The distribution of novel bacterial laccases in alpine paleosols is directly related to soil stratigraphy

Dandare, S., Young, J. M., Kelleher, B. P., & Allen, C. (2019). The distribution of novel bacterial laccases in alpine paleosols is directly related to soil stratigraphy. *Science of the Total Environment*, *671*, 19-27.

Published in:
Science of the Total Environment

Document Version:
Peer reviewed version

The distribution of novel bacterial laccases in alpine paleosols is directly related to soil stratigraphy.

S.U. Dandare[1], *J.M. Young*[1], *B.P. Kelleher*[2], *C.C.R. Allen**[1,3]

[1]*School of Biological Sciences, Queen's University Belfast, UK*

[2]*School of Chemical Sciences, Dublin City University, Ireland*

[3]*Institute for Global Food Security, Queen's University Belfast, UK*

**Corresponding author*

ABSTRACT

Bacterial laccases are now known to be abundant in soil and to function outside of the cell facilitating the bacterial degradation of lignin. In this study we wanted to test the hypotheses that: i) Such enzymes can be identified readily in stratified paleosols using metagenomics approaches, ii) The distribution of these genes as potential 'public good' proteins in soil is a function of the soil environment, iii) Such laccase genes can be readily retrieved and expressed in *E.coli* cloning systems to demonstrate that *de novo* assembly processes can be used to obtain similar metagenome-derived enzyme activities. To test these hypotheses, *in silico* gene-targeted assembly was employed to identify genes encoding novel typeB two-domain bacterial laccases from alpine soil metagenomes sequenced on an Illumina MiSeq sequencer. The genes obtained from different strata were heterologously cloned, expressed and the gene products were shown to be active against two classical laccase substrates. The use of a metagenome-driven pipeline to obtain such active biocatalysts has demonstrated the potential for gene mining to be applied systematically for the disccovery of such enzymes. These data ultimately further demonstrate the application of soil pedology methods to

environmental enzyme discovery. As an interdisciplinary effort, we can now establish that palesols can serve as a useful source of novel biocatalytic enzymes for various applications. We also, for the first time, link soil stratigraphy to enzyme profiling for widespread functional gene activity in paleosols.

Keywords paleosols, stratigraphy, laccase, metagenome

1. INTRODUCTION

Laccases are oxidoreductase enzymes which can act upon a broad range of phenolic and non-phenolic compounds *via* the reduction of molecular oxygen to H_2O. Consequently, they have many applications as 'green' biocatalysts in varied industries - including textile and dye decolourisation, bioremediation, pulp bio-bleaching, microbial fuel cells and biosensors (Couto et al., 2006). The advent of high throughput sequencing and associated metagenomics has revealed that bacteria possess a huge diversity of laccase-like multicopper-oxidases (LMCOs) genes. Many of these appear to possess N terminal signal peptide sequences (Ausec et al., 2011), suggesting they are secreted by the cell, potentially enabling oxidation of lignin-derived compounds outside the cell. Only a few bacterial LMCOs have been purified and characterised to date; however, data from these studies suggests unique and biotechnologically advantageous properties, including chloride tolerance (Guan et al., 2014), increased stability at extremes of pH and temperature (Chanuhan and Jha, 2018, Fang et al., 2011) and ionic liquid tolerance (Chauhan et al., 2018). Soils are the primary site of terrestrial organic carbon degradation and are some of the most genetically diverse environments on earth (Torsvik 1990). Despite this, very little is known about both the abundance and distribution of Bacterial LMCO's in the soil.

Bacterial laccases can be divided into several categories based upon sequence similarity, the number of cupredoxin domains they possess and their tertiary structure (Ausec et al., 2011). Type B and C two domain laccases are homotrimeric proteins with each monomer consisting of two cupredoxin domains (Martins et al., 2015). By contrast, three-domain bacterial laccases are monomeric and can be

47 subdivided into three categories based upon sequence similarity, 'small' three domain laccases, 'big'
48 three domain laccases and CotA type laccase (similar to cotA from *B. subtilis*). Most Bacterial laccases
49 which have been heterologously expressed and characterised to date are three domain type laccases
50 (Martins et al., 2015), and recombinant three-domain Laccases have found widespread industrial
51 usage.

52 There is presently considerable interest in the degradation of lignin in diverse environments, where
53 different microorganisms compete for lignin and lignin breakdown products (Cragg et al., 2015).
54 However, fundamental questions about the distribution of lignin-degrading bacteria are evident: Do
55 different types of lignin degrading bacteria exist in competing or cooperative bacterial communities?
56 How do these bacteria exist alongside the lignin-degrading fungi? Do various types of extracellular
57 lignin-degrading 'common good' proteins (such as peroxidases and laccases) exhibit functional
58 variation that is related to the environmental niche where they are found as opposed to their
59 phylogenetic origin? Analysis of soil metagenomes may help us to answer some of these questions
60 (Jacquiod et al., 2014). Further gene-mining approaches can be used to retrieve functional laccases
61 from soil metagenomes also (Yang et al., 2018).

