Fungal glycoside hydrolase family 44 xyloglucanases are restricted to the phylum Basidiomycota and show a distinct xyloglucan cleavage pattern

Sun, Peicheng

2022-01-21

Sun , P , Li , X , Dilokpimol , A , Henrissat , B , de Vries , R P , Kabel , M A & Mäkelä , M R 2022 , ' Fungal glycoside hydrolase family 44 xyloglucanases are restricted to the phylum Basidiomycota and show a distinct xyloglucan cleavage pattern ' , iScience , vol. 25 , no. 1 , 103666 . https://doi.org/10.1016/j.isci.2021.103666

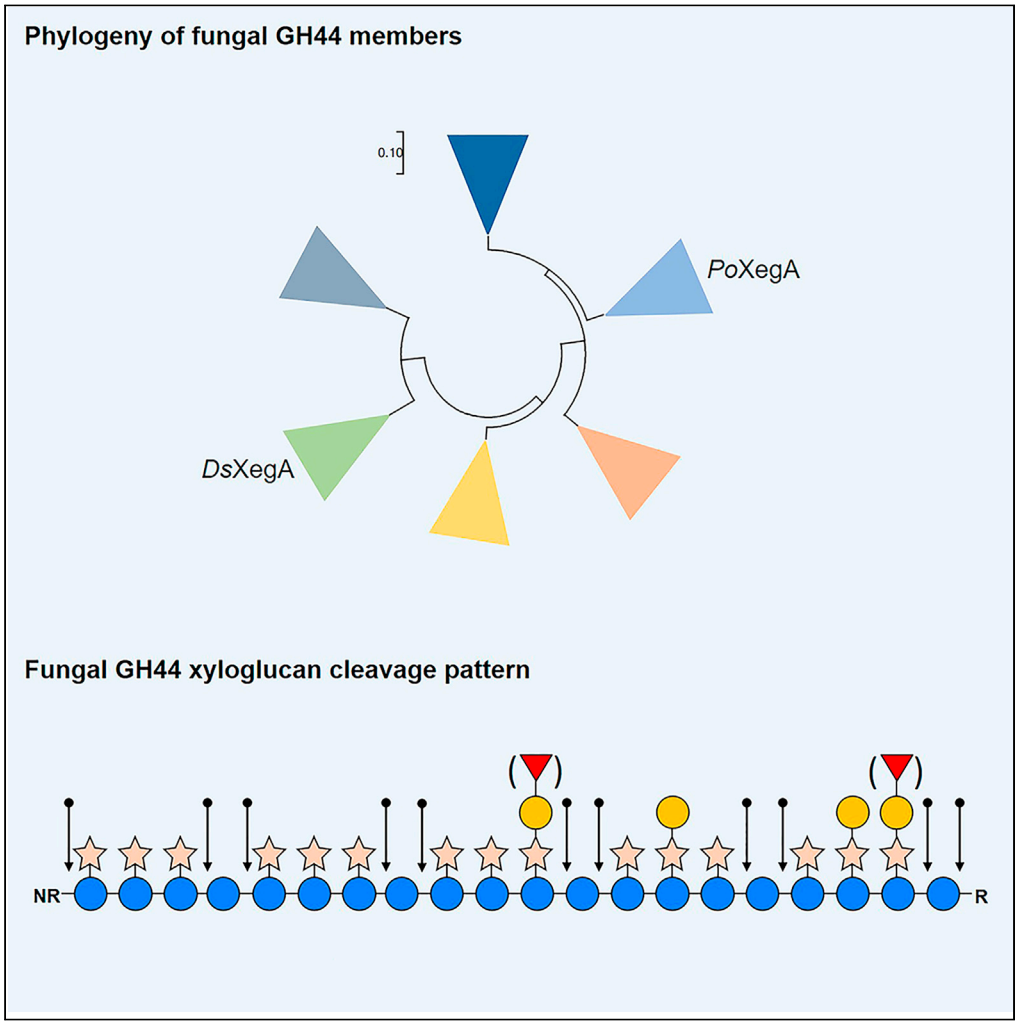
http://hdl.handle.net/10138/341887
https://doi.org/10.1016/j.isci.2021.103666

iScience

CellPress
OPEN ACCESS

Article

Fungal glycoside hydrolase family 44 xyloglucanases are restricted to the phylum Basidiomycota and show a distinct xyloglucan cleavage pattern



Peicheng Sun, Xinxin Li, Adiphol Dilokpimol, Bernard Henrissat, Ronald P. de Vries, Mirjam A. Kabel, Miia R. Mäkelä

miia.r.makela@helsinki.fi

Highlights

The fungal members of GH44 only belong to the phylum Basidiomycota

The characterized fungal GH44 xyloglucanases (XEGs) are xyloglucan-specific

These XEGs cleave xyloglucan at both sides of unsubstituted glucosyl units

Sun et al., iScience 25, 103666
January 21, 2022
https://doi.org/10.1016/j.isci.2021.103666

iScience

CellPress
OPEN ACCESS

Article

Fungal glycoside hydrolase family 44 xyloglucanases are restricted to the phylum Basidiomycota and show a distinct xyloglucan cleavage pattern

Peicheng Sun,[1,6] Xinxin Li,[2,6] Adiphol Dilokpimol,[2] Bernard Henrissat,[3,4] Ronald P. de Vries,[2] Mirjam A. Kabel,[1] and Miia R. Mäkelä[5,7,*]

SUMMARY

Xyloglucan is a prominent matrix heteropolysaccharide binding to cellulose microfibrils in primary plant cell walls. Hence, the hydrolysis of xyloglucan facilitates the overall lignocellulosic biomass degradation. Xyloglucanases (XEGs) are key enzymes classified in several glycoside hydrolase (GH) families. So far, family GH44 has been shown to contain bacterial XEGs only. Detailed genome analysis revealed GH44 members in fungal species from the phylum Basidiomycota, but not in other fungi, which we hypothesized to also be XEGs. Two GH44 enzymes from *Dichomitus squalens* and *Pleurotus ostreatus* were heterologously produced and characterized. They exhibited XEG activity and displayed a hydrolytic cleavage pattern different from that observed in fungal XEGs from other GH families. Specifically, the fungal GH44 XEGs were not hindered by substitution of neighboring glucosyl units and generated various "XXXG-type," "GXXX(G)-type," and "XXX-type" oligosaccharides. Overall, these fungal GH44 XEGs represent a novel class of enzymes for plant biomass conversion and valorization.

INTRODUCTION

Xyloglucan is present in all vascular plants, in particular within the primary cell walls, and is a major component of dicots (up to 25% of total dry weight) and non-commelinid monocots including conifers (10%) (Scheller and Ulvskov, 2010). The rigidity of plant cell walls is largely driven by loosening and tightening of cellulose microfibrils, in which xyloglucan plays a key role owing to its binding ability to cellulose by surface adsorption and chain intercalation (Park and Cosgrove, 2015; Pauly et al., 2001). In addition, xyloglucan is found as a major seed-storage polysaccharide in a number of terrestrial plants including tamarind and nasturtium (Kozioł et al., 2015). Enzymatic hydrolysis of xyloglucan not only facilitates the access of cellulases to cellulose but also contributes to complete saccharification of agro-industrial side streams to drive the production of biofuels and biochemicals (Benkő et al., 2008; Kaida et al., 2009). Furthermore, soluble xyloglucan and its hydrolysis products can serve as relevant food and feed additives, for example, to reduce calorie content, to increase the texture of food products, or due to their prebiotic behavior (Chen et al., 2020; Dutta et al., 2020; Mishra and Malhotra, 2009).

Xyloglucan has a substituted β-(1→4)-D-glucan backbone, of which the exact type and degree of substitution varies for different plant sources and tissues (Scheller and Ulvskov, 2010). Still, most xyloglucan structures have been shown as the "XXGG-type" and "XXXG-type" (Vincken et al., 1997), in which G stands for a glucosyl residue and X for a glucosyl residue substituted with an α-(1→6)-D-xylosyl unit (Fry et al., 1993). In addition, xylosyl residues can be further substituted by β-(1→2)-D-galactosyl units (coded L) or, to a lesser extent, with α-(1→2)-L-arabinosyl units (coded S). The galactosyl residues can be fucosylated at the *O*-2 position (coded F). Other less common substitutions are denoted with other designations described elsewhere in detail (Fry et al., 1993; Kozioł et al., 2015).

