UNIVERSITY of York

This is a repository copy of *A multi-omics approach to lignocellulolytic enzyme discovery reveals a new ligninase activity from Parascedosporium putredinis NO1*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/173215/

Version: Accepted Version

Article:
Oates, Nicola Claire, Abood, Amira, Schirmacher, Alexandra et al. (11 more authors) (2021) A multi-omics approach to lignocellulolytic enzyme discovery reveals a new ligninase activity from Parascedosporium putredinis NO1. Proceedings of the National Academy of Sciences of the United States of America. e2008888118. ISSN 1091-6490

https://doi.org/10.1073/pnas.2008888118

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

PNAS
www.pnas.org

Main Manuscript for

A multi-omics approach to lignocellulolytic enzyme discovery reveals a new ligninase activity from *Parascedosporium putredinis* NO1

Nicola C. Oates[a], Amira Abood[a], Alexandra M. Schirmacher[a], Anna M. Alessi[a], Susannah M. Bird[a], Joseph P. Bennett[a], Daniel R. Leadbeater[a], Yi Li[a], Adam A. Dowle[b], Sarah Liu[c], Vitaliy I. Tymokhin[c], John Ralph[c], Simon J. McQueen-Mason[a], and Neil C. Bruce[a]

[a]Centre for Novel Agricultural Products, Department of Biology, University of York, York, YO10 5DD, UK.

[b]Bioscience Technology Facility, Department of Biology, University of York, York, YO10 5DD, UK.

[c]Department of Biochemistry, and the Department of Energy's Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin, Madison, WI, 53726, USA

Correspondence and requests for materials should be addressed to N.C.B.

Email: neil.bruce@york.ac.uk

Classification

Biological sciences: Applied Biological Sciences

Keywords

Lignocellulose, Lignin, β-etherase, Tricin, *Parascedosporium putredinis*

Author Contributions

N.C.B. and J.R. conceived the idea; N.C.O., S.J.M-M. and N.C.B. designed the experiments; N.C.O., A.A., A.M.S., A.M.A., S.M.B., J.P.B., D.L., Y.L., A.A.D. S.L and V.I.T. performed the research; N.C.O., A.A., A.S., J.P.B., D.R.L., Y.L., J.R., S.J.M-M. and N.C.B. analyzed data; N.C.O and N.C.B. wrote the manuscript with contributions from co-authors. All authors reviewed and approved the final manuscript.

Abstract

Lignocellulose, the structural component of plant cells, is a major agricultural byproduct and the most abundant terrestrial source of biopolymers on Earth. The complex and insoluble nature of lignocellulose limits its conversion into value-added commodities and, currently, efficient transformation requires expensive pretreatments and high loadings of enzymes. Here, we report on a fungus from the *Parascedosporium* genus, isolated from a wheat-straw composting community, that secretes a large and diverse array of carbohydrate-active enzymes (CAZymes) when grown on lignocellulosic substrates. We describe a new oxidase activity that cleaves the major β-ether units in lignin, thereby releasing the flavonoid tricin from monocot lignin, and enhancing the digestion of lignocellulose by polysaccharidase cocktails. We show that the enzyme, which holds potential for the biorefining industry, is widely distributed among lignocellulose-degrading fungi from the Sordariomycetes phylum.

Significance Statement

Lignocellulose, in the form of crop residues, presents an attractive alternative to crude oil for both the production of renewable fuels and chemicals. Its large-scale application as a feedstock, however, remains limited. A bottleneck in its implementation is the presence of lignin, a complex hydrophobic polymer, that envelopes the structure, physically blocking access to sugar-rich polymers that lie beneath. Here we describe the isolation of an exceptional lignocellulose-degrading fungus that produces a new oxidase activity with no cofactor requirements. This enzyme cleaves β-ether units in lignin releasing tricin, a flavonoid of pharmaceutical potential, from the lignin macromolecule. Furthermore, we demonstrate that treatments with this enzyme

can increase the digestibility of lignocellulosic biomass, offering the possibility of producing a valuable product from lignin while decreasing processing costs.

Main Text

Introduction

Photosynthetically-fixed carbon in lignocellulose is produced in vast quantities on the Earth's surface. The abundance of crop residue lignocellulose makes it an attractive alternative to crude oil in the production of renewable, low-carbon, fuels and chemicals (1). Effective utilization of lignocellulose, nevertheless, remains a challenge, as the extraction of sugars for fermentation requires intensive physicochemical pretreatments and high loadings of enzyme cocktails. A key factor in the recalcitrance of lignocellulose to degradation is lignin, a heterogeneous, hydrophobic aromatic polymer that encases the cellulose and hemicellulosic polysaccharides, blocking enzyme accessibility and impeding cellulase activity (2, 3).

Lignin is typically synthesized in secondary cell walls of higher plants through the phenoxy-radical coupling of the differentially methoxylated hydroxycinnamyl alcohols, sinapyl alcohol, coniferyl alcohol, and *p*-coumaryl alcohol, generating β–O–4, 4–O–5, β–5, β–1, 5–5 and β–β inter-unit linkages in β-ether, biphenyl ether, phenylcoumaran, spirodienone, biphenyl, and resinol units, respectively. Tricin [5,7-dihydroxy-2-(4-hydroxy-3,5-dimethoxyphenyl)-4*H*-chromen-4-one], an *O*-methylated flavone recognized for its pharmaceutical potential due to its antioxidant and antibacterial properties, with purported functions ranging from anti-tumor activity to potential diabetes suppression (4, 5). Tricin has been recently described to form part of the structure of lignin from monocotyledonous plants including wheat, rice, and sugarcane (6-9). To date, tricin has only been observed incorporated into the lignin structure via 4–O–β linkages, having arisen from the radical coupling of the flavone at its 4'–O-position with a

74 monolignol at its β-position. It is therefore reported to be a potential nucleation point at which
75 lignification initiates in monocots, locating tricin at the (starting) terminus of the lignin
76 macromolecule (8, 10). The selective extraction of tricin from monocots, which encompass the
77 majority of agricultural biomass, using methods amenable to the lignocellulose biorefinery
78 concept could potentially go some way towards mitigating the cost of second-generation biofuel
79 production (11).

80 Despite extensive research into the biological degradation of lignocellulose and the mining
81 of microbial communities for their ability to break down cellulose and hemicelluloses, questions
82 persist over the biodegradation of lignin, and the mechanisms that facilitate its
83 depolymerization. Although delignification involving manganese superoxidase dismutases (12),
84 laccases, and dye-decolorizing peroxidases (13) by bacteria has been described, it is fungi that
85 are the major lignin degraders in the terrestrial environment. Historically, wood-decaying fungi
86 have been divided into white-, brown-, and soft-rots, depending on the morphology of their
87 decomposition products. White-rot Basidiomycetous fungi, such as *Ceriporiopsis subvermispora*
88 or *Phanerochaete chrysosporium*, are named for their ability to degrade and mineralize the
89 dark-colored lignin and selectively enrich the white cellulose (14). This degradation is catalyzed
90 through the action of oxidative enzymes (oxidoreductases) such as laccases, manganese
91 peroxidases, dye-decolorizing peroxidases, and high-redox-potential heme-peroxidases (15, 16).

92 Conversely, although lignin modification has been recognized in brown-rot species (17, 18),
93 significant solubilization does not occur. Instead, the degradative strategy is typified by the
94 selective removal of the polysaccharide components from the plant cell wall through chemical
95 and enzymatic means (19-21). Soft-rot fungi that, unlike the Basidiomycete white- and brown-
96 rots, are usually Ascomycetes, deploy an alternative strategy. These fungi are capable of

extensively degrading lignocellulose through the secretion of large quantities of enzymes close to the site of attack due to the penetrating nature of their filamentous hyphae (22). This causes characteristic softening of the lignocellulose, as the plant cell walls lose their structural integrity. Ascomycetes are not, however, well known for their ability to solubilize lignin and, although there have been reports that they possess the capacity to modify and degrade lignin (23), it is not clear how this occurs.

Uncovering the mechanisms that govern lignin degradation by Ascomycetes, represents an opportunity to discover new enzyme systems that can be employed in biorefining applications. Here we report the discovery of an Ascomycete, *Parascedosporium putredinis* NO1, isolated from straw-enriched compost, that thrives in the latter stages of decomposition. We identify an extensive arsenal of lignocellulose-degrading enzymes during growth on wheat straw, and report on the abundance of secreted, and as yet, hypothetical proteins. In particular, we report the discovery of a new oxidase that can cleave the β–O–4 inter-unit structural linkage of lignin, releasing tricin from monocot biomass and boosting the digestion of the biomass by cellulases. Importantly, this enzyme does not require a cofactor.

Results

Isolation of *Parascedosporium putredinis* NO1

We inoculated liquid cultures containing wheat straw as the sole carbon source with homogenized samples of wheat straw-enriched compost. From these we tracked the dynamics of the resulting microbial community. Sequencing of 16S ribosomal RNA genes generated over three million reads from the prokaryotic community over the whole timecourse, which clustered together to form 25,304 operational taxonomic units (OTUs) (Fig. 1*A*). The most abundant bacterial phyla identified

were the gram-negative *Bacteroidetes*, *Verrucomicrobia* and *Proteobacteria*, respectively, representing an average of 31, 19.8 and 15.5% of the total reads across the time course. Analysis of the eukaryotic community by sequencing the Internal Transcribed Spacer (ITS) region predominantly yielded reads that had no match within the UNITE fungal rDNA sequence database (24, 25). In total, 96.5% of generated OTUs were not recognized as fungal and instead showed the closest homologies to protozoa. Among the fungi, we noted distinct changes in the composition of the community with time. In particular, a fungus (designated strain NO1) an Ascomycete in the Microascaceae family, showed increased abundance after 4 weeks of incubation (Fig. 1*B*). Reads assigned to genus *Graphium* dominated the eukaryotic community in the shake flasks after four weeks of incubation, representing 84% of the identifiable fungal reads at 8 weeks, a time point by which, we hypothesize, the majority of easily accessible carbon from wheat straw has been depleted (26). A synamorph of *Graphium, Parascedosporium putredinis* strain NO1 identified through ITS and 18S analysis was readily isolated from shake flasks by culturing on both nutrient agar and potato dextrose agar. Interestingly, this fungus could be selectively cultivated when agar plates contained kraft lignin as the sole carbon source.

Omics analysis of wheat straw degradation by *P. putredinis* NO1

We confirmed that *P. putredinis* NO1 could grow on wheat straw as a sole carbon source and optimized the composition of growth media for cellulase and xylanase production (*SI Appendix*, Fig. S1). The deconstruction of wheat straw by *P. putredinis* NO1 over 28 days was tracked by measuring mass loss and carbohydrate-active enzyme (CAZy) activity (*SI Appendix,* Fig. S2). This growth experiment identified the second, fourth and tenth day of incubation on wheat straw as distinct time points to harvest RNA for sequence analysis, chosen because together they represent the first detection of lignocellulolytic activity (day 2), the peak of enzyme activities (day

144 4) and the subsequent reduction of lignocellulolytic activity (day 10) – a point at which the easily
145 accessible sugars in the wheat straw had been utilized. RNA was also harvested from *P. putredinis*
146 grown on glucose for four days.

147 After sequencing, 5,586 unique contiguous DNA sequences (contigs) were assembled from
148 339,854,704 reads, and differential gene analysis identified 2,189 contigs that were upregulated
149 at high confidence and fold-change ($P<0.001$, FC >10) between growth on wheat straw compared
150 to glucose. These highly upregulated genes included those coding for 102 putative CAZy proteins;
151 47 glycoside hydrolases (GH), 41 auxiliary activities (AA), ten carbohydrate esterases (CE) and a
152 polysaccharide lyase (PL), the majority of which were upregulated after four days of growth (Fig.
153 2), in agreement with the peak of the observed enzymatic activities in *P. putredinis* NO1 culture
154 supernatants.

155 As the macromolecular structure of lignocellulose prohibits intracellular degradation,
156 enzymes for its deconstruction are typically secreted. To capture these enzymes, we performed
157 LC-MS/MS analysis on protein samples collected directly from the culture supernatant, and
158 separately, from those bound to insoluble components of the culture using a biotin-labelling
159 method (27). From across all samples, 3,671 proteins were identified, including 1,037 proteins
160 present in only wheat straw conditions (*SI Appendix,* Fig. S3*A*) and 275 sequences that contained
161 a recognizable CAZy domain (*SI Appendix*, Dataset S1). These putative carbohydrate-
162 deconstructing enzymes accounted for 25.2% (192 proteins) of the molar percentage of the
163 supernatant samples and 13.9% (171) of the biotin-labelled samples after four days of growth
164 on wheat straw, compared to 13.3% (97) of the supernatant and 2% (56) of the biotin labelled
165 samples from glucose-grown cultures (*SI Appendix*, Fig. S3*B*).

166 The most abundant CAZy family, 3.7% and 3.6% of the respective supernatant and biotin-
167 labelled fractions on the fourth day, were GH6s (Fig. 3). These, along with GH7s, often constitute
168 the major cellulases in filamentous fungi (28), and may be endoglucanases or processive
169 cellobiohydrolases. Other GH families that are likely active on cellulose, including the GH7
170 (typically cellobiohydrolases or endoglucanases), GH5 and GH45 (often endoglucanases), and
171 GH1 and GH3 (typically glucosidases) families (29), were also prominent within the secretome.
172 Efficient lignocellulose deconstruction, however, demands a combination of cellulolytic and
173 hemicellulolytic enzymes that work cooperatively. Enzymes related to the depolymerization of
174 arabinoxylan (the major hemicellulose in wheat straw), were well represented within the
175 exoproteome. Nine proteins were identified with homology to endo β-1,4-xylanases (GH10 and
176 GH11) that hydrolyze the arabinoxylan backbone, and five proteins were identified as putative
177 β-1,4-xylosidases that act on the resultant fragments to produce xylose monomers (GH3, GH31,
178 GH43_1, GH43_11, GH43_36). Also of note, were the GH43 subfamilies GH43_1, GH43_21,
179 GH43_22, GH43_26 and GH43_36 that were abundant within the secretome, and include
180 putative β-D-xylosidases, α-L-arabinofuranosidase, and β-1,3–galactosidase activities.
181 Three proteins, belonging to the CE1 family, showed significant sequence homology to
182 feruloyl esterases. Ferulate acylates the arabinose side-chain of arabinoxylans, and through the
183 formation of diferulate bridges and ester-ether linkages allows the respective formation of
184 covalent interactions between arabinoxylan chains (with each other) and with lignin. Feruloyl
185 esterases, therefore, are thought to aid the solubilization of plant cell wall polysaccharides by
186 the hydrolysis of the ester link that exists between ferulic acid residues and arabinose, thereby
187 disrupting the cross-linking of cell wall components (30). Putative acetyl xylan esterases (3 in
188 CAZy family CE1 and 3 in CE5) were also observed and are known to facilitate the degradation of

189 xylan through the removal of acetyl substitutions that render the substrates more recognizable
190 by polysaccharidase enzymes (31).
191 The CAZy auxiliary activity (AA) class is classified as containing enzymes that act in conjunction
192 with carbohydrate-active enzymes through redox activities. Interestingly, 69 putative proteins
193 from the AA class were detected in the exosecretome, more than many lignocellulose-degrading
194 fungi contain in their total genome (32), suggesting an important role for the oxidative
195 degradation of lignocellulose in *P. putredinis* NO1. The AA9 family were highly represented
196 within the exosecretome. This family, along with the AA10, AA11, AA13, AA14 and AA15
197 families, constitutes the lytic polysaccharide monooxygenases (LPMOs) – a class of copper
198 metalloenzymes that catalyze the oxidative cleavage of glycosidic bonds in multiple
199 polysaccharide substrates including chitin, cellulose, and xylan (33, 34). In total, we identified
200 nineteen putative LPMOs (16 AA9s; 2 AA11s; 1 AA13), fifteen of which were upregulated tenfold
201 or more between glucose and wheat straw conditions. Fittingly, 16 AA3s (glucose-methanol-
202 choline (GMC) oxidoreductase) and 9 AA7s (glucooligosaccharide oxidase), which have been
203 shown to facilitate the activity of the LPMOs through electron shuttling (35, 36), were also
204 present within wheat straw cultures.
205 Established lignin depolymerizing enzymes associated with the white-rot fungal decay of
206 lignin, including laccases from the AA1_1 subfamily (37, 38), or peroxidases from the AA2 family
207 (14), were not present within the libraries. This is perhaps not surprising given that *P. putredinis*
208 NO1 sits within the Ascomycota phylum, and as such is closer in relation to the soft-rots.
209 Five putative multicopper oxidase proteins were also observed – two from the AA1_3
210 subfamily (Laccase-like multicopper oxidase) and one from the AA1_2 subfamily (Ferroxidase).
211 Laccase-like multicopper oxidases are of unknown function but have been implicated in lignin

degradation, as well as other diverse functions (iron homeostasis, offense/defense) (39), whereas ferroxidases have been reported to be involved in lignocellulose degradation in Ascomycetes, in which they generate hydroxyl radicals via the Fenton reaction (40).

