Hemocyanin facilitates lignocellulose digestion by wood-boring marine crustaceans

Katrin Besser1, Graham P. Malyon2, William S. Eborall1, Giovanni Paro da Cunha3, Jefferson G. Filgueiras3, Adam Dowle4, Lourdes Cruz Garcia2, Samuel J. Page5, Ray Dupree5, Marcelo Kern1, Leonardo D. Gomez1, Yi Li1, Luisa Elias1, Federico Sabbadin1, Shaza E. Mohamad1,6, Giovanna Pesante1, Clare Steele-King1, Eduardo Ribeiro da Azevedo3, Igor Polikarpov3, Paul Dupree7, Simon M. Cragg2, Neil C. Bruce1 & Simon J. McQueen-Mason1

Woody (lignocellulosic) plant biomass is an abundant renewable feedstock, rich in polysaccharides that are bound into an insoluble fiber composite with lignin. Marine crustacean woodboringers of the genus Limnoria are among the few animals that can survive on a diet of this recalcitrant material without relying on gut resident microbiota. Analysis of fecal pellets revealed that Limnoria targets hexose-containing polysaccharides (mainly cellulose, and also glucomannans), corresponding with the abundance of cellulases in their digestive system, but xylans and lignin are largely unconsumed. We show that the limnoriid respiratory protein, hemocyanin, is abundant in the hindgut where wood is digested, that incubation of wood with hemocyanin markedly enhances its digestibility by cellulases, and that it modifies lignin. We propose that this activity of hemocyanins is instrumental to the ability of Limnoria to feed on wood in the absence of gut symbionts. These findings may hold potential for innovations in lignocellulose biorefining.
ignocellulosic, woody tissues provide mechanical support to higher plants. They are formed of a macromolecular composite of cellulose microfibrils coated and bound to one another by hemicellulose and lignin, producing a strong material that is resistant to biological and chemical degradation. Cellulose is made of β-1,4-linked glucan polymers that form strong paracrystalline microfibrils, whereas the hemicellulose may be galactoglucomannans in softwoods (produced by gymnosperms such as pine) or predominantly glucuronoxylans in hardwoods (produced by angiosperms such as willow). These polysaccharides typically comprise ~70% of the wood, but are rendered hard to access within a hydrophobic macromolecular material interpenetrated and bound by the polyphenolic lignin. Lignin is unusual among biological polymers in that it is not made by specific polymerases. Instead, the constituent monolignols are secreted into the cell wall space and then polymerize without enzymatic assistance in situ by oxidative coupling, followed by a re-aro-matization reaction, which can also form bonds between lignin and hemicellulose. This gives rise to a polymer without a regular repeat structure in which specific lytic enzymes can evolve, resulting in high resistance to degradation.

Lignocellulose represents the largest pool of fixed carbon in the terrestrial environment and the abundance and availability of lignocellulosic materials such as timber, crop residues, and dedicated biomass crops makes them an attractive feedstock for the production of renewable fuels, chemicals, and materials without competing with food demand. Despite their great abundance, the recalcitrance of woody materials to degradation presents a challenge for cost-effective biofuels production, as well as for digestion by heterotrophic organisms that feed on this material, such as fungi, bacteria, invertebrates, and other animals.

Fungi are preeminent among terrestrial wood-degrading organisms, owing to their ability to degrade lignin. Fungal lignin degradation follows two general strategies: white-rot fungi such as Phanerochaete chrysosporium are able to fully depolymerize and metabolize lignin, whereas brown-rot fungi such as Postia placenta modify lignin extensively through depolymerization and repolymerization, but do not metabolize it, in order to gain access to the polysaccharides in wood. These different strategies are reflected in the secretome of such fungi: white rots invest in an extensive array of ligninolytic peroxidases and carbohydrate-active enzymes (CAZymes), whereas brown rots use a radical-generating system based on Fenton chemistry to attack lignin and polysaccharides, and deploy a less extensive range of CAZymes to degrade wood.

Besides the well-studied capability of bacteria to degrade cellulose and hemicellulose, there is growing evidence of their ligninolytic capacity by the utilization of peroxidases, laccases, and superoxide dismutases. A number of invertebrates have evolved to live on woody plant materials, particularly insects such as termites and beetles. These animals typically have complex digestive systems, involving several stages of digestion, and rely on populations of specialized microbes resident in the digestive tract. In this context, termites have been best studied and their digestive systems are home to a fascinating diversity of bacteria, archaea, and protists that contribute to highly efficient lignocellulose digestion, whereas modified lignin accumulates in the feces of lower and higher termites without being metabolized by the animal. However, there is limited information about the mechanism for the lignin deconstruction in termites and beetles.

Although lignocellulose is produced by land plants, large quantities enter the marine environment through estuaries, mangroves, and salt marshes, providing a niche for specialist marine arthropods and mollusks. Tereedinid mollusks (shipworms) consume large amounts of wood and rely, at least partly, on endosymbiotic gill bacteria for lignocellulose digestion. In contrast, marine isopod crustacean woodborders of the Limnoriidae (colloquially known as gribble) do not rely on microbial symbionts, and have digestive systems devoid of resident microbiota. Recent work has begun to uncover the genes and enzymes involved in cellulose degradation in Limnoria spp. However, the mechanisms for overcoming the lignin barrier to allow these enzymes to access their substrates remain unknown.

King et al. (2010) reported that hemocyanins are abundant in the digestive tract of Limnoria quadripunctata and speculated that they might be involved in wood digestion. Hemocyanins are copper-containing proteins in the hemolymph of arthropods, in which they have evolved from structurally related phenoloxidases (POs). They are recognized as the major respiratory proteins in the hemolymph of Pancrustacea (Hexapoda and Crustacea) and share a type-3 di-copper active center involved in oxygen binding with POs, tyrosinases, and catechol oxidases. Hemocyanins are shown to exhibit PO activity upon activation, which is thought to involve loosening of the tertiary structure enabling access to the copper-containing active site. PO activity plays a role in immune responses, melanization, and sclerotization by generating quinones, which are cytotoxic and highly reactive. Quinones can undergo redox cycling, which can generate semiquinone radicals that form polymers or adducts during melanization and sclerotization, or reactive oxygen species (ROS) that have antimicrobial properties. Previous work in crayfish showed hemocyanins to be involved in tanning the chitinous cuticle lining the hindgut during molting, but they are confined to gastrotheca (temporary calcium storage organs) to fulfill this function.

Here, we present an investigation into the process of wood digestion in Limnoria, revealing the extent of lignocellulose breakdown and a characterization of the digestive proteome. We show that the oxygen-carrier hemocyanin is found in the site of wood digestion, has ligninolytic activity, and markedly increases the digestibility of wood by cellulases. We therefore suggest that hemocyanins play a key role in isopod woodborer digestion, adding another aspect to the multi-functionality of these proteins in arthropods.

**Results**

**Wood transformation during digestion.** We carried out mass loss studies and compositional analyses comparing wood fed to Limnoria with their excrated fecal pellets (Fig. 1a) to gain insights into the lignocellulose digestion process. This showed that ~22% of ingested material is consumed during gut passage, which takes about a day, both for willow (angiosperm, hardwood) and Scots pine sapwood (gymnosperm, softwood) (Supplementary Table 1).

In willow, most, if not all, of this mass loss is associated with digestion of the partially crystalline cellulose fraction, of which more than half was removed, whereas the typically more accessible hemicellulose fraction, as well as lignin, accumulate in the fecal pellets and show little change in relative amounts when normalized to mass loss (Fig. 1b). Sugar composition analysis of the sequentially hydrolyzed polysaccharide fractions hemicellulose (using trifluoroacetic acid, TFA) and cellulose (using sulfuric acid (H₂SO₄) following TFA hydrolysis) revealed that most of the mass lost during digestion was accounted for by release of glucose, the major sugar of the cellulose fraction (Fig. 1c). Most of the small amount of TFA-resistant hemicellulose in wood was reduced in fecal pellets (either because the hemicellulose is degraded during digestion or becomes more susceptible to hydrolysis by TFA after digestion), as shown by the
marked decrease in xylose and mannose contents of the cellulosic fraction (Fig. 1c). In contrast, very little sugar of the hemicellulose fraction appeared to be mobilized, and consisted mainly of xylose and smaller amounts of glucose (Fig. 1d). This indicates that the mechanism of hardwood digestion is mainly targeting the cellulosic fraction, where most of the glucose is found.