To date, characterized xyloglucanases (XEGs) have been found in glycoside hydrolase (GH) families GH5, GH9, GH12, GH16, GH44, and GH74 in the Carbohydrate-Active enZYme (CAZy) database (http://www.cazy.org) (Lombard et al., 2014). Fungal XEGs are, however, only found in families GH5 (Matsuzawa et al., 2020), GH12, and GH74 (Attia and Brumer, 2016). Fungal GH12 XEGs have been shown to catalyze the hydrolysis of the xyloglucan backbone in an endo-mode, cleaving the glycosidic linkages at the

[1]Laboratory of Food Chemistry, Wageningen University and Research, Bornse Weilanden 9, 6708 WG Wageningen, the Netherlands

[2]Fungal Physiology, Westerdijk Fungal Biodiversity Institute and Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands

[3]DTU Bioengineering, Technical University of Denmark, Søltofts Plads, 2800 Kongens Lyngby, Denmark

[4]Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[5]Department of Microbiology, University of Helsinki, Viikinkaari 9, 00790 Helsinki, Finland

[6]These authors contributed equally

[7]Lead contact

*Correspondence: miia.r.makela@helsinki.fi

https://doi.org/10.1016/j.isci.2021.103666

"reducing end" side of unsubstituted glucosyl units, releasing XXXG-type oligosaccharides (Damásio et al., 2012; Master et al., 2008; Vitcosque et al., 2016). Fungal GH5 and GH74 XEGs have also been shown to display an endo-type of hydrolysis of the xyloglucan backbone, cleaving at the reducing end side of unsubstituted glucosyl units like the GH12 XEGs (Master et al., 2008; Matsuzawa et al., 2020; Vitcosque et al., 2016). However, in contrast to the GH12 XEGs, the GH5 and GH74 XEGs display a processive mode of action (Arnal et al., 2019; Berezina et al., 2017; Desmet et al., 2007; Martinez-Fleites et al., 2006; Matsuzawa et al., 2020; Yaoi and Mitsuishi, 2002). As a result, from the early stage of hydrolysis, XXXG- and XXLG-building blocks are released by GH74 XEGs, which can even be further degraded into XX, XG, and LG (Arnal et al., 2019; Attia et al., 2016; Desmet et al., 2007; Yaoi and Mitsuishi, 2002). GH5 XEG generated XXXG-type oligosaccharides, in addition to several non-XXXG-type oligosaccharides such as X, XG, XX, LG, XL/LX, XXG, and XXX (Matsuzawa et al., 2020). Detailed genome analysis revealed the presence of basidiomycete proteins in enzyme family GH44 of the CAZy database, for which only bacterial endoglucanases and XEGs have been reported so far (Ariza et al., 2011). Interestingly, no members from other fungal phyla were present in GH44. In this study, we evaluated whether the basidiomycete members of GH44 are also XEGs and compared their product profile with that of XEGs of families GH12 and GH74.

RESULTS AND DISCUSSION

Discovery of new fungal GH44 enzymes through genome mining

A large set of fungal genomes (>400) were analyzed for their diversity with respect to the set of genes assigned to CAZy families in the CAZy database, including both genomes present in the public CAZy database and genomes only available at JGI MycoCosm (Grigoriev et al., 2014). This revealed that part of the basidiomycete genomes contained genes that were already assigned to GH44, a family for which, so far, only bacterial enzymes have been characterized (Ariza et al., 2011). Interestingly, our genome analyses also revealed that fungal GH44 sequences could only be identified in the phylum Basidiomycota. GH44 sequences were found in only about 28% of the basidiomycete genomes according to JGI MycoCosm (situation January 22, 2021) (Table S1). This narrow distribution among basidiomycetes suggests that the GH44 enzymes are ancient, possibly of bacterial origin, and that they were lost during evolution for most of the fungal species, including ascomycetes. Alternatively, a basidiomycete ancestor may have acquired this enzyme by horizontal gene transfer from bacteria after the split from ascomycetes.

Labeling the phylogenetic tree of GH44 candidates with the fungal lifestyles (e.g., white rot, gray rot, litter decomposer, mycorrhizae, coprophilous, and endophyte) demonstrated that GH44 enzymes do not appear associated with a particular lifestyle (Figure 1). Some species contain multiple GH44-encoding genes resulting from recent gene duplications, suggesting a more important role for these enzymes in their habitat. One of these fungi is the mycorrhizal fungus *Sebacina vermifera* whose genome encodes five GH44 candidates. It is tempting to speculate that the putative GH44 enzymes may be involved in partial degradation of plant root cell walls that are enriched in xyloglucan, to enable the establishment of symbiosis between fungus and plant roots. Xyloglucan is also released from plant roots (Galloway et al., 2018) and could therefore be a carbon source for fungi that live in close association with roots (e.g., mycorrhizae) or inhabit soil, such as litter-decomposing basidiomycetes, explaining the presence of GH44 enzymes in these fungi (Figure 1).

Furthermore, our analysis of the evolutionary relationships between the GH44 candidates showed that those from the two wood degrading white rot fungi, *Dichomitus squalens* XegA and *Pleurotus ostreatus* XegA, were present in distant clades in the phylogenetic tree (Figure 1). *In situ* expression and production of the *D. squalens* and *P. ostreatus* GH44 genes and proteins have been shown in previously published transcriptome and exoproteome datasets (Daly et al., 2018; Wu et al., 2020, 2021). In short, *D. squalens* CBS464.89 expressed *xegA* and produced the corresponding protein during growth on spruce and birch wood (Daly et al., 2018, 2020), whereas *P. ostreatus* PC9 expressed the XegA encoding gene on rice straw (Wu et al., 2021), and in the Δ*pex1* and Δ*gat1* mutants on beechwood sawdust (Wu et al., 2020), again supporting a role for these genes in plant biomass degradation. Therefore, these two candidates were selected for further characterization to demonstrate the function of basidiomycete GH44 enzymes.

The selected basidiomycete GH44 candidates are xyloglucan-specific XEGs

D. squalens xegA and *P. ostreatus xegA* were heterologously expressed in *Pichia pastoris*, and their recombinant enzymes were produced as C-terminal His-tag fusion proteins and purified from the culture supernatants using immobilized metal affinity chromatography via hexahistidine-tag. The apparent mass of

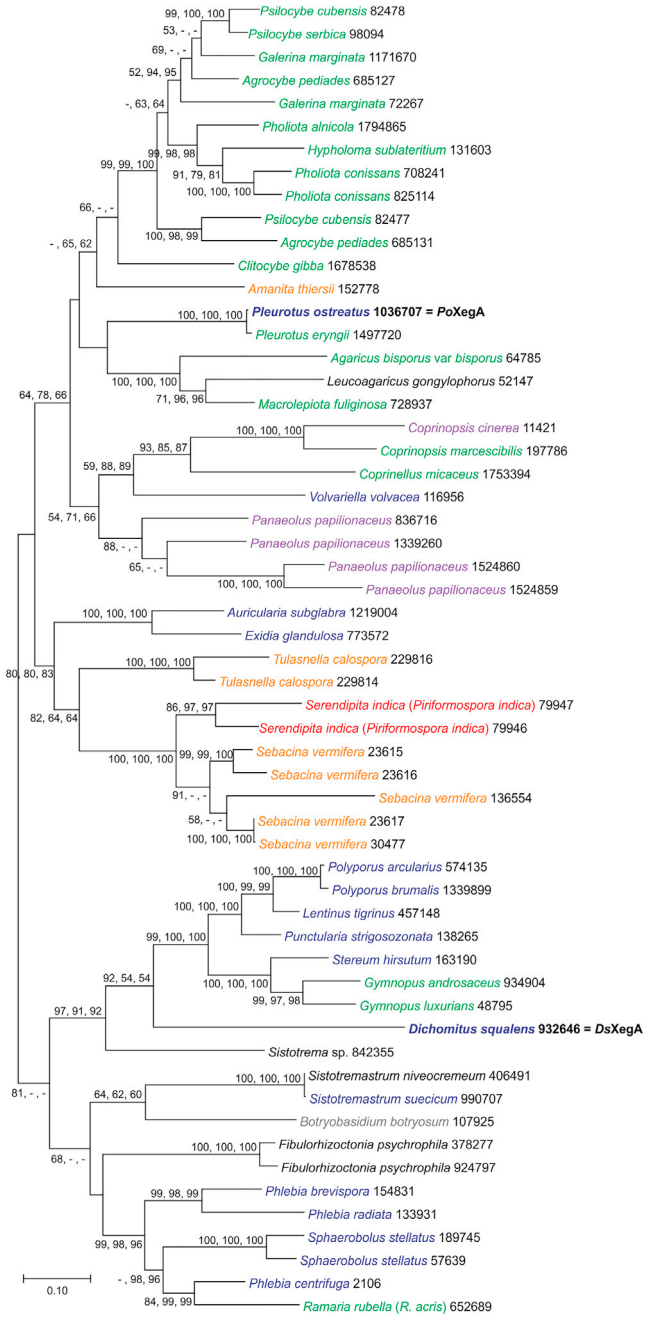


Figure 1. Phylogenetic analysis of GH44 candidates from selected basidiomycete fungi

The figure is a representative Maximum Likelihood tree (500 bootstraps) of a MAFFT alignment of the amino acid sequences of the selected genes. JGI protein IDs are indicated behind the species. Lifestyles of the species are color coded as follows: blue, white rot; gray, gray rot; green, litter decomposer; orange, mycorrhizae; purple, coprophilous; red, endophyte; black, unknown/not clear. Bootstrap values are indicated at the nodes based on Maximum Likelihood, Minimal Evolution, and Neighbor Joining algorithms, respectively. Only bootstrap values that are 50 or higher are displayed. The scale bar corresponds to 0.10 amino acid substitutions per site.

purified enzymes was assessed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and summarized in Table 1. The hydrolytic activity of the purified GH44 enzymes toward four types of polysaccharides, including tamarind xyloglucan (TXG), Avicel PH-101, galactomannan, and beechwood xylan, was validated by high-performance anion-exchange chromatography with pulsed amperometric

