Analysis and optimisation of plant biomass degrading enzyme production in *Aspergillus*

Helena Marie Culleton

Analysis and optimisation of plant biomass degrading enzyme production in *Aspergillus*

Analyse en optimalisatie van de productie van planten biomassa afbrekende enzymen in *Aspergillus*
(met een Nederlandse samenvatting)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 26 februari 2015 des middags te 12.45 uur

door

Helena Marie Culleton

geboren op 3 april 1986 te Wexford, Ireland

Promotor: Prof. Dr. ir. R.P. de Vries

Co-promotor: Dr. V.A. McKie

For my parents and family

The *Aspergillus niger* image on the cover was kindly provided by;
Dr. Nick Reid,
Professor of Fungal Cell Biology,
Director, Manchester Fungal Infection Group,
Institute of Inflammation and Repair,
University of Manchester,
CTF Building,
Grafton Street,
Manchester M13 9NT.

Printed by Snap ™ Printing, www.snap.ie

The research described in this thesis was performed in;
Megazyme International Ireland, Bray Business Park, Bray, Co. Wicklow, Ireland;
Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands;
CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands;
and supported by Megazyme International Ireland, Bray Business Park, Bray, Co. Wicklow, Ireland.

Contents

Chapter 1

General Introduction

Part of this chapter was published in:

Culleton, H.M., McKie, V.A. and de Vries, R.P. 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*? Biotechnology Journal. **8**: 884-894.

Filamentous Fungi

Fungi form a large and diverse group of eukaryotic organisms. To date, about 80,000 fungal species have been described, with the actual number world-wide being estimated at 1.5 million [57]. Fungi have the ability to grow in many diverse and often harsh environments, including extreme heat [73], extreme acidity [107], high pressure [96] and radiation [27]. The major recognized fungal groups are Chytridiomycota, Zygomycota, Glomeromycota, Ascomycota and Basidiomycota [70], with the majority of described species belonging to the latter two phyla. Fungi can be unicellular (yeast) or multicellular (filamentous), but some species have the ability to switch between both forms (dimorphic fungi) [14]. Filamentous fungi are chemoheterotrophic and depend on the presence of organic material for their carbon and energy source. To facilitate this, filamentous fungi grow by means of extending hyphae which branch sub-apically to form a network of hyphae called a mycelium. This hyphal mode of growth supplies a large surface area for the uptake of nutrients but also enables the fungus to colonize its substrate in an efficient manner [52]. The organic material, which acts as a substrate for most fungal species is generally composed of large, complex molecules which need to be broken down before they can serve as a source of nutrition. To this end, filamentous fungi secrete diverse enzymatic mixtures, tailor-made to the polysaccharide which they encounter.

The genus *Aspergillus*

The genus *Aspergillus* is a group of ascomycete filamentous fungi, consisting of more than 250 species. Of these species, the majority are saprotrophic fungi that live off dead or decaying organic matter but there are some e.g. *A. fumigatus* and *A. chevalieri*, which are opportunistic human pathogens [67, 79] and others e.g. *A. niger* and *A. flavus* which are phytopathogens [16, 48]. Some *Aspergillus* species e.g. *A. flavus* and *A. parasiticus*, are known to produce mycotoxins [25, 58], others e.g. *A. niger* and *A. oryzae*, are extensively used in industry for the production of enzymes and metabolites [52] while *A. nidulans* is routinely used as a model organism for lower eukaryotes [50]. The diversity of topics that are relevant to *Aspergillus* has resulted in it becoming one of the most widely studied groups of filamentous fungi. Combined with the fact that many Aspergilli have good fermentation capabilities and high levels of protein secretion [30], it is of no surprise that this fungus has been applied in industries such as food [94] and feed [13], pulp and paper [22], biofuels, biodegradable plastics, textiles [9] and as hosts for heterologous protein production [28]. As common soil fungi, found in many different environments, Aspergilli are also found to be capable of producing an extensive set of enzyme mixes that degrade a very broad range of polysaccharides [23].

Plant biomass degrading enzymes and their substrates

Plant biomass is the most abundant and universally used carbon source by most fungal species on earth. Primarily consisting of plant cell walls, it contains a complex structure of polysaccharides, proteins and lignin which differ greatly in monomeric composition and linkage. Fungi, including *Aspergillus*, degrade these polysaccharides extracellularly by secreting enzymes which release oligo- and monosaccharides that can then be utilized by the fungal cell [52]. Plant polysaccharides can be divided into two main groups; plant cell wall polysaccharides (cellulose, hemicelluloses [xyloglucans, xylan, galacto(gluo)mannan] and pectin) and storage polysaccharides (e.g. starch and inulin.) The plant cell wall consists of the primary and secondary cell wall, which both differ in composition and function. The primary cell wall is rich in polysaccharides (~90% polysaccharides (cellulose, hemicelluloses and pectin) and ~10% proteins) and is formed during growth [3]. The secondary cell wall is deposited on the primary layer after cell elongation stops and consists mainly of cellulose, hemicelluloses and lignin [10]. These cell wall polysaccharides interact with each other as well as the aromatic polymer, lignin, to form a network of linkages and hydrogen bonds that give the plant cell wall its rigidity [34].

Due to the diverse and complex structure of plant biomass, fungi are required to produce a large range of enzyme activities to degrade these polysaccharides into their monomeric components [23, 34]. These enzymes can be divided into families based on the catalytic modules present in their corresponding amino acid sequence i.e. glycoside hydrolases (GH), glycosyl transferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE) and carbohydrate binding modules (CBM). The resulting Carbohydrate-Active enzyme database (CAZy-http://www.cazy.org) has become a powerful tool in providing insight into the carbohydrate potential of a fungus and supports the identification and prediction of the function of novel genes [17, 23]. A full description of the enzyme activities involved in plant polysaccharide degradation are detailed below and corresponding CAZy families and EC numbers are presented in Table 1.

Cellulose

Cellulose is the most abundant plant cell wall polysaccharide. It exists as highly ordered linear polymers of β-1,4-linked D-glucose residues which are bundled together in microfibrils via hydrogen bonds (Fig. 1) [45]. This polymeric structure gives the plant cell wall its strength and definition [34]. Four enzymes groups are involved in the biodegradation of cellulose: endoglucanases, cellobiohydrolases, β-glucosidases and *exo*-glucanases. *Endo*-glucanases and cellobiohydrolases hydrolyze cellulose into gluco-oligosaccharides and cellobiose, respectively. These oligosaccharides are then further degraded into D-glucose molecules by the action of β-glucosidases and *exo*-glucanases (Fig. 1, Table 1).

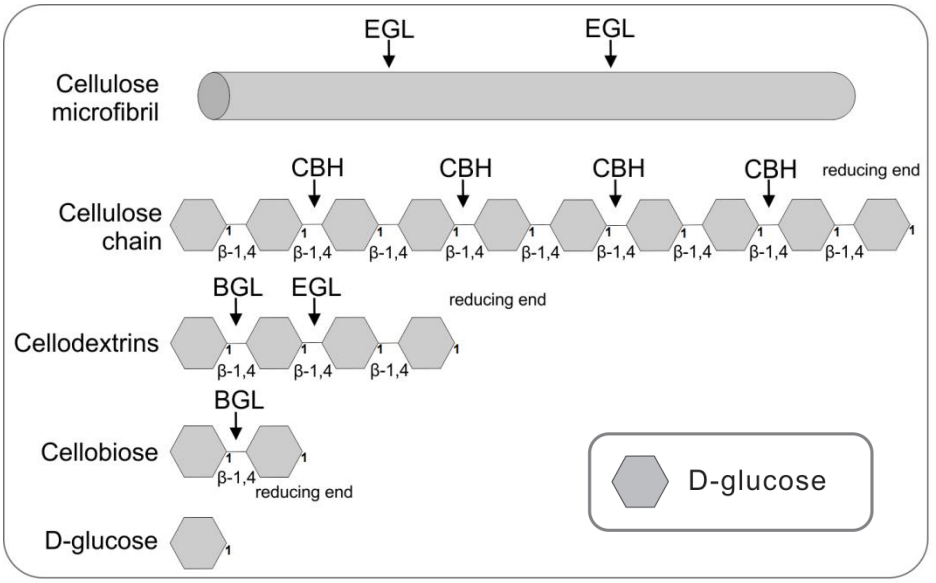


Figure 1: Representation of cellulose and its hydrolysis by cellulose degrading enzymes. The enzymes involved in the complete biodegradation of cellulose are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Hemicelluloses

The hemicellulose group of polysaccharides represent the second most abundant polysaccharide in plant cell walls and form cross-links with cellulose through hydrogen bonds to provide extra rigidity and protection against plant pathogens, such as fungi [20, 127]. Hemicellulose is present in three main forms in the plant cell wall: xylans, xyloglucan and mannans.

Xylan

The xylan group of polysaccharides consists of a backbone of β-1,4-linked D-xylose residues which can be acetylated and/or substituted with either single residues or short side chains of α-1,2- or α-1,3-linked L-arabinose (arabinoxylan) and/or single α-1,2-linked D-glucuronic acid residues (glucuronoxylan) [42, 125]. D-xylose can be found attached to the L-arabinose residues while feruloyl and *p*-coumaroyl residue substitutions can be found in the terminal L-arabinose residues. The D-xylose attached to the L-arabinose residues may be acetylated while methylation is only found in the D-glucuronic acid residues [106, 124] (Fig. 2). The main xylan present in cereals is arabinoxylan while hardwood contains mainly glucuronoxylan [34]. Due to its variable composition, the enzymatic mixture produced to break down xylan can be highly

varied [8, 15, 60, 103]. *Endo*-xylanases cleave the xylan backbone into smaller oligosaccharides which act as a substrate for degradation by *β*-xylosidases to D-xylose [93]. The L-arabinose side chain residues are hydrolyzed by *α*-arabinofuranosidases and arabinoxylan arabinofuranohydrolases, and the glucuronic acid residues by *α*-glucuronidases. The acetyl, feruloyl and *p*-coumaroyl are removed by the action of xylan acetylesterases, feruloyl esterases and *p*-coumaroyl esterases, respectively [93] (Fig. 2, Table 1).

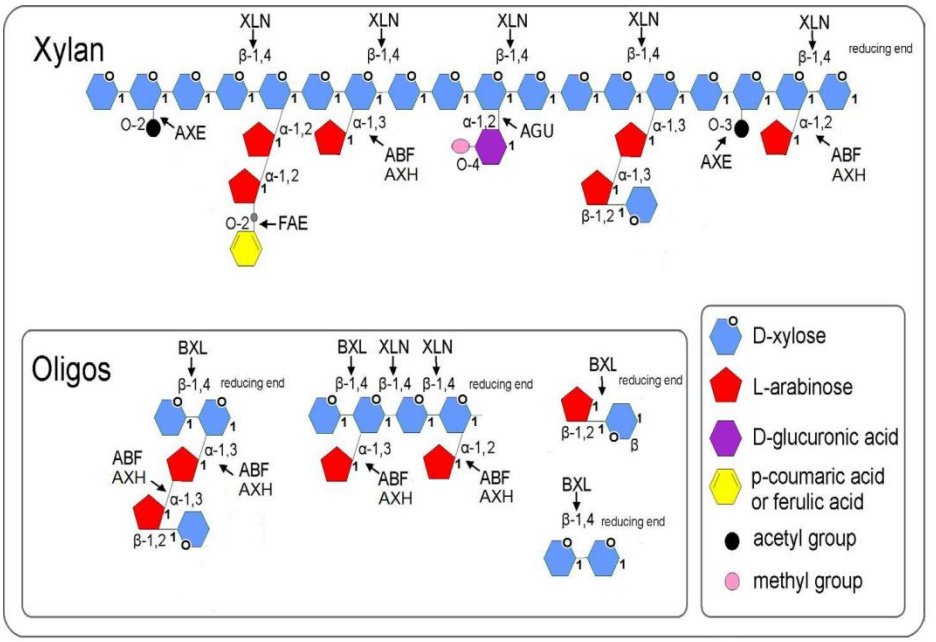


Figure 2: Representation of xylan and its hydrolysis by xylan degrading enzymes. The enzymes involved in the complete biodegradation of xylan are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Xyloglucan

Xyloglucan consists of *β*-1,4-linked D-glucose residues substituted by D-xylose. Two major groups of xyloglucans have been identified [119]. The XXXG-type xyloglucan contains repeating units of three *β*-1,4-linked D-glucose residues which are substituted with D-xylose residues through an *α*-1,6-linkage and are separated by an unsubstituted D-glucose residue (Fig. 3). The XXGG-type xyloglucan consists of two D-glucose residues substituted with *α*-1,6-linked D-xylose separated by two unsubstituted D-glucose residues [99]. The structural features of these structures have been discussed in detail by Vincken *et. al.* (1997) [119]. The D-xylose residues in xyloglucan can be

substituted with D-galactose via *β*-1,2-linkages and in some instances D-galactose has been further decorated with L-galactose or L-fucose via an *α*-1,2-linkage [46]. The D-xylose and unsubstituted D-glucose residues may also be substituted with L-arabinose via *α*-1,2-linkages, which in turn may be further decorated with *β*-1,2-linked D-xylose [59, 61, 129] (Fig. 3). As the basic backbone of xyloglucan is similar to that of cellulose, the same *endo*-glucanases and *β*-glucosidases can be responsible for hydrolyzing both polysaccharides although *endo*-glucanases which are specific to the xyloglucan backbone have also been reported (e.g. from *A. aculeatus*) [85]. Accessory enzymes for xyloglucan degradation include: *α*-xylosidases, *α*-/*β*-galactosidases and *α*-fucosidases. As with xylan, the L-arabinose side chain residues of xyloglucan can be removed by *α*-L-arabinofuranosidases. Xyloglucan acetylesterases are active only if one or more of the D-galactose residues are acetylated [34, 52] (Fig. 3, Table 1).

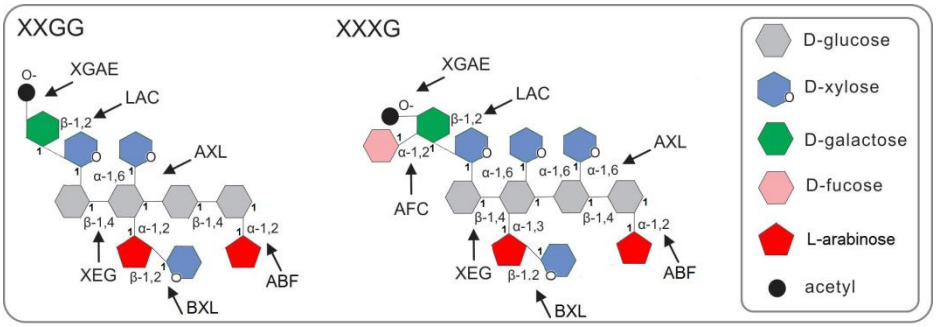


Figure 3: Representation of both the XXGG and XXXG forms of xyloglucan and its hydrolysis by xyloglucan degrading enzymes. The enzymes involved in the complete biodegradation of xyloglucan are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Mannan (galacto(gluco)mannan)

Mannan or galacto(gluco)mannan, consists of a *β*-1,4-linked D-mannose backbone, which can be interrupted by varying amounts of *β*-1,4-linked D-glucose residues and substituted with *α*-1,6-linked D-galactose residues. These D-galactose residues were found to be further decorated with *β*-1,2-linked D-galactose residues in the galacto(gluco)mannan from *Nicotiana plumbaginifolia*, although this is not common [105]. Water-soluble galactoglucomannan contains acetyl groups attached to the mannan backbone and has a higher degree of D-galactose substitution than water-insoluble galactoglucomannan [110] (Fig. 4). The backbone of galactoglucomannan is hydrolyzed to manno-, gluco-manno-, and galacto-gluco-mannooligosaccharides by *β*-endomannanases and *β*-mannosidases [75]. The complete hydrolysis of galactoglucomannan requires the additional action of *β*-glucosidases, *α*-galactosidases and galactomannan acetyl esterases [34, 74] (Fig. 4, Table 1).

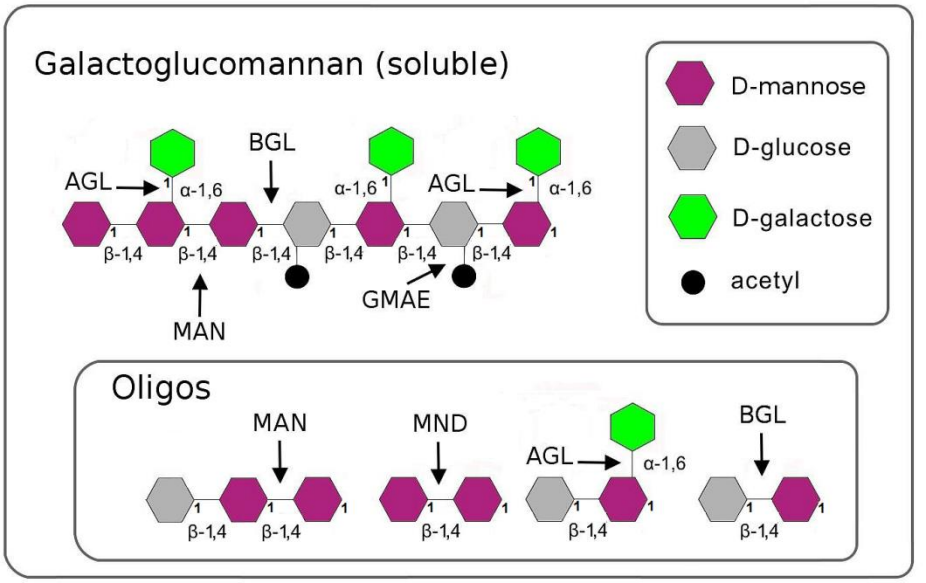


Figure 4: Representation of the water-soluble galactomannan structure and its hydrolysis by galactomannan degrading enzymes. The enzymes involved in the complete biodegradation of galactomannan are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Pectin

Pectin is also responsible for rigidity and definition and plays an important role in the physiology of the plant cell wall including porosity, pH, ion balance and surface charge [52, 126]. It contains several substructures: homogalacturonan (HGA), xylogalacturonan (XGA), rhamnoglacturonan I (RG-I) and rhamnogalacturonan II (RG-II) [126].

Homogalacturonan (HGA)

HGA or pectin's "smooth" region, consists of a linear chain of *α*-1,4-linked D-galacturonic acid residues which can be acetylated at O-2 or O-3 or methylated at O-6 [127] (Fig. 5). Pectin methyl and acetyl esterases act on this substrate to de-esterify the backbone and make it suitable for cross-linking to calcium molecules, thus forming a type of gel which plays an important role in intracellular adhesion [77]. HGA is cleaved by *endo*- and *exo*-polygalacturonases and/or pectin and pectate lyases while the acetyl and methyl groups are removed by pectin acetyl and methyl esterases [26] (Fig. 5, Table 1).

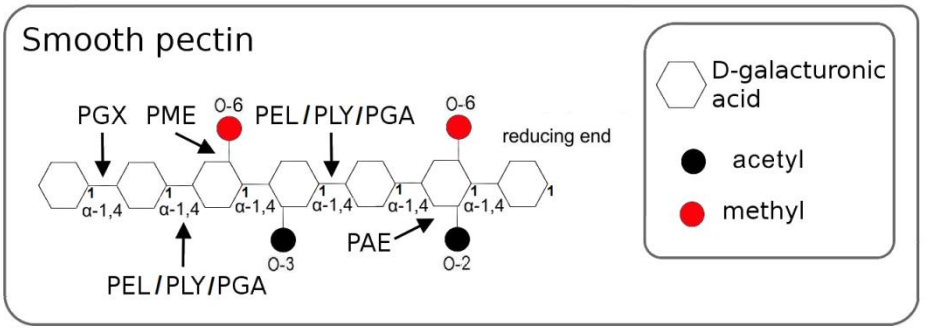


Figure 5: Representation of the homogalacturonan (HGA) structure and its hydrolysis by pectin degrading enzymes. The enzymes involved in the complete biodegradation of homogalacturonan (HGA) are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Xylogalacturonan (XGA) and Rhamnogalacturonan I (RG-I)

XGA and RG-I together make up the "hairy" region of pectin [34, 52]. XGA, like HGA, contains an *α*-1,4-linked D-galacturonic acid backbone which can be substituted with *β*-1,3-linked D-xylose [101], which in turn can be further decorated with *β*-1,2- or *β*-1,4-linked D-xylose (Fig. 6) [69, 80, 131]. RG-I contains an alternating backbone of *α*-1,4-linked D-galacturonic acid and *α*-1,2-linked L-rhamnose residues, the latter with long side chains of L-arabinose (arabinan), D-galactose (galactan) or mixture of both (arabinogalactan) [68] (Fig. 7 and 8). The main chain may also have ester-linked acetyl groups attached to the D-galacturonic acid residues [100, 102] (Fig. 7). The arabinan side chains consist of *α*-1,5-linked L-arabinose residues which can be substituted with *α*-1,3-linked L-arabinose and/or feruloyl residues [63]. The galactan chains exist as a linear chain of *β*-1,4-linked D-galactose residues which also can be substituted with feruloyl residues [63]. The arabinogalactan side chains can either consist of a backbone of *β*-1,4-linked galactan substituted with L-arabinose residues or *β*-1,3-linked galactan which can be substituted with either *β*-1,6-linked D-galactose or *α*-1,3-, *α*-1,5- and *α*-1,6-linked L-arabinose residues [63] (Fig. 8).

While many of the same enzyme activities act on HGA and XGA, some xylogalacturonases which are specific to XGA may also be present [113]. The *β*-linked D-xylose side chains of XGA are removed by *β*-xylosidases (Fig. 6, Table 1). RG-I is cleaved by endorhamnogalacturonan hydrolases, exorhamnogalacturonan hydrolases, rhamnogalacturonan lyases and *α*-rhamnosidases [34] while the acetyl groups are removed by rhamnogalacturonan acetyl esterases [33] (Fig. 7, Table 1). The side-chains of RG-I are hydrolyzed by endo-/exo-arabinases, arabinofuranosidases, *endo*-/*exo*-galactanases and *β*-galactosidases [34] (Fig. 8, Table 1).

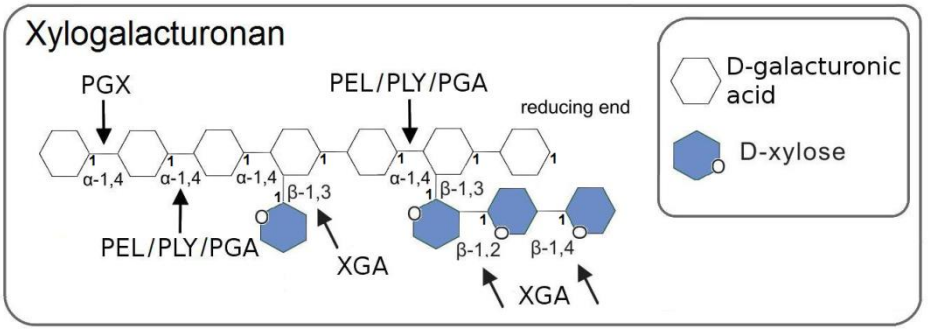


Figure 6: Representation of the xylogalacturonan (XGA) structure and its hydrolysis by pectin degrading enzymes. The enzymes involved in the complete biodegradation of xylogalacturonan (XGA) are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

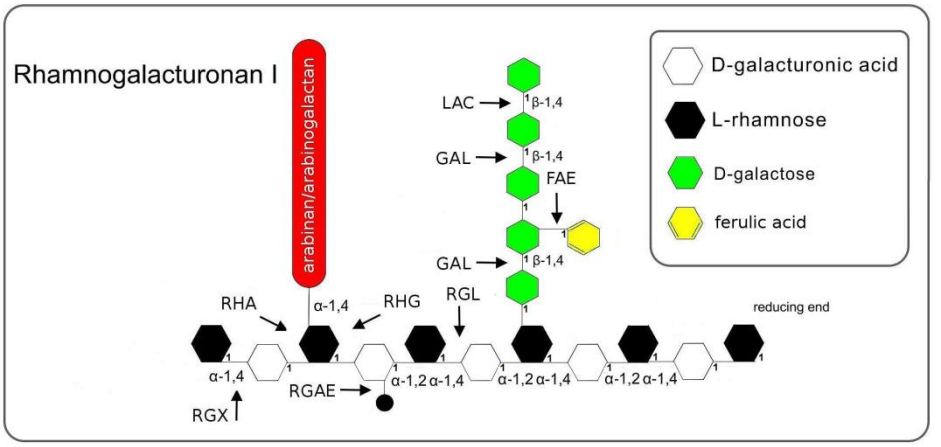


Figure 7: Representation of the rhamnogalacturonan I (RG-I) structure and its hydrolysis by pectin degrading enzymes. RG-I contains an alternating backbone of *α*-1,4-linked D-galacturonic acid and *α*-1,2-linked L-rhamnose residues, the latter to which long side chains of L-arabinose (arabinan), D-galactose (galactan) or mixture of both (arabinogalactan) can be attached (Fig. 8).The enzymes involved in the complete biodegradation of rhamnogalacturonan I (RG-I) are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

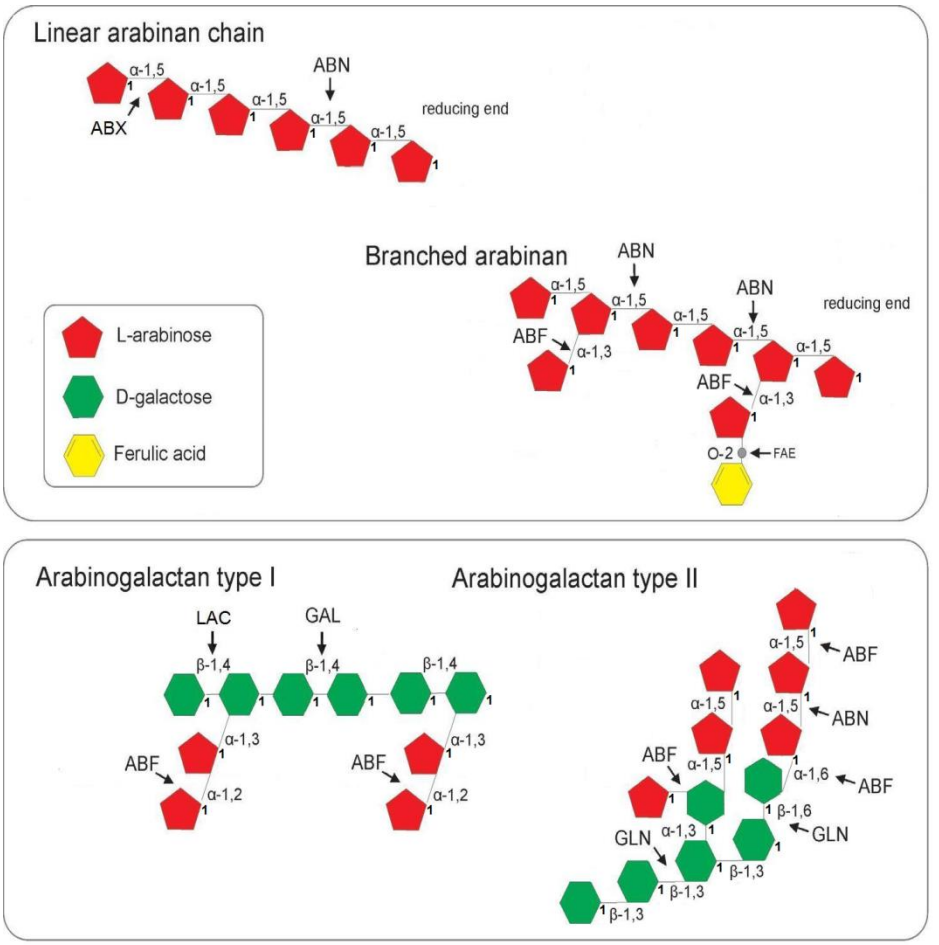


Figure 8: Representation of the rhamnogalacturonan I (RG-I) arabinan and arabinogalactan side chains and their hydrolysis by pectin degrading enzymes. The enzymes involved in the complete biodegradation of rhamnogalacturonan I (RG-I) are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Rhamnogalacturonan II (RG-II)

RG-II is a complex polysaccharide, consisting of approximately 30 monosaccharide units [34] with a backbone of eight or more D-galacturonic acid residues [118] and is substituted with up to five different side-chains [118]. These decorations may be of mono- or oligosaccharide structure and can consist of the following residues; L-rhamnose, D-galactose, L-arabinose, D-glucuronic acid, L-fucose, D-apiose, L-aceric acid, 2-O-methyl L-fucose, 2-O-methyl D-xylose, L-galactose, 2-keto-3-deoxy-D-lyxo-heptulosaric acid (Dha) and 2-keto-3-deoxy-D-manno-octulosonic acid (KDO) [118] (Fig. 9). Due to its extremely complex structure, enzymatic degradation of pectin requires a large set of activities (Table 1). The identification of the enzymes involved in the degradation of RG-II residue however, thus far remain unclear (Fig. 9).

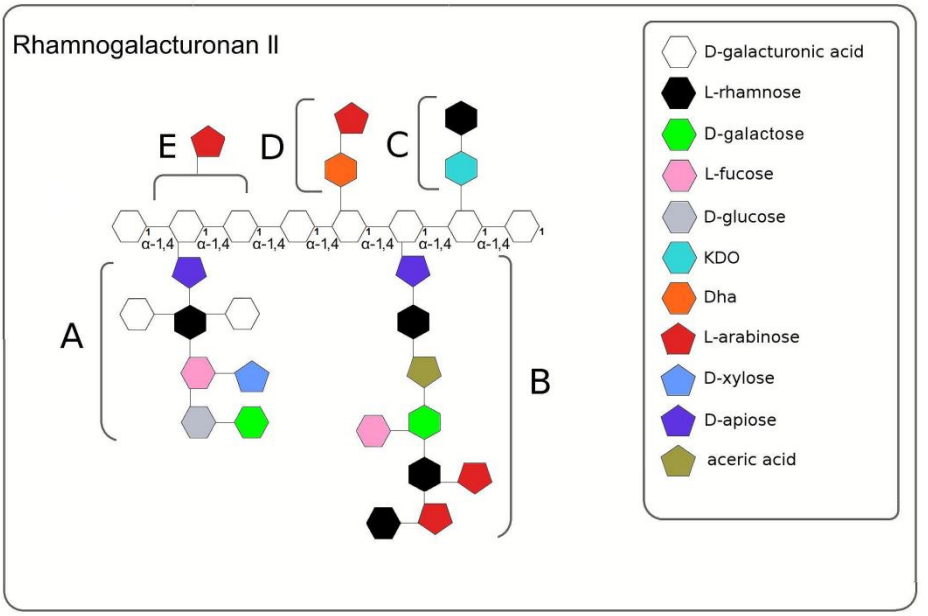


Figure 9: Representation of the rhamnogalacturonan II (RG-II) structure. A-E represent rhamnogalacturonan II (RG-II) side chains, with two possible positions for 'E' being indicated. The enzymes involved in complete hydrolysis of this polysaccharide remain elusive, and are therefore not included in the figure. The figure is based on [91].

Table 1: CAZymes involved in plant biomass degradation. Adapted from [23].

Substrate	Enzyme Class	Enz.	CAZy families	EC no.
Cellulose	*β*-1,4-D-glucosidase*	BGL	GH1,3	3.2.1.21
Cellulose	Cellobiohydrolase	CBH	GH6,7	3.2.1.91
Cellulose	*β*-1,4-D-endoglucanase	EGL	GH5,7,12,61	3.2.1.4
Xylan	*α*-L-arabinofuranosidase*	ABF	GH43,51,54	3.2.1.55
Xylan	*α*-glucuronidase	AGU	GH67,115	3.2.1.139
Xylan	acetyl xylan esterase	AXE	CE1	3.1.1.72
Xylan	arabinoxylan arabinofuranohydrolase	AXH	GH62	3.2.1.55
Xylan	*β*-1,4-D-xylosidase*	BXL	GH3,43	3.2.1.37
Xylan	feruloyl esterase*	FAE	CE1	3.1.1.73
Xylan	*β*-1,4-D-endoxylanase	XLN	GH10,11	3.2.1.8
Xyloglucan	*α*-L-arabinofuranosidase*	ABF	GH43,51,54	3.2.1.55
Xyloglucan	*α*-L-fucosidase	AFC	GH29,95	3.2.1.51
Xyloglucan	*α*-1,4-D-galactosidase*	AGL	GH27,36	3.2.1.22
Xyloglucan	*α*-D-xylosidase	AXL	GH31	3.2.1.-
Xyloglucan	*β*-1,4(1,2)-D-galactosidase*	LAC	GH2,35	3.2.1.23
Xyloglucan	xyloglucan-active *β*-1,4-D-endoglucanase	XEG	GH12,74	3.2.1.151
Xyloglucan	xyloglucan acetylesterase	XGAE	-	-
Xyloglucan	*β*-xylosidase*	BXL	GH3,43	3.2.1.37
Galactomannan	*α*-1,4-D-galactosidase*	AGL	GH27,36	3.2.1.22
Galactomannan	*β*-1,4-D-glucosidase*	BGL	GH1,3	3.2.1.21
Galactomannan	galactomannan acetyl esterase	GMAE	-	-
Galactomannan	*β*-1,4-D-endomannanase	MAN	GH5,26	3.2.1.78
Galactomannan	*β*-1,4-D-mannosidase	MND	GH2	3.2.1.25
Pectin	*α*-L-arabinofuranosidase*	ABF	GH43,51,54	3.2.1.55
Pectin	*endo*-arabinanase	ABN	GH43	3.2.1.99
Pectin	*exo*-arabinanase	ABX	GH93	3.2.1.-
Pectin	*β*-xylosidase*	BXL	GH3,43	3.2.1.37
Pectin	feruloyl esterase*	FAE	CE1	3.1.1.73
Pectin	*β*-1,4-endogalactanase	GAL	GH53	3.2.1.89
Pectin	*β*-1,3/-1,6-endogalactanase	GLN	GH5	3.2.1.164
Pectin	*β*-1,4(1,2)-D-galactosidase *	LAC	GH2,35	3.2.1.23

Pectin	pectin acetyl esterase	PAE	CE12	3.1.1.-
Pectin	pectin lyase	PEL	PL1	4.2.2.10
Pectin	rhamnogalacturonan lyase	RGL	PL4,11	4.2.2.-
Pectin	*endo*-polygalacturonase	PGA	GH28	3.2.1.15
Pectin	*exo*-polygalacturonase	PGX	GH28	3.2.1.67
Pectin	pectate lyase	PLY	PL1,3,9	4.2.2.2
Pectin	pectin methyl esterase	PME	CE8	3.1.1.11
Pectin	rhamnogalacturonan acetyl esterase	RGAE	CE12	3.1.1.-
Pectin	rhamnogalacturonan galaturonohydrolase/exorhamnogalacturonase	RGX	GH28	3.2.1.40
Pectin	α-rhamnosidase/rhamnogalacturonan rhamnohydrolase	RHA	GH78	3.2.1.40
Pectin	rhamnogalacturonan hydrolase/endorhamnogalacturonase	RHG	GH28	3.2.1.-
Pectin	d-4,5-unsaturated glucuronyl hydrolase	UGH	GH88	3.2.1.-
Pectin	unsaturated rhamnogalacturonan hydrolase	URH	GH105	3.2.1.-
Pectin	β-1,4-exogalactanase	XFG	-	-
Pectin	β-1,6-exogalactanase	XSG	-	-
Pectin	β-1,3-exogalactanase	XTG	GH43	3.2.1.145
Starch	Isoamylase	ISA	GH31	3.2.1.68
Starch	α-amylase	AMY	GH13	3.2.1.1
Starch	Glucoamylase	GLA	GH15	3.2.1.3
Inulin	*endo*-inulinase	INU	GH32	3.2.1.7
Inulin	*exo*-inulinase	INX	GH32	3.2.1.80
Inulin	invertase/fructofuranosidase	SUC	GH32	3.2.1.26

*indicates enzyme activities which are involved in the degradation of more than one different substrate.

Storage polysaccharides: Starch and Inulin

In addition to the cell wall polysaccharides outlined above, plants also consist of storage polysaccharides such as starch and inulin [52]. The formation of these polysaccharides allows the plant to store simple sugars such as glucose and fructose.

Starch is located in the plastids and is comprised of a branched α-linked D-glucose backbone that is degraded by the enzymatic action of α-amylase which is an *endo*-acting enzyme that hydrolyses starch into gluco-oligosaccharides. These gluco-oligosaccharides are then further degraded into D-glucose by the action of glucoamylases and isoamylase [112] (Fig. 10, Table 1).

Inulin is found in the roots and rhizomes of many plants and is composed of a linear polymer of β-2,1-linked D-fructose, connected to a terminal sucrose residue [86]. This polysaccharide is hydrolyzed by *endo*- and *exo*-inulinases to D-fructose [40], while the sucrose disaccharide is broken down into D-glucose and D-fructose by the action of fructofuranosidase (invertase) [120] (Fig. 10, Table 1).

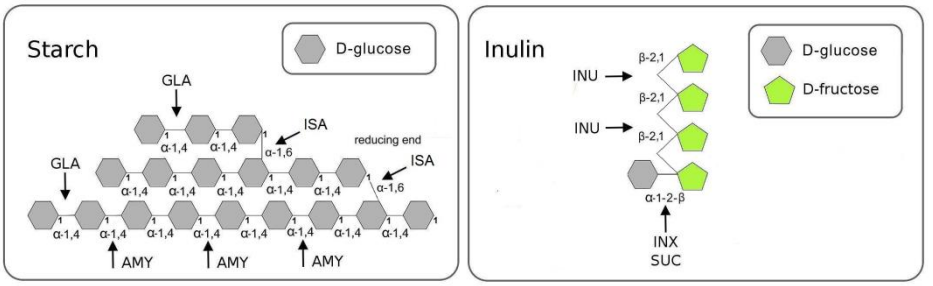


Figure 10: Representation of the starch and inulin structures and their hydrolysis by starch/inulin degrading enzymes. The enzymes involved in the complete biodegradation of starch/inulin are indicated. Abbreviations of enzyme names are explained in Table 1. The figure is modified from [5].

Plant polysaccharide degradation by *Aspergillus*

Plant polysaccharide degradation has been studied in detail in *Aspergillus* [34, 37]. To date, the genome sequences of *A. niger* [6, 90], *A. nidulans* [47], *A. oryzae* [72], *A. fumigatus* [81], *A. flavus* [87], *A. sojae* [98], *A. clavatus* [128], *A. terreus* [128], *A. aculeatus*, *A. carbonarius*, *A. kawachii*, *Neosatorya fischeri*, *A. sydowii*, *A. versicolor*, *A. brasiliensis*, *A. tubingensis*, *A. glaucus*, *A. zonatus*, *A. acidicus* and *A. wentii* are all available; (http://www.ncbi.nlm.nih.gov/bioproject, http://genome.jgi.psf.org/programs/fungi/genome-releases.jsf). Comparative analysis of these genomes increases our understanding of the biology of these species and helps us realize their full potential in industrial applications but also in plant biomass degradation [52].

One such comparative study compared the plant polysaccharide degrading potential of *A. nidulans* to that of two industrially relevant strains: *A. niger* and *A. oryzae* [23]. Comparison of the CAZy content of the genomes of these fungi demonstrated significant differences (Table 2). The *A. oryzae* genome contained a significantly higher number of xylan- and pectin-related genes than the other two species, while this was the case for galactomannan-related genes in *A. nidulans* and for inulin-related genes in *A. niger*. Differences in the genome content don't always correlate with the growth profiles of these species found in the Fung-Growth website (www.fung-growth.org). While a good correlation can be found for inulin this is less clear for the other polysaccharides (Fig. 11, Table 2).

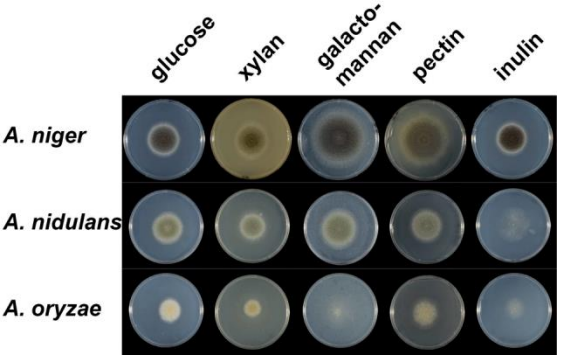


Figure 11: Growth comparison of three Aspergilli on four plant polysaccharides relative to growth on glucose.

Table 2: Comparison of the number of putative genes involved in plant biomass degradation of three Aspergilli. Modified from [23].

Polysaccharide	CAZy families	*A. nidulans*	*A. niger*	*A. oryzae*
cellulose	GH1,3,5,6,7,45,61	37	30	33
xylan	GH3,10,11,43,62,67,115; CE1,15	20	20	37
xyloglucan	GH12,29,31,74,95	6	7	8
galactomannan	GH2,5,26,27,36	19	12	12
pectin	GH2,28,35,43,51,53,54,78,88,93,105; PL1,3,4,9,11; CE1,8,12	73	66	92
starch	GH13,15	16	15	21
inulin	GH32	2	5	4

Efficient degradation of polysaccharides not only relies on the production of a specific set of enzymes but also requires synergistic interactions between the enzymes themselves. Synergy between enzymes in Aspergilli is a well-accepted phenomenon and one which has been shown to enhance its degrading abilities [34]. Addition of acetylxylan esterase increased the release of xylose and associated oligosaccharides for the *endo*-xylanses by up to a factor of 4.4 (*endo*-xylanase I), 14.7 (*endo*-xylanase II) and 16.3 (*endo*-xylanae III) [65].The action of endo-xylanase and *β*-xylosidase on the xylan backbone increased 2- and 2.5-fold through the cooperation of arabinoxylan arabinofuranohydrolase and *α*-L-arabinofuranosidase, respectively [33]. Similarly, in the degradation of pectin, the action of *β*-galactosidase was shown to support the activity of feruloyl esterase A by increasing the amount of ferulic acid released by over two-fold [33].

Regulatory systems of *Aspergillus* related to plant biomass degradation

Due to the complexity of plant biomass and its large variety depending on species, tissue and season, efficient degradation by fungi requires a fine-tuned production of diverse enzyme sets, tailored to the available biomass. This is organized through a set of transcriptional regulators that activate or repress the expression of genes encoding plant biomass degrading enzymes. Regulation of plant biomass degradation has been best studied in *Aspergillus* and will be discussed in the next sections and is summarised in Table 3. The general model for regulation of plant biomass utilization is depicted in Fig. 12 (de Vries, unpublished data).When a monomeric component of a polysaccharide is sensed by the fungus, a signalling pathway results in the activation of a transcription factor (regulator) that enters the nucleus and binds to the promoter of its target genes. These genes are then expressed and produce enzymes involved in metabolism of the compound to enable the cell to grow. As monomeric compounds are rare in the environment, the presence of it suggests the presence of the polymeric compound it originates from. Therefore, genes encoding extracellular polysaccharide degrading enzymes are also expressed resulting in liberation of more of the monomeric compound.

The amylolytic regulator AmyR

AmyR controls the genes involved in starch degradation, such as glucoamylase (*glaA*), alpha-amylase (*amyA*, *amyB*) and *α*-glucosidase (*agdA*) [92, 109]. AmyR contains a Zn_2Cys_6 DNA-binding domain and belongs to the Gal4 regulator-family [83]. AmyR can bind to $CGGN_8CGG$ and CGGAAATTTAA sequences in the promoter of its target genes in the presence of specific compounds (inducers), such as starch and maltose [92]. Maltose was long considered the true inducing compound of the system but recent studies have questioned this. In *A. oryzae*, isomaltose has been shown to be

the best inducer [78], while glucose has also been shown to induce activation of AmyR in *A. niger* and *A. oryzae* [19, 117]. In *A. niger*, AmyR also regulates genes encoding *β*-glucosidases and *α*- and *β*-galactosidases [117].

The (hemi-)cellulolytic regulator XlnR

XlnR is also a member of the Gal4 (Zn_2Cys_6) family of regulators. While this regulator was originally described as a xylanolytic regulator [115], later studies demonstrated that is also controls genes involved in the degradation of cellulose and xyloglucan [31, 49, 114] as well as several genes of the pentose catabolic pathway [11, 12, 29, 55, 116]. XlnR binds to GGCTAR sequences in the promoter of its target genes in the presence of its inducer, xylose [36]. While *xlnR* itself is constitutively expressed, the protein gets activated in the presence of xylose by C-terminal proteolytic cleavage that results in transport of XlnR to the nucleus [56].

XlnR has also been studied in *A. oryzae* and *A. nidulans* where it performs a similar role, although the set of target genes are not identical in the three species. However, approximately 20 genes were shown to be consistently regulated by XlnR in these three species [4].

The arabinanolytic regulator AraR

After the availability of the *A. niger* genome sequence [90] three genes with amino acid homology to XlnR were identified and analysis of these genes demonstrated that the closest homolog encoded AraR also contains a Zn_2Cys_6 DNA binding domain [12]. AraR activates the expression of genes encoding enzymes releasing arabinose from arabinoxylan and pectin, as well as genes encoding other pectinases and genes of the pentose catabolic pathway. The binding site of AraR has not yet been determined, but *araR* itself is induced in the presence of arabinose [12], suggesting autoregulation. Analysis of the function of AraR in *A. nidulans* revealed differences in its effect on specific target genes, such as *ladA* (encoding L-arabitol dehydrogenase) [11].

The inulinolytic regulator InuR

InuR was identified through analysis of the surrounding region of genes encoding inulin-degrading enzymes (invertase, *endo*- and *exo*-inulinase) in *A. niger* [130], based on the gene organization for *amyR* in *Aspergillus* [92]. Interestingly, InuR has sequence similarity to AmyR and is likely to have originated from a common ancestor, similar to XlnR and AraR.

InuR is also a member of the Zn_2Cys_6 family of transcriptional regulators, but its DNA binding site has not yet been determined. The inducer of InuR has been suggested to be sucrose [130].

The galactose-related regulators GalR and GalX

GalR and GalX form a two-factorial regulatory system for galactose utilization in *A. nidulans* [21]. GalR was identified by homology to XlnR and is also a member of the Zn_2Cys_6 family of transcriptional regulators. Detailed analysis demonstrated a second gene encoding a Zn_2Cys_6 regulator (GalX) next to GalR in the genome of *A. nidulans*. GalX regulates the expression of GalR and some D-galactose metabolic genes, while the majority of the metabolic genes as well as at least one α-galactosidase encoding gene are under control of GalR. GalR is unique to *A. nidulans*, while GalX is found in several Aspergilli as well as other filamentous ascomycetes [21]. In *A. niger* GalX regulates the oxido-reductive D-galactose catabolic pathway [53].

The pectinolytic regulator RhaR

In *A. niger* the presence of a general galacturonic acid responsive regulator that controls most pectinases was suggested, with additional regulators responding to rhamnose and arabinose (AraR, see above) [35]. Analysis of micro array data on a pectin-related compound identified several candidate regulator encoding genes that were specifically expressed in the presence of rhamnose and rhamnose-containing substrates. One was located next to genes of a recently discovered L-rhamnose catabolic pathway [123] and disruption of this gene (*rhaR*) resulted in reduced growth on L-rhamnose as well as reduced α-rhamnosidase activity compared to the wild [52]. RhaR affects the expression of not only α-rhamnosidase encoding genes, but also genes encoding other pectinolytic enzymes, such as rhamnogalacturonases, rhamnogalacturonan lyases, rhamnogalacturonan and pectin acetyl esterases and β-galactosidases [52]. Most of these genes are related to degradation of the rhamnogalacturonan I.

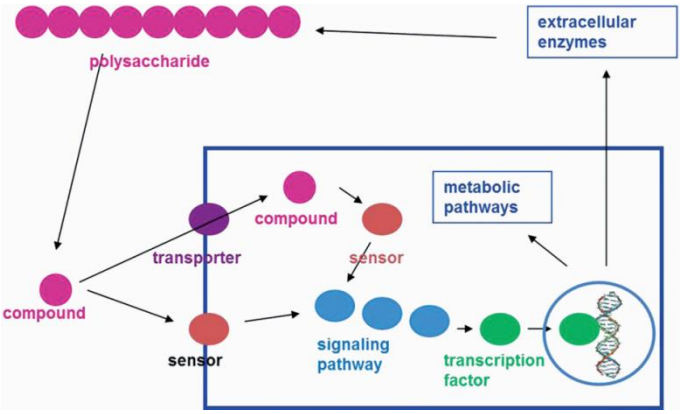


Figure 12: Model for the utilization of plant polysaccharides by fungi.

The mannanolytic regulator ManR

ManR is also a member of the Zn_2Cys_6 family of transcriptional regulators and was identified in *A. oryzae* [82]. ManR was shown to regulate genes encoding *endo*-mannanase, *β*-mannosidase, *α*-galactosidase, acetylmannan esterase and *β*-glucosidase and disruption of this regulator resulted in a significant reduction in growth on galactomannan [82].

The carbon catabolite repressor CreA

CreA was originally discovered in *A. nidulans* as a glucose-repressor that suppresses the expression of its target genes in the presence of glucose [41] and its DNA-binding site was established to be SYGGRT [66]. Later studies demonstrated that CreA suppresses the expression of many genes in the present of sufficient amounts of easily metabolizable carbon sources, such as glucose, but also xylose, mannose, glucuronic acid and others [97]. Its role seems to be in conserving energy by turning off genes for secondary carbon sources when sufficient amount of a primary carbon source is present. This not only includes genes of alternative metabolic pathways, but also genes involved in the release of monomeric carbon sources, such as those encoding plant biomass degrading enzymes [97]. Several studies demonstrated significant increases in the production of plant polysaccharide degrading enzymes by fungi when *creA* is deleted [32, 62, 108]. One of the best cellulose producing strains is the *Trichoderma reesei* RUT-C30 strain, which has been shown to be a *cre1* deletion strain [104]. It was also suggested that a bigger effect on protein production can be achieved by removing the repression on the positively acting regulator (e.g. AmyR) than on its target genes (e.g. amylase-encoding genes) [2].

Interaction between regulators

In recent studies, the interaction between several regulators was examined and demonstrated. XlnR and AraR both regulate the genes associated with the common steps of the L-arabinose and D-xylose catabolic pathways [12]. In addition, they act antagonistically with respect to their target genes. In a wild type strain, XlnR-regulated genes are expressed on xylose while AraR-regulated genes are expressed on arabinose. In a *xlnR*-knockout however, AraR-regulated genes are also expressed on xylose while in an *araR*-knockout, XlnR-regulated genes are also expressed on arabinose [12]. The mechanism of this interaction has not yet been elucidated.

XlnR and CreA, together, determine the expression level of xylanolytic genes in the presence of D-xylose. In a wild type strain, xylose induces the genes through the action of XlnR, but increasing levels of xylose result in a reduction of the expression of these genes [32]. This did not occur in a CreA-mutant, demonstrating that this modulation of the expression level occurs via CreA [32]. In a recent study this data has

been confirmed for a larger group of xylanolytic genes, although the strength of the CreA effect differed for sub-groups of these genes [71].

Table 3: Regulators involved in plant biomass utilisation in *Aspergillus.*

Regulator	Function	Inducer	Reference
Positively acting regulators			
AmyR	Starch degradation, glucose/galactose release	D-glucose, isomaltose	[78, 92, 117]
AraR	L-arabinose release and catabolism	L-arabitol	[11, 12]
GalR	D-galactose catabolism	D-galactose	[21]
GalX	D-galactose catabolism	D-galactose	[21, 53]
InuR	Inulin degradation	Sucrose	[130]
ManR	(galacto-)mannan degradation	Mannobiose	[82]
XlnR	(hemi-)cellulose degradation, D-xylose catabolism	D-xylose	[114, 115]
Negatively acting regulators			
CreA	Carbon catabolite repression	Sufficient levels of monomeric carbon sources	[41, 97]

Improved enzyme production

The initial focus of protein production in filamentous fungi was to increase homologous expression. Examples of such studies are the development of an *A. niger* strain in which secretion of glucoamylase went from 0.5 g/L to levels of up to 30 g/L through classical mutation-selection techniques, media development and fermentation optimizations, over a 20 year period. Finkelstein (1987) showed that recombinant technology could be used as a viable alternative to the classical methods of strain improvement [43] and thus created an interest in these fungi as possible production hosts for heterologous proteins.

The high yields achieved with homologous protein production have not yet been realized with heterologous proteins [51]. Apart from protease activity, which is one of the leading problems for heterologous protein production, this may be due to a number of factors including transcription, translation, secretion and host strain physiology. To improve recombinant protein production in fungi, several strategies have been developed. Genetic strategies include the introduction of multi-copies of the target protein gene, the use of strong promoters and effective secretion signals, gene fusions to genes associated with well-expressed and secreted proteins and development of

protease-deficient host strains [95, 121]. Some success has also been obtained using a bioprocessing approach of optimizing fungal morphology, mycelia immobilization and culture conditions [1, 18].

Transcriptional control

Studying gene expression in filamentous fungi has increased our understanding of the molecular mechanisms controlling transcription initiation and/or regulation along with the selection of strong promoters. The promoter regions of the *Aspergillus* amylase genes consist of four highly conserved regions, which work together to enhance expression levels [76, 122]. A sequence of CCAAT, present in the *amdS* promoter region of *A. nidulans* and responsible for high-level expression of acetamidase, was also found in many other promoter regions in *A. nidulans* [84, 122]. Furthermore, the enolase gene (*enoA*) which is one of the most highly expressed genes in *A. oryzae*, was found to contain a 15-bp sequence, essential for transcriptional regulation of that gene [111, 122].

A number of other studies have shown transcriptional control to be a contributor in production levels of heterologous proteins in filamentous fungi. One such analysis, carried out by Archer *et al*. (1994) showed that hen egg-white lysozyme (HEWL) mRNA levels under the control of the *glaA* promoter in *A. niger* correlated well with HEWL protein production, but comparable HEWL and glucoamylase mRNA levels did not lead to comparable secreted protein levels [7].

Protein synthesis and secretion

A general model for protein synthesis and secretion in fungi was proposed [44]. During synthesis, proteins are transported to the endoplasmic reticulum where folding and glycosylation are initiated. In Aspergilli, folding and maturation of secretory proteins is assisted by protein disulfide isomerase (Pdi). SNARE proteins facilitate vesicle-mediated transport of the completed proteins to the more-porous hyphal tip for extracellular secretion while those which did not fold or glycosylate correctly are sent to the proteosomes or vacuoles for degradation. It has been hypothesized that over-production of recombinant proteins into the ER has the potential of congesting the folding, assembly and secretion pathways of the fungal hosts. The over-production of foldases and chaperones, which catalyze and mediate the folding of emerging polypeptides into functional proteins, could alleviate this bottleneck [122]. After secretion, recombinant proteins often face the risk of degradation by homologous proteases produced by many filamentous fungi. Different heterologous proteins are subject to degradation by different proteases [7] although careful consideration of species such as *A. vadensis* [38] or the use of protease deficient mutants can limit this.

Bioprocessing strategies

Optimization of the yield of secreted heterologous proteins also relies on the response of protein secretion to growth conditions. Alterations to the growth medium, including agitation and dissolved oxygen [24], have been shown to relate directly to the specific responses of a given promoter but also to fungal morphology, protease production [39] and cell viability [7]. Ambient pH directly regulates the level of proteases produced in *A. niger* [39] while carbon and nitrogen sources can have either repressing or inducing effects on enzyme production [88, 89]. Carbon and nitrogen sources can also effect protease activity with most extracellular proteases being inhibited under conditions of high glucose or ammonium in the medium [121]. Furthermore, some studies have shown that gene expression and protein secretion might be quite different in solid-state fermentation than that observed in submerged fermentation e.g. in submerged cultures some enzyme activities are found in the cell wall of the mycelia while in solid-state cultures they are found to be secreted into the media [54, 64].

Aim and outline of this thesis

Much research over the past 25 years has been applied to the development of filamentous fungi, most notably *Aspergillus*, as hosts for recombinant protein production. Their inherent abilities to grow at high rates and to high biomass densities and their exceptional capacity to secrete high levels of homologous product are well recognized. Despite there being many advances made in the hyper-production of heterologous proteins in filamentous fungi, their ability to produce and secrete homologous proteases along with different native protein glycosylation still requires further strain improvement to efficiently produce a wide range of heterologous proteins. The aim of this thesis was to develop an efficient fungal expression system for the production of recombinant proteins, in particular those involved in plant biomass degradation. Due to its potential in biomass degradation and with its favourable fermentation capabilities, *Aspergillus* was chosen for this study.

In **Chapter 2** eight different *Aspergillus* species were compared with respect to their genomic ability to degrade plant cell wall biomass. While all tested Aspergilli had a similar potential to degrade plant biomass, results showed that even in closely related species, their strategies differed markedly. Combining the approaches from different species is likely to result in better enzyme mixtures for industrial applications, such as the saccharification of plant biomass for biofuel production.

In **Chapter 3**, the molecular and phenotypical differences between *A. vadensis* and six other species of black Aspergilli were examined. Growth on varying carbon sources and extracellular enzyme profiles when grown on maltose/starch indicated significant and unique differences between *A. vadensis* and the other black Aspergilli. Further analysis of the partial genome of *A. vadensis* genome, combined with gene

expression data when grown on maltose indicate that its aberrant phenotype is likely caused by the low expression of *prtT* and *amyR* regulators or their associated genes and not a mutation or deletion as was originally concluded.

In **Chapter 4** six novel constitutive promoters from *A. niger* (p*ef1α*, p*tktA*, p*ef1β*, p*tal1*, p*cetA* and p*pgkA*) and a further five from *A. vadensis* (p*ef1α*, p*rps31*, p*gpdA*, p*ubi1* and p*oliC*) were tested in *A. vadensis* using a gene encoding a secreted arabinofuranosidase from *Fusarium oxysporum* as a reporter for heterologous protein production. Of the promoters tested, 3 from *A. niger* (p*ef1α* > p*tal* > p*pgkA*) and 3 from *A. vadensis* (p*ef1α* > p*oliC* > p*rps31*) all resulted in higher ABF activity than for that of the commonly used *gpdA* promoter from *A. nidulans.*

In **Chapter 5** the potential of *A. vadensis* as an expression host was tested by successfully expressing an *α*-L-arabinofuranosidase (*abfB*) (GH54) and an *endo*-1,4-*β*-D-glucanase (*eglA*) (GH12) from *A. vadensis* under the control of the *gpdA* promoter from *A. vadensis*.

In **Chapter 6** an evolutionary screening method was used to improve the inulin degradation potential of *Aspergillus oryzae* through the upregulation of *exo*-inulinase. As an organism with no predicted *endo*-inulinase function, improved inulin degradation would be largely dependent on the overproduction of this enzyme. Subsequent generation growth of *Aspergillus oryzae* (Rib40) on inulin for 9 weeks successfully resulted in *exo*-inulinase overproducing mutants.

References

1. **Agger, T., Spohr, A.B., Carlsen, M. and Nielsen, J.** 1998. Growth and product formation of *Aspergillus oryzae* during submerged cultivations: verification of a morphologically structured model using fluorescent probes. Biotechnol Bioeng. **57**: 321-329.
2. **Agger, T. and Nielsen, J.** 1999. Genetically structured modeling of protein production in filamentous fungi. Biotechnol Bioeng. **66**: 164-170.
3. **Albersheim, P.** 1975. The walls of growing plant cells. Scientific American. **232**: 80-95.
4. **Andersen, M.R., Vongsangnak, W., Panagiotou, G., Salazar, M.P., Lehmann, L. and Nielsen, J.** 2008. A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. Proc Natl Acad Sci U S A. **105**: 4387-4392.
5. **Andersen, M.R.** 2009. Systems Biology Studies of Aspergilli: From Sequence to Science. Technical University of Denmark.
6. **Andersen, M.R., Salazar, M.P., Schaap, P.J., van de Vondervoort, P.J., Culley, D., Thykaer, J., et al.** 2011. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res. **21**: 885-897.
7. **Archer, D.B., Jeenes, D.J. and Mackenzie, D.A.** 1994. Strategies for improving heterologous protein production from filamentous fungi. Antonie Van Leeuwenhoek. **65**: 245-250.
8. **Bajpai, P.** 1997. Microbial xylanolytic enzyme system: properties and applications. Advances in applied microbiology. **43**: 141-194.
9. **Battaglia, E., Benoit, I., Gruben, B.S. and de Vries, R.P.** 2010. Plant cell wall derived sugars as substrates for fungi and industry. In: Jenkins PT, (ed) The Sugar Industry and Cotton Crops, Nova Science Publishers, New York.
10. **Battaglia, E.** 2011. Regulation and diversity of plant polysaccharide utilisation in fungi. Utrecht University.
11. **Battaglia, E., Hansen, S.F., Leendertse, A., Madrid, S., Mulder, H., Nikolaev, I., et al.** 2011. Regulation of pentose utilisation by AraR, but not XlnR, differs in *Aspergillus nidulans* and *Aspergillus niger*. Appl Microbiol Biotechnol. **91**: 387-397.
12. **Battaglia, E., Visser, L., Nijssen, A., van Veluw, G.J., Wosten, H.A. and de Vries, R.P.** 2011. Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in *Eurotiales*. Stud Mycol. **69**: 31-38.
13. **Bedford, M.R. and Classen, H.L.** 1991. The influence of dietary xylanase on intestinal viscosity and molecular weight distribution of carbohydrates in rye-fed broiler chicks. Progress in biotechnology. **7**: 361-370.

14. **Boyce, K.J. and Andrianopoulos, A.** 2006. Morphogenesis: Control of Cell Types and Shape. In: Kües U, Fischer R, (ed). Growth, Differentiation and Sexuality, Springer Berlin Heidelberg, pp. 3-20.
15. **Brillouet, J.-M. and Joseleau, J.-P.** 1987. Investigation of the structure of a heteroxylan from the outer pericarp (beeswing bran) of wheat kernel. Carbohydr Res. **159**: 109-126.
16. **Campbell, K. and White, D.** 1995. Inheritance of Resistance to *Aspergillus* Ear Rot and Anatoxin in Corn Genotypes.
17. **Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. **37**: 5.
18. **Carlsen, M., Spohr, A.B., Nielsen, J. and Villadsen, J.** 1996. Morphology and physiology of an alpha-amylase producing strain of *Aspergillus oryzae* during batch cultivations. Biotechnol Bioeng. **49**: 266-276.
19. **Carlsen, M. and Nielsen, J.** 2001. Influence of carbon source on alpha-amylase production by *Aspergillus oryzae*. Appl Microbiol Biotechnol. **57**: 346-349.
20. **Carpita, N.C. and Gibeaut, D.M.** 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J. **3**: 1-30.
21. **Christensen, U., Gruben, B.S., Madrid, S., Mulder, H., Nikolaev, I. and de Vries, R.P.** 2011. Unique regulatory mechanism for D-galactose utilization in *Aspergillus nidulans*. Appl Environ Microbiol. **77**: 7084-7087.
22. **Christov, L.P., Szakacs, G. and Balakrishnan, H.** 1999. Production, partial characterization and use of fungal cellulase-free xylanases in pulp bleaching. Process Biochemistry. **34**: 511-517.
23. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.
24. **Cui, Y.Q., van der Lans, R.G.J.M., Giuseppin, M.L.F. and Luyben, K.C.A.M.** 1998. Influence of fermentation conditions and scale on the submerged fermentation of *Aspergillus awamori*. Enzyme and Microbial Technology. **23**: 157-167.
25. **D'Mello, J.F., Macdonald, A.M., Postel, D., Dijksma, W.T., Dujardin, A. and Placinta, C.M.** 1998. Pesticide use and mycotoxin production in *Fusarium* and *Aspergillus* phytopathogens. European Journal of Plant Pathology. **104**: 741-751.
26. **Daas, P.J.H., Meyer-Hansen, K., Schols, H.A., De Ruiter, G.A. and Voragen, A.G.J.** 1999. Investigation of the non-esterified galacturonic acid distribution in pectin with endopolygalacturonase. Carbohydr Res. **218**: 135-145.

27. **Dadachova, E., Bryan, R.A., Huang, X., Moadel, T., Schweitzer, A.D., Aisen, P., et al.** 2007. Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. PLoS One. **2**: e457.
28. **Davies, R.W.** 1994. Heterologous gene expression and protein secretion in *Aspergillus*. Prog Ind Microbiol. **29**: 527-560.
29. **de Groot, M.J., van den Dool, C., Wosten, H.A., Levisson, M., vanKuyk, P.A., Ruijter, G.J., et al.** 2007. Regulation of pentose catabolic pathway genes of *Aspergillus niger*. Food Technol Biotechnol. **45**: 134-138.
30. **de Vries, R.P.** 1999. Accessory enzymes from *Aspergillus* involved in xylan and pectin degradation. Wageningen Agricultural University.
31. **de Vries, R.P., van den Broeck, H.C., Dekkers, E., Manzanares, P., de Graaff, L.H. and Visser, J.** 1999. Differential expression of three *α*-galactosidase genes and a single *β*-galactosidase gene from *Aspergillus niger*. Appl Environ Microbiol. **65**: 2453-2460.
32. **de Vries, R.P., Visser, J. and de Graaff, L.H.** 1999. CreA modulates the XlnR-induced expression on xylose of *Aspergillus niger* genes involved in xylan degradation. Res Microbiol. **150**: 281-285.
33. **de Vries, R.P., Kester, H.C., Poulsen, C.H., Benen, J.A. and Visser, J.** 2000. Synergy between enzymes from *Aspergillus* involved in the degradation of plant cell wall polysaccharides. Carbohydr Res. **327**: 401-410.
34. **de Vries, R.P. and Visser, J.** 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev. **65**: 497-522.
35. **de Vries, R.P., Jansen, J., Aguilar, G., Parenicova, L., Joosten, V., Wulfert, F., et al.** 2002. Expression profiling of pectinolytic genes from *Aspergillus niger*. FEBS Lett. **530**: 41-47.
36. **de Vries, R.P., van de Vondervoort, P.J., Hendriks, L., van de Belt, M. and Visser, J.** 2002. Regulation of the *α*-glucuronidase-encoding gene (*aguA*) from *Aspergillus niger*. Mol Genet Genomics. **268**: 96-102.
37. **de Vries, R.P.** 2003. Regulation of *Aspergillus* genes encoding plant cell wall polysaccharide-degrading enzymes; relevance for industrial production. Appl Microbiol Biotechnol. **61**: 10-20.
38. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
39. **Denison, S.H.** 2000. pH Regulation of Gene Expression in Fungi. Fungal Genetics and Biology. **29**: 61-71.

40. **Derycke, D.G. and Vandamme, E.J.** 1984. Production and properties of *Aspergillus niger* inulinase. Journal of Chemical Technology and Biotechnology Biotechnology. **34**: 45-51.
41. **Dowzer, C.E. and Kelly, J.M.** 1989. Cloning of the *creA* gene from *Aspergillus nidulans*: a gene involved in carbon catabolite repression. Curr Genet. **15**: 457-459.
42. **Eriksson, K.-E.L., Blanchette, R.A. and Ander, P.** Microbial and enzymatic degradation of wood and wood components. Springer-verlag, 1990.
43. **Finkelstein, D.** 1987. Improvement of enzyme production in *Aspergillus*. Antonie Van Leeuwenhoek. **53**: 349-352.
44. **Fleissner, A. and Dersch, P.** 2010. Expression and export: recombinant protein production systems for *Aspergillus*. Appl Microbiol Biotechnol. **87**: 1255-1270.
45. **Frey-Wyssling, A.** 1954. The fine structure of cellulose microfibrils. Science. **119**: 80-82.
46. **Fry, S.C.** 1989. The structure and functions of xyloglucan. J Exp Bot. **40**: 1-11.
47. **Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., et al.** 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature. **438**: 1105-1115.
48. **Gautam, A.K., Sharma, S., Avasthi, S. and Bhadauria, R.** 2011. Diversity, pathogenicity and toxicology of *A. niger*: an important spoilage fungi. Research Journal of Microbiology. **6**: 270-280.
49. **Gielkens, M.M., Dekkers, E., Visser, J. and de Graaff, L.H.** 1999. Two cellobiohydrolase-encoding genes from *Aspergillus niger* require D-xylose and the xylanolytic transcriptional activator XlnR for their expression. Appl Environ Microbiol. **65**: 4340-4345.
50. **Goldman, G.H. and Kafer, E.** 2004. *Aspergillus nidulans* as a model system to characterize the DNA damage response in eukaryotes. Fungal Genetics and Biology. **41**: 428-442.
51. **Gouka, R.J., Punt, P.J. and van den Hondel, C.A.** 1997. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. Appl Microbiol Biotechnol. **47**: 1-11.
52. **Gruben, B.S.** 2012. Novel transcriptional activators of *Aspergillus* involved in plant biomass utilization. Utrecht University.
53. **Gruben, B.S., Zhou, M. and de Vries, R.P.** 2012. GalX regulates the D-galactose oxido-reductive pathway in *Aspergillus niger*. FEBS Lett. **586**: 3980-3985.
54. **Hashimoto, T., Morishita, M., Iwashita, K., Shimoi, H., Nakata, Y., Tsuji, Y., et al.** 1999. Production and some properties of salt-tolerant *β*-xylosidases from a shoyu koji mold, *Aspergillus oryzae* in solid and liquid cultures. J Biosci Bioeng. **88**: 479-483.

55. **Hasper, A.A., Visser, J. and de Graaff, L.H.** 2000. The *Aspergillus niger* transcriptional activator XlnR, which is involved in the degradation of the polysaccharides xylan and cellulose, also regulates D-xylose reductase gene expression. Mol Microbiol. **36**: 193-200.
56. **Hasper, A.A., Trindade, L.M., van der Veen, D., van Ooyen, A.J. and de Graaff, L.H.** 2004. Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. Microbiology. **150**: 1367-1375.
57. **Hawksworth, D.L.** 1991. The fungal dimension of biodiversity: magnitude, significance, and conservation. Mycological Research. **95**: 641-655.
58. **Hedayati, M., Pasqualotto, A., Warn, P., Bowyer, P. and Denning, D.** 2007. *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. Microbiology. **153**: 1677-1692.
59. **Hisamatsu, M., York, W.S., Darvill, A.G. and Albersheim, P.** 1992. Characterization of seven xyloglucan oligosaccharides containing from seventeen to twenty glycosyl residues. Carbohydr Res. **227**: 45-71.
60. **Huisman, M., Schols, H. and Voragen, A.** 2000. Glucuronoarabinoxylans from maize kernel cell walls are more complex than those from sorghum kernel cell walls. Carbohydrate Polymers. **43**: 269-279.
61. **Huisman, M., Weel, K., Schols, H. and Voragen, A.** 2000. Xyloglucan from soybean (Glycine max) meal is composed of XXXG-type building units. Carbohydrate Polymers. **42**: 185-191.
62. **Ilyes, H., Fekete, E., Karaffa, L., Sandor, E., Szentirmai, A. and Kubicek, C.P.** 2004. CreA-mediated carbon catabolite repression of *β*-galactosidase formation in *Aspergillus nidulans* is growth rate dependent. FEMS Microbiol Lett. **235**: 147-151.
63. **Ishii, T.** 1997. Structure and functions of feruloylated polysaccharides. Plant Science. **127**: 111-127.
64. **Iwashita, K., Todoroki, K., Kimura, H., Shimoi, H. and Ito, K.** 1998. Purification and characterization of extracellular and cell wall bound *β*-glucosidases from *Aspergillus kawachii*. Biosci Biotechnol Biochem. **62**: 1938-1946.
65. **Kormelink, F.J.M., Lefebvre, B., Strozyk, F. and Voragen, A.G.J.** 1993. Purification and characterization of an acetyl xylan esterase from *Aspergillus niger*. Journal of Biotechnology. **27**: 267-282.
66. **Kulmburg, P., Mathieu, M., Dowzer, C., Kelly, J. and Felenbok, B.** 1993. Specific binding sites in the *alcR* and *alcA* promoters of the ethanol regulon for the CREA repressor mediating carbon catabolite repression in *Aspergillus nidulans*. Mol Microbiol. **7**: 847-857.
67. **Latge, J.P.** 1999. *Aspergillus fumigatus* and aspergillosis. Clin Microbiol Rev. **12**: 310-350.