62 In this study, we used a metagenome analysis of genes for bacterial laccase enzymes in an alpine
63 paleosol. In bacteria, laccases have differing functionalities: whereas some enzymes are clearly
64 associated with non-catabolic roles, such as endospore germination in Firmicutes and are not known
65 to be secreted from the cell (Francisco et al., 2004); others (e.g. the two-domain group of bacterial
66 laccases) appear to play an extracellular function.

67 Here we took samples from Alpine paleosols (Mahaney *et al.,* 2016) that we know have well-defined
68 stratigraphy, with the inceptisols harbouring distinct Ah, Bw and Cox soil horizons. We then tested
69 the hypothesis that there is a differential distribution of the major well-defined groups of functional
70 bacterial laccases across these horizons, which are therefore affected by changing physical and
71 chemical environmental conditions. By investigating the demarcation of different classes of laccase

gene across paleosol soil horizon boundaries, we sought to show for the first time that the different environments present in these soil environments could affect the molecular ecology of such genes and therefore that their functionality could be linked to their ecological niche.

In validating these core questions, we also wanted to show that the two-domain group of bacterial laccases can be cloned and expressed for biochemical analysis when directly derived from such soil metagenome sources.

2. MATERIALS AND METHODS

Materials

The substrates 2, 2' – Azino-bis (3-ethylbenzothiazoline-6-sulfonate) (ABTS) and syringaldazine (SGZ) were products obtained from Sigma-Aldrich (St. Louis, MO, USA). Mo Bio Powersoil DNA extraction kit was obtained from Qiagen. Gel Extraction Kit, Plasmid Mini-Prep Kit, Dream Taq DNA polymerase, Phusion green DNA polymerase, ClonJET cloning kit, aLicator LIC cloning and expression system, SYBR Green qPCR Master Mix, prestained protein ladder, and the TOP10 chemically competent *E. coli* cells were all purchased from Thermo Fisher Scientific (UK). The BL21 (DE3) chemically competent cells were purchased from New England Biolabs, while the restriction enzymes were obtained from Fermentas. All other chemicals used in this study were of analytical reagent grade.

Sampling sites

Soils samples used in the study were as previously described: vertical profiles of soil from a series of glacial moraines in the Guil river valley in the French Alps, (G1, G2, G3a, G9, and G11) and the Po river valley in the Italian Alps (V10) were sampled by their soil horizon classification (Mahaney et al., 2016).

Soil sample collection

Samples were collected as previously described in Mahaney et al. (2016). A map of the sampling site is reproduced in the supplementary section (S6) from Young et al. (2019). Briefly, samples were collected from the mid-point depth of soil horizon profiles for specific horizons (e.g. Ah, Bw, Cox). So for a soil horizon of 5cm depth, samples were collected 2.5 cm into the horizon. The aim was to obtain samples that were representative of each horizon. Samples were collected using sterile implements and storage vials. They were immediately stored on ice, in the dark, throughout transport to the laboratory. They were then frozen at -20^{o}C before further analysis.

Metagenomic library construction

DNA was extracted from alpine paleosol samples using a PowerSoil DNA extraction kit (Mo Bio). For each extraction 0.25g of soil was used, extractions were performed in triplicate for each sample and pooled. Extractions were performed according to the manufacturer's protocol with the following modifications: samples were homogenised using a FastPrep 120 cell disrupter system (Thermo-Fisher) at 5.5 $m.s^{-1}$ for 2 minutes, rather than a standard benchtop vortex. The eluted DNA was further purified *via* two rounds of ethanol precipitation; DNA solution was suspended in 3 volumes of ice-cold 100% ethanol, 0.1 volumes 5 M sodium acetate solution (pH 5.2) and 2 µL of linear polyacrylamide (LPA), the solution was then incubated overnight at -20°C and centrifuged at 18000 x g at 4°C for 30 minutes. The supernatant was discarded, and the pellet washed in 70% ice-cold ethanol and again centrifuged at 18000 x g at 4°C for 5 minutes. Finally, the supernatant was discarded and the pellet allowed to air dry for 15 minutes before resuspension in 50 µL molecular grade H_20. Final DNA concentrations were measured using a Quantus Fluorometer (Promega) in conjunction with the Quantiflour dsDNA dye system (Promega).