A similar situation is evident when the animals are feeding on Scots pine, with the cellulosic fraction being reduced in mass by 40% (Supplementary Fig. 1a, b). In contrast to willow, the hemicellulosic fraction is also reduced in mass by ~20% (Supplementary Fig. 1a). Softwood hemicellulose consists mainly of galactoglucomannan, but also of some arabinogluconuroxylan, and during digestion mostly mannose and some glucose are released from this fraction, with little loss of xylose content (Supplementary Fig. 1c). These data may reflect the ability of many cellulases to partially digest glucomannans which contain β-1,4-glucose units and suggest that polymers containing hexose sugars (mainly cellulose but also some glucomannan) are the main targets for softwood digestion by *Limnoria*.

Solid-state $^{13}$C nuclear magnetic resonance (ssNMR) studies of willow wood and fecal pellets confirm the major loss of cellulose when spectra are normalized to lignin (Supplementary Fig. 2, inset). The spectra also reveal that there is little or no change in the ultrastructure of the residual cellulose following digestion, evidenced by the unaltered ratio of the C4 signals of interior-to-surface or ordered-to-disordered cellulose microfibril regions (89 vs 84 ppm) in wood versus fecal pellets (Supplementary Fig. 2).

Although direct assessments of chemical changes of single molecules in a heterogeneous complex such as lignocellulosic plant biomass cannot be accomplished by Fourier transform infrared (FTIR) spectroscopy alone, analyses of wood and fecal pellets corroborate our findings of the digestion process in *Limnoria* (Supplementary Fig. 3; Supplementary Table 2). Attenuated total reflectance (ATR)-FTIR spectra are consistent with an increase in hemicellulose (1160 and 1045 cm$^{-1}$), which may be less acetylated (1737 cm$^{-1}$), an increase in lignin (1630–1670, 1160, and 1140 cm$^{-1}$), which may be more oxidized (1640 and 1550 cm$^{-1}$), and a decrease in cellulose (985 and 896 cm$^{-1}$) in fecal pellets in comparison with wood (Supplementary Fig. 3a, b; Supplementary Table 2).

Taken together, these data suggest that hexose-containing polysaccharides (cellulose and glucomannan) represent the major target of the *Limnoria* digestive system, with little reduction in the lignin and pentose-containing hemicellulose fractions. This is supported by the dominance of cellulase-like glycosyl hydrolases (GHs) of GH families 7 and 9 found in this animal’s digestive transcriptome$^{21}$. In order to enable access of the cellulases to their substrate, the tight complex of cellulose with hemicelluloses and phenolic lignin must be disrupted. Hemicellulose connects with cellulose by interaction with the crystalline surface of microfibrils$^{32}$ and this association seems to be disrupted as indicated by the decreased amount of hemicellulose in the cellulosic fraction of the biomass (Fig. 1c; Supplementary Fig. 1b). As lignin is still present in the fecal pellets, this polymer presumably has been modified to disrupt the association with the polysaccharide fraction and to facilitate access to the cellulose. To further investigate the mechanisms of wood digestion, we examined the physio-chemical properties and proteome of the digestive system.
A two-compartment digestive system. The digestive system of *Limnoria* is dominated by two structures: a cuticle-lined linear hindgut that is tightly packed with finely chopped wood particles and the hepatopancreas, which is the site of enzyme production and nutrient uptake (Fig. 2a–c). The hepatopancreas has a secretory and absorptive microvilli-lined epithelium facing its lumen that is kept free of wood particles by a complex filter system at the manifold connecting the hepatopancreas and the midgut, which leads into the hindgut (Fig. 2a–c). The hepatopancreas is contractile and, it is assumed, contracts to inject enzyme solutions, then relaxes, when the muscles of the hindgut contract to return breakdown products and enzymes for recycling, whereas spikes in the hindgut lining prevent the return of fecal matter. During feeding, seawater is ingested along with wood particles and is mixed with fluids in the digestive tract. This is evident in the pH gradient from the acidic hepatopancreas (distal region of lobes at pH 5.6 to proximal region of lobes at pH 5.8), to the neutral hindgut (pH 7.3), and to the exterior alkaline seawater (pH 7.9; Fig. 2a; Supplementary Fig. 4). The pH in terrestrial isopods is maintained ~6.0, a slight gradient from the most acidic distal end of the hindgut to the hepatopancreas, and similar trends and values were observed in freshwater amphipods. We show by microprobe analysis that the hindgut has a low oxygen content compared with the rest of the body and the surrounding seawater, suggesting it is an oxygen sink (Fig. 2d). In the absence of gut-resident microbial populations this may reflect the action of an oxygen-consuming biochemical process rather than it being the result of microbial respiration, as suggested in other isopod digestive systems.

To identify the enzymes involved in the degradation of lignocellulosic biomass in the hindgut of the woodborer, we undertook proteomic analyses using label-free emPAI (exponentially modified protein abundance index)-based relative protein quantification of the fluids collected from the hindgut, with spectral data searched against a transcriptomic *Limnoria* library (PRJNA453115; SRP142516). A third of the proteins detected in the hindgut fluid comprise CAZymes (32.9%), which represent almost exclusively GHs and few carbohydrate esterases (0.8% of CAZymes) (PXD009486; MSV000082271). Known cellulase families GH7 and GH9 account for over 50% of the soluble GHs in the gut fluids (Fig. 3a). Hemocyanins, the oxygen-carrier proteins of crustaceans, were also detected to differentially accumulate in the gut fluids, as shown by comparing relative hemocyanin protein abundance in fluids and tissue/solid content of the hindgut (Fig. 3b). Reverse Transcriptase quantitative PCR (RT-qPCR) revealed that GHs and hemocyanins are specifically transcribed in the hepatopancreas, in which the proteins were detected by western analysis (Fig. 3c, d). The prediction of N-terminal signal peptides in GH and hemocyanin sequences (using SignalP 4.1) suggested that proteins are secreted into the hepatopancreas lumen, from where they are moved into the hindgut, in which they accumulate as evidenced by immunoblots and proteomic data (Fig. 3c, d; Supplementary Fig. 5). The digestive proteome of *Limnoria* lacks peptides for known ligninolytic enzymes such as peroxidases, laccases, and other oxidoreductases known to degrade lignin.

A role for hemocyanin in lignocellulose digestion. Our study shows that respiratory hemocyanins are abundant and soluble in the hindgut luminal fluids of the marine woodborer, where they might play a part in wood digestion. Phenoloxidase activity of hemocyanins has previously been shown to occur in the hemolymph and in gastroliths of arthropods to fulfill multiple functions in immune response reactions and cuticle tanning.

To examine a potential role in wood digestion, native hemocyanins from *Limnoria* were purified by gel filtration and their identity confirmed by western blot analysis (Supplementary Fig. 6a, b). Analysis of the purified and concentrated fractions revealed a strong enrichment in hemocyanin, which forms multimers, most likely hexamers and multimers of these (Fig. 3e), as has been observed for terrestrial isopods and arthropods. We tested the purified hemocyanin for potential oxidative activity on small phenolic compounds and found that it oxidized...
pyrogallol to form purpurogallin, as indicated by the increased absorbance at \( \sim 320 \text{ nm} \), and that this reaction required activation of hemocyanin with SDS (sodium dodecyl sulfate), but was independent of hydrogen peroxide supply (Fig. 4a, Supplementary Fig. 6c, d).

