CAZyme Prediction in Ascomycetous Yeast Genomes Guides Discovery of Novel Xylanolytic Species with Diverse Capacities for Hemicellulose Hydrolysis

Jonas L. Ravn
Chalmers University of Technology: Chalmers tekniska hogskola
Martin K. M. Engqvist
Chalmers University of Technology: Chalmers tekniska hogskola
Johan Larsbrink
Chalmers University of Technology: Chalmers tekniska hogskola
Cecilia Geijer (cecilia.geijer@chalmers.se)
Chalmers tekniska hogskola https://orcid.org/0000-0002-4158-2938

Research

Keywords: Ascomycota, non-conventional yeasts, CAZymes, xylanase, xylan, xylanolytic yeasts

Posted Date: April 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-420485/v1

Abstract

Background Ascomycetous yeasts from the kingdom fungi inhabit every biome in Nature. While filamentous fungi have been studied extensively regarding their enzymatic degradation of the complex polymers comprising lignocellulose, yeasts have been largely overlooked. As yeasts are key organisms used in industry, understanding their enzymatic strategies for biomass conversion is an important factor in developing new and more efficient cell factories. The aim of this study was to identify polysaccharide-degrading yeasts by mining CAZymes in 332 yeast genomes from the phylum Ascomycota. Selected CAZyme-rich yeasts were then characterized in more detail through growth and enzymatic activity assays.

Results The CAZyme analysis revealed a large spread in the number of CAZyme-encoding genes in the Ascomycetous yeast genomes. We identified a total of 224 predicted CAZyme families, including several CAZymes likely involved in degradation of plant polysaccharides. Growth characterization of 40 CAZyme-rich yeasts revealed no cellulolytic yeasts, but several species from the Trichomonascaceae and CUG-Ser1 clades were able to grow on xylan, β-glucan and xyloglucan. Blastobotrys mokoenaii, Sugiyamaella lignohabitans, Spencermartinsiella europaea and several Scheffersomyces species displayed superior growth on xylan and well as high enzymatic activities. These species contained several putative xylanolytic enzymes, including the well-studied xylanase-containing glycoside hydrolase families GH10 and GH30 that appear attached to the cell surface. B. mokoenaii was the only species containing a GH11 xylanase, which was shown to be secreted. Surprisingly, no known xylanases were predicted in the xylanolytic species Wickerhamomyces canadensis, suggesting that this yeast possess novel xylanases. In addition, by examining non-sequenced yeasts closely related to the xylanolytic yeasts, we were able to identify novel species with high xylanolytic capacities. **Conclusions** Our approach of combining high-throughput bioinformatic CAZyme-prediction with growth and enzyme characterization proved to be a powerful pipeline for discovery of novel xylan-degrading yeasts and enzymes. The identified yeasts display diverse profiles in terms of growth, enzymatic activities and xylan substrate preferences, pointing towards different strategies for degradation and utilization of xylan. Together, the results provide novel insights into how yeast degrade xylan, which can be used to improve cell factory design and industrial bioconversion processes.

Full Text

This preprint is available for download as a PDF.

Tables

Table 1. Overview of budding yeast growth assessment on agar plates and liquid cultures using different polysaccharides. Growth was scored by visual comparison to a negative control plate not containing a carbon source and by the difference in colony thickness and size (including hyphae, if present).