DNA sequencing libraries for Illumina sequencing prepared using a Nextera NX library preparation kit (Illumina), DNA libraries were sequenced on an Illumina MiSeq DNA sequencer in the paired-end mode for 300 cycles. Library preparation and sequencing were performed at the University of Cambridge

DNA sequencing facility. Raw Sequence data files are available in the NCBI sequence read archive under Bioproject number PRJNA39461.

Metagenomic library screening and sequence analysis

Raw reads were quality trimmed to a minimum mean quality of 20 over a sliding window of 10 bases using bbduk from the bbmap package (Bushnell, 2015). Reads from all 18 samples were co-assembled using Megahit (Li et al., 2016), contigs < 1kbp were discarded, and open reading frames were identified and extracted using Prodigal (Hyatt et al., 2010) with the –meta flag. Resulting in a library of 1154087 ORFs. Putative bacterial laccase sequences were classified from this library of ORFs using previously published profile hidden Markov models for bacterial laccases (Ausec et al., 2011). Only hits with an e-value < 1e-10 were retained, these were manually checked against the NCBI conserved domain database to ensure they possessed the cupredoxin domain characteristic of laccases. The final library of laccases was taxonomically classified using Kaiju (Menzel et al., 2016), against a database of all proteins from all bacteria, archaea, single-celled eukaryotes and viruses in the NCBI non-redundant protein database (NCBI Resource Coordinators, 2017). Raw reads were mapped onto the library of putative laccase genes using bbmap with the flags with the flags k=13 vslow=t. Normalised relative abundances were calculated as TPMS (transcripts per million) as this is the most robust and appropriate method for comparisons of gene sets between samples described by Wagner et al. (2012) and is applicable to metagenomic as well as RNA-Seq data.

In order to isolate novel type B two-domain laccase for downstream expression and characterisation, primers were designed for three type-B two-domain laccases for which an entire ORF including upstream ribosomal binding site was captured in the assembly (**supplementary data S1**). However, only one of these primer sets yielded PCR products (mg_lacc_fw: 5’-ATGAGCAATAGACGTGGCTT-3’, mg_lac_rev: 5’-ATTCATGTTTGTGCTCCATCTC-3’).

PCR was performed using these primers with total extracted DNA from the three Viso 10 site samples, with the following conditions: initial denaturation 95°C for 3 minutes; denaturation 95°C for 30

seconds; annealing 53°C for 90 seconds; extension 72°C for 1 minute; final extension 72°C for 7 min; 30 cycles. DNA gel electrophoresis confirmed fragments of the expected size (1500bp), these were excised from the gel and purified using a GeneJET gel extraction kit (Thermo-Fisher). Purified PCR products were ligated into PJET 1.2/blunt vector using a CloneJET PCR cloning kit (Thermo-Fisher), and transformed into *E.Coli* BL21 (DE3) competent cells by heat shock transformation. Transformed clones were grown overnight on 100 μg mL^{-1} ampicillin LB agar plates at 30°C. Individual colonies were selected and grown overnight in 100 μg mL^{-1} ampicillin LB broth, Plasmids were extracted from each broth culture using a GeneJET plasmid extraction kit (Thermo-scientific), extracted plasmids were sequenced using an Applied Biosystems 3730 DNA analyser at the University of Dundee DNA sequencing facility, yielding a final library of 8 sequenced PCR products of the expected insert size.

Statistical analysis of data

Metagenome sequence data were analysed using the same methods to those described in our earlier report – the identical metagenome dataset was used in the present study (Young et al., 2019). Statistical analysis were carried out in R version 3.4.1 (R CoreTeam, 2011), running on an Ubuntu 16.04 LTS PC with 16 GB RAM and an Intel core i7

6700K processor. The reader is referred to our earlier study for details of methodology used.

Laccase sequence dissimilarity was compared using Krustal-Wallace ANOVA analysis performed using the BASE R: kruskal test package (R Core Team, 2011, Kruskaland Wallis, 1952).

The analysis of 26 abiotic variables – previously described by Mahaney et al. (2016) –a Euclidean distance dissimilarity matrix of soil geochemical variables was produced using the package Vegan (Dixon, 2003) in R. v3.4.1 [R Core Team, 2011] using permutational analysis of variance. Principle component analysis was performed on this matrix using the vegan function cap scale, two-dimensional ordination plots were produced using the first two principle components as axes. For statistical comparison of laccase, the laccase genes were identified as being of 5 distinct types (see

Figure 3 and references in the legend). The laccase genes identified in each sample were annotated using the package Kaiju (Menzel et al., 2016). Statistical analysis of each enzyme subgroup across the metagenome datasets was performed using Kruskal-Wallis one way ANOVA (Kruskal & Wallis, 1952).