Despite the apparent lack of known ligninases in *P. putredinis* NO1, a putative AA6 (1,4-benzoquinone reductase) associated with the intracellular biodegradation of aromatic compounds was present within the supernatant and may have a role in the metabolism of lignin breakdown products (32, 41).

Of key interest to us was the potential of *P. putredinis* NO1 to produce novel lignocellulolytic activities, particularly those able to boost lignocellulose deconstruction via the modification and solubilization of lignin. An unknown protein, c2092_g1_i1, identified in the exosecretome was subsequently found to have β-etherase activity and no CAZy identification.

A new oxidase displaying β-etherase activity

The β-ether motif, with its characteristic β–O–4 inter-unit linkage, is the most abundant in lignin, estimated at representing over 50% of the total inter-unit linkages (42). Enzymes employing β-ether cleavage mechanisms can deconstruct synthetic and extracted lignin (43-45); these bacterial etherases that have been characterized to date, however, are intracellular proteins, and are glutathione- or NAD^+- dependent, suggesting that in nature they are not directly involved in the breakdown of the lignin macromolecule, but rather its smaller, membrane-transportable oligomers. An extracellular fungal protein displaying β-etherase activity was previously purified from the supernatant of the *Chaetomium* sp. 2BW- 1, although its identity remains unknown (46).

Using a synthetic lignin model compound, GGβ4MU (7-[2-hydroxy-2-(4-hydroxy-3-methoxyphenyl)-1-(hydroxymethyl)ethoxy]-4-methyl-2*H*-1-benzopyran-2-one), containing a β-

235 methylumbelliferyl ether, guaiacylglycerol-β-(4-methylumbelliferyl) ether that when cleaved
236 yields the fluorogenic product 4-methylumbelliferone (4MU) (*SI Appendix,* Fig. S4) (47), we
237 detected β-etherase activity within the culture supernatant of *P. putredinis* NO1. This activity
238 was present when *P. putredinis* NO1 was grown on wheat straw but not on glucose, suggesting a
239 possible role in lignocellulose degradation, and appeared to be independent of cofactors such as
240 glutathione or NAD^+. Given its presence in the secretome and its apparent cofactor
241 independence, we hypothesized that this putative ligninase was unlikely to share significant
242 sequence homology to the previously described intracellular β-etherases from sphingomonads,
243 and indeed no proteins with similarity to these enzymes were detected. We, therefore,
244 subjected the culture supernatant of *P. putredinis* NO1 grown on wheat straw to a series of
245 protein fractionation techniques, enriching at each step for β-etherases activity.
246 The putative β-etherase was initially purified by ammonium sulfate precipitation of the
247 proteins in the culture supernatant to decrease sample pigmentation and reduce protein-
248 protein interactions. This treatment facilitated further purification by size-exclusion and anion-
249 exchange chromatography. Using shotgun proteomics, we identified c2092_g1_i1, a 44.5 kDa
250 protein present in the purified fraction that contained a predicted signal peptide. Analysis of the
251 transcriptomic and proteomic data revealed this protein was strongly upregulated in the
252 presence of wheat straw and present in both the supernatant and biotin-labelled proteomic
253 libraries throughout the growth of *P. putredinis* NO1 on wheat straw (*SI Appendix,* Fig. S5). Using
254 profile Hidden Markov models constructed by HMMER3 on using the pFAM database (48), we
255 saw homology to a common central tyrosinase domain (PF00264; Evalue = 7.1e-49) with a
256 characteristic binuclear type-3 copper-binding site consisting of six histidine residues located in
257 a four-helical bundle coordinating the binding of two copper ions (49) (*SI Appendix,* Fig. S6).

Mushroom tyrosinase (*Agaricus bisporus*), has been reported to have promiscuous β-etherase activity on small synthetic compounds but no significant activity has been reported against macromolecular lignin (50) . Fungal tyrosinases (polyphenol oxidases) are predominantly associated with pigmentation and browning; specifically, through melanin production, whereby they catalyze the introduction of a hydroxyl group *ortho* to the phenol in a *para*-substituted monophenol and the subsequent oxidation to the corresponding *o*-quinone (51). However, c2092_g1_i1 lacks both the C- and N-terminal domains that tyrosinases typically contain and instead shows higher homology (170/370 identity (46%)) to a catechol oxidase (AoCO4) from *Aspergillus oryzae* (52), which differs from tyrosinases due to a lack of mono-oxygenase activity (53). Examination of the proteomics library resulted in the identification of seven sequences with significant similarities to c2092_g1_i1 (*SI Appendix,* Table S1), all predicted to be extracellular and soluble, and five upregulated in the presence of wheat straw (*SI Appendix,* Fig. S7). Searches within the NCBI non-redundant database further revealed the presence of proteins of similar sequence (>50% sequence identity) distributed throughout fungal genomes of the Sordariomycetes class of Ascomycetes (*SI Appendix*, Table S2, Fig. S8).

Experimental confirmation of β-etherase activity

To determine if c2092_g1_i1 was responsible for the observed β-etherase activity, we heterologously expressed the codon-optimized sequence in *Escherichia coli*. The recombinant protein was purified (*SI Appendix*, Fig. S9), and the β-etherase activity of the protein was confirmed by determining the level of fluorescence released after incubation with GGβ4MU (*SI Appendix,* Fig. S10*A*). In reaction conditions absent of oxygen, we confirmed the oxidative nature of this protein, seeing a near total reduction in activity when the assay was conducted under anaerobic conditions (*SI Appendix*, Fig. S10B).

The pH and temperature dependency of the enzyme were investigated, revealing maximum activity at pH 10 and 60 °C (*SI Appendix,* Fig. S10*C* and *D*). The *P. putredinis* NO1 oxidase did not display activity against L-tyrosine and L-DOPA, as is characteristic for tyrosinases (*SI Appendix,* Fig. S11) (54). We subsequently assayed for potential oxidase activity against a range of phenolic substrates, including di-phenolics, known to be catechol oxidase substrates (53), and observed no similarities to catechol oxidase in terms of substrate preference (*SI Appendix,* Fig. S12, Table S4). Interestingly, activity was seen with the substrates: 4-hydroxybenzoic acid, vanillic acid, and quercetin, all known to be tyrosinase inhibitors (55).

Release of tricin and lignin units from wheat straw

Tricin has recently been described as a subunit in the lignin of monocot species, incorporated through a 4–O–β linkage (11). As wheat straw contains relatively high concentrations of tricin compared to other agriculturally relevant feedstocks (8), we assessed the ability of the oxidase to release tricin from wheat straw. The oxidase was incubated with wheat straw for sixteen hours under physiological conditions (pH 8.5 and 30 °C). Reaction products were monitored by High-Performance Liquid-Chromatography (HPLC), and a peak corresponding to tricin was identified by reference to an authentic standard and confirmed by mass spectrometry. Under the growth conditions used for *P. putredinis* NO1, a significantly higher concentration of tricin was present in the reaction supernatant of wheat straw with the purified protein fractions compared to incubations with buffer alone (ANOVA, $F(2,12)=44.67$, $p<0.05$) (Fig. 4*A*). We were also able to detect the presence of *p*-coumaric acid, vanillin, and *p*-hydroxybenzaldehyde in the reaction supernatant through comparisons with authentic standards and mass spectrometry; however, unlike tricin, these compounds were not enriched under the enzyme-treated reaction

303 conditions (*SI Appendix,* Fig. S13) and presumably are produced as a result of simple ester
304 cleavage.
305 NMR (Fig. 5) of the enzyme lignins (EL) isolated (following crude polysaccharidase treatment
306 to saccharify most of the polysaccharides) (56), and the product generated from it by a non-
307 optimized treatment with our enzyme showed little change to the actual lignin profile but a
308 strong decrease in the tricin level. Thus, even though integration of correlation contours in the
309 spectra resulting from such 2D-HSQC (heteronuclear single-quantum coherence) experiments
310 does not provide reliable quantification, their relative values are considered to be valid (57, 58).
311 Analysis showed that the relative tricin ether level in the lignin dropped from nearly 12% in the
312 control to about 8.5% after the treatment. We were initially disappointed that we couldn't
313 detect similar reductions in levels of the β-ether units **A** (Fig. 5), but caution that these are
314 'quantified' on an A+B+C=100% basis and it is easy to speculate on how the levels might not
315 significantly change even with some (presumably low-level) β-ether cleavage; in changing the
316 basis to a level per 100 aromatic (S+G+H) units as previously used (7, 23), the values (Fig. 5) do
317 confirm a modest drop (67 down to 65) in β-ether units **A** specifically. In spectra from the whole
318 cell wall component (and not just the isolated lignin, not shown), the trends were similar and
319 the T_6 and T_8 contours were particularly weak in the treated sample whereas the T2′/6′ peak
320 was relatively strong; we have noted this occurrence before in rapidly relaxing samples, and do
321 not fully understand its origin; regardless, the relative tricin level in the treated material was
322 again lower than in the control and obviously consistent with the measured release of tricin
323 noted above.
324 As mushroom tyrosinase has been reported to cleave β-ether linkages promiscuously (50),
325 we tested its β-etherase activity on wheat straw under equivalent conditions. We observed less

tricin production in the reaction mixtures containing mushroom tyrosinase compared to the *P. putredinis* NO1 enzyme treatments. As tricin is a known tyrosinase inhibitor that binds non-competitively to the hydrophobic pocket of the protein (59), and *p*-coumaric acid has been characterized as having a mixed-type inhibition effect (60), inhibition through the non-reversible binding of the reaction products, could go some way to explaining why mushroom tyrosinase displays little activity towards the lignin macromolecule.

Enzyme pretreatment boosts saccharification

To investigate if a pretreatment of wheat straw with the *P. putredinis* NO1 oxidase would improve saccharification rates, we incubated wheat straw with enzyme for sixteen hours before the addition of commercial cellulases, and observed a 20% increase in the level of glucose released compared to wheat straw treated with buffer alone (ANOVA, $F(2,12)=4.47$, $p<0.05$) (Fig. 4*B*).

Discussion

P. putredinis NO1 is able to dominate cultures in the latter stages of wheat straw degradation in a mixed microbial community when easily accessible polysaccharides have been exhausted. Using a combination of 'omics approaches, we have identified a diverse range of potentially industrially relevant carbohydrate-active enzymes, including a large number of enzymes associated with the oxidative attack on lignocellulose. In particular, we have identified a new extracellular oxidase that is preferentially expressed in the presence of wheat straw and demonstrated that this enzyme can release the pharmaceutically relevant flavonoid tricin from monocot lignin. We also demonstrated that pre-treatment with the oxidase can significantly boost the saccharification of wheat straw when used with a commercial cellulase cocktail.

Whether this is a direct effect of increased carbohydrate availability due to lignin removal or a boosting effect of the oxidase on components of the saccharification cocktail remains to be established. We contend that this ability to deconstruct and modify lignin is important for *P. putredinis* NO1 to be able to out-compete other microbial species during the latter stage of plant biomass degradation when easily accessible lignocellulose components are depleted. In a recent report, the ascomycete fungus *Podospora anserina* was observed to possess ligninolytic activities, and analysis of the residual lignin after fungal growth on wheat straw demonstrated a decrease in tricin. *P. anserina* encodes for a number of proteins which share homology to the enzyme described here, two of which were detected in the proteome of *P. anserina* when cultured with wheat straw lignin (23). Preferential removal of tricin subunits has also been described by the white-rot fungi, *Pleurotus eryngii*, during the selective delignification of wheat straw and has been proposed to be key to lignocellulose degradation, although the enzyme activity that facilitated tricin release was not identified (61). When the publicly available genome of *P. eryngii* was examined for the presence of proteins with homology to the oxidase from *P. putredinis* NO1, no significant hits were detected.

As the protein described as being responsible for β-etherase activity from *Chaetomium* sp. 2BW-1 was not identified to sequence level, it is unclear whether it shares homology to the enzyme described here; however, the proteins appear to be distinct as the reported sizes differ by 20 kDa (46). Taken together, these observations suggest that multiple, structurally dissimilar, enzymes in the natural environment may mediate ether linkage disruption in lignocellulose-degrading microbes.

This enzymes ability to release tricin from lignin and could have potential biotechnological applications. To the best of our knowledge, this is the first identification and characterization of

372 an extracellular enzyme capable of β-ether cleavage that has no cofactor requirement for
373 activity.
374
375 **Materials and Methods**
376
377 **Methods**
378 **Wheat straw degradation in shake-flasks inoculated with compost**
379 Wheat straw compost that had developed over a year, was used to inoculate 1 L minimal media
380 cultures that had been supplemented with 5% (w/v) milled wheat straw. Flasks were incubated
381 at 30 °C and shaken at 150 rpm. The minimal media contained KCl 0.52 g/L, KH_2PO_4 0.815 g/L,
382 K_2HPO_4 1.045 g/L, $MgSO_4$ 1.35 g/L, $NaNO_3$ 1.75 g/L, and Hutner's trace elements and was based
383 on *Aspergillus niger* minimal media (62). Spread plates on nutrient agar (NA) and potato
384 dextrose agar (PDA) were created weekly from serial dilutions.
385 **Targeted amplicon sequencing of 16S and ITS region**
386 Genomic DNA was harvested from the compost cultures using a modified CTAB protocol as
387 described by Alessi *et al.* (26). Phusion® High-Fidelity DNA Polymerase (Finnzymes OY, Finland)
388 was used to generate amplicons for sequencing. These were purified using Agencourt AMPure
389 XP (Beckman Coulter). Sequencing was performed using an Ion Torrent platform at the
390 Biorenewable Development Centre, York, U.K. The primer pairs, were as follows; ITS1 Fw –
391 TCCGTAGGTGAACCTGCGG, Rv – CGCTGCGTTCTTCATCG (63), 16S Fw – AYTGGGYDTAAAGNG, Rv
392 – TACNVGGGTATCTAATCC (64), for ITS and 16S sequencing respectively. Resultant sequences
393 were demultiplexed, primer sequences were removed and reads adjusted for orientation. Reads
394 without recognizable primer sequences and under 180 bp were filtered out. Remaining reads
395 were then analyzed using Qiime (65). The open- reference operational taxonomic unit (OTU)
396 picking process was used, before taxonomy was assigned using the greengenes gg_13_8 97 otus
397 database (66, 67) and the UNITE (alpha release 12_11) database (68) for bacterial and fungal
398 identification respectively.
399
400 **Central composite design for media optimization**

A central composite design with rotation (69) was used to optimize media for the production of cellulase and xylanase activity after seven days of incubation in minimal media with 1.5 % wheat straw, as described in Oates (70). Cellulase and xylanase activity was assessed by incubating supernatant on carboxymethylcellulose (CMC) or xylan (beechwood) and measuring reducing sugar release. The sodium nitrate concentration was varied between 0 g/L and 3.5 g/L, and yeast extract was varied between 0% and 1% (w/v). The optimized media for *P. putredinis* NO1 growth consisted of yeast extract 8.55 g/L, KCl 0.52 g/L, KH_2PO_4 0.815 g/L, K_2HPO_4 1.045 g/L, $MgSO_4$ 1.35 g/L, $NaNO_3$ 1.75 g/L and Hutner's trace elements.