Clade	Species	NRRL	Wheat AX	Birch GX	Xyloglucan	β-glucan	GluMan	GalMan	Pectin	Poly-MeGal	Curdlan	Avicel	CMC
Saccharom.	*Saccharomyces cerevisiae* (CENPK)		-	-	-	-	+	+	-	W	-	-	-
Lipomyc.	*Lipomyces arxii, CBS 7333*	*Y-17921*	-	-	-	-	-	-	-	+	-	-	-
Trichomonascaceae	*Spencermartinsiella europaea, CBS 11730*	Y-48265	++	++	+	+	W	-	-	+	-	-	-
	Sugiyamaella lignohabitans, CBS 10342	YB-1473	+	++	-	W	W	-	++	++	-	-	-
	Diddensiella caesifluorescens, CBS 12613	Y-48781	W	W	-	+	-	-	-	W	-	-	-
	Blastobotrys peoriensis, CBS 10340	YB-2290	++	++	-	++	-	-	-	+	-	-	-
	Blastobotrys muscicola, CBS 10338	Y-7993	W	W	-	W	-	W	-	++	-	-	-
	Blastobotrys mokoenaii, CBS 8435	Y-27120	+++	++	+++	W	+	+++	-	+	W	-	-
	Blastobotrys americana, CBS 10337	Y-6844	W	W	-	+	-	W	-	+	W	-	-
	Blastobotrys adeninivorans, CBS 7370	Y-17693	+	+	-	++	-	+	-	+	-	-	-
	Blastobotrys proliferans, CBS 522.75	Y-17577	++	W	W	++	+	++	-	++	W	-	-
	Blastobotrys serpentis, CBS 10541	Y-48249	+	+	+	W	W	W	-	+	-	-	-
	Blastobotrys raffinosifermentans, CBS 6800	Y-27150	+	+	+	+	+	-	-	+	+	-	-
Pichiaceae	*Ambrosiozyma philentoma*, CBS 6276	Y-7523	W	W	-	+	+	-	-	+	W	-	-
	Ambrosiozyma monospora, CBS 6392	Y-7403	W	-	-	+	-	-	W	+	-	-	-
	Ambrosiozyma oregonensis, CBS 5560	Y-6106	W	W	-	++	-	-	W	+	-	-	-
CUG-Ser1	*Spathaspora passalidarum*, CBS 10155	Y-27907	+	++	-	-	W	-	-	++	-	-	-
	Debaryomyces subglobosus, CBS 792	Y-6666	-	W	-	-	-	-	-	+	-	-	-
	Debaryomyces fabryi, CBS 5948	Y-1453	-	W	-	-	-	-	-	+	-	-	-
	Debaryomyces prosopidis, CBS 8450	Y-27369	-	W	-	-	-	-	-	+	-	-	-
	Debaryomyces nepalensis, CBS 5921	Y-7108	+	W	-	-	-	-	-	+	-	W	-
	Candida gorgasii, CBS 9880	Y-27707	W	+	-	-	-	-	-	+	-	-	-
	Debaryomyces maramus, CBS 1958	Y-2171	W	W	-	-	-	-	-	+	-	W	-
	Schefferzomyces lignosus, CBS 4705	Y-12856	++	++	-	-	-	-	-	+	-	-	-
	Scefferzomyces stipitis, CBS 6054	Y-7124	+	+	-	-	-	-	-	+	-	-	-
Phaffomycetaceae	*Wickerhamomyces anomalus*, CBS 5759	Y-17698	W	-	W	-	-	-	-	+	-	-	-
	Wickerhamomyces ciferrii, CBS 111	Y-1031	W	-	-	W	+	-	W	++	-	-	-
	Wickerhamomyces canadensis, CBS 1992	Y-1888	+	+	-	-	-	-	-	+	-	-	-
	Wickerhamomyces hampshirensis, CBS 7208	YB-4128	W	+	-	-	-	W	W	W	W	-	-
CUG-Ser2	*Saccharomycopsis capsularis*, CBS 2519	Y-17639	+	+	W	W	-	W	-	W	W	-	-
	Saccharomycopsis malanga, CBS 6267	Y-7175	-	-	+	+	-	-	-	-	-	-	-

Growth was ranked from + to +++, where + was regular growth and +++ extensive growth, while W indicates weak growth and – no growth. Growth after 72 h in liquid cultures >OD = 0.2 is indicated by a green color. AX = arabinoxylan, GX = glucuronoxylan, GluMan = glucomannan, GalMan = galactomannan, Poly-MeGal = poly-methylgalacturonan, CMC = carboxymethyl cellulose. Saccharom. = Saccharomycetaceae, Lipomyc. = Lipomycetaceae.

Table 2. Xylanolytic CAZyme predicted from whole-genome sequenced xylanolytic yeasts.