Table 1. Physical and biochemical properties of purified XEGs used in this study

GH family	Fungal species	Accession number	Gene name	Enzyme name	Calculated mass (kDa)	Apparent mass (kDa) (−) Endo H	Apparent mass (kDa) (+) Endo H	pH optimum	Temperature optimum (°C)	Positive substrate	Reference
44	*Dichomitus squalens*	JGI|Dicsqu464_1| 932646	*xegA*	*Ds*XegA	78	115	90	5.0	50	TXG	This study
44	*Pleurotus ostreatus*	JGI|PleosPC15_2| 1036707	*xegA*	*Po*XegA	80	90	80	6.0	50	TXG	This study
12	*Aspergillus nidulans*	XP_658056.1, AN0452	*xegA*	*An*XegA	24	25	25	6.5	47	TXG	Bauer et al. (2006)
74	*Aspergillus nidulans*	XP_659146.1, AN1542	*xegB*	*An*XegB	85	100	90	3.0	42	TXG, TXG oligosaccharides	Bauer et al. (2005); Bauer et al. (2006)

TXG, tamarind xyloglucan.

detection (HPAEC-PAD). Both enzymes were only active toward xyloglucan and not on other tested substrates, which suggested that these enzymes were specific XEGs (further referred to as *Ds*XegA and *Po*XegA). The xyloglucan-specific activity of these fungal GH44 XEGs was also observed for fungal XEGs from GH5, GH12, and GH74 (Bauer et al., 2006; Grishutin et al., 2004; Master et al., 2008; Matsuzawa et al., 2020; Sinitsyna et al., 2010; Song et al., 2013; Xian et al., 2016; Yaoi and Mitsuishi, 2002). In contrast, bacterial GH44 XEGs have been reported to exhibit a broad substrate specificity and, in addition to xyloglucan, other polysaccharides can be hydrolyzed, including barley β-glucan, birchwood xylan, wheat arabinoxylan, carob galactomannan, and synthetic soluble cellulose derivatives (Ariza et al., 2011; Hirano et al., 2013; Najmudin et al., 2006; Warner et al., 2011; Ye et al., 2012).

*Ds*XegA and *Po*XegA showed an optimum temperature at 50°C (Figure 2A), which was slightly lower than that of the reported fungal XEGs of families GH12 and GH74, which typically have optimal activity in the range of 55°C–65°C (Damásio et al., 2012; Hakamada et al., 2011; Sinitsyna et al., 2010; Xian et al., 2016; Yaoi and Mitsuishi, 2002), but higher than the optimum temperature of *A. oryzae* GH5 XEG, i.e., 45°C (Matsuzawa et al., 2020). Both XEGs were stable up to 40°C, of which *Po*XegA showed better stability than *Ds*XegA at lower temperature (20°C). For pH optimum, *Ds*XegA and *Po*XegA showed the highest activity toward TXG at pH 5.0 and 6.0, respectively (Figures 2B and 2C). These values are comparable with those of the reported fungal GH5, GH12, and GH74 XEGs that generally have optimal activity in an acidic pH range (3.5–6.0) (Grishutin et al., 2004; Hakamada et al., 2011; Master et al., 2008; Matsuzawa et al., 2020; Sinitsyna et al., 2010; Song et al., 2013; Yaoi and Mitsuishi, 2002). When examined for pH stability, *Ds*XegA retained more than 75% residual activity after 1 hour of incubation at pH 5.0–7.0, whereas *Po*XegA was stable at the pH range from 5.0 to 8.0 and was more stable than *Ds*XegA between pH 6.0 and 8.0 (Figure 2D).

GH44 XEGs generate unique xyloglucan oligosaccharide profiles

To characterize the xyloglucan cleavage patterns by the two GH44 XEGs, the TXG oligosaccharides released by *Ds*XegA and *Po*XegA were profiled by HPAEC-PAD (Figures 3A and 3B). For comparison and annotation of peaks, HPAEC profiles of TXG digests generated by *An*XegA (GH12), by *An*XegB (GH74), and by the previously well-characterized *Nc*LPMO9C (Sun et al., 2020b) and of the reference TXG oligosaccharides (Sun et al., 2020b) and TXG oligosaccharide standards were added in Figures 3C–3F. *An*XegB has previously been studied by Bauer et al. (2005), and *An*XegA has not been extensively characterized so far (Bauer et al., 2006). In general, the HPAEC profiles of the *Ds*XegA and *Po*XegA digests were similar and only the intensity of several peaks was different. It should be noted that the profiles of the *An*XegA- and the *An*XegB-TXG digests were very different, indicating distinct cleavage patterns of TXG by three types of fungal GH44, GH12, and GH74 XEGs used in this study.

Based on the comparison in elution from the standards and reference digests (Figures 3G and 3F), typical XXXG-type oligosaccharides, such as XXXG, XLXG, XXLG, and XLLG, were present in all GH44-, GH12- and GH74-TXG digests (Figures 3A–3D), although the relative abundance of these compounds differed for each of the digests. In contrast, XXX, XLX, XXL, and XLL were only found in the two GH44 digests (Figures

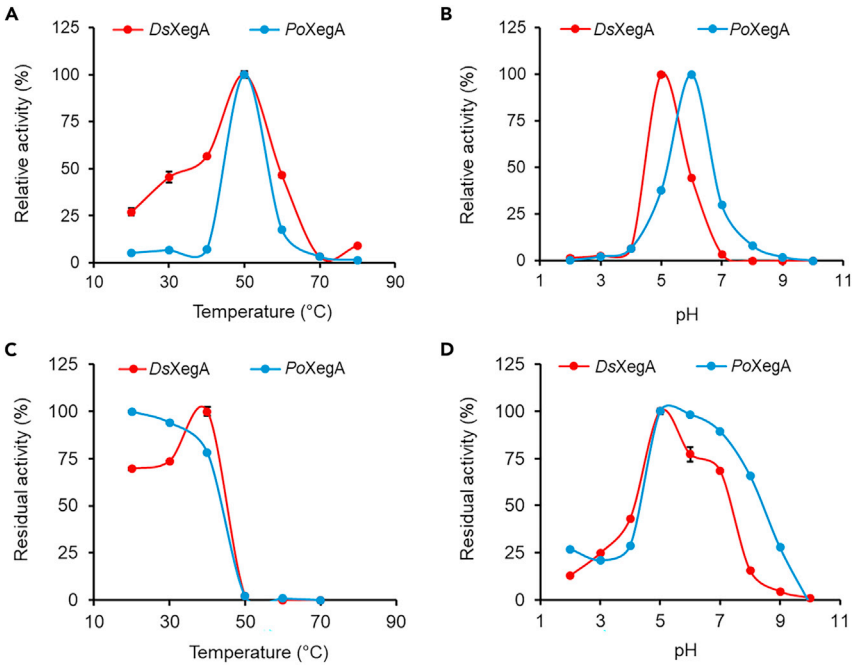


Figure 2. The biochemical properties of *Ds*XegA and *Po*XegA

(A) Temperature optimum was measured at 20°C–80°C in 100 mM phosphate buffer, pH 5.0.

(B) pH optimum was measured at 40°C in 100 mM Britton-Robinson buffer, pH 2.0–10.0.

(C) Temperature stability was measured at 37°C after pretreatment of enzymes in 100 mM phosphate buffer, pH 5.0, at 20°C–70°C for 1 h.

(D) pH stability was measured at 37°C after pretreatment of enzymes in 100 mM Britton-Robinson buffer, pH 2.0–10.0, at 40°C for 1 h. Error bars depict standard deviation of technical duplicate reactions.

3C and 3D). Notably, an unknown peak, eluting between XXXG and XLXG, was identified in both GH44-TXG digests. We propose that this unknown oligomer corresponds to GXXX, based on elution behavior and hydrophilic interaction chromatography coupled with electrospray ionization-collision induced dissociation-mass spectrometry (HILIC-ESI-CID-MS/MS^2) analysis described in the next section. Further unique for the GH44-TXG digests was an additional set of yet unknown compounds, eluting after XLLG. Again, taking into account their elution together with the HILIC-ESI-CID-MS/MS^2 data described in the next section, it could be speculated that these unknown compounds corresponded to GXXXG, GXLX(G), GXXL(G), and GXLL(G).

Furthermore, in the *An*XegB-TXG digest, XG and XX were identified based on comparison with the reference TXG oligosaccharides (Figure 3), suggesting that *An*XegB was able to cleave XXXG into XX + XG. By analogy with this observation, it could be concluded that XXLG was cleaved into XX and LG, substantiated by the relative low abundance of XXLG. The compound eluting at around 14.5 min was therefore annotated as LG.