When hemocyanin was purified in seawater, the expected solvent in the woodborer’s digestive system, its melting temperature (\( T_m \)) was lowered by 5°C compared with the \( T_m \) in buffer, indicating a slight unfolding of the protein (Supplementary Fig. 7a, b, c) sufficient to enable ligninolytic activity. Incubation of hemocyanin in seawater with alkali lignin led to distinctive modifications of the lignin structure, evidenced by ssNMR spectra (Fig. 4b, top panel). The most prominent change upon hemocyanin treatment is the intensity reduction in the spectral region corresponding to aromatic carbons (140–160 ppm), as well as in the signal attributed to aryl methoxyl carbons (56.1 ppm). With the exception of the C1 carbon of guaiacyl units, the spectral region between 140 and 160 ppm is caused by O-aromatic carbons (see caption of Fig. 4b). The intensity of the signal in the region of 110–140 ppm is also decreased and represents contributions from O-aromatic carbons. In addition, the spectrum of the hemocyanin-treated samples shows a signal at 24.9 ppm, which is typical of CH3 groups in aliphatic moieties. Although the NMR results alone cannot be specific about the exact mechanism of the hemocyanin–lignin interaction, the simultaneous relative reduction of the O-aromatic and aryl methoxyl carbon signals, as evidenced by the difference spectrum of hemocyanin- and seawater-treated samples (Fig. 4b inset), suggests that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm). Co-incubation of hemocyanin with the chelator diethylene triamine penta-acetic acid (DTPA) prevented these lignin modifications and NMR signals are very similar in both control and hemocyanin-treated samples (Fig. 4b inset), suggesting that hemocyanin acts mostly on the aromatic-OCH3 sites. This is also supported by the little or no change in the signal of ring carbons not linked to oxygen (105 ppm).
melting curve with a \( T_m \) of 70 °C to below 55 °C (Supplementary Fig. 7a, c), both potentially contributing to inactivation of the protein. It is interesting to note that chelators had no effect on the conformation and stability of (inactive) hemocyanins in buffer without SDS, as evidenced by an unshifted melting curve and stable \( T_m \) with increasing chelator concentration, confirming that the copper in the active center is not exposed and therefore not accessible for sequestration or enzymatic activity in buffer without SDS (Supplementary Fig. 7b, c).

Hemocyanins as effective pretreatment for saccharification. The pretreatment of woody biomass is required to reduce its recalcitrance and to improve accessibility and digestibility of the polysaccharide fractions by hydrolytic enzymes (saccharification)\(^\text{1,8}\). In industrial processes this generally involves thermo- or physicochemical processes, using acidic, alkaline, or organic solvents combined with high temperature and/or pressure\(^\text{5}\), whereas wood-degrading organisms use a range of ligninolytic enzymes and/or Fenton chemistry to achieve this\(^\text{6}\). Our data suggest that hemocyanin in the *Limnoria* digestive system provides the phenooxidative power for lignin disruption, improving access of cellulases to their substrate. To test this hypothesis we pretreated wood with hemocyanin prior to digestion with cellulases, and assessed the impact on its saccharification.

We found that incubation of powdered wood with *Limnoria* hemocyanin in seawater for short periods (10–20 minutes) at room temperature is sufficient to increase cellulase activity by 50–300% compared with controls without hemocyanin pretreatment (Fig. 5, Supplementary Fig. 8a). Similar results were obtained using the recombinant limnoriid *LgCel7B* (*LgGH7B*), a processive cellobiohydrolase (CBH)\(^\text{22}\), or the fungal *HjCel7A* (*Hypocreopsis jecorina* CBH I) in saccharification reactions (Supplementary Fig. 8b), as well as using softwood (Scots pine sapwood, Supplementary Fig. 8c) instead of the hardwood (willow), which have different lignocellulose compositions. This pretreatment effect of hemocyanin was suppressed by addition of the chelator DTPA, which inhibited hemocyanin activity on soluble lignin as shown above (Fig. 5a, Fig. 4b). Similarly, hemocyanin pretreatment in the presence of the strong reducing agent dithionite prevented increased saccharification by CBH, likely owing to the reduction of the enzyme’s copper center (Fig. 5b)\(^\text{18}\). Hemocyanin incubation had only a minor impact on the digestibility of pure phosphoric acid swollen cellulose (PASC) and this was unaffected by chelator treatment (Supplementary Fig. 8d), indicating that the increased digestibility of wood does not arise from direct effects on cellulose. The small increase in cellobiose release from PASC by hemocyanin pretreatment is likely due to the activity of the small amount of cellulases co-purified with hemocyanin (GHs, Supplementary Fig. 6c).

One way that protein addition might enhance cellulase action on wood might be by reducing non-productive binding of cellulases to non-cellulosic components such as lignin\(^\text{39,40}\). To assess if such an effect might be responsible for the enhanced cellulase action, the binding of cellulases to hemocyanin-
pretreated biomass was investigated. A similar proportion of the fungal HjCBH I (containing a cellulose binding module) was found to bind to the biomass during saccharification reactions independent of hemocyanin pretreatment, as shown by Coomassie-stained SDS-polyacrylamide gel electrophoresis (PAGE) gels (Supplementary Fig. 9). A very small proportion of hemocyanin remained on the wood mass after removal of the pretreatment solutions, which is likely to be caused by protein being trapped in the biomass upon pelleting prior to removal (Supplementary Fig. 9). Therefore, it is unlikely that the enhancement of digestibility seen with active hemocyanin is caused by reduced non-productive binding of cellulases to the biomass.

To assess whether hemocyanin treatment led to increased porosity of wood enhancing access to cellulose by cellulosomes, we examined the $T_2$ relaxation time profiles of N,N-Dimethylacetamide (DMAc) in hemocyanin-treated wood using low field time domain NMR. DMAc penetrates the pores in the material and its relaxation time $T_2$ becomes relative to the size of the pore it is filling (the smaller the pore, the shorter $T_2$) and to the pore surface ability to produce spin relaxation of the DMAc protons (so-called surface relativity parameter). This parameter depends on surface properties such as magnetic susceptibility and affinity to the enclosed fluid. Lignocellulose has characteristic interstitial scales (pores), presenting a $T_2$ profile with components that encode the various length scales of the pores as well as their occupancy by the enclosed fluid. The proportion of DMAc in each interstitial scale is estimated from the relative area of the respective component of the $T_2$ distribution, which can be correlated to the accessibility of each type of pore by DMAc molecules.

When comparing willow wood treated with seawater or with hemocyanin, it is apparent that the $T_2$ profiles of DMAc in both samples are similar in shape, i.e., reveal the same type of pore distribution, with three pore populations at similar length scales (Supplementary Fig. 10a, b). Assuming a similar surface relativity of DMAc in willow as in cotton cellulose, the length scales corresponding to these profiles are estimated to be in the order of 1–10 nm, 10–100 nm, and 100–1000 nm (Supplementary Fig. 10b). These pore populations are interpreted as interstitial spaces in the amorphous region of cellulose (surface of cellulose fibrils), voids in the inter-microfibril spaces, and small luminal structures, respectively. However, the relative area of the two lower $T_2$ population curves increased considerably in the hemocyanin-pretreated sample (Supplementary Fig. 10b, c), which indicates that there are more pores within these two length scales (both associated with the accessibility of the cell wall) after hemocyanin incubation. This suggests that there is increased porosity of lignocellulose and potentially better accessibility of cellulose by hydrolytic enzymes upon hemocyanin pretreatment.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of the hemocyanin extract after trypsinolysis revealed that it contained low level contamination with ferritin (1.3%, based on emPAI-derived molar percentages; Supplementary Fig. 6c), which was also detected in the gut fluid proteome (0.5 mol%; PXD009486; MSV00082271). Therefore, we tested the effect of ferritin in pretreatment reactions on willow followed by saccharification with CBH to exclude the possibility that the ferritin co-extracted with hemocyanin in our experiments contributed to the observed digestibility improvement. Ferritin pretreatment had no impact on digestibility of willow when used at a molarity five times higher than the one expected in the hemocyanin extract used in pretreatment reactions (Supplementary Fig. 11).

To compare the pretreatment performance of hemocyanin to other copper-containing phenoloxidases, *Trametes versicolor* laccase and *Agaricus bisporus* mushroom tyrosinase were used at the same molarity as hemocyanin (14 µM) in pretreatment reactions followed by saccharification and revealed no effect on digestibility by the laccase, and a smaller increase (than that by hemocyanin) for the tyrosinase (Supplementary Fig. 11).