Clade	Species	CE (332/332)	GH3 (263/332)	GH5 (324/332)	GH10 (5/332)	GH11 (1/332)	GH30 (11/332)	GH43 (22/332)	GH51 (39/332)	GH6 (1/33
Trichomonascaceae	*Spencermartinsiella europaea*	CE1, CE4, CE15	GH3(11)	GH5_5(2), GH5_9(2), GH5_12(2), GH5_22(4), GH5_49	GH10(2)		GH30_5, GH30_7	GH43_14, GH43_24	GH51	
	Sugiyamaella lignohabitans	CE1(2), CE4(3), CE15	GH3(5)	GH5_9(2), GH5_12(2), GH5_22(2), GH5_49	GH10(2)		GH30_7		GH51	
	Blastobotrys peoriensis	CE1, CE4	GH3(15)	GH5_9(3), GH5_12(2), GH5_22(5), GH5_49	GH10		GH30_3(3)		GH51	
	Blastobotrys mokoenaii	CE1, CE4	GH3(8)	GH5, GH5_5, **GH5_7**(2), GH5_9(2), GH5_12, GH5_22(2), GH5_49		**GH11**	GH30_5, GH30_7	GH43_6, GH43_24	GH51(3)	**GH6**
	Blastobotrys adeninivorans	CE1, CE4(2)	GH3(8)	GH5_9(3), GH5_12(2), GH5_44, GH5_47			GH30_3	GH43_6	GH51	
	Blastobotrys proliferans	CE1(2), CE4	GH3(12)	GH5_5(3), GH5_9(2), GH5_12, GH5_31, GH5_49			GH30_3		GH51(3)	
	Blastobotrys serpentis	CE1, CE4	GH3(5)	GH5, GH5_9(3), GH5_12, GH5_22, GH5_49			GH30_3(2)		GH51(2)	
	Blastobotrys raffinosifermentans	CE1, CE4(2), CE5(6)	GH3(9)	GH5_9(3), GH5_12(2), GH5_49			GH30_3	GH43_6	GH51	
CUG-Ser1	Spathaspora passalidarum	CE1, CE4(2)	GH3(9)	GH5, GH5_9(3), GH5_22, GH5_49						
	Scheffersomyces lignosus	CE1, CE4,	GH3(7)	GH5_5, GH5_9(3), GH5_12, GH5_22(2), GH5_49	GH10(2)					
	Scheffersomyces stipitis	CE1, CE4,	GH3(7)	GH5_9(2), GH5_12, GH5_22 (3), GH5_49	GH10					
Phaffomycetaceae	*Wickerhamomyces canadensis*	CE1(3), CE4	GH3(5)	GH5_9(2), GH5_12, GH5_22, GH5_49						

CAZyme families marked in **bold** are unique enzymes to the species within the 332-yeast dataset. Total number of species containing each CAZyme family is marked by (x/332). Copy number of each enzyme is stated in parenthesis next to the enzyme. CE = carbohydrate esterase; GH = glycoside hydrolase.

Figures

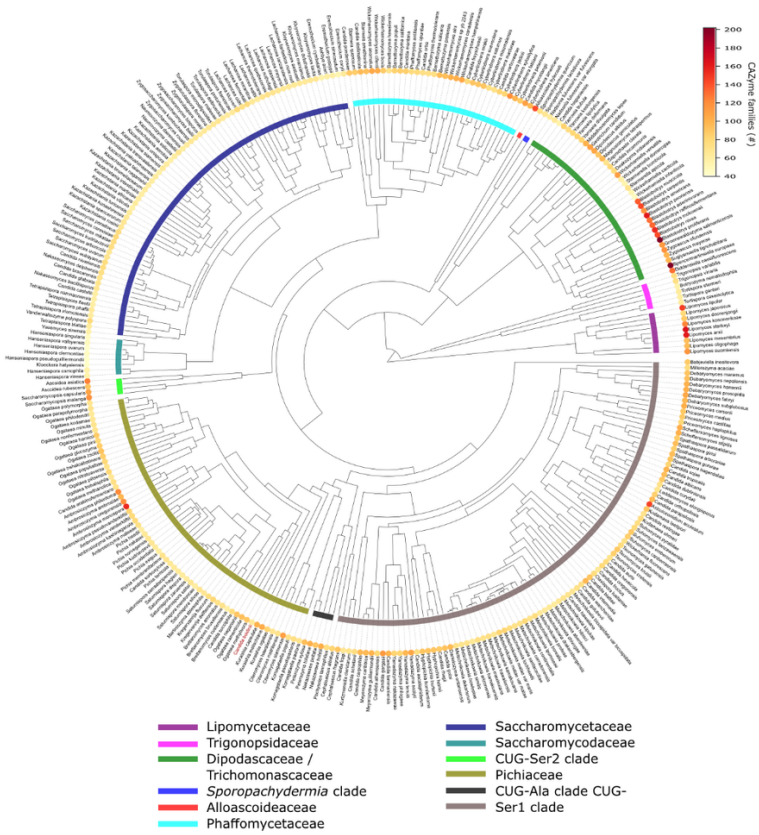

Figure 1

CAZyme abundance in 332 budding yeasts. The total number of predicted CAZymes (GTs excluded) in each yeast species is represented by a heat signature ranging from light yellow to dark red with increasing numbers of predicted CAZymes.

