products from biorefinery carbohydrates—the US department of energy's "top 10" revisited. *Green Chemistry*, 12(4), 539-554.
- Brijwani, K., Oberoi, H. S., & Vadlani, P. V. (2010). Production of a cellulolytic enzyme system in mixed-culture solid-state fermentation of soybean hulls supplemented with wheat bran. *Process Biochemistry*, 45(1), 120-128.
- Busk, P. K., & Lange, L. (2013). Function-based classification of carbohydrate-active enzymes by recognition of short, conserved peptide motifs. *Applied and Environmental Microbiology*, 79(11), 3380-3391.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The carbohydrate-active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research*, 37, D233-D238.
- Chundawat, S. P. S., Beckham, G. T., Himmel, M. E., & Dale, B. E. (2011). Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annual Review of Chemical and Biomolecular Engineering*, Vol 2, 2, 121-145.
- dos Santos Castro, L., Pedersoli, W. R., Antoniêto, A. C. C., Steindorff, A. S., Silva-Rocha, R., Martinez-Rossi, N. M. et al. (2014). Comparative metabolism of cellulose, sophorose and glucose in *Trichoderma reesei* using high-throughput genomic and proteomic analyses. *Biotechnology for Biofuels*, 7(1), 1-18.
- Gerlt JA, Allen KN, Almo SC, Armstrong RN, Babbitt PC, Cronan JE, Dunaway-Mariano D, Imker HJ, Jacobson MP, Minor W (2011) The enzyme function initiative. *Biochemistry* 50: 9950-9962
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I. et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, 29(7), 644-652.
- Häkkinen, M., Arvas, M., Oja, M., Aro, N., Penttilä, M., Saloheimo, M., & Pakula, T. M. (2012). Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. *Microb Cell Fact*, 11(1), 134.
- Henrissat, B., & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *The Biochemical Journal*, 293 ( Pt 3)(Pt 3), 781-788.
- Henrissat, B., & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *The Biochemical Journal*, 316(2), 695-696.
- Henrissat, B., Teeri, T. T., & Warren, R. A. (1998). A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Letters*, 425(2), 352-354.
- Himmel, M. E., Ding, S., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*, 315(5813), 804-807.
- Hölker, U., Höfer, M., & Lenz, J. (2004). Biotechnological advantages of laboratory-scale solid-state fermentation with fungi. *Applied Microbiology and Biotechnology*, 64(2), 175-186.
- Ike, M., Nagamatsu, K., Shioya, A., Nogawa, M., Ogasawara, W., Okada, H., & Morikawa, Y. (2006). Purification, characterization, and gene cloning of 46 kDa chitinase (Chi46) from *Trichoderma reesei* PC-3-7 and its expression in *Escherichia coli*. *Applied Microbiology and Biotechnology*, 71(3), 294-303.
- Iseli, C., Jongeneel, C. V., & Bucher, P. (1999). ESTScan: A program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *ISMB*, 99 138-147.
- Kiyohara, M., Tanigawa, K., Chaiwangsri, T., Katayama, T., Ashida, H., & Yamamoto, K. (2011). An exo-alpha-sialidase from bifidobacteria involved in the degradation of sialyloligosaccharides in human milk and intestinal glycoconjugates. *Glycobiology*, 21(4), 437-447.
- Kubicek, C. P., Herrera-Estrella, A., Seidl-Seiboth, V., Martinez, D. A., Druzhinina, I. S., Thon, M. et al. (2011). Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biology*, 12(4).
- Kullnig-Gradinger, C., Szakacs, G., & Kubicek, C. (2002). Phylogeny and evolution of the genus *Trichoderma*: A multigene approach. *Mycological Research*, 106, 757-767.
- Lan, T., Wei, D., Yang, S., & Liu, X. (2013). Enhanced cellulase production by *Trichoderma viride* in a rotating fibrous bed bioreactor. *Bioresource Technology*, 133, 175-182.
- Limón, M. C., & Codón, A. C. (2004). Biocontrol mechanisms of *Trichoderma* strains. *International Microbiology*, 7, 249-260.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research*, 42(Database issue), D490-5.
- Maruyama, Y., & Nakajima, T. (2000). The aman6 gene encoding a yeast mannan backbone degrading 1, 6- $\alpha$ -mannanase in *Bacillus circulans*: Cloning, sequence analysis, and expression. *Bioscience, Biotechnology, and Biochemistry*, 64(9), 2018-2020.
- Marx, I. J., van Wyk, N., Smit, S., Jacobson, D., Viljoen-Bloom, M., & Volschenk, H. (2013). Comparative secretome analysis of *Trichoderma asperellum* S4F8 and *Trichoderma reesei* rut C30 during solid-state fermentation on sugarcane bagasse. *Biotechnology for Biofuels*, 6, 172.