Maximum likelihood phylogenetic trees were drawn using Mega 7.0 (Kumar et al., 2016). The reference sequences shown were downloaded from the NCBI database. ***Cloning and heterologous expression of novel laccase genes***

The laccase genes were then sub-cloned into pLATE11 (untagged) ligation independent cloning (LIC) and expression vector without their signal peptides. Genes without their signal peptides were obtained by designing new forward primers that excluded the 26 amino acids N-terminal signal peptide during amplification. The pJET 1.2/blunt plasmids containing laccase genes were used as the templates for amplification of laccase genes without their signal peptides. Each Lac-pLATE11 recombinant was transformed into *E. coli* BL21 (DE3), and transformants were selected on LB agar plates supplemented with 100 μg mL^{-1} ampicillin.

For the expression of recombinant laccases, LB (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 5 g L^{-1} NaCl) medium was used (Bertani, 1951). Expression studies were performed in shake flasks; each culture was inoculated with individual overnight pre-cultures (grown at 30 ^{o}C, 140 rpm) and incubated at 30 ^{o}C and 160 rpm until the cells grew to a mid-exponential phase (OD_{600} ≈ 0.6). At mid-growth phase, the temperature and rotation speed were downshifted to 25 ^{o}C and 100 rpm respectively; and the addition of 0.1 mM IPTG induced the expression of recombinant laccases. The culture was also supplemented with copper ions in the form of $CuSO_4$ to a final concentration of 0.25 mM. Oxygen-limiting condition (microaeration) was introduced by switching off rotation (0 rpm) 4 h after induction; static incubation has been reported to increase the amount of copper-loaded bacterial laccases in *E. coli* (Durão et al., 2008). The static incubation was maintained for 20 h, and then cells were harvested by centrifugation (6000 rpm, 15 min). The recovered cell pellets were frozen at -80 ^{o}C.

In order to recover the proteins, cells were thawed and lysed by resuspending in CelLytic B solution (Sigma-Aldrich) containing 0.5 U mL^{-1} Benzonase Nuclease (Novagen) and 0.2 mg mL^{-1} lysozyme. The mixture was incubated at room temperature with shaking (200 rpm) for 20 min according to manufacturer's instructions. After the soluble fraction was separated from the cell debris by centrifugation (18000 x g, 15 min), it was heated at 70 ^{o}C for 15 min to precipitate most of the *E. coli* proteins as a form of partial purification. The thermal activation of laccases has been reported in several studies (Mollania et al., 2018; Papinutti et al., 2008; Koroleva et al., 2001). Samples collected throughout the expression studies including the cell-free extract (CFE) were mixed 1:1 with SDS-PAGE loading buffer, denatured at 95 ^{o}C for 5 min, and electrophoresed on a 10% SDS-PAGE resolving gel. The concentration of protein was estimated using the Bradford assay with bovine serum albumin (BSA) as the standard.

Enzyme assay

The activity of laccases was measured at 25 ^{o}C using ABTS (ε_{420nm} = 36 mM^{-1} cm^{-1}) and syringaldazine (ε_{530nm} = 65 mM^{-1} cm^{-1}) as substrates with a 6705 UV/Vis spectrophotometer (Jenway). Each assay mixture was made up to 1 mL and contained 2 mM ABTS (dissolved in phosphate buffer) or 0.05 mM SGZ (dissolved in absolute methanol), 100 µL of cell-free extract (CFE) and phosphate buffer at pH 4.0 or pH 6.5 for ABTS and SGZ assays respectively. The fungal laccase from *Trametes versicolor* (Sigma Aldrich) was used as a positive control while cell-free extract prepared from an expression experiment with an empty vector was used as a negative control. The oxidation of both ABTS and SGZ was followed by a change in absorbance at 420 nm and 530 nm respectively for 10 min. One unit (U) of laccase activity was defined as the amount of laccase required to oxidise 1 µmol of substrate per minute (1 µmol min^{-1}). All assays were performed in triplicate.

3. RESULTS

3.1 THE STRATIGRAPHY OF SAMPLING SITES IS CONSISTENT WITH SOIL HORIZONS REPRESENTING DISTINCT SETS OF ENVIRONMENTAL PARAMETERS THAT ARE INDEPENDENT OF GEOGRAPHIC SEPARATION.