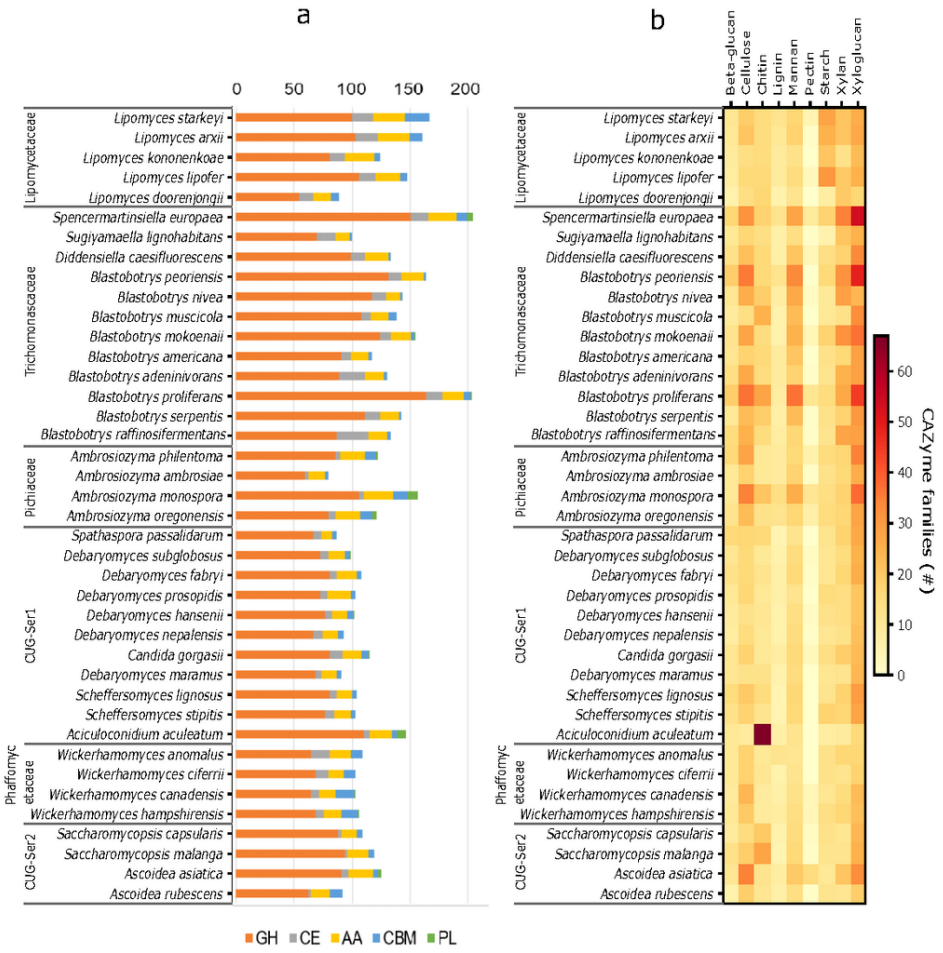

Figure 2

Total number of CAZymes (except GTs) in the 40 selected yeasts and their grouping by function. a Total number of CAZymes in each selected species. b CAZyme families from the same species grouped by predicted function in polysaccharide degradation. Dark red and red-colored squares indicate high number (#) of CAZymes with predicted activity towards the listed polysaccharide. Please note that the heatmap is depicting the total number of CAZyme-encoding genes belongingto families known to degrade specific polysaccharides, and thus heat signatures from polysaccharides with very few CAZymes needed for depolymerization (e.g. β-glucan) may be skewed compared to more complex polysaccharides (such as xylan) requiring many CAZymes. Poly-specific enzyme families such as GH5 and GH3 may also show false positive activities as their members have shown activities on several different β-1,4-linked glycans e.g. xylanase, mannanase, glucanase, glucosidase, galactanase [19]. GH5 enzymes were assigned to cellulose, mannan, xylan, and xyloglucan, while GH3 were assigned to β-glucan, cellulose, xylan and xyloglucan. CBM, carbohydrate binding module; CE, carbohydrate esterase; GH, glycoside hydrolases; PL, polysaccharide lyase.