Structural elucidation of xyloglucan oligosaccharide released by GH44 XEGs

To further elucidate the structures of xyloglucan oligosaccharides generated by the two GH44 XEGs, *Ds*XegA- and *Po*XegA-TXG digests were analyzed by negative ion mode HILIC-ESI-CID-MS/MS^2. The TXG oligosaccharide standards were also analyzed (Figure S1) for the identification and annotation of (unknown) structures. Similar to the HPAEC results, *Ds*XegA- and *Po*XegA-TXG digests showed comparable profiles in their base-peak ion chromatograms and only the relative abundance of each product varied (Figure S1). Owing to separation of α/β-anomers of TXG oligosaccharides in HILIC, each TXG oligosaccharide split into two distinct peaks. Nevertheless, this α/β-anomer separation did not influence the MS/MS^2 elucidation of the TXG oligosaccharide structures. MS^2 fragments of TXG oligosaccharides were annotated following the principle that C-/Z-type, D-type (derived from a double C-/Z-type cleavage of three linked sugar residues), and A-type fragments were generated predominately in negative

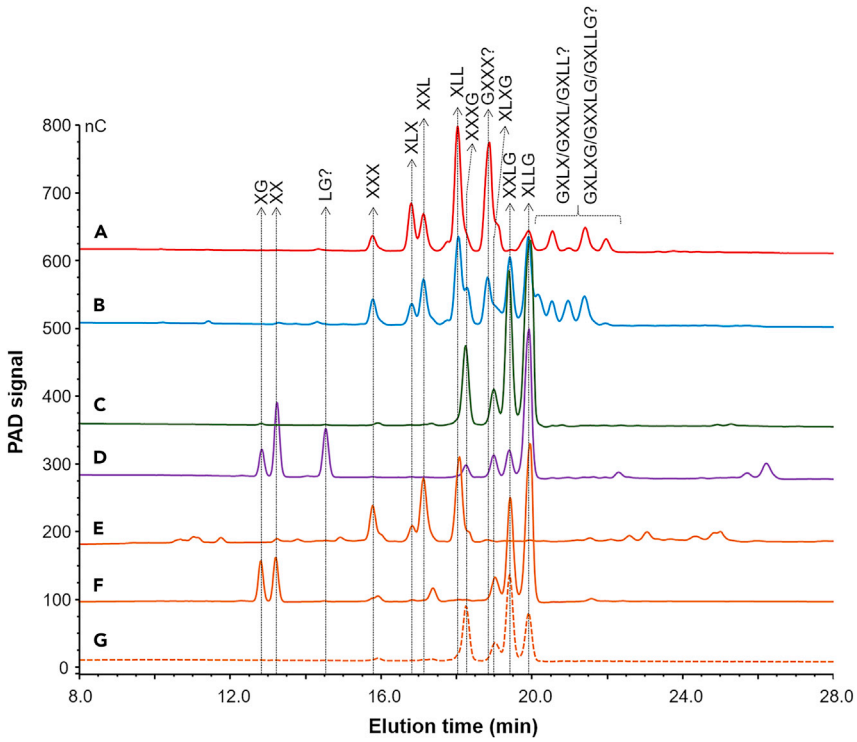


Figure 3. HPAEC elution patterns of tamarind xyloglucan (TXG) oligosaccharides released by end-point (24 h) incubation

TXG oligosaccharides generated by (A) *Ds*XegA, (B) *Po*XegA, (C) *An*XegA, (D) *An*XegB, and (E) *Nc*LPMO9C (in the presence of 1 mM ascorbic acid). (F) Reference TXG oligosaccharides and (G) TXG oligosaccharide standards are also shown for annotation of peaks.

collision-induced dissociation (CID) ion mode (Domon and Costello, 1988; Quéméner et al., 2015; Sun et al., 2020a, 2020b).

In both GH44-TXG digests, regular XXXG-type oligosaccharides were analyzed, including XXXG (*m/z* 1,061.6), XLXG (*m/z* 1,223.6), XXLG (*m/z* 1,223.6), and XLLG (*m/z* 1,385.6). Next, "XXX-type" oligosaccharides (XXX, *m/z* 899.5; XLX and XXL, *m/z* 1,061.6; and XLL *m/z* 1,223.6) and the above suggested "GXXX-type" oligosaccharides (GXXX, *m/z* 1,061.6, GXLX, *m/z* 1,223.6, GXXL, *m/z* 1,223.6, and GXLL, *m/z* 1,385.6) were unambiguously characterized (Figure S1). In addition, several "GXXXG-type" oligosaccharides were found in the two GH44-TXG digests, but to a lesser extent. As an example, Figures 4A and 4B show the annotation of the MS^2 fragmentation of XXL and GXXX, respectively. Identification of these two structures was based on the elution time and considering the TXG structure composed of XXXG-type blocks. As shown in Figure S1, *m/z* 1,061.6 can represent multiple isomeric structures, not only XXL and GXXX but also XXXG and XLX. Based on the elution time of TXG oligosaccharide standards (31.5–32.1 min), XXXG was first identified (peak number 2 in Figure S1) in *Ds*XegA- and *Po*XegA-TXG digests, which was also confirmed by the MS^2 fragments. However, extra MS^2 fragments *m/z* 689, 707, and 767 were present in peak number 2 (Figure S1), which are slightly different from the MS^2 fragments of the XXXG standard and indicated the co-elution of other isomers with XXXG. These three fragments represent a C-type glycosidic bond and ^{0,2}A-type cross-ring cleavage to generate a GXX, an XL, or an LX fragment (three hexaosyl and two pentaosyl residues) from the non-reducing end. Therefore, these two structures can only be released from XLX or GXXX, while GXXX was identified in the following peak number 4 as described later. To summarize this information, the other isomer in peak number 2 (Figure S1) was XLX. Next, we investigated the peak numbers 3 and 4 also showing oligosaccharides with *m/z* 1,061.6. Interestingly, the oligosaccharide in peak number 3 generated comparable fragments as XXXG standards, but with a more abundant *m/z* 353 and a lack of ions *m/z* 899, 839, and 821 (Figure 4A). The abundant former fragment indicated the presence of xylosyl-galactosyl residues, and the lack of latter ions often represented the terminal unsubstituted glucosyl units at the reducing end side (Quéméner et al., 2015; Sun et al., 2020b).

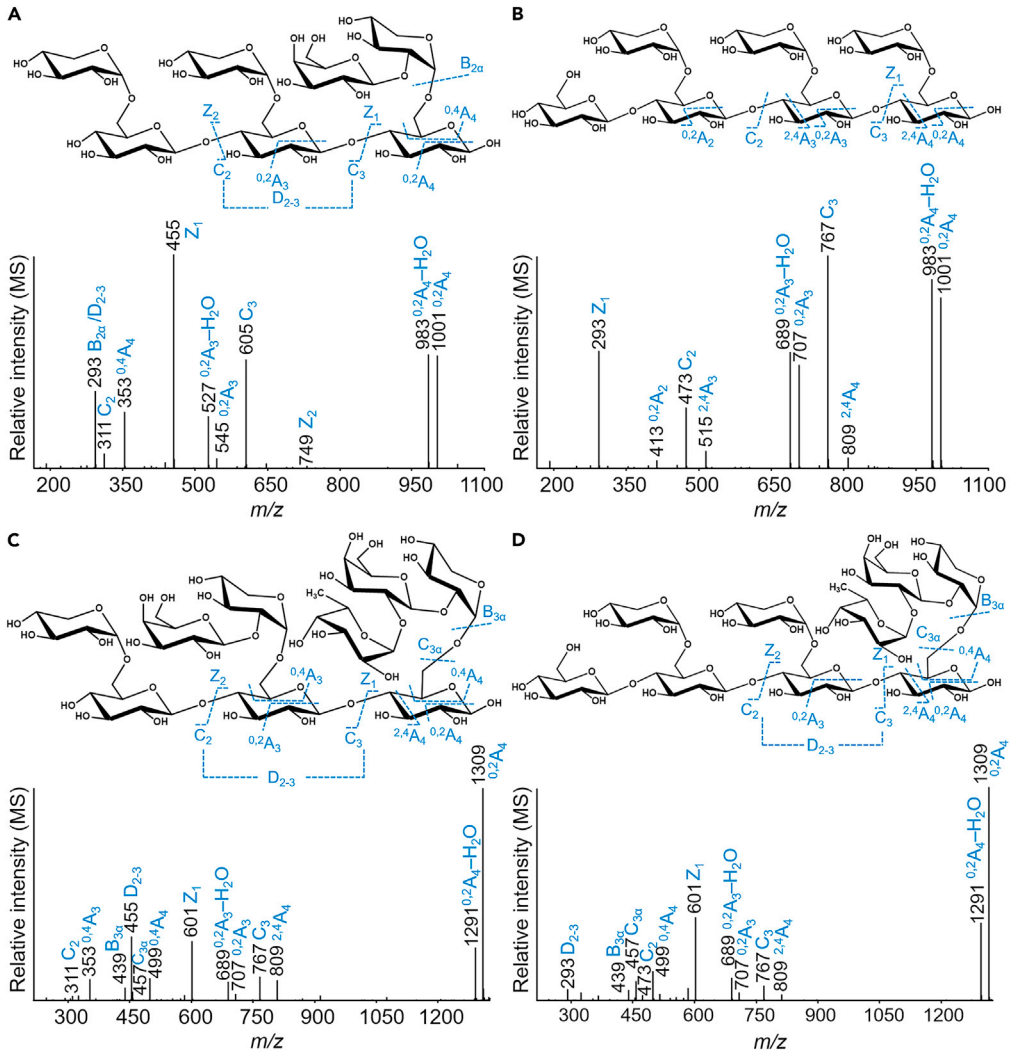


Figure 4. Negative ion mode HILIC-ESI-CID-MS/MS^2 fragmentation patterns of diagnostic TXG oligosaccharides

(A) Diagnostic XXL (*m/z* 1,061.6), (B) GXXX (*m/z* 1,061.6), (C) XLF (*m/z* 1,369.7), and (D) GXXF (*m/z* 1,369.7) both from the end-point *Ds*XegA- and *Po*XegA-TXG digests. The fragments are annotated according to the nomenclature proposed by Domon and Costello (1988).