Previous studies at this site in the Guil valley and Mount Viso region have classified the sampled paleosols as inceptisols with specific soil pedons as defined in the geological literature (Mahaney et al., 2016; Young et al., 2019). While there is a long established convention of describing glacial sediment deposits in terms of Ah, Bw, Cox, Cu, and D horizons - based upon variation in physical and chemical distinctions down the soil profile - there is no detailed attempt to date, to link differences in extra-cellular microbial enzyme distribution to this classification.

We initially used data gathered in an earlier paper (Mahaney et al., 2016) to show that the 27 chemical and physical parameters previously measured in each of the 18 samples – 6 each being designated as Ah, Bw and Cox – did indeed indicate that soil Horizons represent distinct physical/chemical environments. Soil horizons were chemically and physically distinct as indicated by Principle Coordinates Analysis, wherein the first two principle co-ordinate axes explained over 70% of the total variation between samples (Figure 1).

3.2 THE DISTRIBUTION OF BACTERIAL LACCASE GENES IS RELATED TO SOIL STRATIGRAPHY

We searched a library of 1154088 ORFs from all the paleosol samples sequenced from a metagenomic co-assembly of 135365644 raw reads for putative bacterial laccases using previously published profile hidden Markov models for bacterial laccases (Ausec et al., 2011), resulting in 323 new putative laccase-like ORFs being identified. Of these, there were many that include both start and stop codons. The amino acid translation of these sequences were additionally aligned against the NCBI nr protein database revealing that the majority of these sequences had low sequence identity but high sequence

similarity to known LMCO genes [**supplementary data table S2**]. This analysis highlighted the abundance of novel LMCO sequences in these soils.

We hypothesised that the abundances of different classes of bacterial laccases would vary significantly between stratified soil horizons, as their functional role as extracellular biocatalysts would inevitably be linked to localised environmental parameters. In order to test this hypothesis, we applied Kruskal Wallace one way ANOVA (Kruskal & Wallis, 1952) to the relative abundance of putative laccases grouped by soil horizon **(Figure 2),** revealing that all but two classes of bacterial laccases vary significantly in abundance between soil horizon. The type B two domain laccases appear to increase in abundance down a soil profile and show the strongest variation with soil horizon, while the big three domain, small three-domain and CotA –type laccases all decrease in abundance with soil depth. It should be noted that, in terms of relative abundance, the two most abundant classes of enzyme in these samples are the small three-domain and type B two domain laccases (**Figure 3**). These two groups comprise >70% of all laccase-like genes identified.

The predicted taxonomic affiliations for putative laccase ORFs shows strikingly different community structures between laccase types (**Supplementary Figure S4**). Additionally, principal coordinate analysis of the predicted taxonomic affiliations shows clear differences between the Ah Horizon and Bw/Cox horizons for the type B two domain, small three domain and CotA type laccases (**Figure 3**). Indicating that the distribution of laccase genes in the soil horizons is affected by the soil horizon location as well as by the phylogeny of any expressing bacterial cell. Critically, the two most abundant groups of identified genes show the greatest demarcation between soil horizons in these PCA plots. Where relatively few genes were identified in the horizons (e.g. the 'Big 3D' group) there is no clear demarcation. It is important to note here that given the fragmented nature of the assembly from which these ORFs were derived, that we cannot deduce which organisms may be harbouring multiple laccase gene types, although such organisms should certainly exist.

3.3. Further analysis of typeB two-domain laccases, including expression and activity.

Three sets of primers were designed to amplify novel type-B two-domain laccases for which a complete ORF including start and stop codons, as well as a signal peptide (indicating potential extracellular secretion), could be identified. These primers were all tested against multiple samples from the paleosols, to see if laccases could be readily amplified. We found only one set of primers gave reliable amplification and only when using DNA extracted from the V10 Bw and Cox samples. These primers were designated [V10_lac]. Cloning and sequencing of these PCR products yielded eight high-quality sequences of the expected insert size; these showed high homology with the assembled parent sequence; these sequences are compared in a phylogenetic tree (**Figure 5**). Interestingly, none of the PCR amplified sequences were 100% identical to the parent sequence, this may be attributed to the fact that a contiguous sequence derived from a typical metagenomic assembly is often a consensus sequence of closely related strains and variants present in the dataset, other factors may be involved including PCR error, variants introduced during exponential growth of the cloning vectors and sequencing error (both Illumina and Sanger sequencing). Despite this, in this case, it appears that the assembled gene sequence was sufficiently accurate to allow the design of PCR primers which could be used to isolate functionally active proteins from a portion of the soil samples.