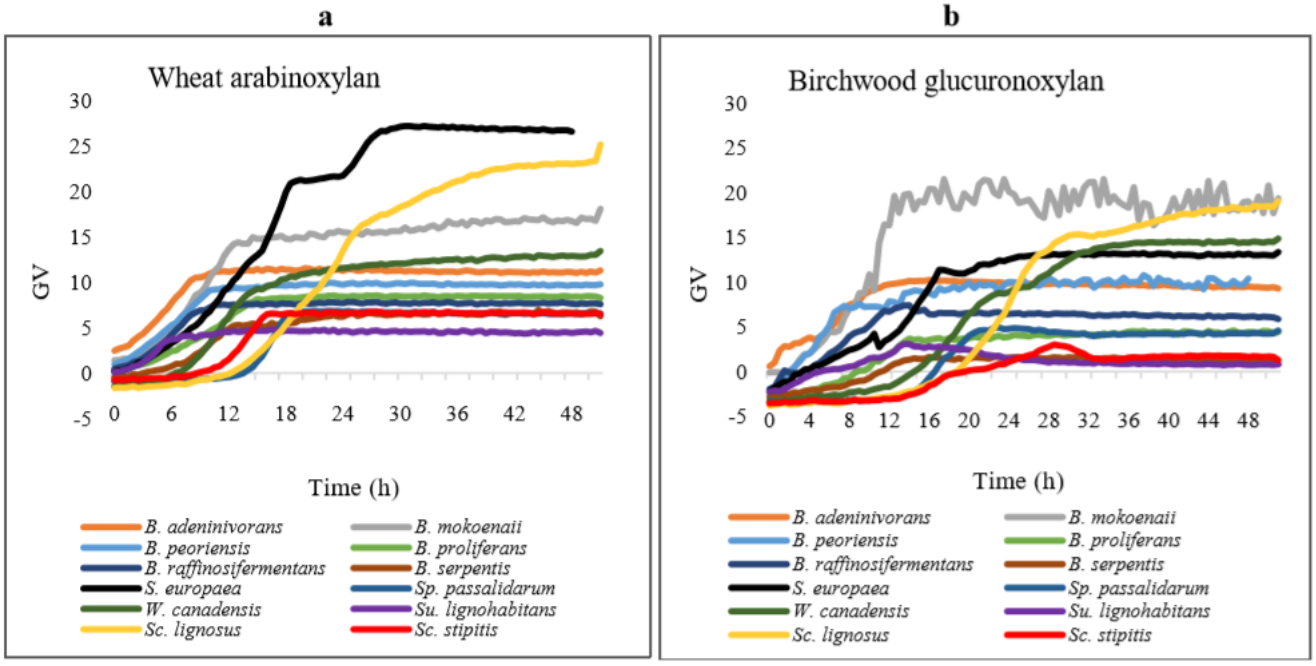

Figure 3

Growth profiles of 12 xylanolytic yeasts in Delft minimal medium containing 10 g/L of either a wheat arabinoxylan or b birchwood glucuronoxylan. GV = Green Value (corresponding to growth based on pixel counts, as determined by a GrowthProfiler instrument). Growth profiles are shown as averages of triplicates.