We also show that the hemocyanin dependent enhancement of cellulase efficiency after ten minutes incubation at ambient temperature is comparable to mild thermochemical pretreatment with sodium hydroxide at 90 °C for 45 min (Supplementary Fig. 8a). To provide evidence for the potential recyclability of hemocyanins in view of industrial applications of ligninolytic agents, the pretreatment solutions from one experiment were removed after the incubation period and re-applied to untreated wood powder of a second cellulase digestion experiment. This repeated use of hemocyanin in pretreatment reactions still led to increased cellobiose release by cellulase from pretreated willow wood (Supplementary Fig. 8a).
Discussion

We have studied the digestion of wood by crustacean woodborers of the genus *Limnoria* that have no detectable microbial symbionts in their guts. Our data show that the digestive process mainly targets the cellulose fraction, and also other hexose-containing polysaccharides of wood, through the action of endogenously produced cellulases of families GH7 and GH9 that are expressed in the hepatopancreas and constitute over 50% of soluble GHS in the hindgut fluids surrounding the wood. We found that hemocyanins are abundant in the digestive proteome of the woodborer, that they can modify lignin, and that pre-incubation of wood with hemocyanins leads to a marked increase in the digestibility of wood by cellulases, which is likely due to an increase in the porosity of the lignocellulose. Ander et al. reported the activity of hemoglobin and other heme compounds towards lignin in the presence of oxidants such as hydrogen peroxide, in a similar fashion to peroxidase-type ligninas. In contrast, we report the discovery that respiratory hemocyanins in the *Limnoria* hindgut enhance lignocellulose digestibility by a mechanism that requires no additional oxidant and appears to function in a manner distinct to that reported for hemoglobin.

The considerable mass loss of 22% during wood digestion in *Limnoria* represents a loss of over 50% of the cellulose per day in vivo (compared with few weeks for similar total mass loss of 10–50% by fungis), leaving the fecal pellets enriched in the hemicellulose and lignin fractions. This differs from the degradation by wood-decay fungis, which first attack the more accessible polysaccharides (non-crystalline cellulose and hemicelluloses) before moving on to the crystalline cellulose fraction. Xylophagous insects, such as beetles and termites, also usually degrade the hemicellulose fraction in addition to cellulose, aided by their microbial gut community. In contrast to their herbivorous hosts, many symbiotic microbes are generally able to metabolize the pentose sugar xylose, resulting in nutritional benefits to the host upon release of these fermentation products. In the absence of gut microbiota in the crustacean woodborer there may be no route to utilize xylose, explaining the accumulation of xylan in fecal pellets of *Limnoria*.

Cellulose occurs as partially crystalline microfibrils of β-1,4 glucans that GHs do not digest efficiently without facilitation through oxidative attack by free radicals or by enzymes like lytic polysaccharide monooxygenases (LPMOs). We could find no putative enzymes of classes recognized for oxidative polysaccharide degradation in the digestive transcriptome or proteome of *Limnoria*. Our ssNMR data suggest that cellulose crystallinity is unchanged in fecal pellets, indicating indiscriminate degradation of amorphous and crystalline cellulose parts during wood digestion. The marine environment of the woodborer may aid the accessibility of cellulose in the absence of LPMOs, as the salts of seawater are thought to partially disrupt the hydrogen bonds between cellulose chains in microfibrils similarly to ionic liquids. The lignin matrix in wood represents a major barrier to cellulose hydrolysis, hindering the access of cellulases to their substrates. Wood-degrading fungis typically utilize extracellular reactive oxygen species for lignin breakdown and/or the action of oxidative enzymes (peroxidases and laccases). Lignin degradation to varying degrees has been reported in the microbe-containing gut systems of wood-feeding beetles, moths and termites, with modifications detected that resemble those caused by wood-decay fungis, including aromatic ring deconstruction and ring demethoxylation. A combined host and symbiont transcriptomic approach in termites revealed that sequences encoding ligninolytic enzymes such as laccases were exclusively host-derived with only one putative peroxidase identified from symbionts. Correlated laccase gene expression and phenoloxidase activity in the foregut and salivary glands, together with confirmation of phenoloxidase activity of laccases from the salivary glands towards lignin-phenolics, suggests a role for laccases in lignocellulose digestion by termites. Transcriptomic profiling of animal and microbial digestive enzymes in wood-feeding beetles suggested a more cooperative approach to lignin degradation, and identified transcripts from both origins for reductases, dehydrogenases, laccases, other multi-copper oxidas, peroxidases, and other auxiliary lignocellulolytic enzymes.

Symbiotic as well as endogenous phenoloxidase activity has been observed in the digestive tracts of some peracarid (a group of crustaceans including isopods and amphipods) detritivores that feed, at least partly, on lignocellulose. More recently, it has been proposed that *sensu strictu* phenoloxidases are lacking in peracarids and that in the hemolymph their function may be replaced by multifunctional hemocyanin, which is involved in respiration and immune responses. Our work shows that hemocyanin from *Limnoria* is abundant inside the digestive system and able to oxidise pyrogallol, a lignin-derived phenolic, but also promotes aromatic ring cleavage of lignin. Lignin peroxidases and laccases involved in lignocellulose digestion in white-rot fungis have been shown to cleave side chains and aromatic rings of lignin model compounds, leading to efficient decomposition of lignin. We speculate that in the absence of such proteins from gut fluids of the marine woodborer their ligninolytic function is provided by hemocyanin in the digestive tract, resulting in lignin modifications that improve accessibility of cellulases and leading to oxygen depletion of the hindgut.

We show that native *Limnoria* hemocyanins provide a highly effective pretreatment of wood, enabling cellulases to hydrolyze the cellulose fraction, whereas hemocyanin has no impact on the digestibility of cellulose itself. We also found that hemocyanin increased the porosity of the biomass. We, therefore, infer that it is the observed lignin-modifying activity that potentiates the hydrolysis of cellulose in wood. It is interesting to note that hemocyanin sequences have been identified in the digestive proteome and transcriptome of a termite and a xylophagous beetle, respectively, but no direct involvement in digestion has been shown. When fungal phenoloxidases, implicated in lignocellulose degradation, were used under the same conditions as for hemocyanin in pretreatment reactions, it became apparent that the laccase had no impact on wood digestibility, maybe owing to the lack of a suitable mediator, whereas the mushroom tyrosinase increased digestibility. In this study, we have shown that mushroom tyrosinase potentially enhances enzymatic saccharification of wood and it is interesting to note that it shares the same type-3 copper center with hemocyanins. In addition to their suggested role as redox partner for LPMOs, tyrosinases may also increase digestibility of wood during decomposition by fungi due to their ligninolytic activity.

Hemocyanins represent a novel class of lignin-modifying proteins that have a pronounced impact on wood digestibility and likely underpin the unusual ability of *Limnoria* to live on a diet of wood in the absence of microbial assistance in digestion. Insights gained from this digestive system may prove useful in developing effective methods for producing sugars from lignocellulose for sustainable biofuel production. The opportunity to exploit seawater rather than valuable and limited freshwater resources may provide a further significant benefit.

Methods

**Animal source.** The animal experiments described in this paper comply with and were approved by the Animal Welfare and Ethical Review Body of the University of Portsmouth (approval number 815 C). Specimens of *Limnoria quadripunctata* Holthuis were collected from a heavily infested piece of balau wood (*Shorea sp.*)
removed from a site in the intertidal zone at Portsmouth, UK. Specimens of the closely related Limnoria tripunctata Menzies were collected from heavily infested unidentified wood removed from the intertidal zone around the Isle of Wight, UK. Both species were used to set up laboratory cultures in seawater tanks.

**Mass balance experiment.** Ten batches of 10 animals (L. quadripunctata) were each fed for 28 days on a stick of willow wood or Scots pine sapwood (20 × 5 × 3 mm) previously leached for 1 week in water, dried for 48 h at 105 °C and accurately weighed. Fecal pellets produced during the experiment were collected from each batch by filtration. After the feeding period, animals were flushed, and the remaining wood was washed and dried at 105 °C before being weighed. Mass loss was calculated by subtracting the sum of weights of fecal pellets plus the remaining wood, from which the pellets had been generated, from the original weight of the wood before feeding.