Characterization of *P. putredinis* NO1 growth on wheat straw

Biomass, that had been gently rinsed with x1 PBS, was flash-frozen in liquid nitrogen and lyophilized to calculate the dried weight of cultures. To estimate total protein, aliquots of 100 µg of this biomass was boiled in 0.2 % (w/v) sodium dodecyl sulfate (SDS) for 5 mins, and vigorously vortexed, before supernatant was collected through centrifugation. This was repeated three times. Five-volumes of ice-cold acetone was used to precipitate protein overnight at -20 °C and protein pellets were collected via centrifugation, washed with 80 % v/v) ice-cold ethanol and resuspended in H_2O, before being quantified using the Bradford assay.

To measure the amount of reducing ends produced after incubation of supernatant on polysaccharides, 10 µL of cultural supernatant was incubated with 2% (w/v) of either CMC or xylan (beechwood) in 200 µL of 50 mM sodium phosphate at 6.8 at 30 °C, before the Lever assay (71) was used to calculate reducing ends at set timepoints. Standard curves were generated using either glucose or xylose.

RNA extraction from *P. putredinis* NO1

RNA was extracted from cultures of *P. putredinis* NO1 that had been incubated at 30 °C with shaking at 180 rpm (70). Optimized growth media was used, supplemented with either 1.5 % wheat straw or 0.5 % glucose. At set time points, aliquots of either 0.5 g, 0.3 g or 0.1 g were taken, and 1 mL of Trizol (Life Technologies) and 3x3 mm tungsten carbide beads were added to disrupt cells in a TissueLyser II (Qiagen) for 2x2 min. Once the cells were disrupted the standard Trizol method was used to purify RNA as per manufacturer's instruction. DNA contamination

was removed with RTS DNase kits (Mobio) followed by treatment with ZymoResearch RNA Clean & Concentrator™ 5 kits. Ribo-Zero™ Magnetic Epidemiology rRNA removal kits (RZE1224/MRZ11124C; Illumina) were used to enrich for mRNA.

Sequencing was performed at The Genome Analysis Centre (TGAC), Norwich, U.K. Using the TruSeq RNA v2 protocol (Illumina 15026495 Rev.B). cDNA libraries were created before being normalized, pooled and diluted to a final concentration of 10 pM with 1% PhiX. Hybridization and first extension was performed using the TruSeq Rapid PE Cluster Kit v1 on the Illumina cBotTemplate, then transferred onto the Illumina HiSeq2500 for the remaining 100 cycles. Reads were trimmed to remove adaptor sequences with the ngsShoRT_2.1 method, and pooled libraries were assembled by Trinity Software and used as reference in subsequent analysis. The raw data was subject to rRNA removal by catching the remaining paired reads after mapping to a modified rRNA_115_tax_silva_v1.0 ribosomal set, using BOWTIE2. The original reads of the individual libraries were mapped to the reference using BWA software package with default parameters and the number of reads counted using SAMtools software package for each contig. Putative open readings frames (ORFs) were selected by translating regions over 300 bp between potential start and stop codons. BLASTp was used to perform searches against the non-redundant protein database, HMMER3 was used to search the Pfam and dbCAN databases (48, 72, 73), BLAST+ 2.3.0 (74) was used to perform local BLAST searches and SignalP 4.0 (75) was used to predict the presence of signal peptides.

Proteomic LC-MS/MS

Supernatant proteins were collected from the culture supernatant in 20 mL aliquots. These were precipitated overnight in five volumes of ice-cold acetone, pelleted through centrifugation at 10,000 xg, washed with 80 % ice-cold acetone and resuspended in 0.5 x PBS with 0.1% SDS. Two-gram samples of biomass were used to extract biomass-bound proteins, as described in Alessi *et al* (27). Both supernatant and biomass bound proteins were loaded into 4-12% (w/v) Bis-Tris acrylamide gels, separated with electrophoresis for 20 min and stained with InstantBlue (Sigma-Aldrich). Gel slices were analyzed as described previously (70). They were washed with 50% (v/v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, reduced with 10 mM DTE and S-carbamidomethylated with 50 mM iodoacetamide and dehydrated with acetonitrile. Digests were performed overnight with 0.2 µg trypsin (Promega) at 37 °C in 25 mM ammonium

460 bicarbonate. After extraction with 50% (v/v) aqueous acetonitrile, peptides were vacuum dried
461 and resuspended in 0.1% (v/v) aqueous trifluoroacetic acid. These were loaded onto a
462 nanoAcquity UPLC system (Waters) with a nanoAcquity Symmetry C18, 5 µm trap (180 µm x 20
463 mm Waters) and a nanoAcquity HSS T3 1.8 µm C18 capillary column (75 mm x 250 mm, Waters).
464 After washing the trap with 0.1% (v/v) aqueous formic acid (solvent A), a gradient solvent A and
465 acetonitrile containing 0.1% (v/v) formic acid (solvent B) was used to separate peptides. The
466 gradient proceeded from 2% to 30% of solvent B over 125 minutes linearly, then up to 50% over
467 five min, before rising to 95% solvent B for 2.5 min. Between each injection, the column was
468 equilibrated to the initial conditions for 25 min. The column temperature was 60 °C and the flow
469 rate was 300 nL min^{-1}. AutoMSMS mode was used to acquire positive ESI- MS & MS/MS spectra.
470 Compass 1.7 software (microTOF control, Hystar and DataAnalysis, Bruker Daltonics) was used
471 for instrument control with the following settings: ion spray voltage: 1,450 V; dry gas: 3 L min^{-1};
472 dry gas temperature 150 °C; collision RF: 1,400 Vpp; transfer time: 120 ms; ion acquisition
473 range: m/z 150-2,000. An absolute threshold of 200 counts and preferred charged state of 2-4
474 was specified. Singly charged ions were excluded. Cycle time: 1 s, MS spectra rate: 5 Hz, MS/MS
475 spectra rate: 5 Hz at 2,500 cts increasing to 20 Hz at 250,000 cts or above. Using the AutoMSMS
476 fragmentation table, collision energy and isolation width settings were automatically calculated.
477 For each precursor, a single MS/MS spectrum was acquired, and unless the precursor intensity
478 increased fourfold a dynamic exclusion for 0.8 min was applied.
479 The ORF library from *P. putredinis* NO1 was used as a reference to match the resultant spectra
480 against, using Mascot (Matrix Science Ltd., version 2.4) run through the Bruker ProteinScape
481 interface (version 2.1). Carbamindomethyl (C) and oxidation (M) were considered as fixed and
482 variable modifications, respectively. MS/MS tolerance was 0.1 Da and peptide tolerance was 10
483 ppm. A decoy database was used to attain a global false discovery rate of 1% using 'Mascot
484 Percolator' and peptides were adjusted to only be accepted with a expect score of 0.05 or
485 lower. Exponentially modified Protein Abundance Index (emPAI) was calculated as described by
486 Ishihama (76), which could be normalized to give molar percentage values by dividing against
487 the sum of all emPAI values for each sample. The R package BioStrings (77) was used to retrieve
488 each protein sequence from the ORF library.
489

Synthesis of synthetic substrate GGβ4MU (7-[2-hydroxy-2-(4-hydroxy-3-methoxyphenyl)-1-(hydroxymethyl)ethoxy]-4-methyl-2*H*-1-benzopyran-2-one).

The synthetic substrate GGβ4MU was synthesized in 6 steps according to the protocol reported by Weinstein and Gold starting from acetovanillone (47). The pure substrate GGβ4MU was obtained as a white solid following purification using plate chromatography on silica-gel (10% v/v MeOH in CH_2Cl_2). The NMR data were in excellent agreement with those previously reported.

Identification of β-etherase from native supernatant

P. putredinis NO1 was cultivated in medium containing 1.5% wheat straw. The supernatant was filtered through mirapore cloth and the protein of interest purified. Briefly, filtered culture supernatant with 0.1% Tween20 was concentrated in a 50 mL stirred Ultracentrifugation Cell (Millipore Corporation, USA) with a Biomax 30 kDa Ultrafiltration Membrane (Millipore Corporation, USA). Ammonium sulfate was slowly added to the filtered culture supernatant to a concentration of 20% while stirring at 4 °C. The solution was centrifuged at 10,000 g for 15 min. The pellet was then resuspended in 2 mL buffer A (50 mM Tris-HCl, 100 mM NaCl, 0.1% Tween 20, pH 8.5). Additional cuts were performed with 30, 40 and 50% ammonium sulfate. After assessing the fractions with the GGβ4MU assay, samples were purified on a Superdex-200 (GE Healthcare, US), using the ÄKTA system and 50 mM Tris-HCl, 100 mM NaCl, 0.1% Tween 20, pH 8.5. The most active sample was further purified using anion exchange chromatography. Anion-exchange chromatography was conducted on a DEAE FF column (GE Healthcare, US) with an increasing salt concentration from 0 to 1 M NaCl in 20 min (5 mL/min). A running buffer of 30 mM Tris-HCl, 0.1% Tween 20, at various pH (7.0/7.4/8.5) was used. The Elution buffer was 30 mM Tris-HCl, 1 M NaCl, 0.1% Tween 20.

Production of recombinant β-etherase

The c2092_g1_i1 gene was codon-optimized for expression in *E. coli* and synthesized into pET151 vector with N-terminal His-tag by Invitrogen. The expression plasmid was transformed into Arctic Express (DE3) competent cells. Auto-induction media was used for protein production. Inoculated cultures were incubated at 30 °C with shaking at 180 rpm until an optical density of 0.6 at 600 nm was reached. The temperature was then reduced to 11 °C for 48 h

519 before cell pellets were collected by centrifugation at 7000 rpm and 4 °C for 15 min.
520 Supernatant was discarded, and pellets were suspended in 5 mL per 100 mL of starting culture
521 20 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (HEPES) pH 8, before sonicated on
522 ice using a Misonix S-4000 sonicator at 70 kHz for 4 min, and a standard program of 3 s off
523 followed by 7 s on. The pellet was collected after centrifugation and washed with 20 mM HEPES,
524 2 M Urea, 0.5 M NaCl, 2% TritonTM X-100, pH 8, using the same volume as before, sonicated
525 and pelleted. The resultant pellet was then resuspended in 20 mM HEPES, 0.5 M NaCl, 5 mM
526 imidazole, 6 M guanidine hydrochloride, 1 mM dithiothreitol (DTT) pH 8, using 10 mL per 100 mL
527 of original cell culture, to solubilize inclusion bodies. After pelleting through centrifugation for a
528 final time, the supernatant was applied to a HisTrap column equilibrated with 20 mM HEPES, 0.5
529 M NaCl, 5 mM imidazole, 6 M guanidine hydrochloride, 1 mM DTT pH 8. The equilibration buffer
530 was then used to wash the column for a total of 5 CV followed by the same volume of 20 mM
531 HEPES, 0.5 M NaCl, 20 mM imidazole, 6 M urea, 1 mM DTT pH 8. A linear gradient from the final
532 wash buffer to 20 mM HEPES, 0.5 M NaCl, 20 mM imidazole, 0.1 mM $CuSO_4$, 1 mM DTT pH 8 was
533 then used to refold the tagged protein on the column. This was applied over 30 mL using a flow
534 rate of 0.5 ml/min. To elute refolded protein another linear gradient was applied over 20 mL,
535 starting with 20 mM HEPES, 0.3 M $MgCl_2$, 20 mM imidazole, 1 mM DTT, pH 8 and ending with
536 the same buffer with the addition of 500 mM imidazole and 10% glycerol. Apart from when
537 otherwise mentioned, the flow rate was kept at 1 mL/ min when using a 1 mL capacity column
538 and 3 mL/min when using a 5 mL capacity column. Fractions of 1.5 mL were collected
539 throughout the elution step, and UV absorbance was used to determine protein content.
540 Fractions with high protein contents were visualised using SDS–polyacrylamide gel
541 electrophoresis (SDS-PAGE) and the presence of the recombinant protein confirmed through
542 western blot analysis. Protein activity was confirmed through the measurement of 4MU from
543 the GGβ4MU assay after removal of imidazole and DTT using Zeba™ Spin Desalting Columns, 7K
544 MWCO (ThermoFisher) or Slide-A-Lyzer™ Dialysis Cassettes 10 K MWCO (ThermoFisher).

545 A second method to produce purified recombinant protein was also employed whereby cell
546 pellets were suspended in 50 mL (50 mM Tris, 1 mm DTT, pH 8.5) and sonicated on ice. After
547 centrifugation at 17,000 rpm for 45 min, supernatant was loaded onto an anion exchange
548 chromatography mono-Q- HP column (5 mL, GE Healthcare) equilibrated with 50 mm Tris, 100

549 mm NaCl, 10% glycerol pH 8. The protein was then eluted with an increasing NaCl gradient (0 to
550 1 M) over 100 mL. Fractions containing the protein of interest were pooled and concentrated
551 using Millipore Vivaspin20 10kDa (Sartorius). These were then injected into a superdex 75
552 (16/60) gel-filtration column (GE Healthcare) equilibrated with 50 mM Tris, 150 mM NaCl, 10%
553 glycerol pH 8.5. SDS-PAGE was used to determine the purity of the eluted protein. Throughout
554 the purification β-etherase activity was calculated (SI Appendix, Table S3). Concentration was
555 calculated spectroscopically using the extinction coefficient at 280 nm.

556 **Fluorescence assay for β-etherase activity**

557 Enzyme activity was measured in 1 mL reaction containing 10 µL 4MU/GGβ4MU (synthetic
558 fluorescent substrate 10 mM) and appropriate concentration of pure protein in 50 mM Tris-HCl,
559 100 mM NaCl, pH 8.5, 5 mM $CuSO_4$. The reaction was incubated at 30 °C for 1 h. Formation of 4-
560 methylumbelliferone (4MU) was monitored using an RF-1500 fluorometric analyzer. After 0 h
561 and 1 h of incubation 100 µL of the reaction mixture was taken and added to 50 µL of 100 mM
562 glycine-NaOH buffer (pH 10.0). One unit of the enzyme was defined as the amount that released
563 1 nmol of 4 MU/h from the substrate. Five replicates were taken for each sample, and control
564 reactions of boiled enzyme and wheat straw treated with buffer only were also performed. The
565 oxidative nature of the protein was investigated by performing this assay in an anaerobic
566 chamber. All buffers were deoxygenated within the chamber overnight prior to experimental
567 set-up. These assays were performed on pure protein.

568 **Enzyme properties**

569 The effect of pH and temperature on enzyme activity was investigated on protein purified by
570 anion-exchange and size-exclusion chromatography, by varying the pH of the reaction mixtures
571 using 50 mM Tris-HCl buffer from pH 7.0 to 9.5, 50 mM glycine-NaOH buffer at pH range 9.0 to
572 10.5 and 50 mM Na_2HPO_4-NaOH buffer at pH range 10.5 to 12. The optimum temperature of
573 enzyme activity was determined at various temperatures ranging from 20 °C to 70 °C. Assays
574 were performed as described in the previous section. Specificity was investigated by incubating
575 1 mM of each substrate of interest with the enzyme in 100 µL Tris pH 8.5 buffer at room
576 temperature. Activity was determined by monitoring the change in Ultraviolet-Visible

absorbance spectra (220 - 750 nm) of aliquots using a NanoDrop 8000 Microvolume UV-Vis spectrophotometer (Thermo Scientific). Scans were performed at regular intervals over 2 h.