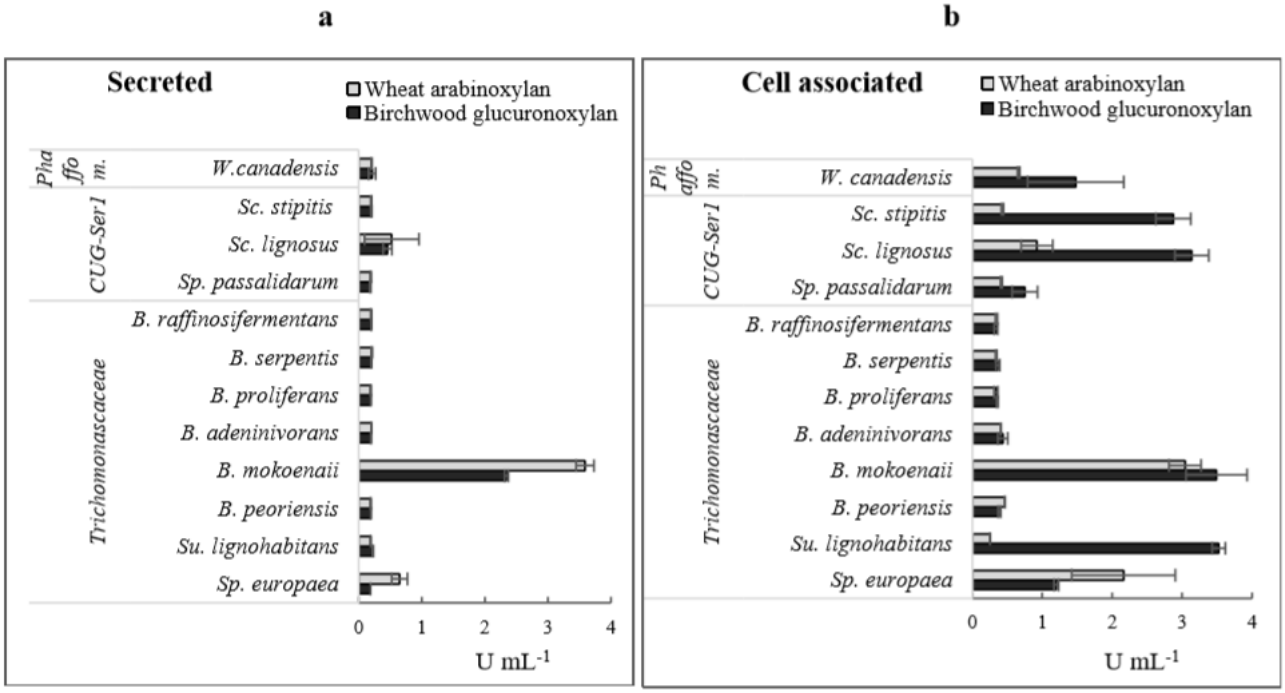

Figure 4

Xylanolytic yeast activities in liquid cultures. Volumetric activities of a secretome xylanases and b yeast cell-associated xylanases in wheat arabinoxylan (grey) and birchwood glucuronoxylan (black) determined at 30 °C after growth on xylan in liquid medium for 72 h. Phaffom. = Phaffomycetaceae clade.