68. **Lau, J.M., McNeil, M., Darvill, A.G. and Albersheim, P.** 1985. Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary cell walls of plants. Carbohydr Res. **137**: 111-125.
69. **Le Goff, A., Renard, C., Bonnin, E. and Thibault, J.-F.** 2001. Extraction, purification and chemical characterisation of xylogalacturonans from pea hulls. Carbohydrate Polymers. **45**: 325-334.
70. **Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., Celio, G., Dentinger, B., et al.** 2004. Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. Am J Bot. **91**: 1446-1480.
71. **Mach-Aigner, A.R., Omony, J., Jovanovic, B., van Boxtel, A.J. and de Graaff, L.H.** 2012. D-Xylose concentration-dependent hydrolase expression profiles and the function of CreA and XlnR in *Aspergillus niger*. Appl Environ Microbiol. **78**: 3145-3155.
72. **Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., et al.** 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature. **438**: 1157-1161.
73. **Maheshwari, R., Bharadwaj, G. and Bhat, M.K.** 2000. Thermophilic fungi: their physiology and enzymes. Microbiol Mol Biol Rev. **64**: 461-488.
74. **McCleary, B. and Matheson, N.** 1974. α-D-Galactosidase activity and galactomannan and galactosylsucrose oligosaccharide depletion in germinating legume seeds. Phytochemistry. **13**: 1747-1757.
75. **McCleary, B. and Matheson, N.** 1975. Galactomannan structure and β-mannanase and β-mannosidase activity in germinating legume seeds. Phytochemistry. **14**: 1187-1194.
76. **Minetoki, T., Kumagai, C., Gomi, K., Kitamoto, K. and Takahashi, K.** 1998. Improvement of promoter activity by the introduction of multiple copies of the conserved region III sequence, involved in the efficient expression of *Aspergillus oryzae* amylase-encoding genes. Appl Microbiol Biotechnol. **50**: 459-467.
77. **Morris, E.R.** 1986. Molecular Interactions in Polysaccharide Gelation. British Polymer Journal. **18**: 14-21.
78. **Murakoshi, Y., Makita, T., Kato, M. and Kobayashi, T.** 2012. Comparison and characterization of *α*-amylase inducers in *Aspergillus nidulans* based on nuclear localization of AmyR. Appl Microbiol Biotechnol. **94**: 1629-1635.
79. **Naidu, J. and Singh, S.M.** 1994. *Aspergillus chevalieri* (Mangin) Thom and Church: a new opportunistic pathogen of human cutaneous aspergillosis. Mycoses. **37**: 271-274.
80. **Nakamura, A., Furuta, H., Maeda, H., TAKAO, T. and NAGAMATSU, Y.** 2002. Analysis of the molecular construction of xylogalacturonan isolated from

soluble soybean polysaccharides. Bioscience, biotechnology, and biochemistry. **66**: 1155-1158.

81. **Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., et al.** 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. **438**: 1151-1156.

82. **Ogawa, M., Kobayashi, T. and Koyama, Y.** 2012. ManR, a novel Zn(II)(2)Cys(6) transcriptional activator, controls the *β*-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol. **49**: 987-995.

83. **Pan, T. and Coleman, J.E.** 1990. GAL4 transcription factor is not a "zinc finger" but forms a Zn(II)2Cys6 binuclear cluster. Proc Natl Acad Sci U S A. **87**: 2077-2081.

84. **Papagiannopoulos, P., Andrianopoulos, A., Sharp, J.A., Davis, M.A. and Hynes, M.J.** 1996. The *hapC* gene of *Aspergillus nidulans* is involved in the expression of CCAAT-containing promoters. Mol Gen Genet. **251**: 412-421.

85. **Pauly, M., Andersen, L.N., Kauppinen, S., Kofod, L.V., York, W.S., Albersheim, P., et al.** 1999. A xyloglucan-specific *endo-β*-1,4-glucanase from *Aspergillus aculeatus*: expression cloning in yeast, purification and characterization of the recombinant enzyme. Glycobiology. **9**: 93-100.

86. **Pavis, N., Chatterton, N.J., Harrison, P.A., Baumgartner, S., Praznik, W., Boucaud, J., et al.** 2001. Structure of fructans in roots and leaf tissues of *Lolium perenne*. New Phytologist. **150**: 83-95.

87. **Payne, G.A., Nierman, W.C., Wortman, J.R., Pritchard, B.L., Brown, D., Dean, R.A., et al.** 2006. Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Medical Mycology. **44**: 9-11.

88. **Pedersen, H., Beyer, M. and Nielsen, J.** 2000. Glucoamylase production in batch, chemostat and fed-batch cultivations by an industrial strain of *Aspergillus niger*. Appl Microbiol Biotechnol. **53**: 272-277.

89. **Pedersen, H. and Nielsen, J.** 2000. The influence of nitrogen sources on the *α*-amylase productivity of *Aspergillus oryzae* in continuous cultures. Appl Microbiol Biotechnol. **53**: 278-281.

90. **Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., et al.** 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol. **25**: 221-231.

91. **Perez, S., Rodriguez-Carvajal, M.A. and Doco, T.** 2003. A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. Biochimie. **85**: 109-121.

92. **Petersen, K.L., Lehmbeck, J. and Christensen, T.** 1999. A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet. **262**: 668-676.

93. **Polizeli, M.L., Rizzatti, A.C., Monti, R., Terenzi, H.F., Jorge, J.A. and Amorim, D.S.** 2005. Xylanases from fungi: properties and industrial applications. Appl Microbiol Biotechnol. **67**: 577-591.
94. **Poutanen, K.** 1997. Enzymes: An important tool in the improvement of the quality of cereal foods. Trends in Food Science & Technology. **8**: 300-306.
95. **Punt, P.J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J. and van den Hondel, C.** 2002. Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol. **20**: 200-206.
96. **Raghukumar, C. and Raghukumar, S.** 1998. Barotolerance of fungi isolated from deep-sea sediments of the Indian Ocean. Aquatic Microbial Ecology. **15**: 153-163p.
97. **Ruijter, G.J.G. and Visser, J.** 1997. Carbon repression in Aspergilli. FEMS Microbiol Lett. **151**: 103-114.
98. **Sato, A., Oshima, K., Noguchi, H., Ogawa, M., Takahashi, T., Oguma, T., et al.** 2011. Draft Genome Sequencing and Comparative Analysis of *Aspergillus sojae* NBRC4239. DNA Research. **18**: 165-176.
99. **Scheller, H.V. and Ulvskov, P.** 2010. Hemicelluloses. Annu Rev Plant Biol. **61**: 263-289.
100. **Schols, H.A. and Voragen, A.G.J.** 1994. Occurrence of pectic hairy regions in various plant cell wall materials and their degradability by rhamnogalacturonase. Carbohydr Res. **256**: 83-95.
101. **Schols, H.A., Bakx, E.J., Schipper, D. and Voragen, A.G.J.** 1995. A xylogalacturonan subunit present in the modified hairy regions of apple pectin. Carbohydr Res. **279**: 265-279.
102. **Schols, H.A. and Voragen, A.G.J.** 1996. Complex Pectins: Structure elucidation using enzymes. In: Visser J, Voragen AGJ, (ed). Progress in Biotechnology, Elsevier, pp. 3-19.
103. **Schooneveld-Bergmans, M., Hopman, A., Beldman, G. and Voragen, A.** 1998. Extraction and partial characterization of feruloylated glucuronoarabinoxylans from wheat bran. Carbohydrate Polymers. **35**: 39-47.
104. **Seidl, V., Gamauf, C., Druzhinina, I.S., Seiboth, B., Hartl, L. and Kubicek, C.P.** 2008. The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome. BMC Genomics. **9**: 327.
105. **Sims, I.M., Craik, D.J. and Bacic, A.** 1997. Structural characterisation of galactoglucomannan secreted by suspension-cultured cells of Nicotiana plumbaginifolia. Carbohydr Res. **303**: 79-92.

106. **Smith, M.M. and Hartley, R.D.** 1983. Occurrence and nature of ferulic acid substitution of cell-wall polysaccharides in graminaceous plants. Carbohydr Res. **118**: 65-80.
107. **Starkey, R.L. and Waksman, S.A.** 1943. Fungi Tolerant to Extreme Acidity and High Concentrations of Copper Sulfate. J Bacteriol. **45**: 509-519.
108. **Tamayo, E.N., Villanueva, A., Hasper, A.A., de Graaff, L.H., Ramon, D. and Orejas, M.** 2008. CreA mediates repression of the regulatory gene *xlnR* which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. Fungal Genet Biol. **45**: 984-993.
109. **Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K., et al.** 2001. Characterization of the *amyR* gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr Genet. **39**: 10-15.
110. **Timell, T.E.** 1967. Recent progress in the chemistry of wood hemicelluloses. Wood Science and Technology. **1**: 45-70.
111. **Toida, J., Fukuzawa, M., Kobayashi, G., Ito, K. and Sekiguchi, J.** 2000. Cloning and sequencing of the triacylglycerol lipase gene of *Aspergillus oryzae* and its expression in *Escherichia coli*. FEMS Microbiol Lett. **189**: 159-164.
112. **van der Maarel, M.J., van der Veen, B., Uitdehaag, J.C., Leemhuis, H. and Dijkhuizen, L.** 2002. Properties and applications of starch-converting enzymes of the α-amylase family. J Biotechnol. **94**: 137-155.
113. **van der Vlugt-Bergmans, C.J., Meeuwsen, P.J., Voragen, A.G. and van Ooyen, A.J.** 2000. Endo-xylogalacturonan hydrolase, a novel pectinolytic enzyme. Appl Environ Microbiol. **66**: 36-41.
114. **van Peij, N.N., Gielkens, M.M., de Vries, R.P., Visser, J. and de Graaff, L.H.** 1998. The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol. **64**: 3615-3619.
115. **van Peij, N.N., Visser, J. and de Graaff, L.H.** 1998. Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol. **27**: 131-142.
116. **vanKuyk, P.A., de Groot, M.J., Ruijter, G.J., de Vries, R.P. and Visser, J.** 2001. The *Aspergillus niger* D-xylulose kinase gene is co-expressed with genes encoding arabinan degrading enzymes, and is essential for growth on D-xylose and L-arabinose. Eur J Biochem. **268**: 5414-5423.
117. **vanKuyk, P.A., Benen, J.A., Wosten, H.A., Visser, J. and de Vries, R.P.** 2012. A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol. **93**: 285-293.
118. **Vidal, S., Doco, T., Williams, P., Pellerin, P., York, W.S., O'Neill, M.A., et al.** 2000. Structural characterization of the pectic polysaccharide rhamnogalacturonan II:

evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain. Carbohydr Res. **326**: 277-294.

119. **Vincken, J.P., York, W.S., Beldman, G. and Voragen, A.G.** 1997. Two general branching patterns of xyloglucan, XXXG and XXGG. Plant Physiol. **114**: 9-13.

120. **Wallis, G.L., Hemming, F.W. and Peberdy, J.F.** 1997. Secretion of two beta-fructofuranosidases by Aspergillus niger growing in sucrose. Arch Biochem Biophys. **345**: 214-222.

121. **Wang, L., Ridgway, D., Gu, T. and Moo-Young, M.** 2005. Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. Biotechnol Adv. **23**: 115-129.

122. **Ward, O.P.** 2012. Production of recombinant proteins by filamentous fungi. Biotechnol Adv. **30**: 1119-1139.

123. **Watanabe, S., Saimura, M. and Makino, K.** 2008. Eukaryotic and bacterial gene clusters related to an alternative pathway of nonphosphorylated L-rhamnose metabolism. J Biol Chem. **283**: 20372-20382.

124. **Wende, G. and Fry, S.C.** 1997. O-feruloylated, O-acetylated oligosaccharides as side-chains of grass xylans. Phytochemistry. **44**: 1011-1018.

125. **Wilkie, K.C.B.** 1979. The hemicelluloses of grasses and cereals. Advances in Carbohydrate Chemistry and Biochemistry.

126. **Willats, W.G., McCartney, L., Mackie, W. and Knox, J.P.** 2001. Pectin: cell biology and prospects for functional analysis. Plant Mol Biol. **47**: 9-27.

127. **Willats, W.G., Orfila, C., Limberg, G., Buchholt, H.C., van Alebeek, G.J., Voragen, A.G., et al.** 2001. Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls. Implications for pectin methyl esterase action, matrix properties, and cell adhesion. J Biol Chem. **276**: 19404-19413.

128. **Wortman, J.R., Fedorova, N., Crabtree, J., Joardar, V., Maiti, R., Haas, B.J., et al.** 2006. Whole genome comparison of the *A. fumigatus* family. Medical Mycology. **44**: 3-7.

129. **York, W.S., Kolli, V.K., Orlando, R., Albersheim, P. and Darvill, A.G.** 1996. The structures of arabinoxyloglucans produced by solanaceous plants. Carbohydr Res. **285**: 99-128.

130. **Yuan, X.L., Roubos, J.A., van den Hondel, C.A. and Ram, A.F.** 2008. Identification of InuR, a new Zn(II)2Cys6 transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol Genet Genomics. **279**: 11-26.

131. **Zandleven, J., Sørensen, S.O., Harholt, J., Beldman, G., Schols, H.A., Scheller, H.V., et al.** 2007. Xylogalacturonan exists in cell walls from various tissues of *Arabidopsis thaliana*. Phytochemistry. **68**: 1219-1226.

Chapter 2

Closely related fungi deploy diverse enzymatic strategies to degrade plant biomass

Benoit, I., Culleton, H.M. (co-first author), Zhou, M., DiFalco, M., Aguilar-Osorio, G., Battaglia, E., Bouzid, O., Brouwer, C.P.J.M., El-Bushari, H.B.O., Coutinho, P.M., Gruben, B.S., Hildén, K.S., Houbraken, J., Jiménez Barboza, L.A, Levasseur, A., Majoor, E., Mäkelä, M.R., Narang, H.M., Trejo-Aguilar, B., van den Brink, J., vanKuyk, P.A., Wiebenga, A., McKie, V.A., McCleary, B.V., Tsang, A., Henrissat, B. and de Vries, R.P. 2014. In preparation.

Abstract

Plant biomass is a major carbon source for many fungi and also an important industrial substrate for the production of food, textiles, fuels and chemicals. Degradation of plant biomass by fungi is a highly complex process that involves a large number of extracellular enzymes as well as non-hydrolytic proteins, whose production is controlled by a set of transcriptional regulators. *Aspergillus* species form one of the best studied fungal genera, largely due to their applications in biotechnology and relevance in human health. In this study we have compared the genomic content and the enzymes produced by eight Aspergilli for the utilization of plant biomass. While all tested Aspergilli have a similar genomic potential to degrade plant biomass, their approaches differ markedly in the overall activities as well as the specific enzymes they employ. An exception is *A. clavatus* that has a strongly reduced pectinolytic ability. Many of the genes have orthologs in (nearly) all tested species, but only very few of the corresponding enzymes are produced by all species during growth on wheat bran or sugar beet pulp. In addition, significant differences were observed between the enzyme sets produced on these feedstocks, largely correlating with their polysaccharide composition. These data demonstrate that *Aspergillus* species employ significantly different approaches to degrade plant biomass. Combining the approaches from different species is expected to result in improved enzyme mixtures for industrial applications, such as the saccharification of plant biomass for biofuel production.

Introduction

Plant biomass is the predominant carbon source for most fungi and consists largely of polymeric compounds, of which polysaccharides are the main components [15, 18]. In addition, lignin encrusts the polysaccharides and acts as a physical barrier that impedes fungal enzymes from gaining access to them. Fungi cannot take up intact polysaccharides, but need to degrade them extracellularly to monomeric and oligomeric compounds using diverse enzymatic mixtures [18, 55]. Plant polysaccharide degradation by fungi has been a topic of study for many decades due to its relevance in many industrial applications, such as paper & pulp, food & feed, beverages, textiles and detergents. More recently, the increasing interest in the production of alternative fuels and chemicals from plant biomass has provided an even greater push for research into fungal decomposition of plant biomass.

Analysis of an increasing number of fungal genome sequences has demonstrated the fundamental differences in the plant polysaccharide degrading machinery of fungi [2, 5, 8, 10, 19, 22]. In addition, the regulatory systems that control plant biomass degradation also differ strongly among fungi, although they are largely conserved among different *Aspergillus* species [29, 53]. Results from a previous study on the utilization of polysaccharides by three Aspergilli [10] suggest that related fungal species may have developed different approaches to plant biomass degradation. In nature, biomass-degrading fungi live in community with other microorganisms. It can be expected that different species target distinct components of the substrate and degrade them using dissimilar enzyme combinations. An enhanced understanding of these strategies will not only increase our knowledge of fungal biodiversity, but will help in the design of efficient industrial enzyme mixtures for plant biomass degradation. In this study, we compared the plant biomass degradation potential and approaches of eight *Aspergillus* species: *A. clavatus*, *A. fischeri*, *A. flavus*, *A. fumigatus*, *A. nidulans*, *A. niger*, *A. oryzae* and *A. terreus* (Suppl. Table 1).

Materials and methods

Media and growth conditions

The fungal strains used in this study are listed in Suppl. Table 1. *Aspergillus* minimal medium was described previously [17]. All monomeric and oligomeric carbon sources were added to a final concentration of 25 mM, while pure polymeric substrates and crude substrates were added to a final concentration of 1% and 3%, respectively. The pH of the medium was adjusted to 6.0. For plate growth, the centre of the plates was inoculated with 2 µl of a suspension of 500 spores/µl and plates were incubated for 5 days at 30°C. All eight species were grown on minimal medium with 35 carbon sources including crude plant biomass, pure plant polysaccharides, oligosaccharides, monosaccharides and control substrates (casein, lignin) (Suppl. Fig. 1, www.fung-

growth.org). To confirm that the detected differences were species specific, a second isolate of each species was examined along with the sequenced strain. Growth on 25 mM D-glucose was used as a reference because the tested strains grow at different rates and D-glucose, among the monosaccharides, supported the fastest growth for all species. Growth on the other substrates relative to growth on D-glucose was then compared among the species. Growth on plates was analyzed by visual inspection by two authors independently after which these were compared and discussed.

Liquid cultures were inoculated with 10^6 spores/ml (final concentration) and incubated at 250 rpm for 3 days. All cultures were incubated at 30°C and performed in duplicate. Two to three strains of all species were grown in liquid cultures with 1% wheat bran or 1% sugar beet pulp. Culture filtrates after three days of cultivation were analyzed for the presence of free monomeric sugars, but no glucose, xylose, galacturonic acid, rhamnose or fructose was detected. SDS-PAGE analysis of the extracellular proteins revealed nearly identical profiles for strains of the same species (Suppl. Fig. 6), indicating that enzyme production is highly conserved within a species. Detailed analysis of the produced enzymes was therefore only performed on a single strain.

Chemicals and media

Glucose, maltose, sucrose, inulin, beechwood xylan, Guar gum, apple pectin and all *p*-nitrophenyl-substrates were from Sigma–Aldrich. Soluble starch was from Difco. Red Debranched Arabinan (S-RDAR), Azo-CM-cellulose (S-ACMC), Azo-galactan (S-AGALP) and AZ-rhamnogalacturonan (S-AZRH), Azo-wheat arabinoxylan (S-AWAXP) and polygalacturonic acid (PGA) were from Megazyme International Ireland.

Composition analysis of plant biomass substrates

Sugar composition was determined by analyzing the sugars as their alditol acetate derivatives using GC-FLD as described previously [30].

CAZy annotation

The identification step of CAZymes followed the procedures previously described [10] where sequences are subject to BlastP analysis [1] against a library composed of modules derived from the CAZy database, the positive hits are then subjected to a modular annotation procedure that maps the individual modules onto the peptide using hits against libraries of catalytic and carbohydrate models derived from CAZy using BlastP or Hidden Markov models [1, 20]. The functional annotation step involves BlastP comparisons against a library of modules derived from biochemically characterized enzymes [10].

Orthology and synteny analysis

Genome scale protein ortholog clusters were constructed using OrthoMCL [34] by inflation factor 1, E-value cutoff 1E-3, percentage match cutoff 60% as for identification of distant homologs [9]. The orthologs clusters were further split according to the synteny detected by the Sybil algorithm [12] at www.aspgd.org. Sequences of genes were manually double checked by multiple sequence alignments with MAFFT[28] and potential errors of gene models were corrected.

Enzyme assays

All exo-acting CAZy enzyme activities were performed in micro titer plates. Reactions were carried out in 100 µL volumes containing 25 mM sodium acetate (pH 5), 0.01% substrate and suitably diluted culture filtrate. The mixture was incubated at 30°C for 2 h and the reaction was terminated by the addition of 100 µL 250 mM sodium carbonate. Enzyme activities (α-arabinofuranosidase, cellobiohydrolase, *α*-galactosidase, *β*-galactosidase, *α*-glucosidase, *β*-glucosidase, glucoamylase (*α*-maltosidase), *β*-mannosidase, *α*-rhamnosidase and *β*-xylosidase) were determined spectrophotometrically at 405 nm by measuring the release of *p*-nitrophenol (*p*NP) from their appropriate *p*NP-substrates and standardized against a known concentration of *p*-nitrophenol (*p*NP). Activities were expressed as nmol *p*NP/mL sample/min.

Endoarabinanase, *endo*-1,4-*β*-glucanase (cellulase), *endo*-1,4-*β*-galactanase and rhamnogalacturonanase activities were measured using 20 mg/mL of Red Debranched Arabinan (S-RDAR), Azo-CM-cellulose (S-ACMC), Azo-galactan (S-AGALP) and AZ-rhamnogalacturonan (S-AZRH), respectively. *Endo*-1,4-*β*-xylanase activity was measured using 10 mg/mL Azo-wheat arabinoxylan (S-AWAXP). 100 µL reactions were carried out containing equal volumes of buffered substrate (pH 4.5) and suitably diluted culture filtrate which were then incubated at 40°C for 1 h in the case of the endoarabinanase, *endo*-1,4-*β*-glucanase and *endo*-1,4-*β*-xylanase activities and 16 h for the *endo*-1,4-*β*-galactanase and rhamnogalacturonan activities. *Endo*-arabinanase reactions were terminated with the addition of 400 µL 95% ethanol, *endo*-1,4-*β*-galactanase, rhamnogalacturonanase and *endo*-1,4-*β*-xylanase reactions with 250 µL 95% ethanol and *endo*-1,4-*β*-glucanase reactions with a 250 µL solution of sodium acetate trihydrate (40 mg/mL) and zinc acetate (4 mg/mL) in 76% ethanol. Precipitated reactions were then centrifuged at 1000×g for 10 min and optical density of supernatants was measured at 590 nm. Endoarabinanase reactions were measured at 520 nm. *Endo*-acting enzyme activities are expressed as amount of dye released (absorbance change)/mL sample/min.

Pectate lyase activity was assayed using polygalacturonic acid (PGA). Reaction mixtures contained equal volumes of 50 mM N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) (pH 10.0) and 2.5 mg/mL PGA, to which suitably diluted culture filtrate

was added. Changes in absorbance at 235 nm were measured for approximately 30 min at 40°C.

Laccase activity was assayed using 2,2'-azino-di-(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS). Reaction mixtures contained 700 µL H_2O, 100 µL 0.5 M glycine-HCl (pH 3.0), 100 µL culture filtrate and 100 µL 14 mM ABTS. The reaction was monitored by measuring the change in absorption at 436 nm at 30°C. The extinction coefficient of 29,300 $M^{-1} \cdot cm^{-1}$ was used for oxidized ABTS. Activity is expressed as is in nmol/min/ml.

Feruloyl esterase activities were determined spectrophotometrically (Shimadzu PharmaSpec UV-1700) at 37°C in 100 mM MOPS (3-(*N*-morpholino)propanesulfonic acid buffer (pH 6.0). Methyl caffeate (MC), methyl ferulate (MF), methyl *p*-coumarate (MpC) and methyl sinapate (MS) (1.18 mM stock solutions in 100 mM MOPS, pH 6.0) were used as substrates. Reaction mixture contained 100 µL of culture liquid, 870 µL MOPS buffer and the reaction was started by the addition of 30 µL substrate. Absorbance was monitored for 5 min at 308 nm for MpC (ε_{308} = 20,390 M^{-1} cm^{-1}), 320 nm for MF (ε_{320} = 29,680 M^{-1} cm^{-1}) and MS (ε_{320} = 15,890 M^{-1} cm^{-1}), and 322 nm for MC (ε_{322} = 14,720 M^{-1} cm^{-1}). FAE activities were expressed as nkat/l (10^{-9} mol/s/L).

Determination of monomeric sugars in the cultures

Presence of monomeric sugars in wheat bran and sugar beet pulp liquid cultures was measured by using Megazyme's Assay kits for D-glucose & D-fructose (K-FRUGL), D-xylose (K-XYLOSE), D-glucuronic acid (K-URONIC) and L-rhamnose (K-RHAMNOSE) using the provider's instruction. All measurements were done with two biological replicates.

SDS-PAGE

Protein profiles were obtained by combining 25 µL of culture supernatant supplemented with 5 µL of 5 x Laemmli Loading Buffer (50 mM Tris–HCl pH 6.8, 2% SDS, 10% glycerol, 0.1 M dithiothreitol, 0.2 mg/mL Bromophenol Blue) and separating this on 12% SDS-PAGE gels. Proteins were visualized by silver staining and a PageRuler Unstained Protein Ladder (Thermo Scientific) was used as protein marker.

Proteomics analysis

Proteins from 3 ml of culture filtrate were precipitated with cold TCA/acetone and the amount of protein recovered was determined using the RCDC kit assay (BioRad, Mississauga, Ont). Five micrograms of protein were digested with trypsin and an aliquot analyzed by LC-MS/MS as previously described [44] on a Velos LTQ-Orbitrap mass spectrometer (Thermo-Fisher, San Jose, CA). MS/MS data were processed using

Proteome Discoverer Quant 1.3 (Thermo-Fisher) and spectral data were searched against *Aspergillus* protein databases downloaded from the Aspergillus Genome Database (AspGD). Search parameters used were 0.80 Da for fragment ion tolerance and 10.0 ppm for parent ion tolerance, fixed iodoacetamide cysteine modification and variable methionine oxidation. Protein and peptide identification confidence filters were applied to satisfy a 1% false discovery rate at the Peptide and Protein level. Protein grouping was applied so as to satisfy the principles of parsimony.

Hierarchical clustering and correlation analysis

Matrix files of presence/absence and activity of CAZy genes and protein production measured by proteomics experiments were generated. The hierarchical clustering of CAZy genes were created by R [48] using the Euclidean distance with complete linkage and visualized by iTOL [31, 32]. The dendrogram and heatmaps of protein abundance were created and visualized using Genesis [50] with Pearson's correlation and complete linkage.

Phylogenetic analysis

Sequences of the RPB1, RPB2 (RNA polymerase II genes), Tsr1 (putative ribosome biogenesis protein), Cct8 (putative chaperonin complex component TCP-1) and AguA (α-glucuronidase) genes were downloaded from the full genome data sets and aligned using the Muscle software in the MEGA5 package [52]. After aligning, the data sets were combined and maximum likelihood analysis was performed using RAxML version 7.2.8 [49]. Each locus was treated as a separate partition. The number of bootstrap replicates was set on 1000 replicates. Sequences of *Penicillium chrysogenum* Wisconsin 54-1255 were used as outgroup.

Results

Genomic potential of the studied Aspergilli related to plant biomass utilization

Based on the Carbohydrate-Active enZymes (CAZy) [35] annotation pipeline, total numbers of Glycoside Hydrolases (GH), Polysaccharide Lyases (PL) and Carbohydrate Esterases (CE) vary among the species (Fig. 1). The percentage of GH genes related to plant polysaccharide degradation (PPD) is 58-66% for all genomes, except that *A. clavatus* has 20-30% less GH genes than the others (Fig. 1), largely due to a reduction in pectinases (GH28, GH54, GH78, GH88) (Suppl. Table 2). *A. clavatus* also contains the lowest percentage of PPD-related PL genes (71% as compared to >86%), which are also all related to pectin degradation. When the genomes were compared for individual CAZy families, significant differences in numbers of genes were observed (Suppl. Table 3A&B). Variations in gene numbers are particularly obvious in certain CAZy families involved in the degradation of mannan (GH26),

pectin (GH28, GH53, GH78, GH88, GH93, PL1, PL3, CE8 and CE12), xyloglucan (GH29 and GH74), starch (GH31), sucrose/inulin (GH32), cellulose (GH45 and GH61), and xylan (GH115 and CE15). Genes encoding lignin peroxidase, manganese peroxidase or versatile peroxidase are not present in any of these genomes, but significant differences are found in the number of laccases and other oxidoreductase enzymes, which may play a role in lignin or polysaccharide degradation (Suppl. Table 3). *A. niger* is richest in laccases (13 in its genome), while the other species have 2-9. The variations in CAZy content are relatively small compared to previous studies with a more diverse set of fungal species [2, 5, 8, 10, 19, 22]. This can be explained by their close phylogenetic relationships and their similar habitats, which would push genome evolution in a similar direction.

Orthologous clustering of the CAZymes showed that only 14.7% of the genes encoding hydrolytic enzymes are shared by all species (Suppl. Table 4A). In contrast 27.5% of the genes are unique to a single species, with the largest number in *A. nidulans*, *A. niger* and *A. terreus*. For the oxidative enzymes, 10.8% of the genes are shared by all species, while 40.8% of the genes are unique to a single species, with again the largest number in *A. nidulans*, *A. niger* and *A. terreus* (Suppl. Table 4B). In general, the CAZyme distribution among the species follows their phylogenetic relationship. In total, this means that only 70 genes are shared by all species, while the number of unique genes differs strongly by species (Fig. 1).

Growth on plant biomass related substrates

Major differences of these eight species grown on mono- oligo- and polysaccharides were observed (Suppl. Fig. 1), indicating different carbon source preferences. Two isolates per species were tested to check that the differences are species specific and not strain specific. These carbon source preferences can be partially linked to their diverse genomic content of CAZyme-encoding genes. For instance, growth of *A. clavatus* was particularly poor on pectin which correlates well with its low number of 30 pectin-targeting genes, as compared to 55-92 for the other species (Suppl. Table 2). Good growth on inulin was observed for *A. niger* ATCC 1015 and *A. fischeri*, while for all other species growth was reduced on inulin, but not on sucrose. This does not correlate with the number of putative inulin/sucrose-targeting genes in the genomes, as *A. fischeri* has only two, while *A. niger* has four, fewer than the six genes for *A. terreus* which grows poorly on inulin (Suppl. Table 2).

All species grew well on wheat bran and also, with the exception of *A. clavatus*, on sugar beet pulp (Suppl. Fig. 1). These substrates were therefore selected to analyze their enzymatic ability in more detail. Composition analysis (Suppl. Table 5) showed that wheat bran contains mainly of cellulose and (arabino)xylan, with xyloglucan and

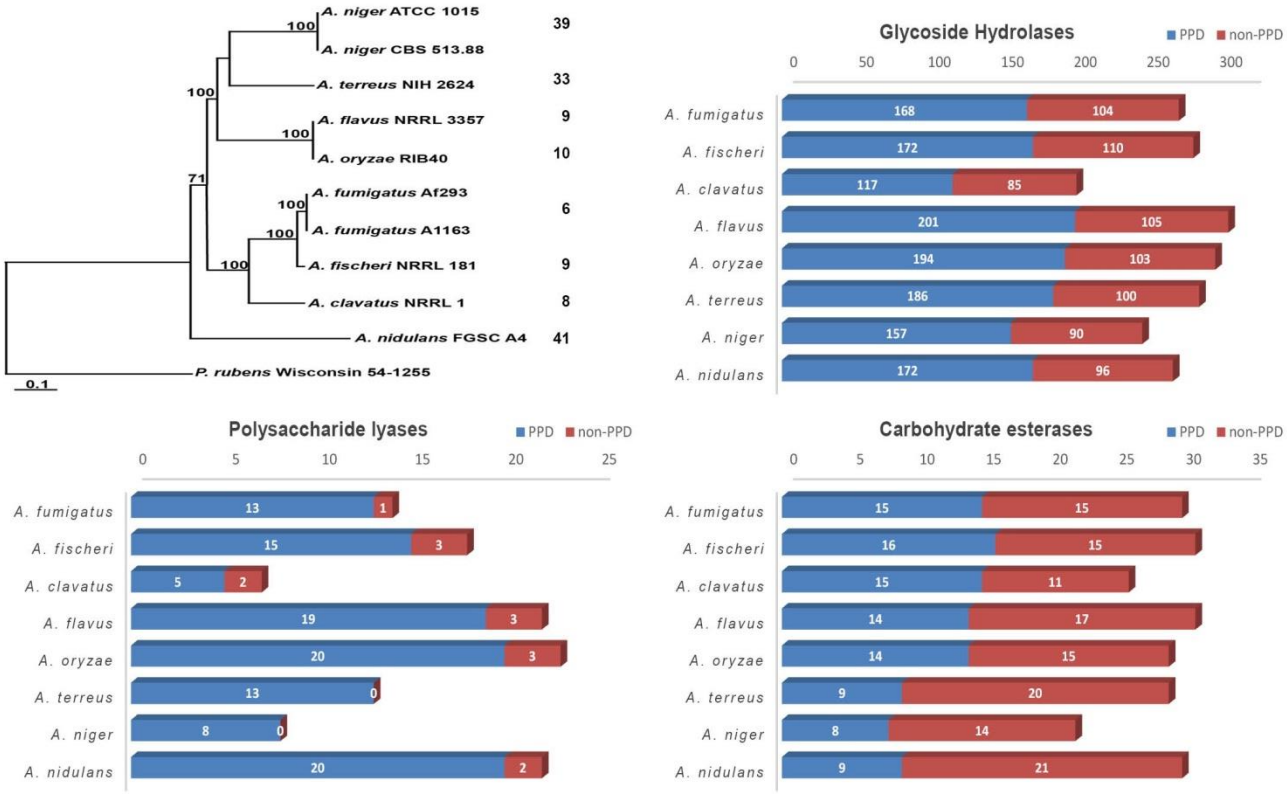


Figure 1: Taxonomic tree of the species used in this study and the numbers of glycoside hydrolases, polysaccharide lyases and carbohydrate esterases detected in their genomes. PPD = plant polysaccharide degradation related. The number of unique genes per species is indicated behind their name in the taxonomic tree.

pectin as minor components. In contrast, sugar beet pulp contains mainly cellulose, xyloglucan and pectin, which explains the reduced growth of *A. clavatus*.

Enzyme profiles during growth on wheat bran and sugar beet pulp

A preliminary test demonstrated that at Day 3 of cultivation, activities for eight plant biomass degrading enzymes were maximal for all fungi (data not shown). This time point was therefore selected for the full enzymatic analysis. Nineteen extracellular, lignocellulose-active enzyme activities of the liquid cultures were measured (Suppl. Fig. 2-4). Comparison of these profiles demonstrated strong differences among the species, not only in the quantities of the activities, but also in the induction of specific enzymes. For instance, the highest activity levels for most enzymes of *A. terreus* were observed during growth on sugar beet pulp, while wheat bran resulted in higher levels of most enzymes for *A. flavus* (Suppl. Fig. 2). When the individual activities were compared across the species, specific differences became noticeable. Wheat bran consists mainly of cellulose and arabinoxylan and the main regulator controlling degradation of these polysaccharides is XlnR, which is present in all Aspergilli [7]. Endoxylanase and *β*-xylosidase were mainly produced on wheat bran, and levels were particularly high for *A. niger* (Suppl. Fig. 2). Endoarabinanase, α-rhamnosidase, pectate lyase and endogalactanases, all related to pectin degradation, were mainly produced on sugar beet pulp, but rarely were all four activities produced by one species. Sugar beet pulp contains mainly cellulose and pectin and therefore pectinases and cellulases would be expected to be the main enzymes produced on this substrate, which is confirmed by our data.

Mass spectrometric analysis of the extracellular proteins confirmed the activity measurements with respect to the enzymes that were detected (Suppl. Table 6A-6D). Figure 2 and Suppl. Fig. 5 show the presence of orthologous enzymes involved in the degradation of different polysaccharides in wheat bran and sugar beet pulp. This analysis demonstrates the high degree of diversity among the species in the production of orthologous enzymes. Only a few orthologous enzymes are produced by all or most species and in most cases they are produced on both wheat bran and sugar beet pulp (Fig. 2) although often with significantly different levels (Suppl. Table 6A-6D). These data highlight the different enzymatic approaches used by the eight species to degrade plant biomass.

Correlation of CAZy profiles, taxonomy and enzyme activity of the eight Aspergilli

Figure 3 shows the correlation of the species for genome content, enzyme activity and production of individual enzymes. Correlating the number of genes per CAZy family demonstrated that with respect to genome content, closely related species (*A. oryzae* – *A. flavus*, *A. fischeri* – *A. fumigatus*) cluster together (Fig. 3A). This indicates that the

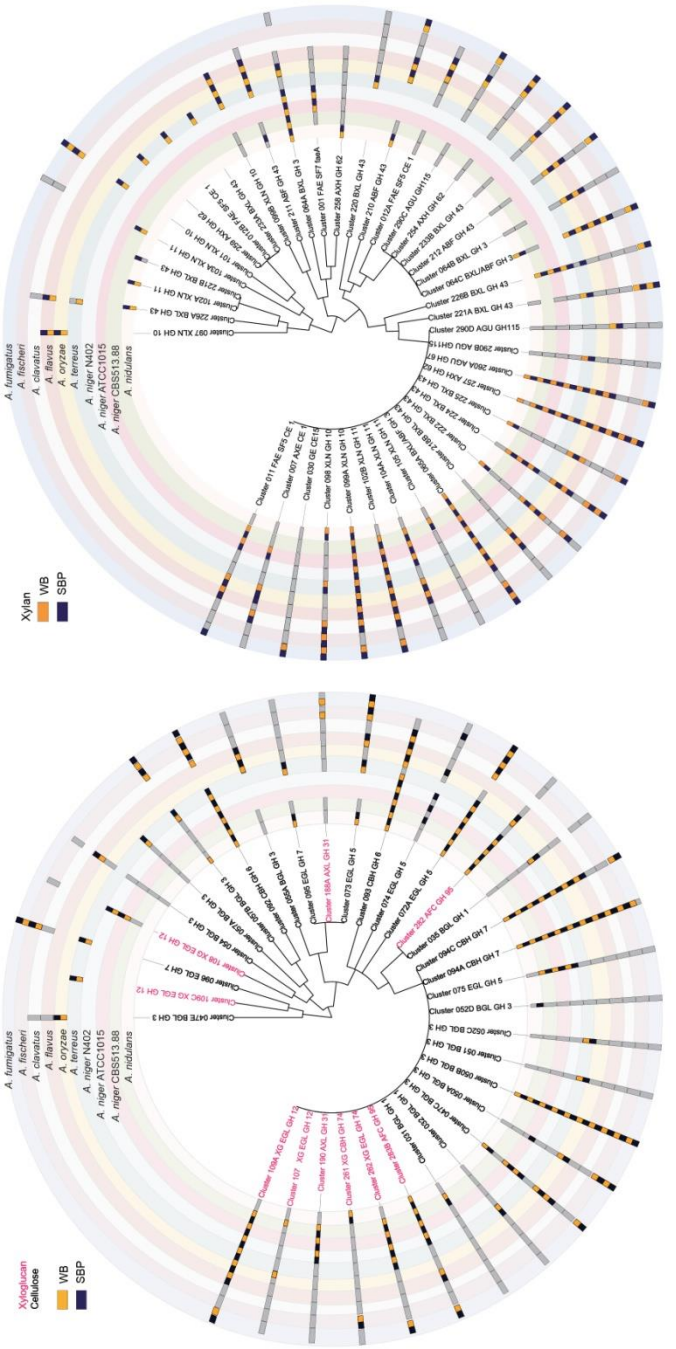


Figure 2: Proteins secreted by the eight *Aspergillus* species during growth on sugar beet pulp (SBP, purple) and wheat bran (WB, orange) as determined by mass spectrometry. Samples were taken after three days and are the same samples used for activity assays. The proteins are plotted using the ortholog clusters (Suppl. Table 4). Presence of the gene in a genome is depicted by a grey box in the circle corresponding to the species/strain.

evolution of their genome content related to plant biomass degradation follows the evolutionary history of the species. However, more distantly related species, *A. nidulans* and *A. terreus*, can display similar CAZy content. A possible explanation for this finding is that natural habitat exerts a stronger influence on genome evolution than phylogenetic relatedness.

No clear correlation was observed between the enzyme activities produced in response to complex substrates and evolutionary relatedness (Fig. 3B), possibly due to the range of non-plant substrates some species are known to consume (e.g. collagen for *A. terreus*, *A. flavus* and *A. nidulans* and insect larvae for *A. clavatus*), resulting in a varying biotope range and dependence on plant biomass. The composition of wheat bran and sugar beet pulp is different and they should elicit different activity profiles. For six of the tested species, the wheat bran and sugar beet pulp activity profiles diverge strongly. Unexpectedly, the sugar beet pulp and wheat bran activity profiles clustered together for *A. flavus* and *A. oryzae*. Two of the three tested *A. niger* strains (N402 and ATCC 1015) clustered together for both substrates, while the third (CBS 513.88) was strongly divergent in the enzyme activity profile. These results show that strains of the same species (CBS 513.88 and ATCC 1015) with near identical genomic content can use dramatically different sets of enzymes to hydrolyze complex biomass. It should be noted that genome sequence analysis suggests that strains ATCC 1015 and N402 are likely descended from the same isolate (A. Tsang and co-workers, unpublished data), which explains the clustering of their activity profiles.

Correlation of the proteomics data did not follow the activity correlation (Fig. 3B and 3C), which can be explained by the production of non-orthologous enzymes for the same general activity by different species (Suppl. Table 6A-6D). This adds an additional dimension to the highly divergent strategies of these Aspergilli. Considering the fairly similar genome content of these species, we conclude that the differences in their plant biomass degrading strategies are mainly at the regulatory level. More detailed studies into the regulation of orthologous CAZy-genes in several species could reveal whether this is due to different sets of target genes of the main regulators or whether additional unknown regulators modulate the influence of the main regulators.

Figure 3: Correlation analysis of the genome (A), enzyme activity (B) and proteomics data (C). Continued on next page.

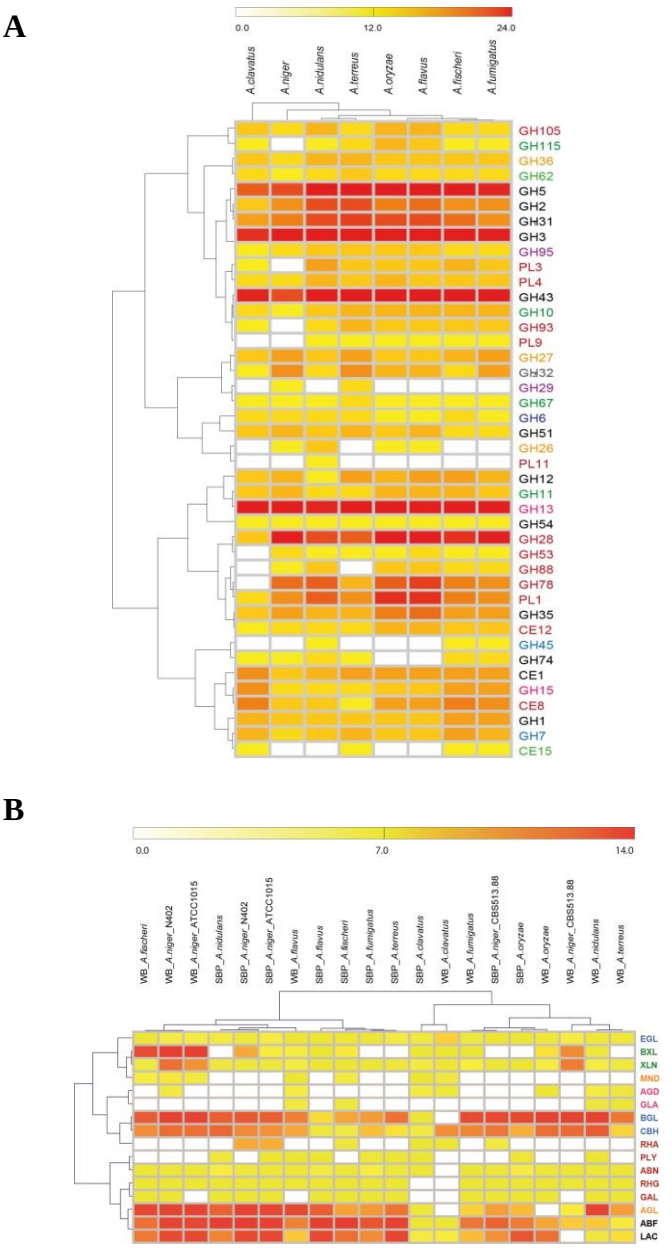
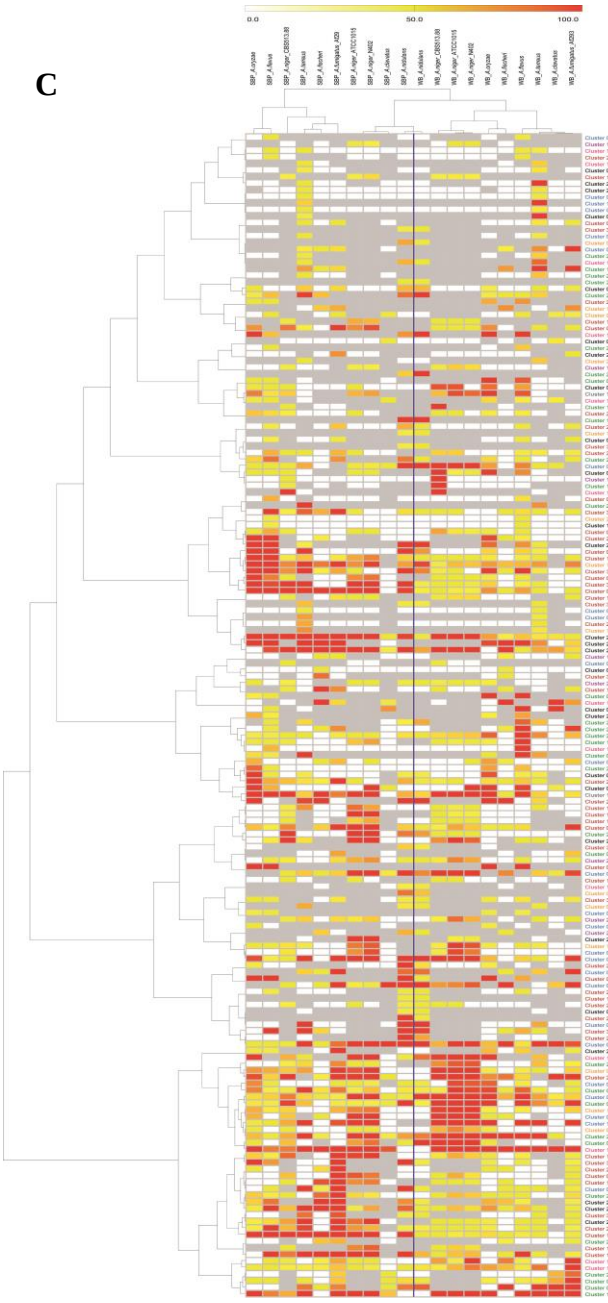


Figure 3 Cont: Correlation analysis of the genome (A), enzyme activity (B) and proteomics data (C). (Please see pdf file to enlargen text).

C



Discussion

In this study we compared eight Aspergilli with respect to plant polysaccharide degradation. The variations in CAZy content between these species was relatively low compared to previous studies in which a more diverse set of fungal species was compared [2, 5, 8, 10, 19, 22]. This can be explained by the close phylogenetic relationships and/or by and the highly similar habitats of these Aspergilli, which would direct genome evolution in a similar direction. Human use of and/or interaction with the species differs markedly; with *A. niger* and *A. oryzae* being widely used industrial fungi, *A. fumigatus* one of the most significant opportunistic fungal human pathogens, and *A. flavus* a plant pathogen. However, all these species are common inhabitants of soil and stored agricultural products, and their spores are widespread in both indoor and outdoor environments. Although some of the sequenced strains are domesticated and not recent natural isolates, the comparison to a second strain that is a natural isolate showed that growth on 35 carbon sources is identical for two strains of the same species. This demonstrates that the sequenced isolates have maintained their natural ability to use the tested carbon sources.

Hierarchical clustering of the plant polysaccharide degrading enzymes of these species demonstrated that in general the species with the most similar CAZy content are also taxonomically close.

The number of unique genes (Suppl. Table 2) per species also correlates well with the phylogenetic distance of the species. The lowest number was found for *A. oryzae*, *A. flavus*, *A. fischeri* and *A. fumigatus*. As the first two and the last two species, respectively, are closely related, their high similarity explains this low number. The more distant species (*A. nidulans*, *A. terreus*, *A. niger*, *A. clavatus*) have higher numbers of unique genes.

A high level of variation was detected in the enzyme activities of the tested species during growth on sugar beet pulp and wheat bran. While all species grew well on these substrates, with the exception of somewhat less growth of *A. clavatus* on sugar beet pulp, the enzyme profiles of the species showed strong differences. Wheat bran consists mainly of cellulose and arabinoxylan, while sugar beet pulp contains mainly cellulose and pectin. Enzymes able to degrade these different combinations of polysaccharides would therefore be expected to be prominent in the culture filtrate of all species grown on these substrates. Our study confirmed this as nearly all enzymes activities detected on wheat bran in all species are involved in xylan or cellulose degradation, while mainly pectinolytic and cellulolytic enzymes were detected on sugar beet pulp. The main regulator controlling the production of xylanolytic and cellulolytic enzymes in *Aspergillus* is XlnR, which has been studied in detail in *A. niger*, *A. oryzae* and *A. nidulans* [14, 27, 38, 39, 43, 51, 56, 57]. XlnR activates the expression of xylanolytic and cellulolytic genes in response to the presence of xylan or

xylose, the latter being the actual inducer. Indications for similar regulation have been reported for the other species [4, 13, 21, 25, 40, 45]. Regulation studies in *A. niger* have previously demonstrated that pectinolytic genes are induced by galacturonic acid, rhamnose, polygalacturonic acid or pectin [16, 26, 37]. Differences in pectinolytic gene content between *A. niger*, *A. nidulans* and *A. oryzae* may be influenced by the pH of their natural habitat [46]. An acidic pH favors pectin hydrolases, while a neutral to alkaline pH favors pectin lyases, supported by the finding that all fungal GH28 pectin hydrolases have activity optimum between pH 2 and pH 5, while pectin lyases have optimum between pH 7 and pH 10 (https://mycoclap.fungalgenomics.ca) [41]. The pH of most samples was 7 except for *A. nidulans* on sugar beet pulp (pH = 8), *A. niger* on wheat bran (pH = 5.5) and sugar beet pulp (pH = 4.5), and *A. clavatus* on sugar beet pulp (pH = 6). The pH in the sugar beet pulp cultures correlates well with the pectin hydrolase and lyase activities and with the proteomics results (Suppl. Table 6A-6D). Therefore, the differences in enzyme levels are likely caused by regulatory variation. Since the major regulators are shared by all tested species [53], their function or range of target genes in the tested Aspergilli is different and/or additional non-shared regulators are involved in the utilization of complex biomass. A difference in the function of the arabinanolytic regulator AraR in *A. niger* and *A. nidulans* was recently described [6], and the inducers for activation of AmyR also appear to differ between *A. niger*, *A. nidulans* and *A. oryzae* [58]. More detailed analysis of the set of target genes, and function and mechanism of the polysaccharide related regulators in the other species will be required to understand the mechanism responsible for these differences. Interestingly, the production of several cellulases appears to be conserved among the species, suggesting that this may be a core-activity for all species. In contrast, the production of hemicellulases is highly varied, suggesting specific adaptations of the species in their biomass degrading approach.

Laccase activity was detected for most species, with the highest activity on wheat bran for *A. flavus* and on sugar beet pulp for *A. fumigatus*. This does not correlate with the numbers of putative laccases detected in the genomes, suggesting significantly different regulation of the production of these enzymes among the species. Induction of laccase-encoding genes was mainly studied in basidiomycetes in which transcription is modulated by metal ions (Cu^{2+}, Ag^{+}, Mn^{2+}), aromatic compounds, nitrogen and carbon sources (nature and ratio) [47]. In ascomycetes, regulatory elements such as HSE (Heat shock elements), MRE (Metal response elements) and nitrogen metabolite regulation elements (NIT-2 like) were identified in the promoter region of laccase-like multicopper oxidase [33]. In addition, laccases are also involved in other biological processes, such as spore pigment formation [54], and not only in lignin degradation, so the total number of laccases likely does not reflect the number of laccases which play a role in plant biomass degradation.

Although the fungi tested in this study produce diverse enzyme sets, they all grow well on the crude plant biomass substrates. This suggests that different strategies for the degradation of plant biomass may be equally efficient (as measured by fungal growth). In biotechnological applications, such as biofuel production, complete hydrolysis of the plant biomass is difficult to achieve with currently available enzyme cocktails. This may in part be explained by the absence of specific activities in these mixtures. The data obtained in this study show the existence of distinctly different enzymatic approaches to degrade biomass. A judicious mix of these approaches is likely to result in improved enzyme cocktails for biomass hydrolysis. Recently it was shown that addition of *Podospora anserina* hydrolases increases the efficiency of a *Trichoderma reesei* enzyme mixture [11]. In this study we provide indications that similar results could be obtained with more closely related fungi. The advantage of using enzymes from other Aspergilli to improve enzyme cocktails of *A. niger* or *A. oryzae* is that heterologous production of these enzymes is not likely to cause problems due to the high similarity in gene structure of these species.

Acknowledgments

IB, GAO, EB and BSG were supported by grants of the Dutch Technology Foundation STW, Applied Science division of NWO and the Technology Program of the Ministry of Economic Affairs UGC 07063 & 07938 to RPdV. IB was also supported by a grant of Kluyver Centre for Genomics of Industrial Fermentation KCII 2.2.27 to RPdV. GAO was also supported by a grants from DGAPA-UNAM number IN225710 and IN 219813. JvdB was supported by a grant from the Netherlands Organisation for Scientific Research (NWO) JSTP.10.005 to RPdV. MZ was supported by a grant from the Netherlands Organisation for Scientific Research (NWO) and the Netherlands Genomics Initiative 93511035 to RPdV. The comparative genomics study was carried out on the Dutch national e-infrastructure with the support of SURF Foundation (e-infra130078). The proteomic work was supported by Genome Canada and Génome Québec.

References

1. **Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389-3402.
2. **Amselem, J., Cuomo, C.A., van Kan, J.A., Viaud, M., Benito, E.P., Couloux, A., et al.** 2011. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet. **7**: e1002230.
3. **Andersen, M.R., Salazar, M.P., Schaap, P.J., van de Vondervoort, P.J., Culley, D., Thykaer, J., et al.** 2011. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res. **21**: 885-897.
4. **Bakri, Y., Masson, M. and Thonart, P.** 2010. Isolation and identification of two new fungal strains for xylanase production. Appl Biochem Biotechnol. **162**: 1626-1634.
5. **Battaglia, E., Benoit, I., van den Brink, J., Wiebenga, A., Coutinho, P.M., Henrissat, B., et al.** 2011. Carbohydrate-active enzymes from the zygomycete fungus *Rhizopus oryzae*: a highly specialized approach to carbohydrate degradation depicted at genome level. BMC Genomics. **12**: 38.
6. **Battaglia, E., Hansen, S.F., Leendertse, A., Madrid, S., Mulder, H., Nikolaev, I., et al.** 2011. Regulation of pentose utilisation by AraR, but not XlnR, differs in *Aspergillus nidulans* and *Aspergillus niger*. Appl Microbiol Biotechnol. **91**: 387-397.
7. **Battaglia, E., Visser, L., Nijssen, A., van Veluw, J., Wösten, H.A.B. and de Vries, R.P.** 2011. Analysis of regulation of pentose utilisation in *Aspergillus niger* reveals evolutionary adaptations in the Eurotiales. Stud Mycol. **69**: 31-38.
8. **Berka, R.M., Grigoriev, I.V., Otillar, R., Salamov, A., Grimwood, J., Reid, I., et al.** 2011. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. Nat Biotechnol. **29**: 922-927.
9. **Boekhorst, J. and Snel, B.** 2007. Identification of homologs in insignificant blast hits by exploiting extrinsic gene properties. BMC Bioinformatics. **8**: 356.
10. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.
11. **Couturier, M., Haon, M., Coutinho, P.M., Henrissat, B., Lesage-Meessen, L. and Berrin, J.G.** 2010. *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. Appl Environ Microbiol. **77**: 237-246.

12. **Crabtree, J., Angiuoli, S.V., Wortman, J.R. and White, O.R.** 2007. Sybil: methods and software for multiple genome comparison and visualization. Methods Mol Biol. **408**: 93-108.
13. **de Souza, C.G., Girardo, N.S., Costa, M.A. and Peralta, R.M.** 1999. Influence of growth conditions on the production of xylanolytic enzymes by *Aspergillus flavus*. J Basic Microbiol. **39**: 155-160.
14. **de Vries, R.P., Visser, J. and de Graaff, L.H.** 1999. CreA modulates the XlnR-induced expression on xylose of *Aspergillus niger* genes involved in xylan degradation. Res Microbiol. **150**: 281-285.
15. **de Vries, R.P. and Visser, J.** 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microb Mol Biol Rev. **65**: 497-522.
16. **de Vries, R.P., Jansen, J., Aguilar, G., Parenicová, L., Benen, J.A.E., Joosten, V., et al.** 2002. Expression profiling of pectinolytic genes from *Aspergillus niger*. FEBS Lett. **530**: 41-47.
17. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J.I., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
18. **de Vries, R.P., Nayak, V., van den Brink, J., Vivas Duarte, A.D. and Stalbrand, H.** 2012. Fungal degradation of plant oligo- and polysaccharides. In: Grunwald P, (ed) Carbohydrate Modifying Biocatalysts, Pan Stanford Publishing Pte, Ltd., Singapore, pp. 693-759.
19. **Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., et al.** 2011. The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science. **333**: 762-765.
20. **Eddy, S.R.** 1998. Profile hidden Markov models. Bioinformatics. **14**: 755-763.
21. **Elshafei, A.M., Hassan, M.M., Haroun, B.M., Abdel-Fatah, O.M., Atta, H.M. and Othman, A.M.** 2009. Purification and properties of an endoglucanase of *Aspergillus terreus* DSM 826. J Basic Microbiol. **49**: 426-432.
22. **Espagne, E., Lespinet, O., Malagnac, F., Da Silva, C., Jaillon, O., Porcel, B.M., et al.** 2008. The genome sequence of the model ascomycete fungus *Podospora anserina*. Genome Biol. **9**: R77.
23. **Fedorova, N.D., Khaldi, N., Joardar, V.S., Maiti, R., Amedeo, P., Anderson, M.J., et al.** 2008. Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet. **4**: e1000046.
24. **Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., et al.** 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature. **438**: 1105-1115.

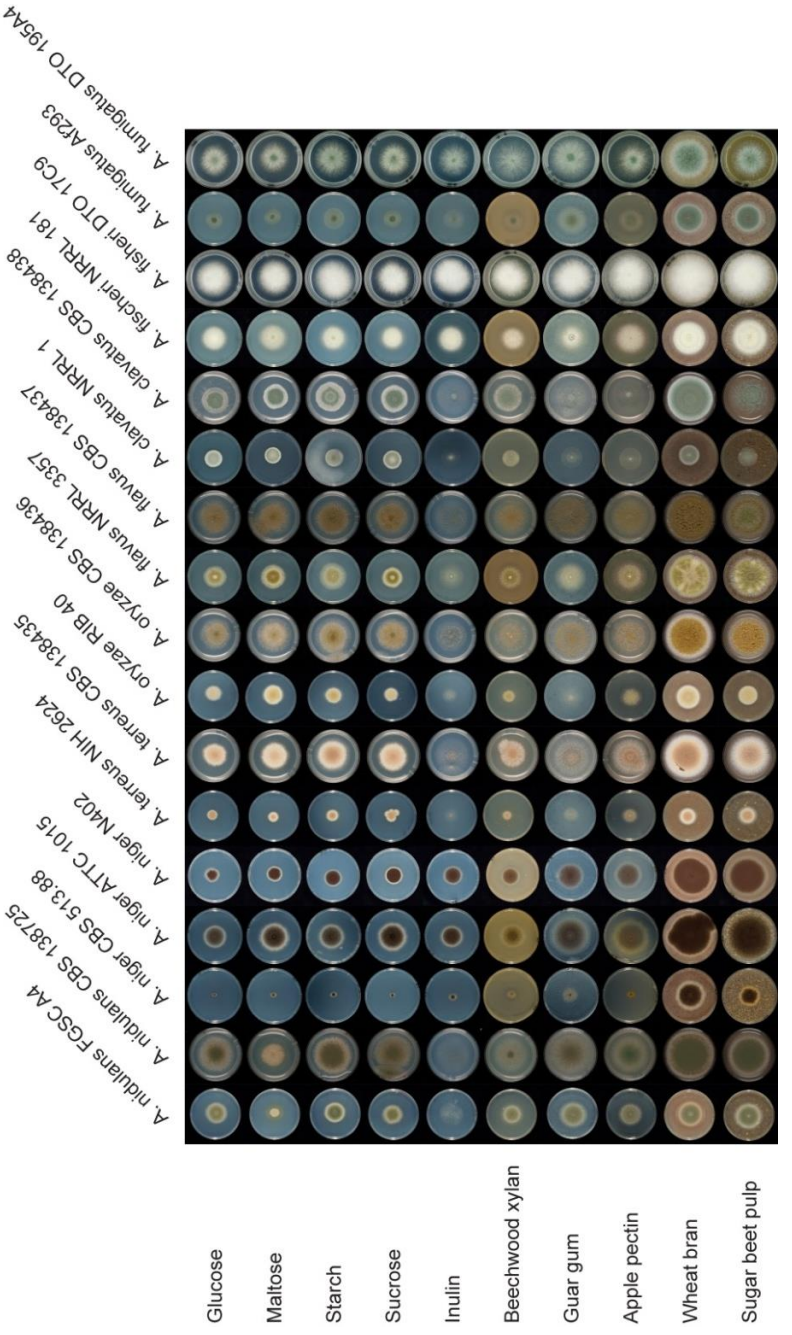
25. **Gawande, P.V. and Kamat, M.Y.** 1999. Production of *Aspergillus* xylanase by lignocellulosic waste fermentation and its application. J Appl Microbiol. **87**: 511-519.
26. **Gruben, B.S., Zhou, M., Wiebenga, A., Ballering, J., Overkamp, K.M., Punt, P.J., et al.** 2014. *Aspergillus niger* RhaR, a regulator involved in L-rhamnose release and catabolism. Appl Microbiol Biotechnol. **98**: 5531-5540.
27. **Hasper, A.A., Trindade, L.M., van der Veen, D., van Ooyen, A.J. and de Graaff, L.H.** 2004. Functional analysis of the transcriptional activator XlnR from *Aspergillus niger*. Microbiology. **150**: 1367-1375.
28. **Katoh, K., Kuma, K.-I., Toh, H. and Miyata, T.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. **33**: 511-518.
29. **Kowalczyk, J.E., Benoit, I. and de Vries, R.P.** 2014. Regulation of plant biomass utilization in *Aspergillus*. Adv Appl Microbiol. **88**: 31-56.
30. **Kuhnel, S., Schols, H.A. and Gruppen, H.** 2011. Aiming for the complete utilization of sugar-beet pulp: Examination of the effects of mild acid and hydrothermal pretreatment followed by enzymatic digestion. Biotechnol Biofuels. **4**: 14.
31. **Letunic, I. and Bork, P.** 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. **23**: 127-128.
32. **Letunic, I. and Bork, P.** 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. **39**: W475-478.
33. **Levasseur, A., Saloheimo, M., Navarro, D., Andberg, M., Pontarotti, P., Kruus, K., et al.** 2010. Exploring laccase-like multicopper oxidase genes from the ascomycete Trichoderma reesei: a functional, phylogenetic and evolutionary study. BMC Biochem. **11**: 32.
34. **Li, L., Stoeckert, C.J., Jr. and Roos, D.S.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13**: 2178-2189.
35. **Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B.** 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. **42**: D490-495.
36. **Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., et al.** 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature. **438**: 1157-1161.
37. **Martens-Uzunova, E.S. and Schaap, P.J.** 2009. Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. Fungal Genet Biol. **46 Suppl 1**: S170-S179.
38. **Marui, J., Kitamoto, N., Kato, M., Kobayashi, T. and Tsukagoshi, N.** 2002. Transcriptional activator, AoXlnR, mediates cellulose-inductive expression of the xylanolytic and cellulolytic genes in *Aspergillus oryzae*. FEBS Lett. **528**: 279-282.

39. **Marui, J., Tanaka, A., Mimura, S., de Graaff, L.H., Visser, J., Kitamoto, N., et al.** 2002. A transcriptional activator, AoXlnR, controls the expression of genes encoding xylanolytic enzymes in *Aspergillus oryzae*. Fungal Genet Biol. **35**: 157-169.
40. **Mellon, J.E., Cotty, P.J., Callicott, K.A. and Abbas, H.** 2011. Identification of a major xylanase from *Aspergillus flavus* as a 14-kD protein. Mycopathologia. **172**: 299-305.
41. **Murphy, C., Powlowski, J., Wu, M., Butler, G. and Tsang, A.** 2011. Curation of characterized glycoside hydrolases of fungal origin. Database (Oxford). **2011**: bar020.
42. **Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., et al.** 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. **438**: 1151-1156.
43. **Noguchi, Y., Sano, M., Kanamaru, K., Ko, T., Takeuchi, M., Kato, M., et al.** 2009. Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol. **85**: 141-154.
44. **Ozturkoglu Budak, S., Zhou, M., Brouwer, C., Wiebenga, A., Benoit, I., Di Falco, M., et al.** 2014. A genomic survey of proteases in Aspergilli. BMC Genomics. **15**: 523.
45. **Peixoto-Nogueira Sde, C., Michelin, M., Betini, J.H., Jorge, J.A., Terenzi, H.F. and Polizeli Mde, L.** 2009. Production of xylanase by Aspergilli using alternative carbon sources: application of the crude extract on cellulose pulp biobleaching. J Ind Microbiol Biotechnol. **36**: 149-155.
46. **Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., et al.** 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol. **25**: 221-231.
47. **Piscitelli, A., Giardina, P., Lettera, V., Pezzella, C., Sannia, G. and Faraco, V.** 2011. Induction and transcriptional regulation of laccases in fungi. Curr Genomics. **12**: 104-112.
48. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2009.
49. **Stamatakis, A., Hoover, P. and Rougemont, J.** 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol. **57**: 758-771.
50. **Sturn, A., Quackenbush, J. and Trajanoski, Z.** 2002. Genesis: cluster analysis of microarray data. Bioinformatics. **18**: 207-208.
51. **Tamayo, E.N., Villanueva, A., Hasper, A.A., de Graaff, L.H., Ramon, D. and Orejas, M.** 2008. CreA mediates repression of the regulatory gene *xlnR* which controls the production of xylanolytic enzymes in *Aspergillus nidulans*. Fungal Genet Biol. **45**: 984-993.
52. **Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S.** 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood,

evolutionary distance, and maximum parsimony methods. Mol Biol Evol. **28**: 2731-2739.
53. **Todd, R.B., Zhou, M., Ohm, R.A., Leeggangers, H.A., Visser, L. and de Vries, R.P.** 2014. Prevalence of transcription factors in ascomycete and basidiomycete fungi. BMC Genomics. **15**: 214.
54. **Tsai, H.F., Wheeler, M.H., Chang, Y.C. and Kwon-Chung, K.J.** 1999. A developmentally regulated gene cluster involved in conidial pigment biosynthesis in *Aspergillus fumigatus*. J Bacteriol. **181**: 6469-6477.
55. **van den Brink, J. and de Vries, R.P.** 2011. Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol. **91**: 1477-1492.
56. **van Peij, N., Gielkens, M.M.C., de Vries, R.P., Visser, J. and de Graaff, L.H.** 1998. The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol. **64**: 3615-3619.
57. **van Peij, N.N., Visser, J. and de Graaff, L.H.** 1998. Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. Mol Microbiol. **27**: 131-142.
58. **vanKuyk, P.A., Benen, J.A., Wosten, H.A., Visser, J. and de Vries, R.P.** 2012. A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol. **93**: 285-293.
59. **Yu, J., Payne, G.A., Nierman, W.C., Machida, M., Bennett, J.W., Campbell, B.C., et al.** 2008. *Aspergillus flavus* genomics as a tool for studying the mechanism of aflatoxin formation. Food Addit Contam Part A Chem Anal Control Expo Risk Assess. **25**: 1152-1157.

Supplementary Figure 1: Growth profiling of eight Aspergilli on plant-biomass related carbon sources. Growth of the eight *Aspergillus* species was evaluated on 35 plant biomass related carbon sources (full profiles are available at www.fung-growth.org). Minimal medium (MM) [17] was used supplemented with 25 mM of mono- or disaccharide, 1% polysaccharide or 3% crude plant biomass. Strains were grown for five days after which pictures were taken immediately. A second strain was also assessed for each species to confirm that the differences were species specific. The general growth speed differed between the strains of a species, but for most species no significant carbon source related differences were observed between the strains (data not shown). An exception to this is *A. niger* CBS513.88 that grew poorly on all pure carbon sources and was shown to have an amino acid auxotrophy (unpublished data), which explains this phenotype. Apparently, both wheat bran (WB) and sugar beet pulp (SBP) contain sufficient protein/amino acids to supplement this deficiency. All other strains grew well on MM + glucose and glucose was therefore used as an internal reference to compare the strains, to avoid misleading differences caused by general differences in growth speed between the species. Growth on the other substrates relative to growth on D-glucose was then compared between the species.

Growth on pure cellulose was zero to very poor for all species. Most fungi had similar growth on glucose, maltose, starch and xylan, with the exception of *A. nidulans*, for which poor growth on maltose was observed for one strain. Growth on sucrose was similar to growth on glucose for nearly all strains, but strong differences were observed on inulin. Good growth on this substrate was observed for *A. niger* ATCC1015 and *A. fischeri*, while for all other species growth was reduced compared to sucrose. This does not correlate with the number of putative inulin-targeting genes in the genomes as *A. fischeri* has only two (Table 2), while *A. niger* has six just like *A. terreus*, which grows poorly on inulin. This suggests that not all GH32 proteins may be involved in inulin degradation. Good growth compared to the other species was observed for *A. niger*, *A. nidulans*, *A. fumigatus* and *A. flavus* on guar gum (galactomannan). While *A. nidulans* has the highest number of galactomannan-targeting genes (Table 2), *A. niger* has in fact the lowest number and *A. terreus* again grows poorly, even though it has the second highest number of galactomannan-targeting genes. Growth of *A. clavatus* was particularly poor on pectin which correlates well with its very low number of pectin-targeting genes (only 30 putative genes), which was less than half of the other species (Suppl. Table 2).
The monomeric composition of the plant biomass substrates was determined and the polysaccharides present in the substrates were estimated based on this analysis (Suppl. Table 5). Wheat bran and cotton seed pulp contain mainly cellulose and (arabino)xylan, with xyloglucan and pectin as minor components of cotton seed pulp. In contrast, sugar beet pulp contains mainly cellulose, xyloglucan and pectin. (Continued on next page).


A. nidulans FGSC A4
A. nidulans CBS 138725
A. niger CBS 513.88
A. niger ATTC 1015
A. niger N402
A. terreus NIH 2624
A. terreus CBS 138435
A. oryzae RIB 40
A. oryzae CBS 138436
A. flavus NRRL 3357
A. flavus CBS 138437
A. clavatus NRRL 1
A. clavatus CBS 138438
A. fischeri NRRL 181
A. fischeri DTO 17C9
A. fumigatus AF293
A. fumigatus DTO 195A4
Glucose
Maltose
Starch
Sucrose
Inulin
Beechwood xylan
Guar gum
Apple pectin
Wheat bran
Sugar beet pulp

Supplementary Figure 2: Hydrolytic enzyme activity profiles of the eight species. Enzyme profiles of the eight *Aspergillus* species during growth on sugar beet pulp (SBP, red) and wheat bran (WB, blue). Samples were taken after three days and are identical to the samples used for proteomics. Vertical lines separate the activities related to the same substrate; from left to right: cellulose (C), xylan (X), galactomannan (G), starch (S), pectin (P). BGL = *β*-glucosidase, CBH = cellobiohydrolase, EGL = endoglucanase, BXL = *β*-xylosidase, XLN = endoxylanase, MND = *β*-mannosidase, AGL = *α*-galactosidase, AGD = *α*-glucosidase, GLA = glucoamylase, PLY = pectate lyase, RHG = endorhamnogalacturonase, RHA = α-rhamnosidase, ABN = endoarabinanase, ABF = *α*-arabinofuranosidase, GAL = endogalactanases, LAC = *β*-galactosidase.

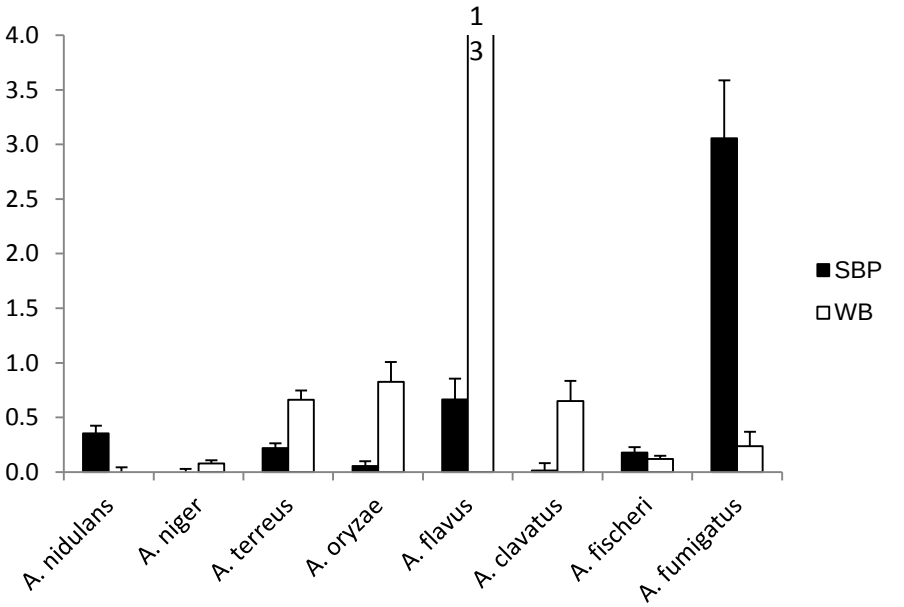
Activity units were: For all *exo*-acting enzyme activities (ABF, CBH, AGL, LAC, AGD, BGL, GLA, MND, RHA, BXL) are expressed as nmol *p*NP released/ml sample/min. *Endo*-acting enzyme activities (ABN, EGL, GAL, XLN and RHG) are expressed as amount of dye released (absorbance change)/ml sample/min. Pectate lyase (PLY) activity is expressed as absorbance change/ml sample/min.