3.4 Characterisation of cloned laccases

There is a limited detailed biochemical analysis of two domain laccase enzymes from Acidobacteria in the literature, despite the importance of these bacteria in soil (Jones et al., 2009). Further, there is just one account of a metagenome laccase analysed to date (Ausec et al., 2017). To validate the metagenomic enzyme discovery pipeline used in this study, we selected four Acidobacteria two-domain laccase enzymes from our earlier gene mining activities (**Supplementary S3**) that were cloned and expressed for biochemical characterisation. The aim here was to demonstrate bona fide laccase like activity for the identified gene products.

In all the laccases tested, the enzymes showed higher activity against SGZ than ABTS (Figure 6). This observation has been made with other laccases (Guan et al., 2014; Lu et al., 2013; Ye et al., 2010). Also, enzymes mined from the Bw horizon (LacB3 and LacB6) showed more activity against both substrates tested than enzymes from the C horizon (LacC12 and LacC15).

This experiment confirmed that complete functional genes for the typeB 2D group could be isolated from the paleosols. It effectively validated both the use of our metagenomics methods for novel laccase gene discovery and complete gene isolation, from these paleosols. It also validated the designation of our metagenome analysis methods for the identification of putative laccase genes.

Further experimentation confirmed that these enzymes were in fact functionally distinct laccases. For example, copper dependence is a key characteristic of laccases, and we showed that one of the enzymes (LacB6, that had the highest activity with both substrate of the four clones) required $CuSO_4$ addition for full activity (**Figure 7**). We also showed that the activity of LacB6 against syringaldazine increased with increasing concentration of copper. The addition of 10-, 50- and 100 µM $CuSO_4$ in the assay mix increased the enzyme activity 2.5-, 5, and 8-fold respectively. These observations confirm that LacB6 is a copper-dependent enzyme and that copper was only partially incorporated into the enzyme's active site during expression. Cloned and expressed laccases are usually partially copper-depleted as a result of the difficulty of the passage of copper ions from the medium into the cell during expression (Harris, 2000; Vulpe and Packman, 1995).

Molecular modelling was performed on the Laccase_B_6 amino acid sequence using three publically available protein modelling servers, ITASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/), Phyre2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) and Swiss-model (https://swissmodel.expasy.org/). All three tools found the closest available crystal structures for homology modelling had between just 14-32% amino acid identity to the query sequence (data not shown). This suggests that homology models produced for this protein sequence are unlikely to be representative of the true structure.

4. DISCUSSION

The relationship between soil stratigraphy nomenclature and abiotic chemical/physical variables has been long established in the geological literature – and has also previously been considered in the context of Alpine glacial deposition (Birkland et al., 1979). In this study, we were able to confirm a clear differentiation between the designated soil horizon samples previously described by Mahaney et al. (2016) and the physical/chemical composition of the samples using PCA plots. The data form a framework for later comparison with the distribution of laccases in the same samples.

Our main objective in this study was to see if there was a connection between soil pedon classification and the distribution of bacterial laccases in a model soil ecosystem. We chose to look at bacterial laccases for a number of reasons: i) they are enzymes with diverse function. Some are required for intracellular endospore germination/formation processes, such as the Cot A proteins (Francisco et al., 2004), while there is clear evidence that the small three domain and typeB two-domain enzymes are extracellular enzymes are used for lignin degradation (Bugg et al., 2004); ii) The extracellular bacterial laccases are secreted by one taxon into the environment, but then can also benefit other components in a microbial ecosystem such as soil. As such, we might expect the distribution of these genes to be related to cooperative behaviours between different bacterial taxa in localised communities (West et al. 2006); iii) As potential industrial biocatalysts they are particularly useful, while also being structurally simple as single component enzymes – making them good candidates to demonstrate the potential for paleosol metagenome-mining as a source of useful functional genes (Call & Mucke, 1997).

Looking at the differences in the distribution of the genes encoding five different 'classes' of enzymes, that are already established in the literature, two key observations become evident: Firstly, there does appear to be present distinct sub-populations of laccase enzymes in soil pedons from our data. Not only do the relative abundance of laccase genes in the stratified soils clearly demarcate between Ah, Bw and Cox soil horizons, but we also observed that the beta diversity