Taking this information into account, peak number 3 is annotated as an XXL structure. Regarding peak number 4 (XLX or GXXX), we detected the diagnostic fragment *m/z* 413, which resulted from a ^{0,2}A-type cross-ring cleavage on a glucosyl residue from either a GX or an L unit at the non-reducing end. As a non-reducing end terminal L unit with parent *m/z* 1,061.6 cannot be present in XLX and GXXX (not even in TXG structure), this diagnostic fragment *m/z* 413 can only represent a terminal GX structure at the non-reducing end, and therefore, peak number 4 was annotated as GXXX (Figure 4B).

This distinct presence of identified XXXG-type, XXX-type, and "GXXX(G)-type" oligosaccharides suggests that the two GH44 XEGs cleaved at both sides of unsubstituted glucosyl units in TXG. To further investigate whether the two GH44 XEGs also cleaved next to unsubstituted glucosyl units neighbored by more heavily branched F units, we searched for the *m/z* of 1,369.6–1,369.8 in the black currant xyloglucan (BCXG) digests. The *m/z* 1,369 represents oligosaccharides having five hexaosyl, three pentaosyl, and one fucosyl residues, which can be XXFG, XFGX, FGXX, GXXF, and XLF. XXFG was first identified based on the diagnostic fragments C_3 (*m/z* 605) and $^{0,2}A_3$ (–H_2O) (*m/z* 1,147 and 1,129), which represented an XX fragment at the non-reducing end side (Figure S2). The non-reducing end side XX fragment can only result from an XXFG unit. Other fragments such as C_4 (*m/z* 1,207), $^{0,2}A_4$ (–H_2O) (*m/z* 1,147 and 1,129), $D_{2\text{-}4}$ (*m/z* 601),

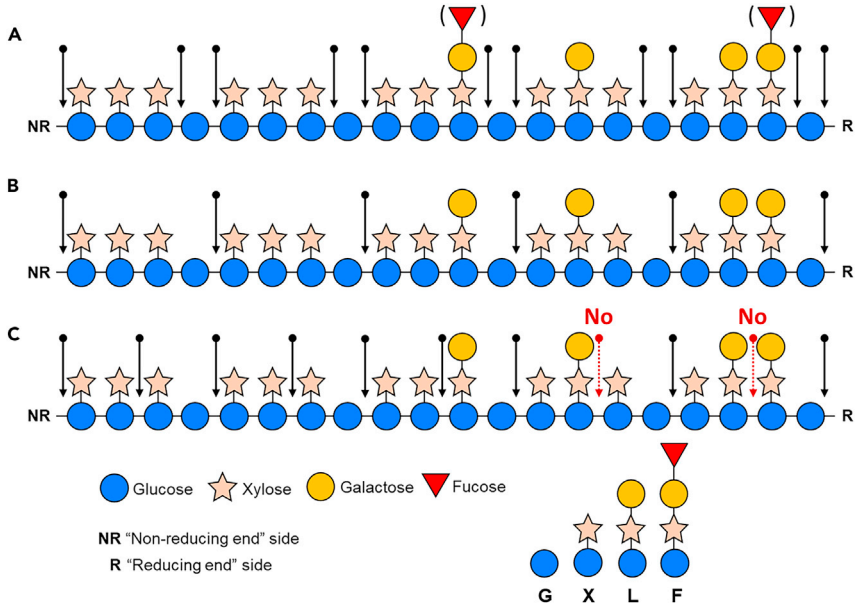


Figure 5. Schematic representation of xyloglucan cleavage patterns by two GH44s (*Ds*XegA and *Po*XegA), one GH12 (*An*XegA), and one GH74 (*An*XegB)

(A) GH44 XEGs cleaved xyloglucan at both sides of unsubstituted glucosyl units to release "XXXG-type," "GXXX(G)-type," and "XXX-type" blocks. The cleavage by GH44 XEGs was not hindered by the substitution of neighboring units.

(B) GH12 XEG cleaved in xyloglucan at the "reducing end" side of unsubstituted glucosyl units to release "XXXG-type" blocks.

(C) GH74 XEG cleaved xyloglucan at the reducing end side of unsubstituted glucosyl units to release XXXG-type blocks and also cleaved within XXXG and XXLG building blocks to release XX + XG and XX + LG. However, GH74 XEG did not cleave at the reducing end side of L units in XLXG and XLLG.

$^{0,4}A_4$ (*m/z* 499), and $C_{3\alpha}$ (*m/z* 457) further confirmed the XXFG structure. In addition to XXFG, XXX-type XLF and GXXX-type GXXF units were released (Figures 4C and 4D). In these MS^2 spectra (Figures 4C and 4D), fragments $^{2,4}A_4$ (*m/z* 809), C3 (*m/z* 767), and $^{0,2}A_3$ ($-H_2O$) (*m/z* 707 and 689) represented fragments composed of three hexaosyl and two pentaosyl residues at the non-reducing end side, which can only be XL and GXX structures in BCXG, and thus XLF and GXXF are the only two possibilities. As shown in the MS^2 spectrum in Figure 4C, the diagnostic ion *m/z* 455 can only result from an internal L residue (regarding XLF and GXXF), and thus XLF was identified as such, whereas Figure 4D represented a GXXF unit.

For comparison, *An*XegA- and *An*XegB-TXG digests were also analyzed by negative ion mode HILIC-ESI-CID-MS/MS^2. In brief, *An*XegA (GH12) only released XXXG-type building blocks (Figure S3). *An*XegB (GH74) not only generated XXXG-type building blocks but also released XG (*m/z* 473.4), XX (*m/z* 605.4), and LG (*m/z* 635.4) units explained by the cleavage of XXXG and XXLG (Figure S4).

Distinct xyloglucan cleavage patterns of GH44 XEGs

Based on all identified products in the GH44-TXG and -BCXG digests (series of XXXG-type, XXX-type, and "GXXX[G]-type" oligosaccharides), it was concluded that the GH44 XEGs cleave xyloglucan at both sides of unsubstituted glucosyl units and that they are not hindered by substitution of neighboring units (Figure 5A). Different from the GH44 XEGs, the GH12 XEG used in this study (*An*XegA) only cleaved within xyloglucan at the reducing end side of unsubstituted glucosyl units (Figure 5B), which is comparable with other fungal GH12 XEGs (Damásio et al., 2012; Master et al., 2008; Vitcosque et al., 2016). The GH74 enzyme (*An*XegB) cleaved xyloglucan similar to the GH12 enzyme, but with additional cleavages within XXXG and XXLG building blocks to release XX, XG, and LG. However, GH74 did not cleave at the reducing end side of L units in XLXG and XLLG (Figure 5C). Similar xyloglucan cleavage patterns have also been found for other fungal GH74 XEGs (Desmet et al., 2007; Yaoi and Mitsuishi, 2002). Slightly different from GH74 XEGs, although not studied in this work, the one characterized fungal GH5 XEG can also cleave next to L units to release even G,

X, XG, LG, XL/LX, XXG, XXX, and LLG units suggested by authors (Matsuzawa et al., 2020). So far, no fungal GH44 XEG has been reported. For a limited number of bacterial GH44, XEG degrading activity has been reported, but the extent of product profile identification differs and does not match the detail reached in our research. For example, for one bacterial GH44 XEG (*Pp*XG44), it has been shown that the compound XXXGXXXG could be hydrolyzed into XXX and GXXXG (Ariza et al., 2011). Similarly, for the bacterial Cel44O, identified as an endocellulase, xyloglucan degrading activity and cleavage between XXX and GXXX have been suggested (Ravachol et al., 2016).

In conclusion, detailed genome analysis led to the discovery of fungal XEGs in family GH44. Our study provided a comprehensive characterization of two of these fungal XEGs and compared their xyloglucan cleavage patterns with that of fungal XEGs from other GH families. These basidiomycete XEGs are xyloglucan-specific but exhibited distinct xyloglucan cleavage patterns compared with fungal XEGs from families GH12 and GH74. Our study also gave insights on phylogenetic diversity of fungal GH44 and understanding of their roles in biomass (xyloglucan) conversion and valorization.

Limitations of the study

We think that the data of our study demonstrated clearly that the two characterized enzymes are xyloglucan-specific with a distinctive product profile compared with other XEG-containing CAZy families. The two characterized enzymes are distantly similar within the phylogenetic tree of fungal GH44 protein sequences, but we cannot exclude that there are other fungal members of GH44 that differ in their substrate specificity or product profile. Characterization of additional fungal GH44 enzymes will shed more light on this.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Microbial strains and growth conditions
- METHOD DETAILS
  - Bioinformatics
  - Cloning, protein production and purification of the selected candidates
  - Enzyme substrate screening
  - Temperature and pH profiles
  - Enzymatic digestion of xyloglucan substrates
  - HPAEC-PAD analysis for substrate screening and for xyloglucan oligosaccharide profiling
  - HILIC-ESI-CID-MS/MS^2 for structural elucidation of xyloglucan oligosaccharides
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.103666.

ACKNOWLEDGMENTS

Alisa Shubina, BSc, is acknowledged for her skillful technical assistance. This work was supported by the China Scholarship Council (grant no. 201803250066 to X.L.) and the Academy of Finland (grant no. 308284 to M.R.M.).