Extraction of tricin

Wheat straw was ground to <1 mm using a cyclone mill (Retsch) and washed several times with 50 mM Tris pH 8 to remove surface sugars. In 1 mL reactions, 100 mg of washed wheat straw was incubated with purified enzyme in 50 mM Tris buffer at pH 8 with 5 mM $CuSO_4$. Reactions were incubated overnight at 30 °C with shaking. Control reactions were performed using wheat straw incubated with buffer only. Tricin was extracted based upon the method described by Karambelkar (77). Briefly, 1 mL of ethyl acetate was added to 100 µL of the reaction supernatant. This was homogenized before being centrifuged for 5 min at 13,000 rpm. The ethyl acetate layer was transferred into new tubes and evaporated using a centrifugal evaporator at 55 °C before being resuspended in 100 µL 50% H_2O, 50% acetonitrile. This was analyzed with a Waters 2996 photodiode array detector Separations Module HPLC system, with a C18-5 µM preparative column (4.6 × 250 mm, Waters). The mobile phase consisted of 0.1% acetic acid in H_2O (A), and methanol (B). The following linear gradient was used; 95% A (5 min), 70% A (25 min), 0% A (30 min), 95% A (5 min), the flow rate was 1.0 mL/min. After identification through comparisons with authentic standards, based on retention time and UV spectrum, peaks were manually collected and the mass confirmed with mass spectroscopy.

Oxidase activity boosting saccharification with cellulase enzymes

For saccharification reactions, biomass pretreated with the oxidase was incubated with 1.2 µg/mL enzyme cocktail (4:1 Celluclast: novo 188 (Novozymes)) in 50 mM sodium acetate at pH 4.5 and incubated overnight at 37–40 °C with shaking. This was performed alongside a control reaction with buffer only. Solids were removed by centrifugation, and residual protein was precipitated with 80% ethanol. The supernatant, containing mono- and oligosaccharides, was dried with a centrifugal evaporator before samples were resuspended in ultra-pure water and filtered through a 0.2 µm polytetrafluoroethylene (PTFE) filter.

High-performance anion-exchange chromatography (HPAEC)

HPAEC was used to analyze monosaccharide release after saccharification. Briefly, 5 µL of samples or standards were injected on a CarboPac PA20 3 × 150 mm analytical column via a CarboPac PA20 3×0 mm guard column using Chromeleon 6.8 Chromatography Data Systems software (Dionex). Sugars were separated at a flow rate of 0.4-0.5 mL min^{-1} as follows: a linear gradient of 100% H_2O to 99%–1% of H_2O–0.2 M NaOH for 5 min, then constant for 10 min, followed by a 7 min linear gradient to 47.5%:22.5%:30% of H_2O:0.2 M NaOH:0.5 M NaOAc/0.1 M NaOH then kept constant for 15 min. Between injections the column was washed with 0.2 M NaOH for 8 min and re-equilibrated with 100% H_2O for 10 min. Carbohydrates were detected by ICS-3000 PAD system with an electrochemical gold electrode, identified by comparison with retention times of external standards, and quantified through the integration of those known standards.

Lignin isolation

Enzyme lignins, representing essentially all of the lignin in the sample, were prepared following ball-milling of the cell wall isolate as previously described (56-58, 78).

NMR analysis

2D NMR of enzyme lignins (EL) in 4:1 v/v DMSO-d_6:pyridine-d_5 were acquired on a Bruker Biospin (Billerica, MA) Avance 700 MHz spectrometer equipped with a 5-mm $^{1}H/^{31}P/^{13}C/^{15}N$ QCI cryoprobe with inverse geometry (proton coils closest to the sample), as described previously (57, 58). Volume-integration of contours in HSQC plots used TopSpin 4.07 (Mac version) software, and no correction factors were used. The data represent volume-integrals only, and data are presented on an S + G + H = 100% or an **A** + **B** + **C** + **C´** basis, although the latter are also provided on a 'number of linkages per 100 aromatic rings' basis also (Fig. 5); *p*CA, and tricin T units are always terminal and are, therefore, likely overestimated (58). Data assignments here were made by comparison with published data from other samples from our lab, including in the various tricin-related papers (7-10, 79, 80).

Statistical Analysis

Where mentioned two tail ANOVAs were performed using R core package "stats" (81). Differential gene analysis was performed using generalized linear models from R package "edgeR".

Data availability

Metaproteomic and metatranscriptomic databases generated during this research are available at MassIVE MSV000084758 and ProteomeXchange PXD016952. A curated dataset, which includes annotations, is available in the Dataset S1. Amplicon sequences are deposited at the European Nucleotide Archive under the accession PRJEB38167.

Acknowledgments

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), UK (Grant BB/1018492/1 and BB/P027717/1), the BBSRC Network in Biotechnology and Bioenergy LBNET. N.C.O and D.R.L. were supported by studentships from the BBSRC Doctoral Training Programme (BB/J014443/1). A.A. acknowledges a Royal Society International Newton Fellowship. J.R., S.L and V. T. acknowledge funding by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science, DE-SC0018409). The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/ K039660/1; EP/M028127/1).

References

1. L. Caspeta, N. A. A. Buijs, J. Nielsen, The role of biofuels in the future energy supply. *Energy Environ. Sci.* **6**, 1077-1082 (2013).
2. J. V. Vermaas *et al.*, Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnol. Biofuels* **8**, 16 (2015).
3. S.-Y. Ding *et al.*, How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* **338**, 1055-1060 (2012).
4. Y. Ogo, K. Ozawa, T. Ishimaru, T. Murayama, F. Takaiwa, Transgenic rice seed synthesizing diverse flavonoids at high levels: a new platform for flavonoid production with associated health benefits. *Plant Biotechnol. J.* **11**, 734-746 (2013).
5. J. M. Duarte-Almeida, G. Negri, A. Salatino, J. E. de Carvalho, F. M. Lajolo, Antiproliferative and antioxidant activities of a tricin acylated glycoside from sugarcane (*Saccharum officinarum*) juice. *Phytochemistry* **68**, 1165-1171 (2007).

6. J. C. del Río *et al.*, Lignin monomers from beyond the canonical monolignol biosynthetic pathway: Another brick in the wall. *ACS Sustainable Chemistry & Engineering* **8**, 4997-5012 (2020).
7. J. C. del Río *et al.*, Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *Journal of Agricultural and Food Chemistry* **60**, 5922-5935 (2012).
8. W. Lan *et al.*, Tricin-lignins: occurrence and quantitation of tricin in relation to phylogeny. *Plant J.* **88**, 1046-1057 (2016).
9. W. Lan *et al.*, Tricin, a flavonoid monomer in monocot lignification. *Plant Physiol.* **167**, 1284-U1265 (2015).
10. W. Lan *et al.*, Maize tricin-oligolignol metabolites and their implications for monocot lignification. *Plant Physiol.* **171**, 810-820 (2016).
11. M. Li, Y. Pu, C. G. Yoo, A. Ragauskas, The occurrence of tricin and its derivatives in plants. *Green Chem.* **18** (2016).
12. G. M. Rashid *et al.*, Identification of manganese superoxide dismutase from *Sphingobacterium* sp. T2 as a novel bacterial enzyme for lignin oxidation. *ACS Chem. Biol.* **10**, 2286-2294 (2015).
13. K. Min, G. Gong, H. M. Woo, Y. Kim, Y. Um, A dye-decolorizing peroxidase from *Bacillus subtilis* exhibiting substrate-dependent optimum temperature for dyes and β-ether lignin dimer. *Sci. Rep.* **5**, 8245 (2015).
14. G. Janusz *et al.*, Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. *FEMS Microbiol. Rev.* **41**, 941-962 (2017).
15. F. J. Ruiz-Dueñas, Á. T. Martínez, Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. *Microb. Biotechnol.* **2**, 164-177 (2009).
16. Z. Duan, R. Shen, B. Liu, M. Yao, R. Jia, Comprehensive investigation of a dye-decolorizing peroxidase and a manganese peroxidase from *Irpex lacteus* F17, a lignin-degrading basidiomycete. *AMB Express* **8**, 119 (2018).
17. B. Goodell *et al.*, Modification of the nanostructure of lignocellulose cell walls via a non-enzymatic lignocellulose deconstruction system in brown rot wood-decay fungi. *Biotechnol. Biofuels* **10**, 179 (2017).
18. B. Goodell *et al.*, Low molecular weight chelators and phenolic compounds isolated from wood decay fungi and their role in the fungal biodegradation of wood. *J. Biotechnol.* **53**, 133-162 (1997).
19. V. Arantes, B. Goodell, *Current understanding of brown-rot fungal biodegradation mechanisms: A review*, Deterioration and Protection of Sustainable Biomaterials (American Chemical Society, 2014), vol. 1158, pp. 3-21.
20. R. Cohen, M. R. Suzuki, K. E. Hammel, Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot basidiomycete *Gloeophyllum trabeum*. *Appl. Microbiol. Biotechnol.* **71**, 2412-2417 (2005).
21. D. J. Yelle, J. Ralph, F. Lu, K. E. Hammel, Evidence for cleavage of lignin by a brown rot basidiomycete. *Environ. Microbiol.* **10**, 1844-1849 (2008).
22. F. Barras, F. Van Gijsegem, A. K. Chatterjee, Extracellular enzymes and pathogenesis of soft-rot *Erwinia*. *Annu. Rev. Phytopathol.* **32**, 201-234 (2003).
23. G. van Erven *et al.*, Evidence for ligninolytic activity of the ascomycete fungus *Podospora anserina*. *Biotechnol. Biofuels* **13**, 75 (2020).

709 24. U. Kõljalg *et al.*, UNITE: a database providing web-based methods for the molecular
710 identification of ectomycorrhizal fungi. *New Phytol.* **166**, 1063-1068 (2005).
711 25. K. Abarenkov *et al.*, The UNITE database for molecular identification of fungi – recent
712 updates and future perspectives. *New Phytol.* **186**, 281-285 (2010).
713 26. A. M. Alessi *et al.*, Defining functional diversity for lignocellulose degradation in a
714 microbial community using multi-omics studies. *Biotechnol. Biofuels* **11**, 166 (2018).
715 27. A. M. Alessi *et al.*, Revealing the insoluble metasecretome of lignocellulose-degrading
716 microbial communities. *Sci. Rep.* **7**, 2356 (2017).
717 28. H. Jun, H. Guangye, C. Daiwen, Insights into enzyme secretion by filamentous fungi:
718 Comparative proteome analysis of *Trichoderma reesei* grown on different carbon sources.
719 *J. Proteomics* **89**, 191-201 (2013).
720 29. N. L. Glass, M. Schmoll, J. H. D. Cate, S. Coradetti, Plant cell wall deconstruction by
721 ascomycete fungi. *Annu. Rev. Microbiol.* **67**, 477-498 (2013).
722 30. D. M. de Oliveira *et al.*, Ferulic acid: a key component in grass lignocellulose recalcitrance
723 to hydrolysis. *Plant Biotechnol. J.* **13**, 1224-1232 (2015).
724 31. J. Zhang, M. Siika-Aho, M. Tenkanen, L. Viikari, The role of acetyl xylan esterase in the
725 solubilization of xylan and enzymatic hydrolysis of wheat straw and giant reed. *Biotechnol.*
726 *Biofuels* **4**, 60 (2011).
727 32. A. Levasseur, E. Drula, V. Lombard, P. M. Coutinho, B. Henrissat, Expansion of the
728 enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes.
729 *Biotechnol. Biofuels* **6**, 41 (2013).
730 33. G. Vaaje-Kolstad *et al.*, An oxidative enzyme boosting the enzymatic conversion of
731 recalcitrant polysaccharides. *Science* **330**, 219-222 (2010).
732 34. R. J. Quinlan *et al.*, Insights into the oxidative degradation of cellulose by a copper
733 metalloenzyme that exploits biomass components. *PNAS* **108**, 15079-15084 (2011).
734 35. C. Laurent, E. Breslmayr, D. Tunega, R. Ludwig, C. Oostenbrink, Interaction between
735 cellobiose dehydrogenase and lytic polysaccharide monooxygenase. *Biochemistry* **58**,
736 1226-1235 (2019).
737 36. T.-C. Tan *et al.*, Structural basis for cellobiose dehydrogenase action during oxidative
738 cellulose degradation. *Nat. Commun.* **6**, 7542 (2015).
739 37. L. Munk, A. K. Sitarz, D. C. Kalyani, J. D. Mikkelsen, A. S. Meyer, Can laccases catalyze bond
740 cleavage in lignin? *Biotechnology Advances* **33**, 13-24 (2015).
741 38. R. Hilgers, J.-P. Vincken, H. Gruppen, M. A. Kabel, Laccase/Mediator Systems: Their
742 Reactivity toward Phenolic Lignin Structures. *ACS Sustainable Chemistry & Engineering* **6**,
743 2037-2046 (2018).
744 39. A. Levasseur *et al.*, Exploring laccase-like multicopper oxidase genes from the ascomycete
745 *Trichoderma reesei*: a functional, phylogenetic and evolutionary study. *BMC Biochemistry*
746 **11** (2010).
747 40. P. Kersten, D. Cullen, Extracellular oxidative systems of the lignin-degrading
748 basidiomycete *Phanerochaete chrysosporium*. *Fungal Genet. Biol.* **44**, 77-87 (2007).
749 41. P. Daly *et al.*, Expression of *Aspergillus niger* CAZymes is determined by compositional
750 changes in wheat straw generated by hydrothermal or ionic liquid pretreatments.
751 *Biotechnol. Biofuels* **10**, 35 (2017).
752 42. W. Schutyser *et al.*, Chemicals from lignin: an interplay of lignocellulose fractionation,
753 depolymerisation, and upgrading. *Chem. Soc. Rev.* **47**, 852-908 (2018).

43. D. L. Gall *et al.*, *In vitro* enzymatic depolymerization of lignin with release of syringyl, guaiacyl, and tricin units. *Appl. Environ. Microbiol.* **84**, e02076-02017 (2018).
44. W. S. Kontur *et al.*, A heterodimeric glutathione S-transferase that stereospecifically breaks lignin's β(R)-aryl ether bond reveals the diversity of bacterial β-etherases. *J. Biol. Chem.* **294**, 1877-1890 (2019).
45. M. Marinović *et al.*, Selective cleavage of lignin β-O-4 aryl ether bond by β-etherase of the white-rot fungus *Dichomitus squalens*. *ACS Sustainable Chemistry & Engineering* **6**, 2878-2882 (2018).
46. Y. Otsuka *et al.*, Detection and characterization of a novel extracellular fungal enzyme that catalyzes the specific and hydrolytic cleavage of lignin guaiacylglycerol β-aryl ether linkages. *Eur. J. Biochem.* **270**, 2353-2362 (2003).
47. D. A. Weinstein, M. H. Gold, Synthesis of guaiacylglycol and glycerol-β-O-(β-methylumbelliferyl) ethers: lignin model substrates for the possible fluorometric assay of β-etherases. *Holzforschung* **33**, 134-135 (1979).
48. R. D. Finn *et al.*, The Pfam protein families database. *Nucleic Acids Res.* **38**, D211-D222 (2010).
49. M. Kanteev, M. Goldfeder, A. Fishman, Structure-function correlations in tyrosinases. *Protein Sci.* **24**, 1360-1369 (2015).
50. K. Min *et al.*, Perspectives for biocatalytic lignin utilization: cleaving 4-O-5 and Cα–Cβ bonds in dimeric lignin model compounds catalyzed by a promiscuous activity of tyrosinase. *Biotechnol. Biofuels* **10**, 212 (2017).
51. S. Halaouli, M. Asther, J. C. Sigoillot, M. Hamdi, A. Lomascolo, Fungal tyrosinases: new prospects in molecular characteristics, bioengineering and biotechnological applications. *J. Appl. Microbiol.* **100**, 219-232 (2006).
52. N. Hakulinen, C. Gasparetti, H. Kaljunen, K. Kruus, J. Rouvinen, The crystal structure of an extracellular catechol oxidase from the ascomycete fungus Aspergillus oryzae. *J. Biol. Inorg. Chem.* **18**, 917-929 (2013).
53. C. Gasparetti *et al.*, Discovery of a new tyrosinase-like enzyme family lacking a C-terminally processed domain: production and characterization of an *Aspergillus oryzae* catechol oxidase. *Applied Microbiology and Biotechnology* **86**, 213-226 (2010).
54. Z. Yang, D. A. Robb, Comparison of tyrosinase activity and stability in aqueous and nearly nonaqueous environments. *Enzyme Microb. Technol.* **15**, 1030-1036 (1993).
55. S. Zolghadri *et al.*, A comprehensive review on tyrosinase inhibitors. *J. Enzyme Inhib.* **34**, 279-309 (2019).
56. H.-M. Chang, E. B. Cowling, W. Brown, E. Adler, G. Miksche, Comparative studies on cellulolytic enzyme lignin and milled wood lignin of sweetgum and spruce. *Holzforschung* **29**, 153-159 (1975).
57. H. Kim, J. Ralph, Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6/pyridine-d_5. *Org Biomol Chem* **8**, 576-591 (2010).
58. S. D. Mansfield, H. Kim, F. Lu, J. Ralph, Whole plant cell wall characterization using solution-state 2D-NMR. *Nature protocols* **7**, 1579-1589 (2012).
59. Y. Mu, L. Li, S.-Q. Hu, Molecular inhibitory mechanism of tricin on tyrosinase. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **107**, 235-240 (2013).
60. J. Y. Lim, K. Ishiguro, I. Kubo, Tyrosinase inhibitory *p*-coumaric acid from ginseng leaves. *Phytother. Res.* **13**, 371-375 (1999).