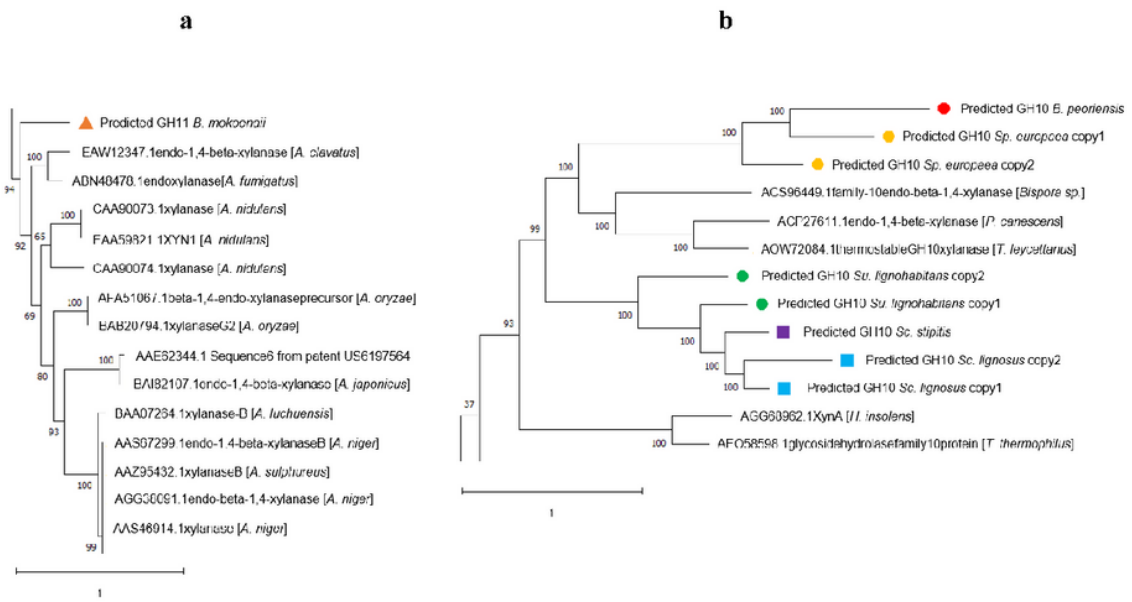

Figure 5

Phylogenetic analysis of GH11 and GH10. a Phylogenetic placement of the GH11 xylanase from B. mokoenaii (orange triangle) and b Phylogenetic placement of the GH10 xylanases from Sp. europaea (yellow circles), Su. lignohabitans (green circles), B. peoriensis (red circle), Sc. lignosus (blue squares) and Sc. stipitis (purple square). The molecular phylogenetic analysis was performed using full protein sequences from 259 GH10 and 208 GH11 characterized enzymes using Newick tree model from MUSCLE alignment with 1000 boot strap replicates. The numbers at each branch indicate bootstrap values and tree topology confidence. Trees are drawn to scale with branch lengths measured in numbers of substitutions per site. Scale bars represents 1.0 substitutions per nucleotide position.

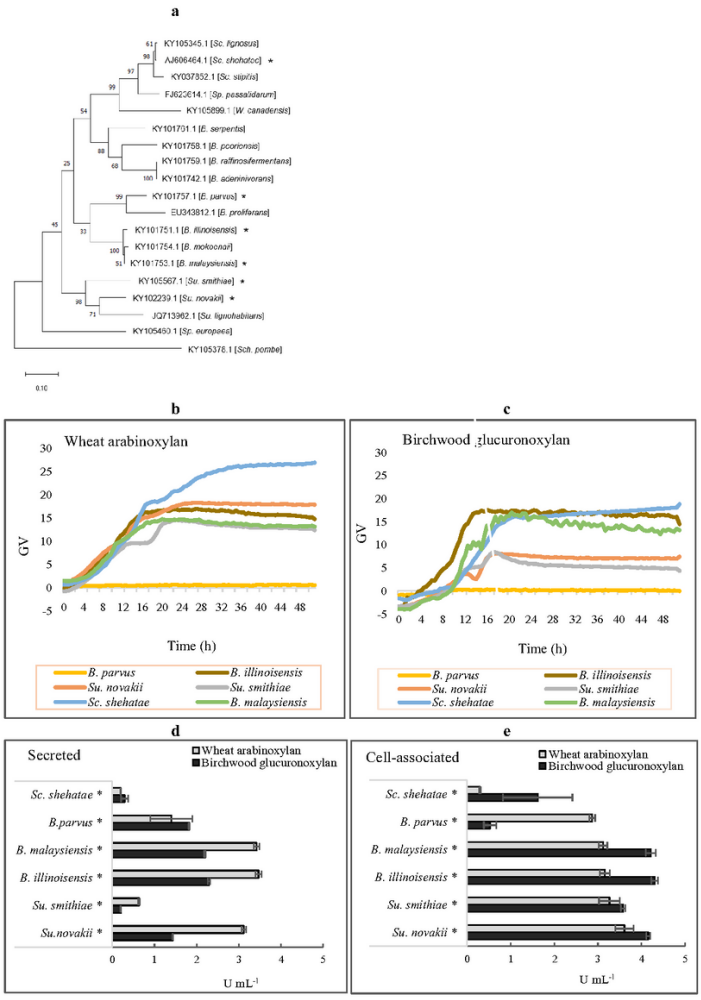

Figure 6

Characterization of non-sequenced xylanolytic yeasts. a Phylogenetic analysis of 19 Blastobotrys, Sugiyamaella and Scheffersomyces species as well as Schizosaccharomyces pombe serving as outgroup. The molecular phylogenetic analysis was based on ITS sequences using maximum-likelihood model from ClustalW alignment with 1000 bootstrap replicates. The numbers at each branch indicate bootstrap values and tree topology confidence. The tree is drawn to scale, with branch lengths measured in the number (0.2) of substitutions per site. Growth profiles of xylanolytic yeasts grown in Delft medium containing 10 g L-1 of b wheat arabinoxylan and c birchwood glucuronoxylan. Yeasts were grown for 48 h at 30 °C. GV = Green Value (corresponding to growth based on pixel counts, as determined by a GrowthProfiler instrument). d Secretome and e cell-associated volumetric xylanase activities on wheat arabinoxylan (grey) and birchwood glucuronoxylan (black) determined at 30 °C after growth in xylan-containing liquid medium for 72 h. Stars (*) symbolizes non-sequenced species.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.pdf
- SupplementaryFigureS2.pdf
- SupplementaryFigureS3.pdf