Significant differences can be observed between the profiles both with respect to the relative activity of the different enzymes as well as to the production of the different activities on SBP and/or WB. Related species do not have high similarity with respect to enzyme activity profiles. Significant differences were observed between *A. oryzae* and *A. flavus*, and between *A. fumigatus* and *A. fischeri*. The reduction of pectinase encoding genes in the *A. clavatus* genome is reflected in very low pectinolytic enzyme activity. (Continued on next page).


C X G S P
A. terreus
A. flavus
A. niger (ATCC 1015)
A. oryzae
A. niger (CBS513.88)
A. clavatus
A. niger (N402)
A. fumigatus
A. nidulans
A. fischeri
BGL CBH EGL BXL XLN MND AGL AGD GLA PLY RHG RHA ABN ABF GAL LAC

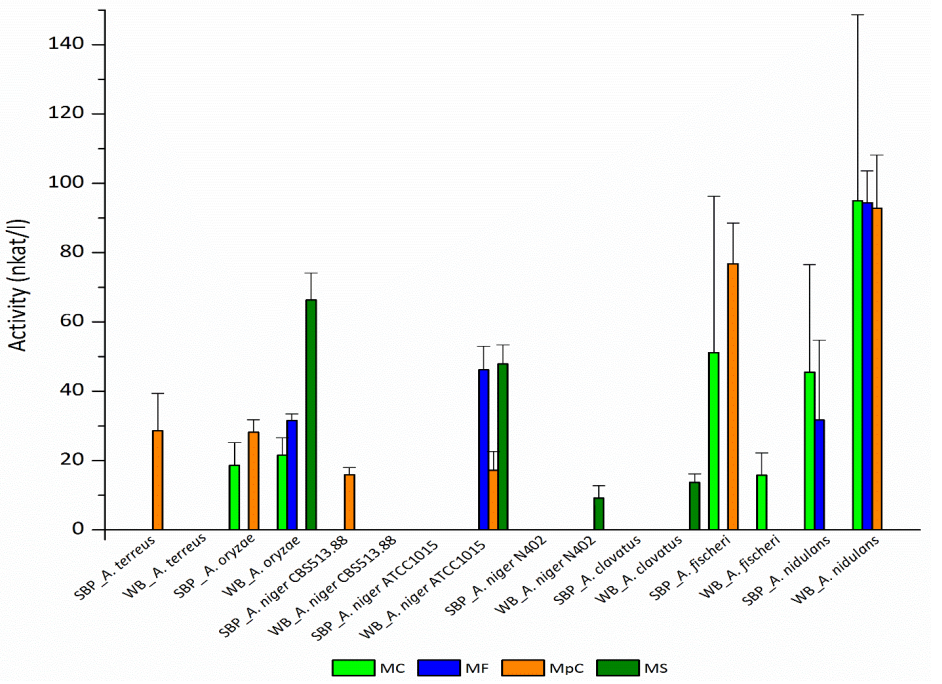
Supplementary Figure 3: Laccase activity of the eight species. Laccase activity of the eight *Aspergillus* species during growth on sugar beet pulp (SBP, red) and wheat bran (WB, blue). Samples were taken after three days and are identical to the samples used for proteomics. Laccase activity is in nmol/min/ml. *A. niger* N402 was used for these assays.

The highest laccase activity was observed for *A. flavus* during growth on WB and for *A. fumigatus* during growth on SBP. Activities in the other species were significantly lower.



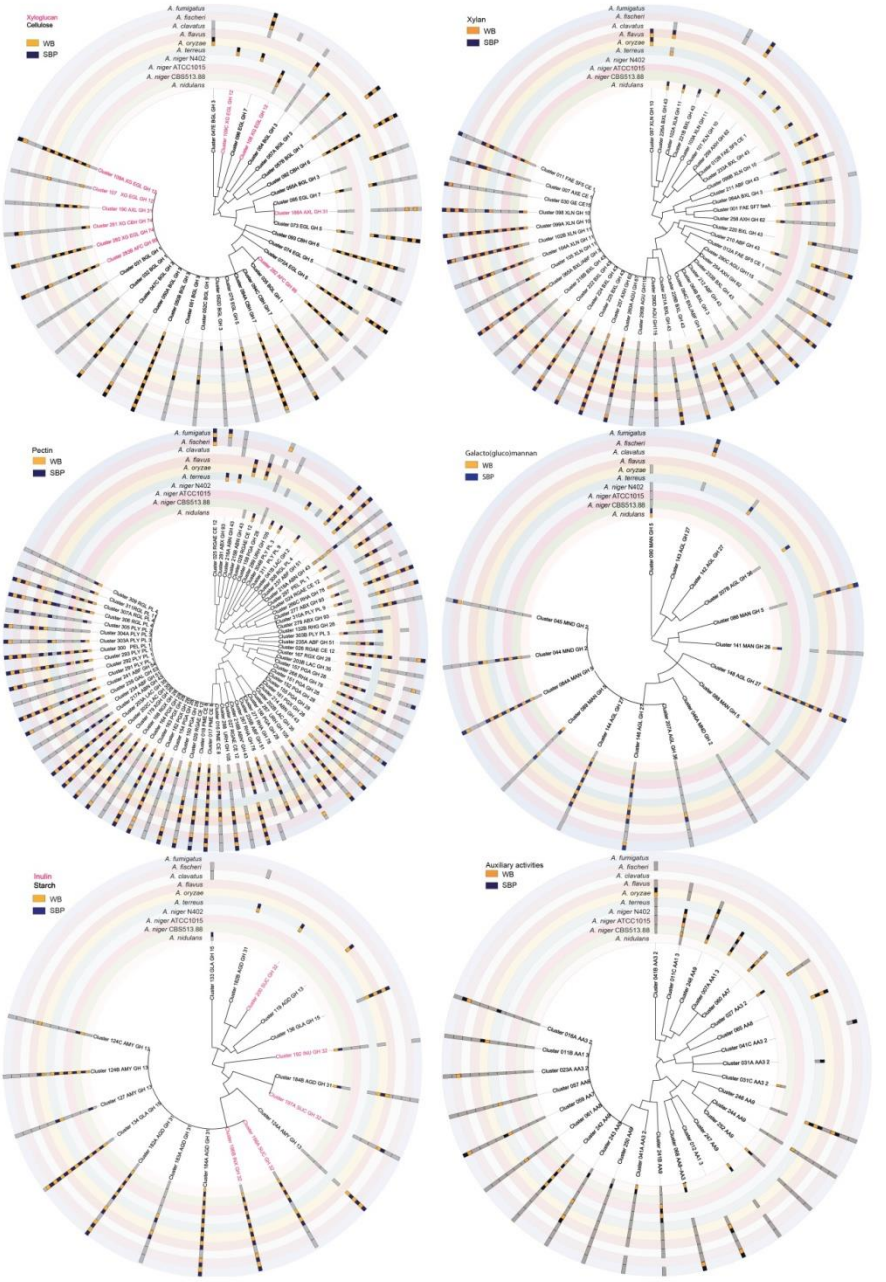
Supplementary Figure 4: Differences in feruloyl esterase production. Fungal strains were grown in 50 ml liquid minimal medium [17] with 1% wheat bran (WB) or sugar beet pulp (SBP) in 250 ml Erlenmeyer flasks. Culture filtrate samples were taken on day 3 and used for enzyme assays. Feruloyl esterase activities were determined spectrophotometrically at 37 °C in 100 mM MOPS buffer (pH 6) using methyl caffeate (MC), methyl ferulate (MF), methyl *p*-coumarate (MpC) and methyl sinapate (MS) as substrates. Absorbance was monitored for 5 min at 308 nm for MpC (ε_{308} = 20,390 M^{-1} cm^{-1}), 320 nm for MF (ε_{320} = 29,680 M^{-1} cm^{-1}) and MS (ε_{320} = 15,890 M^{-1} cm^{-1}), and 322 nm for MC (ε_{322} = 14,720 M^{-1} cm^{-1}).

Strongly divergent FAE activity profiles were observed for the studied Aspergilli. Differences were observed with respect to the carbon source that induced the activities as well as the substrate that was converted in the assays. The highest FAE activities were detected in *A. nidulans* WB cultures, but no activity was observed against MS. Both *A. terreus* and *A. niger* CBS513.88 strains produced FAE activity only in SBP cultures with M*p*C as substrate. In contrast, *A. niger* N402 and *A. clavatus* produced FAE activity only in the WB cultures with MS as a substrate. While no activity was detected in the SBP cultures of *A. niger* ATCC1015, it produced activities against MF, M*p*C and MS in WB cultures. FAE activity against MC was detected in both SBP and WB cultures of *A. oryzae*, *A. fischeri* and *A. nidulans*. For other substrates, the activity profiles differed between these strains. As it is unlikely that the substrate specificity of orthologous enzymes would differ this much with respect to these four substrates, the data implies that different feruloyl esterases are produced by the strains.

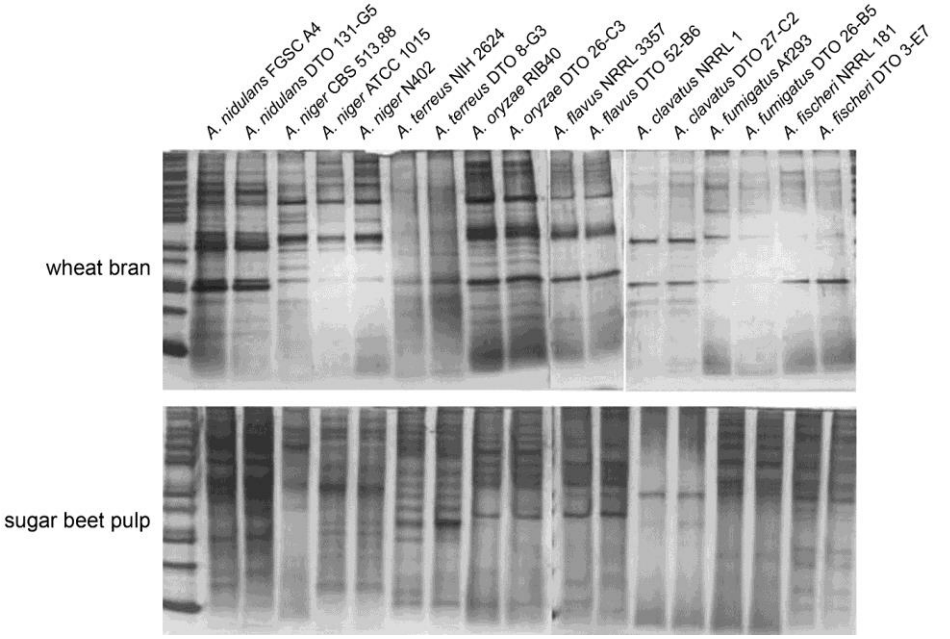


Supplementary Figure 5: Proteomic analysis of the presence of CAZy enzymes during growth on wheat bran and sugar beet pulp. Proteomics data of the eight *Aspergillus* species during growth on sugar beet pulp (SBP, purple) and wheat bran (WB, orange). Samples were taken after three days and are identical to the samples used for activity assays. The proteins are plotted using the ortholog clusters (Suppl. Table 4). The dendrogram was created with the presence/absence pattern of orthologous genes by R using Euclidean distance with complete linkage. The resulting tree was visualized by iTOL. Presence of the gene in a genome is depicted by a grey box in the circle corresponding to the species/strain. An orange colour indicates that the protein was detected during growth on wheat bran, while a purple colour indicates presence in sugar beet cultures.

These figures demonstrate the high diversity of the *Aspergillus* species, not only with respect to presence/absence of orthologs, but also with respect to the production of orthologous enzymes during growth on plant biomass. (Continued on next page; Please see pdf file to enlargen text).


Xyloglucan
Cellulose
WB
SBP
A. fumigatus
A. fischeri
A. clavatus
A. flavus
A. oryzae
A. terreus
A. niger N402
A. niger ATCC1015
A. niger CBS513.88
A. nidulans
Xylan
Pectin
Galacto(gluco)mannan
Inulin
Starch
Auxiliary activities

Supplementary Figure 6: Conserved SDS-PAGE profiles for isolates of the same species.
SDS-PAGE profiles of the strains used in this study. Extracellular culture samples from the wheat bran and sugar beet cultures were separated by SDS-PAGE and the gels were stained using silver staining. Two or three isolates per species were analysed, which demonstrated high conservation of the extracellular protein profile within a species. Larger differences were visible between the species. The profiles of more closely related species (*A. oryzae* – *A. flavus* and *A. fumigatus* – *A. fischeri*) were more similar to each other than to the profiles of the other species.



Supplemental Table 1: Strains used in this study.

Species	Strain number	Alternative strain number	Genome sequence reference
A. nidulans	FGSC A4	ATCC 38163	[24]
A. nidulans		DTO 131-G5	n/a
A. niger	CBS 513.88		[46]
A. niger	ATTC 1015	CBS 113.46, NRRL 328, FGSC A1144	[3]
A. niger	N402		n/a
A. terreus	NIH 2624	FGSC A1156	Unpublished
A. terreus		DTO 8-G3	n/a
A. oryzae	RIB 40	ATCC 42149	[36]
A. oryzae		DTO 26-C3	n/a
A. flavus	NRRL 3357	CBS 128202, ATCC 200026	[59]
A. flavus		DTO 52-B6	n/a
A. clavatus	NRRL 1	CBS 513.65, ATCC 1007	[23]
A. clavatus		DTO 27-C2	n/a
A. fischeri	NRRL 181	CBS 544.65, ATCC 1020	[23]
A. fischeri		DTO 3-E7	n/a
A. fumigatus	Af293	FGSC A1435	[42]
A. fumigatus		DTO 26-B5	n/a

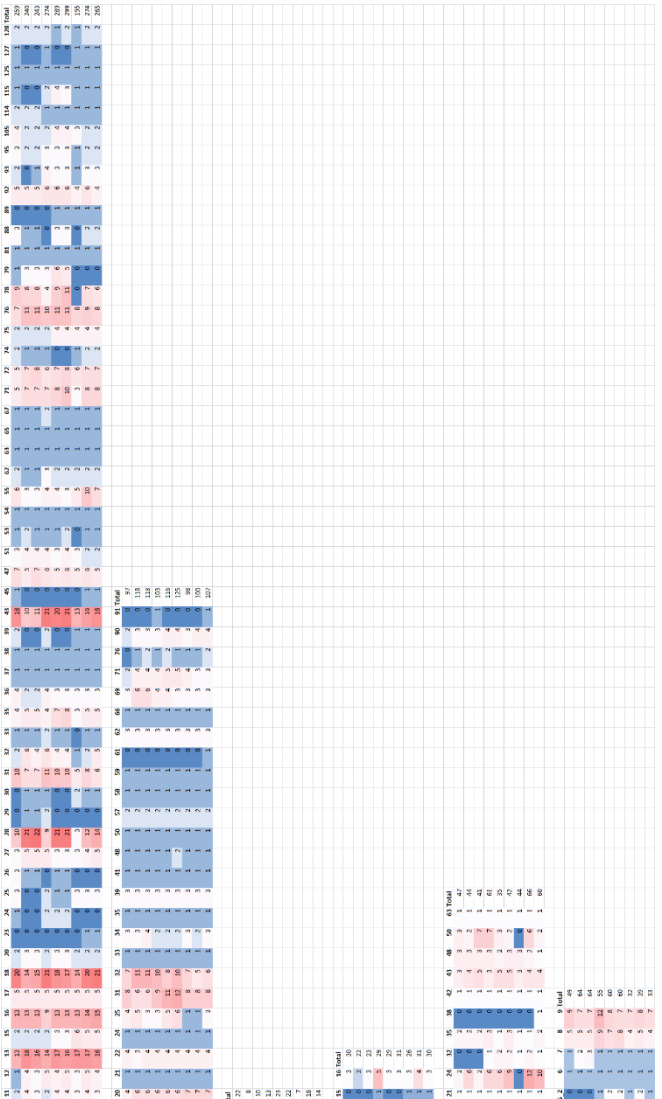
n/a = not available

Supplemental Table 2: Comparison of the polysaccharide degradation potential of eight Aspergilli based on their genome content. The potential per polysaccharide was determined by adding up the number of genes per polysaccharide-related (sub-)family.

Species	Cellulose*	Xyloglucan	Xylan	Galactomannan	Pectin	Starch	Inulin
	GH1, GH12[1], GH5[1], GH6, GH7, GH45, AA9	GH12[2], GH29, GH31[3], GH74, GH95	CE1[4], CE15, GH3[5], GH10, GH11, GH43[6], GH62, GH67, GH115	GH2[7], GH5[8], GH26, GH27, GH36,	CE8, CE12, GH2[9], GH28, GH35, GH43[10], GH51, GH53, GH54, GH78, GH88, GH93, GH105, PL1, PL3, PL4, PL9, PL11	GH13[11], GH15, GH31[12]	GH32
A. nidulans	22 (36)	7	29	19	71	21	2
A. niger ATCC1015	19 (33)	8	14	12	64	15	4
A. terreus	30 (43)	11	33	18	55	19	6
A. oryzae	22 (39)	7	34	14	89	23	4
A. flavus	22 (39)	7	34	14	92	22	4
A. clavatus	22 (28)	4	21	11	30	23	1
A. fischeri	30 (44)	8	29	14	66	24	2
A. fumigatus Af293	26 (37)	6	28	14	65	22	4

*In brackets the numbers including putative GH3 BGLs are given. BGLs are also involved in other processes than cellulose degradation and their high number in the genomes could hide the real difference in gene numbers related to cellulose degradation between the species. [1]Only endoglucanases of this family. [2]Only xyloglucan-active endoglucanases of this family. [3]Only α-xylosidases of this family. [4]Only acetyl xylan esterases of this family. [5]Only β-xylosidases of this family. [6]Only β-xylosidases and α-arabinofuranosidases of this family. [7]Only β-mannosidases of this family. [8]Only endomannanases of this family. [9]Only β-galactosidases of this family. [10]Only endoarabinanases of this family. [11]Only α-galactosidases and α-amylases of this family. [12]Only α-galactosidases of this family.

Supplemental Table 3A: Numbers of putative genes per CAZy family for the 10 genomes addressed in this study. Colour coding is meant to visualize the relative number of genes in the species and families. (Please see pdf file to enlargen text).



Supplemental Table 3B: Numbers of putative genes per Plant Polysaccharide Degradation-related CAZy family for the 10 genomes addressed in this study. Colour coding is meant to visualize the relative numbe rof genes in the species and families.

GH family	1	2	3	5	6	7	10	11	12	13	15	26	27	28	29	31	32	35	36	43	45	51	53	54	62	67	74	78	88	93	95	105	115	Total
Aspergillus nidulans FGSC A4	3	10	21	16	2	3	3	2	1	13	2	3	3	10	0	10	2	4	4	18	1	3	1	1	2	1	2	9	3	2	3	4	1	163
Aspergillus niger CBS 513.88	3	6	17	10	2	2	1	4	4	18	2	1	5	21	1	7	6	5	2	10	0	4	2	1	1	1	1	8	1	0	2	2	0	150
Aspergillus niger ATCC 1015	3	6	19	10	2	2	2	3	3	16	2	1	5	22	1	7	4	5	2	11	0	4	1	1	1	1	1	8	1	1	2	2	0	149
Aspergillus terreus NIH2624	3	10	21	18	2	4	4	2	5	14	2	0	5	9	2	11	6	4	4	21	0	4	1	1	3	2	1	4	0	4	3	2	2	174
Aspergillus oryzae RIB40	3	7	23	14	1	3	4	4	4	17	3	1	3	21	0	10	4	7	3	20	0	3	1	1	2	1	0	9	3	3	3	4	4	186
Aspergillus flavus NRRL3357	3	8	24	15	1	3	4	4	5	16	3	1	3	21	0	10	4	8	3	21	0	4	2	1	2	1	0	11	3	3	3	4	3	194
Aspergillus clavatus NRRL 1	4	3	12	9	2	4	2	3	3	17	6	0	3	3	0	5	1	3	3	13	0	3	0	1	2	1	1	0	0	1	1	3	1	110
Aspergillus fischeri NRRL 181	5	6	19	16	2	5	4	4	5	17	5	0	4	12	0	8	2	5	3	16	1	2	1	1	2	1	2	7	2	3	2	2	1	165
Aspergillus fumigatus Af293	5	6	18	13	1	4	4	3	4	16	5	0	5	14	0	6	5	5	3	18	1	2	1	1	2	1	2	6	2	3	2	2	1	161

PL family	1	3	4	9	11	Total
Aspergillus nidulans FGSC A4	9	5	4	1	1	20
Aspergillus niger CBS 513.88	6	0	2	0	0	8
Aspergillus niger ATCC 1015	6	0	2	0	0	8
Aspergillus terreus NIH2624	6	3	3	1	0	13
Aspergillus oryzae RIB40	12	3	4	1	0	20
Aspergillus flavus NRRL3357	12	3	3	1	0	19
Aspergillus clavatus NRRL 1	2	1	2	0	0	5
Aspergillus fischeri NRRL 181	7	4	3	1	0	15
Aspergillus fumigatus Af293	6	3	3	1	0	13

CE family	1	8	12	15	Total
Aspergillus nidulans FGSC A4	4	3	2	0	9
Aspergillus niger CBS 513.88	3	3	2	0	8
Aspergillus niger ATCC 1015	3	3	2	0	8
Aspergillus terreus NIH2624	5	1	2	1	9
Aspergillus oryzae RIB40	5	5	4	0	14
Aspergillus flavus NRRL3357	5	5	4	0	14
Aspergillus clavatus NRRL 1	6	7	1	1	15
Aspergillus fischeri NRRL 181	5	7	3	1	16
Aspergillus fumigatus Af293	5	6	3	1	15

Auxiliary acitivities	1_2	1_3	3_1	7	9	Total
Aspergillus nidulans FGSC A4	1	7	1	1	9	19
Aspergillus niger CBS 513.88	3	13	2	2	7	27
Aspergillus niger ATCC 1015	3	12	2	3	7	27
Aspergillus terreus NIH2624	2	7	2	1	12	24
Aspergillus oryzae RIB40	2	7	2	1	8	20
Aspergillus flavus NRRL3357	2	9	3	1	7	22
Aspergillus clavatus NRRL 1	3	2	2	1	7	15
Aspergillus fischeri NRRL 181	2	4	2	1	8	17
Aspergillus fumigatus Af293	2	4	2	1	7	16

Supplemental Table 4A: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families.

Colour codes: Yellow = at least one of the orthologs has biochemical support
Green = only present in one species
Orange = present in all species
Blue = present in only one strain of this species

The number of species in which an ortholog is present is listed after the orthologous cluster and a sum of this is given at the bottom.

The number of unique genes per species is given at the bottom of the species column.

(Continued on next 6 pages).

enzyme code	function	CAZY	*A. nidulans*	*A. niger* CBS513.88	*A. niger* ATCC 1015	*A. terreus*	*A. oryzae*	*A. flavus*	*A. clavatus*	*A. fischeri*	*A. fumigatus* A1163	*A. fumigatus* Af293	orthologous clusters	8 species	7 species	6 species	5 species	4 species	3 species	2 species	1 species
FAE SF7	feruloyl esterase	faeA		An09g00120	51662	ATEG_08907	AO090001000207	AFL2G_07436					Cluster 001					1			
FAE SF1	feruloyl esterase	SF1	AN1772	An12g10390	51478	ATEG_06663	AO090001000582	AFL2G_09228	ACLA_083360	NFIA_054700	AFUB_074990	Afu6g09040	Cluster 002A	1							
FAE SF1	feruloyl esterase	SF1				ATEG_02212	AO090001000066	AFL2G_07310		NFIA_047590	AFUB_097070	Afu6g00450	Cluster 002B				1				
FAE SF3	feruloyl esterase	SF3					AO090102000013						Cluster 003A								1
FAE SF3	feruloyl esterase	SF3					AO090102000014	AFL2G_09470					Cluster 003B							1	
FAE SF3	feruloyl esterase	SF3				ATEG_02415							Cluster 004								1
FAE SF4	feruloyl esterase	SF4		An04g02780	190471		AO090010000573	AFL2G_11725					Cluster 005						1		
FAE SF4	feruloyl esterase	SF4		An09g05350	43194								Cluster 006								1
AXE	acetyl xylan esterase	CE 1	AN6093	An12g05010	211544	ATEG_09843	AO090011000745	AFL2G_05471	ACLA_081220	NFIA_099230	AFUB_081200	Afu8g06570	Cluster 007	1							
AXE	acetyl xylan esterase	CE 1	AN8320									Afu1g17510	Cluster 008							1	
AXE	acetyl xylan esterase	CE 1								NFIA_084920			Cluster 009								1
FAE SF5	feruloyl esterase	CE 1					AO090005000945	AFL2G_00922	ACLA_065130	NFIA_115130	AFUB_088950	Afu7g02380	Cluster 010				1				
FAE SF5	feruloyl esterase	CE 1	AN5267	An12g02550	43785	ATEG_08112	AO090023000158	AFL2G_04047	ACLA_055050	NFIA_089720	AFUB_030150	Afu2G14530	Cluster 011	1							
FAE SF5	feruloyl esterase	CE 1				ATEG_06644	AO090701000884	AFL2G_06446	ACLA_061520				Cluster 012A					1			
FAE SF5	feruloyl esterase	CE 1				ATEG_01914							Cluster 012B								1
FAE SF5	feruloyl esterase	CE 1							ACLA_017480		AFUB_025300	Afu2g09440	Cluster 013							1	
?		CE 1	AN8782	An07g03100	53315	ATEG_06438	AO090005000277	AFL2G_00281	ACLA_012870	NFIA_077450	AFUB_057420	Afu5g09860	Cluster 014	1							
PME	pectin methyl esterase	CE 8					AO090003001514	AFL2G_01578					Cluster 015							1	
PME	pectin methyl esterase	CE 8		An03g06310	44585		AO090012000749	AFL2G_03618	ACLA_044240	NFIA_095020	AFUB_085090	Afu8g01520	Cluster 016			1					
PME	pectin methyl esterase	CE 8	AN4860	An02g12505	174365	ATEG_01704	AO090020000404	AFL2G_10916	ACLA_035610	NFIA_069500	AFUB_041450	Afu3g07650	Cluster 017	1							
PME	pectin methyl esterase	CE 8	AN3390	An04g09690	214857		AO090102000010	AFL2G_09467		NFIA_099600	AFUB_080890	Afu8g06880	Cluster 018		1						
PME	pectin methyl esterase	CE 8	AN7566										Cluster 019								1
PME	pectin methyl esterase	CE 8					AO090113000039	AFL2G_08528					Cluster 020							1	
PME	pectin methyl esterase	CE 8								NFIA_100100	AFUB_080500	Afu8g07250	Cluster 021							1	
PME	pectin methyl esterase	CE 8							ACLA_059970				Cluster 022								1
PME	pectin methyl esterase	CE 8							ACLA_059980				Cluster 023								1
RGAE	rhamnogalacturonan acetyl esterase	CE 12					AO090003001268	AFL2G_01797					Cluster 024							1	
RGAE	rhamnogalacturonan acetyl esterase	CE 12								NFIA_099110		Afu8g06480	Cluster 025							2	
RGAE	rhamnogalacturonan acetyl esterase	CE 12		An04g09360	51400	ATEG_10016	AO090102000092			NFIA_092640	AFUB_032900	Afu2g17250	Cluster 026				1				
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN7240	An05g02270	190025		AO090102000091	AFL2G_09543		NFIA_092660	AFUB_032920	Afu2g17270	Cluster 027			1					
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2834				AO090113000155	AFL2G_08631					Cluster 028						1		
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2528	An09g02160	189254	ATEG_03511	AO090701000556	AFL2G_06177	ALCL_041970	NFIA_062750	AFUB_034720	Afu3g14510	Cluster 029	1							
GE	glucuronoyl esterase	CE15				ATEG_00945			ACLA_087520	NFIA_060260	AFUB_000370	Afu6g14390	Cluster 030					1			
BGL	beta glucosidase	GH 1		An11g02100	140573								Cluster 031								1
BGL	beta glucosidase	GH 1	AN10124	An03g03740	213437	ATEG_00687	AO090003000497	AFL2G_02496	ACLA_020660	NFIA_010690	AFUB_014260	Afu1g14710	Cluster 032	1							
BGL	beta glucosidase	GH 1							ACLA_019180	NFIA_009040	AFUB_015740	Afu1g16400	Cluster 033						1		
BGL	beta glucosidase	GH 1					AO090113000148	AFL2G_08626		NFIA_060550	AFUB_000170	Afu6g14600	Cluster 034					1			
BGL	beta glucosidase	GH 1	AN9183	An04g03170	131747	ATEG_02657	AO090120000075	AFL2G_08111		NFIA_099670	AFUB_080820	Afu8g06970	Cluster 035		1						
BGL	beta glucosidase	GH 1	AN10375			ATEG_04135			ACLA_040420	NFIA_064710	AFUB_036580	Afu3g12600	Cluster 036				1				
BGL	beta glucosidase	GH 1							ACLA_064280				Cluster 037								1
GUS	beta glucuronidase	GH 2	AN3200	An05g02410	189620	ATEG_09745		AFL2G_05262					Cluster 038					1			

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families. (Continued on next 5 pages).

GUS	beta glucuronidase	GH 2	AN2395	An02g00610	52111	ATEG_01031	AO090023000053	AFL2G_03956		NFIA_089700	AFUB_030140	Afu2g14520	Cluster 039			1													
GUS	beta glucuronidase	GH 2	AN5361	An01g01260	46827	ATEG_10255	AO090038000009	AFL2G_08828					Cluster 040							1									
LAC	beta galactosidase	GH 2				ATEG_04784							Cluster 041A																1
LAC	beta galactosidase	GH 2	AN2463			ATEG_00712							Cluster 041B															1	
LAC	beta galactosidase	GH 2	AN3201			ATEG_10243	AO090012000389	AFL2G_03296		NFIA_072410	AFUB_062230	Afu5g14550	Cluster 041C				1												
LAC	beta galactosidase	GH 2	AN6388										Cluster 042																1
LAC	beta galactosidase	GH 2	AN1107							NFIA_072890	AFUB_061780	Afu5g14090	Cluster 043											1					
MND	beta mannosidase	GH 2	AN1742	An11g06540	138876	ATEG_06636	AO090001000556	AFL2G_09201	ACLA_083570	NFIA_054490	AFUB_074800	Afu6g08840	Cluster 044	1															
MND	beta mannosidase	GH 2	AN11680	An01g06630	172587	ATEG_07339	AO090003001410	AFL2G_01666	ACLA_078510	NFIA_045250	AFUB_100920	Afu4g00390	Cluster 045	1															
MND	beta mannosidase	GH 2				ATEG_09890	AO090005000740	AFL2G_00728					Cluster 046A											1					
MND	beta mannosidase	GH 2	AN3368	An12g01850	212893	ATEG_08684	AO090010000208	AFL2G_11464	ACLA_066240	NFIA_114030	AFUB_087900	Afu7g01320	Cluster 046B	1															
BGL	beta glucosidase	GH 3				ATEG_10274							Cluster 047A																1
BGL	beta glucosidase	GH 3		An14g01770	210981	ATEG_02806	AO090003001511	ALF2G_01582		NFIA_095760	AFUB_084510	Afu8g02100	Cluster 047B				1												
BGL	beta glucosidase	GH 3	AN7396	An11g00200	179265	ATEG_10320	AO090012000135	AFL2G_03066		NFIA_007920	AFUB_016780	Afu1g17410	Cluster 047C			1													
BGL	beta glucosidase	GH 3					AO090103000127	AFL2G_12245					Cluster 047D														1		
BGL	beta glucosidase	GH 3					AO090701000841	AFL2G_06408	ACLA_007810				Cluster 047E											1					
BGL	beta glucosidase	GH 3		An07g07630	139037								Cluster 048																1
BGL	beta glucosidase	GH 3		An07g09760	39613								Cluster 049A																1
BGL	beta glucosidase	GH 3	AN7915	An08g08240	38077		AO090001000266	AFL2G_07497					Cluster 049B									1							
BGL	beta glucosidase	GH 3	AN10482	An11g06080	208871	ATEG_06617	AO090001000544	AFL2G_09187	ACLA_083710	NFIA_054350	AFUB_074660	Afu6g08700	Cluster 050A	1															
BGL	beta glucosidase	GH 3	AN4102	An18g03570	56782	ATEG_03047	AO090009000356	AFL2G_10322	ACLA_028810	NFIA_018950	AFUB_006160	Afu1g05770	Cluster 050B	1															
BGL	beta glucosidase	GH 3	AN6652	An15g01890	182309	ATEG_07121	AO090009000554	AFL2G_10164	ACLA_096980	NFIA_050080	AFUB_094720	Afu6g03570	Cluster 050C	1															
BGL	beta glucosidase	GH 3				ATEG_02724	AO090005000337	AFL2G_00334					Cluster 051											1					
BGL	beta glucosidase	GH 3		An03g05330	44520		AO090010000034	AFL2G_11300					Cluster 052A											1					
BGL	beta glucosidase	GH 3	AN3903			ATEG_04069	AO090166000090	AFL2G_09413	ACLA_087610	NFIA_060370	AFUB_000280	Afu6g14490	Cluster 052B			1													
BGL	beta glucosidase	GH 3	AN2227	An17g00520	129891	ATEG_09329	AO090701000244	AFL2G_05886	ACLA_010450	NFIA_080070	AFUB_054750	Afu5g07190	Cluster 052C	1															
BGL	beta glucosidase	GH 3	AN2612								AFUB_086800	Afu7g00240	Cluster 052D														1		
BGL	beta glucosidase	GH 3	AN7865							NFIA_057590	AFUB_077900	Afu6g11910	Cluster 052E											1					
BGL	beta glucosidase	GH 3					AO090012000003	AFL2G_02949					Cluster 053														1		
BGL	beta glucosidase	GH 3	AN1804				AO090026000123	AFL2G_07119					Cluster 054											1					
BGL	beta glucosidase	GH 3					AO090038000223	AFL2G_09023					Cluster 055A														1		
BGL	beta glucosidase	GH 3	AN5976	An15g04800	181816	ATEG_02713	AO090038000425	AFL2G_07763		NFIA_100430			Cluster 055B							1									
BGL	beta glucosidase	GH 3								NFIA_098520			Cluster 055C																1
BGL	beta glucosidase	GH 3		An06g02040	176601		AO090166000048	AFL2G_09452		NFIA_112660			Cluster 056									1							
BGL	beta glucosidase	GH 3				ATEG_07931							Cluster 057A																1
BGL	beta glucosidase	GH 3	AN2828			ATEG_07419				NFIA_027390	AFUB_091720	Afu7g06140	Cluster 057B									1							
BGL	beta glucosidase	GH 3	AN0712										Cluster 058																1
BGL	beta glucosidase	GH 3	AN3949										Cluster 059																1
BGL	beta glucosidase	GH 3			129779	ATEG_00157							Cluster 060														1		
BGL	beta glucosidase	GH 3								NFIA_000750	AFUB_048210	Afu3g00230	Cluster 061																
BGL	beta glucosidase	GH 3								NFIA_057910	AFUB_077990	Afu6g12010	Cluster 062A														1		
BGL/BXL	beta glucosidase/beta xylosidase	GH 3	AN0479										Cluster 063																1
BXL	beta xylosidase	GH 3	AN2359	An01g09960	205670	ATEG_05106	AO090005000986	AFL2G_00957			AFUB_016310	Afu1g16920	Cluster 064A				1												
BXL	beta xylosidase	GH 3				ATEG_07383	AO090011000140	AFL2G_04928	ACLA_018590				Cluster 064B									1							
BXL/ABF	beta xylosidase/alpha arabinofuranosidase	GH 3	AN8401			ATEG_09052	AO090103000120	AFL2G_12252	ACLA_062400	NFIA_003180	AFUB_046310	Afu3g02090	Cluster 064C			1													
BXL/ABF	beta xylosidase/alpha arabinofuranosidase	GH 3	AN2217	An17g00300	50997	ATEG_09314	AO090701000274	AFL2G_05912	ACLA_010340	NFIA_080180	AFUB_054640	Afu5g07080	Cluster 065A	1															
BXL/ABF	beta xylosidase/alpha arabinofuranosidase	GH 3				ATEG_08027							Cluster 065B																1
?		GH 3	AN1416	An16g09090	45461	ATEG_00018	AO090103000019	AFL2G_12338	ACLA_057390	NFIA_096770	AFUB_083500	Afu8g04060	Cluster 066A	1															
?		GH 3									AFUB_097970	Afu1g00540	Cluster 066B																1
?		GH 3	AN3360	An06g02460	37673	ATEG_04963			ACLA_052760	NFIA_102600	AFUB_070710	Afu4g13770	Cluster 067				1												
?		GH 3	AN2870	An02g07590	120104	ATEG_04729	AO090003000741	AFL2G_02272	ACLA_039660	NFIA_065550	AFUB_037410	Afu3g11780	Cluster 068	1																

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families. (Continued on next 4 pages).

?		GH 3						AFL2G_08686					Cluster 069												1
?		GH 3	AN2599										Cluster 070												1
EGL	endoglucanase	GH 5	AN3013	An16g02100	210716	ATEG_01592	AO090005001389	AFL2G_01298	ACLA_036770	NFIA_068300	AFUB_040280	Afu3g08820	Cluster 071	1											
EGL	endoglucanase	GH 5	AN1285	An01g11670	205580	ATEG_05003			ACLA_085250	NFIA_057290	AFUB_077620	Afu6g11610	Cluster 072A				1								
EGL	endoglucanase	GH 5								NFIA_095570			Cluster 072B												1
EGL	endoglucanase	GH 5	AN8068			ATEG_09802	AO090003001341+	AFL2G_01726		NFIA_040280	AFUB_050360	Afu5g01830	Cluster 073				1								
EGL	endoglucanase	GH 5	AN5214	An07g08950	209376		AO090005001553	AFL2G_01447	ACLA_081650	NFIA_053150	AFUB_073450	Afu6g07480	Cluster 074			1									
EGL	endoglucanase	GH 5		An16g06800	214608	ATEG_04390	AO090011000715	AFL2G_05447	ACLA_081310	NFIA_085010	AFUB_025390	Afu2g09520	Cluster 075			1									
EXG	exo 1,3 galactanase	GH 5				ATEG_08371				NFIA_056040			Cluster 076											1	
EXG	exo 1,3 galactanase	GH 5	AN8947	An06g02060	175759	ATEG_07719							Cluster 077A									1			
EXG	exo 1,3 galactanase	GH 5	AN1332	An08g01100	52811		AO090012000917	AFL2G_03782		NFIA_113230			Cluster 077B						1						
EXG	exo 1,3 galactanase	GH 5	AN7533	An11g07660	123981	ATEG_06686	AO090001000604	AFL2G_09249	ACLA_083150	NFIA_054930	AFUB_075220	Afu6g09250	Cluster 078	1											
EXG	exo 1,3 galactanase	GH 5	AN4052	An18g04100	202490	ATEG_03849	AO090003000990	AFL2G_02039	ACLA_031040	NFIA_021060	AFUB_004010	Afu1g03600	Cluster 079	1											
EXG	exo 1,3 galactanase	GH 5				ATEG_06369	AO090005000423	AFL2G_00412	ACLA_007330	NFIA_026860	AFUB_091190	Afu7g05610	Cluster 080				1								
EXG	exo 1,3 galactanase	GH 5		An18g03330	42805	ATEG_03062	AO090009000373	AFL2G_10308					Cluster 081							1					
EXG	beta 1,6 glucanase	GH 5	AN3777			ATEG_09844	AO090011000757	AFL2G_05484		NFIA_084850	AFUB_025210	Afu2g09350	Cluster 082				1								
GLN	exo 1,6 galactanase	GH 5	AN9166	An03g01050	194447	ATEG_10242	AO090012000046	AFL2G_02982		NFIA_072400	AFUB_062240	Afu5g14560	Cluster 083			1									
MAN	endomannanase	GH 5	AN3358	An05g01320	50378	ATEG_08654	AO090010000122	AFL2G_11381	ACLA_066420	NFIA_113780	AFUB_087650	Afu7g01070	Cluster 084A	1											
MAN	endomannanase	GH 5				ATEG_02659							Cluster 084B												1
MAN	endomannanase	GH 5				ATEG_10292	AO090012000006	AFL2G_02951		NFIA_041960	AFUB_049020	Afu5g00550	Cluster 085									1			
MAN	endomannanase	GH 5	AN7639				AO090038000444	AFL2G_07781	ACLA_044470	NFIA_099770	AFUB_080770	Afu8g07030	Cluster 086				1								
MAN	endomannanase	GH 5	AN2709										Cluster 087												1
MAN	endomannanase	GH 5	AN3297										Cluster 088												1
MAN	endomannanase	GH 5	AN6427			ATEG_09991							Cluster 089											1	
MAN	endomannanase	GH 5	AN9276										Cluster 090												1
MAN	endomannanase	GH 5		An11g06040	124912	ATEG_01374				NFIA_101450		Afu1g00640	Cluster 091							1				
CBH	cellobiohydrolase	GH 6	AN5282	An12g02220	54490	ATEG_07493			ACLA_062560	NFIA_002990	AFUB_046510	Afu3g01910	Cluster 092				1								
CBH	cellobiohydrolase	GH 6	AN1273	An08g01760	133986	ATEG_00193	AO090038000439	AFL2G_07776	ACLA_025560	NFIA_015680			Cluster 093			1									
CBH	cellobiohydrolase	GH 7	AN0494	An01g11660	51773	ATEG_05002	AO090001000348	AFL2G_07571	ACLA_085260	NFIA_057290	AFUB_077620	Afu6g11600	Cluster 094A	1											
CBH	cellobiohydrolase	GH 7								NFIA_095570			Cluster 094B												1
CBH	cellobiohydrolase	GH 7	AN5176	An07g09330	53159	ATEG_03727	AO090012000941	AFL2G_03805	ACLA_088870	NFIA_052720	AFUB_073010	Afu6g07070	Cluster 094C	1											
EGL	endoglucanase	GH 7	AN3418			ATEG_08700	AO090010000314	ALF2G_11497	ACLA_056030	NFIA_114250	AFUB_088110	Afu7g01540	Cluster 095			1									
EGL	endoglucanase	GH 7				ATEG_08705			ACLA_098940	NFIA_047960	AFUB_096550	Afu6g01800	Cluster 096							1					
XLN	endoxylanase	GH 10					AO090001000208	AFL2G_07437					Cluster 097											1	
XLN	endoxylanase	GH 10	AN7401			ATEG_03410	AO090103000326	AFL2G_12071	ACLA_086910	NFIA_059570	AFUB_001130	Afu6g13610	Cluster 098			1									
XLN	endoxylanase	GH 10	AN1818	An03g00940	57436	ATEG_00809	AO090701000887	AFL2G_06449	ACLA_048770	NFIA_106540	AFUB_066600	Afu4g09480	Cluster 099A	1											
XLN	endoxylanase	GH 10				ATEG_08906	AO090103000423	AFL2G_11983					Cluster 099B									1			
XLN	endoxylanase	GH 10	AN2356		50977					NFIA_057510	AFUB_077830	Afu6g11820	Cluster 100							1					
XLN	endoxylanase	GH 10				ATEG_07190				NFIA_061880	AFUB_034010	Afu3g15210	Cluster 101									1			
XLN	endoxylanase	GH 11	AN3613										Cluster 102A												1
XLN	endoxylanase	GH 11	AN9365	An01g00780	52071	ATEG_07461	AO090120000026	AFL2G_08066	ACLA_063140	NFIA_000850	AFUB_048130	Afu3g00320	Cluster 102B	1											
XLN	endoxylanase	GH 11		An14g07390						NFIA_001010	AFUB_047990	Afu3g00470	Cluster 103A									1			
XLN	endoxylanase	GH 11		An01g14600	171269	ATEG_04943	AO090001000111	AFL2G_07347	ACLA_085410	NFIA_058160	AFUB_078210	Afu6g12210	Cluster 104A			1									
XLN	endoxylanase	GH 11							ACLA_064270				Cluster 104B												1
XLN	endoxylanase	GH 11		An15g04550	183088		AO090026000103	AFL2G_07138		NFIA_055240			Cluster 105							1					
XLN	endoxylanase	GH 11					AO090103000141	AFL2G_12233					Cluster 106											1	
XG EGL	xyloglucan active endoglucanase	GH 12	AN0452	An01g03340	52011	ATEG_03755	AO090003000905	AFL2G_02120	ACLA_029940	NFIA_020020	AFUB_005070	Afu1g04730	Cluster 107	1											
XG EGL	xyloglucan active endoglucanase	GH 12		An03g05530	191511								Cluster 108												1
XG EGL	endoglucanase	GH 12		An14g02760	211053	ATEG_07420	AO090026000102	AFL2G_07140	ACLA_007820	NFIA_027400	AFUB_091730	Afu7g06150	Cluster 109A			1									
XG EGL	endoglucanase	GH 12				ATEG_05519							Cluster 109B												1

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families. (Continued on next 3 pages).

AGL	alpha galactosidase	GH 27				ATEG_01905							Cluster 142											1
AGL	alpha galactosidase	GH 27				ATEG_03427				NFIA_048850	AFUB_095770	Afu6g02560	Cluster 143							1				
AGL	alpha galactosidase	GH 27		An06g00170	37736	ATEG_04382	AO090005000217	AFL2G_00225	ACLA_016820	NFIA_073100		Afu5g13830	Cluster 144		1									
AGL	alpha galactosidase	GH 27	AN0022	An01g01320	172232					NFIA_023390	AFUB_079600	Afu1g01200	Cluster 145						1					
AGL	alpha galactosidase	GH 27	AN7624	An02g11150	207264	ATEG_09830	AO090003001305	AFL2G_01765	ACLA_003130	NFIA_039990	AFUB_050660	Afu5g02130	Cluster 146	1										
AGL	alpha galactosidase	GH 27		An11g06330	39180								Cluster 147											1
AGL	alpha galactosidase	GH 27	AN7152	An14g01800	41606	ATEG_02160	AO090023000151	AFL2G_04039		NFIA_029860	AFUB_099470	Afu4g03580	Cluster 148		1									
AGL	alpha galactosidase	GH 27							ACLA_097990				Cluster 149											1
PGA	endopolygalacturonase	GH 28	AN8327	An01g14670	46255	ATEG_07748				NFIA_008150	AFUB_016610	Afu1g17220	Cluster 150					1						
PGA	endopolygalacturonase	GH 28		An05g02440	43957								Cluster 151											1
PGA	endopolygalacturonase	GH 28		An15g05370	182156								Cluster 152											1
PGA	endopolygalacturonase	GH 28		An16g06990	214596								Cluster 153											1
PGA	endopolygalacturonase	GH 28	AN6656	An09g03260	50161		AO090005000186	AFL2G_00201		NFIA_099410			Cluster 154					1						
PGA	endopolygalacturonase	GH 28		An02g12450	172944	ATEG_01601	AO090005001400	AFL2G_01310	ACLA_036670	NFIA_068440	AFUB_040410	Afu3g08680	Cluster 155		1									
PGA	endopolygalacturonase	GH 28	AN4372	An01g11520	141677	ATEG_04991	AO090023000401	AFL2G_04252	ACLA_052860	NFIA_102450	AFUB_070830	Afu4g13920	Cluster 156	1										
PGA	endopolygalacturonase	GH 28		An02g04900	52219		AO090023000161	AFL2G_04049		NFIA_095620	AFUB_084640	Afu8g01970	Cluster 157					1						
PGA	endopolygalacturonase	GH 28					AO090138000085						Cluster 158											1
PGA	endopolygalacturonase	GH 28									AFUB_081060	Afu8g06730	Cluster 159											1
PGA	endopolygalacturonase	GH 28						AFL2G_08764		NFIA_023290	AFUB_079470	Afu1g01320	Cluster 160							1				
PGX	exopolygalacturonase	GH 28		An11g04040	178172								Cluster 161											1
PGX	exopolygalacturonase	GH 28	AN8891	An03g06740	191158	ATEG_10357	AO090010000753	AFL2G_11892	ACLA_043100	NFIA_096340	AFUB_083960	Afu8g02630	Cluster 162	1										
PGX	exopolygalacturonase	GH 28	AN8761	An12g07500	42184	ATEG_07152	AO090026000784	AFL2G_06533		NFIA_049320	AFUB_095310	Afu6g02980	Cluster 163		1									
PGX	exopolygalacturonase	GH 28	AN9045										Cluster 164											1
PGX	exopolygalacturonase	GH 28								NFIA_100120		A6u8g07265	Cluster 165										1	
RGX	exorhamnogalacturonase	GH 28		An01g14650	172236		AO090001000133	AFL2G_07371					Cluster 166							1				
RGX	exorhamnogalacturonase	GH 28		An18g04810	42917		AO090009000470	AFL2G_10228		NFIA_018590	AFUB_006520	Afu1g06140	Cluster 167					1						
RGX	exorhamnogalacturonase	GH 28	AN10274	An03g02080	194461	ATEG_06408	AO090102000139	AFL2G_09582		NFIA_027700	AFUB_091970	Afu7g06410	Cluster 168		1									
RGX	exorhamnogalacturonase	GH 28				ATEG_09025	AO090113000199	AFL2G_08671					Cluster 169							1				
RGX	exorhamnogalacturonase	GH 28					AO090138000066	AFL2G_03102					Cluster 170										1	
RGX	exorhamnogalacturonase	GH 28	AN11626										Cluster 171											1
RGX	exorhamnogalacturonase	GH 28					AO090138000067	AFL2G_08746					Cluster 172										1	
RHG	endorhamnogalacturonase	GH 28		An11g06320	178191								Cluster 173											1
RHG	endorhamnogalacturonase	GH 28		An07g01000	180922		AO090003000524	AFL2G_02475			AFUB_100610	Afu4g00100	Cluster 174						1					
RHG	endorhamnogalacturonase	GH 28					AO090005000067	AFL2G_00087					Cluster 175										1	
RHG	endorhamnogalacturonase	GH 28		An06g02070	123651		AO090026000252	AFL2G_06999					Cluster 176							1				
RHG	endorhamnogalacturonase	GH 28		An11g08700	39337		AO090124000009	AFL2G_08037					Cluster 177							1				
XGH	xylogalacturonase	GH 28					AO090026000120	AFL2G_07122					Cluster 178										1	
XGH	xylogalacturonase	GH 28	AN3389	An04g09700	46065		AO090102000011	AFL2G_09468		NFIA_099610	AFUB_080880	Afu8g06890	Cluster 179				1							
AFC	alpha fucosidase	GH 29		An13g02110	44822	ATEG_08111							Cluster 180										1	
AFC	alpha fucosidase	GH 29				ATEG_06637							Cluster 181											1
AGD	alpha glucosidase	GH 31	AN2017	An04g06920	214233	ATEG_00723	AO090003001209	AFL2G_01842	ACLA_049370	NFIA_105900	AFUB_067270	Afu4g10150	Cluster 182A	1										
AGD	alpha glucosidase	GH 31				ATEG_08278				NFIA_060440			Cluster 182B										1	
AGD	alpha glucosidase	GH 31				ATEG_08472							Cluster 182C											1
AGD	alpha glucosidase	GH 31	AN0280	An01g04880	55419	ATEG_02528	AO090005000767	AFL2G_00750	ACLA_031260	NFIA_021450	AFUB_003550	Afu1g03140	Cluster 183A	1										
AGD	alpha glucosidase	GH 31	AN7320										Cluster 183B											1
AGD	alpha glucosidase	GH 31	AN0941	An01g10930	119858	ATEG_05177	AO090005001084	AFL2G_01038	ACLA_019300	NFIA_009180	AFUB_015590	Afu1g16250	Cluster 184A											
AGD	alpha glucosidase	GH 31	AN8963				AO090038000471	AFL2G_07812					Cluster 184B							1				
AGD	alpha glucosidase	GH 31	AN7345										Cluster 184C											1
AGD	alpha glucosidase	GH 31	AN10935				AO090026000111	AFL2G_07131					Cluster 185							1				
AGD	alpha glucosidase	GH 31	AN3504	An18g05620	49940	ATEG_02966	AO090023000288	AFL2G_04152		NFIA_018130	AFUB_006940	Afu1g06560	Cluster 186A		1									
AGD	alpha glucosidase	GH 31					AO090023000290						Cluster 186B											1

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families. (Continued on next 2 pages).

AGD	alpha glucosidase	GH 31	AN11054	An09g05880	50055	ATEG_08065	AO090102000559	AFL2G_09934	ACLA_001670	NFIA_038610	AFUB_052020	Afu5g03500	Cluster 187	1							
AXL	alpha xylosidase	GH 31	AN7505			ATEG_06730	AO090001000649	AFL2G_09290	ACLA_089300	NFIA_082140	AFUB_022430	Afu2g05400	Cluster 188A		1						
AXL	alpha xylosidase	GH 31				ATEG_00207							Cluster 188B								1
AXL	alpha xylosidase	GH 31		An07g00350	40261		AO090701000558	AFL2G_06180					Cluster 189						1		
AXL	alpha xylosidase	GH 31		An09g03300	43342	ATEG_08390	AO090701000639	AFL2G_06239		NFIA_032680			Cluster 190				1				
AXL	alpha xylosidase	GH 31				ATEG_05169							Cluster 191								1
INU	endo inulinase	GH 32		An11g03200	52928						AFUB_048930	Afu5g00530	Cluster 192							1	
INU	endo inulinase	GH 32								NFIA_041980	AFUB_049000	Afu5g00480	Cluster 193							1	
INU	endo inulinase	GH 32				ATEG_07356							Cluster 194								1
INX	exo inulinase	GH 32		An11g03210									Cluster 195								1
INX	exo inulinase	GH 32	AN11778	An12g08280	56664	ATEG_08155	AO090701000400	AFL2G_06028	ACLA_094550	NFIA_033540	AFUB_018320	Afu2g01240	Cluster 196A	1							
INU	endo inulinase	GH 32									AFUB_048930	Afu5g00530	Cluster 196B								1
SUC	invertase/beta fructofuranosidase	GH 32	AN3837	An15g00320		ATEG_05860	AO090701000038	AFL2G_05693		NFIA_051560	AFUB_093290	Afu6g05000	Cluster 197A		1						
SUC	invertase/beta fructofuranosidase	GH 32		An06g02420	176039								Cluster 197B								1
SUC	invertase/beta fructofuranosidase	GH 32		An08g11070	198063	ATEG_07479	AO090020000640	AFL2G_10707					Cluster 198A					1			
SUC	invertase/beta fructofuranosidase	GH 32				ATEG_04996							Cluster 198B								1
SUC	invertase/beta fructofuranosidase	GH 32					AO090103000043	AFL2G_12317					Cluster 199							1	
SUC	invertase/beta fructofuranosidase	GH 32				ATEG_10253							Cluster 200								1
LAC	beta galactosidase	GH 35	AN0194										Cluster 201								1
LAC	beta galactosidase	GH 35		An07g04420	180727		AO090001000259	AFL2G_07489					Cluster 202A						1		
LAC	beta galactosidase	GH 35		An06g00290	177434	ATEG_07446	AO090003000042	AFL2G_02901	ACLA_088440	NFIA_052310	AFUB_072590	Afu6g06660	Cluster 202B		1						
LAC	beta galactosidase	GH 35	AN0980	An01g10350	46429	ATEG_05131	AO090120000158	AFL2G_08182		NFIA_008690	AFUB_016080	Afu1g16700	Cluster 202C		1						
LAC	beta galactosidase	GH 35	AN0756	An01g12150	51764	ATEG_00616	AO090012000445	AFL2G_03352	ACLA_021260	NFIA_011250	AFUB_013710	Afu1g14170	Cluster 203A	1							
LAC	beta galactosidase	GH 35		An14g05820	41910		AO090012000746	AFL2G_03616		NFIA_000910	AFUB_048070	Afu3g00380	Cluster 203B				1				
LAC	beta galactosidase	GH 35					AO090701000770						Cluster 203C								1
LAC	beta galactosidase	GH 35						AFL2G_06343					Cluster 203D								1
LAC	beta galactosidase	GH 35	AN2804			ATEG_08255	AO090023000076	AFL2G_03977		NFIA_041490	AFUB_049140	Afu5g00670	Cluster 204			1					
LAC	beta galactosidase	GH 35							ACLA_018830				Cluster 205								1
LAC	beta galactosidase	GH 35						AFL2G_12387					Cluster 206								1
AGL	alpha galactosidase	GH 36	AN8138	An09g00260/A	212736	ATEG_07929	AO090010000684	AFL2G_11824	ACLA_044620	NFIA_094580	AFUB_085470	Afu8g01130	Cluster 207A	1							
AGL	alpha galactosidase	GH 36	AN9035			ATEG_02312							Cluster 207B							1	
AGL	alpha galactosidase	GH 36	AN8392			ATEG_07935	AO090011000063	AFL2G_04865	ACLA_044560	NFIA_094540	AFUB_085500	Afu8g01100	Cluster 208		1						
AGL	alpha galactosidase	GH 36	AN3874	An04g02700	190816	ATEG_05379	AO090023000743	AFL2G_04555	ACLA_047600	NFIA_107920	AFUB_065350	Afu4g08250	Cluster 209	1							
ABF	alpha arabinofuranosidase	GH 43	AN8472				AO090012000356	AFL2G_03265		NFIA_032900	AFUB_017710	Afu2g00650	Cluster 210				1				
ABF	alpha arabinofuranosidase	GH 43	AN7781			ATEG_08386	AO090005000065	AFL2G_00086					Cluster 211					1			
ABF	alpha arabinofuranosidase	GH 43	AN2533				AO090701000838	AFL2G_06400		NFIA_002750	AFUB_046730	Afu3g01660	Cluster 212				1				
ABN	endoarabinanase	GH 43		An07g04930	128077	ATEG_05780				NFIA_078690	AFUB_056130	Afu5g08590	Cluster 213					1			
ABN	endoarabinanase	GH 43		An15g03550	182100	ATEG_07817	AO090026000804	AFL2G_06513	ACLA_072730	NFIA_089980	AFUB_030390	Afu2g14750	Cluster 214		1						
ABN	endoarabinanase	GH 43	AN3044	An16g02730	184195	ATEG_01562	AO090005001320	AFL2G_01244	ACLA_037080	NFIA_068020	AFUB_039990	Afu3g09140	Cluster 215A	1							
ABN	endoarabinanase	GH 43					AO090005000064	AFL2G_00085					Cluster 215B							1	
ABN	endoarabinanase	GH 43				ATEG_03688			ACLA_018180	NFIA_008030	AFUB_016700	Afu1g17320	Cluster 216A					1			
ABN*	endoarabinanase	GH 43	AN8007	An02g10550	197735		AO090138000055	AFL2G_08737	ACLA_098980	NFIA_047920	AFUB_096700	Afu6g00770	Cluster 216B		1						
ABN	endoarabinanase	GH 43	AN6352	An02g01400	134398	ATEG_01407	AO090023000165	AFL2G_04053	ACLA_073460	NFIA_089320	AFUB_029770	Afu2g14150	Cluster 217A	1							
ABN	endoarabinanase	GH 43				ATEG_07787							Cluster 217B								1
ABN	endoarabinanase	GH 43	AN2534	An09g01190	203143	ATEG_03520	AO090701000481	AFL2G_06106	ACLA_042100	NFIA_062660	AFUB_034610	Afu3g14620	Cluster 218A	1							
BXL	beta xylosidase	GH 43	AN1870	An11g03120	179682	ATEG_06059	AO090003000239	AFL2G_02739	ACLA_090090	NFIA_081240	AFUB_021510	Afu2g04480	Cluster 218B	1							
BXL	beta xylosidase	GH 43	AN6751	An02g00140	174379	ATEG_06306	AO090005000476	AFL2G_00456		NFIA_027350	AFUB_091690	Afu7g06110	Cluster 219		1						
BXL	beta xylosidase	GH 43				ATEG_05083	AO090103000268	AFL2G_12122	ACLA_087680	NFIA_060490	AFUB_000220	Afu6g14550	Cluster 220			1					
BXL	beta xylosidase	GH 43	AN10199			ATEG_00093	AO090005000698	AFL2G_00689/(	ACLA_058100/	NFIA_097490/(	AFUB_082790/	Afu8g04710/047	Cluster 221A		1						
BXL	beta xylosidase	GH 43	AN1477										Cluster 221B								1
BXL	beta xylosidase	GH 43	AN10919			ATEG_10072	AO090010000029	AFL2G_11296		NFIA_093350			Cluster 222				1				

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families. (Continued on next page).

BXL	beta xylosidase	GH 43				ATEG_07784	AO090010000494	AFL2G_11656		NFIA_023360	AFUB_079570	Afu1g01230	Cluster 223				1				
BXL	beta xylosidase	GH 43	AN7313	An08g10780	122978	ATEG_01188	AO090010000562	AFL2G_11714	ACLA_043260	NFIA_096240	AFUB_084070	Afu8g02510	Cluster 224	1							
BXL	beta xylosidase	GH 43	AN2664			ATEG_06107	AO090012000350	AFL2G_03257		NFIA_072780	AFUB_061880	Afu5g14190	Cluster 225			1					
BXL	beta xylosidase	GH 43	AN8477			ATEG_01292		AFL2G_06974	ACLA_018160				Cluster 226A					1			
BXL	beta xylosidase	GH 43		An08g01900	47677	ATEG_01292	AO090102000331	AFL2G_09750	ACLA_072000	NFIA_088370	AFUB_028830	Afu2g13190	Cluster 226B		1						
BXL	beta xylosidase	GH 43	AN7864				AO090113000059						Cluster 227							1	
BXL	beta xylosidase	GH 43				ATEG_07517							Cluster 228								1
BXL	beta xylosidase	GH 43	AN2633										Cluster 229								1
BXL	beta xylosidase	GH 43	AN1043										Cluster 230								1
EGL	endoglucanase	GH 45	AN6786							NFIA_028020	AFUB_092320	Afu7g06740	Cluster 231						1		
BXL	beta xylosidase	GH 43				ATEG_10193							Cluster 233A								1
BXL	beta xylosidase	GH 43	AN7275			ATEG_06643	AO090701000886	AFL2G_06448	ACLA_078600	NFIA_033220	AFUB_018010	Afu2g00930	Cluster 233B		1						
ABF	alpha arabinofuranosidase	GH 51	AN1277	An08g01710	38549	ATEG_00198		AFL2G_07796	ACLA_025610	NFIA_015730	AFUB_009340	Afu1g09900	Cluster 234		1						
ABF	alpha arabinofuranosidase	GH 51			50979	ATEG_03540							Cluster 235A							1	
ABF	alpha arabinofuranosidase	GH 51	AN2541	An01g00330	206387	ATEG_02882	AO090124000023	AFL2G_08019	ACLA_099110				Cluster 235B			1					
ABF	alpha arabinofuranosidase	GH 51	AN9439	An09g00880	131891								Cluster 235C							1	
ABF	alpha arabinofuranosidase	GH 51					AO090020000712	AFL2G_10643					Cluster 236							1	
ABF	alpha arabinofuranosidase	GH 51				ATEG_07868	AO090012000298	AFL2G_03217	ACLA_074120	NFIA_090410	AFUB_030820	Afu2g15160	Cluster 237			1					
GAL	endogalactanase	GH 53		An16g06590									Cluster 238								1
GAL	endogalactanase	GH 53	AN5727	An18g05940	187227	ATEG_02927	AO090001000492	AFL2G_09135		NFIA_017780	AFUB_007290	Afu1g06910	Cluster 239		1						
GAL	endogalactanase	GH 53						AFL2G_07488					Cluster 240								1
ABF	alpha arabinofuranosidase	GH 54	AN1571	An15g02300	200605	ATEG_07939	AO090023000001	AFL2G_03901	ACLA_066470	NFIA_060630		Afu6g14620	Cluster 241	1							
AXH	arabinoxylan arabinofuranohydrolase	GH 62				ATEG_00186	AO090103000088	AFL2G_12281	ACLA_071560	NFIA_087900	AFUB_028400	Afu2g12770	Cluster 256			1					
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN2632	An03g00960	55136	ATEG_10071	AO090701000885	AFL2G_06447	ACLA_061510	NFIA_033210	AFUB_018000	Afu2g00920	Cluster 257		1						
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN7908										Cluster 258								1
AXH	arabinoxylan arabinofuranohydrolase	GH 62				ATEG_10379							Cluster 259								1
AGU	alpha glucuronidase	GH 67	AN9286	An14g05800	56619	ATEG_06085	AO090026000127	AFL2G_07114	ACLA_017270	NFIA_072510	AFUB_062100	Afu5g14380	Cluster 260A	1							
AGU	alpha glucuronidase	GH 67				ATEG_09975							Cluster 260B								1
XG CBH	xyloglucan active cellobiohydrolase	GH 74	AN1542							NFIA_096000	AFUB_084270	Afu8g02330	Cluster 261						1		
XG EGL	xyloglucan active endoglucanase	GH 74	AN5061	An01g01870	206333	ATEG_04708			ACLA_044310	NFIA_094990	AFUB_085120	Afu8g01490	Cluster 262			1					
RHA	alpha rhamnosidase	GH 78	AN8465	An07g00240	40264			AFL2G_05364					Cluster 263						1		
RHA	alpha rhamnosidase	GH 78		An15g04530									Cluster 264								1
RHA	alpha rhamnosidase	GH 78	AN2631	An08g09140	176718	ATEG_05089	AO090001000105	AFL2G_07340		NFIA_004560	AFUB_045370	Afu3g02880	Cluster 265		1						
RHA	alpha rhamnosidase	GH 78	AN11954	An04g09070	51410		AO090003001016	AFL2G_02014					Cluster 266A					1			
RHA	alpha rhamnosidase	GH 78	AN12368				AO090113000149/A	AFL2G_08627		NFIA_057930	AFUB_078020	Afu6g14610	Cluster 266B				1				
RHA	alpha rhamnosidase	GH 78	AN7151			ATEG_02922	AO090010000561	AFL2G_11713		NFIA_057930	AFUB_078020	Afu6g12030	Cluster 266C			1					
RHA	alpha rhamnosidase	GH 78						AFL2G_10644		NFIA_060560	AFUB_000160		Cluster 266D						1		
RHA	alpha rhamnosidase	GH 78	AN10277	An01g06620	170172		AO090003001291	AFL2G_01780		NFIA_022970	AFUB_002060	Afu1g01660	Cluster 267			1					
RHA	alpha rhamnosidase	GH 78		An18g04800	42916	ATEG_03018	AO090009000471	AFL2G_10227		NFIA_018620	AFUB_006490	Afu1g06130	Cluster 268			1					
RHA	alpha rhamnosidase	GH 78	AN3780										Cluster 269A								1
RHA	alpha rhamnosidase	GH 78	AN10867	An10g00290	44977	ATEG_04706	AO090012000058	AFL2G_02993		NFIA_026130	AFUB_090610	Afu7g05040	Cluster 269B		1						
RHA	alpha rhamnosidase	GH 78					AO090103000432	AFL2G_11972					Cluster 270							1	
RHA	alpha rhamnosidase	GH 78	AN6929	An12g05700	131668		AO090005001416	AFL2G_03939					Cluster 271					1			
UGH	unsaturated galacturonyl hydrolase	GH 88					AO090005000324	AFL2G_00324		NFIA_026210	AFUB_090670	Afu7g05090	Cluster 272					1			
UGH	unsaturated galacturonyl hydrolase	GH 88	AN3991	An01g01340	36414		AO090138000087	AFL2G08766		NFIA_023340	AFUB_079550	Afu1g01250	Cluster 273			1					
UGH	unsaturated galacturonyl hydrolase	GH 88	AN11078					AFL2G_06465					Cluster 274							1	
UGH	unsaturated galacturonyl hydrolase	GH 88					AO090701000907						Cluster 275								1
UGH	unsaturated galacturonyl hydrolase	GH 88	AN4629										Cluster 276								1
ABX	exoarabinanase	GH 93	AN2060			ATEG_06045	AO090003001017	AFL2G_02013		NFIA_081320	AFUB_021620	Afu2G04570	Cluster 277			1					
ABX	exoarabinanase	GH 93					AO090011000141	AFL2G_04929					Cluster 278							1	

Supplemental Table 4A Cont: Orthology clusters of feruloyl esterase (SF), glycoside hydrolase (GH), carbohydrate esterase and polysaccharide lyase (PL) families.

ABX	exoarabinanase	GH 93	AN5231		49311	ATEG_07909	AO090012000101	AFL2G_03035	ACLA_008170	NFIA_027720	AFUB_091990	Afu7g06430	Cluster 279	1							
ABX	exoarabinanase	GH 93				ATEG_08029							Cluster 280								1
ABX	exoarabinanase	GH 93				ATEG_00891				NFIA_058060	AFUB_078110	Afu6g12120	Cluster 281						1		
AFC	alpha fucosidase	GH 95	AN8149	An16g02760	184037	ATEG_01560	AO090005000382	AFL2G_00373					Cluster 282					1			
AFC	alpha fucosidase	GH 95					AO090005000512	AFL2G_00515					Cluster 283A							1	
AFC	alpha fucosidase	GH 95	AN6673	An16g00540	53702	ATEG_09768	AO090009000086	AFL2G_10565	ACLA_004070	NFIA_040960	AFUB_049660	Afu5g01190	Cluster 283B	1							
AFC	alpha fucosidase	GH 95	AN10376			ATEG_04136				NFIA_064720	AFUB_036590	Afu3g12590	Cluster 284						1		
URH	unsaturated rhamnogalacturonyl hydrolase	GH 105	AN3196	An14g02920	41703	ATEG_07907	AO090001000174	AFL2G_07407	ACLA_056060	NFIA_030670	AFUB_100260	Afu4g02860	Cluster 285	1							
URH	unsaturated rhamnogalacturonyl hydrolase	GH 105					AO090003000153	AFL2G_02808					Cluster 286							1	
URH	unsaturated rhamnogalacturonyl hydrolase	GH 105	AN9383	An14g05340	41877		AO090113000146	AFL2G_08624	ACLA_077810				Cluster 287						1		
URH	unsaturated rhamnogalacturonyl hydrolase	GH 105	AN10505			ATEG_02892	AO090001000063	AFL2G_07308	ACLA_072870	NFIA_089840	AFUB_030270	Afu2g14630	Cluster 288		1						
URH	unsaturated rhamnogalacturonyl hydrolase	GH 105	AN7828										Cluster 289								1
AGU	alpha glucuronidase	GH115				ATEG_04355	AO090005001415	AFL2G_01323					Cluster 290A						1		
AGU	alpha glucuronidase	GH115	AN9329			ATEG_09974	AO090010000038	AFL2G_11304	ACLA_006360	NFIA_025630	AFUB_090220	Afu7g04680	Cluster 290B		1						
AGU	alpha glucuronidase	GH115					AO090001000267	AFL2G_07498					Cluster 290C							1	
AGU	alpha glucuronidase	GH115					AO090113000058						Cluster 290D								1
PLY	pectate lyase	PL 1	AN0741			ATEG_08834	AO090011000673	AFL2G_05417					Cluster 291						1		
PLY	pectate lyase	PL 1	AN5333			ATEG_05467	AO090102000072	AFL2G_09523		NFIA_060270	AFUB_000360	Afu6g14400	Cluster 292			1					
PLY	pectate lyase	PL 1	AN7646	An10g00870	45021	ATEG_08123	AO090701000321	AFL2G_05954		NFIA_033040	AFUB_017840	Afu2g00760	Cluster 293		1						
PLY	pectate lyase	PL 1	AN9367/AN9368				AO090011000030	AFL2G_04835					Cluster 294						1		
PEL	pectin lyase	PL 1				ATEG_00950	AO090003001295	AFL2G_01776					Cluster 295A						1		
PEL	pectin lyase	PL 1					AO090103000463	AFL2G_11948					Cluster 295B							1	
PEL	pectin lyase	PL 1	AN4882	An15g07160	40837	ATEG_7577	AO090012000121	AFL2G_03052		NFIA_077100	AFUB_057770	Afu5g10170	Cluster 295C		1						
PEL	pectin lyase	PL 1		An19g00270	55212		AO090010000030	AFL2G_11297		NFIA_026110	AFUB_090600	Afu7g05030	Cluster 296				1				
PEL	pectin lyase	PL 1	AN2331	An03g00190	45821		AO090010000504	AFL2G_11666	ACLA_013470	NFIA_076850	AFUB_057970	Afu5g10380	Cluster 297		1						
PEL	pectin lyase	PL 1	AN10147	An11g04030	208760		AO090012000451	AFL2G_03359					Cluster 298					1			
PEL	pectin lyase						AO090138000204	AFL2G_08823					Cluster 299							1	
PEL	pectin lyase	PL 1	AN2569	An14g04370	41815	ATEG_01216	AO090010000087	AFL2G_11352	ACLA_094210	NFIA_033080	AFUB_017880	Afu2g00800	Cluster 300	1							
PEL	pectin lyase	PL 1								NFIA_023100			Cluster 301								1
PEL	pectin lyase	PL 1	AN9439										Cluster 302								1
PLY	pectate lyase	PL 3	AN6748			ATEG_06314	AO090005000472	AFL2G_00461		NFIA_027690	AFUB_091960	Afu7g06400	Cluster 303A			1					
PLY	pectate lyase	PL 3	AN8453			ATEG_06285							Cluster 303B							1	
PLY	pectate lyase	PL 3	AN3337				AO090010000706	AFL2G_11846	ACLA_059210	NFIA_098670	AFUB_081630	Afu8g05910	Cluster 304A			1					
PLY	pectate lyase	PL 3	AN2542										Cluster 304B								1
PLY	pectate lyase	PL 3	AN6106			ATEG_08626	AO090038000502	AFL2G_07839		NFIA_023470	AFUB_079680	Afu1g01120	Cluster 305			1					
RGL	rhamnogalacturonan lyase	PL 4	AN7135	An14g01130	210947	ATEG_02193	AO090011000349	AFL2G_05136	ACLA_054660	NFIA_029620	AFUB_099240	Afu4g03780	Cluster 306	1							
RGL	rhamnogalacturonan lyase	PL 4	AN6395	An11g00390	47780	ATEG_10327	AO090012000147/4	AFL2G_03075	ACLA_018320	NFIA_008140	AFUB_016620	Afu1g17230	Cluster 307A	1							
RGL	rhamnogalacturonan lyase	PL 4	AN12097				AO090113000057						Cluster 307B							1	
RGL	rhamnogalacturonan lyase	PL 4				ATEG_08610	AO090138000119	AFL2G_08794		NFIA_094270	AFUB_085750	Afu8g00820	Cluster 308				1				
RGL	rhamnogalacturonan lyase	PL 4	AN3950										Cluster 309								1
PLY	pectate lyase	PL 9	AN2537			ATEG_04635	AO090038000131	AFL2G_08953		NFIA_062360	AFUB_034350	Afu3g14890	Cluster 310A			1					
PLY	pectate lyase	PL 9					AO090038000132						Cluster 310B								1
PLY	pectate lyase	PL 9				ATEG_03526							Cluster 311								1
RGL	rhamnogalacturonan lyase	PL 11	AN2543										Cluster 312								1
														57	44	32	26	35	37	51	106
	unique genes per species		29	16	16	25		7	7	6	7	4	4	14.7	11.3	8.2	6.7	9.0	9.5	13.1	27.2
	strain specific genes			9	5							2	7								

Supplemental Table 4B: Orthology clusters of auxiliary activities (AA).