AUTHOR CONTRIBUTIONS

Investigation, P.S., X.L., A.D., and M.R.M.; data curation, P.S., A.D., and B.H.; formal analysis, P.S., X.L., A.D., and B.H.; methodology, P.S., A.D., R.P.d.V., M.A.K., and M.R.M.; software, P.S.; validation, X.L. and A.D.; visualization, X.L.; supervision, A.D., R.P.d,V., and M.A.K.; writing - original draft, P.S., X.L., A.D., M.A.K., and M.R.M.; writing - review & editing, A.D., B.H., R.P.d.V., M.A.K., and M.R.M., conceptualization, R.P.d.V., M.A.K., and M.R.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 7, 2021
Revised: November 23, 2021
Accepted: December 16, 2021
Published: January 21, 2022

REFERENCES

Ariza, A., Eklöf, J.M., Spadiut, O., Offen, W.A., Roberts, S.M., Besenmatter, W., Friis, E.P., Skjøt, M., Wilson, K.S., Brumer, H., and Davies, G. (2011). Structure and activity of *Paenibacillus polymyxa* xyloglucanase from glycoside hydrolase family 44. J. Biol. Chem. *286*, 33890–33900. https://doi.org/10.1074/jbc.M111.262345.

Arnal, G., Stogios, P.J., Asohan, J., Attia, M.A., Skarina, T., Viborg, A.H., Henrissat, B., Savchenko, A., and Brumer, H. (2019). Substrate specificity, regiospecificity, and processivity in glycoside hydrolase family 74. J. Biol. Chem. *294*, 13233–13247. https://doi.org/10.1074/jbc.RA119.009861.

Attia, M., Stepper, J., Davies, G.J., and Brumer, H. (2016). Functional and structural characterization of a potent GH 74 endo-xyloglucanase from the soil saprophyte *Cellvibrio japonicus* unravels the first step of xyloglucan degradation. FEBS J. *283*, 1701–1719. https://doi.org/10.1111/febs.13696.

Attia, M.A., and Brumer, H. (2016). Recent structural insights into the enzymology of the ubiquitous plant cell wall glycan xyloglucan. Curr. Opin. Struct. Biol. *40*, 43–53. https://doi.org/10.1016/j.sbi.2016.07.005.

Bauer, S., Vasu, P., Mort, A.J., and Somerville, C.R. (2005). Cloning, expression, and characterization of an oligoxyloglucan reducing end-specific xyloglucanobiohydrolase from *Aspergillus nidulans*. Carbohydr. Res. *340*, 2590–2597. https://doi.org/10.1016/j.carres.2005.09.014.

Bauer, S., Vasu, P., Persson, S., Mort, A.J., and Somerville, C.R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc. Natl. Acad. Sci. U S A *103*, 11417–11422. https://doi.org/10.1073/pnas.0604632103.

Benkő, Z., Siika-aho, M., Viikari, L., and Réczey, K. (2008). Evaluation of the role of xyloglucanase in the enzymatic hydrolysis of lignocellulosic substrates. Enzyme Microb. Technol. *43*, 109–114. https://doi.org/10.1016/j.enzmictec.2008.03.005.

Berezina, O.V., Herlet, J., Rykov, S.V., Kornberger, P., Zavyalov, A., Kozlov, D., Sakhibgaraeva, L., Krestyanova, I., Schwarz, W.H., Zverlov, V.V., et al. (2017). Thermostable multifunctional GH74 xyloglucanase from *Myceliophthora thermophila*: high-level expression in *Pichia pastoris* and characterization of the recombinant protein. Appl. Microbiol. Biotechnol. *101*, 5653–5666. https://doi.org/10.1007/s00253-017-8297-2.

Britton, H.T.S., and Robinson, R.A. (1931). CXCVIII.—universal buffer solutions and the dissociation constant of veronal. J. Chem. Soc. 1456–1462. https://doi.org/10.1039/JR9310001456.

Chen, H., Jiang, X., Li, S., Qin, W., Huang, Z., Luo, Y., Li, H., Wu, D., Zhang, Q., and Zhao, Y. (2020). Possible beneficial effects of xyloglucan from its degradation by gut microbiota. Trends Food Sci. Technol. *97*, 65–75. https://doi.org/10.1016/j.tifs.2020.01.001.

Coconi Linares, N., Dilokpimol, A., Stålbrand, H., Mäkelä, M.R., and de Vries, R.P. (2020). Recombinant production and characterization of six novel GH27 and GH36 α-galactosidases from *Penicillium subrubescens* and their synergism with a commercial mannanase during the hydrolysis of lignocellulosic biomass. Bioresour. Technol. *295*, 122258. https://doi.org/10.1016/j.biortech.2019.122258.

Daly, P., Casado López, S., Peng, M., Lancefield, C.S., Purvine, S.O., Kim, Y.M., Zink, E.M., Dohnalkova, A., Singan, V.R., Lipzen, A., et al. (2018). *Dichomitus squalens* partially tailors its molecular responses to the composition of solid wood. Environ. Microbiol. *20*, 4141–4156. https://doi.org/10.1111/1462-2920.14416.

Daly, P., Peng, M., Casado López, S., Lipzen, A., Ng, V., Singan, V.R., Wang, M., Grigoriev, I.V., de Vries, R.P., and Mäkelä, M.R. (2020). Mixtures of aromatic compounds induce ligninolytic gene expression in the wood-rotting fungus *Dichomitus squalens*. J. Biotechnol. *308*, 35–39. https://doi.org/10.1016/j.jbiotec.2019.11.014.

Damásio, A.R.L., Ribeiro, L.F.C., Ribeiro, L.F., Furtado, G.P., Segato, F., Almeida, F.B.R., Crivellari, A.C., Buckeridge, M.S., Souza, T.A.C.B., Murakami, M.T., et al. (2012). Functional characterization and oligomerization of a recombinant xyloglucan-specific endo-β-1,4-glucanase (GH12) from *Aspergillus niveus*. Biochim. Biophys. Acta Proteins Proteom. *1824*, 461–467. https://doi.org/10.1016/j.bbapap.2011.12.005.

Desmet, T., Cantaert, T., Gualfetti, P., Nerinckx, W., Gross, L., Mitchinson, C., and Piens, K. (2007). An investigation of the substrate specificity of the xyloglucanase Cel74A from *Hypocrea jecorina*. FEBS J. *274*, 356–363. https://doi.org/10.1111/j.1742-4658.2006.05582.x.

Domon, B., and Costello, C.E. (1988). A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate J. *5*, 397–409. https://doi.org/10.1007/BF01049915.

Dutta, P., Giri, S., and Giri, T.K. (2020). Xyloglucan as green renewable biopolymer used in drug delivery and tissue engineering. Int. J. Biol. Macromol. *160*, 55–68. https://doi.org/10.1016/j.ijbiomac.2020.05.148.

Fry, S.C., York, W.S., Albersheim, P., Darvill, A., Hayashi, T., Joseleau, J.P., Kato, Y., Lorences, E.P., Maclachlan, G.A., and McNeil, M. (1993). An unambiguous nomenclature for xyloglucan-derived oligosaccharides. Physiol. Plant *89*, 1–3. https://doi.org/10.1111/j.1399-3054.1993.tb01778.x.

Galloway, A.F., Pedersen, M.J., Merry, B., Marcus, S.E., Blacker, J., Benning, L.G., Field, K.J., and Knox, J.P. (2018). Xyloglucan is released by plants and promotes soil particle aggregation. New Phytol. *217*, 1128–1136. https://doi.org/10.1111/nph.14897.

Grigoriev, I.V., Roman, N., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., et al. (2014). MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res. *42*, D699–D704. https://doi.org/10.1093/nar/gkt1183.

Grishutin, S.G., Gusakov, A.V., Markov, A.V., Ustinov, B.B., Semenova, M.V., and Sinitsyn, A.P. (2004). Specific xyloglucanases as a new class of polysaccharide-degrading enzymes. Biochimi. Biophys. Acta Gen. Subj. *1674*, 268–281. https://doi.org/10.1016/j.bbagen.2004.07.001.

Hakamada, Y., Arata, S., and Ohashi, S. (2011). Purification and characterization of a xyloglucan-specific glycosyl hydrolase from *Aspergillus oryzae* RIB40. J. Appl. Glycosci. *58*, 47–51. https://doi.org/10.5458/jag.jag.JAG-2010_010.

Hilz, H., de Jong, L.E., Kabel, M.A., Schols, H.A., and Voragen, A.G. (2006). A comparison of liquid chromatography, capillary electrophoresis, and mass spectrometry methods to determine xyloglucan structures in black currants. J. Chromatogr. *1133*, 275–286. https://doi.org/10.1016/j.chroma.2006.08.024.

Hirano, N., Hasegawa, H., Nihei, S., and Haruki, M. (2013). Cell-free protein synthesis and substrate specificity of full-length endoglucanase CelJ (Cel9D-Cel44A), the largest multi-enzyme subunit of the *Clostridium thermocellum* cellulosome. FEMS Microbiol. Lett. *344*, 25–30. https://doi.org/10.1111/1574-6968.12149.