- Mbarga, J. B., Ten Hoopen, G. M., Kuate, J., Adiobo, A., Ngonkeu, M. E. L., Ambang, Z., et al. (2012). *Trichoderma asperellum*: A potential biocontrol agent for *Pythium myriotylum*, causal agent of cocoyam (*Xanthosoma sagittifolium*) root rot disease in Cameroon. *Crop Protection*, 36, 18-22.
- Mekala, N. K., Singhania, R. R., Sukumaran, R. K., & Pandey, A. (2008). Cellulase production under solid-state fermentation by *Trichoderma reesei* RUT C30: Statistical optimization of process parameters. *Applied Biochemistry and Biotechnology*, 151(2-3), 122-131.
- Merino, S. T., & Cherry, J. (2007). Progress and challenges in enzyme development for biomass utilization. *Biofuels* (95-120) Springer.

30. Molinuevo-Salces, B., Fernández-Varela, R., & Uellendahl, H. (2014). Key factors influencing the potential of catch crops for methane production. *Environmental Technology*, 35(13), 1685-1694.
31. Monte, E. (2010). Understanding trichoderma: Between biotechnology and microbial ecology. *International Microbiology*, 4(1), 1-4.
32. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods*, 5(7), 621-628.
33. Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S. et al. (2003). TIGR gene indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinformatics (Oxford, England)*, 19(5), 651-652.
34. Popper, Z. A. (2008). Evolution and diversity of green plant cell walls. *Current Opinion in Plant Biology*, 11(3), 286-292.
35. Ragauskas, A., Williams, C., Davison, B., Britovsek, G., Cairney, J., Eckert, C. et al. (2006). The path forward for biofuels and biomaterials. *Science*, 311(5760), 484-489.
36. Raghuvanshi, S., Deswal, D., Karp, M., & Kuhad, R. C. (2014). Bioprocessing of enhanced cellulase production from a mutant of *Trichoderma asperellum* RCK2011 and its application in hydrolysis of cellulose. *Fuel*, 124, 183-189.
37. Rossman, A. Y., Samuels, G. J., Rogerson, C. T., & Lowen, R. (1999). *Genera of Bionectriaceae, Hypocreaceae and Nectriaceae (hypocreales, ascomycetes)*. Centraalbureau voor Schimmelcultures (CBS).
38. Schnoes AM, Brown SD, Dodevski I, Babbitt PC (2009) Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. *PLoS computational biology* 5: e1000605
39. Schuster, A., & Schmoll, M. (2010). Biology and biotechnology of *Trichoderma*. *Applied Microbiology and Biotechnology*, 87(3), 787-799.
40. Singhanian, R. R., Patel, A. K., Soccol, C. R., & Pandey, A. (2009). Recent advances in solid-state fermentation. *Biochemical Engineering Journal*, 44(1), 13-18.
41. Singhanian, R. R., Patel, A. K., Sukumaran, R. K., Larroche, C., & Pandey, A. (2013). Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. *Bioresource Technology*, 127, 500-507.
42. Smits, J., Rinzema, A., Tramper, J., VanSonsbeek, H., & Knol, W. (1996). Solid-state fermentation of wheat bran by *Trichoderma reesei* QM9414: Substrate composition changes, C balance, enzyme production, growth and kinetics. *Applied Microbiology and Biotechnology*, 46(5-6), 489-496.
43. Sun, X., Liu, Z., Qu, Y., & Li, X. (2008). The effects of wheat bran composition on the production of biomass-hydrolyzing enzymes by *Penicillium decumbens*. *Applied Biochemistry and Biotechnology*, 146(1-3), 119-128.
44. Tseng, S., Liu, S., Yang, H., Lo, C., & Peng, K. (2008). Proteomic study of biocontrol mechanisms of *Trichoderma harzianum* ETS 323 in response to *Rhizoctonia solani*. *Journal of Agricultural and Food Chemistry*, 56(16), 6914-6922.
45. Uversky, V. N., Horta, M. A. C., Vicentini, R., Delabona, P. d. S., Laborda, P., Crucello, A. et al. (2014). Transcriptome profile of *Trichoderma harzianum* IOC-3844 induced by sugarcane bagasse. *PLoS ONE*, 9(2), e88689.
46. Vazquez-Garciduenas, S., Leal-Morales, C. A., & Herrera-Estrella, A. (1998). Analysis of the beta-1,3-glucanolytic system of the biocontrol agent *Trichoderma harzianum*. *Applied and Environmental Microbiology*, 64(4), 1442-1446.
47. Xie, B. -, Qin, Q. -, Shi, M., Chen, L. -, Shu, Y. -, Luo, Y., . . . Song, X. -. (2014). Comparative genomics provide insights into evolution of *Trichoderma* nutrition style. *Genome Biology and Evolution*, 6(2), 379 -390.