Colour codes: Yellow = at least one of the orthologs has biochemical support
Green = only present in one species
Orange = present in all species
Blue = present in only one strain of this species

The number of species in which an ortholog is present is listed after the orthologous cluster and a sum of this is given at the bottom.

The number of unique genes per species is given at the bottom of the species column.

(Continued on next 2 pages).

enzyme code	function	Module	*A. nidulans*	*A. niger* CBS513.88	*A. niger* ATCC 1015	*A. terreus*	*A. oryzae*	*A. flavus*	*A. clavatus*	*A. fischeri*	*A. fumigatus* A1163	*A. fumigatus* Af293	Sequence orthologous clusters	8 species	7 species	6 species	5 species	4 species	3 species	2 species	1 species
	candidate ferroxidase	AA1_2		An01g08960	171927	ATEG_08032	AO090102000589	AFL2G_09962	ACLA_001340	NFIA_038330	AFUB_052300	Afu5g03790	Cluster 001A		1						
	candidate ferroxidase	AA1_2		An15g05520	40704		AO090124000036	AFL2G_08008					Cluster 001B						1		
MCO	multicopper oxidase	AA1_2		An14g05370	185409				ACLA_076450	NFIA_092960	AFUB_033230	Afu2g17540	Cluster 002					1			
	distantly related to ferroxidase	AA1_2				ATEG_07961							Cluster 003								1
	distantly related to ferroxidase	AA1_2							ACLA_062650				Cluster 004A								1
MCO	laccase-like multicopper oxidase	AA1_3	AN5397	An11g03580	178468	ATEG_03639	AO090090000067	AFL2G_10583		NFIA_072100			Cluster 005A			1					
MCO	laccase-like multicopper oxidase	AA1_3	AN9170	An05g02340	189367	ATEG_08458	AO090010000607	AFL2G_11750		NFIA_101880	AFUB_071740	Afu4g14490	Cluster 005B		1						
MCO	laccase-like multicopper oxidase	AA1_3		An05g08450	47386								Cluster 005C								1
MCO	candidate laccase-like multicopper oxidase	AA1_3	AN6830										Cluster 006								1
MCO	laccase-like multicopper oxidase	AA1_3		An18g02690	131352								Cluster 007A								1
MCO	candidate laccase-like multicopper oxidase	AA1_3		An01g00860	131667								Cluster 007B								1
MCO	laccase-like multicopper oxidase	AA1_3		An16g02020	123717								Cluster 008								1
MCO	laccase-like multicopper oxidase	AA1_3		An12g05810	123557								Cluster 009								1
MCO	laccase-like multicopper oxidase	AA1_3		An01g11120	35672	ATEG_03000	AO090005001096	AFL2G_01047		NFIA_009370			Cluster 010				1				
MCO	related to laccase-like multicopper oxidase	AA1_3		An05g02540									Cluster 011A								1
MCO	laccase-like multicopper oxidase	AA1_3	AN0901	An01g14010	171385	ATEG_05229	AO090166000081	AFL2G_09420	ACLA_019700	NFIA_009660	AFUB_015200	Afu1g15670	Cluster 011B	1							
MCO	laccase-like multicopper oxidase	AA1_3	AN0878	An01g13660	171880		AO090005001177	AFL2G_01113					Cluster 011C					1			
MCO	laccase-like multicopper oxidase	AA1_3		An03g03750	44408								Cluster 011D								1
MCO	distantly related to laccase-like multicopper oxidase	AA1_3		An04g10400									Cluster 011E								1
MCO	distantly related to laccase-like multicopper oxidase	AA1_3				ATEG_05254						Afu4g14280	Cluster 011F							1	
MCO	laccase-like multicopper oxidase	AA1_3					AO090102000546	AFL2G_09924	ACLA_076440	NFIA_092950	AFUB_033220	Afu2g17530	Cluster 012				1				
MCO	distantly related to laccase-like multicopper oxidase	AA1_3	AN7389			ATEG_07309		AFL2G_08693					Cluster 012A						1		
MCO	distantly related to ferroxidase	AA1_2	AN6581										Cluster 012B								1
MCO	distantly related to laccase-like multicopper oxidase	AA1_3				ATEG_00141							Cluster 013								1
MCO	laccase-like multicopper oxidase	AA1_3	AN6635				AO090011000755	AFL2G_05482					Cluster 014						1		
MCO	distantly related to laccase-like multicopper oxidase	AA1_3						AFL2G_05032					Cluster 015								1
GOX	candidate glucose oxidase	AA3_2	AN7408	An07g00450	40254	ATEG_07516	AO090005000449	AFL2G_00439	ACLA_093090	NFIA_034090	AFUB_018840	Afu2g01770	Cluster 016A	1							
GOX	candidate glucose oxidase	AA3_2	AN2175	An11g05580	179579	ATEG_08295	AO090103000214	AFL2G_12171			AFUB_030480	Afu2g14840	Cluster 016B			1					
GOX	glucose oxidase	AA3_2		An01g14740	55227	ATEG_06246	AO090012000544	AFL2G_03439					Cluster 017A					1			
GOX	candidate glucose oxidase	AA3_2				ATEG_08474							Cluster 017B								1
GOX	candidate glucose oxidase	AA3_2				ATEG_07840		AFL2G_07249					Cluster 017C							1	
GOX	candidate glucose oxidase	AA3_2		An12g03440	50376								Cluster 018								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN8329	An02g00400	36617	ATEG_04852	AO090003001409	AFL2G_01667	ACLA_078430	NFIA_045070	AFUB_101160	Afu4g00610	Cluster 019A	1							
AAO	related to aryl-alcohol oxidase	AA3_2		An08g08250	207831								Cluster 019B								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN8547	An07g06310	209587	ATEG_09074	AO090103000085	AFL2G_12283	ACLA_062810	NFIA_002550	AFUB_046830	Afu3g01580	Cluster 020	1							
AAO	related to aryl-alcohol oxidase	AA3_2	AN3206	An14g02940	53801	ATEG_07905	AO090113000007	AFL2G_08503	ACLA_056040	NFIA_036190			Cluster 021A		1						
AAO	related to aryl-alcohol oxidase	AA3_2						AFL2G_12263					Cluster 021B								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN10931	An15g03100	40542	ATEG_07855	AO090012000422	AFL2G_03327	ACLA_072460	NFIA_090260	AFUB_030670	Afu2g15020	Cluster 022A	1							
AAO	related to aryl-alcohol oxidase	AA3_2		An18g01120	131188		AO090005000970	AFL2G_00942					Cluster 022B						1		
AAO	related to aryl-alcohol oxidase	AA3_2	AN10931				AO090003001420	AFL2G_01658					Cluster 022C						1		
AAO	related to aryl-alcohol oxidase	AA3_2	AN6445	An02g11560	37397	ATEG_01660	AO090005001472	AFL2G_01375	ACLA_036030	NFIA_069070	AFUB_041050	Afu3g08070	Cluster 023A	1							

Supplemental Table 4B Cont: Orthology clusters of auxiliary activities (AA). (Continued on next page)

AAO	related to aryl-alcohol oxidase	AA3_2		An05g02380	190068	ATEG_01896			ACLA_093600	NFIA_093820			Cluster 023B					1			
AAO	related to aryl-alcohol oxidase	AA3_2							ACLA_077840	NFIA_014210	AFUB_010800	Afu1g11370	Cluster 024						1		
AAO	related to aryl-alcohol oxidase	AA3_2		An03g05670	44543	ATEG_05456	AO090020000390	AFL2G_10926	ACLA_077970				Cluster 025A				1				
AAO	related to aryl-alcohol oxidase	AA3_2		An15g03820			AO090012000349	AFL2G_03256					Cluster 025B						1		
AAO	related to aryl-alcohol oxidase	AA3_2		An04g02980	190238		AO090701000527	AFL2G_06148	ACLA_060160				Cluster 026A					1			
AAO	distantly related to aryl-alcohol oxidase	AA3_2					AO090023000847						Cluster 026B								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN4006			ATEG_09491	AO090005000530	AFL2G_00531					Cluster 027					1			
AAO	related to aryl-alcohol oxidase	AA3_2				ATEG_08635	AO090701000511	AFL2G_06134					Cluster 028						1		
AAO	related to aryl-alcohol oxidase	AA3_2	AN9011	An12g10020	51489	ATEG_06402							Cluster 029A						1		
AAO	related to aryl-alcohol oxidase	AA3_2		An13g01960	50882		AO090103000128	AFL2G_12244					Cluster 0330						1		
AAO	related to aryl-alcohol oxidase	AA3_2				ATEG_06274							Cluster 031A								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN4212										Cluster 031B								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN3229										Cluster 031C								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN7812				AO090026000028	AFL2G_07213					Cluster 032						1		
AAO	related to aryl-alcohol oxidase	AA3_2					AO090023000328	AFL2G_04181					Cluster 033							1	
AAO	related to aryl-alcohol oxidase	AA3_2	AN7267	An11g05600	208861		AO090001000619	AFL2G_09264					Cluster 034A					1			
AAO	related to aryl-alcohol oxidase	AA3_2	AN3531										Cluster 034B								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN9348										Cluster 035								1
AAO	related to aryl-alcohol oxidase	AA3_2	AN2704										Cluster 036								1
AAO	related to aryl-alcohol oxidase	AA3_2		An12g07280	42202		AO090010000215						Cluster 037							1	
AAO	related to aryl-alcohol oxidase	AA3_2		An04g08890	45887								Cluster 038								1
AAO	related to aryl-alcohol oxidase	AA3_2		An12g04530	131581								Cluster 039								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2	AN1429	An16g09050	51246	ATEG_00021	AO090103000022	AFL2G_12335	ACLA_057420	NFIA_096800	AFUB_083470	Afu8g04090	Cluster 040A	1							
AAO	distantly related to aryl-alcohol oxidase	AA3_2										Afu1g00460	Cluster 040B								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2	AN7832	An12g10770	45958	ATEG_00487	AO090003000032	AFL2G_02909		NFIA_041860	AFUB_049100	Afu5g00630	Cluster 041A		1						
AAO	distantly related to aryl-alcohol oxidase	AA3_2		An16g05660	40960	ATEG_00801	AO090120000309	AFL2G_08315		NFIA_084880			Cluster 041B				1				
AAO	distantly related to aryl-alcohol oxidase	AA3_2								NFIA_005130			Cluster 041C								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2								NFIA_093020			Cluster 042								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2					AO090103000110						Cluster 043								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2					AO090026000068						Cluster 044								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2	AN7998										Cluster 045								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2	AN7056										Cluster 046								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2	AN3960										Cluster 047								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2		An03g00130	205156								Cluster 048								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2		An18g00530	139484								Cluster 049								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2		An15g04820	183069								Cluster 050								1
AAO	distantly related to aryl-alcohol oxidase	AA3_2		An18g00940	187819								Cluster 051								1
	alcohol oxidase	AA3_3	AN0567	An18g05480	42956	ATEG_09188	AO090001000447	AFL2G_09093					Cluster 052				1				
	candidate alcohol oxidase	AA3_3		An15g02200	56311		AO090003000309	AFL2G_02671					Cluster 053						1		
	candidate pyranose oxidase	AA3_4	AN5281				AO090001000158	AFL2G_07394					Cluster 054						1		
VAO	related to vanillyl-alcohol oxidase	AA4	AN7068	An09g01380	189030		AO090003000878				AFUB_039670	Afu3g09500	Cluster 055					1			
	candidate galactose oxidase	132-AA5_2					AO090011000020	AFL2G_04827	ACLA_076740	NFIA_023620			Cluster 058					1			
	candidate galactose oxidase	132-AA5_2				ATEG_06260	AO090308000021	AFL2G_00002		NFIA_028120	AFUB_092400	Afu7g06820	Cluster 071				1				
	candidate pyranose oxidase	AA6	AN5281				AO090001000158	AFL2G_07394					Cluster 056						1		
	candidate benzoquinone reductase	AA6	AN0297	An01g05340	206038	ATEG_01946	AO090005000812	AFL2G_00793	ACLA_031610	NFIA_021810	AFUB_092400	Afu7g062820	Cluster 057	1							
	candidate galactose oxidase	3M32-AA7					AO090011000020	AFL2G_04827	ACLA_076740	NFIA_023620			Cluster 058					1			
	related to glucooligosaccharide oxidase	AA7	AN10388	An03g05210	44512	ATEG_01295	AO090026000299	AFL2G_06956	ACLA_057620	NFIA_060220	AFUB_000410	Afu6g14340	Cluster 059	1							
	related to glucooligosaccharide oxidase	AA7		An01g14520	35412								Cluster 060								1
	candidate benzoquinone reductase	AA8	AN0297	An01g05340	206038	ATEG_01946	AO090005000812	AFL2G_00793	ACLA_031610	NFIA_021810	AFUB_092400	Afu7g062820	Cluster 061	1							
	candidate iron reductase domain	AA8								NFIA_098570	AFUB_081740	Afu8g05805	Cluster 060A							1	

Supplemental Table 4B Cont: Orthology clusters of auxiliary activities (AA).

	candidate iron reductase domain	AA8	AN8890	An14g02900	134460	ATEG_10370	AO090010000757	AFL2G_11894	ACLA_059690	NFIA_045610			Cluster 060B		1							
	candidate iron reductase domain	AA8				ATEG_02996	AO090026000220	AFL2G_07032		NFIA_063710	AFUB_035540	Afu3g13650	Cluster 061A				1					
	candidate iron reductase domain	AA8				ATEG_10246			ACLA_054830				Cluster 061B								1	
	candidate iron reductase domain	AA8				ATEG_08150							Cluster 062A									1
	candidate iron reductase domain	AA8	AN10820	An02g06090	173811	ATEG_01526	AO090011000593	AFL2G_05344					Cluster 062B				1					
	candidate iron reductase domain	AA8	AN12326			ATEG_04326	AO090023000095	AFL2G_03990					Cluster 063						1			
	candidate iron reductase domain	AA8				ATEG_08151							Cluster 064									1
	candidate iron reductase domain	AA8					AO090003001458	AFL2G_01626					Cluster 065								1	
	candidate iron reductase domain	AA8	AN10488										Cluster 066									1
	candidate iron reductase domain	AA8		An11g04100	39013								Cluster 067									1
CDH	candidate cellobiose dehydrogenase	AA8-AA3_1	AN7230			ATEG_09993	AO090102000058	AFL2G_09511	ACLA_076510	NFIA_093020	AFUB_033310	Afu2g17620	Cluster 068		1							
CDH	cellobiose dehydrogenase	AA8-AA3_1		An10g00390	192610	ATEG_08151		AFL2G_06016	ACLA_094490	NFIA_033480	AFUB_018620	Afu2g01180	Cluster 069			1						
CDH	candidate cellobiose dehydrogenase	AA8-AA3_1		An02g09270	173720	ATEG_00720	AO090113000054	AFL2G_08542					Cluster 070						1			
LPMO	lytic polysaccharide monooxygenase	AA9	AN10419										Cluster 242									1
LPMO	lytic polysaccharide monooxygenase	AA9		An12g04610	211595	ATEG_04210			ACLA_060890	NFIA_006140	AFUB_044090	Afu3g03870	Cluster 243A				1					
LPMO	lytic polysaccharide monooxygenase	AA9	AN9524	An14g02670	53797	ATEG_07920	AO090012000090	AFL2G_03026		NFIA_044390	AFUB_101610		Cluster 243B		1							
LPMO	lytic polysaccharide monooxygenase	AA9	AN1041	An08g05230	52688	ATEG_00448	AO090001000221	AFL2G_07454	ACLA_022980	NFIA_012990	AFUB_012030	Afu1g12560	Cluster 244	1								
LPMO	lytic polysaccharide monooxygenase	AA9	AN1602	An15g04900	182430	ATEG_07790	AO090005000531	AFL2G_00532	ACLA_059790	NFIA_099510	AFUB_080950	Afu8g06830	Cluster 245	1								
LPMO	lytic polysaccharide monooxygenase	AA9	AN2388			ATEG_01035	AO090023000056		ACLA_073030	NFIA_089670	AFUB_030110	Afu2g14490	Cluster 246			1						
LPMO	lytic polysaccharide monooxygenase	AA9	AN3046	An12g02540	43784	ATEG_08113	AO090023000159	AFL2G_04048	ACLA_055060	NFIA_089730	AFUB_030160	Afu2g14540	Cluster 247A	1								
LPMO	lytic polysaccharide monooxygenase	AA9							ACLA_017470				Cluster 247B									1
LPMO	lytic polysaccharide monooxygenase	AA9				ATEG_05416	AO090023000787	AFL2G_04596	ACLA_047220	NFIA_108320	AFUB_064960	Afu4g07850	Cluster 248			1						
LPMO	lytic polysaccharide monooxygenase	AA9	AN6428			ATEG_06077							Cluster 249								1	
LPMO	lytic polysaccharide monooxygenase	AA9		An15g04570	56338	ATEG_10194							Cluster 250								1	
LPMO	lytic polysaccharide monooxygenase	AA9				ATEG_05081	AO090103000087	AFL2G_12282		NFIA_055220	AFUB_075590	Afu6g09540	Cluster 251				1					
LPMO	lytic polysaccharide monooxygenase	AA9	AN7891	An04g08550	194765		AO090138000004	AFL2G_08699					Cluster 252						1			
LPMO	lytic polysaccharide monooxygenase	AA9				ATEG_08942							Cluster 253									1
LPMO	lytic polysaccharide monooxygenase	AA9	AN3860										Cluster 254									1
LPMO	lytic polysaccharide monooxygenase	AA9				ATEG_01456							Cluster 255									1
	unique genes per species		13	15	15	8	3	2	2		0	0		13	7	5	10	13	14	9	49		
	strain specific genes			3	0						1	2		10.8	5.8	4.2	8.3	10.8	11.7	7.5	40.8		

Supplemental Table 5: Composition of the plant biomass substrates used in this study.
Values are given in mol%. Polysaccharide composition is inferred from the monomer values.

	Rhamnose	Arabinose	Xylose	Mannose	Galactose	Glucose	Uronic acid	Polysaccharides
Wheat bran	0	17	35	1	2	42	3	cellulose, (arabino)xylan
Sugar beet pulp	1	28	2	2	7	33	26	cellulose, pectin, xyloglucan

Supplemental Table 6A: Detection of proteins in cultures grown on wheat bran sorted by CAZy family. The colour codes indicate which percentage of the total number of detected peptides was of the specific protein. (Continued on next 4 pages).

			0-0.1%	0.1-0.5%	0.5-1%	1-2%	2-5%	5-10%	>10%	
enzyme code	**function**	**CAZY**	***A. nidulans***	***A. niger* ATCC 1015**	***A. terreus***	***A. oryzae***	***A. flavus***	***A. clavatus***	***A. fischeri***	***A. fumigatus* Af293**
FAE SF7	feruloyl esterase	faeA		51662	ATEG_08907	AO090001000207	AFL2G_07436			
FAE SF1	feruloyl esterase	SF1	AN1772	51478	ATEG_06663	AO090001000582	AFL2G_09228	ACLA_083360	NFIA_054700	Afu6g09040
FAE SF1	feruloyl esterase	SF1			ATEG_02212	AO090001000066	AFL2G_07310		NFIA_047590	Afu6g00450
FAE SF3	feruloyl esterase	SF3				AO090102000013				
FAE SF3	feruloyl esterase	SF3			ATEG_02415					
FAE SF4	feruloyl esterase	SF4		190471		AO090010000573	AFL2G_11725			
FAE SF4	feruloyl esterase	SF4		43194						
AXE	acetyl xylan esterase	CE 1	AN6093	211544	ATEG_09843	AO090011000745	AFL2G_05471	ACLA_081220	NFIA_099230	Afu8g06570
AXE	acetyl xylan esterase	CE 1	AN8320							Afu1g17510
AXE	acetyl xylan esterase	CE 1							NFIA_084920	
FAE SF5	feruloyl esterase	CE 1				AO090005000945	AFL2G_00922	ACLA_065130	NFIA_115130	
FAE SF5	feruloyl esterase	CE 1	AN5267	43785	ATEG_08112	AO090023000158	AFL2G_04047	ACLA_055050	NFIA_089720	Afu2g14530
FAE SF5	feruloyl esterase	CE 1			ATEG_06644	AO090701000884	AFL2G_06446	ACLA_061520		
FAE SF5	feruloyl esterase	CE 1								Afu2g09440
FAE SF5	feruloyl esterase	CE 1			ATEG_01914					
FAE SF5	feruloyl esterase	CE 1						ACLA_017480		
?		CE 1	AN8782	53315	ATEG_06438	AO090005000277	AFL2G_00281	ACLA_012870	NFIA_077450	Afu5g09860
PME	pectin methyl esterase	CE 8				AO090003001514	AFL2G_01578			
PME	pectin methyl esterase	CE 8		44585		AO090012000749	AFL2G_03618	ACLA_044240	NFIA_095020	Afu8g01520
PME	pectin methyl esterase	CE 8	AN4860	174365	ATEG_01704	AO090020000404	AFL2G_10916	ACLA_035610	NFIA_069500	Afu3g07650
PME	pectin methyl esterase	CE 8	AN3390	214857		AO090102000010	AFL2G_09467		NFIA_099600	Afu8g06880
PME	pectin methyl esterase	CE 8	AN7966							
PME	pectin methyl esterase	CE 8				AO090113000039			NFIA_100100	Afu8g07250
PME	pectin methyl esterase	CE 8						ACLA_059970		
PME	pectin methyl esterase	CE 8					AFL2G_08528			
PME	pectin methyl esterase	CE 8						ACLA_059980		
RGAE	rhamnogalacturonan acetyl esterase	CE 12				AO090003001268	AFL2G_01797			
RGAE	rhamnogalacturonan acetyl esterase	CE 12							NFIA_099110	Afu8g06480
RGAE	rhamnogalacturonan acetyl esterase	CE 12		51400	ATEG_10016	AO090102000092	AFL2G_09543			
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2834			AO090113000155	AFL2G_08631		NFIA_092640	Afu2g17250
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2528	189254	ATEG_03511	AO090701000556	AFL2G_06177	ALCL_041970	NFIA_062750	Afu3g14510
GE	glucuronoyl esterase	CE15			ATEG_00945			ACLA_087520	NFIA_060260	Afu6g14390
BGL	beta glucosidase	GH 1		140573						
BGL	beta glucosidase	GH 1	AN10124	213437	ATEG_00687	AO090003000497	AFL2G_02496	ACLA_020660	NFIA_010690	Afu1g14710
BGL	beta glucosidase	GH 1						ACLA_019180	NFIA_009040	Afu1g16400
BGL	beta glucosidase	GH 1				AO090113000148	AFL2G_08626		NFIA_060550	Afu6g14600
BGL	beta glucosidase	GH 1	AN9183	131747	ATEG_02657	AO090120000075	AFL2G_08111		NFIA_099670	Afu8g06970
BGL	beta glucosidase	GH 1	AN10375		ATEG_04135				NFIA_064710	Afu3g12600
BGL	beta glucosidase	GH 1						ACLA_064280		
BGL	beta glucosidase	GH 1						ACLA_040420		
GUS	beta glucuronidase	GH 2	AN3200	189620	ATEG_09745		AFL2G_05262			
GUS	beta glucuronidase	GH 2	AN2395	52111	ATEG_01031	AO090023000053	AFL2G_03956		NFIA_089700	Afu2g14520
LAC	beta galactosidase	GH 2			ATEG_10255					
LAC	beta galactosidase	GH 2			ATEG_04784					
LAC	beta galactosidase	GH 2	AN2463		ATEG_00712					
GUS	beta glucuronidase	GH 2	AN5361	46827		AO090038000009	AFL2G_08828			
LAC	beta galactosidase	GH 2	AN3201		ATEG_10243	AO090012000389	AFL2G_03296		NFIA_072410	Afu5g14550
LAC	beta galactosidase	GH 2	AN6388							
LAC	beta galactosidase	GH 2	AN1107						NFIA_072890	Afu5g14090
MND	beta mannosidase	GH 2	AN1742	138876	ATEG_06636	AO090001000556	AFL2G_09201	ACLA_083570	NFIA_054490	Afu6g08840
MND	beta mannosidase	GH 2	AN11680	172587	ATEG_07339	AO090003001410	AFL2G_01666	ACLA_078510	NFIA_045250	Afu4g00390
MND	beta mannosidase	GH 2				AO090005000740	AFL2G_00728			
MND	beta mannosidase	GH 2	AN3368	212893	ATEG_08684	AO090010000208	AFL2G_11464	ACLA_066240	NFIA_114030	Afu7g01320
MND	beta mannosidase	GH 2			ATEG_09890					
BGL	beta glucosidase	GH 3			ATEG_10274					
BGL	beta glucosidase	GH 3		139037						
BGL	beta glucosidase	GH 3								
BGL	beta glucosidase	GH 3		39613						
BGL	beta glucosidase	GH 3	AN7915	38077		AO090001000266	AFL2G_07497			
BGL	beta glucosidase	GH 3	AN10482	208871	ATEG_06617	AO090001000544	AFL2G_09187	ACLA_083710	NFIA_054350	Afu6g08700
BGL	beta glucosidase	GH 3								
BGL	beta glucosidase	GH 3		210981	ATEG_02806	AO090003001511	ALF2G_01582		NFIA_095760	Afu8g02100
BGL	beta glucosidase	GH 3			ATEG_02724	AO090005000337	AFL2G_00334			
BGL	beta glucosidase	GH 3	AN4102	56782	ATEG_03047	AO090009000356	AFL2G_10322	ACLA_028810	NFIA_018950	Afu1g05770
BGL	beta glucosidase	GH 3	AN6652	182309	ATEG_07121	AO090009000554	AFL2G_10164	ACLA_096980	NFIA_050080	Afu6g03570
BGL	beta glucosidase	GH 3		44520		AO090010000034	AFL2G_11300			
BGL	beta glucosidase	GH 3				AO090012000003	AFL2G_02949			
BGL	beta glucosidase	GH 3	AN7396	179265	ATEG_10320	AO090012000135	AFL2G_03066		NFIA_007920	Afu1g17410
BGL	beta glucosidase	GH 3	AN1804			AO090026000123	AFL2G_07119			
BGL	beta glucosidase	GH 3				AO090038000223	AFL2G_09023			
BGL	beta glucosidase	GH 3	AN5976	181816	ATEG_02713	AO090038000425	AFL2G_07763		NFIA_100430	
BGL	beta glucosidase	GH 3				AO090103000127	AFL2G_12245			
BGL	beta glucosidase	GH 3		176601		AO090166000048	AFL2G_09452		NFIA_112660	
BGL	beta glucosidase	GH 3	AN3903		ATEG_04069	AO090166000090	AFL2G_09413	ACLA_087610	NFIA_060370	Afu6g14490
BGL	beta glucosidase	GH 3	AN2227	129891	ATEG_09329	AO090701000244	AFL2G_05886	ACLA_010450	NFIA_080070	Afu5g07190
BGL	beta glucosidase	GH 3			ATEG_07931					
BGL	beta glucosidase	GH 3	AN0712							
BGL	beta glucosidase	GH 3	AN2612							Afu7g00240

Supplemental Table 6A Cont: Detection of proteins in cultures grown on wheat bran sorted by CAZy family. (Continued on next 3 pages).

BGL	beta glucosidase	GH 3	AN2828		ATEG_07419	AO090701000841	AFL2G_06408	ACLA_007810	NFIA_027390	Afu7g06140
BGL	beta glucosidase	GH 3	AN3949							
BGL	beta glucosidase	GH 3	AN7865						NFIA_057590	Afu6g11910
BGL	beta glucosidase	GH 3		129779	ATEG_00157					
BGL	beta glucosidase	GH 3							NFIA_000750	Afu3g00230
BGL	beta glucosidase	GH 3							NFIA_098520	
BGL	beta glucosidase	GH 3							NFIA_057910	
BGL/BXL	beta glucosidase/beta xylosidase	GH 3	AN0479							
BXL	beta xylosidase	GH 3	AN2359	205670	ATEG_05106	AO090005000986	AFL2G_00957			Afu1g16920
BXL	beta xylosidase	GH 3			ATEG_07383	AO090011000140	AFL2G_04928			
BXL	beta xylosidase	GH 3						ACLA_018590		
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN8401		ATEG_09052	AO090103000120	AFL2G_12252	ACLA_062400	NFIA_003180	Afu3g02090
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN2217	50997	ATEG_09314	AO090701000274	AFL2G_05912	ACLA_010340	NFIA_080180	Afu5g07080
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3			ATEG_08027					
?		GH 3	AN1416	45461	ATEG_00018	AO090103000019	AFL2G_12338	ACLA_057390	NFIA_096770	Afu8g04060
?		GH 3								Afu1g00540
?		GH 3	AN3360	37673	ATEG_04963			ACLA_052760	NFIA_102600	Afu4g13770
?		GH 3		120104	ATEG_04729	AO090003000741	AFL2G_02272	ACLA_039660	NFIA_065550	Afu3g11780
?		GH 3					AFL2G_08686			
?		GH 3	AN2599							
EGL	endoglucanase	GH 5	AN3013	210716	ATEG_01592	AO090005001389	AFL2G_01298	ACLA_036770	NFIA_068300	Afu3g08820
EGL	endoglucanase	GH 5	AN1285	205580	ATEG_05002			ACLA_085250	NFIA_057290	Afu6g11600
EGL	endoglucanase	GH 5							NFIA_095570	
EGL	endoglucanase	GH 5	AN8068		ATEG_09802	AO090003001341+A	AFL2G_01726		NFIA_040280	Afu5g01830
EGL	endoglucanase	GH 5	AN5214	209378		AO090005001553	AFL2G_01447	ACLA_081650	NFIA_053150	Afu6g07480
EGL	endoglucanase	GH 5		214608	ATEG_04390	AO090011000715	AFL2G_05447	ACLA_081310	NFIA_085010	Afu2g09520
EGL	endoglucanase	GH 5			ATEG_05003					
GLN	exo 1,6 galactanase	GH 5	AN9166	194447	ATEG_10242	AO090012000046	AFL2G_02982		NFIA_072400	Afu5g14560
EXG	exo 1,3 galactanase	GH 5			ATEG_08371				NFIA_056040	
EXG	exo 1,3 galactanase	GH 5	AN8947	175759	ATEG_07719					
EXG	exo 1,3 galactanase	GH 5	AN7533	123981	ATEG_06686	AO090001000604	AFL2G_09249	ACLA_083150	NFIA_054930	Afu6g09250
EXG	exo 1,3 galactanase	GH 5	AN4052	202490	ATEG_03849	AO090003000990	AFL2G_02039	ACLA_031040	NFIA_021060	Afu1g03600
EXG	exo 1,3 galactanase	GH 5			ATEG_06369	AO090005000423	AFL2G_00412	ACLA_007330	NFIA_026860	Afu7g05610
EXG	exo 1,3 galactanase	GH 5			ATEG_03062	AO090009000373	AFL2G_10308			
EXG	exo 1,3 galactanase	GH 5	AN1332	52811		AO090012000917	AFL2G_03782		NFIA_113230	
EXG	beta 1,6 glucanase	GH 5	AN3777		ATEG_09844	AO090011000757	AFL2G_05484		NFIA_084850	Afu2g09350
MAN	endomannanase	GH 5	AN3358	50378	ATEG_08654	AO090010000122	AFL2G_11381	ACLA_066420	NFIA_113780	Afu7g01070
MAN	endomannanase	GH 5			ATEG_10292	AO090012000006	AFL2G_02951		NFIA_041960	Afu5g00550
MAN	endomannanase	GH 5	AN7639			AO090038000444	AFL2G_07781	ACLA_044470	NFIA_099770	Afu8g07030
MAN	endomannanase	GH 5			ATEG_02669					
MAN	endomannanase	GH 5	AN2709							
MAN	endomannanase	GH 5	AN3297							
MAN	endomannanase	GH 5	AN6427		ATEG_09991					
MAN	endomannanase	GH 5	AN9276							
MAN	endomannanase	GH 5			ATEG_01374					
		GH 5			ATEG_03677					
CBH	cellobiohydrolase	GH 6	AN5282	54490	ATEG_07493			ACLA_062560	NFIA_002990	Afu3g01910
CBH	cellobiohydrolase	GH 6	AN1273	133986	ATEG_00193	AO090038000439	AFL2G_07776	ACLA_025560	NFIA_015680	
CBH	cellobiohydrolase	GH 7	AN0494	51773	ATEG_05002	AO090001000348	AFL2G_07571	ACLA_085260	NFIA_057300	Afu6g11610
CBH	cellobiohydrolase	GH 7							NFIA_095570	
CBH	cellobiohydrolase	GH 7	AN5176	53159	ATEG_03727	AO090012000941	AFL2G_03805	ACLA_088870	NFIA_052720	Afu6g07070
EGL	endoglucanase	GH 7	AN3418		ATEG_08700	AO090010000314	ALF2G_11497	ACLA_066030	NFIA_114250	Afu7g01540
EGL	endoglucanase	GH 7			ATEG_08705			ACLA_098940	NFIA_047960	Afu6g01800
XLN	endoxylanase	GH 10				AO090001000208	AFL2G_07437			
XLN	endoxylanase	GH 10	AN7401		ATEG_03410	AO090103000326	AFL2G_12071	ACLA_088910	NFIA_059570	Afu6g13610
XLN	endoxylanase	GH 10	AN1818	57436	ATEG_00809	AO090701000887	AFL2G_06449	ACLA_048770	NFIA_106540	Afu4g09480
XLN	endoxylanase	GH 10			ATEG_08906	AO090103000423	AFL2G_11983			
XLN	endoxylanase	GH 10	AN2356	50977					NFIA_057510	Afu6g11820
XLN	endoxylanase	GH 10			ATEG_07190				NFIA_061880	Afu3g15210
XLN	endoxylanase	GH 11	AN3613							
XLN	endoxylanase	GH 11								
XLN	endoxylanase	GH 11		171269	ATEG_04943	AO090001000111	AFL2G_07347	ACLA_085410	NFIA_058160	Afu6g12210
XLN	endoxylanase	GH 11		183088		AO090026000103	AFL2G_07138		NFIA_055240	
XLN	endoxylanase	GH 11				AO090103000141	AFL2G_12233			
XLN	endoxylanase	GH 11	AN9365	52071	ATEG_07461	AO090120000026	AFL2G_08066	ACLA_063140	NFIA_000850	Afu3g00320
XLN	endoxylanase	GH 11							NFIA_001010	Afu3g00470
XLN	endoxylanase	GH 11						ACLA_064270		
XG EGL	xyloglucan active endoglucanase	GH 12	AN0452	52011	ATEG_03755	AO090003000905	AFL2G_02120	ACLA_029940	NFIA_020020	Afu1g04730
XG EGL	xyloglucan active endoglucanase	GH 12		211053	ATEG_07420	AO090026000102	AFL2G_07140	ACLA_007820	NFIA_027400	Afu7g06150
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_05519					
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_09894					
XG EGL	xyloglucan active endoglucanase	GH 12				AO090038000175	AFL2G_08984		NFIA_002040	
XG EGL	xyloglucan active endoglucanase	GH 12				AO090701000185	AFL2G_05831	ACLA_060930	NFIA_005810	Afu3g03610
XG EGL	xyloglucan active endoglucanase	GH 12		191511					NFIA_100330	
XG EGL	xyloglucan active endoglucanase	GH 12					AFL2G_04893			
XG EGL	xyloglucan active endoglucanase	GH 12								Afu3g01160
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_07466					
AGT	4 alpha glucanotransferase	GH 13		57002						
?		GH 13		182162						
GBA	glycogen debranching enzyme	GH 13	AN2314	211162	ATEG_07608	AO090010000483	AFL2G_11646	ACLA_013550	NFIA_076670	Afu5g10540
AGD	alpha glucosidase	GH 13		50927						
AGD	alpha glucosidase	GH 13	AN4843	52452	ATEG_01729	AO090020000176	AFL2G_11111	ACLA_035550	NFIA_069880	Afu3g07380
AGD	alpha glucosidase	GH 13				AO090026000034	AFL2G_07207			
AGD	alpha glucosidase	GH 13				AO090038000234	AFL2G_09036	ACLA_008120	NFIA_027660	Afu7g06380
AGD	alpha glucosidase	GH 13	AN10420		ATEG_01029	AO090103000129	AFL2G_12243	ACLA_059930	NFIA_099830	Afu8g07070
AGD	alpha glucosidase	GH 13				AO090103000378	AFL2G_12021			

Supplemental Table 6A Cont: Detection of proteins in cultures grown on wheat bran sorted by CAZy family. (Continued on next 2 pages).

AGD	alpha glucosidase	GH 13					AFL2G_08694			
AGD	alpha glucosidase	GH 13						ACLA_070570	NFIA_086860	Afu2g11620
AGD	alpha glucosidase	GH 13						ACLA_070560	NFIA_086850	Afu2g11610
AGS	alpha glucan synthase	GH 13		212915						
AGS	alpha glucan synthase	GH 13		40878						
AGS	alpha glucan synthase	GH 13	AN3307	54378	ATEG_03622	AO090003001500	AFL2G_01593	ACLA_063450	NFIA_001720	Afu3g00910
AGS	alpha glucan synthase	GH 13		55204	ATEG_10371	AO090010000106	AFL2G_11365	ACLA_042430	NFIA_009960	Afu1g15440
AGS	alpha glucan synthase	GH 13	AN5885		ATEG_01449	AO090026000523	AFL2G_06761	ACLA_070240	NFIA_086550	Afu2g11270
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13				AO090120000196		ACLA_094070	NFIA_032970	Afu2g00710
AMY	alpha amylase	GH 13	AN3402	47911	ATEG_10103	AO090023000944		ACLA_052920		
AMY	alpha amylase	GH 13	AN2018	45304	ATEG_08279	AO090003001210	AFL2G_01841	ACLA_049350	NFIA_105920	Afu4g10130
AGT	4 alpha glucanotransferase	GH 13	AN3308	188489	ATEG_03623	AO090003001498	AFL2G_01594	ACLA_063440	NFIA_001710	Afu3g00900
AMY	alpha amylase	GH 13				AO090120000263	AFL2G_08276			
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13	AN4507							
AMY	alpha amylase	GH 13	AN3388		ATEG_02515					
AMY	alpha amylase	GH 13	AN6324					ACLA_072270	NFIA_088650	Afu2g13460
AMY	alpha amylase	GH 13		122069	ATEG_03624	AO090003001497	AFL2G_01595			
AMY	alpha amylase	GH 13	AN10060	46621	ATEG_04879	AO090005000884	AFL2G_00860	ACLA_032290	NFIA_022510	Afu1g02140
AMY	alpha amylase	GH 13	AN3309	46290	ATEG_00838	AO090005001193		ACLA_020060	NFIA_010270	Afu1g15150
AMY	alpha amylase	GH 13			ATEG_00724					
AMY	alpha amylase	GH 13					AFL2G_01127			
AMY	alpha amylase	GH 13						ACLA_091300	NFIA_035590	Afu2g03230
AMY	alpha amylase	GH 13							NFIA_035580	
GLA	glucoamylase	GH 15	AN7402					ACLA_049360	NFIA_105910	
GLA	glucoamylase	GH 15								Afu4g10140
GLA	glucoamylase	GH 15			ATEG_05980	AO090003000321	AFL2G_02658	ACLA_089470		
GLA	glucoamylase	GH 15	AN11143	213597	ATEG_04375	AO090010000746	AFL2G_11885	ACLA_094080	NFIA_032960	Afu2g00690
GLA	glucoamylase	GH 15		189911		AO090138000105	AFL2G_08782	ACLA_044430	NFIA_094880	Afu8g01390
GLA	glucoamylase	GH 15						ACLA_082570	NFIA_053760	Afu6g08080
GLA	glucoamylase	GH 15						ACLA_078620	NFIA_001210	Afu3g00610
MAN	endomannanase	GH 26	AN3336	40875		AO090011000055	AFL2G_04859			
MAN	endomannanase	GH 26	AN3326							
MAN	endomannanase	GH 26	AN7413							
AGL	alpha galactosidase	GH 27			ATEG_01905				NFIA_048850	Afu6g02560
AGL	alpha galactosidase	GH 27		37736	ATEG_04382			ACLA_016820	NFIA_073100	Afu5g13830
AGL	alpha galactosidase	GH 27	AN0022	172232	ATEG_03427	AO090005000217	AFL2G_00225		NFIA_023390	Afu1g01200
AGL	alpha galactosidase	GH 27	AN7624	207264	ATEG_09830	AO090003001305	AFL2G_01765	ACLA_003130	NFIA_039990	Afu5g02130
AGL	alpha galactosidase	GH 27		39180						
AGL	alpha galactosidase	GH 27	AN7152	41606	ATEG_02160	AO090023000151	AFL2G_04039		NFIA_029860	Afu4g03580
AGL	alpha galactosidase	GH 27						ACLA_097900		
PGA	endopolygalacturonase	GH 28	AN8327	46255					NFIA_008150	Afu1g17220
PGA	endopolygalacturonase	GH 28		43957						
PGA	endopolygalacturonase	GH 28		182156						
PGA	endopolygalacturonase	GH 28		214598						
PGA	endopolygalacturonase	GH 28	AN6656	50161		AO090005000186	AFL2G_00201		NFIA_099410	
PGA	endopolygalacturonase	GH 28		172944	ATEG_01601	AO090005001400	AFL2G_01310	ACLA_036670	NFIA_068440	Afu3g08680
PGA	endopolygalacturonase	GH 28	AN4372	141677	ATEG_04991	AO090023000401	AFL2G_04252	ACLA_052860	NFIA_102450	Afu4g13920
PGA	endopolygalacturonase	GH 28		52219		AO090023000161	AFL2G_04049		NFIA_095620	Afu8g01970
PGA	endopolygalacturonase	GH 28				AO090138000086				
PGA	endopolygalacturonase	GH 28								Afu8g06730
PGA	endopolygalacturonase	GH 28			ATEG_07748					
PGA	endopolygalacturonase	GH 28					AFL2G_08764		NFIA_023290	Afu1g01320
PGX	exopolygalacturonase	GH 28		178172						
PGX	exopolygalacturonase	GH 28	AN8891	191158	ATEG_10357	AO090010000753	AFL2G_11892	ACLA_043100	NFIA_096340	Afu8g02630
PGX	exopolygalacturonase	GH 28	AN8761	42184	ATEG_07152	AO090026000784	AFL2G_06533		NFIA_049320	Afu6g02980
PGX	exopolygalacturonase	GH 28	AN9045							
RGX	exorhamnogalacturonase	GH 28		172236		AO090001000133	AFL2G_07371			
RGX	exorhamnogalacturonase	GH 28		42917		AO090009000470	AFL2G_10228		NFIA_018590	Afu1g06140
RGX	exorhamnogalacturonase	GH 28	AN10274	194461	ATEG_09025	AO090102000139	AFL2G_09582		NFIA_027700	Afu7g06410
RGX	exorhamnogalacturonase	GH 28			ATEG_06408	AO090113000199	AFL2G_08671			
RGX	exorhamnogalacturonase	GH 28				AO090138000066	AFL2G_03102			
RGX	exorhamnogalacturonase	GH 28	AN11626							
RGX	exorhamnogalacturonase	GH 28				AO090138000067	AFL2G_08746			
RHG	endorhamnogalacturonase	GH 28		178393						
RHG	endorhamnogalacturonase	GH 28		180922		AO090003000524	AFL2G_02475			Afu4g00100
RHG	endorhamnogalacturonase	GH 28				AO090005000067	AFL2G_00087			
RHG	endorhamnogalacturonase	GH 28		211163	ATEG_07607	AO090010000484	AFL2G_11646			
RHG	endorhamnogalacturonase	GH 28		123651		AO090026000252	AFL2G_06999			
RHG	endorhamnogalacturonase	GH 28	AN9134	189722		AO090038000552	AFL2G_07883			
RHG	endorhamnogalacturonase	GH 28		39337		AO090124000009	AFL2G_08037			
RHG	endorhamnogalacturonase	GH 28							NFIA_076680	Afu5g10530
XGH	xylogalacturonase	GH 28				AO090026000120	AFL2G_07122			
XGH	xylogalacturonase	GH 28	AN3389	46065		AO090102000011	AFL2G_09468		NFIA_099610	Afu8g06890
PGX	exopolygalacturonase	GH 28							NFIA_100120	Afu8g07265
AFC	alpha fucosidase	GH 29		44822	ATEG_08111					
AFC	alpha fucosidase	GH 29			ATEG_05691					
AGD	alpha glucosidase	GH 31	AN2017	214233	ATEG_00723	AO090003001209	AFL2G_01842	ACLA_049370	NFIA_105900	Afu4g10150
AGD	alpha glucosidase	GH 31	AN0280	55419	ATEG_02528	AO090005000767	AFL2G_00750	ACLA_031260	NFIA_021450	Afu1g03140
AGD	alpha glucosidase	GH 31	AN0941	119858	ATEG_05177	AO090005001084	AFL2G_01038	ACLA_019300	NFIA_009180	Afu1g16250
AGD	alpha glucosidase	GH 31	AN3504	49940	ATEG_02966	AO090023000288	AFL2G_04152		NFIA_018130	Afu1g06560
AGD	alpha glucosidase	GH 31	AN10935			AO090026000111	AFL2G_07131			
AGD	alpha glucosidase	GH 31	AN8953			AO090038000471	AFL2G_07812			
AGD	alpha glucosidase	GH 31	AN11054	50055	ATEG_08065	AO090102000559	AFL2G_09934	ACLA_001670	NFIA_038610	Afu5g03500

Supplemental Table 6A Cont: Detection of proteins in cultures grown on wheat bran sorted by CAZy family. (Continued on next page).

AGD	alpha glucosidase	GH 31	AN7120							
AGD	alpha glucosidase	GH 31	AN7345							
AGD	alpha glucosidase	GH 31			ATEG_08278					
AGD	alpha glucosidase	GH 31			ATEG_08472					
AGD	alpha glucosidase	GH 31							NFIA_060440	
AXL	alpha xylosidase	GH 31	AN7505		ATEG_06730	AO090001000649	AFL2G_09290	ACLA_089300	NFIA_082140	Afu2g05400
AXL	alpha xylosidase	GH 31		40261		AO090701000558	AFL2G_06180			
AXL	alpha xylosidase	GH 31		43342	ATEG_08390	AO090701000639	AFL2G_06239		NFIA_032680	
AXL	alpha xylosidase	GH 31			ATEG_00207					
AXL	alpha xylosidase	GH 31			ATEG_05169					
INU	endo inulinase	GH 32		52928						Afu5g00530
INU	endo inulinase	GH 32								Afu5g00480
INU	endo inulinase	GH 32			ATEG_07356					
INX	exo inulinase	GH 32								
INX	exo inulinase	GH 32	AN11778	56664	ATEG_08155	AO090701000400	AFL2G_06028	ACLA_094550	NFIA_033540	Afu2g01240
SUC	invertase/beta fructofuranosidase	GH 32	AN3837		ATEG_05860	AO090701000038	AFL2G_05693		NFIA_051560	Afu6g05000
SUC	invertase/beta fructofuranosidase	GH 32		198063	ATEG_07479	AO090020000640	AFL2G_10707			
SUC	invertase/beta fructofuranosidase	GH 32				AO090103000043	AFL2G_12317			
SUC	invertase/beta fructofuranosidase	GH 32		176039						
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_04996					
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_10253					
LAC	beta galactosidase	GH 35	AN0194							
LAC	beta galactosidase	GH 35		180727						
LAC	beta galactosidase	GH 35				AO090001000259	AFL2G_07489			
LAC	beta galactosidase	GH 35		177434	ATEG_07446	AO090003000042	AFL2G_02901	ACLA_088440	NFIA_052310	Afu6g06660
LAC	beta galactosidase	GH 35	AN0756	51764	ATEG_00616	AO090012000445	AFL2G_03352	ACLA_021260	NFIA_011250	Afu1g14170
LAC	beta galactosidase	GH 35				AO090012000746				
LAC	beta galactosidase	GH 35	AN2804		ATEG_08255	AO090023000076	AFL2G_03977		NFIA_041490	Afu5g00670
LAC	beta galactosidase	GH 35	AN0980	46429	ATEG_05131	AO090120000158	AFL2G_08182		NFIA_008690	Afu1g16700
LAC	beta galactosidase	GH 35		41910			AFL2G_03616		NFIA_000910	Afu3g00380
LAC	beta galactosidase	GH 35				AO090701000770				
LAC	beta galactosidase	GH 35						ACLA_018830		
LAC	beta galactosidase	GH 35					AFL2G_06343			
LAC	beta galactosidase	GH 35					AFL2G_12387			
AGL	alpha galactosidase	GH 36	AN8138	212736	ATEG_07929	AO090010000684	AFL2G_11824	ACLA_044620	NFIA_094580	Afu8g01130
AGL	alpha galactosidase	GH 36	AN8392		ATEG_07935	AO090011000063	AFL2G_04865	ACLA_044560	NFIA_094540	Afu8g01100
AGL	alpha galactosidase	GH 36	AN3874	190816	ATEG_05379	AO090023000743	AFL2G_04555	ACLA_047600	NFIA_107920	Afu4g08250
AGL	alpha galactosidase	GH 36	AN9035		ATEG_02312					
ABF	alpha arabinofuranosidase	GH 43	AN8472			AO090012000356	AFL2G_03265		NFIA_032900	Afu2g00650
ABF	alpha arabinofuranosidase	GH 43	AN7781		ATEG_08386	AO090050000065	AFL2G_00086			
ABF	alpha arabinofuranosidase	GH 43	AN2533			AO090701000838	AFL2G_06400		NFIA_002750	Afu3g01660
ABN	endoarabinanase	GH 43		128077	ATEG_05780				NFIA_078690	Afu5g08590
ABN	endoarabinanase	GH 43		182100	ATEG_07817	AO090026000804	AFL2G_06513	ACLA_072730	NFIA_089980	Afu2g14750
ABN	endoarabinanase	GH 43	AN3044	184195	ATEG_01562	AO090005001320	AFL2G_01244	ACLA_037080	NFIA_068020	Afu3g09140
ABN	endoarabinanase	GH 43						ACLA_018180	NFIA_008030	Afu1g17320
ABN	endoarabinanase	GH 43	AN6352	134398	ATEG_01407	AO090023000165	AFL2G_04053	ACLA_073460	NFIA_089320	Afu2g14150
ABN*	endoarabinanase	GH 43	AN8007	197735		AO090138000055	AFL2G_08737	ACLA_098980	NFIA_047920	Afu6g00770
ABN	endoarabinanase	GH 43	AN2534	203143	ATEG_03520	AO090701000481	AFL2G_06106	ACLA_042100	NFIA_062660	Afu3g14620
ABN	endoarabinanase	GH 43								
ABN	endoarabinanase	GH 43				AO090005000064	AFL2G_00085			
ABN	endoarabinanase	GH 43			ATEG_03688					
ABN	endoarabinanase	GH 43			ATEG_07787					
BXL	beta xylosidase	GH 43	AN1870	179682	ATEG_06059	AO090003000239	AFL2G_02739	ACLA_090090	NFIA_081240	Afu2g04480
BXL	beta xylosidase	GH 43	AN6751	174379	ATEG_06306	AO090005000476	AFL2G_00456		NFIA_027350	Afu7g06110
BXL	beta xylosidase	GH 43			ATEG_05083	AO090103000268	AFL2G_12122	ACLA_087680	NFIA_060490	Afu6g14550
BXL	beta xylosidase	GH 43	AN10199		ATEG_00893	AO090005000698	AFL2G_00689/0	ACLA_058100/(	NFIA_097490/0	Afu8g04710/047
BXL	beta xylosidase	GH 43	AN10919		ATEG_10072	AO090010000029	AFL2G_11296		NFIA_093350	
BXL	beta xylosidase	GH 43			ATEG_07784	AO090010000494	AFL2G_11656		NFIA_023360	Afu1g01230
BXL	beta xylosidase	GH 43	AN7313	122978	ATEG_01188	AO090010000562	AFL2G_11714	ACLA_043260	NFIA_096240	Afu8g02510
BXL	beta xylosidase	GH 43	AN2664		ATEG_06107	AO090012000350	AFL2G_03257		NFIA_072780	Afu5g14190
BXL	beta xylosidase	GH 43	AN8477		ATEG_01292		AFL2G_06974	ACLA_018160		
BXL	beta xylosidase	GH 43				AO090113000059				
BXL	beta xylosidase	GH 43			ATEG_10193					
BXL	beta xylosidase	GH 43	AN7275		ATEG_06643	AO090701000886	AFL2G_06448	ACLA_078600	NFIA_033220	Afu2g00930
BXL	beta xylosidase	GH 43	AN7864	47677	ATEG_01292	AO090102000331	AFL2G_09750	ACLA_072000	NFIA_088370	Afu2g13190
BXL	beta xylosidase	GH 43			ATEG_07517					
BXL	beta xylosidase	GH 43	AN2633							
BXL	beta xylosidase	GH 43	AN1043							
BXL	beta xylosidase	GH 43					AFL2G_08528			
?	?	GH 43	AN1477							
EGL	endoglucanase	GH 45	AN6786						NFIA_028020	Afu7g06740
ABF	alpha arabinofuranosidase	GH 51	AN1277	38549	ATEG_00198		AFL2G_07796	ACLA_025610	NFIA_015730	Afu1g09900
ABF	alpha arabinofuranosidase	GH 51		50979	ATEG_03540					
ABF	alpha arabinofuranosidase	GH 51				AO090020000712	AFL2G_10643			
ABF	alpha arabinofuranosidase	GH 51	AN2541	206387	ATEG_02882	AO090124000023	AFL2G_08019	ACLA_099110		
ABF	alpha arabinofuranosidase	GH 51	AN9439	131891	ATEG_07868	AO090012000298	AFL2G_03217	ACLA_074120	NFIA_090410	Afu2g15160
GAL	endogalactanase	GH 53								
GAL	endogalactanase	GH 53	AN5727	187227	ATEG_02927	AO090001000492	AFL2G_09135		NFIA_017780	Afu1g06910
GAL	endogalactanase	GH 53					AFL2G_07488			
ABF	alpha arabinofuranosidase	GH 54	AN1571	200605	ATEG_07939	AO090023000001	AFL2G_03901	ACLA_066470	NFIA_060630	Afu6g14620
CO	cellulose oxidase	GH 61	AN10419							
CO	cellulose oxidase	GH 61		211595	ATEG_04210			ACLA_060890	NFIA_006140	Afu3g03870
CO	cellulose oxidase	GH 61	AN1041	52688	ATEG_00448	AO090001000221	AFL2G_07454	ACLA_022980	NFIA_012990	Afu1g12560
CO	cellulose oxidase	GH 61	AN1602	182430	ATEG_07790	AO090005000531	AFL2G_00532	ACLA_059790	NFIA_099510	Afu8g06830
CO	cellulose oxidase	GH 61	AN9524	53797	ATEG_07920	AO090012000090	AFL2G_03026		NFIA_044390	
CO	cellulose oxidase	GH 61	AN2388		ATEG_01035	AO090023000056		ACLA_073030	NFIA_089670	Afu2g14490

Supplemental Table 6A Cont: Detection of proteins in cultures grown on wheat bran sorted by CAZy family.

CO	cellulose oxidase	GH 61	AN3046	43784	ATEG_08113	AO090023000159	AFL2G_04048	ACLA_055060	NFIA_089730	Afu2g14540
CO	cellulose oxidase	GH 61			ATEG_05416	AO090023000787	AFL2G_04596	ACLA_047220	NFIA_108320	Afu4g07850
CO	cellulose oxidase	GH 61	AN6428		ATEG_06077					
CO	cellulose oxidase	GH 61		56338	ATEG_05081	AO090103000087	AFL2G_12282		NFIA_055220	Afu6g09540
CO	cellulose oxidase	GH 61			ATEG_10194					
CO	cellulose oxidase	GH 61	AN7891	194765		AO090138000004	AFL2G_08699			
CO	cellulose oxidase	GH 61	AN3860							
CO	cellulose oxidase	GH 61			ATEG_08942					
CO	cellulose oxidase	GH 61			ATEG_01456					
CO	cellulose oxidase	GH 61						ACLA_017470		
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_00186	AO090103000088	AFL2G_12281	ACLA_071560	NFIA_087900	Afu2g12770
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN2632	55136	ATEG_10071	AO090701000885	AFL2G_06447	ACLA_061510	NFIA_033210	Afu2g00920
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN7908							
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_10379					
AGU	alpha glucuronidase	GH 67	AN9286	56619	ATEG_06085	AO090026000127	AFL2G_07114	ACLA_017270	NFIA_072510	Afu5g14380
AGU	alpha glucuronidase	GH 67			ATEG_09975					
XG CBH	xyloglucan active cellobiohydrolase	GH 74	AN1542						NFIA_096000	Afu8g02330
XG EGL	xyloglucan active endoglucanase	GH 74	AN5061	206333	ATEG_04708			ACLA_044310	NFIA_094990	Afu8g01490
RHA	alpha rhamnosidase	GH 78	AN8465	40264			AFL2G_05364			
RHA	alpha rhamnosidase	GH 78								
RHA	alpha rhamnosidase	GH 78	AN2631	176718	ATEG_05089	AO090001000105	AFL2G_07340		NFIA_004560	Afu3g02880
RHA	alpha rhamnosidase	GH 78	AN11954	51410		AO090003001016	AFL2G_02014			
RHA	alpha rhamnosidase	GH 78	AN10277	170172		AO090003001291	AFL2G_01780		NFIA_022970	Afu1g01660
RHA	alpha rhamnosidase	GH 78		42916	ATEG_03018	AO090009000471	AFL2G_10227		NFIA_018620	Afu1g06130
RHA	alpha rhamnosidase	GH 78	AN7151		ATEG_02922	AO090010000561	AFL2G_11713		NFIA_057930	Afu6g12030
RHA	alpha rhamnosidase	GH 78	AN3780							
RHA	alpha rhamnosidase	GH 78				AO090103000432	AFL2G_11972			
RHA	alpha rhamnosidase	GH 78				AO090113000149/A	AFL2G_08627		NFIA_057930	Afu6g14610
RHA	alpha rhamnosidase	GH 78	AN6929	131668		AO090005001416				
RHA	alpha rhamnosidase	GH 78							NFIA_060560	
RHA	alpha rhamnosidase	GH 78					AFL2G_10644			
RHA	alpha rhamnosidase	GH 78	AN12368							
RHA	alpha rhamnosidase	GH 78					AFL2G_03939			
RHA	alpha rhamnosidase	GH 78	AN10867	44977	ATEG_04706	AO090012000058	AFL2G_02993		NFIA_026130	Afu7g05040
UGH	unsaturated galacturonyl hydrolase	GH 88				AO090005000324	AFL2G_00324		NFIA_026210	Afu7g05090
UGH	unsaturated galacturonyl hydrolase	GH 88	AN3991	36414		AO090138000087	AFL2G08766		NFIA_023340	Afu1g01250
UGH	unsaturated galacturonyl hydrolase	GH 88	AN11078				AFL2G_06465			
UGH	unsaturated galacturonyl hydrolase	GH 88				AO090701000907				
UGH	unsaturated galacturonyl hydrolase	GH 88	AN4629							
ABX	exoarabinanase	GH 93	AN2060		ATEG_06045	AO090003001017	AFL2G_02013		NFIA_081320	Afu2G04570
ABX	exoarabinanase	GH 93				AO090011000141	AFL2G_04929			
ABX	exoarabinanase	GH 93	AN5231	49311	ATEG_07909	AO090012000101	AFL2G_03035	ACLA_008170	NFIA_027720	Afu7g06430
ABX	exoarabinanase	GH 93			ATEG_08029					
ABX	exoarabinanase	GH 93			ATEG_00891				NFIA_058060	Afu6g12120
AFC	alpha fucosidase	GH 95	AN8149	184037	ATEG_01560	AO090005000382	AFL2G_00373			
AFC	alpha fucosidase	GH 95				AO090005000512	AFL2G_00515			
AFC	alpha fucosidase	GH 95	AN6673	53702	ATEG_09768	AO090009000086	AFL2G_10565	ACLA_004070	NFIA_040960	Afu5g01190
AFC	alpha fucosidase	GH 95	AN10376		ATEG_04136				NFIA_064720	Afu3g12590
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN3196	41703	ATEG_07907	AO090001000174	AFL2G_07407	ACLA_056060	NFIA_030670	Afu4g02860
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105				AO090003000153	AFL2G_02808			
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN10505			AO090001000063	AFL2G_07308	ACLA_072870	NFIA_089840	Afu2g14630
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN9383	41877	ATEG_02892	AO090113000146	AFL2G_08624	ACLA_077810		
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN7828							
AGU	alpha glucuronidase	GH115				AO090005001415	AFL2G_01323			
AGU	alpha glucuronidase	GH115	AN9329		ATEG_09974	AO090010000038	AFL2G_11304	ACLA_006360	NFIA_025630	Afu7g04680
AGU	alpha glucuronidase	GH115				AO090001000267	AFL2G_07498			
AGU	alpha glucuronidase	GH115				AO090113000058				
AGU	alpha glucuronidase	GH 115			ATEG_04355					
PEL	pectin lyase	PL 1			ATEG_00950	AO090003001295	AFL2G_01776			
PEL	pectin lyase	PL 1		55212		AO090010000030	AFL2G_11297		NFIA_026110	Afu7g05030
PEL	pectin lyase	PL 1		45821						
PEL	pectin lyase	PL 1	AN4882	40837	ATEG_7577	AO090012000121	AFL2G_03052		NFIA_077100	Afu5g10170
PEL	pectin lyase	PL 1				AO090012000451	AFL2G_03359			
PEL	pectin lyase	PL 1				AO090103000463	AFL2G_11948			
PEL	pectin lyase	PL 1	AN10147	208760		AO090138000204	AFL2G_08823			
PEL	pectin lyase	PL 1						ACLA_094210	NFIA_033080	Afu2g00800
PEL	pectin lyase	PL 1	AN2331	41815		AO090010000504	AFL2G_11666	ACLA_013470	NFIA_076850	Afu5g10380
PEL	pectin lyase	PL 1	AN2569		ATEG_01216	AO090010000087	AFL2G_11352			
PEL	pectin lyase	PL 1							NFIA_023100	
PEL	pectin lyase	PL 1	AN9439							
PLY	pectate lyase	PL 1	AN0741		ATEG_08834	AO090011000673	AFL2G_05417			
PLY	pectate lyase	PL 1	AN5333		ATEG_05467	AO090102000072	AFL2G_09523		NFIA_060270	Afu6g14400
PLY	pectate lyase	PL 1	AN7646	45021	ATEG_08123	AO090701000321	AFL2G_05954		NFIA_033040	Afu2g00760
PLY	pectate lyase	PL 1	AN9367/AN9368			AO090011000030	AFL2G_04835			
PLY	pectate lyase	PL 3	AN6748		ATEG_06314	AO090005000472	AFL2G_00461		NFIA_027690	Afu7g06400
PLY	pectate lyase	PL 3	AN3337		ATEG_06285					
PLY	pectate lyase	PL 3	AN2542			AO090010000706	AFL2G_11846	ACLA_059210	NFIA_098670	Afu8g05910
PLY	pectate lyase	PL 3	AN6106		ATEG_08626	AO090038000502	AFL2G_07839		NFIA_023470	Afu1g01120
PLY	pectate lyase	PL 3	AN8453							
RGL	rhamnogalacturonan lyase	PL 4	AN7135	210947	ATEG_02193	AO090011000349	AFL2G_05136	ACLA_054660	NFIA_029620	Afu4g03780
RGL	rhamnogalacturonan lyase	PL 4	AN6395	47780	ATEG_10327	AO090012000147/A	AFL2G_03075	ACLA_018320	NFIA_008140	Afu1g17230
RGL	rhamnogalacturonan lyase	PL 4	AN12097			AO090113000057				
RGL	rhamnogalacturonan lyase	PL 4			ATEG_08610	AO090138000119	AFL2G_08794		NFIA_094270	Afu8g00820
RGL	rhamnogalacturonan lyase	PL 4	AN3950							
PLY	pectate lyase	PL 9	AN2537			AO090038000131	AFL2G_08953		NFIA_062360	Afu3g14890
PLY	pectate lyase	PL 9			ATEG_03526					
RGL	rhamnogalacturonan lyase	PL 11	AN2543							

Supplemental Table 6B: Detection of proteins in cultures grown on sugar beet pulp sorted by CAZy family. The colour codes indicate which percentage of the total number of detected peptides was of the specific protein. (Continued on next 4 pages).

			0-0.1%	0.1-0.5%	0.5-1%	1-2%	2-5%	5-10%	>10%	
enzyme code	**function**	**CAZY**	***A. nidulans***	***A. niger*** **ATCC 1015**	***A. terreus***	***A. oryzae***	***A. flavus***	***A. clavatus***	***A. fischeri***	***A. fumigatus*** **Af293**
FAE SF7	feruloyl esterase	faeA		51662	ATEG_08907	AO090001000207	AFL2G_07436			
FAE SF1	feruloyl esterase	SF1	AN1772	51478	ATEG_06663	AO090001000582	AFL2G_09228	ACLA_083360	NFIA_054700	Afu6g09040
FAE SF1	feruloyl esterase	SF1			ATEG_02212	AO090001000066	AFL2G_07310		NFIA_047590	Afu6g00450
FAE SF3	feruloyl esterase	SF3				AO090102000013				
FAE SF3	feruloyl esterase	SF3			ATEG_02415					
FAE SF4	feruloyl esterase	SF4		190471		AO090010000573	AFL2G_11725			
FAE SF4	feruloyl esterase	SF4		43194						
AXE	acetyl xylan esterase	CE 1	AN6093	211544	ATEG_09843	AO090011000745	AFL2G_05471	ACLA_081220	NFIA_099230	Afu8g06570
AXE	acetyl xylan esterase	CE 1	AN8320							Afu1g17510
AXE	acetyl xylan esterase	CE 1							NFIA_084920	
FAE SF5	feruloyl esterase	CE 1				AO090005000945	AFL2G_00922	ACLA_065130	NFIA_115130	
FAE SF5	feruloyl esterase	CE 1	AN5267	43785	ATEG_08112	AO090023000158	AFL2G_04047	ACLA_055050	NFIA_089720	Afu2G14530
FAE SF5	feruloyl esterase	CE 1			ATEG_06644	AO090701000884	AFL2G_06446	ACLA_081520		
FAE SF5	feruloyl esterase	CE 1								Afu2g09440
FAE SF5	feruloyl esterase	CE 1			ATEG_01914					
FAE SF5	feruloyl esterase	CE 1						ACLA_017480		
?		CE 1	AN8782	53315	ATEG_06438	AO090005000277	AFL2G_00281	ACLA_012870	NFIA_077450	Afu5g09860
PME	pectin methyl esterase	CE 8				AO090003001514	AFL2G_01578			
PME	pectin methyl esterase	CE 8		44585		AO090012000749	AFL2G_03618	ACLA_044240	NFIA_095020	Afu8g01520
PME	pectin methyl esterase	CE 8	AN4660	174365	ATEG_01704	AO090020000404	AFL2G_10918	ACLA_035610	NFIA_069500	Afu3g07850
PME	pectin methyl esterase	CE 8	AN3390	214857		AO090102000010	AFL2G_09467		NFIA_099600	Afu8g06880
PME	pectin methyl esterase	CE 8	AN7966							
PME	pectin methyl esterase	CE 8				AO090113000039			NFIA_100100	Afu8g07250
PME	pectin methyl esterase	CE 8						ACLA_059970		
PME	pectin methyl esterase	CE 8					AFL2G_08528			
PME	pectin methyl esterase	CE 8						ACLA_059980		
RGAE	rhamnogalacturonan acetyl esterase	CE 12				AO090003001268	AFL2G_01797			
RGAE	rhamnogalacturonan acetyl esterase	CE 12							NFIA_099110	Afu8g06480
RGAE	rhamnogalacturonan acetyl esterase	CE 12		51400	ATEG_10016	AO090102000092	AFL2G_09543			
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2834			AO090113000155	AFL2G_08631		NFIA_092640	Afu2g17250
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2528	189254	ATEG_03511	AO090701000556	AFL2G_06177	ALCL_041970	NFIA_062750	Afu3g14510
GE	glucuronoyl esterase	CE15			ATEG_00945			ACLA_087520	NFIA_060260	Afu6g14390
BGL	beta glucosidase	GH 1		140573						
BGL	beta glucosidase	GH 1	AN10124	213437	ATEG_00687	AO090003000497	AFL2G_02496	ACLA_020660	NFIA_010690	Afu1g14710
BGL	beta glucosidase	GH 1						ACLA_019180	NFIA_009040	Afu1g16400
BGL	beta glucosidase	GH 1				AO090113000148	AFL2G_08626		NFIA_060550	Afu6g14600
BGL	beta glucosidase	GH 1	AN9183	131747	ATEG_02657	AO090120000075	AFL2G_08111		NFIA_099670	Afu8g06970
BGL	beta glucosidase	GH 1	AN10375		ATEG_04135				NFIA_064710	Afu3g12600
BGL	beta glucosidase	GH 1						ACLA_064280		
BGL	beta glucosidase	GH 1						ACLA_040420		
GUS	beta glucuronidase	GH 2	AN3200	189620	ATEG_09745		AFL2G_05262			
GUS	beta glucuronidase	GH 2	AN2395	52111	ATEG_01031	AO090023000053	AFL2G_03956		NFIA_089700	Afu2g14520
LAC	beta galactosidase	GH 2			ATEG_10255					
LAC	beta galactosidase	GH 2			ATEG_04784					
LAC	beta galactosidase	GH 2	AN2463		ATEG_00712					
GUS	beta glucuronidase	GH 2	AN5361	46827		AO090038000009	AFL2G_08828			
LAC	beta galactosidase	GH 2	AN3201		ATEG_10243	AO090012000389	AFL2G_03296		NFIA_072410	Afu5g14550
LAC	beta galactosidase	GH 2	AN6388							
LAC	beta galactosidase	GH 2	AN1107						NFIA_072890	Afu5g14090
MND	beta mannosidase	GH 2	AN1742	138876	ATEG_06636	AO090001000556	AFL2G_09201	ACLA_083570	NFIA_054490	Afu6g08840
MND	beta mannosidase	GH 2	AN11680	172587	ATEG_07339	AO090003001410	AFL2G_01666	ACLA_078510	NFIA_045250	Afu4g00390
MND	beta mannosidase	GH 2				AO090005000740	AFL2G_00728			
MND	beta mannosidase	GH 2	AN3368	212893	ATEG_08684	AO090010000208	AFL2G_11464	ACLA_066240	NFIA_114030	Afu7g01320
MND	beta mannosidase	GH 2			ATEG_09890					
BGL	beta glucosidase	GH 3			ATEG_10274					
BGL	beta glucosidase	GH 3		139037						
BGL	beta glucosidase	GH 3								
BGL	beta glucosidase	GH 3		39613						
BGL	beta glucosidase	GH 3	AN7915	38077		AO090001000266	AFL2G_07497			
BGL	beta glucosidase	GH 3	AN10482	208871	ATEG_06617	AO090001000544	AFL2G_09187	ACLA_083710	NFIA_054350	Afu6g08700
BGL	beta glucosidase	GH 3								
BGL	beta glucosidase	GH 3		210981	ATEG_02806	AO090003001511	ALF2G_01582		NFIA_095760	Afu8g02100
BGL	beta glucosidase	GH 3			ATEG_02724	AO090005000337	AFL2G_00334			
BGL	beta glucosidase	GH 3	AN4102	58782	ATEG_03047	AO090090000356	AFL2G_10322	ACLA_028810	NFIA_018950	Afu1g05770
BGL	beta glucosidase	GH 3	AN6652	182309	ATEG_07121	AO090090000554	AFL2G_10164	ACLA_096980	NFIA_050080	Afu6g03570
BGL	beta glucosidase	GH 3		44520		AO090010000034	AFL2G_11300			
BGL	beta glucosidase	GH 3				AO090012000003	AFL2G_02949			
BGL	beta glucosidase	GH 3	AN7396	179265	ATEG_10320	AO090012000135	AFL2G_03066		NFIA_007920	Afu1g17410
BGL	beta glucosidase	GH 3	AN1804			AO090026000123	AFL2G_07119			
BGL	beta glucosidase	GH 3				AO090038000223	AFL2G_09023			
BGL	beta glucosidase	GH 3	AN5976	181816	ATEG_02713	AO090038000425	AFL2G_07763		NFIA_100430	
BGL	beta glucosidase	GH 3				AO090103000127	AFL2G_12245			
BGL	beta glucosidase	GH 3		176601		AO090166000048	AFL2G_09452		NFIA_112660	
BGL	beta glucosidase	GH 3	AN3903		ATEG_04069	AO090166000090	AFL2G_09413	ACLA_087610	NFIA_060370	Afu6g14490
BGL	beta glucosidase	GH 3	AN2227	129891	ATEG_09329	AO090701000244	AFL2G_05886	ACLA_010450	NFIA_080070	Afu5g07190
BGL	beta glucosidase	GH 3			ATEG_07931					
BGL	beta glucosidase	GH 3	AN0712							
BGL	beta glucosidase	GH 3	AN2612							Afu7g00240

Supplemental Table 6B Cont: Detection of proteins in cultures grown on sugar beet pulp sorted by CAZy family. (Continued on next 3 pages).