**Biomass fraction analysis.** For biomass analysis the same ten samples of willow wood, Scots pine sapwood and fecal pellets (of 10 animals each) from the mass balance experiment were used. One aliquot of (3–4 mg accurately weighed) milled wood of nine mono-or oligosaccharides (arabinose, fucose, galactose, galacturonic acid, glucose, glucuronic acid, mannose, rhamnose, and xylose) each at 100 µl and were dispensed at 250, 500, and 700 µl into tubes in duplicates before drying in a speed vacuum concentrator (Savant I115D, ThermoSientific), used for all drying steps mentioned hereafter. One set of standards was hydrolyzed with TFA and the other with the with the 2-propanol acidification. All samples were milled with a ball mill with three cycles of 5 min milling each. After mashing samples were centrifuged and a 550 µl aliquot partially neutralized by adding 0.1 M NaOH and promptly incubated in 2 mm of 0.1 M NaOH for 1 h. TFA-FTIR spectra of wood powders and fecal pellets were obtained between 850–1850 cm⁻¹ using a Spectrum One spectrometer equipped with a diamond that allows collection of spectra directly on the sample without any sample preparation (PerkinElmer). Three spectra for each sample were acquired with Spectrum version 2.05.0.1 using 256 scans at a resolution of 4 cm⁻¹, and the triplicate-averaged spectrum was used for principal component analysis (PCA) of three wood and fecal pellet samples, respectively. PCA was carried out using the Unscrambler X software v 10.5 (CAMO) after peak normalization, linear baseline correction, and area normalization. Spectral assignments were made according to the literature (see Supplementary Table 2).

**Scanning electron microscopy (SEM).** Fecal pellets were obtained from a laboratory culture of L. quadripunctata and hindguts were removed by dissection and promptly incubated in fixative for 1 h. Fecal pellets were fixed in 4% glutaraldehyde in 0.2 M sodium cacodylate and 2 M sodium chloride at pH 7.4. Hindguts were fixed in 0.1 M sodium cacodylate, 3% paraformaldehyde and 0.5% glutaraldehyde pH 7.4. Tissues were rinsed in 0.2 M sodium cacodylate pH 7.4 before post-fixation with 1% (w/v) aqueous osmium tetroxide for 1 h. After 3 × 30 min rinses in reverse Osmosis water, the samples were dehydrated through a graded ethanol series, into acetone and transferred into hexamethyldisilazane, then air-dried onto adhesive tape. They were then cut into 40 µm sections, sputter coated with gold. SEM images were obtained with a JEOL 6060 LV microscope operating at 15 kV in secondary electron mode.

**Solid-state 13C NMR on willow and fecal pellets.** Samples of L. tripunctata were kept in seawater on willow sticks for 2 weeks and fecal pellets were collected daily, and used for centrifugation and then kept in freshwater with 0.02% 3 M sodium chloride. Control wood was milled using a cyclone mill (Retech) with a 0.5 mm mesh and soaked in the same seawater before being transferred into HO2/azide and kept in the same way as the fecal pellets. Magic Angle Spinning (MAS) ssNMR experiments on wood and fecal pellets were performed on a Bruker Advance III NMR spectrometer operating with TopSpin software version 3.5 at 1H and 13C Larmor frequencies of 398.8 and 100.3 MHz, respectively at 9.4 mm MAS. Precise probe experiments were conducted at room temperature and a MAS frequency of 12 kHz. The 13C chemical shift was determined using the carbonyl peak at 177.8 ppm of l-alanine as an external reference with respect to tetramethylsilane; the s2 pulse lengths were 2.5 µs (1H) and 4.2 µs (13C). 1H-13C cross-polarization (CP) with humped (70–100%) 1H RF amplitude was used to obtain the spectra with 100 kHz 1H decoupling during acquisition and a contact time of 1 ms. The recycle delay was 5 s with two blocks of 12,000 acquisitions added for each sample.

**Transmission electron microscopy (TEM).** Hepatopancreatic tissues were removed from L. quadripunctata by dissection and fixed and dehydrated as described for fecal pellets used for SEM (s.a.) before infiltration with low viscosity resin (Agar Scientific) and polymerization at 60 °C for 16 h. For TEM, gold sections (70 nm) were cut using a Leica Ultracut UCT and mounted on 200 mesh copper grids. Sections were post-stained with 2% aqueous uranyl acetate (10 min) and lead citrate (5 min in a carbon dioxide-depleted chamber). Sections were viewed with an FEI Tecnai G2 TEM operating at 120 kV. Measurements of pH values within the digestive tract. Animals (L. quadripunctata) were kept in sticks of pine wood (Pinus sylvestris) (20 × 2 × 4 mm) at 18–20 °C in seawater (taken from Langstone Harbour, Portsmouth, UK) for a month. At the day of the experiment, the animals were removed from the sticks and fixed by the cephalon and the telson with acupuncture needles to agarose plates (2% in Langstone Harbour seawater). The thoracic segments were removed with the help of tweezers to expose the hindgut and then the plate was filled up with seawater. Immediately afterwards, a 25 µm-tip pH microsensor (Unisense, Denmark) was introduced from above into the hindgut lumen. A pH profile was set to measure the pH every 25 µm, with 3 s of acquisition and 3 s of rest between measurements, starting inside the hindgut lumen going onwards into the gastrohepatic duct and pyloric caeca. Measurements were made in the proximal region of the small and large regions, near the connection to the hindgut, and in a more distal region in the larger lobes. The sensor tip was introduced into the tissue and the pH was recorded after the signal was stable.
Measurements of pH were compared by Welch’s one-way analysis of variance with region of measurement as the factor (distal region of large lobe of hepatopancreas, proximal region of large lobe, proximal region of small lobe, hindgut, seawater/agarose support for hindgut). Comparisons were made without assuming equality of variances using the Games-Howell multiple comparison method. Comparisons between the measurements in various regions of the hepatopancreas were performed with Tukey’s multiple comparison method. P-values were adjusted for multiple comparisons using the Bonferroni correction. A common mean pool was created by taking equal aliquots of all samples.

For LC-MS/MS analysis, samples were loaded onto an UltiMate 3000 RSLCnano HPLC system (Thermo) equipped with a PepMap 100 Å C18, 5 μm trap column (300 μm i.d., 150 μm, Thermo) and a PepMap, 2 μm, 100 Å, c5, EasyNano nanofrit capillary column (75 μm i.d., 500 μm, Thermo). The flow rate was 300 μL/min, the electric field was 50% (v/v) trifluoroacetic acid and the trapping flow rate was 15 μL/min. The trap was washed for 3 min before switching flow to the capillary column. Separation was achieved by gradient elution of two solvents: solvent A, aqueous 1% (v/v) formic acid; solvent B, aqueous 80% (v/v) acetonitrile containing 1% (v/v) formic acid. The flow rate for the capillary column was 300 nL/min and the column temperature was 40°C. Analyses were performed over 1 h (hemocyanin extracts) or 3 h (tissue and fluid extracts) acquisitions. For 1 h the linear multi-step gradient profile was: 3–10% B over 7 min, 10–35% B over 30 min then 35–99% B over 5 min. For 3 h acquisitions the gradient profile was: 3–10% B over 7 min, 10–33% B over 30 min, 33–99% B over 5 min then process 35–99% B for 4 min. The column was returned to initial conditions and re-equilibrated for 15 min before subsequent injections. The nanoLC system was interfaced with an Orbitrap Fusion hybrid mass spectrometer (Thermo) with an EasyNano ionization source (Thermo). Positive ESI-MS and MS2 spectra were acquired using Xcalibur software (version 4.0, Thermo). Instrument source settings were: ion spray voltage, 1,900 V; sweep, 0 Ar; ion transfer tube temperature, 275°C. MS spectra were acquired in the Orbitrap with: 120,000 resolution, scan range, m/z 375–1,750; AGC target, 4×106; max fill time, 100 ms. Data dependent acquisition were performed in top speed mode using a fixed 1 s cycle, selecting the most intense precursors with charge state 2+–4+. Easy-IC was used for internal calibration. DFs for the top5 ions per scan and extracted ions were performed for 50 s post precursor selection and a minimum threshold for fragmentation was set at 5e. MS2 spectra were acquired in the linear ion trap with: scan, turbo; quadrupole isolation, 1.6 m/z; activation type, HCD; activation energy: 32% AGC target, 5e4; first mass, 110 m/z; max fill time, 100 ms. All acquisitions were Xcalibur to inject ions for all available parallelizable time.