between populations of laccase genes in soil horizons is suggestive of greater dissimilarity of gene populations between soil horizons in any individual paleosol than between geographical locations – even when these locations are kilometres apart (such as between the site V10C and any of the G sites studied here). This is especially evident for the typeB two-domain and small three-domain enzyme groups that were generally abundant in all the soil pedons (**Figure 3**). Clearly, from the perspective of bacterial laccase gene function and phylogenetic diversity, we can say that Ah and Bw soil horizon - though just separated physically by centimetres - are fundamentally different environments . One explanation we can propose for this marked difference in laccase gene diversity across soil horizons down profile is their potential importance as 'public good' proteins in complex bacterial communities of heterotrophic bacteria that rely upon the degradation of plant-derived lignin as a carbon and energy source. The principal component analysis of the 27 physical parameters measured in the soil horizons (**Figure 1**) showed clear demarcation of the three types of soil horizon across all 18 soil pedons. However, differences between the populations of the most abundant typeB two-domain laccases and the small three-domain laccases across the boundary of the Ah and Bw/Cox pedons follow a different distribution: clear dissimilarity is evident between Ah and Bw gene populations for PCoA1, whereas no dissimilarity is seen between the populations of these genes in the Bw and Cox horizons. This would support a hypothesis that biological and perhaps even societal factors consistently lead to polarisation of these groups of enzymes between Ah and Bw horizons. This conclusion is certainly speculative at this stage, but it does highlight the possibility of using paleosol horizon metagenome analysis of such 'public goods protein' functional genes for the study of developing field of bacterial sociobiology (Griffin et al., 2004). Critically, the different classes of laccase studied do not partition in an identical fashion between the pedons. For example, the CotA type three domain laccase gene populations differentiate as well between Bw and Cox soil horizons as they do between Ah and Bw horizons from the PCA analysis (**Figure 3**). Furthermore, the fact that CotA enzymes are not likely to be public goods proteins could be important. Enzymes may

have the same functionality but different roles in a bacterial community, and laccases in particular may serve a community function in soil microbial ecosystems (Freedman and Zak, 2015).

As initially noted, the use of metagenome gene-mining as a process to source new industrial enzymes is an important possibility for the generation of new biocatalysts. We wanted to demonstrate here that with our paleosol data set we could successfully obtain, clone and validate expression of novel typeB two-domain laccase enzymes. We chose to isolate genes for the most abundant 2D two component enzymes – that have not been studied extensively as biocatalytic enzymes – due to their relative novelty. The methodology developed here shows clearly that such enzymes can be readily obtained and assayed for enzyme activity. The observed activity against model substrates is consistent with other extracellular bacterial laccases of biotechnological importance (Couto et al., 2006).

5. Acknowledgements

We thank the REMEDIATE ITN project (BPK, CCRA), Invest Northern Ireland (JMY) and the Commonwealth Scholarships Commission (SUD) for financial support. We also acknowledge a reviewer of the manuscript for useful suggestions.

6. SUPPLEMENTARY DATA

S1: Complete laccase ORF for which PCR primers were designed. This was assembled using the Megahit assembly approach described in the methods.

ATGAGCAATAGACGTGGCTTTTTGCGAAATATTCTCGCCGGAGCTGGCGCTATGGTTTCCGCGAGAGTTCTTT
CAGCCCAAGAGACGGACACGTCTCATGGGATGCACGGAATGAAAATGAAAGGTAGGCAGGAGTCGGAGCGC
GGCCACGGTTCTCCCGTGCTTATGGAAACGCCGGACGTTCCGCAACTTCCGTGGCGCATGGATGGGGGCGTA
AAGGAATTCCATCTCATTGCCGAGCCGGTGAAACAGGAGATTTTTCCTGGCCGGATCGTTGACCTCTGGGGTT
ACAACGCCAGCGCTCCGGGGCCGACGATTCAAGTCAACCAGGGCGATCGCGTTCGGATCATCGTGGACAACC
ACCTTCCCGAAGCGACGTCGATGCACTGGCACGGCTTCGAAATCCCGAGCGAGATGGACGGAGCGCCCGGTT
CGAGTCAAGACCCGATCCCTCCCGGGGGCCGCTTTGTCTATGAATTCACGCTCCACCAGGAAGGCACTTATTTC
TACCACTCGCACATGGCGATGCAGGAAATGATGGGCATGATCGGCGCGTTCATCATGCATCCCAAGGAGCCA
TACAAGCCGCGCGCCGATAAAGACTTCGCCATCATTATGCAGGAGTACGCCATCTTGCCGAACATCAAGGTCC
CGAACTCCATGAACATGGAGTTCAACTGGCTTACCTTCAACGGAAAATCCGGTCCAGCCAACACTCCGCTCATT
GTGCGGCACGGAGAGCGCGTCCGCATCCGCCTTATCAATCTGGGGATGGACCACCATCCGATCCACCTTCATG
GGCACCAGTTCGTAGTTACAGGTACTGAAGGAGGCCGACAGCCCGAAACGACTTGGGGGCCGGGCAATACT
GTTCTCGTTGGAGTCGCACAGTCCCGCGACGTGGAGTTCGTTGCCAACAACCCCGGCGACTGGATGCTGCACT
GCCATCTTCCGCACCACATGATGAACCAGATGTCATCGATGGTCGGCCCGATGTCACGCCGAAATGGAATGCC
TGCGGGTCTCGACATGGAGCGGGGCATGGGAATGCTGCGACAAGGAAGCGCGACGTCTGAGGAGAACGGT
CCGAGCCTGGGTCGCGGAATGGGCGTCGGTTCGACAGCGGAGCAAACCATGTCGAATTCTCCGCTCAAAGCC
GGGAGTCCGATGCAGCACCAAGACATGCCCGACATGCAGCAGCAAGGGATGCAGGGGAGGCAGCCCGACAA
GCGCGATGACACGAAGGACGCGCATTGCTTGCCGGAGGATGCCTTGGTGGAAGGTATCATGATGGCGATGG
ATCAAATGGTCGACAAACCTGAAAACTTCGGCCTGCGTCCCGGCTGGAGCGGCTTCATGGCAGGAATGATGA
CGTTCGTACGCGTGCTGCCTCAAGACAAGTACGACCACAGCATGGAGCTTCGAAAAAAGCAGGAAGGCAATA
AGCCAATGAAAATGGACATGCCGGAGATGGAGCACAAACATGAATAA