BGL	beta glucosidase	GH 3	AN2828		ATEG_07419	AO090701000841	AFL2G_06408	ACLA_007810	NFIA_027390	Afu7g06140
BGL	beta glucosidase	GH 3	AN3949							
BGL	beta glucosidase	GH 3	AN7865						NFIA_057590	Afu6g11910
BGL	beta glucosidase	GH 3		129779	ATEG_00157					
BGL	beta glucosidase	GH 3							NFIA_000750	Afu3g00230
BGL	beta glucosidase	GH 3							NFIA_098520	
BGL	beta glucosidase	GH 3							NFIA_057910	
BGL/BXL	beta glucosidase/beta xylosidase	GH 3	AN0479							
BXL	beta xylosidase	GH 3	AN2359	205670	ATEG_05106	AO090005000986	AFL2G_00957			Afu1g16920
BXL	beta xylosidase	GH 3			ATEG_07383	AO090011000140	AFL2G_04928			
BXL	beta xylosidase	GH 3						ACLA_018590		
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN8401		ATEG_09052	AO090103000120	AFL2G_12252	ACLA_062400	NFIA_003180	Afu3g02090
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN2217	50997	ATEG_09314	AO090701000274	AFL2G_05912	ACLA_010340	NFIA_080180	Afu5g07080
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3			ATEG_08027					
?		GH 3	AN1416	45461	ATEG_00018	AO090103000019	AFL2G_12338	ACLA_057390	NFIA_096770	Afu8g04060
?		GH 3								Afu1g00540
?		GH 3	AN3360	37673	ATEG_04963			ACLA_052760	NFIA_102600	Afu4g13770
?		GH 3		120104	ATEG_04729	AO090003000741	AFL2G_02272	ACLA_039660	NFIA_065550	Afu3g11780
?		GH 3					AFL2G_08686			
?		GH 3	AN2599							
EGL	endoglucanase	GH 5	AN3013	210716	ATEG_01592	AO090005001389	AFL2G_01298	ACLA_036770	NFIA_068300	Afu3g08820
EGL	endoglucanase	GH 5	AN1285	205580	ATEG_05002			ACLA_085250	NFIA_057290	Afu6g11600
EGL	endoglucanase	GH 5							NFIA_095570	
EGL	endoglucanase	GH 5	AN8068		ATEG_09802	AO090003001341+A	AFL2G_01726		NFIA_040280	Afu5g01830
EGL	endoglucanase	GH 5	AN5214	209376		AO090005001553	AFL2G_01447	ACLA_081650	NFIA_053150	Afu6g07480
EGL	endoglucanase	GH 5		214608	ATEG_04390	AO090011000715	AFL2G_05447	ACLA_081310	NFIA_085010	Afu2g09520
EGL	endoglucanase	GH 5			ATEG_05003					
GLN	exo 1,6 galactanase	GH 5	AN9166	194447	ATEG_10242	AO090012000046	AFL2G_02982		NFIA_072400	Afu5g14560
EXG	exo 1,3 galactanase	GH 5			ATEG_08371				NFIA_056040	
EXG	exo 1,3 galactanase	GH 5	AN8947	175759	ATEG_07719					
EXG	exo 1,3 galactanase	GH 5	AN7533	123981	ATEG_06686	AO090001000604	AFL2G_09249	ACLA_083150	NFIA_054930	Afu6g09250
EXG	exo 1,3 galactanase	GH 5	AN4052	202490	ATEG_03849	AO090003000990	AFL2G_02039	ACLA_031040	NFIA_021060	Afu1g03600
EXG	exo 1,3 galactanase	GH 5			ATEG_06369	AO090005000423	AFL2G_00412	ACLA_007330	NFIA_026860	Afu7g05610
EXG	exo 1,3 galactanase	GH 5			ATEG_03062	AO090009000373	AFL2G_10308			
EXG	exo 1,3 galactanase	GH 5	AN1332	52811		AO090012000917	AFL2G_03782		NFIA_113230	
EXG	beta 1,6 glucanase	GH 5	AN3777		ATEG_09844	AO090011000757	AFL2G_05484		NFIA_084850	Afu2g09350
MAN	endomannanase	GH 5	AN3358	50378	ATEG_08654	AO090010000122	AFL2G_11381	ACLA_066420	NFIA_113780	Afu7g01070
MAN	endomannanase	GH 5			ATEG_10292	AO090012000006	AFL2G_02951		NFIA_041960	Afu5g00550
MAN	endomannanase	GH 5	AN7639			AO090038000444	AFL2G_07781	ACLA_044470	NFIA_099770	Afu8g07030
MAN	endomannanase	GH 5			ATEG_02669					
MAN	endomannanase	GH 5	AN2709							
MAN	endomannanase	GH 5	AN3297							
MAN	endomannanase	GH 5	AN6427		ATEG_09991					
MAN	endomannanase	GH 5	AN9276							
MAN	endomannanase	GH 5			ATEG_01374					
		GH5			ATEG_03677					
CBH	cellobiohydrolase	GH 6	AN5282	54490	ATEG_07493			ACLA_062560	NFIA_002990	Afu3g01910
CBH	cellobiohydrolase	GH 6	AN1273	133986	ATEG_00193	AO090038000439	AFL2G_07776	ACLA_025560	NFIA_015680	
CBH	cellobiohydrolase	GH 7	AN0494	51773	ATEG_05002	AO090001000348	AFL2G_07571	ACLA_085260	NFIA_057300	Afu6g11610
CBH	cellobiohydrolase	GH 7							NFIA_095570	
CBH	cellobiohydrolase	GH 7	AN5176	53159	ATEG_03727	AO090012000941	AFL2G_03805	ACLA_088870	NFIA_052720	Afu6g07070
EGL	endoglucanase	GH 7	AN3418		ATEG_08700	AO090010000314	ALF2G_11497	ACLA_066030	NFIA_114250	Afu7g01540
EGL	endoglucanase	GH 7			ATEG_08705			ACLA_098940	NFIA_047960	Afu6g01800
XLN	endoxylanase	GH 10				AO090001000208	AFL2G_07437			
XLN	endoxylanase	GH 10	AN7401		ATEG_03410	AO090103000326	AFL2G_12071	ACLA_086910	NFIA_059570	Afu6g13610
XLN	endoxylanase	GH 10	AN1818	57436	ATEG_00809	AO090701000887	AFL2G_06449	ACLA_048770	NFIA_106540	Afu4g09480
XLN	endoxylanase	GH 10			ATEG_08906	AO090103000423	AFL2G_11983			
XLN	endoxylanase	GH 10	AN2356	50977					NFIA_057510	Afu6g11820
XLN	endoxylanase	GH 10			ATEG_07190				NFIA_061880	Afu3g15210
XLN	endoxylanase	GH 11	AN3613							
XLN	endoxylanase	GH 11								
XLN	endoxylanase	GH 11		171269	ATEG_04943	AO090001000111	AFL2G_07347	ACLA_085410	NFIA_058160	Afu6g12210
XLN	endoxylanase	GH 11		183088		AO090026000103	AFL2G_07138		NFIA_055240	
XLN	endoxylanase	GH 11				AO090103000141	AFL2G_12233			
XLN	endoxylanase	GH 11	AN9365	52071	ATEG_07461	AO090120000026	AFL2G_08066	ACLA_063140	NFIA_000850	Afu3g00320
XLN	endoxylanase	GH 11							NFIA_001010	Afu3g00470
XLN	endoxylanase	GH 11						ACLA_064270		
XG EGL	xyloglucan active endoglucanase	GH 12	AN0452	52011	ATEG_03755	AO090003000905	AFL2G_02120	ACLA_029940	NFIA_020020	Afu1g04730
XG EGL	xyloglucan active endoglucanase	GH 12		211053	ATEG_07420	AO090026000102	AFL2G_07140	ACLA_007820	NFIA_027400	Afu7g06150
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_05519					
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_09894					
XG EGL	xyloglucan active endoglucanase	GH 12				AO090038000175	AFL2G_08984		NFIA_002040	
XG EGL	xyloglucan active endoglucanase	GH 12				AO090701000185	AFL2G_05831	ACLA_060930	NFIA_005810	Afu3g03610
XG EGL	xyloglucan active endoglucanase	GH 12		191511					NFIA_100330	
XG EGL	xyloglucan active endoglucanase	GH 12					AFL2G_04893			
XG EGL	xyloglucan active endoglucanase	GH 12								Afu3g01160
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_07466					
AGT	4 alpha glucanotransferase	GH 13		57002						
?		GH 13		182162						
GBA	glycogen debranching enzyme	GH 13	AN2314	211162	ATEG_07608	AO090010000483	AFL2G_11646	ACLA_013550	NFIA_076670	Afu5g10540
AGD	alpha glucosidase	GH 13		50927						
AGD	alpha glucosidase	GH 13	AN4843	52452	ATEG_01729	AO090020000176	AFL2G_11111	ACLA_035550	NFIA_069880	Afu3g07380
AGD	alpha glucosidase	GH 13				AO090026000034	AFL2G_07207			
AGD	alpha glucosidase	GH 13				AO090038000234	AFL2G_09036	ACLA_008120	NFIA_027660	Afu7g06380
AGD	alpha glucosidase	GH 13	AN10420		ATEG_01029	AO090103000129	AFL2G_12243	ACLA_059930	NFIA_099830	Afu8g07070
AGD	alpha glucosidase	GH 13				AO090103000378	AFL2G_12021			

Supplemental Table 6B Cont: Detection of proteins in cultures grown on sugar beet pulp sorted by CAZy family. (Continued on next 2 pages).

AGD	alpha glucosidase	GH 13					AFL2G_08694			
AGD	alpha glucosidase	GH 13						ACLA_070570	NFIA_086860	Afu2g11620
AGD	alpha glucosidase	GH 13						ACLA_070560	NFIA_086850	Afu2g11610
AGS	alpha glucan synthase	GH 13		212915						
AGS	alpha glucan synthase	GH 13		40878						
AGS	alpha glucan synthase	GH 13	AN3307	54378	ATEG_03622	AO090003001500	AFL2G_01593	ACLA_063450	NFIA_001720	Afu3g00910
AGS	alpha glucan synthase	GH 13		55204	ATEG_10371	AO090010000106	AFL2G_11365	ACLA_042430	NFIA_009960	Afu1g15440
AGS	alpha glucan synthase	GH 13	AN5885		ATEG_01449	AO090026000523	AFL2G_06761	ACLA_070240	NFIA_086550	Afu2g11270
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13				AO090120000196		ACLA_094070	NFIA_032970	Afu2g00710
AMY	alpha amylase	GH 13	AN3402	47911	ATEG_10103	AO090023000944		ACLA_052920		
AMY	alpha amylase	GH 13	AN2018	45304	ATEG_08279	AO090003001210	AFL2G_01841	ACLA_049350	NFIA_105920	Afu4g10130
AGT	4 alpha glucanotransferase	GH 13	AN3308	188489	ATEG_03623	AO090003001498	AFL2G_01594	ACLA_063440	NFIA_001710	Afu3g00900
AMY	alpha amylase	GH 13				AO090120000263	AFL2G_08276			
AMY	alpha amylase	GH 13								
AMY	alpha amylase	GH 13	AN4507							
AMY	alpha amylase	GH 13	AN3388		ATEG_02515					
AMY	alpha amylase	GH 13	AN6324					ACLA_072270	NFIA_088650	Afu2g13460
AMY	alpha amylase	GH 13		122069	ATEG_03624	AO090003001497	AFL2G_01595			
AMY	alpha amylase	GH 13	AN10060	46621	ATEG_04879	AO090005000884	AFL2G_00860	ACLA_032290	NFIA_022510	Afu1g02140
AMY	alpha amylase	GH 13	AN3309	46290	ATEG_00838	AO090005001193		ACLA_020060	NFIA_010270	Afu1g15150
AMY	alpha amylase	GH 13			ATEG_00724					
AMY	alpha amylase	GH 13					AFL2G_01127			
AMY	alpha amylase	GH 13						ACLA_091300	NFIA_035590	Afu2g03230
AMY	alpha amylase	GH 13							NFIA_035580	
GLA	glucoamylase	GH 15	AN7402					ACLA_049360	NFIA_105910	
GLA	glucoamylase	GH 15								Afu4g10140
GLA	glucoamylase	GH 15			ATEG_05980	AO090003000321	AFL2G_02658	ACLA_089470		
GLA	glucoamylase	GH 15	AN11143	213597	ATEG_04375	AO090010000745	AFL2G_11885	ACLA_094080	NFIA_032960	Afu2g00690
GLA	glucoamylase	GH 15		189911		AO090138000105	AFL2G_08782	ACLA_044430	NFIA_094880	Afu8g01390
GLA	glucoamylase	GH 15						ACLA_082570	NFIA_053760	Afu6g08080
GLA	glucoamylase	GH 15						ACLA_078620	NFIA_001210	Afu3g00610
MAN	endomannanase	GH 26	AN3336	40875		AO090011000055	AFL2G_04859			
MAN	endomannanase	GH 26	AN3326							
MAN	endomannanase	GH 26	AN7413							
AGL	alpha galactosidase	GH 27			ATEG_01905				NFIA_048850	Afu8g02560
AGL	alpha galactosidase	GH 27		37736	ATEG_04382			ACLA_016820	NFIA_073100	Afu5g13830
AGL	alpha galactosidase	GH 27	AN0022	172232	ATEG_03427	AO090005000217	AFL2G_00225		NFIA_023390	Afu1g01200
AGL	alpha galactosidase	GH 27	AN7624	207264	ATEG_09830	AO090003001305	AFL2G_01765	ACLA_003130	NFIA_039990	Afu5g02130
AGL	alpha galactosidase	GH 27		39180						
AGL	alpha galactosidase	GH 27	AN7152	41606	ATEG_02160	AO090023000151	AFL2G_04039		NFIA_029860	Afu4g03580
AGL	alpha galactosidase	GH 27						ACLA_097900		
PGA	endopolygalacturonase	GH 28	AN8327	46255					NFIA_008150	Afu1g17220
PGA	endopolygalacturonase	GH 28		43957						
PGA	endopolygalacturonase	GH 28		182156						
PGA	endopolygalacturonase	GH 28		214598						
PGA	endopolygalacturonase	GH 28	AN6656	50161		AO090005000186	AFL2G_00201		NFIA_099410	
PGA	endopolygalacturonase	GH 28		172944	ATEG_01601	AO090005001400	AFL2G_01310	ACLA_036670	NFIA_068440	Afu3g08680
PGA	endopolygalacturonase	GH 28	AN4372	141677	ATEG_04991	AO090023000401	AFL2G_04252	ACLA_052860	NFIA_102450	Afu4g13920
PGA	endopolygalacturonase	GH 28		52219		AO090023000161	AFL2G_04049		NFIA_095620	Afu8g01970
PGA	endopolygalacturonase	GH 28				AO090138000086				
PGA	endopolygalacturonase	GH 28								Afu8g06730
PGA	endopolygalacturonase	GH 28			ATEG_07748					
PGA	endopolygalacturonase	GH 28					AFL2G_08764		NFIA_023290	Afu1g01320
PGX	exopolygalacturonase	GH 28		178172						
PGX	exopolygalacturonase	GH 28	AN8891	191158	ATEG_10357	AO090010000753	AFL2G_11892	ACLA_043100	NFIA_096340	Afu8g02630
PGX	exopolygalacturonase	GH 28	AN8761	42184	ATEG_07152	AO090026000784	AFL2G_08533		NFIA_049320	Afu6g02980
PGX	exopolygalacturonase	GH 28	AN9045							
RGX	exorhamnogalacturonase	GH 28		172236		AO090001000133	AFL2G_07371			
RGX	exorhamnogalacturonase	GH 28		42917		AO090009000470	AFL2G_10228		NFIA_018590	Afu1g06140
RGX	exorhamnogalacturonase	GH 28	AN10274	194461	ATEG_09025	AO090102000139	AFL2G_09582		NFIA_027700	Afu7g06410
RGX	exorhamnogalacturonase	GH 28			ATEG_06408	AO090113000199	AFL2G_08671			
RGX	exorhamnogalacturonase	GH 28				AO090138000066	AFL2G_03102			
RGX	exorhamnogalacturonase	GH 28	AN11626							
RGX	exorhamnogalacturonase	GH 28				AO090138000067	AFL2G_08746			
RHG	endorhamnogalacturonase	GH 28		178393						
RHG	endorhamnogalacturonase	GH 28		180922		AO090003000524	AFL2G_02475			Afu4g00100
RHG	endorhamnogalacturonase	GH 28				AO090005000067	AFL2G_00087			
RHG	endorhamnogalacturonase	GH 28		211163	ATEG_07607	AO090010000484	AFL2G_11646			
RHG	endorhamnogalacturonase	GH 28		123651		AO090026000252	AFL2G_06999			
RHG	endorhamnogalacturonase	GH 28	AN9134	189722		AO090038000552	AFL2G_07883			
RHG	endorhamnogalacturonase	GH 28		39337		AO090124000009	AFL2G_08037			
RHG	endorhamnogalacturonase	GH 28							NFIA_076680	Afu5g10530
XGH	xylogalacturonase	GH 28				AO090026000120	AFL2G_07122			
XGH	xylogalacturonase	GH 28	AN3389	46065		AO090102000011	AFL2G_09468		NFIA_099610	Afu8g06890
PGX	exopolygalacturonase	GH 28							NFIA_100120	Afu8g07265
AFC	alpha fucosidase	GH 29		44822	ATEG_08111					
AFC	alpha fucosidase	GH 29			ATEG_05691					
AGD	alpha glucosidase	GH 31	AN2017	214233	ATEG_00723	AO090003001209	AFL2G_01842	ACLA_049370	NFIA_105900	Afu4g10150
AGD	alpha glucosidase	GH 31	AN0280	55419	ATEG_02528	AO090005000767	AFL2G_00750	ACLA_031260	NFIA_021450	Afu1g03140
AGD	alpha glucosidase	GH 31	AN0941	119858	ATEG_05177	AO090005001084	AFL2G_01038	ACLA_019300	NFIA_009180	Afu1g16250
AGD	alpha glucosidase	GH 31	AN3504	49940	ATEG_02966	AO090023000288	AFL2G_04152		NFIA_018130	Afu1g06560
AGD	alpha glucosidase	GH 31	AN10935			AO090026000111	AFL2G_07131			
AGD	alpha glucosidase	GH 31	AN8953			AO090038000471	AFL2G_07812			
AGD	alpha glucosidase	GH 31	AN11054	50055	ATEG_08065	AO090102000559	AFL2G_09934	ACLA_001670	NFIA_038610	Afu5g03500

Supplemental Table 6B Cont: Detection of proteins in cultures grown on sugar beet pulp sorted by CAZy family. (Continued on next page).

AGD	alpha glucosidase	GH 31	AN7120							
AGD	alpha glucosidase	GH 31	AN7345							
AGD	alpha glucosidase	GH 31			ATEG_08278					
AGD	alpha glucosidase	GH 31			ATEG_08472					
AGD	alpha glucosidase	GH 31							NFIA_060440	
AXL	alpha xylosidase	GH 31	AN7505		ATEG_06730	AO090001000649	AFL2G_09290	ACLA_089300	NFIA_082140	Afu2g05400
AXL	alpha xylosidase	GH 31		40261		AO090701000558	AFL2G_06180			
AXL	alpha xylosidase	GH 31		43342	ATEG_08390	AO090701000639	AFL2G_06239		NFIA_032680	
AXL	alpha xylosidase	GH 31			ATEG_00207					
AXL	alpha xylosidase	GH 31			ATEG_05169					
INU	endo inulinase	GH 32		52928						Afu5g00530
INU	endo inulinase	GH 32								Afu5g00480
INU	endo inulinase	GH 32			ATEG_07356					
INX	exo inulinase	GH 32								
INX	exo inulinase	GH 32	AN11778	56664	ATEG_08155	AO090701000400	AFL2G_06028	ACLA_094550	NFIA_033540	Afu2g01240
SUC	invertase/beta fructofuranosidase	GH 32	AN3837		ATEG_05860	AO090701000038	AFL2G_05693		NFIA_051560	Afu6g05000
SUC	invertase/beta fructofuranosidase	GH 32		198063	ATEG_07479	AO090020000640	AFL2G_10707			
SUC	invertase/beta fructofuranosidase	GH 32				AO090103000043	AFL2G_12317			
SUC	invertase/beta fructofuranosidase	GH 32		176039						
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_04996					
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_10253					
LAC	beta galactosidase	GH 35	AN0194							
LAC	beta galactosidase	GH 35		180727						
LAC	beta galactosidase	GH 35				AO090001000259	AFL2G_07489			
LAC	beta galactosidase	GH 35		177434	ATEG_07446	AO090003000042	AFL2G_02901	ACLA_088440	NFIA_052310	Afu6g06660
LAC	beta galactosidase	GH 35	AN0756	51764	ATEG_00618	AO090012000445	AFL2G_03352	ACLA_021260	NFIA_011250	Afu1g14170
LAC	beta galactosidase	GH 35				AO090012000746				
LAC	beta galactosidase	GH 35	AN2804		ATEG_08255	AO090023000076	AFL2G_03977		NFIA_041490	Afu5g00670
LAC	beta galactosidase	GH 35	AN0980	46429	ATEG_05131	AO090120000158	AFL2G_08182		NFIA_008690	Afu1g18700
LAC	beta galactosidase	GH 35		41910			AFL2G_03616		NFIA_000910	Afu3g00380
LAC	beta galactosidase	GH 35				AO090701000770				
LAC	beta galactosidase	GH 35						ACLA_018830		
LAC	beta galactosidase	GH 35					AFL2G_06343			
LAC	beta galactosidase	GH 35					AFL2G_12387			
AGL	alpha galactosidase	GH 36	AN8138	212736	ATEG_07929	AO090010000684	AFL2G_11824	ACLA_044620	NFIA_094580	Afu8g01130
AGL	alpha galactosidase	GH 36	AN8392		ATEG_07935	AO090011000063	AFL2G_04865	ACLA_044560	NFIA_094540	Afu8g01100
AGL	alpha galactosidase	GH 36	AN3874	190816	ATEG_05379	AO090023000743	AFL2G_04555	ACLA_047600	NFIA_107920	Afu4g08250
AGL	alpha galactosidase	GH 36	AN9035		ATEG_02312					
ABF	alpha arabinofuranosidase	GH 43	AN8472			AO090012000356	AFL2G_03265		NFIA_032900	Afu2g00650
ABF	alpha arabinofuranosidase	GH 43	AN7781		ATEG_08386	AO090005000065	AFL2G_00086			
ABF	alpha arabinofuranosidase	GH 43	AN2533			AO090701000838	AFL2G_06400		NFIA_002750	Afu3g01660
ABN	endoarabinanase	GH 43		128077	ATEG_05780				NFIA_078690	Afu5g08590
ABN	endoarabinanase	GH 43		182100	ATEG_07817	AO090026000804	AFL2G_06513	ACLA_072730	NFIA_089980	Afu2g14750
ABN	endoarabinanase	GH 43	AN3044	184195	ATEG_01562	AO090005001320	AFL2G_01244	ACLA_037080	NFIA_068020	Afu3g09140
ABN	endoarabinanase	GH 43						ACLA_018180	NFIA_008030	Afu1g17320
ABN	endoarabinanase	GH 43	AN6352	134398	ATEG_01407	AO090023000165	AFL2G_04053	ACLA_073460	NFIA_089320	Afu2g14150
ABN*	endoarabinanase	GH 43	AN8007	197735		AO090138000055	AFL2G_08737	ACLA_098980	NFIA_047920	Afu6g00770
ABN	endoarabinanase	GH 43	AN2534	203143	ATEG_03520	AO090701000481	AFL2G_06106	ACLA_042100	NFIA_062660	Afu3g14620
ABN	endoarabinanase	GH 43								
ABN	endoarabinanase	GH 43				AO090005000064	AFL2G_00085			
ABN	endoarabinanase	GH 43			ATEG_03688					
ABN	endoarabinanase	GH 43			ATEG_07787					
BXL	beta xylosidase	GH 43	AN1870	179682	ATEG_06059	AO090003000239	AFL2G_02739	ACLA_090090	NFIA_081240	Afu2g04480
BXL	beta xylosidase	GH 43	AN6751	174379	ATEG_06306	AO090005000476	AFL2G_00456		NFIA_027350	Afu7g06110
BXL	beta xylosidase	GH 43			ATEG_05083	AO090103000268	AFL2G_12122	ACLA_087680	NFIA_060490	Afu6g14550
BXL	beta xylosidase	GH 43	AN10199		ATEG_00093	AO090005000698	AFL2G_00689/0	ACLA_058100/(	NFIA_097490/0	Afu8g04710/04720
BXL	beta xylosidase	GH 43	AN10919		ATEG_10072	AO090010000029	AFL2G_11296		NFIA_093350	
BXL	beta xylosidase	GH 43			ATEG_07784	AO090010000494	AFL2G_11656		NFIA_023360	Afu1g01230
BXL	beta xylosidase	GH 43	AN7313	122978	ATEG_01188	AO090010000562	AFL2G_11714	ACLA_043260	NFIA_096240	Afu8g02510
BXL	beta xylosidase	GH 43	AN2664		ATEG_06107	AO090012000350	AFL2G_03257		NFIA_072780	Afu5g14190
BXL	beta xylosidase	GH 43	AN8477		ATEG_01292		AFL2G_06974	ACLA_018160		
BXL	beta xylosidase	GH 43				AO090113000059				
BXL	beta xylosidase	GH 43			ATEG_10193					
BXL	beta xylosidase	GH 43	AN7275		ATEG_06643	AO090701000886	AFL2G_06448	ACLA_078600	NFIA_033220	Afu2g00930
BXL	beta xylosidase	GH 43	AN7864	47677	ATEG_01292	AO090102000331	AFL2G_09750	ACLA_072000	NFIA_088370	Afu2g13190
BXL	beta xylosidase	GH 43			ATEG_07517					
BXL	beta xylosidase	GH 43	AN2633							
BXL	beta xylosidase	GH 43	AN1043							
BXL	beta xylosidase	GH 43					AFL2G_08528			
?	?	GH 43	AN1477							
EGL	endoglucanase	GH 45	AN6786						NFIA_028020	Afu7g06740
ABF	alpha arabinofuranosidase	GH 51	AN1277	38549	ATEG_00198		AFL2G_07796	ACLA_025610	NFIA_015730	Afu1g09900
ABF	alpha arabinofuranosidase	GH 51		50979	ATEG_03540					
ABF	alpha arabinofuranosidase	GH 51				AO090020000712	AFL2G_10643			
ABF	alpha arabinofuranosidase	GH 51	AN2541	298387	ATEG_02882	AO090124000023	AFL2G_08019	ACLA_099110		
ABF	alpha arabinofuranosidase	GH 51	AN9439	131891	ATEG_07868	AO090012000298	AFL2G_03217	ACLA_074120	NFIA_090410	Afu2g15160
GAL	endogalactanase	GH 53								
GAL	endogalactanase	GH 53	AN5727	187227	ATEG_02927	AO090001000492	AFL2G_09135		NFIA_017780	Afu1g08910
GAL	endogalactanase	GH 53					AFL2G_07488			
ABF	alpha arabinofuranosidase	GH 54	AN1571	200605	ATEG_07939	AO090023000001	AFL2G_03901	ACLA_066470	NFIA_060830	Afu6g14620
CO	cellulose oxidase	GH 61	AN10419							
CO	cellulose oxidase	GH 61		211595	ATEG_04210			ACLA_060890	NFIA_006140	Afu3g03870
CO	cellulose oxidase	GH 61	AN1041	52688	ATEG_00448	AO090001000221	AFL2G_07454	ACLA_022980	NFIA_012990	Afu1g12560
CO	cellulose oxidase	GH 61	AN1602	182430	ATEG_07790	AO090005000531	AFL2G_00532	ACLA_059790	NFIA_099510	Afu8g06830
CO	cellulose oxidase	GH 61	AN9524	53797	ATEG_07920	AO090012000090	AFL2G_03026		NFIA_044390	
CO	cellulose oxidase	GH 61	AN2388		ATEG_01035	AO090023000056		ACLA_073030	NFIA_089670	Afu2g14490

Supplemental Table 6B Cont: Detection of proteins in cultures grown on sugar beet pulp sorted by CAZy family.

CO	cellulose oxidase	GH 61	AN3046	43784	ATEG_08113	AO090023000159	AFL2G_04048	ACLA_055060	NFIA_089730	Afu2g14540
CO	cellulose oxidase	GH 61			ATEG_05416	AO090023000787	AFL2G_04596	ACLA_047220	NFIA_108320	Afu4g07850
CO	cellulose oxidase	GH 61	AN6428		ATEG_06077					
CO	cellulose oxidase	GH 61		56338	ATEG_05081	AO090103000087	AFL2G_12282		NFIA_055220	Afu6g09540
CO	cellulose oxidase	GH 61			ATEG_10194					
CO	cellulose oxidase	GH 61	AN7891	194765		AO090138000004	AFL2G_08699			
CO	cellulose oxidase	GH 61	AN3860							
CO	cellulose oxidase	GH 61			ATEG_08942					
CO	cellulose oxidase	GH 61			ATEG_01456					
CO	cellulose oxidase	GH 61						ACLA_017470		
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_00186	AO090103000088	AFL2G_12281	ACLA_071560	NFIA_087900	Afu2g12770
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN2632	55138	ATEG_10071	AO090701000885	AFL2G_06447	ACLA_061510	NFIA_033210	Afu2g00920
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN7908							
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_10379					
AGU	alpha glucuronidase	GH 67	AN9286	56619	ATEG_06085	AO090026000127	AFL2G_07114	ACLA_017270	NFIA_072510	Afu5g14380
AGU	alpha glucuronidase	GH 67			ATEG_09975					
XG CBH	xyloglucan active cellobiohydrolase	GH 74	AN1542						NFIA_096000	Afu8g02330
XG EGL	xyloglucan active endoglucanase	GH 74	AN5061	206333	ATEG_04708			ACLA_044310	NFIA_094990	Afu8g01490
RHA	alpha rhamnosidase	GH 78	AN8465	40264			AFL2G_05364			
RHA	alpha rhamnosidase	GH 78								
RHA	alpha rhamnosidase	GH 78	AN2631	176718	ATEG_05089	AO090001000105	AFL2G_07340		NFIA_004560	Afu3g02880
RHA	alpha rhamnosidase	GH 78	AN11954	51410		AO090003001016	AFL2G_02014			
RHA	alpha rhamnosidase	GH 78	AN10277	170172		AO090003001291	AFL2G_01780		NFIA_022970	Afu1g01660
RHA	alpha rhamnosidase	GH 78		42916	ATEG_03018	AO090009000471	AFL2G_10227		NFIA_018620	Afu1g06130
RHA	alpha rhamnosidase	GH 78	AN7151		ATEG_02922	AO090010000561	AFL2G_11713		NFIA_057930	Afu6g12030
RHA	alpha rhamnosidase	GH 78	AN3780							
RHA	alpha rhamnosidase	GH 78				AO090103000432	AFL2G_11972			
RHA	alpha rhamnosidase	GH 78				AO090113000149/A	AFL2G_08627		NFIA_057930	Afu6g14610
RHA	alpha rhamnosidase	GH 78	AN6929	131668		AO090005001416				
RHA	alpha rhamnosidase	GH 78							NFIA_060560	
RHA	alpha rhamnosidase	GH 78					AFL2G_10644			
RHA	alpha rhamnosidase	GH 78	AN12368							
RHA	alpha rhamnosidase	GH 78					AFL2G_03939			
RHA	alpha rhamnosidase	GH 78	AN10867	44977	ATEG_04706	AO090012000058	AFL2G_02993		NFIA_026130	Afu7g05040
UGH	unsaturated galacturonyl hydrolase	GH 88				AO090005000324	AFL2G_00324		NFIA_026210	Afu7g05090
UGH	unsaturated galacturonyl hydrolase	GH 88	AN3991	36414		AO090138000087	AFL2G08766		NFIA_023340	Afu1g01250
UGH	unsaturated galacturonyl hydrolase	GH 88	AN11078				AFL2G_06465			
UGH	unsaturated galacturonyl hydrolase	GH 88				AO090701000907				
UGH	unsaturated galacturonyl hydrolase	GH 88	AN4629							
ABX	exoarabinanase	GH 93	AN2060		ATEG_06045	AO090003001017	AFL2G_02013		NFIA_081320	Afu2G04570
ABX	exoarabinanase	GH 93				AO090011000141	AFL2G_04929			
ABX	exoarabinanase	GH 93	AN5231	49311	ATEG_07909	AO090012000101	AFL2G_03035	ACLA_008170	NFIA_027720	Afu7g06430
ABX	exoarabinanase	GH 93			ATEG_08029					
ABX	exoarabinanase	GH 93			ATEG_00891				NFIA_058060	Afu6g12120
AFC	alpha fucosidase	GH 95	AN8149	184037	ATEG_01560	AO090005000382	AFL2G_00373			
AFC	alpha fucosidase	GH 95				AO090005000512	AFL2G_00515			
AFC	alpha fucosidase	GH 95	AN6673	53702	ATEG_09768	AO090009000086	AFL2G_10565	ACLA_004070	NFIA_040960	Afu5g01190
AFC	alpha fucosidase	GH 95	AN10376		ATEG_04136				NFIA_064720	Afu3g12590
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN3196	41703	ATEG_07907	AO090001000174	AFL2G_07407	ACLA_056060	NFIA_030670	Afu4g02860
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105				AO090003000153	AFL2G_02808			
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN10505			AO090001000063	AFL2G_07308	ACLA_072870	NFIA_089840	Afu2g14630
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN9383	41877	ATEG_02892	AO090113000146	AFL2G_08624	ACLA_077810		
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN7828							
AGU	alpha glucuronidase	GH115				AO090005001415	AFL2G_01323			
AGU	alpha glucuronidase	GH115	AN9329		ATEG_09974	AO090010000038	AFL2G_11304	ACLA_006360	NFIA_025630	Afu7g04680
AGU	alpha glucuronidase	GH115				AO090001000267	AFL2G_07498			
AGU	alpha glucuronidase	GH115				AO090113000058				
AGU	alpha glucuronidase	GH 115			ATEG_04355					
PEL	pectin lyase	PL 1			ATEG_00950	AO090003001295	AFL2G_01776			
PEL	pectin lyase	PL 1		55212		AO090010000030	AFL2G_11297		NFIA_026110	Afu7g05030
PEL	pectin lyase	PL 1		45821						
PEL	pectin lyase	PL 1	AN4882	40837	ATEG_7577	AO090012000121	AFL2G_03052		NFIA_077100	Afu5g10170
PEL	pectin lyase	PL 1				AO090012000451	AFL2G_03359			
PEL	pectin lyase	PL 1				AO090103000463	AFL2G_11948			
PEL	pectin lyase	PL 1	AN10147	208760		AO090138000204	AFL2G_08823			
PEL	pectin lyase	PL 1						ACLA_094210	NFIA_033080	Afu2g00800
PEL	pectin lyase	PL 1	AN2331	41815		AO090010000504	AFL2G_11666	ACLA_013470	NFIA_076850	Afu5g10380
PEL	pectin lyase	PL 1	AN2569		ATEG_01216	AO090010000087	AFL2G_11352			
PEL	pectin lyase	PL 1							NFIA_023100	
PEL	pectin lyase	PL 1	AN9439							
PLY	pectate lyase	PL 1	AN0741		ATEG_08834	AO090011000673	AFL2G_05417			
PLY	pectate lyase	PL 1	AN5333		ATEG_05467	AO090102000072	AFL2G_09523		NFIA_060270	Afu6g14400
PLY	pectate lyase	PL 1	AN7646	45021	ATEG_08123	AO090701000321	AFL2G_05954		NFIA_033040	Afu2g00760
PLY	pectate lyase	PL 1	AN9367/AN9368			AO090011000030	AFL2G_04835			
PLY	pectate lyase	PL 3	AN6748		ATEG_06314	AO090005000472	AFL2G_00461		NFIA_027690	Afu7g06400
PLY	pectate lyase	PL 3	AN3337		ATEG_06285					
PLY	pectate lyase	PL 3	AN2542			AO090010000706	AFL2G_11846	ACLA_059210	NFIA_098670	Afu8g05910
PLY	pectate lyase	PL 3	AN8166		ATEG_08626	AO090038000502	AFL2G_07839		NFIA_023470	Afu1g01120
PLY	pectate lyase	PL 3	AN8453							
RGL	rhamnogalacturonan lyase	PL 4	AN7135	210947	ATEG_02193	AO090011000349	AFL2G_05136	ACLA_054660	NFIA_029620	Afu4g03780
RGL	rhamnogalacturonan lyase	PL 4	AN6395	47780	ATEG_10327	AO090012000147/A	AFL2G_03075	ACLA_018320	NFIA_008140	Afu1g17230
RGL	rhamnogalacturonan lyase	PL 4	AN12097			AO090113000057				
RGL	rhamnogalacturonan lyase	PL 4			ATEG_08610	AO090138000119	AFL2G_08794		NFIA_094270	Afu8g00820
RGL	rhamnogalacturonan lyase	PL 4	AN3950							
PLY	pectate lyase	PL 9	AN2537			AO090038000131	AFL2G_08953		NFIA_062360	Afu3g14890
PLY	pectate lyase	PL 9			ATEG_03526					
RGL	rhamnogalacturonan lyase	PL 11	AN2543							

Supplemental Table 6C: Detected proteins in cultures grown on wheat bran sorted by number of species that contain an orthologue. The colour codes indicate which percentage of the total number of detected peptides was of the specific protein. (Continued on next page; Please see pdf file for enlargened text).

			0-0.1%	0.1-0.5%	0.5-1%	1-2%	2-5%	5-10%	>10%					
enzyme code	function	CAZY	*A. nidulans*	*A. niger*	*A. terreus*	*A. oryzae*	*A. flavus*	*A. clavatus*	*A. fischeri*	*A. fumigatus*	no. species with detected protein	no. of species with orthologue of gene	percentage detected proteins per no. of orthologs
BGL	beta glucosidase	GH 3	AN4102	56762	ATEG_03047	A0900090000356	AFL2G_10322	ACLA_028810	NFIA_018950	Afu1g05770	8	8	100
CBH	cellobiohydrolase	GH 7	AN0024	51773	ATEG_00692	A0900001000348	AFL2G_07571	ACLA_065260	NFIA_057300	Afu5g11870	8	8	100
LAC	beta galactosidase	GH 35	AN0758	51784	ATEG_00616	A0900012000445	AFL2G_03352	ACLA_021260	NFIA_011250	Afu1g14170	8	8	100
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN2217	50997	ATEG_09314	A0900701000274	AFL2G_05912	ACLA_010340	NFIA_080180	Afu5g07080	7	8	88
XLN	endoxylanase	GH 10	AN1818	57436	ATEG_00809	A0900701000887	AFL2G_06449	ACLA_048770	NFIA_106540	Afu4g09460	7	8	88
GLA	glucoamylase	GH 15	AN11143	213597	ATEG_04375	A0900010000746	AFL2G_11685	ACLA_094060	NFIA_032960	Afu2g00690	7	8	88
AGD	alpha glucosidase	GH 31	AN0941	119858	ATEG_05177	A0900005001084	AFL2G_01036	ACLA_019300	NFIA_009180	Afu1g16250	7	8	88
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN2632	53176	ATEG_10073	A0900701000885	AFL2G_06447	ACLA_061510	NFIA_013210	Afu2g00820	7	8	88
CBH	cellobiohydrolase	GH 6	AN5282	54490	ATEG_07493			ACLA_062560	NFIA_002990	Afu3g01910	6	6	100
ABX	exoarabinanase	GH 93	AN2060		ATEG_06045	A0900003001017	AFL2G_02013		NFIA_081320	Afu2g04570	6	6	100
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN8401		ATEG_09052	A0900103000120	AFL2G_12252	ACLA_062400	NFIA_003180	Afu3g02090	6	7	86
AGL	alpha galactosidase	GH 27	AN7152	41606	ATEG_02160	A0900023000151	AFL2G_04039		NFIA_029860	Afu4g03580	6	7	86
ABN*	endoarabinanase	GH 43	AN8007	197735		A0900138000055	AFL2G_08737	ACLA_098980	NFIA_047920	Afu6g00770	6	7	86
XLN	endoxylanase	GH 11	AN9365	52071	ATEG_07461	A0900120000026	AFL2G_08066	ACLA_063140	NFIA_000850	Afu3g00320	6	8	75
ABF	alpha arabinofuranosidase	GH 54	AN1571	200605	ATEG_07939	A0900023000001	AFL2G_03901	ACLA_066470	NFIA_060830	Afu6g14620	6	8	75
EGL	endoglucanase	GH 5	AN1355	205580	ATEG_05002			ACLA_085250	NFIA_057290	Afu6g11600	5	6	83
BGL	beta glucosidase	GH 3	AN7396	179265	ATEG_10320	A0900012000135	AFL2G_03066		NFIA_007920	Afu1g17410	5	7	71
XLN	endoxylanase	GH 10	AN7401		ATEG_03416	A0900103000326	AFL2G_12071	ACLA_068910	NFIA_058570	Afu5g13610	5	7	71
XG EGL	xyloglucan active endoglucanase	GH 12		211053	ATEG_07420	A0900026000102	AFL2G_07140	ACLA_007820	NFIA_027400	Afu7g06150	5	7	71
PGX	exopolygalacturonase	GH 28	AN8761	42184	ATEG_07152	A0900026000784	AFL2G_06533		NFIA_049320	Afu6g02980	5	7	71
LAC	beta galactosidase	GH 35	AN0980	46429	ATEG_05131	A0900120000158	AFL2G_08182		NFIA_008690	Afu1g16700	5	7	71
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2528	189254	ATEG_03511	A0900701000556	AFL2G_06177	ACLA_041970	NFIA_062750	Afu3g14510	5	8	63
CBH	cellobiohydrolase	GH 7	AN5178	53159	ATEG_03727	A0900012000941	AFL2G_03805	ACLA_088870	NFIA_052720	Afu6g07070	5	8	63
PGX	exopolygalacturonase	GH 28	AN8891	191158	ATEG_10357	A0900010000753	AFL2G_11892	ACLA_043100	NFIA_096340	Afu8g02630	5	8	63
INX	exo inulinase	GH 32	AN11778	58664	ATEG_08155	A0900701000409	AFL2G_06026	ACLA_094550	NFIA_033540	Afu2g01240	5	8	63
ABF	alpha arabinofuranosidase	GH 51	AN9439	131691	ATEG_07866	A0900012000296	AFL2G_03217	ACLA_074120	NFIA_095410	Afu2g15160	5	8	63
PEL	pectin lyase	PL 1	AN2569		ATEG_01216	A0900010000087	AFL2G_11352				4	4	100
AMY	alpha amylase	GH 13	AN3402	47911	ATEG_10103	A0900023000944		ACLA_052920			4	5	80
ABF	alpha arabinofuranosidase	GH 43	AN8472			A0900012000356	AFL2G_03265		NFIA_032900	Afu2g00650	4	5	80
PLY	pectate lyase	PL 9	AN2537			A0900038000131	AFL2G_08953		NFIA_062360	Afu3g14890	4	5	80
PME	pectin methyl esterase	CE 8	AN3390	214857		A0900102000010	AFL2G_09467		NFIA_099600	Afu8g06880	4	6	67
XGH	xyloglucanase	GH 28	AN3389	46065		A0900102000011	AFL2G_09468		NFIA_099610	Afu8g06890	4	6	67
GUS	beta glucuronidase	GH 2	AN2395	52111	ATEG_01031	A0900023000053	AFL2G_03956		NFIA_089700	Afu2g14520	4	7	57
BGL	beta glucosidase	GH 3	AN2828		ATEG_07419	A0900701000841	AFL2G_06408	ACLA_007810	NFIA_027390	Afu7g06140	4	7	57
CBH	cellobiohydrolase	GH 6	AN1793	133986	ATEG_00193	A0900038000439	AFL2G_07776	ACLA_025560	NFIA_015680		4	7	57
ABN	endoarabinanase	GH 43		182100	ATEG_07817	A0900026000804	AFL2G_06513	ACLA_072730	NFIA_089980	Afu2g14750	4	7	57
GAL	endogalactanase	GH 53	AN5727	187227	ATEG_02927	A0900001000492	AFL2G_09135		NFIA_017780	Afu1g08910	4	7	57
PLY	pectate lyase	PL 1	AN7646	45021	ATEG_08123	A0900701000321	AFL2G_05954		NFIA_033040	Afu2g00760	4	7	57
FAE SF5	feruloyl esterase	CE 1	AN5267	43785	ATEG_08112	A0900023000158	AFL2G_04047	ACLA_055050	NFIA_089720	Afu2g14530	4	8	50
PME	pectin methyl esterase	CE 8	AN4860	174365	ATEG_01704	A0900026000404	AFL2G_10916	ACLA_035610	NFIA_069500	Afu3g07650	4	8	50
MAN	endomannanase	GH 5	AN3358	50376	ATEG_08654	A0900010000122	AFL2G_11381	ACLA_066420	NFIA_113780	Afu7g01070	4	8	50
AGL	alpha galactosidase	GH 27	AN7624	207264	ATEG_09830	A0900003001305	AFL2G_01765	ACLA_003130	NFIA_039990	Afu5g02130	4	8	50
AGD	alpha glucosidase	GH 31	AN2017	214233	ATEG_00725	A0900003001209	AFL2G_01842	ACLA_049370	NFIA_105900	Afu4g10160	4	8	50
AGU	alpha glucuronidase	GH 67	AN8286	56619	ATEG_06085	A0900026000127	AFL2G_07114	ACLA_017270	NFIA_072510	Afu5g14360	4	8	50
RGL	rhamnogalacturonan lyase	PL 4	AN7135	210947	ATEG_02193	A0900011000349	AFL2G_05136	ACLA_054660	NFIA_029620	Afu4g03780	4	8	50
XLN	endoxylanase	GH 10			ATEG_07190				NFIA_081880	Afu3g15210	3	3	100
AGL	alpha galactosidase	GH 27			ATEG_01905				NFIA_048850	Afu6g02560	3	3	100
AGD	alpha glucosidase	GH 31	AN8953			A0900038000471	AFL2G_07812				3	3	100
EGL	endoglucanase	GH 7			ATEG_08705			ACLA_098940	NFIA_047960	Afu6g01800	3	4	75
AMY	alpha amylase	GH 13				A0900120000196		ACLA_094070	NFIA_032970	Afu2g00710	3	4	75
GLA	glucoamylase	GH 15			ATEG_05980	A0900003000321	AFL2G_02658	ACLA_089470			3	4	75
SUC	invertase/beta fructofuranosidase	GH 32		198063	ATEG_07479	A0900023000840	AFL2G_10707				3	4	75
ABF	alpha arabinofuranosidase	GH 43	AN7781		ATEG_08386	A0900005000065	AFL2G_00096				3	4	75
BXL	beta xylosidase	GH 43	AN8477		ATEG_01292		AFL2G_06974	ACLA_018160			3	4	75
BXL	beta xylosidase	GH 3	AN2559	205670	ATEG_05106	A0900005000986	AFL2G_00957			Afu1g16920	3	5	60
ABF	alpha arabinofuranosidase	GH 43	AN2533			A0900701000838	AFL2G_06400		NFIA_002750	Afu3g01660	3	5	60
BXL	beta xylosidase	GH 43	AN10618		ATEG_10072	A0900010000029	AFL2G_11296		NFIA_093350		3	5	60
EGL	endoglucanase	GH 5	AN8068		ATEG_09802	A0900023001341+	AFL2G_01726		NFIA_040280	Afu5g01830	3	6	50
MAN	endomannanase	GH 5	AN7639			A0900038000444	AFL2G_07781	ACLA_044470	NFIA_099770	Afu8g07030	3	6	50
BXL	beta xylosidase	GH 43	AN10199		ATEG_00893	A0900005000698	AFL2G_00689/0	ACLA_058100/	NFIA_097490/0	Afu8g04710/047	3	6	50
PLY	pectate lyase	PL 3	AN2542			A0900010000706	AFL2G_11846	ACLA_059210	NFIA_098670	Afu8g05910	3	6	50
RGX	exorhamnogalacturonase	GH 28	AN10274	194461	ATEG_09025	A0900102000139	AFL2G_09582		NFIA_027700	Afu7g06410	3	7	43
PEL	pectin lyase	PL 1	AN2331	41815		A0900010000504	AFL2G_11666	ACLA_013470	NFIA_076850	Afu5g10380	3	7	43
AXE	acetyl xylan esterase	CE 1	AN6093	211544	ATEG_09643	A0900011000745	AFL2G_05471	ACLA_081220	NFIA_099230	Afu8g06570	3	8	38
BGL	beta glucosidase	GH 3	AN10482	208871	ATEG_06617	A0900001000544	AFL2G_09187	ACLA_083710	NFIA_054350	Afu6g08700	3	8	38
AGD	alpha glucosidase	GH 31	AN0280	55419	ATEG_02528	A0900005000767	AFL2G_06750	ACLA_031260	NFIA_021450	Afu1g03140	3	8	38
BGL	beta glucosidase	GH 3				A0900038000223	AFL2G_09023				2	2	100
MAN	endomannanase	GH 5	AN8427		ATEG_09991						2	2	100
XLN	endoxylanase	GH 10				A0900010000208	AFL2G_07457				2	2	100
ABN	endoarabinanase	GH 43				A0900005000064	AFL2G_00085				2	2	100
ABX	exoarabinanase	GH 93				A0900011000141	AFL2G_04929				2	2	100
BXL	beta xylosidase	GH 3			ATEG_07383	A0900011000140	AFL2G_04928				2	3	67
XLN	endoxylanase	GH 10			ATEG_08906	A0900103000423	AFL2G_11963				2	3	67
ABX	exoarabinanase	GH 93			ATEG_00891				NFIA_058060	Afu6g12120	2	3	67
PEL	pectin lyase	PL 1						ACLA_094210	NFIA_033080	Afu2g00800	2	3	67
FAE SF5	feruloyl esterase	CE 1			ATEG_06644	A0900701000884	AFL2G_06446	ACLA_061520			2	4	50
RGAE	rhamnogalacturonan acetyl esterase	CE 12		51400	ATEG_10016	A0900102000092	AFL2G_09543				2	4	50
PME	pectin methyl esterase	CE 8		44585		A0900012000749	AFL2G_03618	ACLA_044240	NFIA_095020	Afu8g01520	2	5	40

Supplemental Table 6C Cont: Detected proteins in cultures grown on wheat bran sorted by number of species that contain an orthologue. (Please see pdf file for enlargened text).

RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2834			A0090113000155	AFL2G_08631		NFIA_092640	AfuZg17250	2	5	40
AGL	alpha galactosidase	GH 27		37736	ATEG_04382			ACLA_016820	NFIA_073100	AfuSg13830	2	5	40
RGL	rhamnogalacturonan lyase	PL 4			ATEG_08610	A0090138000119	AFL2G_08794		NFIA_094270	AfuBg00820	2	5	40
EXG	exo 1,3 galactanase	GH 5			ATEG_06369	A0090005000423	AFL2G_00412	ACLA_007330	NFIA_026860	AfuTg05610	2	6	33
CO	cellulose oxidase	GH 61			ATEG_05416	A0090023000787	AFL2G_04596	ACLA_047220	NFIA_108320	AfuHg07850	2	6	33
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_00186	A0090103000088	AFL2G_12281	ACLA_071560	NFIA_087900	AfuZg12770	2	6	33
PLY	pectate lyase	PL 3	AN6106		ATEG_08626	A0090038000502	AFL2G_07839		NFIA_023470	AfuTg01120	2	6	33
RHA	alpha rhamnosidase	GH 78		42916	ATEG_03018	A0090098000471	AFL2G_10227		NFIA_018620	AfuTg06130	2	6	33
EGL	endoglucanase	GH 5	AN5214	209376		A0090005001553	AFL2G_01447	ACLA_081650	NFIA_053150	AfuBg07480	2	7	29
GLN	exo 1,6 galactanase	GH 5	AN9166	194447	ATEG_10242	A0090012000046	AFL2G_02982		NFIA_072400	AfuSg14560	2	7	29
EGL	endoglucanase	GH 7	AN3418		ATEG_08709	A0090010000314	AFL2G_11497	ACLA_066030	NFIA_114250	AfuTg01540	2	7	29
AGL	alpha galactosidase	GH 27	AN0022	172232	ATEG_03427	A0090005002217	AFL2G_00225		NFIA_023390	AfuTg01200	2	7	29
AXL	alpha xylosidase	GH 31	AN7505		ATEG_06730	A0090001000649	AFL2G_09290	ACLA_089300	NFIA_082140	AfuZg05400	2	7	29
LAC	beta galactosidase	GH 35		177434	ATEG_07446	A0090003000042	AFL2G_02901	ACLA_088440	NFIA_052310	AfuBg06660	2	7	29
BXL	beta xylosidase	GH 43	AN7275		ATEG_06643	A0090701000886	AFL2G_06448	ACLA_078600	NFIA_033220	AfuZg00930	2	7	29
AGU	alpha glucuronidase	GH115	AN9329		ATEG_09974	A0090010000038	AFL2G_11304	ACLA_006360	NFIA_025630	AfuTg04680	2	7	29
MND	beta mannosidase	GH 2	AN1742	138878	ATEG_06638	A0090001000556	AFL2G_09201	ACLA_083570	NFIA_054490	AfuBg08840	2	8	25
BXL	beta xylosidase	GH 43	AN1870	179682	ATEG_06059	A0090003000239	AFL2G_02739	ACLA_090090	NFIA_081240	AfuZg04480	2	8	25
BXL	beta xylosidase	GH 43	AN7313	122978	ATEG_01188	A0090010000562	AFL2G_11714	ACLA_043260	NFIA_096240	AfuBg02510	2	8	25
BXL	beta xylosidase	GH 43	AN7864	47677	ATEG_01292	A0090102000331	AFL2G_09750	ACLA_072000	NFIA_088370	AfuZg13190	2	8	25
CO	cellulose oxidase	GH 61	AN1041	52688	ATEG_05448	A0090001000221	AFL2G_07454	ACLA_022980	NFIA_012990	AfuTg12560	2	8	25
CO	cellulose oxidase	GH 61	AN1602	182430	ATEG_07790	A0090005000531	AFL2G_00532	ACLA_059790	NFIA_099510	AfuBg06830	2	8	25
EXG	exo 1,3 galactanase	GH 5	AN4052	202490	ATEG_03849	A0090003000990	AFL2G_02039	ACLA_031040	NFIA_021060	AfuTg03600	2	8	25
FAE SF5	feruloyl esterase	CE 1			ATEG_01914						1	1	100
BGL	beta glucosidase	GH 3			ATEG_07931						1	1	100
EGL	endoglucanase	GH 5			ATEG_05003						1	1	100
MAN	endomannanase	GH 5	AN3297								1	1	100
MAN	endomannanase	GH 5	AN9276								1	1	100
XLN	endoxylanase	GH 11	AN3613								1	1	100
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_09894						1	1	100
AGD	alpha glucosidase	GH 13					AFL2G_08694				1	1	100
PGA	endopolygalacturonase	GH 28		182156							1	1	100
PGX	exopolygalacturonase	GH 28		178172							1	1	100
AGD	alpha glucosidase	GH 31			ATEG_08278						1	1	100
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_10253						1	1	100
ABN	endoarabinanase	GH 43			ATEG_03688						1	1	100
BXL	beta xylosidase	GH 43			ATEG_10193						1	1	100
?	?	GH 43	AN1477								1	1	100
CO	cellulose oxidase	GH 61			ATEG_10194						1	1	100
CO	cellulose oxidase	GH 61	AN3860								1	1	100
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN7908								1	1	100
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_10379						1	1	100
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN7828								1	1	100
AGU	alpha glucuronidase	GH 115			ATEG_04355						1	1	100
PLY	pectate lyase	PL 3	AN8453								1	1	100
RGAE	rhamnogalacturonan acetyl esterase	CE 12				A0090003001268	AFL2G_01797				1	2	50
AMY	alpha amylase	GH 13	AN3388		ATEG_02515						1	2	50
AGL	alpha galactosidase	GH 36	AN9035		ATEG_02312						1	2	50
ABF	alpha arabinofuranosidase	GH 51		50979	ATEG_03540						1	2	50
AGU	alpha glucuronidase	GH115				A0090001000267	AFL2G_07498				1	2	50
CO	cellulose oxidase	GH 61	AN6428		ATEG_06077						1	2	50
BGL	beta glucosidase	GH 3	AN1804			A0090026000123	AFL2G_07119				1	3	33
BGL	beta glucosidase	GH 3			ATEG_02724	A0090005000337	AFL2G_00334				1	3	33
GLA	glucoamylase	GH 15	AN7402					ACLA_049360	NFIA_105910		1	3	33
FAE SF7	feruloyl esterase	faeA		51662	ATEG_08907	A0090001000207	AFL2G_07436				1	4	25
GE	glucuronoyl esterase	CE15			ATEG_00945			ACLA_087520	NFIA_060260	AfuBg14390	1	4	25
PLY	pectate lyase	PL 1	AN0741		ATEG_08834	A0090011000673	AFL2G_05417				1	4	25
PGA	endopolygalacturonase	GH 28		52219		A0090023000161	AFL2G_04049		NFIA_095620	AfuBg01970	1	5	20
RGX	exorhamnogalacturonase	GH 28		42917		A0090090000470	AFL2G_10228		NFIA_018590	AfuTg06140	1	5	20
AXL	alpha xylosidase	GH 31		43342	ATEG_08390	A0090701000639	AFL2G_06239		NFIA_032680		1	5	20
XG EGL	xyloglucan active endoglucanase	GH 74	AN5061	206333	ATEG_04708			ACLA_044310	NFIA_094990	AfuBg01490	1	5	20
AFC	alpha fucosidase	GH 95	AN8149	184037	ATEG_01560	A0090005000382	AFL2G_00373				1	5	20
?	?	GH 3	AN3360	37673	ATEG_04963			ACLA_052760	NFIA_102600	AfuHg13770	1	6	17
EXG	beta 1,6 glucanase	GH 5	AN3777		ATEG_09844	A0090011000757	AFL2G_05484		NFIA_084850	AfuZg09350	1	6	17
BXL	beta xylosidase	GH 43			ATEG_05083	A0090103000268	AFL2G_12122	ACLA_087680	NFIA_060490	AfuSg14550	1	6	17
BXL	beta xylosidase	GH 43	AN2664		ATEG_06107	A0090012000358	AFL2G_03257		NFIA_072780	AfuSg14190	1	6	17
ABF	alpha arabinofuranosidase	GH 51	AN2541	206387	ATEG_02882	A0090124000023	AFL2G_08019	ACLA_099110			1	6	17
CO	cellulose oxidase	GH 61	AN9524	53797	ATEG_07920	A0090012000090	AFL2G_03026		NFIA_044390		1	6	17
CO	cellulose oxidase	GH 61	AN2388		ATEG_01035	A0090023000056		ACLA_073030	NFIA_089670	AfuZg14490	1	6	17
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN10505			A0090091000063	AFL2G_07308	ACLA_072870	NFIA_089840	AfuZg14630	1	6	17
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN8383	41877	ATEG_02892	A0090113000146	AFL2G_08624	ACLA_077810			1	6	17
PLY	pectate lyase	PL 3	AN6748		ATEG_06314	A0090005000472	AFL2G_00461		NFIA_027690	AfuTg06400	1	6	17
EGL	endoglucanase	GH 5		214608	ATEG_04390	A0090011000715	AFL2G_05447	ACLA_081310	NFIA_085010	AfuZg09520	1	7	14
XLN	endoxylanase	GH 11		171269	ATEG_04943	A0090001000111	AFL2G_07347	ACLA_085410	NFIA_058160	AfuSg12210	1	7	14
ABF	alpha arabinofuranosidase	GH 51	AN1277	38549	ATEG_00198		AFL2G_07796	ACLA_025610	NFIA_015730	AfuTg09900	1	7	14
BGL	beta glucosidase	GH 1	AN9183	131747	ATEG_02657	A0090120000075	AFL2G_08111		NFIA_099670	AfuBg06970	1	7	14
BGL	beta glucosidase	GH 3	AN2227	129891	ATEG_09329	A0090701000244	AFL2G_05886	ACLA_010450	NFIA_080070	AfuSg07190	1	8	13
XG EGL	xyloglucan active endoglucanase	GH 12	AN0452	52011	ATEG_03755	A0090003000905	AFL2G_02120	ACLA_029940	NFIA_020020	AfuTg04730	1	8	13
AMY	alpha amylase	GH 13	AN2018	45304	ATEG_08279	A0090003001210	AFL2G_01941	ACLA_049350	NFIA_105920	AfuHg10130	1	8	13
AGT	4-alpha glucanotransferase	GH 13	AN3308	188489	ATEG_03623	A0090003001498	AFL2G_01594	ACLA_063440	NFIA_001710	AfuSg00900	1	8	13
PGA	endopolygalacturonase	GH 28	AN4372	141677	ATEG_04991	A0090023000401	AFL2G_04252	ACLA_052860	NFIA_102450	AfuHg13920	1	8	13
ABN	endoarabinanase	GH 43	AN2534	203143	ATEG_03520	A0090701000481	AFL2G_06106	ACLA_042100	NFIA_062660	AfuSg14620	1	8	13
AFC	alpha fucosidase	GH 95	AN6673	53762	ATEG_09768	A0090090000086	AFL2G_10565	ACLA_004070	NFIA_040960	AfuSg01190	1	8	13
RGL	rhamnogalacturonan lyase	PL 4	AN6395	47780	ATEG_10327	A0090012000147/A	AFL2G_03075	ACLA_018320	NFIA_008140	AfuTg17230	1	8	13
BGL	beta glucosidase	GH 1	AN10124	213437	ATEG_00687	A0090003000497	AFL2G_02496	ACLA_020660	NFIA_010690	AfuTg14710	1	8	13

Supplemental Table 6D: Detected proteins in cultures grown on sugar beet pulp sorted by number of species that contain an orthologue. The colour codes indicate which percentage of the total number of detected peptides was of the specific protein. (Continued on next page; Please see pdf file for enlargened text).

			0-0.1%	0.1-0.5%	0.5-1%	1-2%	2-5%	5-10%	>10%					
enzyme code	**function**	**CAZY**	***A. nidulans***	***A. niger***	***A. terreus***	***A. oryzae***	***A. flavus***	***A. clavatus***	***A. fischeri***	***A. fumigatus***	no. species with detected protein	no. of species with orthолog of gene	percentage detected proteins per no. of orthologs
FAE SFS	feruloyl esterase	CE 1			ATEG_01914						1	1	100
PME	pectin methyl esterase	CE 8	AN3390	214867		AO090102000010	AFL2G_09467		NFIA_099600	Afu8g06880	6	6	100
RGAE	rhamnogalacturonan acetyl esterase	CE 12				AO090005001268	AFL2G_01797				2	2	100
RGAE	rhamnogalacturonan acetyl esterase	CE 12							NFIA_099110	Afu8g06480	2	2	100
BGL	beta glucosidase	GH 3				AO090038000223	AFL2G_09023				2	2	100
BGL	beta glucosidase	GH 3			ATEG_07931						1	1	100
EGL	endoglucanase	GH 5			ATEG_05003						1	1	100
MAN	endomannanase	GH 5	AN3297								1	1	100
MAN	endomannanase	GH 5	AN6427		ATEG_09991						2	2	100
MAN	endomannanase	GH 5	AN9276								1	1	100
		GH5			ATEG_03877						1	1	100
XLN	endoxylanase	GH 10				AO090001000208	AFL2G_07437				2	2	100
XLN	endoxylanase	GH 11	AN3813								1	1	100
XG EGL	xyloglucan active endoglucanase	GH 12			ATEG_09894						1	1	100
AGD	alpha glucosidase	GH 13					AFL2G_08694				1	1	100
MAN	endomannanase	GH 26	AN7413								1	1	100
AGL	alpha galactosidase	GH 27			ATEG_01905				NFIA_048850	Afu6g02560	3	3	100
AGL	alpha galactosidase	GH 27	AN7152	41606	ATEG_02160	AO090020000151	AFL2G_04039		NFIA_029860	Afu4g03580	7	7	100
PGA	endopolygalacturonase	GH 28		43957							1	1	100
PGA	endopolygalacturonase	GH 28		182156							1	1	100
SUC	invertase/beta fructofuranosidase	GH 32			ATEG_10253						1	1	100
ABN	endoarabinanase	GH 43				AO090005000064	AFL2G_00085				2	2	100
ABN	endoarabinanase	GH 43			ATEG_03688						1	1	100
BXL	beta xylosidase	GH 43			ATEG_10193						1	1	100
CO	cellulose oxidase	GH 61			ATEG_10194						1	1	100
CO	cellulose oxidase	GH 61	AN3860								1	1	100
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN7908								1	1	100
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_10379						1	1	100
RHA	alpha rhamnosidase	GH 78					AFL2G_03939				1	1	100
ABX	exoarabinanase	GH 93	AN2060		ATEG_06045	AO090003001017	AFL2G_02013		NFIA_081320	Afu2g04570	6	6	100
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN7828								1	1	100
PLY	pectate lyase	PL 3	AN3337		ATEG_06285						2	2	100
PLY	pectate lyase	PL 3	AN8453								1	1	100
RGL	rhamnogalacturonan lyase	PL 4	AN3950								1	1	100
PLY	pectate lyase	PL 9			ATEG_03526						1	1	100
RGL	rhamnogalacturonan lyase	PL 11	AN2543								1	1	100
PGX	exopolygalacturonase	GH 28		178172							1	1	100
PGX	exopolygalacturonase	GH 28	AN9045								1	1	100
AGU	alpha glucuronidase	GH 115			ATEG_04355						1	1	100
PME	pectin methyl esterase	CE 8	AN4660	174385	ATEG_01794	AO090020000404	AFL2G_10916	ACLA_035610	NFIA_089500	Afu2g07550	7	8	88
BXL/ABF	beta xylosidase/alpha arabinofuranosidas	GH 3	AN2217	50997	ATEG_09314	AO090701000274	AFL2G_05912	ACLA_010340	NFIA_080180	Afu5g07680	7	8	88
LAC	beta galactosidase	GH 35	AN0756	51764	ATEG_00613	AO090012000445	AFL2G_03352	ACLA_021260	NFIA_011250	Afu1g14170	7	8	88
ABF	alpha arabinofuranosidase	GH 54	AN1571	200605	ATEG_07939	AO090023000001	AFL2G_03901	ACLA_066470	NFIA_060630	Afu5g14620	7	8	88
AXH	arabinoxylan arabinofuranohydrolase	GH 62	AN2632	55136	ATEG_10071	AO090701000885	AFL2G_06447	ACLA_061510	NFIA_033210	Afu2g00920	7	8	88
PGX	exopolygalacturonase	GH 28	AN6761	42184	ATEG_07153	AO090026000784	AFL2G_06332		NFIA_042020	Afu5g02580	6	7	86
CBH	cellobiohydrolase	GH 6	AN8282	54490	ATEG_07493			ACLA_062560	NFIA_002990	Afu3g01910	5	6	83
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2834			AO090113000155	AFL2G_09831		NFIA_092640	Afu2g17250	4	5	80
ABF	alpha arabinofuranosidase	GH 43	AN8472			AO090012000356	AFL2G_03265		NFIA_032900	Afu2g00650	4	5	80
RGL	rhamnogalacturonan lyase	PL 4			ATEG_08610	AO090138000119	AFL2G_08794		NFIA_094270	Afu5g00820	4	5	80
CBH	cellobiohydrolase	GH 7	AN5176	53159	ATEG_03727	AO090012000941	AFL2G_03805	ACLA_088870	NFIA_052720	Afu6g07070	6	8	75
XLN	endoxylanase	GH 10	AN1815	57436	ATEG_00809	AO090701000887	AFL2G_06449	ACLA_048770	NFIA_106540	Afu4g09480	6	8	75
XLN	endoxylanase	GH 11	AN9365	52071	ATEG_07461	AO090120000026	AFL2G_08066	ACLA_063140	NFIA_000850	Afu3g06320	6	8	75
AMY	alpha amylase	GH 13				AO090120000196		ACLA_094070	NFIA_032970	Afu2g00710	3	4	75
GLA	glucoamylase	GH 15	AN11143	213597	ATEG_04375	AO090010000746	AFL2G_11885	ACLA_094080	NFIA_032960	Afu2g00690	6	8	75
PGX	exopolygalacturonase	GH 28	AN8891	191158	ATEG_10357	AO090010000753	AFL2G_11892	ACLA_043100	NFIA_096340	Afu8g02630	6	8	75
INX	exo inulinase	GH 32	AN11778	58664	ATEG_08155	AO090701000400	AFL2G_06028	ACLA_094550	NFIA_033540	Afu2g01240	6	8	75
SUC	invertase/beta fructofuranosidase	GH 32		198063	ATEG_07479	AO090020000640	AFL2G_10707				3	4	75
ABF	alpha arabinofuranosidase	GH 43	AN7781		ATEG_08386	AO090005000065	AFL2G_00088				3	4	75
PEL	pectin lyase	PL 1	AN2569		ATEG_01216	AO090010000087	AFL2G_11352				3	4	75
LAC	beta galactosidase	GH 35	AN0980	46429	ATEG_05131	AO090120000158	AFL2G_08182		NFIA_060680	Afu1g16700	5	7	71
EGL	endoglucanase	GH 5	AN8068		ATEG_09602	AO090003001341	AFL2G_01726		NFIA_040280	Afu5g01830	4	6	67
XLN	endoxylanase	GH 10			ATEG_08908	AO090103000423	AFL2G_11983				2	3	67
XLN	endoxylanase	GH 10			ATEG_07190				NFIA_061880	Afu3g15210	2	3	67
XGH	xyloglucanase	GH 28	AN3389	46065		AO090102000011	AFL2G_09468		NFIA_099610	Afu8g06890	4	6	67
AGD	alpha glucosidase	GH 31	AN8953			AO090038000471	AFL2G_07812				2	3	67
XG CBH	xyloglucan active cellobiohydrolase	GH 74	AN1542						NFIA_096000	Afu8g02330	2	3	67
PEL	pectin lyase	PL 1						ACLA_094210	NFIA_033080	Afu2g00800	2	3	67
PLY	pectate lyase	PL 3	AN6106		ATEG_08029	AO090038000502	AFL2G_07839		NFIA_023470	Afu1g01120	4	6	67
BGL	beta glucosidase	GH 3	AN4102	56234	ATEG_03047	AO090009000356	AFL2G_10322	ACLA_028810	NFIA_018950	Afu1g05770	5	8	63
MAN	endomannanase	GH 5	AN3358	50378	ATEG_09654	AO090010000122	AFL2G_11381	ACLA_066420	NFIA_113780	Afu7g01070	5	8	63
CBH	cellobiohydrolase	GH 7	AN0494	51773	ATEG_05002	AO090001000348	AFL2G_07571	ACLA_085260	NFIA_057300	Afu6g11610	5	8	63
BXL	beta xylosidase	GH 43	AN7313	122978	ATEG_01188	AO090010000562	AFL2G_11714	ACLA_043260	NFIA_096240	Afu8g02510	5	8	63
ABF	alpha arabinofuranosidase	GH 51	AN9439	131891	ATEG_07869	AO090012000290	AFL2G_09211	ACLA_074120	NFIA_090410	Afu2g15780	5	8	63
RGL	rhamnogalacturonan lyase	PL 4	AN7135	210947	ATEG_02193	AO090011002349	AFL2G_05136	ACLA_054660	NFIA_029620	Afu4g03790	5	8	63
RGL	rhamnogalacturonan lyase	PL 4	AN6395	47780	ATEG_10327	AO090012000147	AFL2G_03075	ACLA_018320	NFIA_008140	Afu1g17230	5	8	63
AMY	alpha amylase	GH 13	AN3402	47911	ATEG_10103	AO090021000944		ACLA_052920			3	5	60
BXL	beta xylosidase	GH 43	AN10919		ATEG_10072	AO090010000029	AFL2G_11296		NFIA_093350		3	5	60
PLY	pectate lyase	PL 1	AN5333		ATEG_05467	AO090102000072	AFL2G_09523		NFIA_060270	Afu5g14400	3	5	60
PLY	pectate lyase	PL 9	AN2537			AO090038000131	AFL2G_08963		NFIA_062360	Afu3g14690	3	5	60
XG EGL	xyloglucan active endoglucanase	GH 12		211053	ATEG_07428	AO090026000102	AFL2G_07140	ACLA_007820	NFIA_027400	Afu7g06150	4	7	57
RGX	exorhamnogalacturonase	GH 28	AN10274	194461	ATEG_09025	AO090102000139	AFL2G_09582		NFIA_027700	Afu7g06410	4	7	57
LAC	beta galactosidase	GH 35		177434	ATEG_07446	AO090003000042	AFL2G_02901	ACLA_088440	NFIA_052310	Afu5g06560	4	7	57
ABN*	endoarabinanase	GH 43	AN8007	197735		AO090138000055	AFL2G_08737	ACLA_098980	NFIA_047920	Afu6g00770	4	7	57

Supplemental Table 6D Cont: Detected proteins in cultures grown on sugar beet pulp sorted by number of species that contain an orthologue. (Please see pdf file for enlargened text).