For the data analysis, peak lists were converted from raw to.mgf format using MStCat (ProteoWizard 3.0.9974) before submitting to a locally-running copy of the Mascot program using Mascot Daemon (version 2.5.1, Matrix Science). Data were searched against an in-house database generated from manually assembled contigs (see above) using the Mascot program using Mascot Daemon (version 2.5.1, Matrix Science). Search criteria specified: Enzyme, trypsin; Max missed cleavages, 2; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M); Peptide tolerance, 3 ppm; MS/MS tolerance, 0.5 Da; Instrument, ESI-TRAP. Search results were passed through Mascot Percolator to achieve a 1% peptide false discovery rate and filtered to require a minimum expect score of 0.05 for individual matches. Protein identifications were filtered to require a minimum of two peptide matches in the purified hemocyanin extract, and at least in one sample when comparing organ-specific abundances, with samples being triplicates each of gut tissue/solids, gut fluids, hepatopancreas tissue/solids, and hepatopancreas fluids. Relative protein molar abundances were calculated from Mascot derived emPAI scores as described by Ishihama et al.66.

**Protein ID.** The digestion was the same as for LC-MS/MS workflow (see above), Label-free quantitative proteomic analysis, but protein ID used MALDI-MS/MS for analysis. A µL aliquot of each peptide mixture was applied directly to the glass steel MALDI target plate, followed immediately by an equal volume of a freshly-prepared 5 mg mL-1 solution of 4-hydroxy-a-cyano-cinnamic acid (Sigma) in 50% aqueous (v/v) acetonitrile/methanol (for hemocyanin) or 65% aqueous (v/v) acetonitrile/methanol (for hemocyanin-extract) washed with 3 min before switching to the capillary column. Separation was achieved by gradient elution of two solvents: solvent A, aqueous 1% (v/v) formic acid; solvent B, aqueous 80% (v/v) acetonitrile containing 1% (v/v) formic acid. The flow rate for the capillary column was 300 nL/min and the column temperature was 40°C. Analyses were performed over 1 h (hemocyanin extracts) or 3 h (tissue and fluid extracts) acquisitions. For 1 h the linear multi-step gradient profile was: 3–10% B over 7 min, 10–35% B over 30 min then 35–99% B over 5 min. For 3 h acquisitions the gradient profile was: 3–10% B over 7 min, 10–33% B over 30 min, 33–99% B over 5 min then process 35–99% B for 4 min. The column was returned to initial conditions and re-equilibrated for 15 min before subsequent injections. The nanoLC system was interfaced with an Orbitrap Fusion hybrid mass spectrometer (Thermo) with an EasyNano ionization source (Thermo). Positive ESI-MS and MS2 spectra were acquired using Xcalibur software (version 4.0, Thermo). Instrument source settings were: ion spray voltage, 1,900 V; sweep, 0 Ar; ion transfer tube temperature, 275°C. MS spectra were acquired in the Orbitrap with: 120,000 resolution, scan range, m/z 375–1,750; AGC target, 4×106; max fill time, 100 ms. Data dependent acquisition were performed in top speed mode using a fixed 1 s cycle, selecting the most intense precursors with charge state 2+–4+. Easy-IC was used for internal calibration. DFs for the top5 ions per scan and extracted ions were performed for 50 s post precursor selection and a minimum threshold for fragmentation was set at 5e. MS2 spectra were acquired in the linear ion trap with: scan, turbo; quadrupole isolation, 1.6 m/z; activation type, HCD; activation energy: 32% AGC target, 5e4; first mass, 110 m/z; max fill time, 100 ms. All acquisitions were Xcalibur to inject ions for all available parallelizable time.

For the data analysis, peak lists were converted from raw to.mgf format using MStCat (ProteoWizard 3.0.9974) before submitting to a locally-running copy of the Mascot program using Mascot Daemon (version 2.5.1, Matrix Science). Data were searched against an in-house database generated from manually assembled contigs (see above) using the Mascot program using Mascot Daemon (version 2.5.1, Matrix Science). Search criteria specified: Enzyme, trypsin; Max missed cleavages, 2; Fixed modifications, Carbamidomethyl (C); Variable modifications, Oxidation (M); Peptide tolerance, 3 ppm; MS/MS tolerance, 0.5 Da; Instrument, ESI-TRAP. Search results were passed through Mascot Percolator to achieve a 1% peptide false discovery rate and filtered to require a minimum expect score of 0.05 for individual matches. Protein identifications were filtered to require a minimum of two peptide matches in the purified hemocyanin extract, and at least in one sample when comparing organ-specific abundances, with samples being triplicates each of gut tissue/solids, gut fluids, hepatopancreas tissue/solids, and hepatopancreas fluids. Relative protein molar abundances were calculated from Mascot derived emPAI scores as described by Ishihama et al.66.
using the Bioanalyzer (Agilent). Total RNA (400 [GHS]–500 [Hc] ng) was treated with RQI DNase (Promega) to remove genomic DNA, and subjected to reverse transcription (RT) (RT random primers nested 3′ RT (Life Technologies) according to the manufacturer’s instructions. Quantitative polymerase chain reaction (qPCR) of diluted cDNA and no-RT controls (equivalent to 10 ng total RNA per reaction) was performed with Fast SYBR® Green Master Mix according to the manufacturer’s protocol on a StepOnePlus™ Real-Time PCR System (Life Technologies).

Antibody production and purification. Specific primers for L. quadruplicata genes coding for hemocyanin (GenBank accessions: GU166295 [Hc1], GU166296 [Hc2], GU166297 [Hc3], GU166298 [Hc4]), GF family 5, 7 and 9 (GenBank accessions: GU068826 [GHSA], GU068827 [GHSA], FJ940756 [GH7A], F940757 [GH7B], F940758 [GH5A], long and short isoforms), ubiquitin (UB2) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were used: Hc1, TCTCCTGATTTCCGCATCACT; GACTCATATCCGCTTCGACT; Hc2, CACATTACCAGAAAATCAGACAATCAT; TCTG; Hc4, AAGGACCCATGTTTCTGATGATGACT; CACGAGTTGAAGGATTGTA; CAATGCTATTAGCTTCCCAAATA; TC; CTGCACATTTTCTTGGTATGCT; GGCCTGT; TCTGGGACACTA; GH5C, GCTCTACATCTTACGTGATAT; GATCCTAGTCATTCCAAAATA; GH7B, TGGCTTGCGAAGAATCAGTCTG; GH9A, TGCTGACGCCTCACCTCACTGCAACAGTTGAGGGA; GH9B, GGGTTGATCTTTGCCGGAAAG, CACCGCTATTACGTCTTGGCAGGTGAAGTGTTG; GAPDH, 1, TGTAATTTTCCTTCCATCAGCAGACAC, CCCACACACAGGCTGCTACA; GAPDH2, CTCTACCTGCGGCTACCATGCTGAAAGA. Quality controls have been employed according to the MIQE-guidelines67 and included verification of primer specificity, testing for contamination of RNA with genomic DNA and of the PCR mix and evaluation of consistency of cDNA synthesis, all as previously described22. Gene expression levels are shown as previously described relative ubiquitin-normalized efficiency-corrected values22.