61. G. van Erven *et al.*, Mechanistic insight in the selective delignification of wheat straw by three white-rot fungal species through quantitative ^{13}C-IS py-GC–MS and whole cell wall HSQC NMR. *Biotechnol. Biofuels* **11**, 262 (2018).
62. S. Hutner, L. Provasoli, A. Schatz, C. Haskins, Some approaches to the study of the role of metals in the metabolism of microorganisms. *Proceedings of the American Philosophical Society* **94**, 152-170 (1950).
63. White, T. Bruns, S. Lee, J. Taylor, "Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics" in PCR Protocols: A guide to methods and applications. (1990), pp. 315-322.
64. A. Klindworth *et al.*, Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**, e1 (2013).
65. J. G. Caporaso *et al.*, QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335-336 (2010).
66. D. McDonald *et al.*, An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME* **6**, 610-618 (2012).
67. T. Z. DeSantis *et al.*, Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069-5072 (2006).
68. K. Abarenkov *et al.*, The UNITE database for molecular identification of fungi-recent updates and future perspectives. *New Phytol.* **186**, 281-285 (2010).
69. M. A. Bezerra, R. E. Santelli, E. P. Oliveira, L. S. Villar, L. A. Escaleira, Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta* **76**, 965-977 (2008).
70. N. C. Oates, "Mining microbial compost communities for lignocellulose degrading proteins", University of York, York, U.K, (2016).
71. M. Lever, Colorimetric and fluorometric carbohydrate determination with *p*-hydroxybenzoic acid hydrazide. *Biochemical Medicine* **7**, 274-281 (1973).
72. H. Zhang *et al.*, dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95-W101 (2018).
73. R. D. Finn, J. Clements, S. R. Eddy, HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29-W37 (2011).
74. C. Camacho *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
75. T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature methods* **8**, 785-786 (2011).
76. Y. Ishihama *et al.*, Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* **4**, 1265-1272 (2005).
77. P. Karambelkar, Jadhav, V.M. , Kadam, V., Isolation and characterization of flavonoid tricin from sugarcane sludge. *Indo American Journal of Pharmaceutical Research* **4**, 7 (2014).
78. H. Kim *et al.*, Monolignol benzoates incorporate into the lignin of transgenic *Populus trichocarpa* depleted in C3H and C4H. *ACS Sustainable Chemistry & Engineering* **8**, 3644-3654 (2020).
79. J. Rencoret *et al.*, Structural characterization of lignin isolated from coconut (*Cocos nucifera*) Coir Fibers. *J. Agric. Food Chem.* **61**, 2434-2445 (2013).

80. W. Lan *et al.*, Elucidating tricin-lignin structures: Assigning correlations in HSQC spectra of monocot lignins. *Polymers (Basel)* **10**, 916 (2018).
81. R. C. Team (2019) R: A Language and Environment for Statistical Computing. (Vienna, Austria).

Figures and Tables

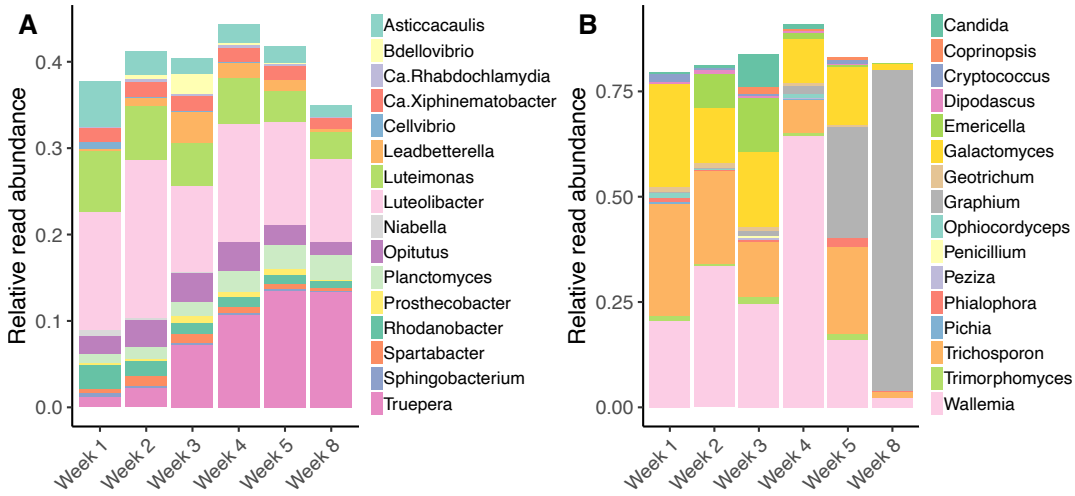

Figure 1. Composition of prokaryotic and fungal genera during wheat straw degradation. Sequences were generated on an ion torrent platform after amplification of the 16S and ITS for (*A*) prokaryotic and (*B*) eukaryotic identification, respectively. Operational taxonomic units were identified to genus level N=1.

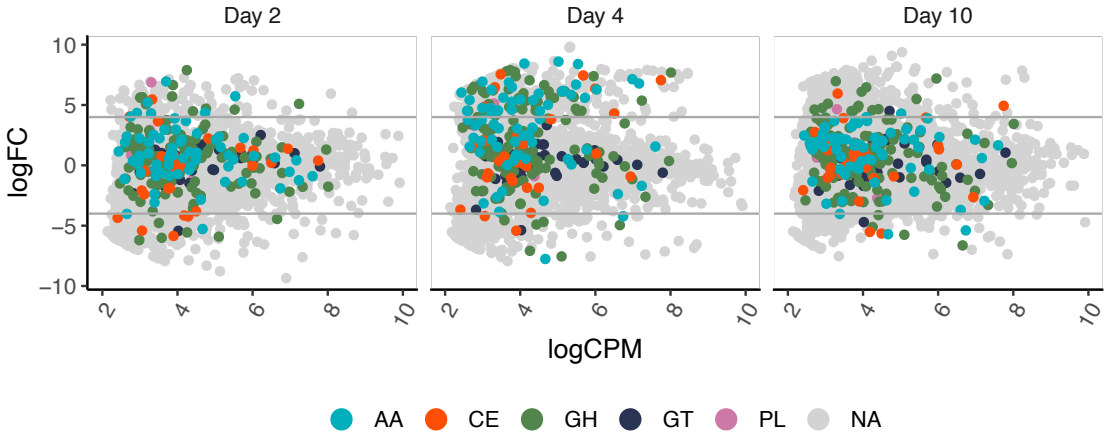

Figure 2. Expression change of contigs between glucose and wheat straw conditions. RNA was extracted and sequenced after two, four and ten days of *P. putredinis* NO1 incubation on wheat straw and four days of growth on glucose. Points represent the log fold change (FC) and average counts per million (CPM) of contigs, between the wheat straw and glucose conditions. Carbohydrate-active enzymes were annotated using dbCAN namely auxiliary activities (AA), glycoside hydrolases (GH), polysaccharide lyases (PL), carbohydrate esterases (CE) and glycosyltransferases (GT). Points are the average of three biological replicates.

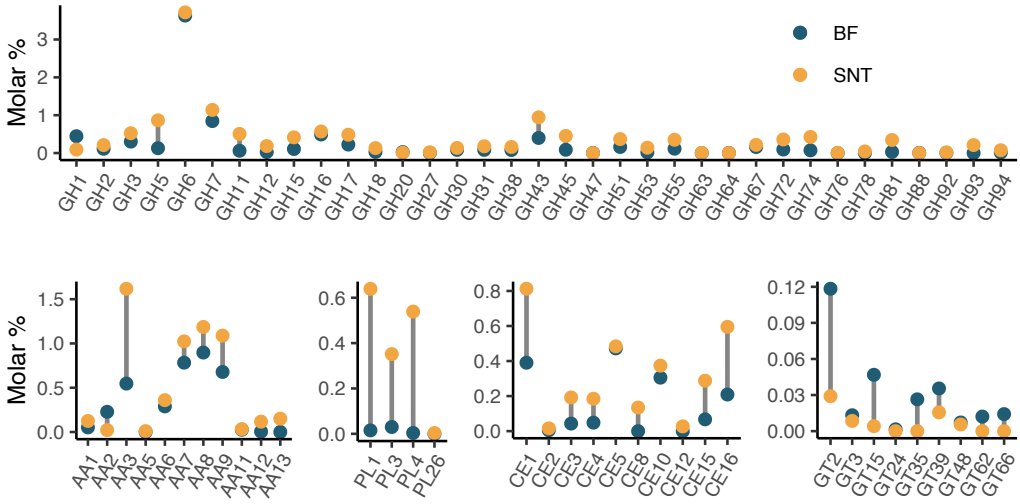

Figure 3. Molar percentages of supernatant (SNT) and biotin-labelled (BF) proteins after four days of incubation on wheat straw. Molar percentages of carbohydrate-active families, GH: Glycoside hydrolase, AA: Auxiliary activity, PL: Polysaccharide lyase, CE: Carbohydrate esterase and GT Glycosyl transferase, were calculated as the sum of contigs annotated and taken as an average for each biological replicate.

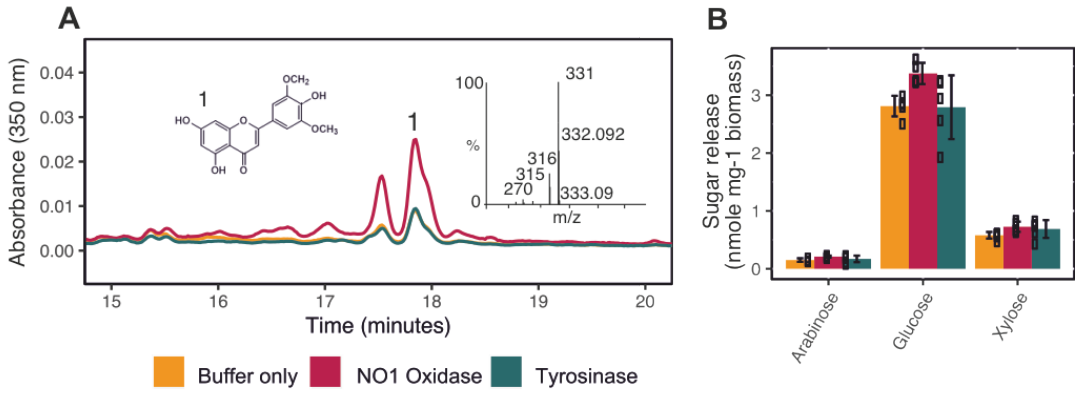

Figure 4. Treatment of Biomass with recombinant oxidase. (*A*) Tricin **1** release from wheat straw was observed and compared to an authentic standard using a High-Performance Liquid-Chromatography (HPLC), and mass was confirmed by time-of-flight mass spectrometry. (*B*) wheat straw was treated with recombinant oxidase, commercial mushroom tyrosinase, and buffer only for 16 h prior to the application of Celluclast® commercial saccharification cocktail. Sugar release was calculated from the reaction mixture using High-Performance Anion-Exchange chromatography. Error bars represent the standard deviation of five biological replicates.

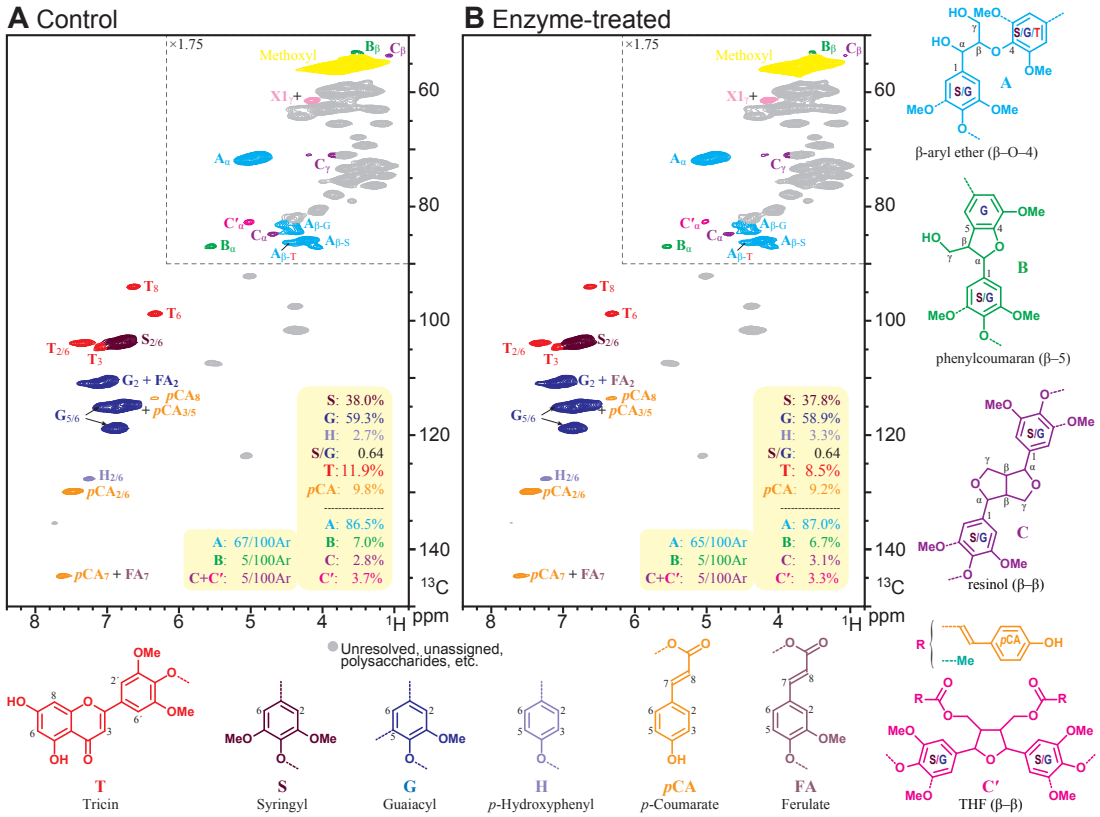

Figure 5. Lignin aromatic and side-chain region of 2D HSQC NMR spectra (DMSO-d_6:pyridine-d_5, 4:1, v/v) of enzyme lignins (EL) from (A) the wheat control, and (B) the enzyme-treated wheat. The quantification values in the yellow boxes are for relative comparisons of the lignin components determined from NMR contour volume-integrals based on **S** + **G** + **H** = 100% or an A + B + C + C′ basis, although the latter are also provided on a 'number of linkages per 100 aromatic rings' basis also. The ***p*CA** and **T** units are lignin appendages; their levels were estimated and expressed based on the total lignin (**S** + **G** + **H**). The chemical structures of the lignin monomeric subunits are color-coded to match their signal assignments in the spectra. Assignments are from papers noted in the Experimental Section, along with the new $A_{\beta\text{-}T}$ assignment (80). Note that, to allow the crucial lignin side-chain contours to be more clearly seen, the boxed lignin side-chain region was vertically scaled by ~1.75×.