Kaida, R., Kaku, T., Baba, K.i., Oyadomari, M., Watanabe, T., Nishida, K., Kanaya, T., Shani, Z., Shoseyov, O., and Hayashi, T. (2009). Loosening xyloglucan accelerates the enzymatic degradation of cellulose in wood. Mol. Plant *2*, 904–909. https://doi.org/10.1093/mp/ssp060.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

Kozioł, A., Cybulska, J., Pieczywek, P.M., and Zdunek, A. (2015). Evaluation of structure and assembly of xyloglucan from tamarind seed

(*Tamarindus indica* L.) with atomic force microscopy. Food. Biophys. *10*, 396–402. https://doi.org/10.1007/s11483-015-9395-2.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Li, X., Griffin, K., Langeveld, S., Frommhagen, M., Underlin, E.N., Kabel, M.A., de Vries, R.P., and Dilokpimol, A. (2020). Functional validation of two fungal subfamilies in Carbohydrate Esterase family 1 by biochemical characterization of esterases from uncharacterized branches. Front. Bioeng. Biotechnol. *8*, 694. https://doi.org/10.3389/fbioe.2020.00694.

Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490–D495. https://doi.org/10.1093/nar/gkt1178.

Martinez-Fleites, C., Guerreiro, C.I., Baumann, M.J., Taylor, E.J., Prates, J.A., Ferreira, L.M., Fontes, C.M., Brumer, H., and Davies, G.J. (2006). Crystal structures of *Clostridium thermocellum* xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. J. Biol. Chem. *281*, 24922–24933. https://doi.org/10.1074/jbc.M603583200.

Master, E.R., Zheng, Y., Storms, R., Tsang, A., and Powlowski, J. (2008). A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: recombinant expression, purification and characterization. Biochem. J. *411*, 161–170. https://doi.org/10.1042/Bj20070819.

Matsuzawa, T., Kameyama, A., Nakamichi, Y., and Yaoi, K. (2020). Identification and characterization of two xyloglucan-specific endo-1,4-glucanases in *Aspergillus oryzae*. Appl. Microbiol. Biotechnol. *104*, 8761–8773. https://doi.org/10.1007/s00253-020-10883-7.

Mishra, A., and Malhotra, A.V. (2009). Tamarind xyloglucan: a polysaccharide with versatile application potential. J. Mater. Chem. *19*, 8528–8536. https://doi.org/10.1039/B911150F.

Najmudin, S., Guerreiro, C.I., Carvalho, A.L., Prates, J.A., Correia, M.A., Alves, V.D., Ferreira, L.M., Romão, M.J., Gilbert, H.J., Bolam, D.N., and Fontes, C.M. (2006). Xyloglucan is recognized by carbohydrate-binding modules that interact with β-glucan chains. J. Biol. Chem. *281*, 8815–8828. https://doi.org/10.1074/jbc.M510559200.

Park, Y.B., and Cosgrove, D.J. (2015). Xyloglucan and its interactions with other components of the growing cell wall. Plant Cell Physiol. *56*, 180–194. https://doi.org/10.1093/pcp/pcu204.

Pauly, M., Qin, Q., Greene, H., Albersheim, P., Darvill, A., and York, W.S. (2001). Changes in the structure of xyloglucan during cell elongation. Planta *212*, 842–850. https://doi.org/10.1007/s004250000448.

Quéméner, B., Vigouroux, J., Rathahao, E., Tabet, J.C., Dimitrijevic, A., and Lahaye, M. (2015). Negative electrospray ionization mass spectrometry: a method for sequencing and determining linkage position in oligosaccharides from branched hemicelluloses. J. Mass. Spectrom. *50*, 247–264. https://doi.org/10.1002/jms.3528.

Ravachol, J., de Philip, P., Borne, R., Mansuelle, P., Maté, M.J., Perret, S., and Fierobe, H.-P. (2016). Mechanisms involved in xyloglucan catabolism by the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. Sci. Rep. *6*, 1–17. https://doi.org/10.1038/srep22770.

Scheller, H.V., and Ulvskov, P. (2010). Hemicelluloses. Annu. Rev. Plant Biol. *61*, 263–289. https://doi.org/10.1146/annurev-arplant-042809-112315.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675. https://doi.org/10.1038/nmeth.2089.

Sinitsyna, O., Fedorova, E., Pravilnikov, A., Rozhkova, A., Skomarovsky, A., Matys, V.Y., Bubnova, T., Okunev, O., Vinetsky, Y.P., and Sinitsyn, A. (2010). Isolation and properties of xyloglucanases of *Penicillium* sp. Biochemistry (Mosc.) *75*, 41–49. https://doi.org/10.1134/S0006297910010062.

Song, S., Tang, Y., Yang, S., Yan, Q., Zhou, P., and Jiang, Z. (2013). Characterization of two novel family 12 xyloglucanases from the thermophilic *Rhizomucor miehei*. Appl. Microbiol. Biotechnol. *97*, 10013–10024. https://doi.org/10.1007/s00253-013-4770-8.

Sun, P., Frommhagen, M., Kleine Haar, M., van Erven, G., Bakx, E., van Berkel, W., and Kabel, M. (2020a). Mass spectrometric fragmentation patterns discriminate C1-and C4-oxidised cello-oligosaccharides from their non-oxidised and reduced forms. Carbohydr. Polym. *234*, 115917. https://doi.org/10.1016/j.carbpol.2020.115917.

Sun, P., Laurent, C., Scheiblbrandner, S., Frommhagen, M., Kouzounis, D., Sanders, M.G., van Berkel, W.J.H., Ludwig, R., and Kabel, M.A. (2020b). Configuration of active site segments in lytic polysaccharide monooxygenases steers oxidative xyloglucan degradation. Biotechnol. Biofuels *13*, 95. https://doi.org/10.1186/s13068-020-01731-x.

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. https://doi.org/10.1104/pp.114.1.9.

Vitcosque, G.L., Ribeiro, L.F.C., de Lucas, R.C., da Silva, T.M., Ribeiro, L.F., Damasio, A.R.D., Farinas, C.S., Goncalves, A.Z.L., Segato, F., Buckeridge, M.S., et al. (2016). The functional properties of a xyloglucanase (GH12) of *Aspergillus terreus* expressed in *Aspergillus nidulans* may increase performance of biomass degradation. Appl. Microbiol. Biotechnol. *100*, 9133–9144. https://doi.org/10.1007/s00253-016-7589-2.

Warner, C.D., Go, R.M., García-Salinas, C., Ford, C., and Reilly, P.J. (2011). Kinetic characterization of a glycoside hydrolase family 44 xyloglucanase/endoglucanase from *Ruminococcus flavefaciens* FD-1. Enzyme Microb. Technol. *48*, 27–32. https://doi.org/10.1016/j.enzmictec.2010.08.009.

Wu, H., Nakazawa, T., Takenaka, A., Kodera, R., Morimoto, R., Sakamoto, M., and Honda, Y. (2020). Transcriptional shifts in delignification-defective mutants of the white-rot fungus *Pleurotus ostreatus*. FEBS Lett. *594*, 3182–3199. https://doi.org/10.1002/1873-3468.13890.

Wu, H., Nakazawa, T., Xu, H., Yang, R., Bao, D., Kawauchi, M., Sakamoto, M., and Honda, Y. (2021). Comparative transcriptional analyses of *Pleurotus ostreatus* mutants on beech wood and rice straw shed light on substrate-biased gene regulation. Appl. Microbiol. Biotechnol. *105*, 1175–1190. https://doi.org/10.1007/s00253-020-11087-9.

Xian, L., Wang, F., Yin, X., and Feng, J.-X. (2016). Identification and characterization of an acidic and acid-stable endoxyloglucanase from *Penicillium oxalicum*. Int. J. Biol. Macromol. *86*, 512–518. https://doi.org/10.1016/j.ijbiomac.2016.01.105.

Yaoi, K., and Mitsuishi, Y. (2002). Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. J. Biol. Chem. *277*, 48276–48281. https://doi.org/10.1074/jbc.M208443200.