GAL	endogalactanase	GH 53	AN5727	187227	ATEG_02927	AO090001000492	AFL2G_09135		NFIA_017780	Afu1g06910	4	7	57
PEL	pectin lyase	PL 1	AN2331	41815		AO090010000504	AFL2G_11666	ACLA_013470	NFIA_076850	Afu5g10380	4	7	57
XG EGL	xyloglucan active endoglucanase	GH 12		191511					NFIA_100330		1	2	50
FAE SF5	feruloyl esterase	CE 1			ATEG_06644	AO090701000884	AFL2G_06446	ACLA_061820			2	4	50
RGAE	rhamnogalacturonan acetyl esterase	CE 12		51400	ATEG_10016	AO090102000092	AFL2G_09543				2	4	50
RGAE	rhamnogalacturonan acetyl esterase	CE 12	AN2528	189254	ATEG_02511	AO090701000556	AFL2G_06177	ACLA_041970	NFIA_062750	Afu3g14510	4	8	50
LAC	beta galactosidase	GH 2	AN2463		ATEG_00712						1	2	50
BGL	beta glucosidase	GH 3	AN2612							Afu7g00240	1	2	50
EGL	endoglucanase	GH 5	AN1285	205580	ATEG_05002			ACLA_085250	NFIA_057290	Afu6g11600	3	6	50
AMY	alpha amylase	GH 13	AN3388		ATEG_02515						1	2	50
AGL	alpha galactosidase	GH 27	AN7624	207264	ATEG_09830	AO090003001305	AFL2G_01765	ACLA_003130	NFIA_039990	Afu5g02130	4	8	50
PGA	endopolygalacturonase	GH 28	AN8327	46255					NFIA_008150	Afu1g17220	2	4	50
BXL	beta xylosidase	GH 43	AN2664		ATEG_06107	AO090012000350	AFL2G_03257		NFIA_072780	Afu5g14190	3	6	50
BXL	beta xylosidase	GH 43	AN8477		ATEG_01292		AFL2G_06974	ACLA_018160			2	4	50
ABF	alpha arabinofuranosidase	GH 51		50979	ATEG_03540						1	2	50
CO	cellulose oxidase	GH 61	AN6428		ATEG_06077						1	2	50
AGU	alpha glucuronidase	GH 67	AN9286	56619	ATEG_06085	AO090026000127	AFL2G_07114	ACLA_017270	NFIA_072510	Afu5g14380	4	8	50
XG EGL	xyloglucan active endoglucanase	GH 74	AN5061	206333	ATEG_04708			ACLA_044310	NFIA_094990	Afu8g01490	3	6	50
ABX	exoarabinanase	GH 93				AO090011000141	AFL2G_04929				1	2	50
AFC	alpha fucosidase	GH 95	AN6673	53702	ATEG_09768	AO090009000086	AFL2G_10565	ACLA_004070	NFIA_040960	Afu5g01190	4	8	50
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN10505			AO090001000063	AFL2G_07308	ACLA_072870	NFIA_089840	Afu2g14630	3	6	50
URH	unsaturated rhamnogalacturonyl hydrolas	GH 105	AN9383	41877	ATEG_02892	AO090113000146	AFL2G_06624	ACLA_077810			3	6	50
AGU	alpha glucuronidase	GH115				AO090001000267	AFL2G_07498				1	2	50
PLY	pectate lyase	PL 3	AN6748		ATEG_06314	AO090005000472	AFL2G_00461		NFIA_027690	Afu7g06400	3	6	50
BGL	beta glucosidase	GH 3	AN2828		ATEG_07419	AO090701000841	AFL2G_06408	ACLA_007810	NFIA_027390	Afu7g06140	3	7	43
BXL/ABF	beta xylosidase/alpha arabinofuranosida	GH 3	AN8401		ATEG_09052	AO090103000120	AFL2G_12252	ACLA_062400	NFIA_003180	Afu3g02090	3	7	43
XLN	endoxylanase	GH 10	AN7401		ATEG_03410	AO090103000326	AFL2G_12071	ACLA_086910	NFIA_059570	Afu6g13610	3	7	43
PGA	endopolygalacturonase	GH 28		172944	ATEG_01601	AO090005001400	AFL2G_01310	ACLA_036670	NFIA_068440	Afu3g06630	3	7	43
PLY	pectate lyase	PL 1	AN7646	45021	ATEG_08123	AO090701000321	AFL2G_05954		NFIA_033040	Afu2g00760	3	7	43
AGL	alpha galactosidase	GH 27		37736	ATEG_04382			ACLA_016820	NFIA_073100	Afu5g13830	2	5	40
ABF	alpha arabinofuranosidase	GH 43	AN2533			AO090701000838	AFL2G_06400		NFIA_002750	Afu3g01680	2	5	40
PLY	pectate lyase	PL 3	AN2542			AO090010000706	AFL2G_11846	ACLA_059210	NFIA_098670	Afu4g05910	2	5	40
FAE SF5	feruloyl esterase	CE 1	AN5267	43785	ATEG_08112	AO090023000158	AFL2G_04047	ACLA_055050	NFIA_089720	Afu2G14530	3	8	38
PME	pectin methyl esterase	CE 8		44585		AO090012000749	AFL2G_03618	ACLA_044240	NFIA_095020	Afu8g01520	2	6	33
BGL	beta glucosidase	GH 3	AN1804			AO090026000123	AFL2G_07119				1	3	33
BXL	beta xylosidase	GH 3	AN2359	205670	ATEG_05106	AO090005000986	AFL2G_10957			Afu1g16920	2	6	33
BXL	beta xylosidase	GH 3			ATEG_07383	AO090011000140	AFL2G_04928				1	3	33
MAN	endomannanase	GH 5	AN7639			AO090038000444	AFL2G_07781	ACLA_044470	NFIA_099770	Afu8g07030	2	6	33
EGL	endoglucanase	GH 7			ATEG_08705			ACLA_098940	NFIA_047960	Afu6g01800	1	3	33
GLA	glucoamylase	GH 15	AN7402					ACLA_049360	NFIA_105910		1	3	33
RGX	exorhamnogalacturonase	GH 28			ATEG_06408	AO090113000199	AFL2G_08671				1	3	33
ABF	alpha arabinofuranosidase	GH 51	AN2541	206387	ATEG_02882	AO090124000023	AFL2G_08019	ACLA_099110			2	6	33
CO	cellulose oxidase	GH 61			ATEG_05418	AO090023000787	AFL2G_04596	ACLA_042290	NFIA_108320	Afu4g07850	2	6	33
AXH	arabinoxylan arabinofuranohydrolase	GH 62			ATEG_00186	AO090103000088	AFL2G_12281	ACLA_071560	NFIA_087900	Afu2g12770	2	6	33
RHA	alpha rhamnosidase	GH 78		42916	ATEG_03018	AO090009000471	AFL2G_10227		NFIA_018620	Afu1g06130	2	6	33
RHA	alpha rhamnosidase	GH 78	AN7151		ATEG_02922	AO090010000561	AFL2G_11713		NFIA_057930	Afu6g12030	2	6	33
ABX	exoarabinanase	GH 93			ATEG_00891				NFIA_058060	Afu6g12120	1	3	33
GUS	beta glucuronidase	GH 2	AN2395	52111	ATEG_01031	AO090023000053	AFL2G_03956		NFIA_089700	Afu2g14520	2	7	29
BGL	beta glucosidase	GH 3	AN7396	179265	ATEG_10320	AO090012000135	AFL2G_03066		NFIA_007920	Afu1g17410	2	7	29
EGL	endoglucanase	GH 5	AN5214	209376		AO090005001553	AFL2G_01447	ACLA_081650	NFIA_053150	Afu6g07460	2	7	29
GLN	exo 1,6 galactanase	GH 5	AN9166	194447	ATEG_10242	AO090012000046	AFL2G_02962		NFIA_072400	Afu5g14580	2	7	29
CBH	cellobiohydrolase	GH 6	AN1273	133986	ATEG_00193	AO090038000439	AFL2G_07776	ACLA_025560	NFIA_015680		2	7	29
EGL	endoglucanase	GH 7	AN3416		ATEG_08700	AO090010000314	AFL2G_11497	ACLA_066030	NFIA_114250	Afu7g01540	2	7	29
XLN	endoxylanase	GH 11		171269	ATEG_04943	AO090001000111	AFL2G_07347	ACLA_085410	NFIA_058160	Afu6g12210	2	7	29
BXL	beta xylosidase	GH 43	AN7275		ATEG_06643	AO090701000886	AFL2G_06448	ACLA_078600	NFIA_033220	Afu2g00930	2	7	29
AGU	alpha glucuronidase	GH115	AN9329		ATEG_09974	AO090010000038	AFL2G_11304	ACLA_068360	NFIA_025630	Afu7g04680	2	7	29
RHG	endorhamnogalacturonase	GH 28		211163	ATEG_07607	AO090010000484	AFL2G_11646				1	4	25
AXE	acetyl xylan esterase	CE 1	AN6093	211544	ATEG_09843	AO090011000745	AFL2G_05471	ACLA_081220	NFIA_099230	Afu8g06570	2	8	25
GE	glucuronoyl esterase	CE15			ATEG_00945			ACLA_087520	NFIA_060260	Afu6g14390	1	4	25
BGL	beta glucosidase	GH 3	AN10482	208871	ATEG_06617	AO090001000544	AFL2G_09187	ACLA_083710	NFIA_054350	Afu6g08700	2	8	25
GLA	glucoamylase	GH 15			ATEG_05980	AO090003000321	AFL2G_02658	ACLA_089470			1	4	25
PGA	endopolygalacturonase	GH 28	AN4372	141677	ATEG_04991	AO090023000401	AFL2G_04252	ACLA_052860	NFIA_102450	Afu4g13920	2	8	25
AGD	alpha glucosidase	GH 31	AN2017	214233	ATEG_00723	AO090003001209	AFL2G_01842	ACLA_049370	NFIA_105900	Afu4g10150	2	8	25
AGD	alpha glucosidase	GH 31	AN6941	119658	ATEG_05177	AO090005001084	AFL2G_10338	ACLA_019300	NFIA_009180	Afu1g16250	2	8	25
LAC	beta galactosidase	GH 35		41910			AFL2G_03616		NFIA_000910	Afu3g00380	1	4	25
ABN	endoarabinanase	GH 43	AN2534	203144	ATEG_03520	AO090701000481	AFL2G_06106	ACLA_042100	NFIA_062660	Afu3g14620	2	8	25
BXL	beta xylosidase	GH 43	AN1870	179682	ATEG_06059	AO090003000239	AFL2G_02739	ACLA_090090	NFIA_081240	Afu2g04480	2	8	25
CO	cellulose oxidase	GH 61	AN1041	52688	ATEG_05448	AO090010000221	AFL2G_07454	ACLA_022980	NFIA_012990	Afu1g12560	2	8	25
CO	cellulose oxidase	GH 61	AN1602	182430	ATEG_07790	AO090005000531	AFL2G_00532	ACLA_059790	NFIA_099510	Afu8g06830	2	8	25
PLY	pectate lyase	PL 1	AN0741		ATEG_08834	AO090011000673	AFL2G_05417				1	4	25
CO	cellulose oxidase	GH 61	AN7891	194765		AO090138000004	AFL2G_08699				1	4	25
AXL	alpha xylosidase	GH 31		43342	ATEG_08390	AO090701000639	AFL2G_06239		NFIA_032680		1	5	20
PGA	endopolygalacturonase	GH 28	AN6656	50161		AO090005000186	AFL2G_00201		NFIA_099410		1	5	20
RGX	exorhamnogalacturonase	GH 28		42917		AO090009000470	AFL2G_10228		NFIA_018590	Afu1g06140	1	5	20
AFC	alpha fucosidase	GH 95	AN8149	184037	ATEG_01560	AO090005000382	AFL2G_00373				1	5	20
PGA	endopolygalacturonase	GH 28		52219		AO090023000161	AFL2G_04049		NFIA_095620	Afu8g01970	1	5	20
?		GH 3	AN3360	37673	ATEG_04963			ACLA_052760	NFIA_102600	Afu4g13770	1	6	17
CO	cellulose oxidase	GH 61	AN9524	53797	ATEG_07920	AO090012000090	AFL2G_03026		NFIA_044390		1	6	17
RHA	alpha rhamnosidase	GH 78	AN10277	170172		AO090003001291	AFL2G_01780		NFIA_022970	Afu1g01660	1	6	17
EXG	beta 1,6 glucanase	GH 5	AN3777		ATEG_09844	AO090011000757	AFL2G_05464		NFIA_084850	Afu2g09350	1	6	17
BGL	beta glucosidase	GH 1	AN9183	131747	ATEG_02657	AO090120000075	AFL2G_08111		NFIA_099670	Afu8g06970	1	7	14
EGL	endoglucanase	GH 5		214608	ATEG_04390	AO090011000715	AFL2G_05447	ACLA_081310	NFIA_085010	Afu2g09520	1	7	14
AGL	alpha galactosidase	GH 27	AN0022	172232	ATEG_03427	AO090005000217	AFL2G_00225		NFIA_023390	Afu1g01200	1	7	14
ABN	endoarabinanase	GH 43		182100	ATEG_07817	AO090026000804	AFL2G_06513	ACLA_072730	NFIA_089980	Afu2g14750	1	7	14
EXG	exo 1,3 galactanase	GH 5	AN4052	202490	ATEG_03849	AO090003000990	AFL2G_02039	ACLA_031040	NFIA_021060	Afu1g03600	1	8	13
AGT	4 alpha glucanotransferase	GH 13	AN3308	188489	ATEG_03623	AO090003001498	AFL2G_01594	ACLA_063440	NFIA_001710	Afu3g00900	1	8	13
AGD	alpha glucosidase	GH 31	AN0280	55419	ATEG_02528	AO090005000767	AFL2G_00750	ACLA_031260	NFIA_021450	Afu1g03140	1	8	13
AGL	alpha galactosidase	GH 36	AN8138	212736	ATEG_07529	AO090010000684	AFL2G_11824	ACLA_044620	NFIA_094580	Afu8g01130	1	8	13
BXL	beta xylosidase	GH 43	AN7864	47677	ATEG_01292	AO090012000331	AFL2G_09750	ACLA_072000	NFIA_088370	Afu2g13190	1	8	13
MND	beta mannosidase	GH 2	AN1742	138876	ATEG_06636	AO090001000556	AFL2G_09201	ACLA_083570	NFIA_054490	Afu6g08840	1	8	13

Chapter 3

Analysis of the molecular basis for the aberrant phenotype of *Aspergillus vadensis* compared to other black Aspergilli

Culleton, H.M., Zhou, M., Kowalczyk, J., Patyshakuliyeva, A., de Lely, P., Czech, A.S., van Bers, N.E.M., de Haan, J.R., McKie, V.A. and de Vries, R.P. 2014.
In preparation.

Abstract

In the recent past, much research has been applied to the development of *Aspergillus*, most notably *A. niger* and *A. oryzae*, as hosts for recombinant protein production. Recently *A. vadensis*, a close relative of *A. niger*, has been suggested as a suitable and more favourable alternative for recombinant protein production as it does not acidify the culture medium and produces very low levels of extracellular proteases. Generated growth profiles of this species on a variety of different carbon sources revealed further phenotypical differences, with *A. vadensis* possessing a distinct inability to degrade maltose or starch compared to the other black Aspergilli. In this study we examined the molecular differences between *A. vadensis* and six other species of black Aspergilli with a particular focus on the *prtT* and *amyR* Zn2Cys6 regulatory gene loci on chromosome VI to examine if the lack of protease and amylase gene expression is due to a mutation/deletion in this area.

Introduction

The black Aspergilli (*Aspergillus* section *Nigri*) form an important subgroup of the genus *Aspergillus,* of which *A. vadensis* is a recently identified and novel member. Over the past 25 years, much research has been dedicated to the development of *Aspergillus*, most notably *A. niger* and *A. oryzae*, as hosts for the industrial production of recombinant proteins. To date however, yields for heterologous protein production still remain low due in part, to the acidification of the media due to the production of organic acids in the case of *A. niger* but also due to the presence of high levels of secreted proteases which effectively degrade many heterologous proteins [33]. Despite considerable progress being made in the development of host strains with reduced levels of secreted proteases [5, 21], these issues remain of significant importance in the area of fungal strain development. As a result, much research has been applied to the identification and assessment of the regulatory pathways involved in the induction of protease gene expression [13, 14, 16, 31]. Based on the available genome sequences from various *Aspergillus* species, the number of protease genes found in the fungal genome was between 100-200 [23], the regulation of which was revealed to be very complicated and interrelated with the global carbon, nitrogen and pH regulatory circuits [5, 15, 16]. Recently, however, Punt *et. al.* (2008) identified the PrtT regulatory protein, a homologue of which was found to be present in many *Aspergillus* species and which was shown to govern expression of many secreted protease genes, including the major alkaline protease *alpA* and neutral protease *Np1* in *A. oryzae* and the aspergillopepsin encoding gene *pepA* in *A. niger* [25]. Furthermore, in all *Aspergillus* species carrying a *prtT* orthologue, a gene cluster related to starch degradation, including the Zn2Cys6 regulatory gene *amyR*, is present directly upstream of the *prtT* locus [25].

A. vadensis, a close relative of *A. niger*, has been suggested as a suitable and more favourable alternative to both *A. niger* and *A. oryzae* for recombinant protein production as it does not acidify the culture medium and produces very low levels of extracellular proteases thus facilitating heterologous protein production and many downstream processes [9]. To date, the use of His-tag affinity chromatography in fungal systems has not been either efficient or practical in fungal systems due to the difficulties experienced with degradation of the Histidine residues by extracellular proteases. Purification of recombinant enzymes from *A. vadensis* through His-tag affinity chromatography has been shown to be successful (Culleton et al. 2014), suggesting its potential as a versatile host in the fundamental research of proteins and for industrial enzyme production.

In this study we examined the molecular and phenotypical differences between *A. vadensis* and six other species of black Aspergilli i.e. *A. niger*, *A. acidus*, *A. aculeatus*, *A. carbonarius*, *A. brasiliensis* and *A. tubingensis*. We analysed the generated growth

profiles of the individual species on a variety of carbon sources to determine if *A. vadensis* has a similar ability for plant biomass degradation as the other black Aspergilli or if its unique phenotype exceeds its protease deficiency. Finally, we examine the *prtT* and *amyR* Zn2Cys6 regulatory gene loci on chromosome VI to determine if the lack of protease expression is due to a mutation/deletion in this area.

Materials and Methods

Strains, media and culture conditions

A. vadensis CBS 113365, *A. niger* ATCC 1015, *A. acidus* CBS 106.47, *A. aculeatus* CBS 172.66, *A. carbonarius* ATCC MYA-4641, *A. brasiliensis* CBS 101740 and *A. tubingensis* CBS 134.48 were used in the growth profiles of this study. *A. vadensis* CBS 137441 and *A. niger* N402 [4] were used for all remaining experiments. All strains were taken from glycerol stocks stored at -45°C and grown on MEA (Malt Extract Agar) prior to use. All plates were grown at 30°C.

Aspergillus minimal medium (MM) and complete medium (CM) were described previously [9]. Agar was added at 2% (w/v) for solid medium. All monomeric and oligomeric carbon sources were added to a final concentration of 25 mM, while pure polymeric substrates were added to a final concentration of 1%. The pH of the medium was adjusted to 6.0. For plate growth, the centre of the plates was inoculated with 2 µl of a suspension of 500 spores/µl and plates were incubated for 5 days. All species were grown on MM with 13 different carbon sources including monosaccharides (glucose, fructose, mannose and xylose), oligosaccharides (cellobiose, maltose and sucrose), pure plant polysaccharides (starch, inulin, beechwood xylan, apple pectin and galactomannan) and phosphoprotein substrate casein. Growth on 25 mM D-glucose was used as a reference for growth rate and growth on the other substrates relative to growth on glucose was then compared among the species. Growth on plates was then analyzed by visual inspection after incubation at 30°C for 5 days. For the analysis of enzyme activities, 400 mL liquid cultures (2.5 L baffled flasks) of MM + 1% (w/v) maltose / 1% (w/v) soluble starch were inoculated in duplicate with 1 x 10^6 spores/mL (final) and incubated at 30°C in an orbital shaker at 250 rpm.

For gene expression studies, *A. vadensis* was pre-grown in CM for 16 h at 30°C and 250 rpm after which the mycelium was harvest over a büchner funnel without suction and washed with MM without carbon source. Aliquots (approx. 1 g wet weight) of the mycelium were transferred to 50 ml MM with either 25 mM sucrose or 25 mM maltose and incubated for 2 h at 30°C and 250 rpm after which mycelium was harvested and frozen immediately in liquid nitrogen.

Enzyme profile comparison by SDS-PAGE

400 ml liquid cultures were inoculated with 1 x 10^6 spores/mL (final concentration) and were grown in duplicate in 2.5 L baffled flasks at 30°C and 250 rpm for 3 days, taking samples on a daily basis. The mycelia were removed from the culture filtrate samples by filtration over a Büchner funnel with nylon gauze. SDS-PAGE was done using a 12% polyacrylamide gel containing 0.1% (w/v) SDS with 15 µL culture filtrate samples being loaded for each expression strain. Protein bands were detected by Coomassie Blue staining [18].

Enzyme activity measurement

For the measurement of enzyme activities from the *A. vadensis* and *A. niger* liquid cultures, an initial 1:10 dilution of the culture filtrates was carried out in 100 mM sodium acetate buffer, pH 4.5 including BSA (1 mg/mL). Para-nitrophenol (*p*NP) assays were used for the measurement of *α*-/*β*-galactosidase and *α*-/*β*-glucosidase activities. Initial *α*-/*β*-galactosidase assays were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 10 mM *p*-nitrophenyl *α*-D-galactopyranoside (pH 4.5) / *β*-D-galactopyranoside (pH 4.5) respectively and incubated at 40 °C for 60 min. Similarly, for the initial measurement of *α*-/*β*-glucosidase activities, 0.2 mL of diluted enzyme was added in duplicate to 0.2 mL of 10 mM *p*-nitrophenyl *α*-D-glucopyranoside (pH 4.5) / *β*-D-glucopyranoside (pH 4.5) respectively and incubated at 40 °C for 60 min. The reactions were stopped with the addition of 3 mL of 2 % (w/v) tri-sodium orthophosphate, pH 12.0 and the absorbance was measured at 400 nm. Activities were expressed as Units/ml where one unit is defined as 1 micromole of *p*-nitrophenol liberated per minute per millilitre of culture filtrate. Depending on results obtained, further time-points were conducted to ensure that activity rates were linear.

Para-nitrophenol (*p*NP) assays were also used for the measurement of *α*-/*β*-amylase activities. Initial *α*-amylase assays were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 10 mM blocked *p*-nitrophenyl-*α*-D-maltoheptaoside + 2.5 U (final) thermostable α-glucosidase (pH 4.5) and incubated at 40°C for 60 min. Initial *β*-amylase activities were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 10 mM blocked *p*-nitrophenyl-*β*-D-maltotrioside + 1.0 U (final) thermostable *β*-glucosidase (pH 4.5) and incubated at 40°C for 60 min. The reactions were stopped with the addition of 3 mL of 2 % (w/v) tri-sodium orthophosphate, pH 12.0 and the absorbance was measured at 400 nm. Enzyme blanks were prepared by adding Stopping Solution to enzyme dilution before the addition of substrate. Activities were expressed as Units/ml where one unit is defined as 1 micromole of *p*-nitrophenol liberated per minute per millilitre of culture

filtrate. Depending on results obtained, further time-points were conducted to ensure that activity rates were linear.

The Megazyme D-Glucose (glucose oxidase/peroxidase; GOPOD) assay kit was used for the measurement of glucoamylase activities using soluble starch (Sigma; Cat. No. S-9765) as the substrate. Initial glucoamylase assays were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.5 mL of 10 mg/mL (final) soluble starch (pH 4.5) and incubated at 40°C for 60 min. The reactions were terminated by placing in a boiling water bath for 2 min. All samples were then treated with 3.0 mL of GOPOD reagent (glucose oxidase, peroxidase and 4-aminoantipyrine (pH 7.4) from Megazyme; Cat. No. KC-GLUC1 + KC-GLUC2) and incubated at 40°C for 20 min before absorbance was read at 510 nm. Enzyme blanks were prepared by boiling enzyme dilution before the addition of substrate. Sugar standards were prepared in quadruplicate by adding 50 µL of 1.0 mg/mL D-glucose (Megazyme; Cat. No. KC-GLUC3) and 150 µL of 100 mM sodium acetate buffer (pH 4.5) to 0.2 mL of 10 mg/mL (final) soluble starch (pH 4.5) and treating as enzyme reactions (above). Activities were expressed as Units/ml where one unit is defined as 1 micromole of glucose liberated per minute per millilitre of culture filtrate. Depending on results obtained, further time-points were conducted to ensure that activity rates were linear.

Genome mining

The genome of *A. vadensis* CBS 137441 was sampled using both short Illumina GAII reads and long PacBio RS reads for the purpose of hybrid assembly by ServiceXS B.V. The Illumina reads were filtered based on their sequence quality by the Q25 phred score (corresponding to a chance of one error in 333 bases). Only reads longer than 45bp after filtering were kept. The filtering was done by the in-house developed Fastq Filter v.2.05 from ServiceXS. The PacBio long read data was generated from two SMRT cells by one 90 minute movie run and one 45 minute run. PacBio SMRT Portal v1.4 was used to process the raw data. Adapter sequences were removed from the reads, which were then split into sub-reads and filtered by overall quality cut-off value of 0.75 and sequence length 50bps. The first attempt of hybrid assembly was performed by PacBio's AHA hybrid assembler and scaffolder. Cerulean v1.0 [10] was then used to merge assembled short read contigs from Abyss [28] (k-mer size of 3) and PacBio long reads. The PacBio reads were not error-corrected and the input file consisted of the data from two runs. In-house scripts were written at the end, to combine the two versions of assemblies. Augustus v3.0.2 [29] and Genemark v2.2a [20] were used to call the potential gene models from the final assembly.

Genome comparison was performed by Mugsy v.1.2.2 [2] and visualized by gmaj [3] and IGV [32]. Comparative genomics analysis of *A. vadensis* and other Aspergilli was done by OrthoMCL [12] with option 1e-20, inflation size 1, and sequence

coverage 60%. Multiple gene alignment was done by BLASTP [1] and Muscle [11], and was visualized by ClustalOmega [27]. Cross-species gene clustering was performed by MultiGeneBlast [22].

RNA isolation and gene expression analysis

Molecular biology methods were performed according to standard procedures [26], unless stated otherwise. Frozen mycelium was ground using a TissueLyser II (Qiagen). RNA was extracted using TRIzol reagent (Invitrogen) [7] and purified with NucleoSpin RNA II (Macherey-Nagel) with DNase treatment. The quantity and quality of RNA samples were checked by NanoDrop-1000 spectrophotometer and by RNA gel electrophoresis. Total RNA in amount of 0.25 µg was converted into cDNA (ThermoScript^{TM} RT-PCR System, Invitrogen) according to the instructions of the manufacturer and the obtained cDNA was diluted 10x and used in qPCR reaction.

Gene expression was assayed by real-time qPCR (Applied Biosystems 7500 Real-time PCR system) using ABI Fast SYBR Master Mix (Applied Biosystems) according to the supplier's instructions. The primers were designed using Primer Express® 3.0 software (Applied Biosystems), optimized and had amplification efficiencies > 97.5 and <102.7% (Supplemental Table 1). Expression levels were normalized against the histon H2B gene expression and calculated according to relative quantification $2^{-\Delta CT}$ method [19].

Results and Discussion

Comparative growth profiles on different carbon sources

The growth profiles of the different *Aspergillus* species showed notable differences between *A. vadensis* and some of the other black Aspergilli (Fig. 1; www.fung-growth.org). All strains grew well on MM + glucose and glucose was therefore used as an internal reference to compare the strains, to avoid misleading differences caused by general differences in growth speed between the species. Growth on the other substrates relative to growth on glucose was then compared between the species. All species had similar growth on the monosaccharides glucose, fructose, mannose and xylose, with the exception of *A. carbonarius*, which was slow to grow on all carbon sources, with especially poor growth on xylose. Growth on cellobiose and sucrose was similar to growth on glucose for all strains, while growth on the pure plant polysaccharides exceeded that generated on the simple sugars in all cases. The most notable growth differences for *A. vadensis*, were observed on maltose and starch, where no growth was observed on either carbon source. This was significantly lower than for the other Aspergilli, with the exception of *A. carbonarius*, which had poor growth on maltose and no growth on starch. The genes involved in maltose and starch degradation, such as glucoamylases (*glaA*), *α*-glucosidases (*agdA*) and to a lesser

extent α-amylase (*amyA*, *amyB*) are all regulated by the amylolytic regulator, AmyR [24, 30]. The unusual growth profile for *A. vadensis* when grown on these carbon sources could be explained by a mutation/deletion of the *amyR* regulatory gene. Interestingly, the *prtT* locus is found close to the *amyR* locus, with only two genes

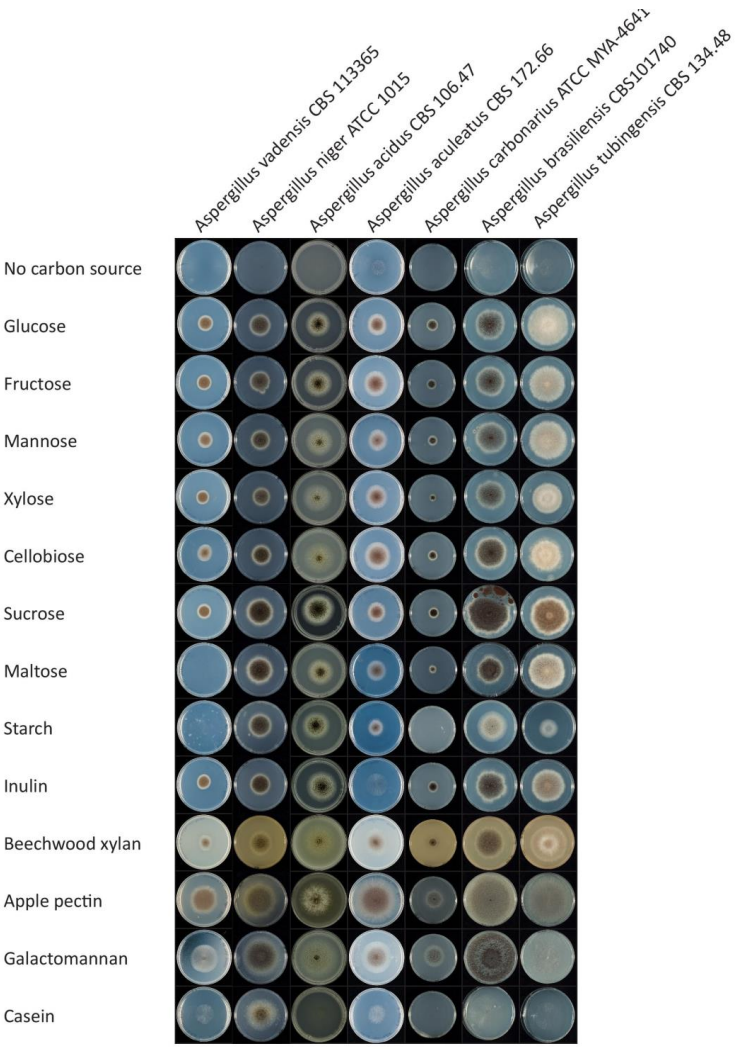


Figure 1: Growth profiles of *A. vadensis* compared to six other black Aspergilli when grown on 13 different carbon sources. All strains were grown on *Aspergillus* Minimal Medium (MM) + 25 mM / 1% carbon sources, pH 6.0 as indicated in Materials and Methods and incubated at 30°C for 5 days.

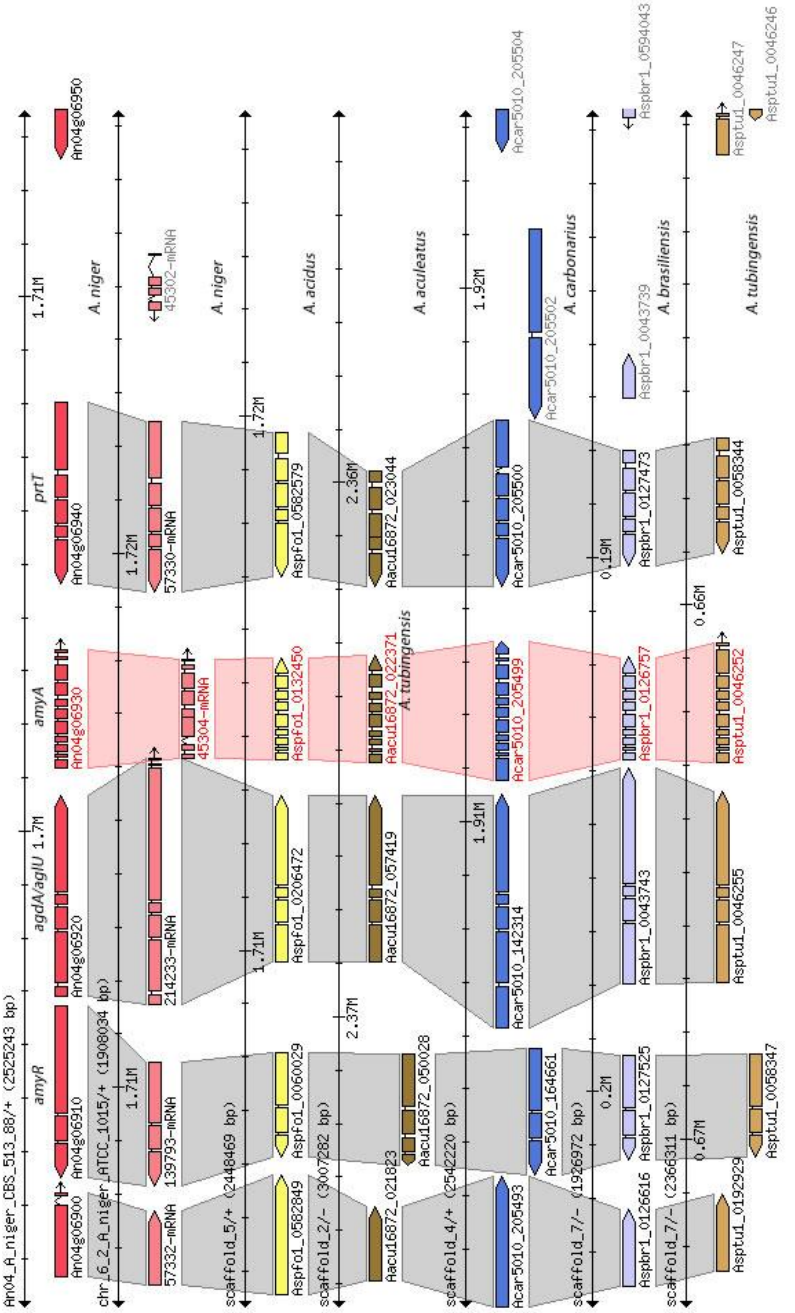


Figure 2: Synergy of the *prtT/amyR* region among the black Aspergilli.

(*agdA/aglU* and *amyA*) between the two regulators, both of which are target genes of AmyR in *A.niger*. Further analysis of the *prtT*/*amyR* region of the genome revealed high conservation in the black Aspergilli (Fig. 2), as supported by the efficient maltose/starch degradation abilities observed for the other black Aspergilli. Growth on casein however, was comparably poor for most the species tested, with the exception of *A. niger*, indicating that this more common protease deficiency may be accounted for at a transcriptional level.

Extracellular enzyme profile and activity comparison

Liquid cultures of *A. vadensis* and *A. niger* were performed to compare enzyme production levels for both species during growth on maltose and starch. Culture filtrate samples were taken on a daily basis to monitor extracellular enzyme production and activities, with Day 3 giving optimum production as determined by SDS-PAGE (Fig. 3). *A. niger* demonstrated a clear increase in enzyme production levels over the 3 days when grown on both maltose and soluble starch compared to *A. vadensis*, which failed to show any detectable protein bands over this incubation period. This was in correlation with the growth for both of these species, with *A. niger* generating more dense mycelia in both maltose and starch liquid cultures compared to *A. vadensis*. Further evaluation of the enzyme profiles in *A. niger* showed an intense band at approximately 80 kDa, which likely corresponds to the extracellular glucoamylase (GlaA), for which the calculated molecular mass in *A. niger* CBS 513.88 (An03g06550) is ~70 kDa [6], but for which molecular weights of between 70-91 kDa in other strains have been reported [17]. The other faint protein band produced by *A. niger* on both maltose and starch is consequently expected to be *α*-glucosidase (AgdA), for which a calculated molecular mass in *A. niger* CBS 513.88 (An04g06920) is ~110 kDa [6].

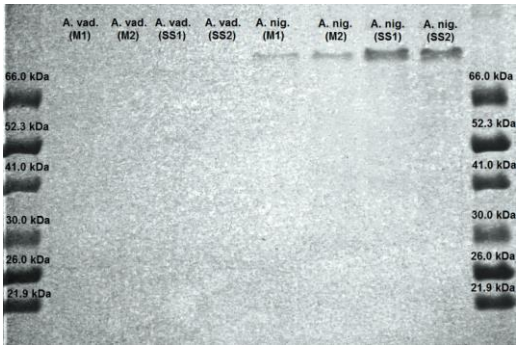


Figure 3: SDS-PAGE profile of *A. vadensis* (CBS 137441) and *A. niger* (N402) culture filtrate samples after 3 days of growth. M = 1% (w/v) maltose; SS = 1% (w/v) soluble starch. Cultures were grown in biological duplicates (1 & 2). Low molecular weight (LMW) ladder produced in house.

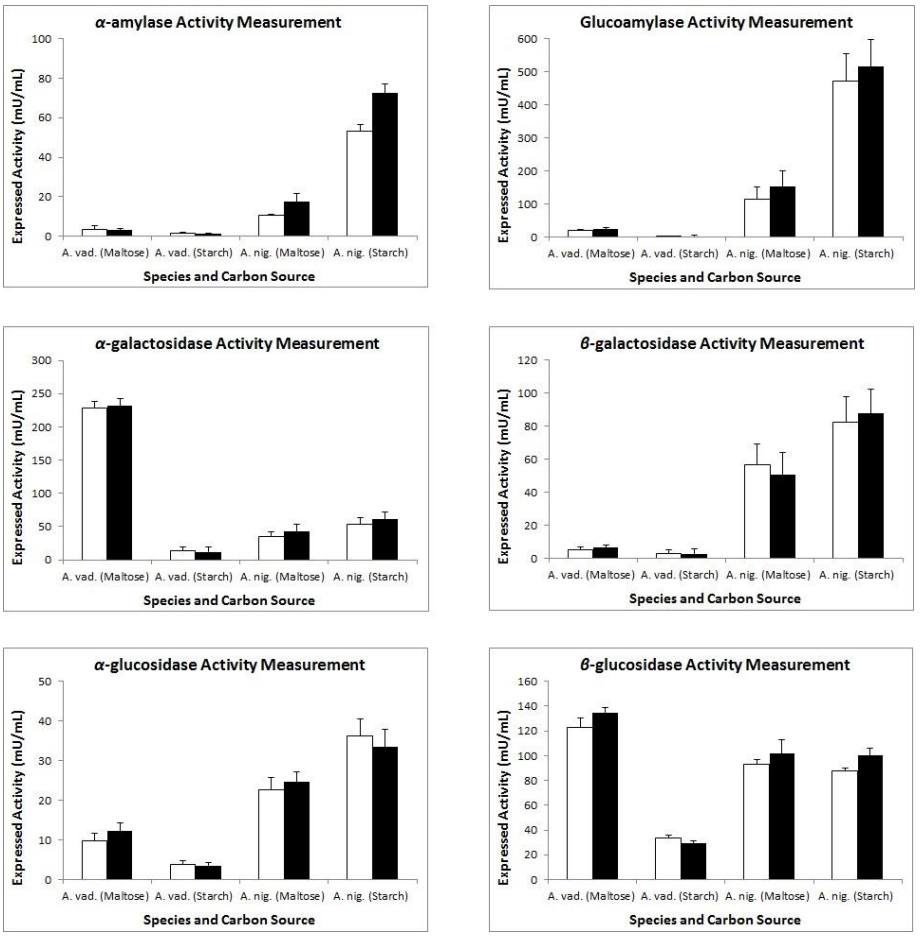


Figure 4: Measured activity values obtained from Day 3 culture filtrate samples from the indicated *A. vadensis* and *A. niger* strains. Error bars calculated on standard deviations between technical duplicates (White/Black). All polysaccharide substrates and *p*-nitrophenol substrates were tested at 40°C, pH 4.5 under conditions as described in Materials and Methods.

In combination with the visual profiles of extracellular enzyme production during growth on maltose/soluble starch, α-amylase, glucoamylase and α-glucosidase activities were measured and compared for both strains. The activities of α- and β-galactosidase and β-glucosidase were also measured as a control and indicative of the level of change in production of un-associated enzyme activities between both species. In correlation to what was discovered with the enzyme profiles, the overall levels of maltose/starch associated enzyme activities were much higher for *A. niger* than for *A. vadensis* on both carbon sources (Fig. 4). This was especially the case for glucoamylase with *A. niger* producing ~6 times the levels of *A. vadensis* on maltose (134 mU/mL vs. 21 mU/L) and ~200 times the levels on soluble starch (494 mU/mL vs. 2mU/mL). This again was in correlation with the growth for both of these species, with *A. niger* generating more dense mycelia in both maltose and especially starch liquid cultures compared to *A. vadensis*, demonstrating a greater degradation and utilisation potential for both carbon sources. Interestingly and in contrast to the general activity profile trend observed, the levels of α-galactosidase and β-glucosidase were higher for *A. vadensis* on maltose but not on soluble starch. Maltose is comprised of α-1,4-linked glucose molecules but is not highly branched with α-1,6-linked subunits as is the case with starch so only requires the action of α-glucosidase to hydrolyse it into glucose molecules. As indicated by the recorded enzyme activities (Fig. 3), some α-glucosidase expression was evident in *A. vadensis* cultures, especially when grown on maltose.

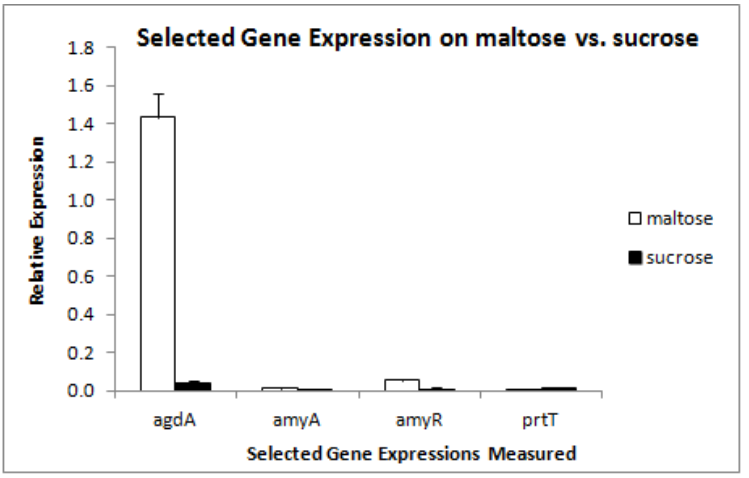


Figure 5: Measured gene expression levels for *agdA*, *amyA*, *amyR* and *ptrT* when grown on maltose and the control substrate, sucrose.

Maltose is degraded by the enzymatic action of this enzyme to give glucose molecules. AmyR controls genes involved in starch degradation but it has been reported for *A. niger* that it also regulates other genes, including those encoding α-galactosidases and β-glucosidases [34]. While starch and maltose are the most common inducers of this regulator, glucose also induces AmyR activation in *A. niger* [34]. In *A. niger* fructose is a relatively weak repressing carbon source [8], while glucose at higher concentrations can cause significant repression of many genes encoding poly- and oligosaccharide degrading enzymes. The increased levels of α-galactosidase and β-glucosidase on maltose for *A. vadensis* could be the result of regulation from a different regulator to AmyR that does not affect amylase gene expression.

Aspergillus vadensis genome and gene expression analysis

Liquid cultures of *A. vadensis* were performed on maltose and sucrose to compare gene expression levels for *amyR* and *prtT* as well as for *agdA* and *amyA* that are lying in between them on the genome and are regulated by *amyR* (Figure 5). Sucrose was used as a control, because it supports good growth of *A. vadensis* and *amyR* and its associated genes should not be up-regulated on this carbon source. In correlation to what was observed in both the growth profiles and liquid expression studies, *amyR*, *prtT* and *amyA* are all very poorly expressed on maltose, giving signals comparable to that observed on sucrose. Interestingly, *agdA* is expressed specifically on maltose and this correlates with the observed α-glucosidase activities observed in *A. vadensis* when grown on maltose (Fig. 3).

Analysis of the *A. vadensis* genome demonstrated that the region of +/-15k bp around *prtT* is well preserved, with no missing or additional genes and no large insertions of non-coding regions. The amino acid sequences of PrtT and AmyR are highly similar to those of the other black Aspergilli and do not contain internal STOP codons (data not shown). This combined data suggests that the phenotype of *A. vadensis* is therefore most likely caused by the low expression of *prtT* and *amyR* and not a mutation or deletion of these regulators. The cause of this low expression however, remains unclear at this point, but this is likely due to a dysfunctional signal transduction, activation or sensing mechanism.

References

1. **Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.** 1990. Basic local alignment search tool. Journal of molecular biology. **215**: 403-410.
2. **Angiuoli, S.V. and Salzberg, S.L.** 2011. Mugsy: fast multiple alignment of closely related whole genomes. Bioinformatics. **27**: 334-342.
3. **Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., et al.** 2004. Aligning multiple genomic sequences with the threaded blockset aligner. Genome research. **14**: 708-715.
4. **Bos, C.J., Debets, A.J., Swart, K., Huybers, A., Kobus, G. and Slakhorst, S.M.** 1988. Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet. **14**: 437-443.
5. **Braaksma, M. and Punt, P.J.** 2008. *Aspergillus* as a cell factory for protein production: controlling protease activity in fungal production. In: Goldman GH, Osmani, S.A., (ed) The Aspergilli: Genomics, Medial Aspects, Biotechnology and Research Methods, CRC Press, Boca Raton, pp. 441–456.
6. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.
7. **de Vries, R.P., Flipphi, M.J., Witteveen, C.F. and Visser, J.** 1994. Characterization of an *Aspergillus nidulans* L-arabitol dehydrogenase mutant. FEMS Microbiol Lett. **123**: 83-90.
8. **de Vries, R.P., vanKuyk, P.A., Kester, H.C. and Visser, J.** 2002. The *Aspergillus niger faeB* gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. Biochem J. **363**: 377-386.
9. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
10. **Deshpande, V., Fung, E.D., Pham, S. and Bafna, V.** 2013. Cerulean: A hybrid assembly using high throughput short and long reads. Algorithms in Bioinformatics, Springer, pp. 349-363.
11. **Edgar, R.C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research. **32**: 1792-1797.
12. **Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., et al.** 2011. Using OrthoMCL to Assign Proteins to OrthoMCL-DB Groups or to Cluster

Proteomes Into New Ortholog Groups. Current Protocols in Bioinformatics. 6.12. 11-16.12. 19.

13. **Katz, M.E., Flynn, P.K. and Cheetham, B.F.** 1996. Mutations affecting extracellular protease production in the filamentous fungus *Aspergillus nidulans*. Molecular and General Genetics MGG. **250**: 715-724.

14. **Katz, M.E., Masoumi, A., Burrows, S.R., Shirtliff, C.G. and Cheetham, B.F.** 2000. The *Aspergillus nidulans xprF* gene encodes a hexokinase-like protein involved in the regulation of extracellular proteases. Genetics. **156**: 1559-1571.

15. **Katz, M.E., Gray, K.-A. and Cheetham, B.F.** 2006. The *Aspergillus nidulans xprG* (*phoG*) gene encodes a putative transcriptional activator involved in the response to nutrient limitation. Fungal Genetics and Biology. **43**: 190-199.

16. **Katz, M.E., Bernardo, S.M. and Cheetham, B.F.** 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. Current genetics. **54**: 47-55.

17. **Kumar, P. and Satyanarayana, T.** 2009. Microbial glucoamylases: characteristics and applications. Critical reviews in biotechnology. **29**: 225-255.

18. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage *T4*. Nature. **227**: 680-685.

19. **Livak, K.J. and Schmittgen, T.D.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}$T method. Methods. **25**: 402–408.

20. **Lomsadze, A., Burns, P.D. and Borodovsky, M.** 2014. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. Nucleic Acids Res. **42**: 2.

21. **Mattern, I.E., van Noort, J.M., van den Berg, P., Archer, D.B., Roberts, I.N. and van den Hondel, C.A.** 1992. Isolation and characterization of mutants of *Aspergillus niger* deficient in extracellular proteases. Molecular and General Genetics MGG. **234**: 332-336.

22. **Medema, M.H., Takano, E. and Breitling, R.** 2013. Detecting sequence homology at the gene cluster level with MultiGeneBlast. Mol Biol Evol. **30**: 1218-1223.

23. **Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., et al.** 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol. **25**: 221-231.

24. **Petersen, K.L., Lehmbeck, J. and Christensen, T.** 1999. A new transcriptional activator for amylase genes in *Aspergillus*. Mol Gen Genet. **262**: 668-676.

25. **Punt, P.J., Schuren, F.H.J., Lehmbeck, J., Christensen, T., Hjort, C. and van den Hondel, C.A.M.J.J.** 2008. Characterization of the *Aspergillus niger* PrtT, a

unique regulator of extracellular protease encoding genes. Fungal Genetics and Biology. **45**: 1591-1599.
26. **Sambrook, J., Fritsch, E.F. and Maniatis, T.** Molecular cloning -a laboratory manual. 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y., 1989.
27. **Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., et al.** 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. **7**:
28. **Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J. and Birol, I.** 2009. ABySS: a parallel assembler for short read sequence data. Genome research. **19**: 1117-1123.
29. **Stanke, M., Schöffmann, O., Morgenstern, B. and Waack, S.** 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC bioinformatics. **7**: 62.
30. **Tani, S., Katsuyama, Y., Hayashi, T., Suzuki, H., Kato, M., Gomi, K., et al.** 2001. Characterization of the *amyR* gene encoding a transcriptional activator for the amylase genes in *Aspergillus nidulans*. Curr Genet. **39**: 10-15.
31. **te Biesebeke, R., van Biezen, N., De Vos, W., Van Den Hondel, C. and Punt, P.** 2005. Different control mechanisms regulate glucoamylase and protease gene transcription in *Aspergillus oryzae* in solid-state and submerged fermentation. Appl Microbiol Biotechnol. **67**: 75-82.
32. **Thorvaldsdóttir, H., Robinson, J.T. and Mesirov, J.P.** 2012. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics. bbs017.
33. **van den Hombergh, J.P.T.W., van de Vondervoort, P.J.I., Fraissinet-Tachet, L. and Visser, J.** 1997. *Aspergillus* as a host for heterologous protein production: the problem of proteases. Trends in Biotechnology. **15**: 256-263.
34. **vanKuyk, P.A., Benen, J.A., Wosten, H.A., Visser, J. and de Vries, R.P.** 2012. A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol. **93**: 285-293.

Supplemental Table 1: Primers used in this study.

Primers	**Sequence**	**Direction**	**Opt. Conc. (nM)**
agdA_F	TTGGCCGCTCAACTTTCG	Forward	900
agdA_R	ATTGTCACCGCCCCAGTGT	Reverse	900
amyA_F	CCGCAACATTCACCATCCTA	Forward	900
amyA_R	TGCTCCTGACCGGCGTAT	Reverse	900
amyR_F	TCCACCGGGAATCAAGTTATG	Forward	900
amyR_R	ACCCTGCGCCGTTCCT	Reverse	900
prtT_F	GCATCTCTCCCCCAACGA	Forward	300
prtT_R	GGCGACCATGCACGAAA	Reverse	300
H2B_F	GGTATCTCGACTCGCGCTATGT	Forward	300
H2B_R	TCGCGACACGCTCAAAGATAT	Reverse	900

Chapter 4

New promoters to improve heterologous protein production in *Aspergillus vadensis*

Culleton, H.M., Bouzid. O., McKie, V.A. and de Vries, R.P. 2014. Curr Biotechnol. **3**: 1-8.

Abstract

Aspergillus is a widely used host organism for the industrial production of homologous and heterologous proteins. Although *Aspergillus niger* is most commonly used, a close relative of this species, *Aspergillus vadensis*, has been suggested as a suitable and more favourable alternative due in part, to the low levels of extracellular proteases which it produces. Despite much progress being made in the hyper production of homologous proteins, the yields obtained for heterologous proteins have still not reached a comparable level. Genetic strategies, including the development of strong constitutive promoters have been shown to lead to an increase in the levels of recombinant protein production. In this study, six novel constitutive promoters from *A. niger* (p*ef1α*, p*tktA*, p*ef1β*, p*tal1*, p*cetA* and p*pgkA*) and a further five from *A. vadensis* (p*ef1α*, p*rps31*, p*gpdA*, p*ubi1* and p*oliC*) were tested in *A. vadensis* using a gene encoding a secreted arabinofuranosidase from *Fusarium oxysporum* as a reporter for heterologous protein production. Remarkably, 9 of the 11 promoter constructs tested all resulted in higher ABF activity than the commonly used *gpdA* promoter. While this could partly be assigned to the number of copies of the expression cassette in the transformants, clear differences in productivity of the promoters could be observed.

Introduction

Over the past 25 years, much research has been dedicated to the development of filamentous fungi, most notably *Aspergillus*, as hosts for the industrial production of recombinant proteins. Although many *Aspergillus* species have the capacity to grow at high rates and to high biomass densities, the protein secretion levels observed in *Aspergillus niger* and *Aspergillus oryzae* make them the most commonly used industrial hosts among the Aspergilli [7]. While significant progress has been made in the hyper-production of heterologous proteins, yields have not yet reached the levels obtained for homologous proteins [3]. This may in part, be caused by the acidification of the media due to the production of organic acids in the case of *A. niger* but also by the presence of high levels of secreted proteases which effectively degrade heterologous proteins [25]. *Aspergillus vadensis*, a close relative of *A. niger*, has been suggested as a suitable and more favourable alternative for recombinant protein production as it does not acidify the culture medium and produces very low levels of extracellular proteases [5]. In fact, this species has been shown to produce higher levels of *A. niger* FaeB than *A. niger* itself [1] and also produced a basidiomycete esterase that could not be produced in *A. niger* [6].

Studying gene expression in filamentous fungi has greatly increased our understanding of the molecular mechanisms controlling transcription initiation and/or regulation of homologous proteins. Such genetic strategies, including the use of strong homologous promoters, have been developed to improve the yields associated with heterologous protein production in fungi. To date, the set of promoters which have been described for recombinant gene expression in *A. niger* have included several highly inducible or strong constitutively active promoters with the focus being on those which have shown high expression levels of their target genes [9]. Inducible expression systems, of which the glucoamylase (*glaA*) promoter from *A. niger* [11], the TAKA-amylase A (*amyB*) promoter from *A. oryzae* [13] and the alcohol dehydrogenase (*alcA*) from *A. nidulans* [27] are the most commonly used, allow for high and low expression from the same construct but are usually dependent on strictly defined conditions and media requirements. Expression of the *glaA* gene for example, is repressed in the presence of xylose but is highly induced if maltose or starch is used as the single carbon source [11]. Although transcriptional control under such inducible promoters allows for the fine-tuned regulation of expression yields and subsequent high biomass formation, the necessity for certain media components or expensive inducers in conjunction with a finely tuned timing of induction can make it both inflexible and costly for some industrial applications. The more favoured option often involves the use of constitutive promoters e.g. the glyceraldehyde-3-phosphate dehydrogenase promoter (*gpdA*) from *A. nidulans* [22], the protein kinase A (*pkiA*) and the alcohol dehydrogenase (*adhA*) promoters from *A. niger* [23] and the glutamate

dehydrogenase promoter (*gdhA*) from *A. awamori* [19], which allow for the development of an expression cassette which is not media/component dependent. In recent studies, Blumhoff *et. al.* (2013) identified six novel constitutive promoters in *A. niger*, of which a promoter with a strong similarity to multiprotein bridging factor 1 (*Mbf1*) from *S. cerevisiae* (*PmbfA*) was shown to give higher expression levels of the heterologous *gusA* reporter gene than the *gpdA* promoter, which is the most common promoter of choice for heterologous gene expression in *Aspergillus* [8, 10, 17, 18].

To further improve protein production in *A. vadensis*, six novel constitutive promoters from *A. niger* and a further five from *A. vadensis* were selected and tested in *A. vadensis* in comparison to the *gpdA* promoter from *A. nidulans*. A gene encoding a secreted α-arabinofuranosidase from *Fusarium oxysporum* was used as a reporter for protein production.

Materials and Methods

Strains, media and culture conditions

Escherchia coli DH5α chemically competent strain (Invitrogen) was used as a host for recombinant DNA manipulation. Microarray analysis was carried out on RNA obtained from *A. niger* CBS 120.49 (N400) and *A. vadensis* Ref 7 and promoter constructs were obtained by PCR from the gDNA of *A. niger* N402 [2] and from *A. vadensis* Ref 7. *A. vadensis* CBS 113365 (*pyrG* -) was used as the parental strain for transformation. All other *A. vadensis* strains were derived from the parental strain and are listed in Table 1.

Aspergillus minimal medium (MM) and complete medium (CM) were described previously [5]. Agar was added at 2% (w/v) for solid medium. Pre-cultures for protoplast formation were grown overnight at 30°C in 200 mL MM supplemented with 0.5% (w/v) yeast extract, 0.2% (w/v) casamino acids, 2% (w/v) glucose and 1.22 mg/mL (final) uridine after inoculating with 5 x 10^6 spores/mL. For the analysis of promoter activities, 50 mL liquid cultures (250 mL flasks) of MM + sucrose (2%) were inoculated with 10^6 spores/mL (final) and incubated at 30°C in an orbital shaker at 250 rpm. For chromosomal DNA isolation, *A. niger* and *A. vadensis* transformants were cultivated in CM supplemented with 1% glucose.

Molecular biology methods

Standard methods were used for DNA manipulations, subcloning, DNA digestion reactions, DNA isolations and Southern Analysis [24]. Chromosomal DNA was isolated as previously described by de Graaff *et al.* 1988 [12].

Table 1: Strains used for this study. *A. niger* CBS 120.49 (N400) and *A. vadensis* CBS Ref 7 were used in microarray analysis to identify promoter constructs. gDNA from *A. niger* N402 and *A. vadensis* CBS Ref 7 were used in PCR reactions to obtain promoter constructs. *A. vadensis* CBS 113365 was used as the parental strain for transformation. All other *A. vadensis* strains marked with an asterisk (*) were derived from the parental strain as listed.

Species	**Strain**	**Description**	**Reference**
A. niger	CBS120.49		
A. niger	N402	*cspA1* mutant of CBS120.49	Bos e*t al*.,1988
A. vadensis	CBS113365	*pyrA5* mutant	de Vries e*t al.*, 2004
A. vadensis	Ref7	Reference strain	This study *
A. vadensis	FP-401	p*ef1α* of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-402	p*tkt*A of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-403	p*ef1β* of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-404	p*tal* of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-405	p*cetA* of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-406	p*pgk*A of *A. niger*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-407	p*gpdA* of *A. nidulans*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-450	p*ef1α* of *A. vadensis*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-451	p*rps31* of *A. vadensis* , *abf* of *F. oxysporum*	This study *
A. vadensis	FP-452	p*gpdA* of *A. vadensis*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-453	p*ubi1* of *A. vadensis*, *abf* of *F. oxysporum*	This study *
A. vadensis	FP-454	p*oliC* of *A. vadensis*, *abf* of *F. oxysporum*	This study *

Selection of constitutive promoters for A. niger and A. vadensis using microarray expression data

A. vadensis Ref 7 was grown in a 200 mL MM + sucrose (2%) culture at 30°C and 250 rpm. After 48, 72 and 96 hrs the mycelia were gently harvested by filtration over a Büchner funnel with nylon gauze, dried, frozen in liquid nitrogen and stored at -80°C. RNA for microarray analysis was extracted using TRIzol reagent (Invitrogen) and purified using TRIzol® Plus RNA Purification Kit (Sigma-Aldrich) according to the instructions of the manufacturer. The quality of the RNA was analysed with an Agilent 2100 bioanalyzer using an RNA6000 LabChip kit (Agilent Technology). Expression data was analysed using the Bioconductor Affy tool package (http://www.bioconductor.org) under the statistical environment R. The probe intensities were normalized by using the Robust Multiarray Average (RMA) algorithm for the Affimetrix. CEL files, quartiles algorithm was used to perform normalization and gene expression values were calculated by the medianpolish summary method

(Bolstad BM, unpublished) with only the perfect match (PM) probes. The normalized data was analysed manually for genes with a high expression value. These were then compared to an in-house dataset of expression data from *A. niger* to identify genes that were highly expressed in *A. vadensis* as well as under most growth conditions in *A. niger*. This resulted in a selection of five genes for *A. vadensis*. In addition, six *A. niger* genes that were highly expressed under all conditions of the in-house dataset were also selected.

PCR and expression vectors

Promoters of the selected genes were amplified from *A. niger* and *A. vadensis* genomic DNA using primers targeting regions 1-1.3Kb upstream of the genes (Supplemental Table 1). PCR's were carried out using AccuTaq™ from Sigma and conditions supplied. PCR products were inserted into pGEMT easy vector (Promega) and plasmids were verified by sequencing.

The pGPDGFP expression vector was used to test the individual promoters in this study. As described by Lagopodi *et al.* (2002) [16], the pGPDGFP vector is composed of the *gfp* gene under the control of the *gpdA* promoter [22] and terminated with the *trpC* terminator [20]. A new expression vector was built based on the pGPDGFP vector where *gfp* was replaced with the arabinofuranosidase gene (*abf*) from *Fusarium oxysporum* using restriction enzymes *NcoI/HindIII*. The *gpdA* promoter was then replaced with the new promoter candidates using restrictions enzymes *NotI/NcoI* (Fig. 1).

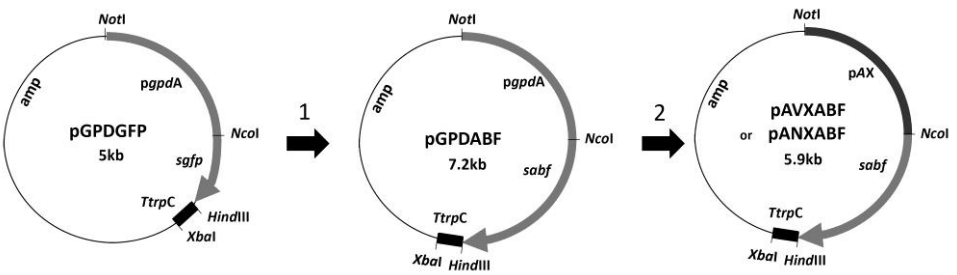


Figure 1: Expression plasmid construction steps. (1) Replacement of green fluorescence protein gene (sgfp) by arabinofuranosidase from *Fusarium oxysporum* (*sabf*). (2) Insertion of *A. niger* or *A. vadensis* selected promoters in place of glyceraldehydes-3-phosphate dehydrogenase promoter (*pgpdA*) from *A. nidulans*.

Transformation of A. vadensis and selection of expression strains

The formation of protoplasts by *Aspergillus* strains was based on the protocols by Peraza *et al.* (2003) [21] and de Bekker *et al.* (2009) [4]. Strains were grown for 16 hrs, after which time the mycelia were gently harvested by filtration over a Büchner funnel with nylon gauze. After washing with 0.9% NaCl (w/v), 2.5 g (wet weight) mycelium was resuspended in 20 mL stabilization buffer (0.2 M phosphate buffer (pH 6.0) and 0.8 M sorbitol). Lytic enzymes were added to the following final concentrations; 5mg/mL lysing enzymes from *Trichoderma harzianum* (L1412, Sigma), 460 units/mL β-glucuronidase from *Helix pomatia* (G0751, Sigma) and 0.15 units/mL chitinase from *Streptomyces griseus* (C6137, Sigma). The mixture was incubated in an orbital shaker at 37°C for 1 to 2 hrs with gentle shaking (120 rpm). Protoplasts were separated from the mycelium by filtering over glass wool. The protoplasts were recovered by centrifugation in a swing-out rotor (10 min 2,200 rpm) and were washed twice with STC (1.33 M sorbitol, 50 mM $CaCl_2$ and 10 mM Tris/HCL, pH 7.5). Transformation was performed as described by Kusters-van Someren *et al.* (1991) [15], with 2 x 10^6 protoplasts, 0.5 µg of pGW635 (carrying the *A. niger pyrG* gene for selection) and 20 µg of the different expression vectors (carrying *Fox-abf* under the control of the studied promoters).

To test the resulting transformants for *abfB* expression, a fluorimetric ABF screen based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-*α*-L-arabinopyranoside (0.02 mg/mL (final) in MM + 2% (w/v) sucrose + 2% (w/v) agar) when 1000 spores (2µL) from each transformant were grown at 30°C for 2 days and viewed under UV light. Based on fluorescence, two transformants per strain were brought forward for liquid expression studies.

α-Arabinofuranosidase activity measurement

Liquid cultures were inoculated with 1 x 10^6 spores/mL (final concentration) and grown at 30°C and 250 rpm with aliquot samples being taken on a daily basis for up to 5 days, centrifuged to remove mycelium and stored at -20°C. For the measurement of α-arabinofuranosidase (ABF) activity, culture filtrate samples (20 µL) were incubated with 10 µL of a 0.1% (w/v) *p*-nitrophenyl α-L-arabinofuranoside solution, 50 µL of a sodium acetate buffer (pH 5.0) and 20 µL water for 1 hr at 30°C in microtiter plates. The reaction was stopped by the addition of 100 µL 0.25 M bisodium carbonate and the optical density at 405 nm was measured on a microplate reader. Activities were expressed as nanomoles of *p*-nitrophenol liberated per minute per millilitre of culture filtrate.

Southern blotting

For each of the selected strains, 5 µg of chromosomal DNA was digested with *ApaI* and with *XhoI* for 6hr at 37°C. *ApaI* has two restriction sites in the *abf* gene of *F. oxysporum*, so was used as a target. *XhoI* has two restriction sites in the pectin lyase gene (*pel*) of *A. vadensis*, which is present in a single copy in the *A. vadensis* genome and was therefore used as a control. For Southern analysis, digested DNA was transferred to a Hybond-N+ membrane (Amersham Biosciences). The hybridization of the DIG-labelled probes was performed according to the DIG user's manual.

Results

Selection of constitutive promoters for A. niger and A. vadensis

Preliminary data demonstrated that during growth on MM with sucrose, very few proteins are produced by *A. vadensis* (data not shown), so this condition was selected for heterologous protein production. Microarray analysis from RNA isolated from sucrose grown mycelium was performed to identify highly expressed genes. Mycelium was harvested at 48, 72 and 96 hrs to select for genes that were continuously highly expressed. The resulting genes were then used to identify the orthologs from *A. niger* and the expression of these orthologs was then evaluated in an in-house micro array database for constitutively high expression. Based on this analysis five promoters from *A. vadensis* and a further six from *A. niger* were selected to be tested as promoters to drive gene expression and protein production (Table 2).

Table 2: Promoters tested in this study.

Source	Promoter	Acc/gene No.	Description
A. niger CBS513.88	p*ef1α*	An18g04840	Similar to elongation factor 1 alpha
	p*tktA*	An08g06570	Putative transketolase
	p*ef1β*	An08g03490	Elongation factor 1 beta
	p*tal1*	An07g03850	Putative transaldolase
	p*cetA*	An16g03330	Secreted thaumatin like protein
	p*pgkA*	An08g02260	Phosphoglycerate kinase
A. vadensis Ref 7	p*ef1α*	KJ420614	Similar to elongation factor 1 alpha
	p*rps31*	KJ420615	Similar to cytoplasmic ribosomal subunit
	p*gpdA*	KJ420616	Glyceraldehyde-3-phosphate dehydrogenase
	p*ubi1*	KJ420617	Similar to ubiquitin
	p*oliC*	KJ420618	Similar to mitochondrial ATP synthase

Development of expression strains

Of the genes identified from the microarrays, 1-1.3 kb fragments upstream of the coding regions were PCR amplified, inserted into pGEMT easy vector (Promega) and sent for sequencing. Sequencing results confirmed that the correct promoter regions were amplified. To test the efficiency of the selected promoters, the corresponding fragments were cloned into a vector containing an α-arabinofuranosidase encoding gene (*abf*) from *Fusarium oxysporum* as the reporter for protein production. This approach was adopted to enable us to determine the strength of the selected promoters by accurate measurement of the expression of a heterologous gene. Once cloned the expression vector containing the promoter sequence was then transformed into *A. vadensis* CBS113365 (Table 2). To test the resulting transformants for *abf* expression, a fluorimetric screen based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-α-L-arabinopyranoside was applied to select the strongest expressers for each strain (Fig. 2).

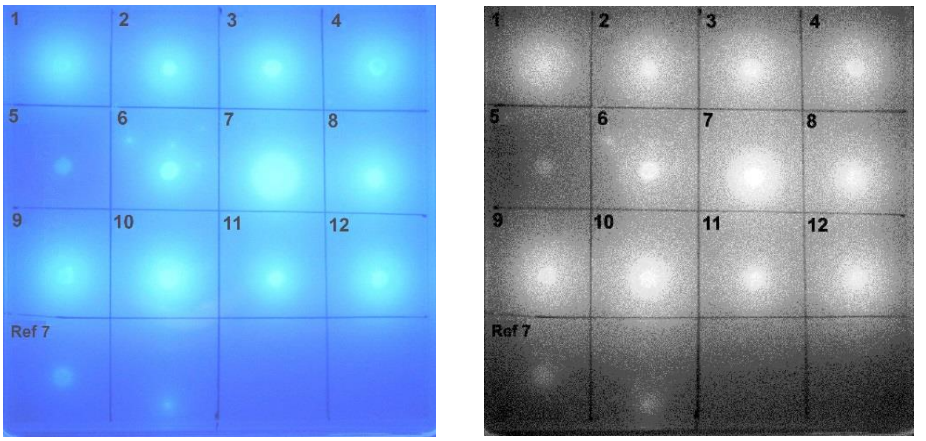


Figure 2: Fluorimetric screen to select the strongest ABF producers among transformants containing *pgpdA* from *A. nidulans*. The screen was based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-*α-L*- arabinopyranoside (0.02 mg/mL (final)) when 1000 spores (2μL) from each transformant were grown at 30°C for 2 days and viewed under UV light. CBS 137441 was grown as a negative control. In this instance colonies no. 7 and no. 10 gave the strongest fluorescence and hence were chosen for liquid expression studies. Fluorimetric screens were conducted in the same way for the other promoter constructs (Fig. not shown).