SDS-PAGE and western blot analysis. For tissue-specific immunoblot detection 50 L. quadruplicata specimens were dissected to remove hepatopancreas and hindguts from which proteins were extracted by grinding material in 50 mM Tris–Hcl pH 8, 300 mM NaCl, 0.1% Tween 20, 1 x Halt protease inhibitor cocktail (Pierce) before pelleting any cell debris by centrifugation. The protein concentration was estimated using Bradford reagent (Pierce). After separation of 5 µg protein (organ) per lane on 10% acrylamide gel, proteins were transferred onto a nitrocellulose membrane (Bio-Rad) before being blocked with 5% (w/v) dry milk for 1 h at room temperature. Specific antibodies in rabbit (Covalab S.A.S., France). fractions of peak 1 and 2 containing hemocyanin were subjected to SDS-PAGE and sample concentration was determined by Bradford (Pierce). About 1 mg of recombinant proteins LqGH5A and LqGH9A was used to raise polyclonal anti-hemocyanin antibodies in rabbits (Covalab S.A.S., France). The purified LqHc1 was dialyzed against PBS and ~ 200 µg of protein was used to raise polyclonal anti-hemocyanin antibodies in rabbits (Covalab S.A.S., France).

To enrich the sera for antigen-specific antibodies, affinity columns were made using recombinant LqHc1 (Gu166297) purified from a bacterial culture under denaturing conditions on Ni-NTA resin using decreasing pH for wash and elution steps (The Qiagen). Puriﬁed recombinant protein preparations were dialyzed against coupling buffer (0.1 M NaHCO3, 0.5 mM NaCl, pH 8.3) and bound to CNBr-activated Sepharose™ 4 Fast Flow resin (GE Healthcare Life Sciences) followed by afﬁnity puriﬁcation of an aliquot of each crude antibody serum according to the resin manufacturer’s instructions. Puriﬁed antibody fractions were characterized for their afﬁnity to each Lq antigen by dot and western blotting and glycosylation degrees (1:5, 1:10, 1:20) and included veriﬁcation of primer specificity by PCR. The puriﬁed LqHc1 was dialyzed against PBS and ~ 200 µg of protein was used to raise polyclonal anti-hemocyanin antibodies in rabbits (Covalab S.A.S., France).

To enrich the sera for antigen-specific antibodies, affinity columns were made using recombinant LqHc1 (Gu166297) purified from a bacterial culture under denaturing conditions on Ni-NTA resin using decreasing pH for wash and elution steps (The Qiagen). Puriﬁed recombinant protein preparations were dialyzed against coupling buffer (0.1 M NaHCO3, 0.5 mM NaCl, pH 8.3) and bound to CNBr-activated Sepharose™ 4 Fast Flow resin (GE Healthcare Life Sciences) followed by afﬁnity puriﬁcation of an aliquot of each crude antibody serum according to the resin manufacturer’s instructions. Puriﬁed antibody fractions were characterized for their afﬁnity to each Lq antigen by dot and western blotting and glycosylation degrees (1:5, 1:10, 1:20) and included veriﬁcation of primer speciﬁcity by PCR. The puriﬁed LqHc1 was dialyzed against PBS and ~ 200 µg of protein was used to raise polyclonal anti-hemocyanin antibodies in rabbits (Covalab S.A.S., France).

RSF-5000 CL column (GE Healthcare) pre-equilibrated with ﬁlter-sterilized seawater or 0.05 M NaPO4, pH 7 at a ﬂow rate of 0.2–0.4 mL min−1, and fractions (0.5 mL) were collected after 0.3–1 CV. Fractions of peak 1 and 2 containing hemocyanin were pooled and concentrated before protein concentration was estimated using Bradford reagent (Pierce) or before proteins were subjected to LC-MS/MS analysis.

Uranyl acetate negative stain. A few µL of native L. limnoria hemocyanin extract were left on a copper grid (with formvar/carbon support film) for 4 min then negative stained with 1% uranyl acetate in water being imaged with a FEI Tecnai 12 TEM, operating at 120 kV, and using a Ceta camera.

UV-Vis spectroscopy. Native L. limnoria hemocyanin (160 µg mL−1, 2.2 µM) or Trametes versicolor laccase (184 µg mL−1, 3.3 µM, equivalent to 1 U; Sigma) were incubated with 1 µM pyrogallol in 100 µL of 0.1 M NaPO4, pH 6.8 at room temperature for up to 45 min. Enzyme activity of hemocyanin was induced with 2.75 mM SDS and buffer controls only lacked the protein. Ultra violet-Visible absorption spectra of aliquots were scanned from 220 to 750 nm in a NanoDrop 8000 Microvolume UV-Vis spectrophotometer (Thermo Scientiﬁc) at regular intervals.

To monitor the effect of chelators on oxygen binding to the copper center of hemocyanin, 50 µM of either ethylene diamine tetra-acetic acid (EDTA), ethylene glycol-bis(2-aminoethyl ether)-N,N′-N′-tetra-acetic acid (EGTA), or diethylene triamine penta-acetic acid (DTPA) was mixed with hemocyanin extract in seawater and scanned as above.

Solid-state 13C NMR spectroscopy of lignin. After incubation of 3% (v/v) alkali lignin (Sigma #471003 low sulfonate content, 3.3% sulfur and 50.1% carbon) in seawater supplemented with 0.18M FeSO4 and with or without 50 mM DTPA with native L. limnoria hemocyanin (~14 µM) or seawater (control) in triplicates each for 1 h at room temperature, reactions were freeze-dried before analysis by solid-state CP/MAS 13C NMR spectroscopy. 3.5 mm diameter, equipped with a Bruker 4.0 mm double-resonance MAS probe, at 13C and 1H frequencies of 100.5 and 400.0 MHz, respectively. The spinning frequency at 14 kHz was controlled by a pneumatic system that ensures rotation stability higher than ~1 Hz. Typical π/2 pulse lengths of 4 and 3.5 µs were applied for 13C and 1H, respectively. 1H decoupling field strength of π/2z = 100 kHz was used. 13C Cross-Polarization (CP)-MAS NMR spectra was measured using Multi-CP excitation29 with nine cross-polarization blocks of 1 ms and one last cross-polarization of 0.8 ms, 90–100% increment in the RF amplitude, repolarization period t2 of 0.9 s, and recycle delay of 2 s. The same experimental conditions were used for the experiments. BRUKER LoRes and LoRes 64 as described was used for data collection, and OriginLab OriginPro 9.0 SR2 for data analysis. Chemical shifts were assigned based on literature data (Supplementary Note 1).

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07575-2 | www.nature.com/naturecommunications

ARTICLE

(2018)5125 | DOI: 10.1038/s41467-018-07575-2 | www.nature.com/naturecommunications
Thermal shift assays (thermofluor). Thermal shift assays were conducted on purified proteins with SYPRO® Orange Protein Gel Stain (Life Technologies) using an Mx3000P qPCR System (Agilent Technologies). The intensity of the fluorescence was measured against a temperature gradient of 25–95 °C and values plotted to determine the melting temperature ($T_m$) by curve fitting using a five parameter sigmoid equation with the $T_m$ measured as the midpoint at http://paulsbond.co.uk/jfat^31.

Digestibility assays. Willow wood and Scots pine sapwood was ground using a cyclone mill (Retach) with a 1 mm mesh and a ball mill with three cycles of 5 min milling and then washed by aliquoting into 2 ml screw cap tubes. After shaking incubation of wood powder (10% w/v) with or without 50 mM DTPA, or 50 mM sodium dithionite, with native Linnormia hemocyanin (~14 µM), horse spleen ferritin (Sigma; 1.9 µM) or seawater alone (control) for 20–20 min at room temperature, the solids were pelleted and pretreatment solutions removed. For pretreatments of wood powder with 14 µM Trametes versicolor laccase (Sigma) or Agaricus bisporus mushroom tyrosinase (Sigma) the reactions were incubated and processed as above but in 0.05 M NaPO₄ at pH 6. For hemocyanin-re-use experiments, the hemocyanin- or seawater-containing pretreatment solutions were removed from replicates after pelleting the solids, then pooled and re-aliquoted onto new willow wood samples in equal volumes for the next pretreatment reaction; as some of the liquid had been taken up by the wood the total volume applied in the re-use experiment was slightly less than in the previous one (86 µL instead of 100 µL). Mild thermochemical pretreatments consisted of incubation of wood powder (10% w/v) in 0.5 M NaPO₄ for 45 min at 90 °C followed by extensive washing in buffer to adjust the pH. For saccharification reactions, the pretreated wood powders and Dbers were incubated with 10 µg (10% w/v) recombinant cellobiohydrolase I from Hypocrea jecorina (HyHiCH1 or Cel7A, Sigma) or L. quadriflumata cellulose (Novozymes®) in 0.05 M NaPO₄ at pH 5–7, with buffer only as no GH controls, for 5 h (wood) or 2 h (PASC), respectively, at 37–40 °C with shaking. Carbohydrate composition of the hydrolysation reaction was determined by HPAEC as described above after removal of solids by centrifugation, precipitation with 80% ethanol, drying of the supernatant containing mono- and oligosaccharides, resuspension in water and filtering through 0.2 µm polytetrafluoroethylene (PTFE) filters.