PNAS
www.pnas.org

Supplementary information for

A multi-omics approach to lignocellulolytic enzyme discovery: uncovering a new ligninase activity from *Parascedosporium putredinis* NO1

Nicola C. Oates[1], Amira Abood[1], Alexandra M. Schirmacher[1], Anna M. Alessi[1], Susannah M. Bird[1], Joseph P. Bennett[1], Daniel R. Leadbeater[1], Yi Li[1], Adam A. Dowle[2], Vitaliy I. Tymokhin[3], John Ralph[3], Simon J. McQueen-Mason[1], Neil C. Bruce[1]

[1]Centre for Novel Agricultural Products, Department of Biology, University of York, York, YO10 5DD, UK.

[2]Bioscience Technology Facility, Department of Biology, University of York, York, YO10 5DD, UK.

[3]Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, 53706, USA

Correspondence and requests for materials should be addressed to N.C.B.

Email: neil.bruce@york.ac.uk

This PDF file includes:

Figures S1 to S12

Table S1 to S4

Other supplementary materials for this manuscript include the following:

Dataset S1

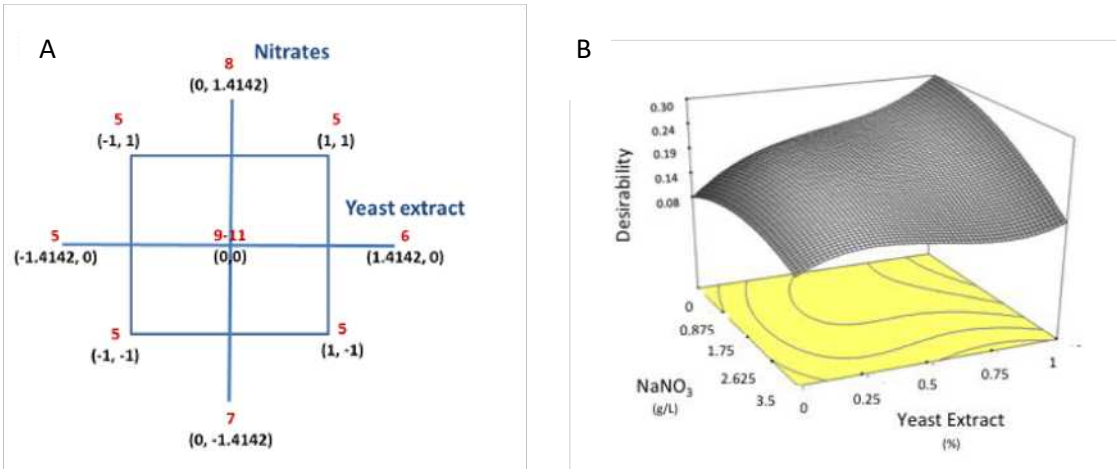

Fig. S1. Optimisation of *P. putredinis* NO1 growth media. (*A*) central composite design was used to create a response surface morphology to yeast extract and sodium nitrate concentrations. (*B*) Both cellulase and xylanase production was improved with a high yeast extract and low nitrate concentrations.

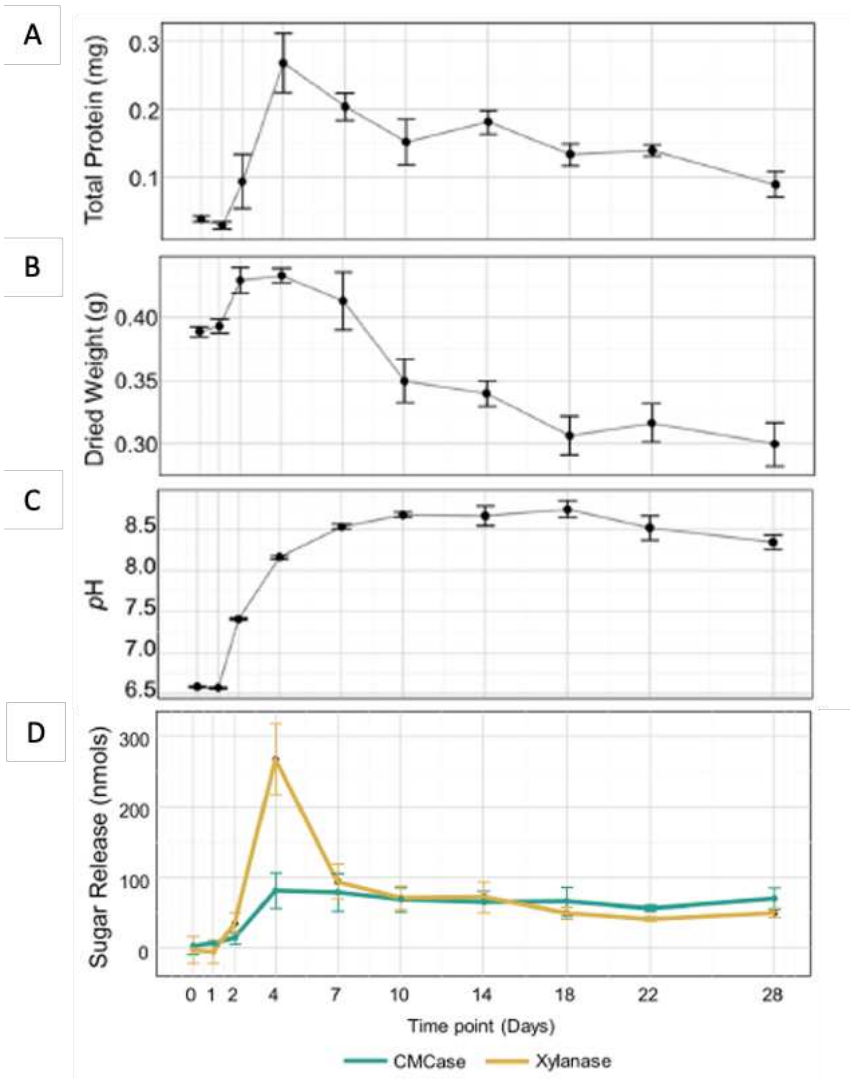

Fig. S2. Growth of *P. putredinis* NO1 on wheat straw over a period of one month. Growth of *P. putredinis* NO1 on wheat straw estimated by (*A*) the total protein present in the culture and (*B*) the dried weight of the total biomass within the culture. (*C*) The pH of the culture was also monitored alongside (*D*) the release of sugar after 1 h from 10 % supernatant loading on carboxymethylcellulose and beechwood xylan.

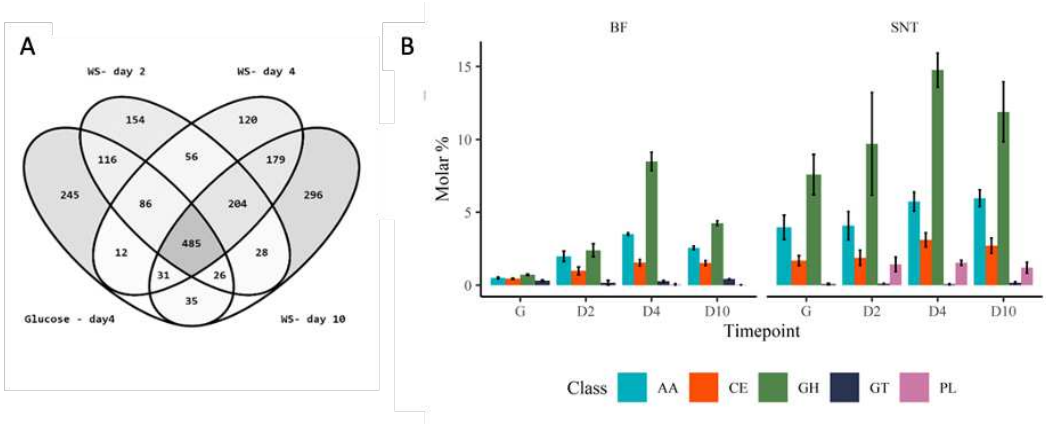

Fig. S3. Proteomics of *P. putredinis* NO1 grown on wheat straw. (*A*) Total proteins recovered from *P. putredinis* NO1 exoproteome across timepoints. (*B*) Total molar percentage of CAZy class across timepoints in the biotin labelled protein sample and supernatant.

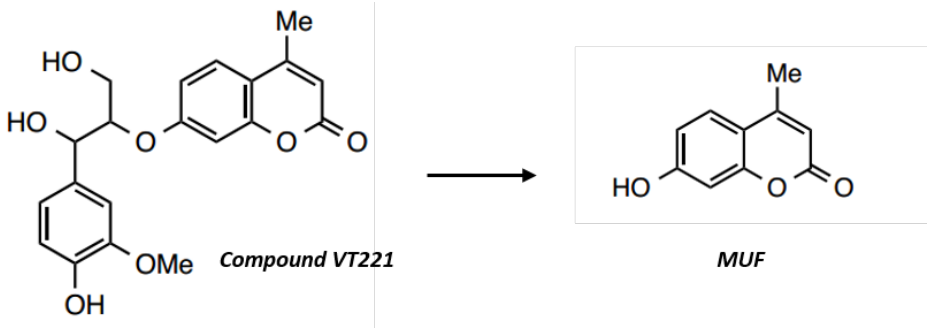

Fig. S4. GGβ4MU β-etherase assay. Under the action of a β-etherase the 4-O-β-ether linkage is cleaved liberating the product MUF. Upon excitement at 372 nm MUF will fluoresce at 445 nm.

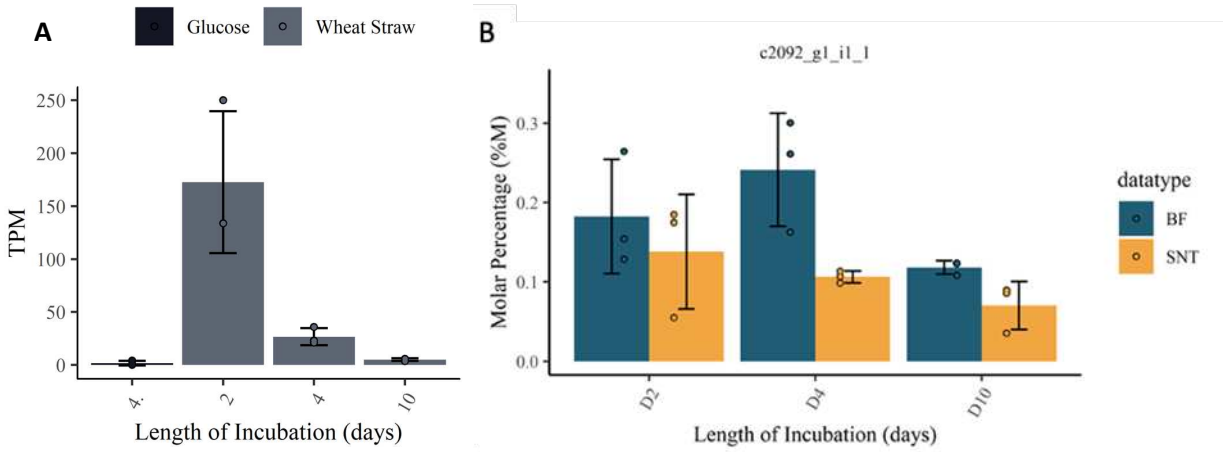

Fig. S5. c2092_g1_i1 abundance within the transcriptomic and proteomic libraries. (*A*) Transcript per million (TPM) of c2092_g1_i1 after growth on wheat straw and glucose and (*B*) molar percent abundance. Circles represent sample values of biological replicates (N=3), and error bars ± SD of the mean.

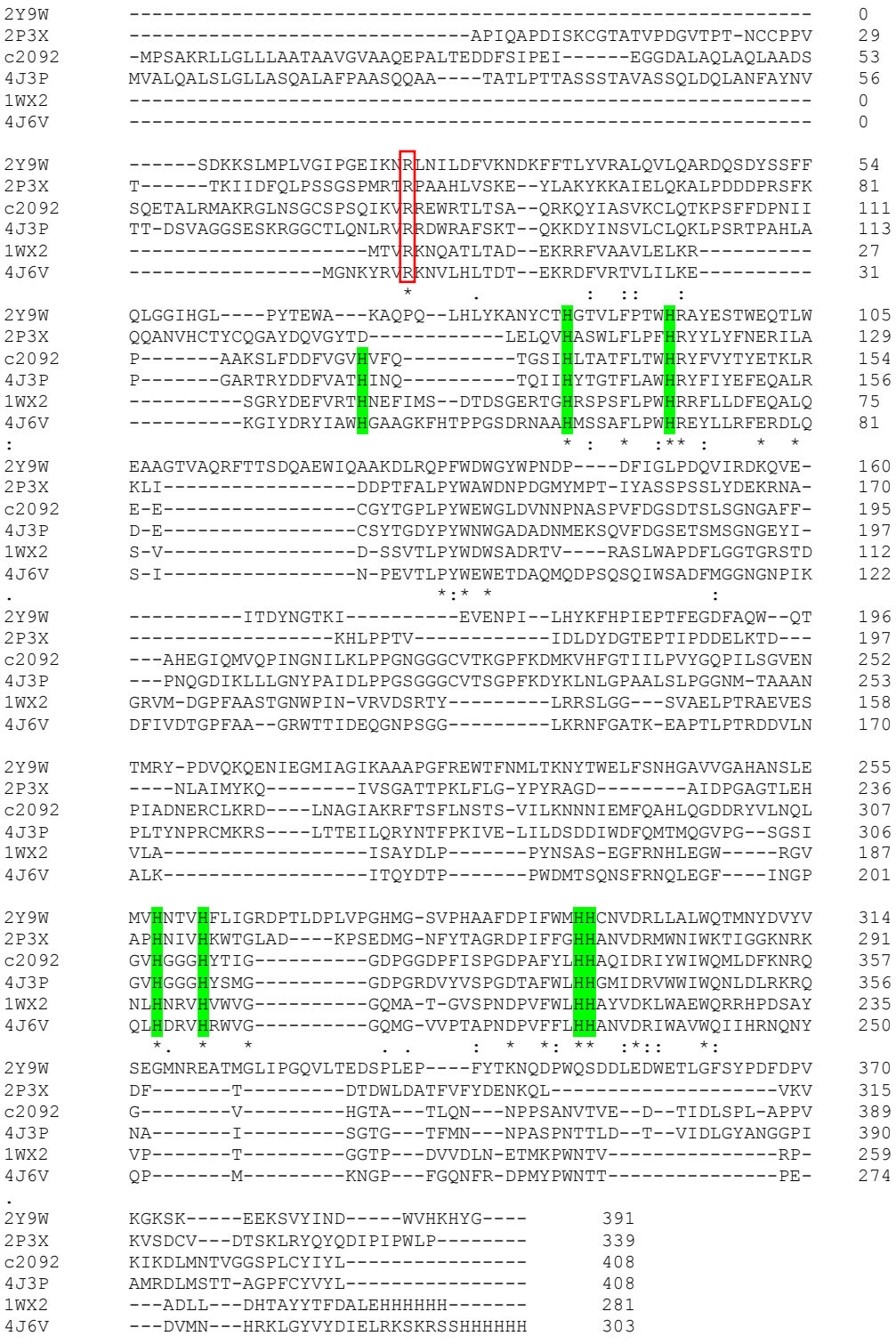

Fig. S6. Alignment of the putative β-etherase amino acid sequence (c2092_g1_i1) with structurally related enzymes. Alignment with 2Y9W; tyrosinase from *Agaricus bisporus* (common mushroom), 2P3X; *Vitis vinifera* Polyphenol Oxidase, 4J3P; catechol oxidase *Aspergillus oryzae*, 1WX2; *Streptomyces castaneoglobisporus* tyrosinase, 4J6V; *Bacillus megaterium* N205D tyrosinase. Identical amino acids are indicated by asterisks and amino acids similarity by dots. The conserved N-terminal arginine residue is highlighted with a red box; copper-binding regions are in green.