Promoter characterization using α-arabinofuranosidase (abf) from Fusarium oxysporum as a reporter for protein production.

Of the transformants screened for ABF expression, two different transformants i.e. biological duplicates (labelled 01 and 02 in Fig. 3) from both the *A. vadensis* and *A. niger* promoter strains were chosen for liquid expression studies and grown in duplicate, with samples being taken on a daily basis for up to 5 days. Arabinofuranosidase (ABF) activities were measured for all samples, with day 4 samples showing optimum activity for both *A. niger* and *A. vadensis* transformants and activities being compared to FP-407 strain which contained *abf* under the control of p*gpdA* from *A. nidulans* (Fig. 3).

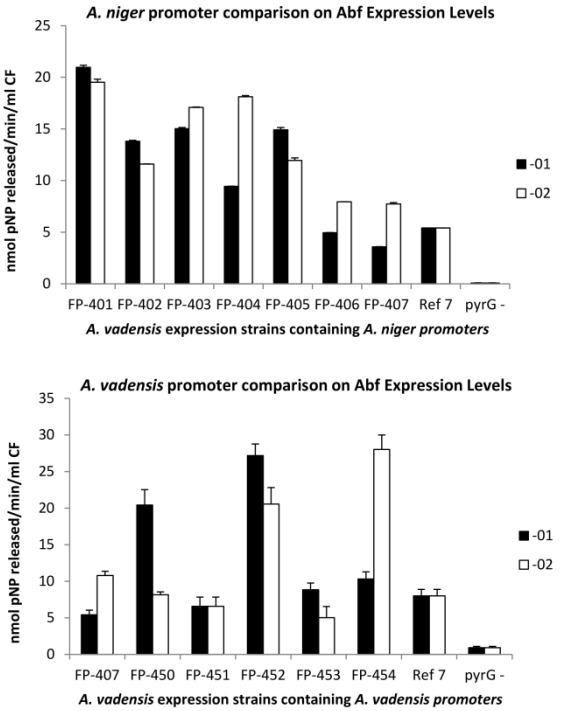


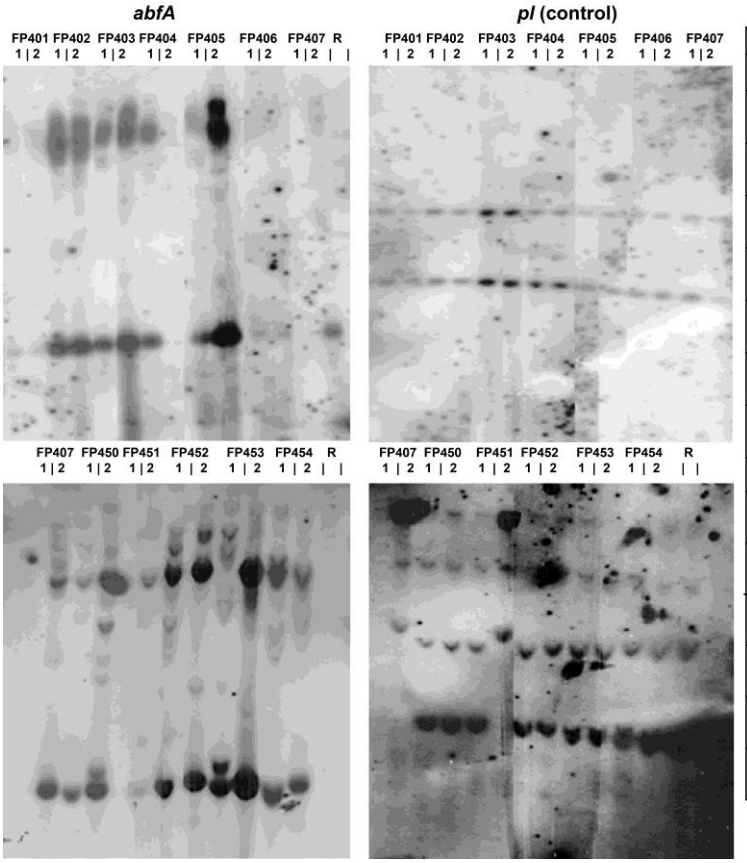
Figure 3: Comparison of arabinofuranosidase (ABF) activity levels in *A. vadensis* transformants strains expressing the heterologous *abf* of *Fusarium oxysporum* under the control of the selected promoters from both *A. niger* and *A. vadensis*. Cultures grown at 30°C and at 250 rpm, with culture filtrate samples being taken on day 4 (optimum) and assayed against *p*-nitrophenyl-α-L-arabinofuranoside substrate with absorbance measured at 405 nm. Activity levels expressed as nmol *p*-nitrophenol released per minute per mL of culture filtrate. FP-407 contains *abf* under the control of p*gpdA* from *A. nidulans*. *A. vadensis* CBS 137441 and CBS 113365 (*pyrG* -) strains were grown and assayed as negative controls.

ABF activity was detected for all transformants tested and was significantly above the basal level of the parental strains which were grown as negative controls (Fig. 3). Expression duplicates were measured and the average obtained to give a standard deviation for each transformant and strain. The *gpdA* promoter from *A. nidulans* (FP-407) was used as a positive control and comparative standard in both sets of expression studies. Remarkably, all 6 selected promoters from *A. niger* and 3 of the 5 promoters selected from *A. vadensis* resulted in higher ABF activity than that expressed under the *gpdA* promoter. In the case of the *A. niger* promoters, the activity levels obtained portray the order of p*ef1*α (FP-401) > p*ef1*β (FP-403) > p*tal* (FP-404) > p*cetA* (FP-405) > p*tktA* (FP-402) > p*pgkA* (FP-406) > p*gpdA* (FP-407). ABF activities of the p*ef1*α transformants were more than 4 times the ABF activity of the p*gpdA* transformants. The *A. vadensis* promoters resulted in ABF activity in the order of p*gpdA* (FP-452) > p*oliC* (FP-454) > p*ef1α* (FP-450) > p*gpdA* (FP-407) > p*ubi1* (FP-453) > p*rps31* (FP-451) where the highest producer also displayed over four times the ABF activity of the lowest. In this instance, p*gpdA* did not result in the lowest level of ABF activity, as the levels of the p*ubi1* and p*rps31* transformants were approximately 14% and 19% lower, respectively. The p*gpdA* (FP-407) control gave consistent activity levels in both sets of expression studies with p*ef1*α from both *A. niger* (FP-401) and *A. vadensis* (FP-450) also giving similar results in both sets of expression studies signifying the close relationship between the two species. The p*gpdA* promoter from *A. vadensis* itself (FP-452) gave approximately three times the level of ABF activity of the p*gpdA* promoter from *A. nidulans* indicating a possible preference between the two promoter constructs. The corresponding activity levels for the two different transformants of each promoter construct were generally quite comparable in both sets of expression studies with over half of the promoter constructs giving < 30% variance between biological duplicates, however, in some cases e.g. the p*gpdA* (FP-407) control presented with ~ 50% difference in monitored *abf* activity levels between both strains. An exceptional difference of 64% was also noted for p*oliC* (FP-454) in the *A. vadensis* expression study with one transformant generating 10.3 nmol *p*-nitrophenol released/min/mL and the other giving a value of 28 nmol *p*-nitrophenol released/min/mL.

Southern blots

To confirm the presence and determine the copy number of the promoter constructs in the genome of *A. vadensis*, the chromosomal DNA from selected transformants was digested with *ApaI* and *XhoI*. *ApaI* has two restriction sites in the *abf* gene of *F. oxysporum* at positions 534/530 and 1216/1212, which when separated generates two specific target bands. *XhoI* has two restriction sites in the pectin lyase gene (*pel*) of *A. vadensis* at positions 433/437 and 1152/1156, also generating two specific bands when

analysed. As mentioned above, the pectin lyase gene (*pel*) of *A. vadensis* is present in a single copy in the *A. vadensis* genome and was therefore used as a control to estimate the copy number of expression constructs present in each of the transformants (Fig. 4). The additional bands observed in the *A. vadensis* blot are likely due to heterologous hybridisation to other *pel* genes.



Transformants	Copy No.
A. niger	
FP401-01	1
FP401-02	1
FP402-01	10
FP402-02	10
FP403-01	6
FP403-02	10
FP404-01	8
FP404-02	1
FP405-01	10
FP405-02	20
FP406-01	1
FP406-02	1
FP407-01	1
FP407-02	3
A. vadensis	
FP407-01	1
FP407-02	3
FP450-01	2
FP450-02	10
FP451-01	1
FP451-02	2
FP452-01	8
FP452-02	10
FP453-01	8
FP453-02	15
FP454-01	6
FP454-02	4

Figure 4: Estimated copy number of expression constructs based on Southern analysis of *A. vadensis* transformants FP401 - FP407 (*A. niger* promoters) and FP450 - FP454 (*A. vadensis* promoters), using an *abf* probe fragment of *Fusarium oxysporum*. CBS 113365 (R) was used as a negative control. A fragment of the *pel* gene (encoding a pectin lyase) was used as an internal control for *A. vadensis.* Copy-numbers were determined by comparing the intensity of the *abf* signal to that of the *pel* signal.

Discussion

To date, many strategies including the use of strong homologous promoters have been applied for the development of filamentous fungi as hosts for the industrial production of recombinant proteins. In this study we identified six novel constitutive promoters from *A. niger*, namely p*ef1α*, p*tktA*, p*ef1β*, p*tal1*, p*cetA* and p*pgkA* and a further five from *A. vadensis*; p*ef1α,* p*rps31*, p*gpdA*, p*ubi1* and p*oliC*, and tested them in *A. vadensis* using a gene encoding a secreted α-arabinofuranosidase from *Fusarium oxysporum* as a reporter for heterologous protein production. In both sets of expression studies comparing the levels of ABF activity produced under firstly, the *A. niger* and secondly, the *A. vadensis* promoters, all *A. niger* promoters and 3 out of 5 *A. vadensis* promoters resulted in higher ABF activity levels than those observed under the control of the widely characterized *gpdA* promoter, which is currently the promoter of choice for recombinant gene expression in *Aspergillus* [8, 10, 17, 18]. In the case of the *A. niger* promoters, the highest expressing promoter, p*ef1α*, gave greater than 4 times the level of *abf* activity of p*gpdA*. Interestingly, when transformed into *A. vadensis*, the p*gpdA* promoter from *A. vadensis* itself gave approximately three times the level of ABF activity of the p*gpdA* promoter from *A. niger*, indicating the possibility of species differences and preferences between the two promoter constructs. The p*ef1α* from both *A. niger* (FP-401) and *A. vadensis* (FP-450) gave comparable results in both sets of expression studies, signifying not only the close relationship between the two species but also demonstrating that this gene is likely expressed similarly in both species as well.

Expression vectors for recombinant protein expression are generally integrated into the chromosomal DNA of *Aspergillus* by either homologous or non-homologous recombination [9]. Multiple insertions of up to and greater than 10 copies are common and are often found to increase recombinant protein production in *Aspergillus* [26], but this is not always the case possibly due to the pleiotropic effect of random integrations or the titration of endogenous transcription factors [14]. To gain a better understanding of the actual strength of the promoter constructs tested in this study, we felt it was necessary to determine the copy number of expression vectors transformed into the individual strains and see if the differences in ABF activity between biological duplicates and promoter constructs could be explained this way. In the *A. niger* expression study p*ef1α* (FP-401), p*pgkA* (FP-406) and p*gpdA* (FP-407) all contain low copy numbers of the expression vector compared to p*tktA* (FP-402) where the duplicates both contain ~10 copies and p*cetA* (FP-405) which has an average of ~15 copies. With this information the promoter strengths can now be put into the order of p*ef1α* (FP-401) > p*tal* (FP-404) >> p*pgkA* (FP-406) > p*gpdA* (FP-407) > p*ef1β* (FP-403) > p*cetA* (FP-405) > p*tktA* (FP-402) with p*ef1α* expressing a 10 fold increase in ABF activity per copy number compared to p*tktA*. The *A. vadensis* strains presented

much more variances in copy numbers between the biological duplicates and promoter constructs which goes a long way to explain the noteworthy differences in activity levels measured between transformant duplicates in this study. Taking copy number into account the *A. vadensis* promoter constructs can be written in the order of p*ef1α* (FP-450) > p*oliC* (FP-454) > p*rps31* (FP-451) > p*gpdA* (FP-407) > p*gpdA* (FP-452) > p*ubi1* (FP-453). Taking the new orders of strength into consideration we can see a different pattern forming than what was portrayed from the measurement of ABF activity alone. p*ef1α* is now not only performing comparably between both sets of expression study but is also the strongest expressing promoter of both the *A. niger* and *A. vadensis* promoters tested. p*gpdA* from *A. nidulans* also remains consistent as a control between both sets of data but now we see a much closer association in ABF expression per copy number between it and p*gpdA* from *A. vadensis* (FP-452) which are now both lying mid-table in terms of promoter strength.

In summary, we identified six novel constitutive promoters from *A. niger* and a further five from *A. vadensis* and tested them in *A. vadensis* using the *abf* gene from *Fusarium oxysporum* as a reporter for heterologous protein production. Of the promoters tested, 3 from *A. niger* (p*ef1α*, p*tal* and p*pgkA*) and 3 from *A. vadensis* (p*ef1α,* p*oliC and* p*rps31*) all resulted in higher ABF activity than for that of the commonly used *gpdA* promoter from *A. nidulans* signifying their potential for the industrial production of recombinant proteins.

Acknowledgements

The authors thank Chema Messlaki for technical assistance.

References

1. **Alberto, F., Navarro, D., de Vries, R.P., Asther, M. and Record, E.** 2009. Technical advance in fungal biotechnology: development of a miniaturized culture method and an automated high-throughput screening. Lett Appl Microbiol. **49**: 278-282.

2. **Bos, C.J., Debets, A.J., Swart, K., Huybers, A., Kobus, G. and Slakhorst, S.M.** 1988. Genetic analysis and the construction of master strains for assignment of genes to six linkage groups in *Aspergillus niger*. Curr Genet. **14**: 437-443.

3. **Culleton, H.M., McKie, V.A. and de Vries, R.P.** 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*? Biotechnology Journal. **8**: 884-894.

4. **de Bekker, C., Wiebenga, A., Aguilar, G. and Wosten, H.A.** 2009. An enzyme cocktail for efficient protoplast formation in *Aspergillus niger*. J Microbiol Methods. **76**: 305-306.

5. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.

6. **Ďuranová, M., Špániková, S., Wösten, H.B., Biely, P. and Vries, R.** 2009. Two glucuronoyl esterases of Phanerochaete chrysosporium. Arch Microbiol. **191**: 133-140.

7. **Finkelstein, D.** 1987. Improvement of enzyme production in *Aspergillus*. Antonie Van Leeuwenhoek. **53**: 349-352.

8. **Flaherty, J.E. and Payne, G.** 1997. Overexpression of aflR Leads to Upregulation of Pathway Gene Transcription and Increased Aflatoxin Production in *Aspergillus flavus*. Appl Environ Microbiol. **63**: 3995-4000.

9. **Fleißner, A. and Dersch, P.** 2010. Expression and export: recombinant protein production systems for *Aspergillus*. Appl Microbiol Biotechnol. **87**: 1255-1270.

10. **Flipphi, M., Mathieu, M., Cirpus, I., Panozzo, C. and Felenbok, B.** 2001. Regulation of the aldehyde dehydrogenase gene (*aldA*) and its role in the control of the coinducer level necessary for induction of the ethanol utilization pathway in *Aspergillus nidulans*. Journal of Biological Chemistry. **276**: 6950-6958.

11. **Fowler, T., Berka, R.M. and Ward, M.** 1990. Regulation of the *glaA* gene of *Aspergillus niger*. Curr Genet. **18**: 537-545.

12. **Graaff, L., Broek, H. and Visser, J.** 1988. Isolation and transformation of the pyruvate kinase gene of *Aspergillus nidulans*. Current genetics. **13**: 315-321.

13. **Kanemori, Y., Gomi, K., Kitamoto, K., Kumagai, C. and Tamura, G.** 1999. Insertion analysis of putative functional elements in the promoter region of the

Aspergillus oryzae Taka-amylase A gene (*amyB*) using a heterologous *Aspergillus nidulans amdS-lacZ* fusion gene system. Biosci Biotechnol Biochem. **63**: 180-183.
14. **Kelly, J.M. and Hynes, M.** 1987. Multiple copies of the *amdS* gene of *Aspergillus nidulans* cause titration of trans-acting regulatory proteins. Current genetics. **12**: 21-31.
15. **Kusters-van Someren, M.A., Harmsen, J.A., Kester, H.C. and Visser, J.** 1991. Structure of the *Aspergillus niger pelA* gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. Curr Genet. **20**: 293-299.
16. **Lagopodi, A.L., Ram, A.F.J., Lamers, G.E.M., Punt, P.J., Van den Hondel, C.A.M.J.J., Lugtenberg, B.J.J., et al.** 2002. Novel Aspects of Tomato Root Colonization and Infection by *Fusarium oxysporum* f. sp. *radicis-lycopersici* Revealed by Confocal Laser Scanning Microscopic Analysis Using the Green Fluorescent Protein as a Marker. Molecular Plant-Microbe Interactions. **15**: 172-179.
17. **Lombraña, M., Moralejo, F.J., Pinto, R. and Martín, J.F.** 2004. Modulation of *Aspergillus awamori* thaumatin secretion by modification of *bipA* gene expression. Appl Environ Microbiol. **70**: 5145-5152.
18. **Lubertozzi, D. and Keasling, J.D.** 2008. Expression of a synthetic Artemesia annua amorphadiene synthase in *Aspergillus nidulans* yields altered product distribution. J Ind Microbiol Biotechnol. **35**: 1191-1198.
19. **Moralejo, F.-J., Cardoza, R.-E., Gutierrez, S. and Martin, J.F.** 1999. Thaumatin production in *Aspergillus awamori* by use of expression cassettes with strong fungal promoters and high gene dosage. Appl Environ Microbiol. **65**: 1168-1174.
20. **Mullaney, E., Hamer, J., Roberti, K., Yelton, M.M. and Timberlake, W.** 1985. Primary structure of the *trpC* gene from *Aspergillus nidulans*. Molecular and General Genetics MGG. **199**: 37-45.
21. **Peraza, L., Ortiz, M.A., Peberdy, J.F. and Aguilar, G.** 2003. Growth and pectinase production by *Aspergillus* Mexican strain protoplast regenerated under acidic stress. Appl Biochem Biotechnol. **111**: 15-27.
22. **Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D.M., Pouwels, P.H. and van den Hondel, C.A.M.J.J.** 1990. Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. Gene. **93**: 101-109.
23. **Roth, A.F.J. and Dersch, P.** 2010. A novel expression system for intracellular production and purification of recombinant affinity-tagged proteins in *Aspergillus niger*. Appl Microbiol Biotechnol. **86**: 659-670.
24. **Sambrook, J. and Russell, D.W.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 2001.

25. **van den Hombergh, J.P.T.W., van de Vondervoort, P.J.I., Fraissinet-Tachet, L. and Visser, J.** 1997. *Aspergillus* as a host for heterologous protein production: the problem of proteases. Trends in Biotechnology. **15**: 256-263.
26. **Verdoes, J.C., Punt, P. and Van Den Hondel, C.** 1995. Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. Appl Microbiol Biotechnol. **43**: 195-205.
27. **Waring, R.B., May, G.S. and Morris, N.R.** 1989. Characterization of an inducible expression system in *Aspergillus nidulans* using *alcA* and tubulincoding genes. Gene. **79**: 119-130.

Supplemental Table 1: Primers used in this study.

Primers [*A. niger*]	**Sequence**	**Direction***	**Amplicon ****
An18g04840F An18g04840R	GCGGCCGCATGGCAAAAAAAGACG CCATGGGATGACGGTTGTGAATGAAC	Forward Reverse	1033
An08g06570F An08g06570R	GCGGCCGCCTGGCTTGAAAATTGG CCATGGGATTGCTGTGTCTAGAGAAGG	Forward Reverse	1138
An08g03490F An08g03490R	GCGGCCGCTTCCCAAGCCAAGTAAG CCATGGTGTGAATGTGTGTTTGGGGG	Forward Reverse	1017
An07g03850F An07g03850R	GCGGCCGCCACTGTCACAATGCTG CCATGGGGTGGAGGTAGAGAATGAG	Forward Reverse	1021
An16g03330F An16g03330R	GCGGCCGCGTAAAAGAAAGGAACG CCATGGTTTGATGAGATTTGTTTTAAGAGATTG	Forward Reverse	1043
An08g02260F An08g02260R	GCGGCCGCGCTAATATTATTACCAGG CCATGGGATGAACGACGGTTCTAC	Forward Reverse	1117
Primers [*A. vadensis*]	**Sequence**	**Direction***	**Amplicon ****
KJ420614F KJ420614R	GGGCGGCCGCAAAAAAAGACGTGG GGCCATGGCCGTTAGCCCATGTCC	Forward Reverse	1125
KJ420615F KJ420615R	GGGCGGCCGCGAAGCGTCAAGCTCC GGCCATGGTGACTGTCTGATGTAGC	Forward Reverse	1033
KJ420616F KJ420616R	GGGCGGCCGCAGAGGCCAGAATAATAAG GGCCATGGTTTAGATGTGTCTATG	Forward Reverse.	1052
KJ420617F KJ420617R	GGGCGGCCGCATGGAGCCCGGCGATTTG GGCCATGGTGATGAAGGTCTGG	Forward Reverse	1038
KJ420618F KJ420618R	GGGCGGCCGCAGGGACCCACCCGCACG GGCCATGGTGATGGTTTGTGG	Forward Reverse	1007

* Forward primers contain a restriction site of *Not*I and reverse primers contain a restriction site of *Nco*I

****** Amplification size upstream of gene

Chapter 5

Overexpression, purification and characterization of homologous α-L-arabinofuranosidase and *endo*-1,4-β-D-glucanase in *Aspergillus vadensis*

Culleton, H.M., McKie, V.A. and de Vries, R.P. 2014. Ind. Micro. & Biotech. **41**: 1697-1708

Abstract

In the recent past, much research has been applied to the development of *Aspergillus*, most notably *A. niger* and *A. oryzae*, as hosts for recombinant protein production. In this study, the potential of another species, *A. vadensis*, was examined. The full length gDNA encoding two plant biomass degrading enzymes i.e. *α*-L-arabinofuranosidase (*abfB*) (GH54) and *endo*-1,4-*β*-D-glucanase (*eglA*) (GH12) from *A. vadensis* were successfully expressed using the *gpdA* promoter from *A. vadensis*. Both enzymes were produced extracellularly in *A. vadensis* as soluble proteins and successfully purified by affinity chromatography. The effect of culture conditions on the expression of *abfB* in *A. vadensis* were examined and optimised to give a yield of 30 mg/L when grown on a complex carbon source such as wheat bran. Characterization of the purified *α*-L-arabinofuranosidase from *A. vadensis* showed an optimum pH and temperature of pH 3.5 and 60°C which concur with those previously reported for *A. niger* AbfB. Comparative analysis to *A. niger* AbfA demonstrated interesting differences in temperate optima, pH stability and substrate specificities. The *endo*-1,4-*β*-D-glucanase from *A. vadensis* exhibited a pH and temperature optimum of pH 4.5 and 50°C respectively. Comparative biochemical analysis to the orthologous EglA from *A. niger* presented similar pH and substrate specificity profiles. However, significant differences in temperature optima and stability were noted.

Introduction

Over the past two decades, Aspergilli have become the most widely studied group of filamentous fungi due, in part, to their potential in many industrial applications [3, 4, 7, 11, 30]. Aspergilli have adapted to their lifestyle as common soil fungi that are found in many different environments by producing an extensive set of enzyme mixes in order to degrade the broad range of plant polysaccharides which they encounter [8]. Plant cell walls consist mainly of the polysaccharides cellulose, hemicelluloses (xyloglucans, xylan and galacto(gluco)mannan) and pectin which interact with each other as well as the aromatic polymer lignin to form a network of polymers with linkages and hydrogen bonds that give the plant cell wall its rigidity [13]. Fungi such as Aspergilli degrade these polysaccharides extracellularly by secreting diverse enzymatic mixtures which release utilizable oligo- and monosaccharides from the polysaccharide that is present [20]. The complete degradation of cellulose for instance, requires the action of at least three enzymes: *β*-1,4-D-glucosidase, cellobiohydrolase and *β*-1,4-D-endoglucanase. In contrast, the hydrolysis of xylan requires the combined action of at least nine different enzymes: *α*-L-arabinofuranosidase, *α*-1,4-D-galactosidase, *α*-glucuronidase, acetylxylan esterase, arabinoxylan arabinofuranohydrolase, *β*-1,4-D-xylosidase, feruloyl esterase, *β*-1,4-D-galactosidase and *β*-1,4-D-endoxylanase [9]. The ability to produce such a broad range of enzymes combined with their good fermentation capabilities have resulted in many studies being dedicated to the development of *Aspergillus* as hosts for the industrial production of recombinant proteins [15]. To date, much of this research was performed with *A. niger* and *A. oryzae*, but recently *A. vadensis*, a close relative of *A. niger*, has been suggested as a possibly more favourable alternative due to the low levels of extracellular proteases which it produces and the fact that it does not acidify the culture medium [14]. In this study, the potential of *A. vadensis* as a host for recombinant protein production was examined by cloning and expressing two homologous genes encoding cell wall polysaccharide degrading enzymes i.e. *α*-L-arabinofuranosidase (*abfB*) (GH54) and *endo*-1,4-*β*-D-glucanase (*eglA*) (GH12) in this prospective industrial strain.

α-L-arabinofuranosidases (non-reducing end *α*-L-arabinofuranosidases; EC 3.2.1.55) act by hydrolysing the terminal non-reducing *α*-L-1,2-, *α*-L-1,3-, and *α*-L-1,5-arabinofuranosyl residues in *α*-L-arabinosides and can act synergistically with other hemicellulases and pectic enzymes for the complete hydrolysis of xylans and pectins [26]. In recent years *α*-L-arabinofuranosidases have attracted considerable interest due to their potential in industrial applications such as in oligosaccharide synthesis [34, 35], the pre-treatment of lignocelluloses for bioethanol production [17, 37] and in the chlorination of paper products [19, 25]. To date, *α*-L-arabinofuranosidases have been isolated from various bacterial and fungal organisms

such as *Streptomyces sp.*[39], *Thermotoga maritima* [27] and *A. niger* [36] and their genes have been cloned and expressed in developed expression systems. *A. niger* itself produces two main extracellular *α*-L-arabinofuranosidases; *α*-L-arabinofuranosidase A (AbfA) which is specifically active towards *p*-nitrophenyl-*α*-L-arabinofuranoside and 1,5-*α*-L-arabinofuranose oligosaccharides, and *α*-L-arabinofuranosidase B (AbfB) which has a broader activity range and is active on both these substrates but also has activity towards 1,5-*α*-L-arabinan, beet and apple arabinan and arabinoxylan [36]. Additional genes encoding putative *α*-L-arabinofuranosidases were detected in the genome of *A. niger*, but have not been biochemically characterised [8]. In this study, a homologous *α*-L-arabinofuranosidase B encoding gene (*abfB*) from glycosyl hydrolase family 54 (GH54) was cloned and expressed in *A. vadensis* to examine the potential of this expression system. This *α*-L-arabinofuranosidase B (AvAbfB) was then characterized and compared to the commercially available and biochemically different *α*-L-arabinofuranosidase A from *A. niger* (AnAbfA) (Megazyme; Cat. No. E-AFASE) from glycosyl hydrolase family 51 (GH51).

Endoglucanases (*endo*-1,4-*β*-D-glucanases, EC 3.2.1.4) are a group of enzymes which combined with cellobiohydrolases (EC 3.2.1.91) and *β*-glucosidases (EC 3.2.1.21) are responsible for the effective degradation of cellulose into glucose. Cellulose exists as highly ordered linear polymers of *β*-1,4-linked D-glucose residues which are bundled together in microfibrils via hydrogen bonds [18]. It is believed that endoglucanases act by initiating random attacks at multiple sites in the non-crystalline regions of the cellulose fibre [33]. This in turn, opens up sites for subsequent attack by cellobiohydrolases, which cleave cellulose chains at the ends and release cellobiose. These oligosaccharides are then further degraded into D-glucose molecules by the action of *β*-glucosidases and exoglucanases [9, 33]. Among the most efficient producers of cellulolytic enzymes, in particular *endo*-1,4-*β*-D-glucanase, is the filamentous fungus *A. niger* which has gained it significant interest especially in the food, textile and pharmaceutical industries [13]. Genes encoding endoglucanases from *A. niger*, such as *eglA*, *eglB* and *eglC*, have been cloned and characterized, with EglA demonstrating the highest activity towards *β*-glucan compared to EglB and EglC [21, 40]. Thus, for this study the *endo*-1,4-*β*-D-glucanase A encoding gene (*eglA*) from glycosyl hydrolase family 12 (GH12) was cloned and expressed in *A. vadensis* and the activity of the corresponding enzyme (AvEglA) was compared to the commercially available orthologous enzyme, AnEglA from *A. niger* (Megazyme; Cat. No. E-CELAN), which was also previously expressed through *Pichia* and described by Quay *et.al*. (2011) [33].

Materials and Methods

Strains, media and culture conditions

Escherchia coli XL1-Blue chemically competent strain (Fisher Scientific) was used as a host for recombinant DNA manipulation. Gene constructs were obtained by PCR from the gDNA of *Aspergillus vadensis* CBS 137441. *Aspergillus vadensis* CBS 113365 (*pyrG* -) was used as the parental strain for transformation.

Aspergillus minimal medium (MM) and complete medium (CM) were described previously [14]. Agar was added at 2% (w/v) for solid medium. Pre-cultures for protoplast formation were grown overnight at 30°C in 200 mL MM supplemented with 0.5% (w/v) yeast extract, 0.2% (w/v) casamino acids, 2% (w/v) glucose and 1.22 mg/mL (final) uridine after inoculating with 5 x 10^6 spores/mL. For the analysis and characterization of enzyme activities, 400 mL liquid cultures (2.5 L baffled flasks) of MM + 1% (w/v) wheat bran (for production of AvAbfB) or 2% (w/v) sucrose (for production of AvEglA) + 0.1% (w/v) TWEEN® 80 were inoculated with 1 x 10^6 spores/mL (final) and incubated at 35°C in an orbital shaker at 250 rpm.

Molecular biology methods

Standard methods were used for DNA manipulations, subcloning, DNA digestion reactions and DNA isolations [38].

PCR and expression vectors

Genes were amplified from the then partially sequenced *A. vadensis* genome (Culleton and de Vries, unpublished data) with primers including 5' *Nco*I and 3' *Afl*II restriction sites necessary for the cloning of the *α*-L-arabinofuranosidase encoding *abfB* gene (1500 bp) and 5' *Sph*I and 3' *Afl*II restriction sites necessary for the *endo*-1,4-*β*-D-glucanase encoding *eglA* gene (836 bp). The nucleotide sequences of the primers utilised (incorporating the restriction sites indicated) are as follows: *abfB* 5' end oligonucleotide 5'-GGCCATGGTCTCCCGCCGAAACC-3' and *abfB* 3' oligonucleotide 3'-GGCTTAAGCGAAGAAAACGCCGTCTC-5'. *eglA* 5' end oligonucleotide 5'-GGGCATGCAGCTCGCAGTGACAC-3' and *eglA* 3' oligonucleotide 3'-GGCTTAAGGTTGACACTAGCGGTCC-5'. PCR reactions were carried out using KOD DNA Polymerase (Merck Biosciences) and conditions supplied. PCR products were inserted into pCR2.1 TOPO vector (Invitrogen) following instructions provided and plasmids were verified by sequencing.

The pGPDGFP expression vector was used to make the new expression constructs in this study. As described by Lagopodi *et al.* (2002) [24], the pGPDGFP vector is composed of the *gfp* gene under the control of the *gpdA* promoter from *A. nidulans* [31] and terminated with the *trpC* terminator [28]. A new expression vector was built based on the pGPDGFP vector where p*gpdA* from *A. nidulans* was replaced with

p*gpdA* from *A. vadensis* [10] using restrictions enzymes *Not*I/*Nco*I. Due to the presence of an internal *Hin*dIII within p*gpdA* from *A. vadensis*, the *Hin*dIII site at the 3' end of the *gfp* gene was replaced with *Afl*II and a 6 X Histidine tag and stop codon were added to assist in the downstream purification of the enzymes once expressed. Due to the presence of an *Nco*I site within the gene sequence of Av*eglA*, an additional *Sph*I site was added alongside the *Nco*I site at 3' end of promoter to facilitate the cloning on of this gene. The *gfp* gene was then replaced with the new gene candidate from *A. vadensis* using restriction enzymes *Nco*I/*Afl*II for Av*abfB* and *Sph*I/*Afl*II for Av*eglA* (Fig. 1).

Transformation of A. vadensis

The formation of protoplasts by *Aspergillus* strains was based on the protocols by Peraza *et al.* (2003) [29] and de Bekker *et al.* (2009) [12]. Strains were grown for 16 hrs, after which time the mycelia were gently harvested by filtration over a Büchner funnel with nylon gauze. After washing with 0.9% NaCl (w/v), 2.5 g (wet weight) mycelium was resuspended in 20 mL stabilization buffer (0.2 M phosphate buffer (pH 6.0), 0.8 M sorbitol). Lytic enzymes were added to the following final concentrations; 5 mg/mL lysing enzymes from *Trichoderma harzianum* (Sigma; Cat. No. L1412), 460 units/mL β-glucuronidase from *Helix pomatia* (Sigma; Cat. No. G0751) and 0.15 units/mL chitinase from *Streptomyces griseus* (Sigma; Cat. No. C6137). The mixture was incubated in an orbital shaker at 37°C for 1 to 2 hrs with gentle shaking (120 rpm). Protoplasts were separated from the mycelium by filtering over glass wool. The protoplasts were recovered by centrifugation in a swing-out rotor (10 min; ~ 800 rcf) and were washed twice with STC (1.33 M sorbitol, 50 mM $CaCl_2$ and 10 mM Tris/HCl, pH 7.5). Transformation was performed as described by Kusters-van Someren *et al.* (1991) [22], with 2 x 10^6 protoplasts, 0.5 µg of pGW635 (carrying the *A. niger pyrG* gene for selection) and 20 µg of the different expression vectors (carrying the studied genes under the control of the p*gpdA* from *A. vadensis*).

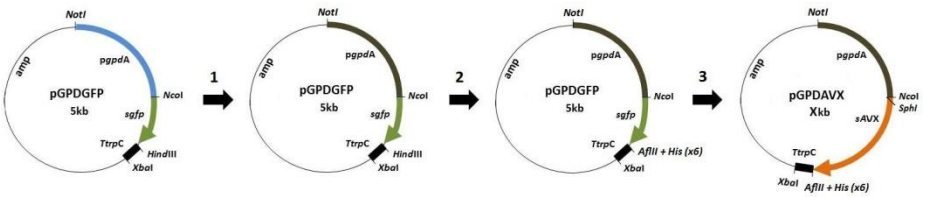


Figure 1: **Construction of the new expression plasmid.** (1) Replacement of glyceraldehyde-3-phosphate dehydrogenase promoter (p*gpdA*) from *A. nidulans* with p*gpdA* from *A. vadensis*. (2) Replacement of *Hin*dIII site at 3' end of gene construct with *Afl*II and addition of 6 x His-tag. (3) Insertion of *Sph*I site alongside *Nco*I site and *A. vadensis* genes (*sAVX*) in place of green fluorescence protein gene (*sgfp*).

Screening of transformants and selection of expression strains

To test the resulting transformants for *abfB* expression, a fluorimetric ABF screen based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-α-L-arabinopyranoside (0.02 mg/mL (final) in MM + 2% (w/v) sucrose + 2% (w/v) agar) when 1000 spores (2µL) from each transformant were grown at 30°C for 2 days and viewed under UV light. To test for *eglA* expression, resulting transformants were grown on agar plates containing high purity dyed and crosslinked insoluble AZCL-Barley beta-Glucan (Megazyme; Cat. No. I-AZBGL) (0.1% (w/v) in MM + 2% (w/v) sucrose + 2% (w/v) agar) when 1000 spores (2µL) from each transformant were grown at 30°C for 2 days. In both screens, CBS 137441 was grown as a negative control.

Production and purification of recombinant enzymes

400 ml liquid cultures were inoculated with 1 x 10^6 spores/mL (final) and grown in 2.5L baffled flasks at 35°C, 250 rpm with production levels being monitored daily by SDS polyacrylamide gel electrophoresis (SDS-PAGE). SDS-PAGE was done using a 12% polyacrylamide gel containing 0.1% (w/v) SDS and protein bands were detected by Coomassie Blue staining [23]. Cultures were grown for the time indicated in Results, after which time the mycelium was removed by filtration over a Büchner funnel with nylon gauze. Culture filtrates were purified by His-tag affinity chromatography, equilibrated with buffer (10mM imidazole, 10 mM HEPES and 500 mM NaCl, pH 7.5) and eluted with a stepwise gradient in same buffer containing 10 mM to 500mM imidazole. Eluent from the column was monitored for levels of protein concentration by testing 10 µL of sample in 150 µL of Bio-Rad protein assay dye reagent (Bio-Rad; Cat. No. 500-0006). 15 µL samples of culture filtrates (C.F.) and column fractions were viewed by SDS-PAGE (as above) and fractions containing purified protein were pooled and precipitated with 50% ammonium sulphate. Enzyme fractions were then concentrated by centrifugation and the enzyme pellet was re-suspended in 3.2 M ammonium sulphate solution for characterization.

Enzyme activity measurement

For the initial measurement of the *A. vadensis* α-arabinofuranosidase B (AvAbfB) and *A. niger* α-arabinofuranosidase A (AnAbfA) activities, serial dilutions of the purified enzymes were carried out in 100 mM sodium acetate buffer, pH 4.0 including BSA (1 mg/mL). 0.2 mL of diluted enzyme were added to 0.2 mL of 10 mM *p*-nitrophenyl α-L-arabinofuranoside (pH 4.0) and incubated at 40 °C for 10 min. The reaction was stopped with the addition of 3 mL of 2 % (w/v) tri-sodium orthophosphate, pH 12.0 and the absorbance was measured at 400 nm. Activities were expressed as Units/mg where one unit is defined as 1 micromole of *p*-nitrophenol liberated per minute per milligram of enzyme.

For the initial measurement of the *A. vadensis endo*-1,4-*β*-D-glucanase A (AvEglA) and the orthologous *A. niger endo*-1,4-*β*-D-glucanase A (AnEglA) activities, serial dilutions of the purified enzymes were performed in 100 mM sodium acetate buffer, pH 4.5 including BSA (1 mg/mL). Nelson-Somogyi reducing sugar assays were performed by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 10 mg/mL Barley *β*-glucan (Megazyme; Cat. No. P-BGBM) (pH 4.5) and incubated at 40°C for 10 min. The reaction was stopped by the addition of 0.5 mL of Stopping Solution (25 mL of Solution A (2.5% (w/v) sodium carbonate anhydrous, 2.5% (w/v) potassium sodium tartrate and 20% (w/v) sodium sulphate), to which 1 mL of Solution B (3% (w/v) copper sulphate pentahydrate) was added). Enzyme reactions were then boiled for 20 min and allowed to cool to room temperature for 5 min before the addition of 3.0 mL of a 1:5 dilution of Solution C (5% (w/v) ammonium molybdate, 4.2% (v/v) concentrated sulphuric acid and 0.6% (w/v) sodium arsenate heptahydrate).

pH optima assays were conducted in duplicate measurements using the optimum enzyme dilution as determined from the initial activity assays and using the following pH buffers covering from pH 1.0 to pH 9.0 including BSA (1 mg/mL); potassium chloride (pH 1.0 and pH 2.0), glycine (pH 2.0, pH 2.5, pH 3.0), citrate phosphate (pH 3.0, pH 3.5, pH 4.0, pH 4.5, pH 5.0, pH 5.5, pH 6.0, pH 6.5 and pH 7.0) and sodium phosphate (pH 6.0, pH 7.0, pH 8.0 and pH 9.0). The remaining assay conditions were consistent with the assay conditions used for the measurement of initial activity. For pH stability assays an initial 10-fold dilution of the enzyme was carried out in the buffers outlined above (pH 1.0 – pH 9.0) including BSA (1 mg/mL) and incubated at 4°C for 48 hrs. Activity assays were then conducted in duplicate measurements using the optimum enzyme concentration and pH as determined by previous assays with the enzyme activity at the optimum pH set at 100%.

For temperature stability assays an initial 10-fold dilution of the enzyme in optimum pH buffer was performed and aliquots incubated at the following temperatures for 15 min; 25°C, 30°C, 40°C, 50°C, 60°C, 70°C, 80°C, 90°C and 100°C. Following incubation, the enzyme was further diluted in optimum pH buffer to an optimum concentration and assayed in duplicate as described above. Temperature optima assays were performed in duplicate measurements using optimum enzyme concentration and pH as determined in previous assays and at temperatures that the enzyme proved stable.

A. vadensis α-arabinofuranosidase B (AvAbfB) and *A. niger α*-arabinofuranosidase A (AnAbfA) substrate specificity assays were performed using the following polysaccharide substrates at a final concentration of 5 mg/mL; wheat flour arabinoxylan (Megazyme; Cat. No. P-WAXYL), sugar beet arabinan (Megazyme; Cat. No. P-ARAB) and debranched sugar beet arabinan (Megazyme; Cat. No. P-DBAR). Activities were determined under optimum conditions using Nelson-Somogyi reducing

sugar assays as described above. To test *A. vadensis* (AvEglA) and *A. niger* (AnEglA) endoglucanase A for substrate specificity, the following polysaccharide substrates and *p*-nitrophenol substrates were used at final concentrations of 5 mg/mL and 5 mM, respectively; carboxymethyl cellulose 4M (Megazyme; Cat. No. P-CMC4M), galactomannan (Megazyme; Cat. No. P-GALML), glucomannan (Megazyme; Cat. No. P-GLCML), pachyman (Megazyme; Cat. No. P-CMPAC), soluble starch (Sigma; Cat. No. S9765), xyloglucan (Megazyme; Cat. No. P-XYGLN), *p*-nitrophenol *α*-D-glucopyranoside, *p*-nitrophenol *β*-D-glucopyranoside and *p*-nitrophenol *β*-D-xylopyranoside. Assays were performed under optimum pH and temperature conditions as determined in previous assays and using methods as described above.

Results

Development of expression strains

The selected genes were PCR amplified from *A. vadensis* gDNA, inserted into pCR2.1 TOPO vector (Invitrogen) and sequenced using the Sanger method at LGC Genomics. Sequencing results confirmed that the correct gene regions were amplified. Identification and comparison of the corresponding *α*-L-arabinofuranosidase B protein (499 amino acids) and *endo*-1,4-*β*-D-glucanase A protein (239 amino acids) from *A. vadensis* with other characterized protein sequences was performed using the BLAST program at NCBI. The AvAbfB protein sequence has 99% identity to *Aspergillus kawachii* IFO 4308 (Q8NK89.1), *Aspergillus awamori* (Q9C4B1.1) and *Aspergillus niger* AbfB (XP_001396769.1). On the basis of the similarities to these other *α*-L-arabinofuranosidases, *A. vadensis abfB* was assigned to GH54 (CAZy-http://www.cazy.org) [5]. The AvEglA protein sequence showed high identity (96%) to *A. niger* CBS 513.88 (XP_001400902.1) and 95% to *Aspergillus usamii* (AEL12376.1), thus providing basis for its inclusion into the GH12 family.

For expression and characterization of the selected genes, the corresponding fragments were cloned behind the *gpdA* promoter from *A. vadensis* [10]. The resulting expression vectors containing the gene sequence were then transformed into *A. vadensis* CBS113365. To test the resulting transformants for *abfB* expression, a fluorimetric screen based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-*α*-L-arabinopyranoside was applied to select the strongest expressing transformants (Fig. 2A) i.e. transformant no. 7 (*abfB*_t7). To test for *eglA* expression, resulting transformants were grown on agar plates containing high purity dyed and crosslinked insoluble AZCL-Barley beta-Glucan (Megazyme; Cat. No. I-AZBGL), from which the strongest expressing clones were selected based on ability to hydrolyse the AZCL-Barley beta-Glucan substrate (Fig. 2B) i.e. transformant no. 9 (*eglA*_t9).

Effect of culture conditions on the production of α-arabinofuranosidase (AbfB)

*A. vadensis abfB*_t7 was cultivated under several different conditions for investigation of *α*-arabinofuranosidase production. Glucose, mannose and sucrose were used as simple carbon sources for comparison purposes, with sucrose giving optimum *abfB* expression with yields of ~7 mg/L. The addition of 0.1% (w/v) TWEEN® 80 to media doubled these yields to ~15 mg/L. Growth temperatures from 25°C - 37°C were examined with *A. vadensis abfB*_t7 favouring higher incubation temperatures i.e. 35°C - 37°C, for the expression of recombinant *α*-arabinofuranosidase. The use of a complex carbon source such as wheat bran (Odlums) under optimum conditions increased native and recombinant protein production alike, giving a purified AvAbfB yield of 30 mg/L.

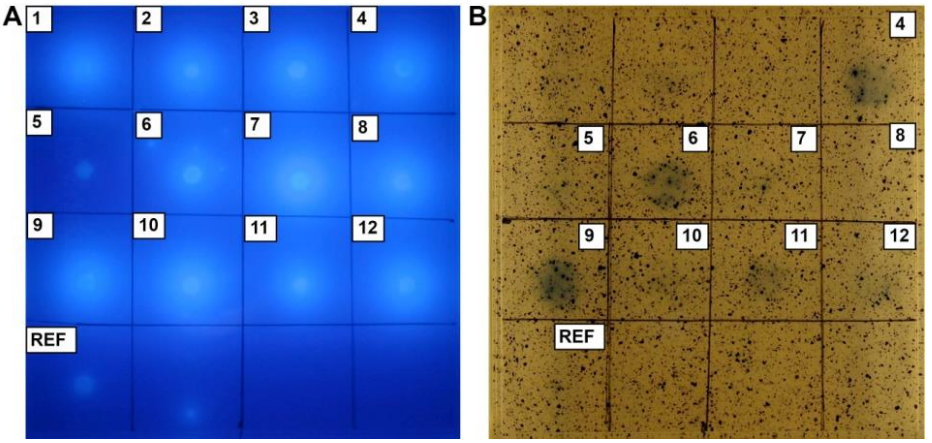


Figure 2: **Substrate plate screens to select the strongest AbfB producers (A) and EglA producers (B).** Fluorimetric ABF screen based on the conversion of the fluorogenic substrate 4-methylumbelliferyl-α-L-arabinopyranoside (0.02 mg/mL (final)) when 1000 spores (2µL) from each transformant were grown at 30°C for 2 days and viewed under UV light. EGL screen based on the ability to hydrolyse the dye particles of the crosslinked and insoluble AZCL-Barley beta-Glucan (I-AZBGL, Megazyme) (0.1% (w/v)) when 1000 spores (2µL) from each transformant were grown at 30°C for 2 days. CBS 137441 was grown as a negative control. In this study colony no. 7 (*abfB*_t7) was shown to give the strongest fluorescence and colony no. 9 (*eglA*_t9) was shown to give the greatest amount of dye release and hence both were chosen for liquid expression studies.

α-arabinofuranosidase (AbfB) production and characterization

Liquid cultures were inoculated with 1 x 10^6 spores/mL (final) from chosen transformants, grown at 35°C and 250 rpm and expression levels were monitored on a daily basis by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Optimum AvAbfB production was observed after 3 days incubation and the enzyme was purified from these culture filtrates using His-tag affinity chromatography. 15 µL samples of culture filtrates (C.F.) and column fractions were viewed by SDS-PAGE (Figure 3A) with purified protein visible in the eluent fractions containing 50 mM, 100 mM and 500 mM imidazole. These were pooled and precipitated with 50% ammonium sulphate. Enzyme fractions were then concentrated and quantified giving a total yield of 30 mg/L purified protein (Table 1A).

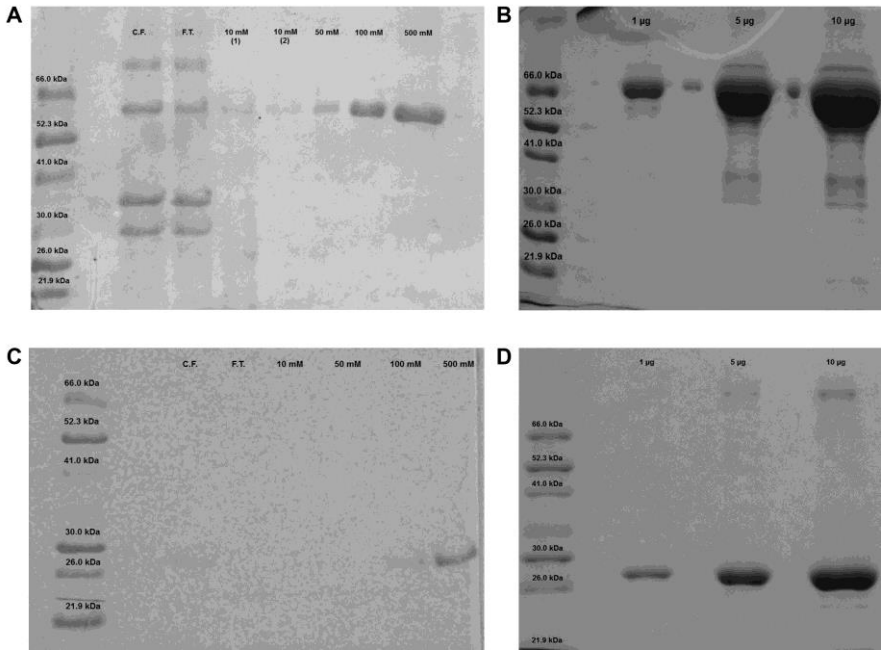


Figure 3: SDS-PAGE analysis and purification of *A. vadensis* α-L-arabinofuranosidase B (AvAbfB) (A) compared to increasing amounts i.e. 1 µg, 5µg and 10µg, of industrially purified α-L-arabinofuranosidase A from *A. niger* (AnAbfA) (B) and *endo*-1,4-β-D-glucanase (AvEglA) (C) compared to increasing amounts i.e. 1 µg, 5µg and 10µg, of industrially purified and orthologous *A. niger* EglA (D). Low molecular weight (LMW) ladder produced in house. Culture filtrates (C.F.) were purified by affinity chromatography, equilibrated with buffer (10mM imidazole, 10 mM HEPES and 500 mM NaCl, pH 7.5) and eluted with a stepwise gradient in same buffer containing 10 mM to 500mM imidazole. Flow through (F.T.) contained proteins which had no specific binding capacity to the resin.

The enzymatic properties of the recombinant *α*-arabinofuranosidase B from *A. vadensis* (AvAbfB) compared to that of the commercially available and industrially purified *α*-arabinofuranosidase A from *A. niger* (AnAbfA) were examined. The optimal pH for both enzymes was pH 3.5 (activity measured at pH 3.5 equalled that at pH 4.0 for AnAbfA) with both retaining ≥ 90% activity at pH 3.0 and pH 4.0 (Figure 4A). AvAbfB demonstrated greater pH stability than AnAbfA, with ≥ 92% of AvAbfB activity remaining at all pH's tested i.e. pH 1.0 – pH 9.0, compared to AnAbfA which resulted in 0% activity at pH 1.0, even with immediate testing after 1 hr incubation at this pH (Figure 4B).

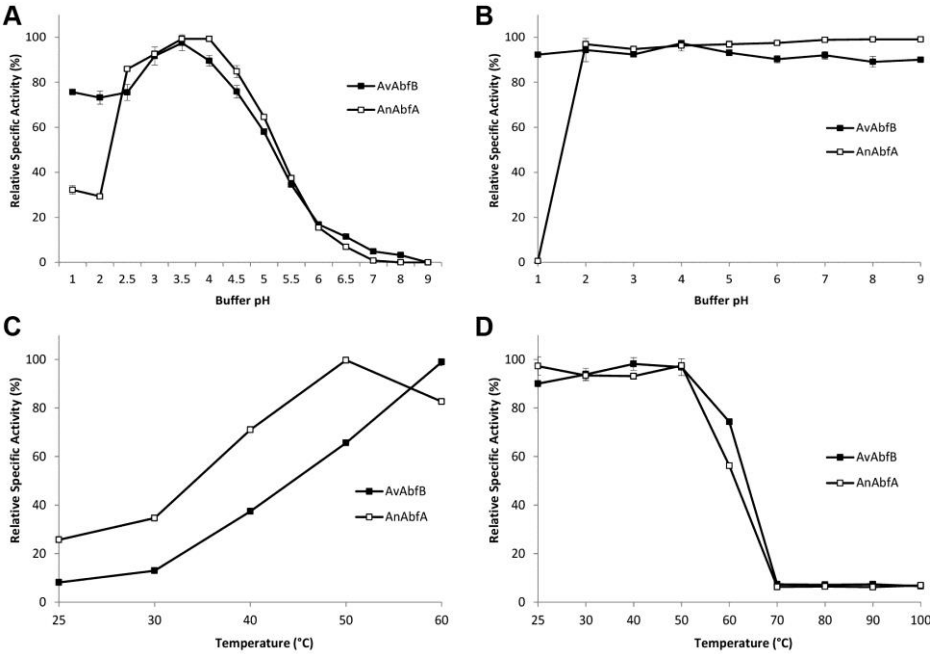


Figure 4: Biochemical properties of purified recombinant *α*-arabinofuranosidase from *A. vadensis* (AvAbfB) compared to those of commercially available *α*-arabinofuranosidase A from *A. niger* (AnAbfA). Values expressed as relative specific activity (%) of maximum activity obtained for that assay. In the case of pH optima and stability, average values were taken where there was an overlap in pH with the use of different buffers. Error bars calculated on standard deviations between technical duplicates. (A) pH optima of AvAbfB and AnAbfA; (B) pH stability of AvAbfB and AnAbfA; (C) Temperature optima of AvAbfB and AnAbfA; (D) Temperature stability of AvAbfB and AnAbfA.

AvAbfB was also more stable at higher temperatures than AnAbfA, with 100% of activity being preserved for both enzymes up to 50°C but with 77% of AvAbfB activity being maintained after a 15 min incubation at 60°C compared to just 58% in the case of AnAbfA (Figure 4D). The temperature optima for these enzymes therefore differed with AvAbfB giving maximum activity at 60°C and AnAbfA having an optimum of 50°C (Figure 4C). Specific activities for both enzymes in 100mM citrate phosphate buffer pH 3.5 at 40°C were 44 U/mg for AvAbfB and 127 U/mg for AnAbfA. These values increased to 100 U/mg for AvAbfB and 221 U/mg for AnAbfA when assayed under optimum pH and at 50°C.

Additional enzyme activities were also measured for both enzymes on wheat flour arabinoxylan, sugar beet arabinan and debranched sugar beet arabinan relative to the specific activities obtained on *p*-nitrophenyl *α*-L-arabinofuranoside at 40°C (Table 2). In these experiments AvAbfB demonstrated greater activity on all three tested polysaccharides than AnAbfA.

endo-1,4-β-D-glucanase (EglA) production and characterization

Liquid cultures were inoculated with 1 x 10^6 spores/mL from chosen transformants, grown at 35°C and 250 rpm and expression levels were monitored on a daily basis by SDS polyacrylamide gel electrophoresis (SDS-PAGE). Optimum AvEglA production was observed after 4 days incubation and resulting culture filtrates were purified by

Table 1: Purification of *α*-L-arabinofuranosidase (AbfB) and *endo*-1,4-*β*-D-glucanase (EglA) from *Aspergillus vadensis.* Specific activities calculated at 40°C / 50°C under optimum conditions for each enzyme as outlined in Material and Methods.

A) Purification of *α*-L-arabinofuranosidase (AbfB)				
	Purification-Fold	Concentration (mg/mL)	Specific Yield (mg/L)	Specific Activity (U/mg) 40°C / 50°C
Crude (1500 mL)	-	-	-	-
Purified (20 mL)	75	2.27	30	44 / 100
B) Purification of *endo*-1,4-*β*-D-glucanase (EglA)				
	Purification-Fold	Concentration (mg/mL)	Specific Yield (mg/L)	Specific Activity (U/mg) 40°C / 50°C
Crude (370 mL)	-	-	-	-
Purified (10 mL)	37	0.92	25	204 / 280

His-tag affinity chromatography. 15µL samples of culture filtrates (C.F.) and column fractions were viewed by SDS-PAGE (Figure 3B) with the 100 mM and 500 mM fractions containing purified protein. These were pooled and precipitated with 50% ammonium sulphate. Enzyme fractions were then concentrated and quantified giving a total yield of 25 mg/L purified protein (Table 1B).

The enzymatic properties of the recombinant *endo*-1,4-*β*-D-glucanase from *A. vadensis* (AvEglA) were compared to the commercially available and industrially purified orthologous *endo*-1,4-*β*-D-glucanase from *A. niger* (AnEglA). The optimal pH for both enzymes peaked at pH 4.5 with AvEglA demonstrating a narrower optima curve than AnEglA, losing 55% / 40% relative activity at pH 4.0 / pH 5.0 respectively compared to AnEglA which lost 25% / 6% activity at same (Fig. 5A). Both enzymes showed comparable stability at all pH's tested i.e. stable at pH 1.0 – pH 9.0 after 48 hrs incubation at this pH (Fig. 5B).

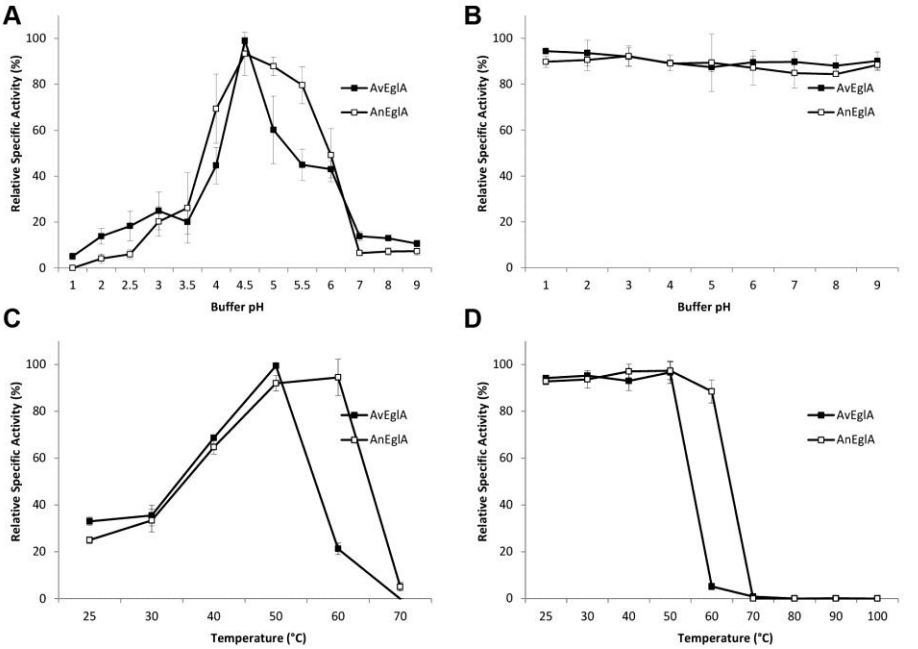


Figure 5: Biochemical properties of purified recombinant endoglucanase from *A. vadensis* (AvEglA) compared to those of commercially available *endo*-1,4-*β*-D-glucanase from *A. niger* (AnEglA). Values expressed as relative specific activity (%) of maximum activity obtained for that assay. In the case of pH optima and stability, average values were taken where there was an overlap in pH with the use of different buffers. Error bars calculated on standard deviations between technical duplicates. (A) pH optima of AvEglA and AnEglA; (B) pH stability of AvEglA and AnEglA; (C) Temperature optima of AvEglA and AnEglA; (D) Temperature stability of AvEglA and AnEglA activity.

The temperature optimum of AvEglA was 50°C, while AnEglA gave a broader optimum curve i.e. between 50°C and 60°C but which favoured the higher temperature of 60°C (50°C giving ~2.5% less relative specific activity compared to 60°C). AnEglA also demonstrating a greater stability profile in maintaining 91% relative activity at 60°C compared to AvEglA where activity at this temperature measured only 5% of that obtained at 50°C (Fig. 5D and 5C).

Specific activities for both enzymes in 100 mM sodium acetate buffer pH 4.5 at 40°C were 204 U/mg for AvEglA and 179 U/mg for AnEglA. These values increased to 280 U/mg for AvEglA and 230 U/mg for AnEglA when assayed under optimum pH at 50°C. No activity was detected for either enzyme on galactomannan, glucomannan, pachyman, soluble starch, xyloglucan, *p*-nitrophenol *α*- D-glucopyranoside, p-nitrophenol β-D-glucopyranoside and *p*-nitrophenol *β*-D-xylopyranoside. A lower activity on cellulose was observed for AvEglA than for AnEglA relative to the specific activities obtained on Barley *β*-glucan at 40°C (Table 2).

Table 2: Relative activities of the purified *α*-arabinofuranosidase and *endo*-glucanase against different substrates. All polysaccharide substrates and *p*-nitrophenol substrates were tested at 40°C and at concentrations of 5 mg/mL (w/v final) and 5 mM (final) respectively, under conditions as described in Materials and Methods.

Substrate	**Relative activity (%)**	
	***A. vadensis* AbfB**	***A. niger* AbfA**
p-nitrophenol *α*-L-arabinofuranoside	100.0	100.0
Wheat flour arabinoxylan	1.5	1.4
Sugar beet arabinan	19.2	5.9
Debranched sugar beet arabinan	3.2	1.0
Substrate	***A. vadensis* EglA**	***A. niger* EglA**
Barley *β*-Glucan	100.0	100.0
Cellulose	24.0	65.0
Galactomannan	0.0	0.0
Glucomannan	0.2	0.2
Pachyman	0.0	0.0
Soluble Starch	0.0	0.0
Xyloglucan	0.0	0.0
p-nitrophenol *α*-D-glucopyranoside	0.0	0.0
p-nitrophenol *β*-D-glucopyranoside	0.0	0.0
p-nitrophenol *β*-D-xylopyranoside	0.0	0.0

Discussion

Over the past 25 years, much research has been devoted to the development of *Aspergillus* as a host for homologous and heterologous protein production. Not only does *Aspergillus* growth excel under fermentation conditions with an exceptional capacity of secreting high levels of homologous product but their potential in plant polysaccharide degradation and the extensive set of enzyme mixes which they secrete for this purpose is also well recognized [9]. To date, much of this research has been focused on *A. niger* and *A. oryzae*, while the capacity of *A. vadensis*, a close relative of *A. niger*, has remained largely unexplored. In this study, the potential of *A. vadensis* as a host for recombinant protein production was examined by cloning and over-expressing two homologous genes encoding plant biomass degrading enzymes i.e. *α*-L-arabinofuranosidase (*abfB*) and *endo*-1,4-*β*-D-glucanase (*eglA*).

As shown in Figure 3A the *A. vadensis α*-L-arabinofuranosidase B (AvAbfB) protein was purified at a molecular weight of ~ 53 kDa, which corresponds to the calculated predicted mass of 52.5 kDa, based on the obtained amino acid sequence information and to that of the previously recorded 51.0 kDa for the orthologous AbfB enzyme in *A. niger* [16]. Earlier studies reported that the *α*-L-arabinofuranosidases from various *Aspergillus* species had a molecular weight of about 30 to 118 kDa, depending on the level of *N*- and *O*-glycosylation sites [2]. This appears to be the case for *α*-L-arabinofuranosidase A (AnAbfA) which has a molecular weight of ~ 62 kDa and is visualised as a single major band on SDS-electrophoresis (Figure 3B) as is described on the product's data sheet (http://secure.megazyme.com/Alpha-L-Arabinofuranosidase_A._niger). The biochemical properties of *A. vadensis α*-L-arabinofuranosidase B (AvAbfB) were examined and found to match those of the previously identified orthologous AbfB from *A. niger* with a pH and temperature optimum of pH 3.5 and 60°C but interestingly, is stable at pH's as low as pH 1.0, which other *α*-L-arabinofuranosidases from *A. niger* were not [2]. Comparative biochemical analysis with *A. niger α*-L-arabinofuranosidase A (AnAbfA) demonstrated that AnAbfA had more specific activity towards *p*-nitrophenyl *α*-L-arabinofuranoside but ~ 3 times less than AvAbfB on both the branched and debranched sugar beet arabinan. Sugar beet arabinan consists of a 1,5-*α*-linked backbone to which 1,3-*α*-linked (and possibly some 1,2-*α*-linked) L-arabinofuranosyl residues are attached. In previous studies, AbfA from *A. niger* was reported to be incapable of splitting 1,3-*α*-L - or 1,2-*α*-L linked arabinose substituents or arabinose from longer 1,5-*α*-L-linked arabinose residues whereas AbfB can [36].

Figure 3C shows the purification of *A. vadensis endo*-1,4-*β*-D-glucanase (AvEglA) at a molecular weight of 26 kDa which too, is in keeping with the calculated predicted mass of 27.72 kDa, based on the obtained amino acid sequence information and with that previously obtained for the purification of the orthologous protein from *A. niger*

[33]. This is confirmed with *A. niger endo*-1,4-*β*-D-glucanase (AnEglA) being visualised at a molecular weight of ~27 kDa and as a single band on SDS-electrophoresis (Figure 3D), as is described on the products data sheet (http://secure.megazyme.com/Cellulase_endo-1-4-Beta-D-glucanase_A._niger). The biochemical characteristics of *A. vadensis endo*-1,4-*β*-D-glucanase (AvEglA) are similar to those previously identified in the orthologous enzyme from *A. niger* [33]. Comparative biochemical analysis of AvEglA to the orthologous *endo*-1,4-*β*-D-glucanase from *A. niger* (AnEglA) showed similar characteristics, both giving equivalent pH optima (pH 4.5) and pH stability (pH 1 – pH 9) profiles. Interestingly, a notable difference was displayed for temperature optima and stability however, with AvEglA having a temperature optimum of 50°C, compared to AnEglA which gave a broader optimum curve between 50°C and 60°C but which favoured the higher temperature of 60°C. AnEglA also demonstrating a greater stability profile in maintaining 91% relative activity at 60°C compared to AvEglA where activity at this temperature measured only 5% of that obtained at 50°C. Both AvEglA and AnEglA demonstrated little to no activity on the additional substrates tested, with the exception of cellulose, where AnEglA demonstrated ~ 3 times greater activity against this substrate than AvEglA. Cellulose is a polymer of 1,4-*β*-linked D-glucosyl residues which both *endo*-1,4-*β*-D-glucanases should be able to degrade with similar efficiency. Sequence alignments for both EglA orthologs studied show a 5% difference at the amino acid level (Supplemental Figure 1B) which may lead to slight variations in the kinetic properties of the corresponding enzymes. Although both enzymes are visually pure by SDS comparison (Figures 3C and 3D), differences in production hosts, glycosylation patterns and purification methods may contribute to the variations recorded.

This study has demonstrated the potential of *A. vadensis* as a host for recombinant protein expression. Not only does *A. vadensis* not acidify the culture medium but it also produces very low levels of extracellular proteases and so facilitates many downstream processes. To date, the use of His-tag affinity chromatography in fungal systems has not been either efficient or practical in fungal systems due to the difficulties experienced with degradation of the Histidine residues by extracellular proteases. Purification of recombinant enzymes from *A. vadensis* through His-tag affinity chromatography have worked well thus far, suggesting its potential as a versatile host in the fundamental research of proteins and for industrial enzyme production. With an initial expression yield of ~ 30 mg/L, *A. vadensis* shows positive signs of becoming an industrially significant contender but considerable improvements will be required to rival the secretion abilities of commercial production strains such as *A. niger* and *T. reesei*. Strategies for improving recombinant protein production in fungi, including the use of strong promoters and effective secretion signals or gene

fusions to genes associated with well-expressed and secreted proteins have been implemented successfully to other systems in the past [32, 41]. Successes have also been obtained using a bioprocessing approach of optimizing fungal morphology, mycelia immobilization and culture conditions [1, 6]. It should be noted however, that the *A. vadensis* strain used in these experiments is near wild type and has not undergone the extensive strain improvement strategies that have resulted in the commercial *A. niger* and *T. reesei* strains that secrete much higher levels of enzyme, suggesting room for improvement.

The results of this study demonstrate significant variations in biochemical characteristics in the different *α*-arabinofuranosidases, AnAbfA and AvAbfB but also in the orthologous *endo*-1,4-*β*-D-glucanases (EglA) when produced from *A. niger* and *A. vadensis*, despite these organisms being taxonomically very close. Despite AnAbfA (GH51) and AvAbfB (GH54) having just 22% identity at an amino acid level (Supplemental Figure 1A) and being assigned to different CAZy families, the biochemical differences observed between these two enzymes are not significantly greater than that observed in the orthologous EglA comparison which have 95% identity. In particular, the difference in temperature optima and stability for these orthologous enzymes is noteworthy and would directly impact the applications of these enzymes. Screening of orthologs from related fungi may therefore be worthwhile for many industrial applications and suggests that it is not always necessary to go to distantly related species or enzyme classifications to get significant changes in biochemical properties. Detailed structural comparison of these orthologous enzymes may reveal insight into the molecular basis of this difference in stability and is worth pursuing in follow-up studies.

References

1. **Agger, T., Spohr, A.B., Carlsen, M. and Nielsen, J.** 1998. Growth and product formation of *Aspergillus oryzae* during submerged cultivations: verification of a morphologically structured model using fluorescent probes. Biotechnol Bioeng. **57**: 321-329.

2. **Alias, N.I., Mahadi, N.M., Murad, A.M.A., Bakar, F.D.A., Rabu, A. and Illias, R.M.** 2011. Expression optimisation of recombinant *α*-L-arabinofuranosidase from *Aspergillus niger* ATCC 120120 in *Pichia pastoris* and its biochemical characterisation. African Journal of Biotechnology. **10**: 6700-6710.

3. **Battaglia, E., Benoit, I., Gruben, B.S. and de Vries, R.P.** 2010. Plant cell wall derived sugars as substrates for fungi and industry. In: Jenkins PT, (ed) The Sugar Industry and Cotton Crops, Nova Science Publishers, New York.

4. **Bedford, M.R. and Classen, H.L.** 1991. The influence of dietary xylanase on intestinal viscosity and molecular weight distribution of carbohydrates in rye-fed broiler chicks. Progress in biotechnology. **7**: 361-370.

5. **Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B.** 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. **37**: 5.

6. **Carlsen, M., Spohr, A.B., Nielsen, J. and Villadsen, J.** 1996. Morphology and physiology of an alpha-amylase producing strain of *Aspergillus oryzae* during batch cultivations. Biotechnol Bioeng. **49**: 266-276.

7. **Christov, L.P., Szakacs, G. and Balakrishnan, H.** 1999. Production, partial characterization and use of fungal cellulase-free xylanases in pulp bleaching. Process Biochemistry. **34**: 511-517.

8. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.

9. **Culleton, H.M., McKie, V.A. and de Vries, R.P.** 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*? Biotechnology Journal. **8**: 884-894.

10. **Culleton, H.M., Bouzoid O, McKie V.A., de Vries R.P.** 2014. New promoters to improve heterologous protein production in *Aspergillus vadensis*. Curr Biotechnol. **3**: 1-8.

11. **Davies, R.W.** 1994. Heterologous gene expression and protein secretion in *Aspergillus*. Prog Ind Microbiol. **29**: 527-560.