Ye, L., Su, X., Schmitz, G.E., Moon, Y.H., Zhang, J., Mackie, R.I., and Cann, I.K. (2012). Molecular and biochemical analyses of the GH44 module of CbMan5B/Cel44A, a bifunctional enzyme from the hyperthermophilic bacterium *Caldicellulosiruptor bescii*. Appl. Environ. Microbiol. *78*, 7048–7059. https://doi.org/10.1128/AEM.02009-12.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
6x-His Tag Monoclonal Antibody (clone 3D5), AP	Thermo Fisher Scientific	Cat# R932-25; RRID: AB_2556555
Bacterial and virus strains		
Escherichia coli DH5α	Thermo Fisher Scientific	Cat# EC0112
Chemicals, peptides, and recombinant proteins		
Sodium acetate	Sigma-Aldrich	Cat# S8750
Ammonium acetate	Sigma-Aldrich	Cat# 240192
Britton-Robinson buffer	Britton and Robinson (1931)	N/A
Endoglycosidase H	New England Biolabs	Cat# P0702S
PmeI	New England Biolabs	Cat# R0560S
Zeocin	Thermo Fisher Scientific	Cat# R25001
5-Bromo-4-chloro-3-indolyl phosphate (BCIP)	Sigma-Aldrich	CAS: 6578-06-9
Nitro blue tetrazolium (NBT)	Sigma-Aldrich	CAS: 298-83-9
Recombinant DNA		
pPicZαA-*xegA*^{Ds} (*E. coli* propagation and *P. pastoris* expression)	This paper	N/A
pPicZαA-*xegA*^{Po} (*E. coli* propagation and *P. pastoris* expression)	This paper	N/A
Software and algorithms		
Excel	Microsoft	N/A
Xcalibur 4.3.73.11	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/OPTON-30965#/OPTON-30965
Chromeleon 7.2.10	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/CHROMELEON7#/CHROMELEON7
ChemDraw 18.0.0.231	PerkinElmer	https://perkinelmerinformatics.com/products/research/chemdraw/
ImageJ	Schneider et al. (2012)	https://imagej.nih.gov/ij/
Multiple Alignment using Fast Fourier Transform (MAFFT)	Katoh and Standley (2013)	https://mafft.cbrc.jp/alignment/software/
MEGA 7.0	Kumar et al. (2016)	https://www.megasoftware.net/
Other		
Tamarind xyloglucan (*Tamarind indica*)	Megazyme	Product code: P-XYGLN
Xyloglucan (hepta+octa+nona saccharides)	Megazyme	Product code: O-XGHON
Avicel® PH-101	Sigma-Aldrich	Cat#11365
Black currant xyloglucan (*Ribes nigrum* L.) fraction CASS	Hilz et al. (2006)	N/A
Xyloglucanase (GH5, *Paenibacillus* sp.)-xyloglucan digest	Sun et al. (2020b)	N/A
*Nc*LPMO9C-xyloglucan digest	Sun et al. (2020b)	N/A
Galactomannan	Megazyme	Product code: P-GALML
Beechwood xylan	Carl Roth GmbH + Co. KG	CAS: 9014-63-5
HisTrap FF 1 mL column	GE Healthcare Life Sciences	Lot: 10287614

RESOURCE AVAILABILITY

Lead contact

Further information and request for resources and reagents should be directed to and will be fulfilled by Miia R. Mäkelä (miia.r.makela@helsinki.fi).

Materials availability

Materials generated in this study are available from the lead contact with a completed Materials Transfer Agreement.

Data and code availability

The published article and its supplemental information files include all datasets generated or analyzed during this study. This paper does not report original code.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbial strains and growth conditions

E. coli DH5α competent cells (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA) were employed to propagate pPicZαA-*xegA*^{Ds} and pPicZαA-*xegA*^{Po} plasmids and the transformed cells were grown in low-salt Luria Bertani medium supplemented with 25 μg/mL Zeocin at 37°C. *P. pastoris* strain X-33 (Invitrogen) was transformed with pPicZαA-*xegA*^{Ds} and pPicZαA-*xegA*^{Po} plasmids and the positive transformants were selected on agar plates containing 1% yeast extract, 2% peptone, 2% glucose, 1 M sorbitol, 2% agar, and 100 μg/mL Zeocin. For production of recombinant proteins, the selected *P. pastoris* transformants were grown in Buffered Glycerol Complex medium and induced in Buffered Methanol Complex medium at 30°C (Coconi Linares et al., 2020).

METHOD DETAILS

Bioinformatics

The amino acid sequences of 190 GH44 candidates were retrieved from the JGI MycoCosm portal (https://mycocosm.jgi.doe.gov/mycocosm/home). After removing sequences from unpublished genomes and keeping only one strain per species, the sequences were aligned using Multiple Alignment using Fast Fourier Transform (MAFFT) (Katoh and Standley, 2013). After manually filtering and correcting the sequences for gene model errors, 57 sequences from 42 basidiomycete species were included in phylogenetic analysis (Table S1). Because characterized bacterial GH44 enzymes appeared to be too distant to be used as an outgroup, unrooted Maximum Likelihood, Minimum Evolution (ME) and Neighbor Joining (NJ) trees were computed using 500 bootstraps in MEGA7 (Kumar et al., 2016) to reveal the evolutionary relationship between the fungal GH44 sequences.

Cloning, protein production and purification of the selected candidates

The selected candidate genes (Table 1) without predicted signal peptides and introns were codon optimized and synthesized into pPicZαA plasmid (Genscript Biotech, Leiden, The Netherlands), which were then chemically transformed into *E. coli* DH5α and propagated. The plasmids pPicZαA-*xegA*^{Ds} and pPicZαA-*xegA*^{Po} were extracted and linearized with PmeI (New England Biolabs, Ipswich, MA) and subsequently transformed into *P. pastoris* strain X-33 by electroporation. The positive colonies were selected by colony Western Blot using 6x-His Tag Monoclonal Antibody conjugated to alkaline phosphatase (Invitrogen, Thermo Fisher Scientific, Carlsbad, CA, USA) and detected with a chromogenic substrate 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) (Li et al., 2020).

The selected *P. pastoris* transformants were cultivated for the protein production according to Coconi Linares et al. (2020). After four days of induction, culture supernatants were harvested (8000 × *g*, 4°C, 20 min) and purified using ÄKTA FPLC device (GE Healthcare Life Sciences, Uppsala, Sweden) as described earlier (Coconi Linares et al., 2020). The purified proteins were stored at 4°C prior to further analysis.

Molecular mass and concentration of the purified enzymes were evaluated by SDS-PAGE and ImageJ program, respectively (Li et al., 2020). Deglycosylation was performed using endoglycosidase H (Endo H; New England Biolabs, Ipswich, MA) following the manufacturer's recommendation.

Enzyme substrate screening

Different substrates including TXG, Avicel® PH-101, galactomannan and beechwood xylan were suspended in 50 mM sodium acetate buffer (pH 5.0) to a final concentration of 2 mg/mL. *Ds*XegA and *Po*XegA were added to a final protein concentration of 100 μg/mL. The digestion (reaction volume of 300 μL) was incubated at 40°C with shaking (Multitron shaking incubators, Infors HT, Bottmingen, Switzerland) at

110 rpm for 24 h. Control reactions were performed without the addition of enzymes. All digestions were performed in duplicate. The reactions were stopped by heating at 97°C for 10 min. Afterwards, the digests were centrifuged at 22000 × *g* for 20 min, and the clear supernatants were collected and analyzed by high-performance anion exchange chromatography with HPAEC-PAD after diluting ten times.

Temperature and pH profiles

The temperature and pH optima of GH44 enzymes were determined at 20–80°C in 100 mM phosphate buffer, pH 5.0, or at 40°C in 100 mM Britton-Robinson buffer from pH 2.0 to pH 10.0 (Britton and Robinson, 1931), respectively. The thermal and pH stability were determined by measuring the residual enzyme activity at 37°C after 1 h incubation at 20–70°C in 100 mM phosphate buffer (pH 5.0 for *Ds*XegA and pH 6.0 for *Po*XegA), or at 40°C in 100 mM Britton-Robinson buffer from pH 2.0 to pH 10.0. All digestions were performed in duplicate using TXG as a substrate.

Enzymatic digestion of xyloglucan substrates

Xyloglucan substrates (TXG or BCXG) were dissolved in 50 mM ammonium acetate buffer (pH 5.0) to a final concentration of 2 mg/mL. *Ds*XegA, *Po*XegA, *An*XegA and *An*XegB were dosed in a final protein concentration of 100 μg/mL. The digestion (reaction volume of 300 μL) was incubated in an Eppendorf ThermoMixer® Comfort at 40°C with shaking at 800 rpm for 24 h. Control reactions were performed without the addition of enzymes. All digestions were performed in duplicate. The reactions were stopped by incubating at 97°C for 10 min. Afterwards, the digests were centrifuged at 22000 × *g* for 20 min, and the clear supernatants were collected and stored at −20°C until further usage. Previous well characterized *Nc*LPMO9C- and GH5-TXG digests were used as reference mixtures (Sun et al., 2020b). The TXG oligosaccharides released by bacterial GH5 are further referred as the "reference TXG oligosaccharides". *Ds*XegA-, *Po*XegA-, *An*XegA- and *An*XegB-TXG digests were analyzed by HPAEC-PAD after diluting 20 times, and by HILIC-ESI-CID-MS/MS^2. The TXG oligosaccharide standards were also analyzed by HPAEC-PAD (50 μg/mL) and HILIC-ESI-CID-MS/MS^2 (100 μg/mL). The *Nc*LPMO9C-TXG digest and reference TXG oligosaccharides were analyzed by HPAEC-PAD after diluting five and ten times, respectively.

HPAEC-PAD analysis for substrate screening and for xyloglucan oligosaccharide profiling

For the substrate screening and the xyloglucan oligosaccharide profiling, TXG and other carbohydrates corresponding digests were analyzed by HPAEC-PAD on an ICS5000 (Thermo Scientific, Waltham, MA, USA) system. Instrument settings, column, mobile phases and elution program have been described previously (Sun et al., 2020b).

HILIC-ESI-CID-MS/MS^2 for structural elucidation of xyloglucan oligosaccharides

The TXG and BCXG digests were analyzed by HILIC-ESI-CID-MS/MS^2 on a Vanquish UHPLC system (Thermo Scientific, San Jose, CA, USA) coupled to an LTQ Velos Pro mass spectrometer (Thermo Scientific). The UHPLC settings, column, mobile phases and elution program (Sun et al., 2020b), and the MS (negative ion mode) settings (Sun et al., 2020a) have been described previously. MS^2 was performed under dependent scan mode.

QUANTIFICATION AND STATISTICAL ANALYSIS

All enzymatic digestions were performed in technical duplicates and standard deviations were calculated for temperature and pH profiles.