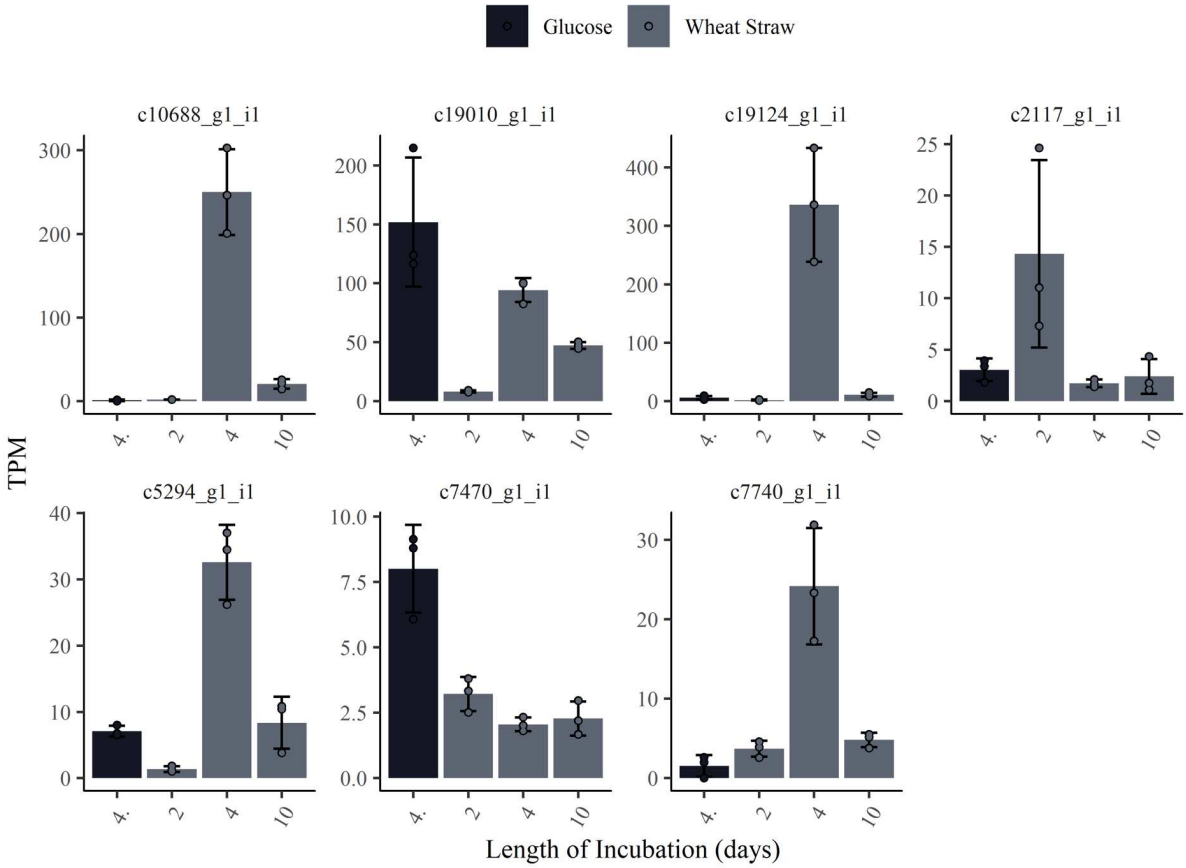

Fig. S7. Transcript per million (TPM) of contigs identified as sharing significant similarity with c2092_g1_i1. Reads with a similarity identity of over 30 % to c2092_g1_i1 were considered as displaying significant homology. Circles represent sample values of biological replicates (N=3), and error bars ± SD of the mean.

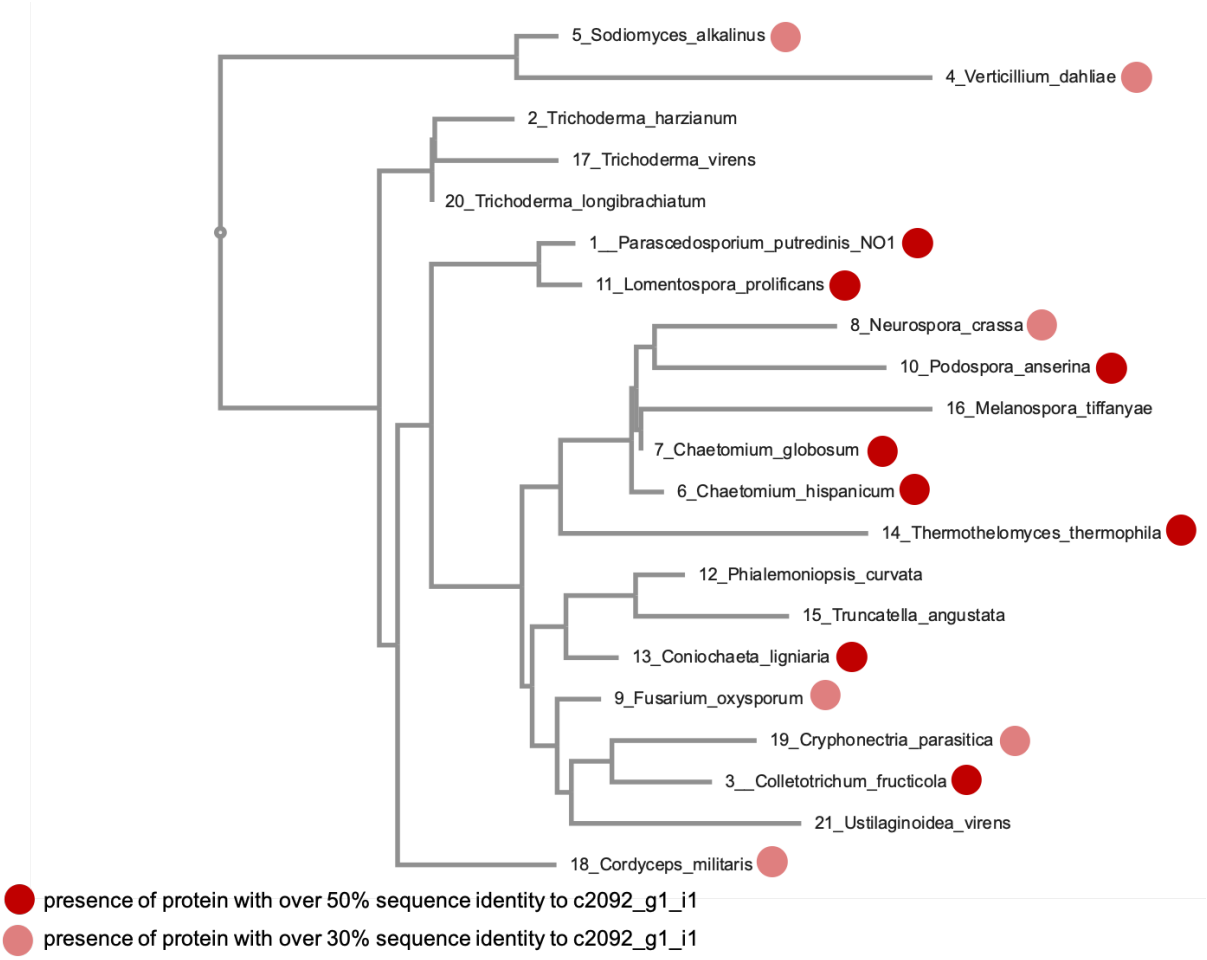

Fig. S8. Distribution of proteins with over 30 % identity to c2092_g1_i1 amongst species from the Sordariomycetes class. The phylogenetic tree was assembled from ITS sequences downloaded from the UNITE database. Multiple sequence alignment was performed using MAFFT (https://mafft.cbrc.jp). The phylogenetic tree was assembled using neighbor joining on conserved sites, and the Jukes-Cantor substitution model was used.

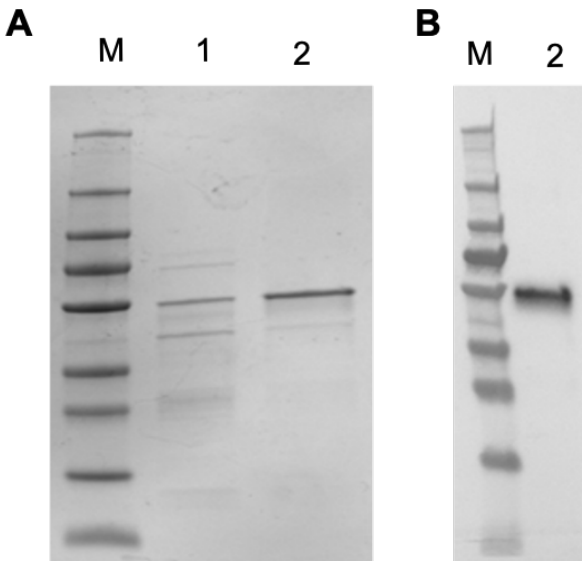

Fig. S9. Identification of the recombinant protein by SDS-PAGE (*A*) and western blot analysis (*B*). Lane M: molecular weight standards (PageRuler Plus Thermo Scientific). 1: denatured protein purified and refolded in the absence of CuSO4, 2: denatured protein refolded in the presence of 0.1 mM CuSO4. Protein identity was confirmed through western blotting using Anti-6-His antibody (*B*) and mass-spectrometry.

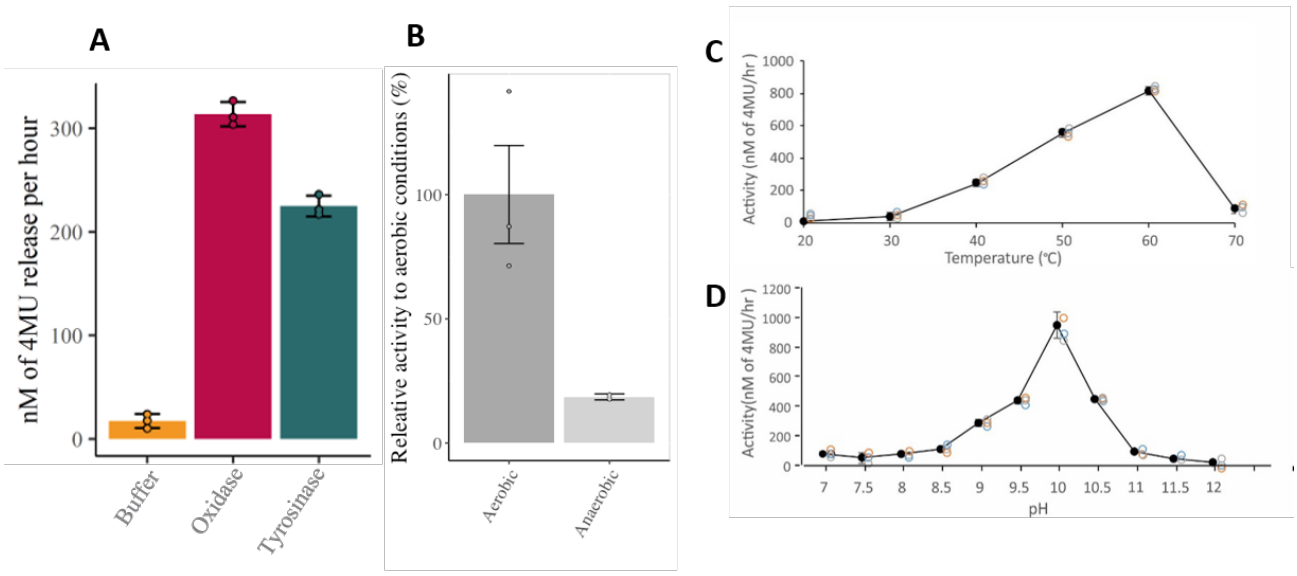

Fig. S10. β-etherase activity of the oxidase against the synthetic substrate GGβ4MU. (*A*) Fluorescence activity of purified protein against commercial mushroom tyrosinase and buffer control reaction. (*B*) GGβ4MU assay performed in presence and absence of oxygen with recombinant protein (*C-D*) optimum temperature and pH as assessed by GGβ4MU assay. Circles represent sample values, and bars sample mean ± SD, N=3.

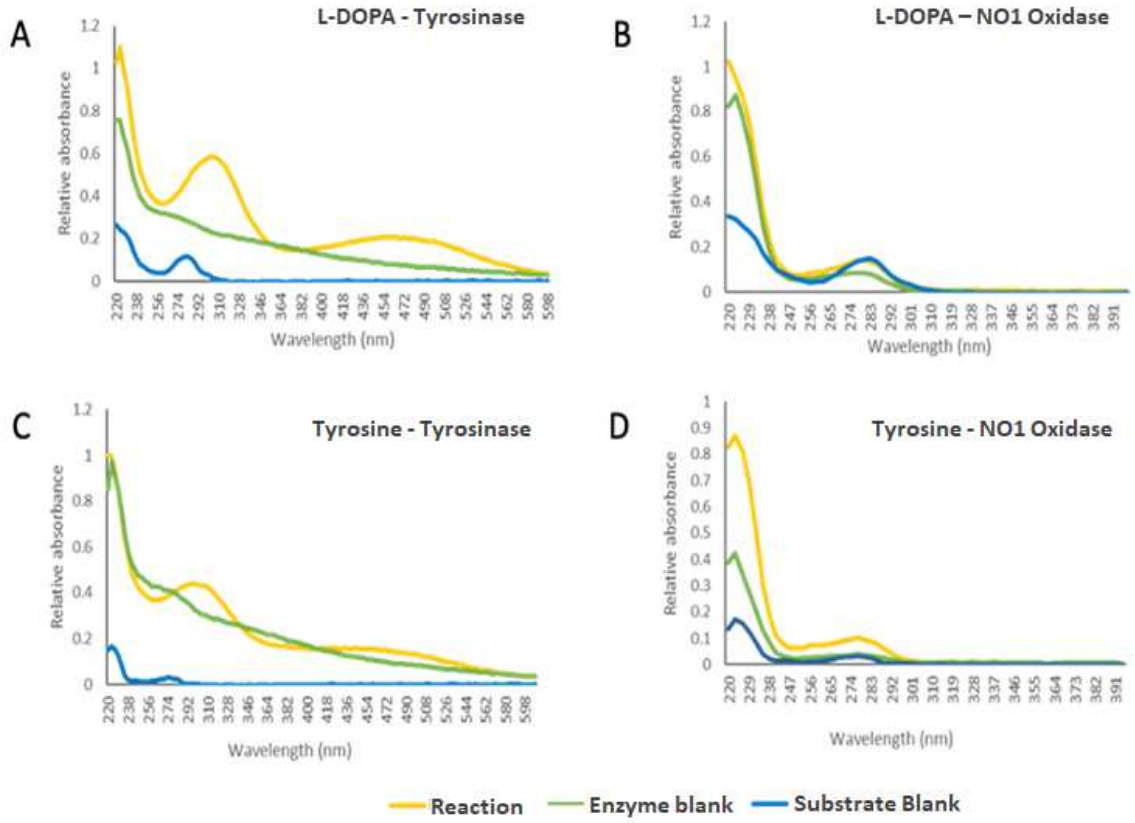

Fig. S11. UV spectra showing activity of the *P. putredinis* NO1 oxidase and tyrosinase against different phenolic compounds. Either was incubated in 50mM Tris pH 8.5 at room temperature with 1mM of substrate against enzyme only or substrate only as controls, (*A*) L-DOPA reaction with tyrosinase, (*B*) L-DOPA reaction with the oxidase, (*C*) tyrosine reaction with tyrosinase, (*D*) tyrosine reaction with the oxidase.