12. **de Bekker, C., Wiebenga, A., Aguilar, G. and Wosten, H.A.** 2009. An enzyme cocktail for efficient protoplast formation in *Aspergillus niger*. J Microbiol Methods. **76**: 305-306.
13. **de Vries, R.P. and Visser, J.** 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev. **65**: 497-522.
14. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
15. **Finkelstein, D.** 1987. Improvement of enzyme production in *Aspergillus*. Antonie Van Leeuwenhoek. **53**: 349-352.
16. **Flipphi, M.J., van Heuvel, M., van der Veen, P., Visser, J. and de Graaff, L.H.** 1993. Cloning and characterization of the *abfB* gene coding for the major *α*-L-arabinofuranosidase (ABF B) of *Aspergillus niger*. Curr Genet. **24**: 525-532.
17. **Foreman, P.K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., et al.** 2003. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. J Biol Chem. **278**: 31988-31997.
18. **Frey-Wyssling, A.** 1954. The fine structure of cellulose microfibrils. Science. **119**: 80-82.
19. **Gomes, J., Gomes, I.I., Terler, K., Gubala, N., Ditzelmuller, G. and Steiner, W.** 2000. Optimisation of culture medium and conditions for *α*-L-arabinofuranosidase production by the extreme thermophilic eubacterium *Rhodothermus marinus*. Enzyme Microb Technol. **27**: 414-422.
20. **Gruben, B.S.** 2012. Novel transcriptional activators of *Aspergillus* involved in plant biomass utilization. Utrecht University.
21. **Hasper, A.A., Dekkers, E., van Mil, M., van de Vondervoort, P.J. and de Graaff, L.H.** 2002. EglC, a new endoglucanase from *Aspergillus niger* with major activity towards xyloglucan. Appl Environ Microbiol. **68**: 1556-1560.
22. **Kusters-van Someren, M.A., Harmsen, J.A., Kester, H.C. and Visser, J.** 1991. Structure of the *Aspergillus niger pelA* gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. Curr Genet. **20**: 293-299.
23. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage *T4*. Nature. **227**: 680-685.
24. **Lagopodi, A.L., Ram, A.F.J., Lamers, G.E.M., Punt, P.J., Van den Hondel, C.A.M.J.J., Lugtenberg, B.J.J., et al.** 2002. Novel Aspects of Tomato Root Colonization and Infection by *Fusarium oxysporum* f. sp. *radicis-lycopersici* Revealed by Confocal Laser Scanning Microscopic Analysis Using the Green Fluorescent Protein as a Marker. Molecular Plant-Microbe Interactions. **15**: 172-179.

25. **Mai, V., Wiegel, J. and Lorenz, W.W.** 2000. Cloning, sequencing, and characterization of the bifunctional xylosidase-arabinosidase from the anaerobic thermophile *Thermoanaerobacter ethanolicus*. Gene. **247**: 137-143.
26. **Margolles-Clark, E., Tenkanen, M., Nakari-Setala, T. and Penttila, M.** 1996. Cloning of genes encoding *α*-L-arabinofuranosidase and *β*-xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. Appl Environ Microbiol. **62**: 3840-3846.
27. **Miyazaki, K.** 2005. Hyperthermophilic *α*-L-arabinofuranosidase from *Thermotoga maritima* MSB8: molecular cloning, gene expression, and characterization of the recombinant protein. Extremophiles. **9**: 399-406.
28. **Mullaney, E., Hamer, J., Roberti, K., Yelton, M.M. and Timberlake, W.** 1985. Primary structure of the *trpC* gene from *Aspergillus nidulans*. Molecular and General Genetics MGG. **199**: 37-45.
29. **Peraza, L., Ortiz, M.A., Peberdy, J.F. and Aguilar, G.** 2003. Growth and pectinase production by *Aspergillus* Mexican strain protoplast regenerated under acidic stress. Appl Biochem Biotechnol. **111**: 15-27.
30. **Poutanen, K.** 1997. Enzymes: An important tool in the improvement of the quality of cereal foods. Trends in Food Science & Technology. **8**: 300-306.
31. **Punt, P.J., Dingemanse, M.A., Kuyvenhoven, A., Soede, R.D.M., Pouwels, P.H. and van den Hondel, C.A.M.J.J.** 1990. Functional elements in the promoter region of the *Aspergillus nidulans gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. Gene. **93**: 101-109.
32. **Punt, P.J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J. and van den Hondel, C.** 2002. Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol. **20**: 200-206.
33. **Quay, D.H.X., Bakar, F.D.A., Rabu, A., Said, M., Illias, R.M., Mahadi, N.M., Hassan, O., and Murad, A.M.A.** 2011. Overexpression, purification and characterization of the *Aspergillus niger* endoglucanase, EglA, in *Pichia pastoris*. African Journal of Biotechnology. **10**: 2101-2111.
34. **Remond, C., Plantier-Royon, R., Aubry, N., Maes, E., Bliard, C. and O'Donohue, M.J.** 2004. Synthesis of pentose-containing disaccharides using a thermostable *α*-L-arabinofuranosidase. Carbohydr Res. **339**: 2019-2025.
35. **Rémond, C., Ferchichi, M., Aubry, N., Plantier-Royon, R., Portella, C. and O'Donohue, M.J.** 2002. Enzymatic synthesis of alkyl arabinofuranosides using a thermostable *α*-L-arabinofuranosidase. Tetrahedron Letters. **43**: 9653-9655.
36. **Rombouts, F.M., Voragen, A.G.J., Searle-van Leeuwen, M.F., Geraeds, C.C.J.M., Schols, H.A. and Pilnik, W.** 1988. The arabinanases of *Aspergillus niger* - Purification and characterisation of two *α*-L-arabinofuranosidases and an endo-1,5-*α*-L-arabinanase. Carbohydrate Polymers. **9**: 25-47.

37. **Saha, B.C.** 2003. Hemicellulose bioconversion. J Ind Microbiol Biotechnol. **30**: 279-291.
38. **Sambrook, J. and Russell, D.W.** Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 2001.
39. **Shi, P., Li, N., Yang, P., Wang, Y., Luo, H., Bai, Y., et al.** 2010. Gene Cloning, Expression, and Characterization of a Family 51 α-L-Arabinofuranosidase from *Streptomyces sp*. S9. Appl Biochem Biotechnol. **162**: 707-718.
40. **van Peij, N.N., Gielkens, M.M., de Vries, R.P., Visser, J. and de Graaff, L.H.** 1998. The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. Appl Environ Microbiol. **64**: 3615-3619.
41. **Wang, L., Ridgway, D., Gu, T. and Moo-Young, M.** 2005. Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. Biotechnol Adv. **23**: 115-129.

Supplemental Figure 1A: **Protein Sequence Alignment for *A. vadensis* α-L-arabinofuranosidase B (AvAbfB) and *A. niger* α-arabinofuranosidase A (AnAbfA).**

A.

```
AvAbfB      MFSRRNLLALGLAATVSAGPCDIYEAGDTPCVAAHSTTRALYSSFSGALYQLQRGSDDTT
AnAbfA      MFSSTL-VSLAV--SWLAGQALAIELTV-SKTGGNSSSPLLYG----VMFEDINNSGDGG
            ***    ::*.:  :  ** .   *      ....*::  **.    .:::  . * *

AvAbfB      TTISPLTAGGVA-DASAQDTFCANTTCLITIIYDQ-----SGNGNHLTQAP---PGGFDG
AnAbfA      IHGQVLRNNGFQGDDPGLTAYSAVGN--VTISQDTANPLSSAITSTLKVAVPSGATGYVG
               . *   *.  *  .  ::.*  .  :**  *      *.  . *. *      *: *

AvAbfB      PDTDGYDNLASAIGAPVTLNGQKAYGVFMSPG--TGYRNNEATGTATGDEAEGMYAVLDG
AnAbfA      FANEGYDGVPV------LGTDYDNY--FYMKGDYSGTVNLRLV----GNSSGIIYADHNI
              .:*** :          .  . *  *   *  :*  * . .    *:.:  :**  :

AvAbfB      THYNDACCFDYGNAETSSTDTGAGHMEAIYLGNSTTWGYGAGDG-------PWIMVDME-
AnAbfA      TVASTSSNFTYYETSFSSSVS---------SDAHNVWQLRFDASKVAGSSLNFGLVQLFP
            *  . :. * * ::. **: :             ..*      .        : :*::

AvAbfB      ---NNLFSGADEGYNSGDPSISYSFVTA------------------------AVKGGAD
AnAbfA      PTYNSRYNGLRDDVASFLADIKPSFLRFPGGNNLEGATPSDRWKWNETIGPVVDRPGREG
               *. :.*  :   *   .*. **:                              *

AvAbfB      KWAIRGGNAASGSLSTYYS-----GARPDYSGYNPMSKEGAIILGIGGDNSNGAQGTFYE
AnAbfA      DWTYPNTDAL--GLDEYLQWCEDMDMEPLLAVWSGLSLGGGIVSGS-------ALDPYVD
            .*:    :*   .*. * .       .*  : :. :*  *.*: *        *   : :

AvAbfB      GVMTSGYPSDDTENSVQENIVAAKYVVGSLVSGPSFTSGEVVSLRVTTPGYTTRYIAHTD
AnAbfA      DI----------LNE------LEQYVLGSADT----TYGSLRAKNGRTEPWDVKYLEVGN
             :           *.        :**:**  :    * *.: : .  *  : .:*:   :

AvAbfB      T-TVNTQVVDDDSSTTL------KEEASWTVVTGLANSQCFSFESVDTPGSYIRHY----
AnAbfA      EDNLNSGCGTYANRFTLIYDAVHAAYPNLTIVASTTDTSCLPSTIPDGVITDIHHYLTPD
              .:*:      .  **          . *:*:. :::.*:     *   : *:**

AvAbfB      NFELLLNANDGTKQFHEDAT--FCPQAALNGEGTSLRSWSYPTRYF--------------
AnAbfA      EFIDLFDE---WDNWSRDWPILVGEYASTTGNDGSTTYWSYMQGSCSEAVYMIGMERNSD
            :*  *::     .:: .*    .   *: .*:  *   ***

AvAbfB      -----------RHYDNVLYAASNGGVQTFDSKTSFNNDVSFEIETAFSS-----------
AnAbfA      IVKMASFAPLLEHFDMAEWSPDLFGLDS--SPDSVTGSTSYYVQKMFSTNRGSTVLPVNT
                       .*:* . :: .  *:::  *  *.. ..*: ::. **:

AvAbfB      ------------------------------------------------------------
AnAbfA      TADFDPLYWVASVSDEGTYYVKLANYGSSSQNVTVNIEGTTSGQLQMLSGGETVSNYPHD


AvAbfB      --------------------------------
AnAbfA      VSITTQSSTVSGSGSFTVDMPAWAVAVLAVS
```

22% amino acid sequence identity.

Supplemental Figure 1B: Protein Sequence Alignment for *A. vadensis endo*-1,4-*β*-D-glucanase A (AvEglA) and *A. niger endo*-1,4-*β*-D-glucanase A (AnEglA).

B.

```
AvEglA      MKLAVTLSMLAATAMGQTMCSQYDSASSPPYSVNQNLWGEYQGTGSQCVYVDKLSSSGAS
AnEglA      MKLPVSLAMLAATAMGQTMCSQYDSASSPPYSVNQNLWGEYQGTGSQCVYVDKLSSSGAS
            ***.*:*:**************************************************

AvEglA      WHTKWTWSGGEGTVKSYSNSGLTFDKKLVSDVSSIPTSVEWKQDNTNVNADVAYDLFTAA
AnEglA      WHTEWTWSGGEGTVKSYSNSGVTFNKKLVSDVSSIPTSVEWKQDNTNVNADVAYDLFTAA
            ***:*****************:**:***********************************

AvEglA      NVDHATSSGDYELMIWLARYGYIQPIGKQIATATVGGKSWEVWYGTSIQAGAEQRTYSFV
AnEglA      NVDHATSSGDYELMIWLARYGNIQPIGKQIATATVGGKSWEVWYGSTTQAGAEQRTYSFV
            ******************** ********************:: ************

AvEglA      SESPINSYSGDINAFFSYLTQNQGFPASSQYLINLQFGTEPFTGGPATFTVDNWTASVN
AnEglA      SESPINSYSGDINAFFSYLTQNQGFPASSQYLINLQFGTEAFTGGPATFTVDNWTASVN
            ************************************.*****************
```

95% amino acid sequence identity.

Chapter 6

An evolutionary screen to increase inulinase production in *Aspergillus oryzae*

Culleton, H.M., Majoor, E., McKie, V.A. and de Vries, R.P. 2014. In preparation.

Abstract

Inulin is found widely distributed in nature as a storage polysaccharide and consists of a linear polymer of *β*-1,2-linked D-fructose molecules which can be hydrolyzed by *endo*-/*exo*-inulinases and fructofuranosidase (invertase) to give D-fructose and fructooligosaccharides. Fructose and fructooligosaccharides have gained considerable interest recently as important ingredients in the food and pharmaceutical industry due to their high sweetening capacity combined with their many functional and nutritional properties. Conventional methods of fructose production have proven to be both costly and inefficient with attention now being directed towards the use of microbial inulinases as a more promising approach for obtaining optimum D-fructose yields. In this study we aimed to improve the inulin degradation potential of *Aspergillus oryzae* through the upregulation of *exo*-inulinase production using an evolutionary screening method. As an organism with no predicted *endo*-inulinase function, improved inulin degradation would be largely dependent on the overproduction of this enzyme. Subsequent generation growth of *Aspergillus oryzae* (Rib40) on inulin for 9 weeks successfully resulted in *exo*-inulinase overproducing mutants.

Introduction

Inulin is widely distributed in nature as a storage carbohydrate in the roots and tubers of several plants such as dandelion, chicory, dahlia, Jerusalem artichoke and burdock [15]. It contains a linear polymer of *β*-1,2-linked D-fructose, connected to a terminal sucrose residue [11] and can be hydrolyzed by inulinases (*endo-* / *exo-*) [5] and fructofuranosidase (invertase) [17], resulting mainly in the formation of D-fructose and fructooligosaccharides. Fructose is the sweetest, naturally occurring sugar compound identified to date and consequently is emerging fast as an important ingredient in the food and pharmaceutical industry. Fructose is a more favourable alternative sweetener to sucrose due to its higher sweetening capacity, low viscosity and beneficial effects associated with diabetic patients and iron absorption in children [12]. Furthermore, fructooligosaccharides have many positive functional and nutritional properties i.e. reduced caloric value, stimulation of intestinal bifidobacteria and as a source of dietary fibre in food preparations [9, 14]. Conventional methods of fructose production involving the enzymatic degradation of starch followed by the conversion of D-glucose to D-fructose using glucose isomerase yields only ~42% D-fructose, with the remaining end-products consisting of glucose (~50%) and oligosaccharides (~8%) [14]. Other techniques, such as ion exchange chromatography, have been developed for fortification of fructose but all add to the production cost [18]. The use of microbial inulinases has therefore been proposed as a more favourable and promising approach for obtaining optimum D-fructose yields.

In a recent study, the plant polysaccharide degrading potential was compared for three different *Aspergillus* species i.e. *A. nidulans*, *A. niger* and *A. oryzae* [2]. Comparison of the CAZy (Carbohydrate-Active enzyme database (CAZy-http://www.cazy.org) [8]) content of the genomes of these fungi demonstrated significant differences, with the number of genes related to particular polysaccharide degradation varying substantially between the three species. *A. niger* has the highest number of inulin-related genes in its genome with 5 open reading frames (ORF's), compared to *A. nidulans* and *A. oryzae*, which have 2 ORF's and 4 ORF's respectively. Growth profiles for these species on inulin, compared on the Fungal Growth Database website (www.fung-growth.org), showed a correlation with genome content, with growth of *A. niger* exceeding that of both *A. nidulans* and *A. oryzae* on this carbon source. While such poor growth could be expected for *A. nidulans*, the genome of which had less than half the number of ORF's of *A. niger*, the results were somewhat surprising for *A. oryzae*. Further analysis into the function predictions of these ORF's showed that of the three species, *A. niger* was the only one to contain an *endo*-inulinase within its genome, thus enhancing its ability for efficient inulin degradation. Based on this information, *A. nidulans* and *A. oryzae* would be required to over-produce *exo*-inulinase in order to utilise this substrate efficiently, thus resulting in

an increased yield of D-fructose. This suggests that these Aspergilli, although closely related, have a unique biological approach towards the enzymatic degradation of the same substrate and presents the question if this approach could be optimised through prolonged cultivation on inulin rather than by genetic manipulation.

The use of this type of evolutionary screen has been very effective in obtaining (industrial) microorganisms with improved phenotypes, such as an expanded substrate range, increased stress tolerance and efficient substrate utilisation [10, 13]. In this study we have looked at an evolutionary approach to develop a strain with increased *exo*-inulinase activity. We have taken the filamentous fungi, *A. oryzae* and *A. nidulans* and generationally grown them on inulin for 9 weeks to obtain *exo*-inulinase overproducing mutants.

Materials and Methods

Strains, media and culture conditions

Aspergillus nidulans (FGSCA4) and *Aspergillus oryzae* (Rib40) were used in this study. Both strains were taken from glycerol stocks stored at -45°C and grown on MEA (Malt Extract Agar) prior to the evolution experiments. All plates were grown at 30°C.

Aspergillus minimal medium (MM) was described previously [4]. Agar was added at 2% (w/v) for solid medium. For the evolution experiments, MM solid agar was supplemented with 1% (w/v) inulin (dahlia tubers) (Sigma; Cat. No. I3754). For the analysis and characterization of enzyme activities, 400 mL liquid cultures (2.5 L baffled flasks) of MM + 1% (w/v) inulin (dahlia tubers) (Sigma; Cat. No. I3754) / inulin (chicory) (Sigma: Cat. No. I2255) were inoculated with 1 x 10^6 spores/mL (final) and incubated at 30°C in an orbital shaker at 250 rpm for 4 days.

Preparation of evolution strains

From the initial regeneration on MEA agar, spores from *A. nidulans* and *A. oryzae* strains were harvested in 10 ml ACES (*N*-(2-Acetamido)-2-aminoethanesulfonic acid) buffer (0.2% (w/v) ACES + 0.02% (w/v) TWEEN® 80, pH 6.0) and 200 µL was used to inoculate the evolution specific agar (MM + 1% (w/v) inulin) in triplicate for each strain. All plates were incubated at 30°C. After one week, growth was examined and 100-200 µL of spores was used to inoculate new inulin plates. This process was repeated for a period of 9 weeks after which time the strains were purified on inulin by diluted plating and selecting the best growing colony. The purified strains were stored in 30% (w/v) glycerol at -80°C.

Enzyme profile comparison by SDS-PAGE

400 ml liquid cultures were inoculated with 1 x 10^6 spores/mL (final) and were grown in 2.5 L baffled flasks at 30°C and 250 rpm for 4 days. Following incubation, the mycelia were removed from the culture filtrate by filtration over a Büchner funnel with nylon gauze. SDS-PAGE was done using a 12% polyacrylamide gel containing 0.1% (w/v) SDS with 15 µL culture filtrate samples being loaded for each expression strain. Protein bands were detected by Coomassie Blue staining [7].

Enzyme activity measurements

For the measurement of enzyme activities from the *A. oryzae* generation strains, an initial 1:10 dilution of the culture filtrates was carried out in 100 mM sodium acetate buffer, pH 4.5 including BSA (1 mg/mL). Para-nitrophenol (*p*NP) assays were used for the measurement of *α*-/*β*-galactosidase and *α*-/*β*-glucosidase activities. Initial *α*-/*β*-galactosidase assays were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 10 mM *p*-nitrophenyl *α*-D-galactopyranoside (pH 4.5) / *β*-D-galactopyranoside (pH 4.5) respectively and incubated at 40 °C for 60 min. Similarly, for the initial measurement of *α*-/*β*-glucosidase activities, 0.2 mL of diluted enzyme was added in duplicate to 0.2 mL of 10 mM *p*-nitrophenyl *α*-D-glucopyranoside (pH 4.5) / *β*-D-glucopyranoside (pH 4.5) respectively and incubated at 40 °C for 60 min. The reactions were stopped with the addition of 3 mL of 2 % (w/v) tri-sodium orthophosphate, pH 12.0 and the absorbance was measured at 400 nm. Activities were expressed as Units/ml where one unit is defined as 1 micromole of *p*-nitrophenol liberated per minute per millilitre of culture filtrate. Depending on results obtained, further time-points were conducted to ensure that activity rates were linear.

Nelson-Somogyi reducing sugar assays were used for the measurement of *exo*-inulinase activities. Two different substrates were used in these experiments for comparison purposes; High MW (Molecular Weight) Inulin (Raftiline; HP Inulin) and Kestose (Megazyme; Cat. No. O-KTR). Initial *exo*-inulinase assays were performed in duplicate by the addition of 0.2 mL of diluted enzyme to 0.2 mL of 20 mg/mL (final) Inulin (pH 4.5) / 5 mg/mL (final) Kestose (pH 4.5) respectively and incubated at 40°C for 60 min. The reactions were stopped by the addition of 0.5 mL of Stopping Solution (25 mL of Solution A (2.5% (w/v) sodium carbonate anhydrous, 2.5% (w/v) potassium sodium tartrate and 20% (w/v) sodium sulphate), to which 1 mL of Solution B (3% (w/v) copper sulphate pentahydrate) was added). Enzyme reactions were then boiled for 20 min and allowed to cool to room temperature for 5 min before the addition of 3.0 mL of a 1:5 dilution of Solution C (5% (w/v) ammonium molybdate, 4.2% (v/v) concentrated sulphuric acid and 0.6% (w/v) sodium arsenate heptahydrate). Depending on results obtained, further time-points were conducted to ensure that activity rates were linear.

Results and Discussion

Development of evolution strains

Aspergillus nidulans (FGSCA4) and *Aspergillus oryzae* (Rib40) were grown on inulin for 9 weeks with weekly re-inoculation over which time the growth rate of *A. oryzae* increased (data not shown). This effect was not seen for *A. nidulans*, indicating that the basal levels of enzymes required to degrade this carbon source had not increased to the same degree as they had in *A. oryzae*.

To further demonstrate this hypothesis, the *exo*-inulinase activity levels for both strains at the beginning and end of the generation experiment were tested. *A. oryzae* week 1 (*Ao*1.1) and 9 (*Ao*9.1) and *A. nidulans* week 1 (*An*1.1) and 9 (*An*9.1) generation strains were chosen, with the same biological clone being used for both time points to give an accurate comparison. Strains were grown in media containing inulin from dahlia tubers and incubated at 30°C for 4 days. Initial *exo*-inulinase activity levels for both generations of each strain were compared, with *Ao*9.1 giving higher than double the activity levels (~125 mU/min/mL) of that obtained for *Ao*1.1 (~60 mU/min/mL). The activity levels measured for *An*1.1 and *An*9.1 remained unchanged and consistent with that obtained for *Ao*1.1, providing a strong correlation with the growth profiles.

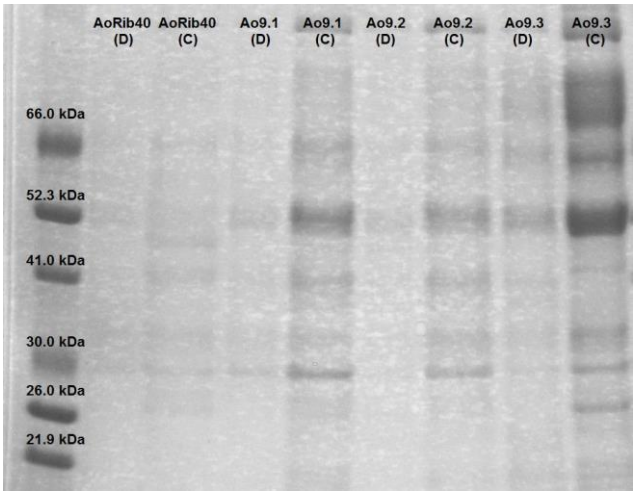


Figure 1: SDS-PAGE analysis of *A. oryzae* (Rib40) and week 9 evolved biological triplicate strains (*Ao*9.1, *Ao*9.2 and *Ao*9.3) when grown on MM + 1% inulin (dahlia tubers (D)/chicory (C)) at 30°C, 250 rpm for 4 days. Low molecular weight (LMW) ladder produced in house.

Extracellular enzyme profile and activity comparison

Based on the results obtained from the initial enzyme activity measurements, the extracellular activities in the *A. oryzae* mutant strains were examined in more detail. *A. oryzae* (Rib40) and the biological triplicates generated from week 9 (*Ao*9.1, *Ao*9.2 and *Ao*9.3) were chosen for this study, with *A. oryzae* (Rib40) acting as a control and indicative of the basal levels of enzyme activities expected when *A. oryzae* wild type (WT) is grown on inulin as a carbon source. Inulin from both dahlia tubers and chicory were examined to determine whether the precise structure of the inulin affects the enzyme mixtures produced by the mutants.

After 4 days of growth, the extracellular enzyme profiles generated for each strain were visualised by SDS-PAGE (Fig. 1). The *A. oryzae* mutant strains all demonstrated a clear increase in enzyme production levels when grown on inulin compared to the wild type (Rib40) with notable differences also being present between the three mutants. This again, was in correlation with the growth profiles of these strains with *Ao*9.3 generating more dense mycelia and a more concentrated enzyme profile in inulin rich liquid cultures than the other strains. Further evaluation of the enzyme profiles showed an intense band at approximately 52 kDa (Fig. 1), which likely corresponds to the extracellular *exo*-inulinase (InuE), for which the calculated molecular mass in *A. oryzae* Rib40 (AO090701000400) is ~55 kDa. The other visible protein bands are consequently expected to be multiple fructofuranosidases (invertases) i.e. AO090020000640, AO090701000038 and AO090103000043, the molecular weights which are ~64 kDa, 69 kDa and 121 kDa respectively [2]. Interestingly, higher enzyme levels were generated on inulin from chicory than from dahlia tubers, the latter being the source of inulin used for the generation studies. Previous studies using analytical HPLC to determine the chain lengths of both these commercial preparations of inulin showed that inulin from dahlia contains a higher molecular weight range which in turn has been known to affect solubility [6]. This indicates that while the basic structure of both carbon sources are identical and require the same enzyme actions for degradation, the inulin from chicory may be more easily utilised by the fungal cell.

In combination with the visual profiles of extracellular enzyme production during growth on inulin, the levels of *exo*-inulinase activity were measured in duplicate and compared for the different strains. The activities of α- and β-galactosidase and α- and β-glucosidase were also measured in duplicate as a control and indicative of the basal level of change in production of un-associated enzyme activities. In correlation to what was discovered with the enzyme profiles, the overall levels of all enzyme activities were higher for the *A. oryzae* mutant strains than for the wild type (Fig. 2). This was especially the case for *Ao*9.3, which gave a 4 and 9 fold increase in *exo*-inulinase activity on inulin and kestose, respectively. Kestose contains more terminal ends than

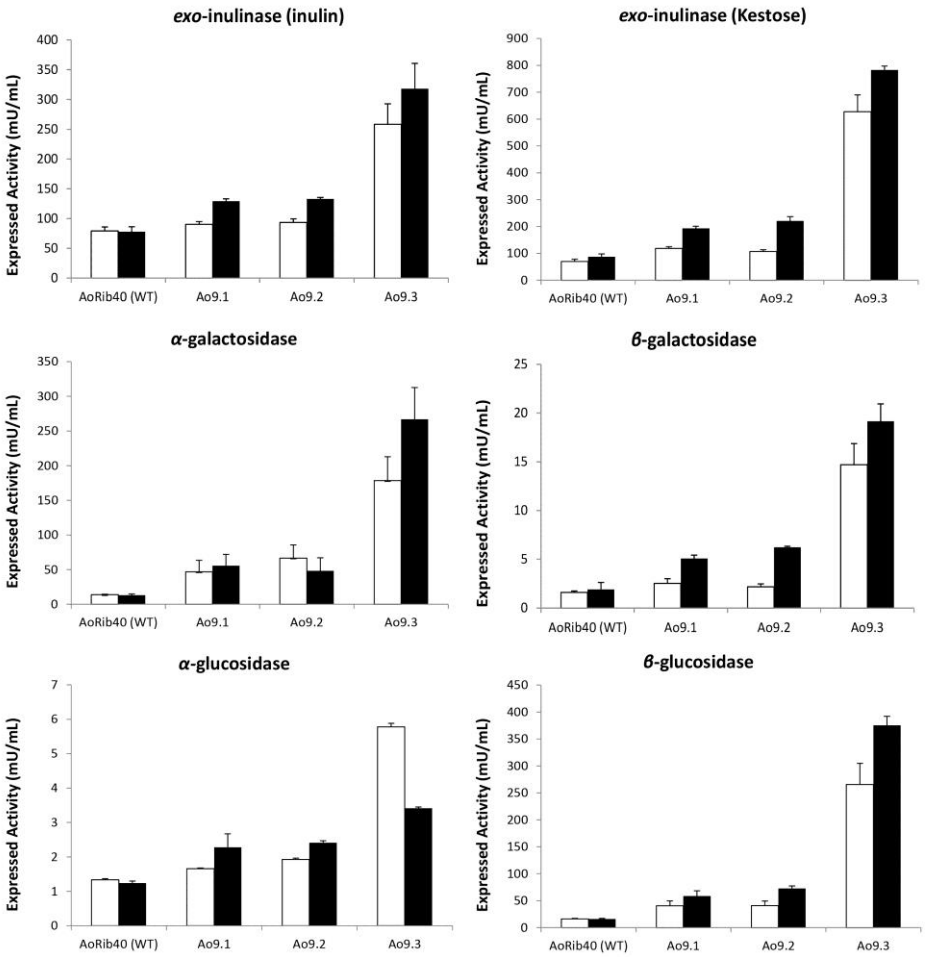


Figure 2: Measured activity values obtained from Day 4 culture filtrate samples from the indicated week 9 *A. oryzae* generation strains. *Ao*Rib40 was used as a control and indicative of the basal levels of activities generated when *A. oryzae* wild type (WT) is grown on inulin as a carbon source. Error bars calculated on standard deviations between technical duplicates. All polysaccharide substrates and *p*-nitrophenol substrates were tested at 40°C, pH 4.5 under conditions as described in Materials and Methods. White = Dahlia inulin, Black = chicory inulin.

inulin, thus confirming the *exo*-acting properties of the enzyme being over-produced. An average of a 12 fold increase in activities for the control enzymes was also noted indicating that the evolution screening not only increased the levels of the required enzyme activities i.e. *exo*-inulinase, but also the overall growth rate and secretion levels of these strains when grown on inulin. This could indicate that instead of a specific *exo*-inulinase increase in the mutants, enzyme secretion as whole has been improved in these strains.

Another hypothesis involves the up-regulation of genes associated with AmyR in response to the complete degradation of inulin to fructose and glucose. AmyR controls genes involved in starch degradation but it has been noted in *A. niger* that it also regulates genes encoding *beta*-glucosidases and *alpha*- and *beta*-galactosidases [16]. While starch and maltose are the most common inducers of this regulator, glucose also induces AmyR activation in *A. niger* and *A. oryzae* [1, 16]. If, like in *A. niger*, AmyR is also responsible for the regulation of these otherwise un-associated genes in *A. oryzae*, the complete degradation of inulin may then also cause an increase in the associated enzyme activities. In *A. niger* fructose is a relatively weak repressing carbon source [3], while glucose at higher concentrations can cause significant repression of many genes encoding poly- and oligosaccharide degrading enzymes. Degradation of inulin would only result in very low glucose levels that may result in AmyR activation.

The mutants were maintained for several generations on MEA and then tested again on inulin without losing their improved properties, signifying that the changes were genetically stable and not dependent on a continuous selection. Future studies will analyse the nature of the mutations in more detail.

References

1. **Carlsen, M. and Nielsen, J.** 2001. Influence of carbon source on alpha-amylase production by *Aspergillus oryzae*. Appl Microbiol Biotechnol. **57**: 346-349.
2. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.
3. **de Vries, R.P., vanKuyk, P.A., Kester, H.C. and Visser, J.** 2002. The *Aspergillus niger faeB* gene encodes a second feruloyl esterase involved in pectin and xylan degradation and is specifically induced in the presence of aromatic compounds. Biochem J. **363**: 377-386.
4. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
5. **Derycke, D.G. and Vandamme, E.J.** 1984. Production and properties of *Aspergillus niger* inulinase. Journal of Chemical Technology and Biotechnology Biotechnology. **34**: 45-51.
6. **Hincha, D.K., Hellwege, E.M., Heyer, A.G. and Crowe, J.H.** 2000. Plant fructans stabilize phosphatidylcholine liposomes during freeze-drying. Eur J Biochem. **267**: 535-540.
7. **Laemmli, U.K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage *T4*. Nature. **227**: 680-685.
8. **Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B.** 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. **42**: 21.
9. **Niness, K.R.** 1999. Inulin and oligofructose: What are they? J Nutrit. **129**: 1402s-1406s.
10. **Patnaik, R.** 2008. Engineering complex phenotypes in industrial strains. Biotechnol Progr. **24**: 38-47.
11. **Pavis, N., Chatterton, N.J., Harrison, P.A., Baumgartner, S., Praznik, W., Boucaud, J., et al.** 2001. Structure of fructans in roots and leaf tissues of *Lolium perenne*. New Phytologist. **150**: 83-95.
12. **Pawan, G.** 1973. Fructose. In: Birch GG, Green, L.F., (ed) Molecular structure and function of food carbohydrates, Applied Science Publishers, Barking, Essex, UK, pp. pp 65-80.
13. **Sauer, U.** 2001. Evolutionary engineering of industrially important microbial phenotypes. Adv Biochem Eng Biotechnol. **73**: 129-169.

14. **Singh, P. and Gill, P.K.** 2006. Production of inulinases: recent advances. Food Technol Biotechnol. **44**: 151-162.
15. **Vandamme, E.J. and Derycke, D.G.** 1983. Microbial inulinases: fermentation process, properties, and applications. Adv Appl Microbiol. **29**: 983.
16. **vanKuyk, P.A., Benen, J.A., Wosten, H.A., Visser, J. and de Vries, R.P.** 2012. A broader role for AmyR in *Aspergillus niger*: regulation of the utilisation of D-glucose or D-galactose containing oligo- and polysaccharides. Appl Microbiol Biotechnol. **93**: 285-293.
17. **Wallis, G.L., Hemming, F.W. and Peberdy, J.F.** 1997. Secretion of two beta-fructofuranosidases by Aspergillus niger growing in sucrose. Arch Biochem Biophys. **345**: 214-222.
18. **Zittan, L.** 1981. Enzymatic hydrolysis of inulin — An alternative way to fructose production. Starch - Stärke. **33**: 373-377.

Chapter 7

Summary and general discussion

Plant biomass is the most abundant and usable carbon source for most fungal species on the planet. Primarily consisting of plant cell walls, it contains two main groups of polysaccharides; plant cell wall polysaccharides (cellulose, hemicelluloses [xyloglucans, xylan, galacto(gluo)mannan] and pectin) and storage polysaccharides (e.g. starch and inulin). Plant cell wall polysaccharides interact with each other as well as the aromatic polymer, lignin, to form a network of linkages and hydrogen bonds that give the plant cell wall its rigidity. Fungi degrade these polysaccharides extracellularly by secreting enzymes which release oligo- and monosaccharides which can then be utilized by the fungal cell [15]. Fungi have optimized this process by producing diverse enzymatic mixtures, specific to the polysaccharides which they encounter. The fine-tuned production of such diverse enzyme sets is controlled through a set of transcriptional regulators, of which there is predicted to be between 400-600 in *Aspergillus* alone, that either activate (e.g. AmyR, XlnR, AraR, InuR, GalR/GalX, RhaR, ManR, ClbR) or repress (CreA) the expression of the genes encoding plant biomass degrading enzymes [25].

Aspergilli have become one of the most widely studied group of filamentous fungi due to their potential in plant biomass degradation and in many industrial applications [4, 5, 9, 13, 35]. As common soil fungi, found in many different environments, Aspergilli, are capable of producing an extensive set of enzyme mixes and degrade a very broad range of polysaccharides [10]. Much research over the past two decades has been applied to the development of *Aspergillus*, most notably *A. niger* and *A. oryzae*, as hosts for recombinant protein production. Their natural ability to grow at high rates and to high biomass densities and their exceptional capacity to secrete high levels of homologous product are well recognized [10, 14]. While significant progress has been made in the hyper-production of heterologous proteins, yields have not yet reached the levels obtained for homologous proteins [12]. This may in part be caused by the acidification of the media due to the production of organic acids in the case of *A. niger*, different native protein glycosylation patterns or by the presence of high levels of secreted proteases which effectively degrade heterologous proteins [41]. Regardless of the broad use of other fungal systems e.g. cellulase production in *Penicillium* sp. [1, 8] and *Trichoderma reesei* [18, 27] and development of new industrial fungal enzyme producers, such as *Myceliophthora thermophila* [40, 44], *Aspergillus* has remained of major interest for applied research. This is reflected by the number of species with a published genome sequence [3, 22, 30, 32-34, 39, 46]; http://genome.jgi-psf.org/programs/fungi/genome-releases.jsf, http://www.ncbi.nlm.nih.gov/ bioproject] and the recent initiation of a project to sequence every species of the genus, as well as comparative studies of these genomes addressing plant biomass degradation [10], **Chapter 2**].

Plant biomass degradation in *Aspergillus*

Plant polysaccharide degradation by fungi has been a topic of study for many decades due to its relevance for many industrial applications, such as paper & pulp [9], food [35], feed [5], beverages, textiles [4] and detergents. More recently, the increasing interest in biomass-based fuels and chemicals has accelerated research into fungal decomposition of plant biomass. The availability of an increasing number of fungal genome sequences has encouraged genome mining of many species to identify new and/or better enzymes than those produced by established enzyme producers, such as *A. niger* and *T. reesei* [11]. By comparison, less attention has been given to the diverse strategies i.e. both enzyme-encoding genes and regulatory systems [16], employed by fungi themselves to efficiently degrade plant biomass.

In **Chapter 2**, eight *Aspergillus* species were compared with respect to their genomic ability to degrade plant cell wall biomass. While all tested Aspergilli had a similar potential to degrade plant biomass, results showed that even in closely related species, their strategies differed markedly. The variation in CAZy content of the genome of these species was relatively small, likely due to their close phylogenetic relationships and their similar habitats. Hierarchical clustering of the plant polysaccharide degrading enzymes of these species demonstrated that species with the most similar CAZy content in their genome were also taxonomically close. In contrast to this correlation, a high level of variation was detected in the enzymes produced by the tested species, especially during growth on sugar beet pulp. Sugar beet pulp contains mainly cellulose and pectin and therefore unsurprisingly, pectinases and cellulases were found to be the main enzymes produced on this substrate. The noted differences in pectinolytic gene content between the species could partially be explained by pH preference of the associated enzymes i.e. *A. niger*, which strongly acidifies the culture medium [34], has an evolutionary preference to pectin hydrolase production (acidic pH optima) over pectin lyases (neutral to alkali pH optima). This genomic pH preference to pectin degradation was also found to be prevalent in the thermophilic fungi *Myceliophthora thermophila* and *Thielavia terrestris*, where *T. terrestris*, which has 3.5 times more pectin hydrolases than lyases in its genome, was found to grow best on pectin at an acidic pH, while *M. thermophila*, which has 3.5 times more lyases than hydrolases in its genome, was found to favour a neutral to alkaline pH for growth on pectin [6]. However, as *A. niger* was the only species in this study to strongly acidify the medium, the variability associated with enzyme production for the remaining species was more likely be due to regulatory differences at the transcriptional level, suggesting that specific regulator function and/or range of target genes could be unique to different *Aspergillus* species. Recent research by Meijer *et. al*. (2011) suggests that this variability and uniqueness between different *Aspergillus* species is indeed accurate. In a study that compared growth and

extracellular enzyme profiles of wild type *A. niger* isolates from different biotopes all over the world and type strains of other black Aspergilli, it was found that irrespective of the environment and carbon source that a particular strain was conditioned to, adaptation to the natural environment does not occur at the genetic level for *A. niger* and its ability to utilise various carbon sources remains consistent across the isolates of this species. In contrast, interspecies variation with respect to carbon source utilisation was significant, especially on D-galactose where *A. brasiliensis* was unique in its ability to grow on this substrate suggesting a distinct difference in sugar transport mechanisms [42] or metabolism.

The unique phenotype of *A. vadensis* compared to other black Aspergilli

A. vadensis, a recently identified member of the black Aspergilli and a close relative of *A. niger*, has been suggested as being beneficial for recombinant protein production as it does not acidify the culture medium and produces very low levels of extracellular proteases [17]. In recent studies, Punt *et. al.* (2008) identified the PrtT regulatory protein, a homologue of which was found to be present in many *Aspergillus* species and which was shown to govern expression of many secreted protease genes, including the major alkaline protease *alpA* and neutral protease *Np1* in *A. oryzae* and the aspergillopepsin encoding gene *pepA* in *A. niger*. In addition, a gene cluster related to starch degradation, including the *amyR* regulatory gene, is present directly upstream of the *prtT* locus, separated by only two genes (*agdA/aglU* and *amyA*), both of which are target genes of AmyR in *A. niger* [37].

In **Chapter 3**, the molecular and phenotypical differences between *A. vadensis* and six other species of black Aspergilli i.e. *A. niger*, *A. acidus*, *A. aculeatus*, *A. carbonarius*, *A. brasiliensis* and *A. tubingensis* were examined. Growth profiles showed significant and unique differences between *A. vadensis* and the other black Aspergilli with no growth being observed for *A. vadensis* on maltose or starch. The unusual growth profile for *A. vadensis* when grown on these carbon sources could be explained by a mutation/deletion within the *amyR* regulatory gene. Growth on casein however, was comparably poor for all the black Aspergilli, with the exception of *A. niger*, indicating that this more common protease deficiency may be accounted for at a transcriptional level. Genome analysis, however, demonstrated that the *prtT* and *amyR* regulatory loci and also the *agdA/aglU* and *amyA* genes were well conserved among all the black Aspergilli, including *A. vadensis*. Combined with gene expression data indicating very poor levels of expression for both *prtT* and *amyR*, along with the associated *amyA* gene, the aberrant phenotype of *A. vadensis* is likely caused by the low expression of *prtT* and *amyR* regulators or their associated genes and not a mutation or deletion of the structural part of these genes as was originally concluded,

thus reinforcing the unique abilities and adaptations of individual and often closely related *Aspergillus* species in the efficient degradation of plant biomass.

Optimisation of recombinant enzyme production in *A. vadensis*

Despite the common use of several *Aspergillus* species, most notably *A. niger* and *A. oryzae* for the industrial production of proteins [19], the yields for heterologous protein production still remain low. This is in part, due to different native protein glycosylation patterns and the presence of high levels of secreted proteases which effectively degrade many heterologous proteins [41]. Genetic strategies, including the development of highly inducible or strong constitutively active promoters, have been shown to lead to an increase in the levels of recombinant protein production.

In **Chapter 4**, six novel constitutive promoters from *A. niger* (p*ef1α*, p*tktA*, p*ef1β*, p*tal1*, p*cetA* and p*pgkA*) and a further five from *A. vadensis* (p*ef1α*, p*rps31*, p*gpdA*, p*ubi1* and p*oliC*) were tested in *A. vadensis* using a gene encoding a secreted arabinofuranosidase from *Fusarium oxysporum* as a reporter for heterologous protein production. When comparing just the ABF expression levels obtained under the control of the different promoters, it was found that 9 out of the 11 promoters tested resulted in higher ABF activity levels than those observed under the control of the widely characterized *gpdA* promoter from *A. nidulans* [20, 21, 28, 29]. However, as multiple insertions of up to and greater than 10 copies are common in *Aspergillus* and are found to affect recombinant protein production, both positively [43] and negatively [24], it was necessary to determine the copy number of expression vectors transformed into the individual strains in order to gain a complete understanding of the actual strength of the tested promoter constructs. Interestingly, the order of promoter strength altered notably with 3 of the promoters from *A. niger* (p*ef1α* > p*tal* > p*pgkA*) and 3 from *A. vadensis* (p*ef1α* > p*oliC* > p*rps31*) then resulting in higher ABF activity than the *gpdA* promoter from *A. nidulans*.

The over-production of target genes by different promoters in *Aspergillus* is also receiving attention in other studies. Only recently, Li *et.al.* (2014) published research accounting the over-expression of ß-glucuronidase (GUS) (*uidA*) when fused to the promoter region encoding the *ß*-glucosidase II gene (*bglII*), with reported GUS specific activity of 189 U/mg. Interestingly however, when fused with a smaller segment of this promoter region, much higher GUS specific activity of 448 U/mg was reported [26]. This method of utilising sections of promoter regions is not a new one however with Minetoki *et. al.* (1998) describing a similar method whereby multiple copies of the conserved sequence region III in the promoter regions of the amylase-encoding genes (*amyB*, *glaA* and *agdA*) of *A. oryzae* were introduced into the *agdA* promoter, resulting in both a significant increase in promoter activity at the transcriptional level and in expression of the *agdA* target gene [31].

In **Chapter 5**, the full length gDNA encoding two plant biomass degrading enzymes i.e. *α*-L-arabinofuranosidase (*abfB*) (GH54) and *endo*-1,4-*β*-D-glucanase (*eglA*) (GH12) from *A. vadensis* were successfully expressed using the *gpdA* promoter from *A. vadensis*. Both enzymes were produced extracellularly in *A. vadensis* as soluble proteins and successfully purified by affinity chromatography, giving a maximum yield of 30 mg/L for AbfB and 25 mg/L for EglA. It should be noted that the *A. vadensis* strain used in these experiments was near wild type and had not undergone the extensive strain improvement strategies that have resulted in the commercial *A. niger* and *T. reesei* strains that secrete much higher levels of enzyme, suggesting room for improvement. With an initial expression yield of ~ 30 mg/L, *A. vadensis* shows positive signs of becoming an industrially significant contender but considerable improvements will be required to rival the secretion abilities of commercial production strains such as *A. niger*, *A. oryzae* and *T. reesei*. Strategies for improving recombinant protein production in fungi, including the use of strong promoters and effective secretion signals or gene fusions to genes associated with well-expressed and secreted proteins have been implemented successfully to other systems in the past [36, 45]. Other methods, such as UV-mediated mutagenesis have been successfully applied to *Aspergillus* strains to generate ultra-high producing mutants [23]. Successes have also been obtained using a bioprocessing approach of optimizing fungal morphology, mycelia immobilization and culture conditions [2, 7].

To date, the use of His-tag affinity chromatography in fungal systems has not been practical due, in part to the difficulties experienced with degradation of the target protein and/or Histidine residues by extracellular proteases [41], especially as was found to be the case with *A. niger* (unpublished data). Purification of recombinant enzymes from *A. vadensis* through His-tag affinity chromatography have worked well thus far (**Chapter 5**; a further eight *A. vadensis* proteins were also produced and purified in this way (unpublished data)), suggesting its potential as a versatile host in the fundamental research of proteins as well as for industrial enzyme production.

AbfB from *A. vadensis* shared characteristics with the previously reported *A. niger* AbfB [38]. However, comparative analysis of EglA from *A. vadensis* to the orthologous EglA from *A. niger* demonstrated significant variations in biochemical characteristics, despite these organisms being taxonomically very close. In particular, the difference in temperature optima and stability for these orthologous enzymes is noteworthy and would directly impact the applications of these enzymes. Screening of orthologs from related fungi may therefore be worthwhile for many industrial applications and suggests that it is not always necessary to go to distantly related species or enzyme classifications to get significant changes in biochemical properties.

Optimisation of native enzyme production in *Aspergillus*

Inulin is found widely distributed in nature as a storage polysaccharide and consists of a linear polymer of *β*-1,2-linked D-fructose molecules which can be hydrolyzed by *endo*-/*exo*-inulinases and fructofuranosidase (invertase) to give D-fructose and fructooligosaccharides. Fructose and fructooligosaccharides have gained considerable interest recently as important ingredients in the food and pharmaceutical industry due to their high sweetening capacity combined with their many functional and nutritional properties. Conventional methods of fructose production have proven to be both costly and inefficient with attention now being directed towards the use of microbial inulinases as a more promising approach for obtaining optimum D-fructose yields.

In **Chapter 6** an evolutionary screening method was used to improve the inulin degradation potential of *Aspergillus oryzae* through the upregulation of *exo*-inulinase. As an organism with no predicted *endo*-inulinase function, improved inulin degradation would be largely dependent on the overproduction of this enzyme. Subsequent generation growth of *Aspergillus oryzae* (Rib40) on inulin for 9 weeks successfully resulted in *exo*-inulinase overproducing mutants. This study demonstrates the genomic flexibility of fungi and their potential for improved enzyme production. The approach used to improve *exo*-inulinase production can be applied for many other enzymes and is mainly dependent on the design of a good screen to push the strains in the right direction. The result that the evolved strains are genetically stable demonstrates that the improvements are genetic rather than pleiotropic. The approach introduced in this study is therefore an attractive alternative to GMO methods, in particular with respect to use in food applications.

Conclusion

The aim of this PhD project was to examine and optimise the plant polysaccharide degradation abilities of *Aspergillus*. It revealed the highly diverse enzymatic strategies for the degradation of plant biomass used by closely related fungi that appeared to have equal efficiency. This demonstrated that it is not always necessary to go to distantly related species or enzyme classifications to get significant changes in biochemical properties between the enzymes produced by these fungi. Therefore, the screening of orthologs from related fungi and/or the identification of enzyme sets employed by different fungi can be used to design better and combined commercial enzyme cocktails for many industrial applications, such as the bio-fuel industry.

While much research has been applied to the development of *Aspergillus* as hosts for recombinant protein production, selecting the most efficient strategy for the construction of a particular expression strain remains a complex task. Factors such as control mechanisms (transcription and regulation) and functional properties (post-translational modifications and toxicity) of the target gene are required before expression strain aspects, such as constitutive vs. inducible promoters, effective secretion signals, carrier proteins and host strain physiology can be considered. In this thesis, the development of a strong constitutive promoter for expression of homologous (*α*-L-arabinofuranosidase and *endo*-1,4-*β*-D-glucanase from *A. vadensis*) and heterologous (*α*-arabinofuranosidase from *Fusarium oxysporum*) proteins on a simple carbon source such as sucrose is defined. The potential of protease deficient hosts, such as *A. vadensis*, is examined and emphasized with successful and efficient production and purification of recombinant enzymes facilitating many downstream processes in industrial enzyme production. Finally, the over-production of native fungal enzymes through generation screening methods offer an alternative to genetic approaches for the optimisation of the plant biomass degrading potential of fungi, such as *Aspergillus.*

References

1. **Adsul, M., Bastawde, K., Varma, A. and Gokhale, D.** 2007. Strain improvement of *Penicillium janthinellum* NCIM 1171 for increased cellulase production. Bioresource Technology. **98**: 1467-1473.

2. **Agger, T., Spohr, A.B., Carlsen, M. and Nielsen, J.** 1998. Growth and product formation of *Aspergillus oryzae* during submerged cultivations: verification of a morphologically structured model using fluorescent probes. Biotechnol Bioeng. **57**: 321-329.

3. **Andersen, M.R., Salazar, M.P., Schaap, P.J., van de Vondervoort, P.J., Culley, D., Thykaer, J., et al.** 2011. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. Genome Res. **21**: 885-897.

4. **Battaglia, E., Benoit, I., Gruben, B.S. and de Vries, R.P.** 2010. Plant cell wall derived sugars as substrates for fungi and industry. In: Jenkins PT, (ed) The Sugar Industry and Cotton Crops, Nova Science Publishers, New York.

5. **Bedford, M.R. and Classen, H.L.** 1991. The influence of dietary xylanase on intestinal viscosity and molecular weight distribution of carbohydrates in rye-fed broiler chicks. Progress in biotechnology. **7**: 361-370.

6. **Berka, R.M., Grigoriev, I.V., Otillar, R., Salamov, A., Grimwood, J., Reid, I., et al.** 2011. Comparative genomic analysis of the thermophilic biomass-degrading fungi *Myceliophthora thermophila* and *Thielavia terrestris*. Nat Biotechnol. **29**: 922-927.

7. **Carlsen, M., Spohr, A.B., Nielsen, J. and Villadsen, J.** 1996. Morphology and physiology of an alpha-amylase producing strain of *Aspergillus oryzae* during batch cultivations. Biotechnol Bioeng. **49**: 266-276.

8. **Chaabouni, S.E., Belguith, H., Hassairi, I., M'Rad, K. and Ellouz, R.** 1995. Optimization of cellulase production by *Penicillium occitanis*. Appl Microbiol Biotechnol. **43**: 267-269.

9. **Christov, L.P., Szakacs, G. and Balakrishnan, H.** 1999. Production, partial characterization and use of fungal cellulase-free xylanases in pulp bleaching. Process Biochemistry. **34**: 511-517.

10. **Coutinho, P.M., Andersen, M.R., Kolenova, K., vanKuyk, P.A., Benoit, I., Gruben, B.S., et al.** 2009. Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. Fungal Genet Biol. **46 Suppl 1**: S161-S169.

11. **Couturier, M., Haon, M., Coutinho, P.M., Henrissat, B., Lesage-Meessen, L. and Berrin, J.G.** 2010. *Podospora anserina* hemicellulases potentiate the *Trichoderma reesei* secretome for saccharification of lignocellulosic biomass. Appl Environ Microbiol. **77**: 237-246.

12. **Culleton, H.M., McKie, V.A. and de Vries, R.P.** 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from *Aspergillus*? Biotechnology Journal. **8**: 884-894.
13. **Davies, R.W.** 1994. Heterologous gene expression and protein secretion in *Aspergillus*. Prog Ind Microbiol. **29**: 527-560.
14. **de Vries, R.P.** 1999. Accessory enzymes from *Aspergillus* involved in xylan and pectin degradation. Wageningen Agricultural University.
15. **de Vries, R.P. and Visser, J.** 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev. **65**: 497-522.
16. **de Vries, R.P.** 2003. Regulation of *Aspergillus* genes encoding plant cell wall polysaccharide-degrading enzymes; relevance for industrial production. Appl Microbiol Biotechnol. **61**: 10-20.
17. **de Vries, R.P., Burgers, K., van de Vondervoort, P.J., Frisvad, J.C., Samson, R.A. and Visser, J.** 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production. Appl Environ Microbiol. **70**: 3954-3959.
18. **Durand, H., Clanet, M. and Tiraby, G.** 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme and Microbial Technology. **10**: 341-346.
19. **Finkelstein, D.** 1987. Improvement of enzyme production in *Aspergillus*. Antonie Van Leeuwenhoek. **53**: 349-352.
20. **Flaherty, J.E. and Payne, G.** 1997. Overexpression of aflR Leads to Upregulation of Pathway Gene Transcription and Increased Aflatoxin Production in *Aspergillus flavus*. Appl Environ Microbiol. **63**: 3995-4000.
21. **Flipphi, M., Mathieu, M., Cirpus, I., Panozzo, C. and Felenbok, B.** 2001. Regulation of the aldehyde dehydrogenase gene (*aldA*) and its role in the control of the coinducer level necessary for induction of the ethanol utilization pathway in *Aspergillus nidulans*. Journal of Biological Chemistry. **276**: 6950-6958.
22. **Galagan, J.E., Calvo, S.E., Cuomo, C., Ma, L.J., Wortman, J.R., Batzoglou, S., et al.** 2005. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature. **438**: 1105-1115.
23. **Ishida, H., Matsumura, K., Hata, Y., Kawato, A., Suginami, K., Abe, Y., et al.** 2001. Establishment of a hyper-protein production system in submerged *Aspergillus oryzae* culture under tyrosinase-encoding gene (*melO*) promoter control. Appl Microbiol Biotechnol. **57**: 131-137.
24. **Kelly, J.M. and Hynes, M.** 1987. Multiple copies of the *amdS* gene of *Aspergillus nidulans* cause titration of trans-acting regulatory proteins. Current genetics. **12**: 21-31.
25. **Kowalczyk, J.E., Benoit, I. and de Vries, R.P.** 2014. Regulation of plant biomass utilization in *Aspergillus*. Advances in Applied Microbiology. **88**: 31-55.

26. **Li, C.-H. and Yan, T.-R.** 2014. Use of *Aspergillus niger β*-glucosidase II gene (*bgl* II) promoter elements to construct an efficient expression vector. Journal of the Taiwan Institute of Chemical Engineers. **45**: 749-754.
27. **Liming, X. and Xueliang, S.** 2004. High-yield cellulase production by *Trichoderma reesei* ZU-02 on corn cob residue. Bioresource Technology. **91**: 259-262.
28. **Lombraña, M., Moralejo, F.J., Pinto, R. and Martín, J.F.** 2004. Modulation of *Aspergillus awamori* thaumatin secretion by modification of *bipA* gene expression. Appl Environ Microbiol. **70**: 5145-5152.
29. **Lubertozzi, D. and Keasling, J.D.** 2008. Expression of a synthetic Artemesia annua amorphadiene synthase in *Aspergillus nidulans* yields altered product distribution. J Ind Microbiol Biotechnol. **35**: 1191-1198.
30. **Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., et al.** 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature. **438**: 1157-1161.
31. **Minetoki, T., Kumagai, C., Gomi, K., Kitamoto, K. and Takahashi, K.** 1998. Improvement of promoter activity by the introduction of multiple copies of the conserved region III sequence, involved in the efficient expression of *Aspergillus oryzae* amylase-encoding genes. Appl Microbiol Biotechnol. **50**: 459-467.
32. **Nierman, W.C., Pain, A., Anderson, M.J., Wortman, J.R., Kim, H.S., Arroyo, J., et al.** 2005. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. Nature. **438**: 1151-1156.
33. **Payne, G.A., Nierman, W.C., Wortman, J.R., Pritchard, B.L., Brown, D., Dean, R.A., et al.** 2006. Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Medical Mycology. **44**: 9-11.
34. **Pel, H.J., de Winde, J.H., Archer, D.B., Dyer, P.S., Hofmann, G., Schaap, P.J., et al.** 2007. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. Nat Biotechnol. **25**: 221-231.
35. **Poutanen, K.** 1997. Enzymes: An important tool in the improvement of the quality of cereal foods. Trends in Food Science & Technology. **8**: 300-306.
36. **Punt, P.J., van Biezen, N., Conesa, A., Albers, A., Mangnus, J. and van den Hondel, C.** 2002. Filamentous fungi as cell factories for heterologous protein production. Trends Biotechnol. **20**: 200-206.
37. **Punt, P.J., Schuren, F.H.J., Lehmbeck, J., Christensen, T., Hjort, C. and van den Hondel, C.A.M.J.J.** 2008. Characterization of the *Aspergillus niger* PrtT, a unique regulator of extracellular protease encoding genes. Fungal Genetics and Biology. **45**: 1591-1599.
38. **Rombouts, F.M., Voragen, A.G.J., Searle-van Leeuwen, M.F., Geraeds, C.C.J.M., Schols, H.A. and Pilnik, W.** 1988. The arabinanases of *Aspergillus niger* -

Purification and characterisation of two α-L-arabinofuranosidases and an endo-1,5-α-L-arabinanase. Carbohydrate Polymers. **9**: 25-47.

39. **Sato, A., Oshima, K., Noguchi, H., Ogawa, M., Takahashi, T., Oguma, T., et al.** 2011. Draft Genome Sequencing and Comparative Analysis of *Aspergillus sojae* NBRC4239. DNA Research. **18**: 165-176.

40. **Singh, B.** 2014. *Myceliophthora thermophila* syn. *Sporotrichum thermophile*: a thermophilic mould of biotechnological potential. Critical reviews in biotechnology. 1-11.

41. **van den Hombergh, J.P.T.W., van de Vondervoort, P.J.I., Fraissinet-Tachet, L. and Visser, J.** 1997. *Aspergillus* as a host for heterologous protein production: the problem of proteases. Trends in Biotechnology. **15**: 256-263.

42. **vankuyk, P., Diderich, J., MacCABE, A., Hererro, O., RUIJTER, G. and Visser, J.** 2004. *Aspergillus niger mstA* encodes a high-affinity sugar/H+ symporter which is regulated in response to extracellular pH. Biochem J. **379**: 375-383.

43. **Verdoes, J.C., Punt, P. and Van Den Hondel, C.** 1995. Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. Appl Microbiol Biotechnol. **43**: 195-205.

44. **Visser, H., Joosten, V., Punt, P.J., Gusakov, A.V., Olson, P.T., Joosten, R., et al.** 2011. Development of a mature fungal technology and production platform for industrial enzymes based on a *Myceliophthora thermophila* isolate, previously known as *Chrysosporium lucknowense* C1. Industrial Biotechnology. **7**: 214-223.

45. **Wang, L., Ridgway, D., Gu, T. and Moo-Young, M.** 2005. Bioprocessing strategies to improve heterologous protein production in filamentous fungal fermentations. Biotechnol Adv. **23**: 115-129.

46. **Wortman, J.R., Fedorova, N., Crabtree, J., Joardar, V., Maiti, R., Haas, B.J., et al.** 2006. Whole genome comparison of the *A. fumigatus* family. Medical Mycology. **44**: 3-7.

Nederlandse samenvatting

Planten biomassa is de meest abundante en bruikbare koolstof bron voor de meeste schimmels op aarde. Schimmels breken de polysacchariden in dit substraat af door enzymen uit te scheiden die mono- en oligosacchariden vrijmaken. Deze worden vervolgens opgenomen en gebruikt dor de schimmel cel. Aspergilli zijn schimmels die in veel verschillende biotopen voorkomen en zij produceren een uitgebreide set enzymen die een breed scala aan polysacchariden kunnen afbreken. In de afgelopen twee decennia is veel onderzoek verricht om *Aspergillus* te ontwikkelen als een gastheer voor de productie van recombinante enzymen. Meer recent is dit onderzoek, in het bijzonder gericht op de afbraak van planten biomassa, in een versnelling gekomen door de sterk toenemende behoefte aan brandstof en chemicaliën gemaakt uit planten biomassa.

In **hoofdstuk 2** zijn acht *Aspergillus* species vergeleken met betrekking tot hun genomisch potentieel voor afbraak van planten biomassa. Ondanks dat dit potentieel vergelijkbaar was in de species, de resultaten lieten zien dat zelfs nauw verwante species hiervoor sterk verschillende strategieën gebruiken. De variatie in de CAZy genen in de genomen was relatief klein en hiërarchische clustering hiervan volgde de taxonomische relaties van de species. In tegenstelling hiermee werd een hoge variatie geobserveerd in de enzymen die door de species geproduceerd werden. The verschillen met betrekking tot pectinolytische genen in de genomen konden deels verklaard worden door de pH voorkeur van de gerelateerde enzymen. Echter, de verschillend werden vooral toegewezen aan regulatoire verschillen, wat suggereert dat de functie van een regulator of de set van genen die deze controleert verschilt in de species.

In **hoofdstuk 3** werden de moleculaire en fysiologische eigenschappen van *Aspergillus vadensis* in vergelijking met andere zwarte Aspergilli bestudeerd. *A. vadensis*, een recentelijk geïdentificeerde species van de zwarte Aspergilli, is een interessante kandidaat als producent van heterologe eiwitten, aangezien deze species het medium niet verzuurd en zeer lage niveaus van extracellulaire proteases heeft. Met uitzondering van de groei op maltose en zetmeel werden geen significante groeiverschillen gevonden met alle andere species. Groei op het eiwit caseine was slecht voor alle species met uitzondering van *Aspergillus niger*. Genoom analyse toonde aan dat de *prtT* en *amyR* regulatoire genen, alsmede de *agdA*/*aglU* en *amyA* genen, geconserveerd waren in alle zwarte Aspergilli. Op basis van expressie data werd duidelijk dat het fenotype van *A. vadensis* op maltose, zetmeel en eiwit hoogstwaarschijnlijk veroorzaakt is door lage expressie van de regulator genen *amyR* en *prtT*. Dit onderstreept de unieke aanpassingen in nauw verwante *Aspergillus* species met betrekking tot efficiënte afbraak van planten biomassa.

In **hoofdstuk 4** zijn zes nieuwe constitutieve promotoren van *A. niger* (p*ef1α*, p*tktA*, p*ef1β*, p*tal1*, p*cetA* and p*pgkA*) en vijf van *A. vadensis* (p*ef1α*, p*rps31*, p*gpdA*, p*ubi1* and p*oliC*) getest voor enzym productie in *A. vadensis* waarbij een gen coderend voor een gesecreteerd arabinofuranosidase van *Fusarium oxysporum* als reporter is gebruikt. Op basis van ABF activiteit bleek dat 9 van de 11 promotoren betere productie gaven in vergelijking met de veel gebruikte *gpdA* promoter van *Aspergillus. nidulans*. Analyse van het copy-nummer van de transformanten had een significant effect op de bepaalde sterkte van de promotoren, waarbij slechts drie promotoren van *A. niger* (p*ef1α* > p*tal* > p*pgkA*) en drie van *A. vadensis* (p*ef1α* > p*oliC* > p*rps31*) tot hogere ABF activiteit leiden dan *gpdA* van *A. nidulans*.

In **hoofdstuk 5** is de succesvolle expressie van genen coderend voor α-L-arabinofuranosidase (*abfB*, GH54) en *endo*-1,4-*β*-D-glucanase (*eglA*, GH12) beschreven. Beide enzymen werden extracellulair als oplosbare eiwitten geproduceerd en succesvol gezuiverd via affiniteit chromatografie, met een opbrengst van 30 mg/L voor AbfB en 25 mg/L voor EglA. AbfB was functioneel vergelijkbaar met de eerder beschreven AbfB van *A. niger*. Echter, vergelijking van *A. vadensis* en *A. niger* EglA toonde significante verschillen aan in hun biochemische eigenschappen, ondanks de nauwe verwantschap van deze twee species. Verschillen werden vooral gevonden in temperatuur optima en stabiliteit wat direct relevant is voor de toepassing van deze enzymen. Dit suggereert dat de vergelijking van orthologen uit verwante schimmels tot relevante data kan leiden voor industriële applicaties en dat het niet altijd nodig is om naar onverwante species te gaan voor enzymen met afwijkende eigenschappen.

In **hoofdstuk 6** is een evolutionaire screening methode gebruikt om het inuline afbraak potentieel van *A. oryzae* door de verhoging van *exo*-inulinase activiteit. Aangezien dit organisme geen *endo*-inulinases lijkt te produceren, zou verbeterde afbraak van inuline vooral afhankelijk zijn van overproductie van *exo*-inulinase. Groei van *Aspergillus oryzae* (RIB40) via herhaaldelijke aanenten op inuline gedurende 9 weken resulteerde in *exo*-inulinase overproducerende mutanten. Deze studie toont de genomische flexibiliteit van schimmels aan en het potentieel hiervan voor verbeterde enzym productie. Aangezien de geëvolueerde stammen stabiel zijn is het waarschijnlijk dat de oorzaak van de verhoogde productie genetisch is en niet pleiotropisch. Dit laat zien dat de aanpak van deze studie een aantrekkelijk alternatief is voor GMO methoden, vooral met betrekking tot gebruik in levensmiddelen toepassingen.

Curriculum Vitae

Helena Marie Culleton was born on April the 3rd, 1986 in Wexford, Ireland. She followed her second education in St. Mary's College, Arklow and graduated in 2004 with an honours Leaving Certificate. In September of the same year she began her study in Genetics and Cell Biology at Dublin City University where she did her first internship in molecular biology at Megazyme International Ireland in Bray, Co. Wicklow under the supervision of Dr. R.M. Lloyd and Dr. S.J. Charnock. As part of her final year, she did her second internship in the National Institute for Cellular Biotechnology (NICB), Dublin under the supervision of Prof. C.E. Loscher. Helena obtained her honours BSc degree in 2008. In June of the same year she began employment as a Molecular Biologist/Biochemist with Megazyme International Ireland and began her PhD in Fungal Physiology in conjunction with Utrecht University, The Netherlands and CBS-KNAW Fungal Biodiversity Centre, The Netherlands under the supervision of Prof. Dr. ir. R.P. de Vries and Dr. V.A. McKie.

List of Publications

Culleton, H.M., McKie, V.A. and de Vries, R.P. 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from Aspergillus? Biotechnology Journal. **8**: 884-894.

Culleton, H.M., Bouzid. O., McKie, V.A. and de Vries, R.P. 2014. New promoters to improve heterologous protein production in *Aspergillus vadensis*. Curr. Biotechnol. **3**: 1-8.

Culleton, H.M., McKie, V.A. and de Vries, R.P. 2014. Overexpression, purification and characterization of homologous α-L-arabinofuranosidase and *endo*-1,4-β-D-glucanase in *Aspergillus vadensis*. Ind. Micro. & Biotech. **41**: 1697-1708.

Abstracts and Presentations

Bouzid, O., **Culleton, H.M.,** McKie, V.A., Mc Cleary, B.V. and de Vries, R.P. – The 9th International *Aspergillus* meeting (ASPERFEST9). March 29-30, 2012, Marburg, Germany. (Poster presentation).

Benoit, I., **Culleton, H.M.,** Wiebenga, A., Coutinho, P.M., Brouwer, C.P.J.M., McKie, V.A., McCleary, B.V., Henrissat, B. and de Vries, R.P. – The 11th European Conference on Fungal Genetics (ECFG11). March 30-April 2, 2012, Marburg, Germany. (Poster presentation).

Acknowledgements

I can't believe it's almost finished!!!! The last words of the thesis....yet I am finding it quite difficult to leave it so. It seems that the closer it gets to putting it all in print, the more I feel the need to change one last figure, restructure one last sentence or correct the page numbers which continuously insist on doing their own thing!!! Yet, as the end draws near and I am reading through the past 4 years' research one last time, it has really emphasised the number of amazing people that I have come in contact with and who have helped me enormously during this time. To each and every one of you, I very humbly say Thank You!!!

Barry and Vincent, I truly am very grateful to you both for giving me the opportunity and confidence to see this project though to the end. Barry, you have such admirable passion for science and industry. Thank you for knowing what questions to ask and for having the drive to be one step ahead with the answers. Megazyme is a wonderful company to work for and do research in – thank you for allowing me the opportunity to do both!! Vincent, thank you for your advice, support and time; for the organisation of the enzyme assays at the beginning to the proof reading of one paper after another at the end. Your practical organisational skills and critical eye were very much appreciated.

Ronald, I think I can very safely say that, without you, this research and thesis could not have happened!! Thank you for your love of science and your enthusiasm for research – it is truly motivational! Thank you for everything; from the countless email assistance and the Dutch translations to the Saturday morning Skype updates and the late night journal submissions. Thank you also for making me feel so welcome during my trips to the Netherlands, both with your research group and with lovely home cooked duck dinners!!! Thank you for your ongoing support in every aspect; it has been a wonderful and rewarding experience!!

Bridget, you have been not only an amazing colleague to me over the past 4 years but a fantastic friend also!! Thank you for the pep talks and for keeping me sane during the tough times but equally, thank you for introducing me to the 'happy science dance' during the good times!! For the time, support, assistance, troubleshooting and encouragement every step of the way – thank you!! Thank you for the hours and evenings you gave to proof-read every single chapter for me, one by one – your genuine interest, well wishes and stamp of approval meant the world to me!!

There are also quite a few other people I wish to thank for both their scientific and social contributions, as both played equally important roles in my life. I would like to thank my colleagues for the great atmosphere in and around the lab, for the inspiring and fun moments and for all the help that I received: Agnija, Alfred, Anna, Bobby, Cat, Claudio, Clodagh, Conor, David, Edd, Eibhlín, Gary, Geraldine, Ida, Iain, Jamie, Jenny, John A., John E., Kim, Maciej, Naomi, Niall, Orlaith, Paraic, Paul, Patricia,

Ruth, Ruth C., Santana, Sinéad (org.), Sinéad, Shirley, Stuart, Sue and Tom. Special thanks to the girls who I shared the office/offices with; Bridget, Cat, Jenny, Naomi, and Orlaith; thank you for the fun, the spontaneous Avoca days and pep talks during the writing period. To Iain, thank you for being such a good friend to me over the past few years – you were always there to lend a bit of support and encouragement, no matter the evening. To Eileen, thank you for your amazing organisational skills; for looking after the travel arrangements to and from the Netherlands and always being such a breath of fresh air!! To Kim, thank you so much for your time and help in getting everything 'print ready'. You're a whizz-kid on Photoshop and InDesign – I would have found this part very difficult without your help!!

I would also like to thank my Dutch colleagues for making me feel so welcome when I was visiting and the help that I received finding my way around: Ad, Alexa, Alexandra, Birgit, Eline, Evy, Isabelle, Joanna, Joost, Manon, Marina, Miaomiao, Ourdia, Sebnem and Sylvia. Special thanks to Ad for helping me so much around the lab; for pointing me in the right direction when I couldn't find what I was looking for and for always being only a click away when I needed samples shipped and advice with protocols. To Isa, thank you for always making me feel so welcome to the group; for your scientific help, especially with Chapter 2 and your lovely liquorice tea on the way to visit the old ruins!! Your warm smile always made me feel very at home!! To Ourdia, thank you for always being available to help no matter when or where. For all your help and work on Chapter 4, for your expressed interest in 'Irish Dancing' and for sharing an appreciation for animated movies with me!! To Miaomiao, Alexa and Joanna, thank you all for your work and help with getting Chapter 3 finished and ready for print. It was a race to the finishing line and I could not have done it without your help. To Manon, thank you for your expert eye and for willingly offering your time to help me with the formatting at the end!!

I would like to thank my friends outside of work for the ongoing support, interest energy and love. To Ciara, Derv, Jenny D. and all the CrossFit gang – thank you for keeping me on an even keel!! To Liam, words cannot express how much I want to thank you!! You have been such a blessing to my life. Thank you for your support and encouragement all those many evenings when I was sitting in the other room, stuck in my laptop. Thank you for not only understanding the craziness but for encouraging it!! I love you!! To my brothers Shane, Mark and Darragh - thank you for your interest and trying to understand the crazy science!! And finally to my parents, thank you for everything!!! For the unconditional love, support, encouragement, positivity and perspective not only in the last four years but in the 24 before that too!! Thank you for the genuine interest in my work and for learning off what I do for a living!! ☺ Thanks for giving me deadlines and goals to achieve so that I kept motivated to the end. Thank you for everything!! I love you both!!