Cellulase binding to hemocyanin-pretreated biomass. Pretreatment and saccharification reactions were performed as described above for digestibility experiments, using 10 mg willow wood powder, native Linnormia hemocyanin (~14 µM) or seawater for 10 min in pretreatment reactions at room temperature, followed by 5 h saccharification with 10 µg (10% w/v) recombinant cellobiohydrolase I from Hypocrea jecorina (HyHiCH1 or Cel7A, Sigma) in 100 µL 0.05 M NaPO₄ pH 6 at 37 °C of 20 mg S. fluvialis SDS-loading buffer of hydrolyzate (16 µL aliquoted from 100 µL hydrolyzate after removal from biomass plus 4 mL of 5x SDS-loading buffer) or biomass (100 µL 1x SDS-loading buffer added to the biomass after pelleting and removal of hydrolyzate) compared with the pretreatment and saccharification solutions prior use (16 µL plus 4 µL 5× SDS-loading buffer each) were loaded onto 4–20% PROTAN * GCO* Precast Protein Gels (Bio-Rad) alongside a molecular weight marker (PageRuler Plus Prestained Protein Ladder, Thermo Fisher Scientific) after boiling at 100 °C for 5 min. Proteins were visualized by Coomassie stain with InstantBlue (Expedeon).

NMR relaxometry of hemocyanin-treated wood. Ten 10 mg willow samples were pretreated with either 100 µL seawater or hemocyanin (~14 µM) each for 3 h at room temperature and solutions removed after pelleting the biomass, which was then dried, pooled and analyzed according to the following methodology.

To assess the pore structure of the materials, we used N,N-Dimethylacetamide (DMAC, high-performance liquid chromatography grade) as a molecular probe. Although water is often used to determine pore structure in relaxometry experiments, it alters the pore distributions in biomass samples. As the interaction of DMAC and the cell wall is rather weak, the use of this molecule does not significantly change the pore structures41. The samples were dried in a vacuum oven at 510 mm Hg and 80 °C during 24 h. Then, they were soaked with DMAc in cotton cellulose T2 relaxation times ranging within the pores result in a multi-exponential decay of the CPMG signal, i.e., a

$$f(x) = ae^{-(t-b)/c}^2$$

(1)

From the model, the relative areas are obtained from the relation

$$A = \sqrt{2}a \pi c$$

(2)

and the error bars defined using error propagation from the standard deviations of the fitted parameters $a$ and $c$. The fluid enclosed in a pore interacts with the pore surface, restraining the molecular mobility of the fluid. As the surface-to-volume ratio $S/V$, defined as the inverse of the characteristic pore size, increases, the strength of the interaction. The exclusion of mobility is reflected by a decrease in the transverse relaxation time $T_2$ and can be quantified, on the fast diffusion regime, by the relation\(^{75,76}\)

$$1/T_2 = \rho(S/V) = 2\rho/r$$

(3)

where $r$ is the surface relaxation and $S$ and $V$ are the pore surface area and the volume, respectively. The surface relaxation constant depends on the particular porous material\(^7\), with its surface properties such as magnetic susceptibility and affinity to the enclosed fluid defining its ability to produce spin relaxation of the DMAC protons. This parameter is usually unknown, turning the estimation of the effective pore sizes, or interstitial scales in this case, into a daunting challenge. In a realistic scenario, the distribution of pore sizes and differences in the fluid mobility within the pores result in a multi-exponential decay of the CPMG signal, i.e., a $T_2$ distribution. For DMAC, the $T_2$ relaxation times range from $10^{-10}$ to 1 s have been reported, corresponding to pore sizes of few nanometers to few micrometers\(^41\). Larger $T_2$ values on the distribution profile correspond to larger pores\(^7\).

Data availability

All transcriptomic raw sequence data were deposited in the NCBI BioProject database (PRJNA453115) and Sequence Read Archive (SRA) (SRP142516). All proteomic data sets, including raw data files, processed peak lists, and database search results have been deposited to the ProteomeXchange Consortium via the MassIVE partner repository with the dataset identifier PXD009486 and the MassIVE accession code MSV000082271. A reporting summary for this Article is available as a Supplementary Information file. All other data are available from the corresponding authors upon reasonable request.

Received: 2 May 2018 Accepted: 7 November 2018 Published online: 03 December 2018

References

1. Marriott, P. E., Gómez, L. D. & McQueen-Mason, S. J. Unlocking the potential of lignocellulosic biomass through plant science. New Phytol. 209, 1366–1381 (2016).
2. Kumar, M., Campbell, L. & Turner, S. Secondary cell walls: biosynthesis and manipulation. J. Exp. Bot. 67, 515–531 (2016).
3. Vanholme, R. et al. Metabolic engineering of novel lignin in biomass crops. New Phytol. 196, 978–1000 (2012).
4. Youngs, H., Somerville, C. Development of feedstocks for cellulosic biofuels. F1000 Biol. Rep. 4, 10 (2012).
5. Sun, S., Sun, S., Cao, X. & Sun, R. The role of pretreatment in improving the enzymatic hydrolysis of lignocellulosic materials. Bioresour. Technol. 199, 49–58 (2015).
6. Grzeby, S. M. et al. Lignocellulose degradation mechanisms across the tree of life. Curr. Opin. Chem. Biol. 29, 108–119 (2015).
7. Floudas, D. et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336, 1715–1719 (2012).
8. Alessi, A. M. et al. Revealing the insoluble metasecretome of lignocellulose-degrading microbial communities. Sci. Rep. 7, 2386 (2017).
9. Watanabe, H. & Tokuda, G. Cellulolytic systems in insects. Annu. Rev. Entomol. 55, 609–632 (2010).
10. Martínez, A. T. et al. Biodegradation of lignocelluloses: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int. Microbiol. 8, 195–204 (2005).
11. Zhang, J. et al. Localizing gene regulation reveals a staggered wood decay mechanism for the brown rot fungus Postia placida. Proc. Natl Acad. Sci. USA 113, 10968–10973 (2016).
12. Goodell, B. 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).
Acknowledgements

From the University of York we thank Jared Cartwright for advice on protein purification and Meg Stark for negative staining and imaging of native hemocyanin, both Bioscience Technology Facility, and Phil Roberts for design of the Limnoria schematic drawing. This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), UK (Grants BB/G016178/1 and BB/L001926/1). I.P., G.P.C., J.G.F., and E.R.A. acknowledge funding from the São Paulo Research Foundation (FAPESP), Brazil (grants 2015/13684-0 and 2017/24463-3) and from the National Council for Scientific and Technological Development (CNPq), Brazil (grants 423693/2016-6 and 312852/2014-2). 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).

Author contributions

K.B., S.M.C., N.C.B. and S.J.M.-M. designed research; K.B., G.P.M., W.S.E., G.P.C., J.G.F., E.R.A., A.D., L.C.G., S.J.P., R.D., M.K., L.D.G., L.E., F.S., S.E.M., G.P. and C.S.K. performed research; K.B., I.P., E.R.A., A.D., Y.L., P.D., R.D., S.M.C., N.C.B. and S.J.M.-M. analyzed data; K.B. and S.J.M.-M. wrote the manuscript with contributions from co-authors. All authors reviewed and approved the final manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07575-2.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018