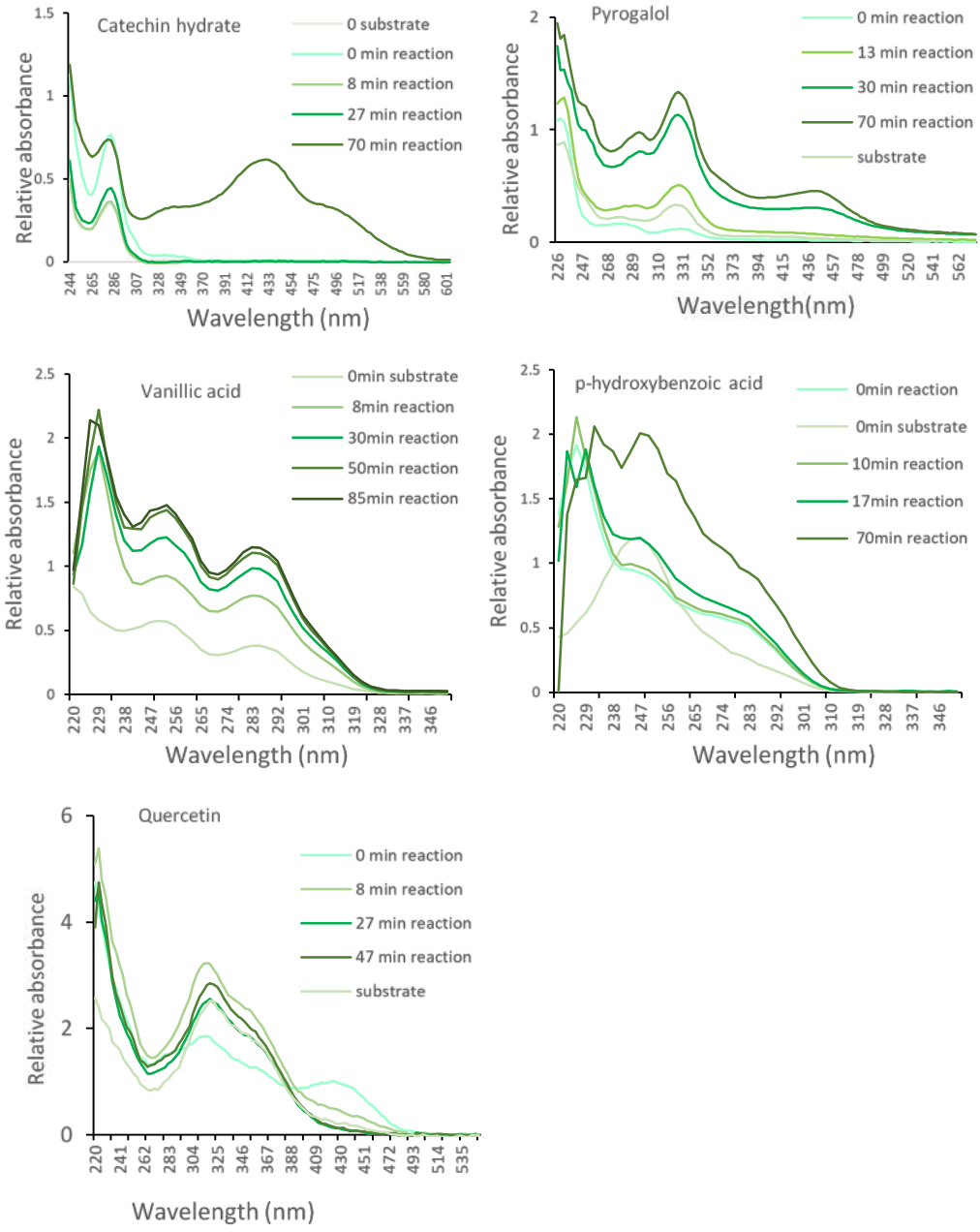

Fig. S12. UV spectra showing activity of the *P. putredinis* NO1 oxidase against different phenolic compounds. 1mg/mL of the enzyme was incubated in 50mM Tris pH 8.5 at room temperature with 1mM of either catechin hydrate, pyrogalol, vanillic acid, p-hydroxybenzoic acid or quercetin.

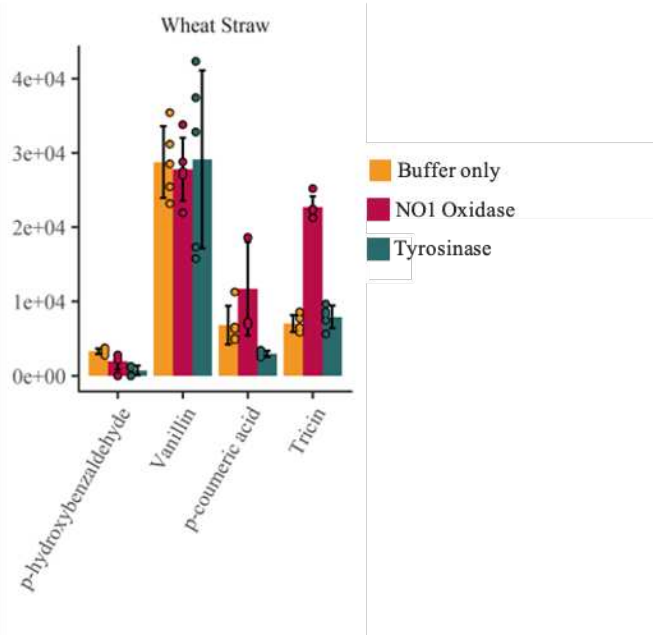

Fig. S13. Release of products from lignocellulosic substrates after incubation with NO1 oxidase, mushroom tyrosinase and buffer only. Reactions were performed at physiological pH 8.5 and 30 °C prior to the reaction products being extracted from the reaction supernatant using ethyl acetate and analysed with high-performance liquid-chromatography. Circles represent the individual sample values (N=5), and error bars ± SD of the mean.

Table S1. Proteins showing homology to the oxidase within the transcriptome of *P. putredinis* NO1. BLASTp searches were performed on the c2092_g1_i1 sequence against the assembled *P. putredinis* NO1 transcriptome.

sseqid	evalue	pident	length	bitscore	Similarity%	Similarity
c19124_g1_i1_4	9.4E-111	43.796	411	330	0.608	256/421
c7740_g1_i1_6	8.17E-77	38.482	382	243	0.508	23/439
c10688_g1_i1_2	1.72E-74	40.395	354	236	0.52	226/435
c5294_g1_i1_3	1.65E-71	37.366	372	229	0.52	223/429
c2117_g1_i1_2	2.9E-57	36.963	349	191	0.422	184/436
c19010_g1_i1_4	2.94E-32	29.254	335	125	0.325	164/505
c7470_g1_i1_2	2.25E-26	23.37	368	108	0.376	169/449

Table S2. Proteins with homology to the oxidase within NCBI non-redundant database. BLASTp searches were performed on the c2092_g1_i1 sequence against the non-redundant protein database held by NCBI. Results were filtered to >50 % identity.

	Description	Max Score	Total Score	Query Cover	E value	Percent identity
gb|PKS12997.1|	hypothetical protein jhhlp_000338 [Lomentospora prolificans]	713	713	100%	0.0	87.50%
ref|XP_016642676.1|	Tyrosinase central domain protein [Scedosporium apiospermum]	674	674	100%	0.0	82.40%
gb|TPX10091.1|	hypothetical protein E0L32_001288 [Phialemoniopsis curvata]	572	572	93%	0.0	67.19%
gb|ELA32929.1|	tyrosinase central domain protein [Colletotrichum fructicola Nara gc5]	506	506	99%	7e-176	57.95%
gb|KZL67883.1|	tyrosinase central domain-containing protein [Colletotrichum tofieldiae]	501	501	97%	8e-174	58.90%
gb|EQB58959.1|	hypothetical protein CGLO_00722 [Colletotrichum gloeosporioides Cg-14]	497	497	92%	3e-172	59.89%
gb|KZL82263.1|	tyrosinase central domain-containing protein [Colletotrichum incanum]	496	496	97%	3e-172	58.15%
gb|KXH49404.1|	tyrosinase central domain-containing protein [Colletotrichum nymphaeae SA-01]	486	486	99%	2e-168	55.88%
gb|KXH35131.1|	tyrosinase central domain-containing protein [Colletotrichum simmondsii]	485	485	99%	1e-167	55.64%
gb|OLN85731.1|	Grixazone synthase 2 [Colletotrichum chlorophyti]	484	484	92%	3e-167	58.99%
ref|XP_018157362.1|	Tyrosinase central domain-containing protein [Colletotrichum higginsianum IMI 349063]	481	481	92%	4e-166	59.37%
gb|EXF76797.1|	tyrosinase central domain-containing protein [Colletotrichum fioriniae PJ7]	479	479	99%	2e-165	55.15%
gb|TDZ75107.1|	Tyrosinase-like protein orsC [Colletotrichum trifolii]	476	476	92%	4e-164	59.95%
gb|TKW48599.1|	hypothetical protein CTA1_467 [Colletotrichum tanaceti]	473	473	92%	7e-163	58.42%
gb|TDZ15437.1|	Tyrosinase-like protein orsC [Colletotrichum orbiculare MAFF 240422]	470	470	92%	4e-162	60.48%
ref|XP_001227696.1|	hypothetical protein CHGG_09769 [Chaetomium globosum CBS 148.51]	469	469	100%	2e-161	55.50%
gb|TDZ29471.1|	Tyrosinase-like protein orsC [Colletotrichum spinosum]	460	460	92%	2e-157	57.00%
ref|XP_022470530.1|	tyrosinase central domain-containing protein [Colletotrichum orchidophilum]	458	458	99%	2e-157	54.66%
gb|OIW32989.1|	tyrosinase central domain-containing protein [Coniochaeta ligniaria NRRL 30616]	447	447	92%	5e-153	53.79%

gb|KXH30586.1|	tyrosinase central domain-containing protein [Colletotrichum salicis]	447	447	97%	3e-152	54.02%
gb|RKU41032.1|	hypothetical protein DL546_002981 [Coniochaeta pulveracea]	442	442	99%	5e-151	51.96%
gb|KZL64229.1|	tyrosinase central domain-containing protein [Colletotrichum incanum]	434	434	92%	4e-145	55.17%
gb|TEA15757.1|	Tyrosinase-like protein orsC [Colletotrichum sidae]	427	427	92%	6e-145	55.00%
gb|OHW92206.1|	tyrosinase central domain-containing protein [Colletotrichum incanum]	420	420	84%	5e-143	57.73%
ref|XP_018162984.1|	Tyrosinase central domain-containing protein [Colletotrichum higginsianum IMI 349063]	425	425	92%	1e-142	54.38%
gb|TID02585.1|	Tyrosinase ustQ [Colletotrichum higginsianum]	425	425	92%	1e-142	54.38%
gb|OLN83361.1|	Tyrosinase 2 [Colletotrichum chlorophyti]	417	417	92%	5e-141	51.97%
emb|CCF32411.1|	hypothetical protein CH063_04807 [Colletotrichum higginsianum]	412	412	84%	7e-140	56.85%
gb|KZL72889.1|	tyrosinase-like protein [Colletotrichum tofieldiae]	412	412	84%	7e-140	57.14%
gb|TKW50870.1|	hypothetical protein CTA1_3684 [Colletotrichum tanaceti]	419	419	92%	7e-140	52.39%
gb|KDN70624.1|	hypothetical protein CSUB01_04485 [Colletotrichum sublineola]	417	417	92%	1e-139	53.58%
gb|EXF84421.1|	hypothetical protein CFIO01_02736 [Colletotrichum fioriniae PJ7]	409	409	92%	1e-136	52.22%
ref|XP_003664995.1|	tyrosinase-like protein [Thermothelomyces thermophilus ATCC 42464]	404	404	92%	3e-136	54.09%
gb|TQN72542.1|	Tyrosinase-like protein orsC [Colletotrichum sp. PG-2018a]	407	407	89%	5e-136	54.77%
ref|XP_003351009.1|	uncharacterized protein SMAC_04313 [Sordaria macrospora k-hell]	399	399	97%	6e-134	50.12%
ref|XP_006692366.1|	hypothetical protein CTHT_0018720 [Chaetomium thermophilum var. thermophilum DSM 1495]	395	395	89%	1e-132	54.67%
gb|TDZ58291.1|	Tyrosinase-like protein orsC [Colletotrichum trifolii]	393	393	79%	6e-132	57.67%
gb|TDZ23501.1|	Nitroalkane oxidase [Colletotrichum orbiculare MAFF 240422]	409	409	80%	8e-132	57.75%
ref|XP_022471338.1|	hypothetical protein CORC01_10513 [Colletotrichum orchidophilum]	397	397	92%	9e-132	50.78%
gb|KXH34366.1|	hypothetical protein CSIM01_00277 [Colletotrichum simmondsii]	396	396	92%	2e-131	50.51%

gb|KXH69104.1|	hypothetical protein CSAL01_01466 [Colletotrichum salicis]	389	389	81%	3e-129	56.19%
ref|XP_008090963.1|	hypothetical protein GLRG_02114 [Colletotrichum graminicola M1.001]	378	378	79%	2e-126	56.44%
ref|XP_001227853.1|	hypothetical protein CHGG_09926 [Chaetomium globosum CBS 148.51]	373	373	92%	5e-124	50.00%
gb|TDZ28941.1|	Tyrosinase-like protein orsC [Colletotrichum spinosum]	371	371	73%	2e-122	58.14%
gb|ELA37064.1|	hypothetical protein CGGC5_3508 [Colletotrichum fructicola Nara gc5]	364	364	72%	1e-121	59.52%
ref|XP_007911158.1|	putative tyrosinase-like protein [Phaeoacremonium minimum UCRPA7]	363	363	68%	2e-121	59.22%
gb|EQB52888.1|	hypothetical protein CGLO_07432 [Colletotrichum gloeosporioides Cg-14]	361	361	72%	2e-120	59.86%
gb|TEA10724.1|	Nitroalkane oxidase [Colletotrichum sidae]	373	373	73%	4e-118	58.33%
ref|XP_024731024.1|	putative tyrosinase [Meliniomyces bicolor E]	331	331	79%	2e-108	51.38%
emb|CDP29730.1|	Putative tyrosinase [Podospora anserina S mat+]	326	326	81%	4e-106	50.15%
emb|VBB81548.1|	Putative tyrosinase [Podospora comata]	326	326	81%	5e-106	50.15%
ref|XP_001273822.1|	tyrosinase, putative [Aspergillus clavatus NRRL 1]	326	326	83%	2e-105	50.00%
ref|XP_001905273.1|	uncharacterized protein PODANS_5_7480 [Podospora anserina S mat+]	323	323	80%	3e-105	50.00%
gb|PGH18781.1|	hypothetical protein AJ79_00194 [Helicocarpus griseus UAMH5409]	325	325	83%	5e-105	50.15%
gb|PBP21500.1|	hypothetical protein BUE80_DR007716 [Diplocarpon rosae]	278	278	68%	4e-88	50.17%

Table S3. Purification of recombinant protein. The heterologously expressed protein was purified using anion exchange (Q) and size-exclusion chromatography (S.E). Protein concentration and VT221 activity was calculated after each purification step.

Purification steps	**Total Protein mg**	**Activity(mU) (nmol/mg/hr)**	**Specific (U/mg)**	**Yield (%)**	**Purification fold**
Culture filtrate	1024	7500	7.32	100	1
Q	29.25	2600	88	34.67	12
S.E	14	1950	139	26	19

Table S4. Substrate specificity of the *P. putredinis* NO1 oxidase.

Substrate	Etherase reactivity	Tyrosinase reactivity
Tyrosine methyl ester	-	+
L-Dopa(3,4-dihydroxy-L-phenylalanine	-	+
Dopamine hydrochloride	-	+
Caffeic acid (catechol oxidase substrate)	-	+
4-Methyl-Catechol (catechol oxidase substrate)	-	+
Tyrosol (catechol oxidase substrate)	-	-
Tannic acid	-	-
(+)-catechin hydrate	+	+
Pyrogallol	+	+
4-hydroxybenzoic acid	+	-
Quercetin	+	-
Vanillic acid	+	-

Dataset S1 (separate file).

Descriptions of sequences containing catalytic carbohydrate active domains.