S2: Sequence statistics from total co-assembly and isolated laccase ORFS. Statistics are shown for both the complete co-assembly and the putative Laccase ORFs. Laccase hits were identified using the HMM search pipeline as described in the text.

	All_contigs	Laccase_hits
# contigs (>= 0 bp)	7171198	323
# contigs (>= 1000 bp)	540891	151
# contigs (>= 5000 bp)	2869	0
# contigs (>= 10000 bp)	165	0
Total length (>= 0 bp)	4127121895	316566
Total length (>= 1000 bp)	804814220	199008
Total length (>= 5000 bp)	19157720	0
Total length (>= 10000 bp)	2202670	0
# contigs	3220893	283
Largest contig	62040	2877
GC (%)	62.22	61.55

N50	775	1143
N75	614	912
L50	1102891	106
L75	2045807	180

S3: Complete amino acid sequences of gene-mined laccases that were shown to have activity. These were obtained using PCR amplification using the V10_las primers as described in the text.

>LacB3

MSNRRGFLRNILAGAGAMASAKVLSAQEMHMSHGTQGMKGMKETERGHVSPTLVETPDVS
QLSWRMGGSVKEFHLIAEPVKQEIFPGRIVDLWGYNGSAPGPTIQVNEGDRVRIIVDNHL
PEATSMHWHGFEIPSAMDGAPGSSQDPIPPGGRFVYEFTLHQEGTYFYHSHMAMQEMMGM
IGAFIMHPKEPHKPRADKDFAIIMQEYAILPNIKVPNSMNMEFNWLTFNGKAGPATTPLI
VRHGDRVRVRLINLGMDHHPIHLHGHTFVVTGTEGGRQPQSTWSPGNTVLVGVAQSRDVE
FVATNPGDWMLHCHLPHHMMNQMSSMVGPMSRRNGTPAGLDMERGMGMLRQESATSEENG
PSLGRGMGMGSTAEQTVSNSPLKAVNPMQHKDMPNMQQPAMQMGKPDVSKDAISVPGFPQ
DAFMEGPMMAMDQMVDKPENFADPPGGSGPTAGTMTFVRVPPPDNYDQIMEHRKKQEGNN
PLQMDMPQMEHNQA

>LacB6

MSNRRGFLRNILAGAGAMVSARVLSAQETDTSHGMHGMKMKGTRESESGHGSPMLVEMPD
VAQLPWRLDGDVKEFHLIAEPVKQEIFPGRIVDLWGYNGSVPGPTIQVNQGDRVRIIVDN
HLPEATSMHWHGFEIPNEMDGAPGSSQDPIPPGGRFVYEFTLHQEGTYFYHSHMAMQEMM
GMIGAFIMHPKEPYKPRADKDFAIIMQEYAILPNIKVPNPMNMEFNWLTFNGKAGPATTP
LIVRHGERVRIRLINLGMDHHPIHLHGHQFVVTGTEGGRQPQSTWGPGNTVLVGVAQSRD
VEFVASNPGDWMLHCHLPHHMMNQMSSMVGPMSRRNEMPAGLDMERGMGMLRQGSATSEE
NGPSLGRGMGVGSTAEQTMSNSPLKAENPMQHQDMPDMQPQGMQMGNPDVSKDANSVPGF
PQDAFMEGPMMAMDQMVDKPENFGLRPGWSGFMAGMMTFVRVLPQDKYNQIMELRKKQEG
NKPMKMDMPEMEHKHE

>LacC12

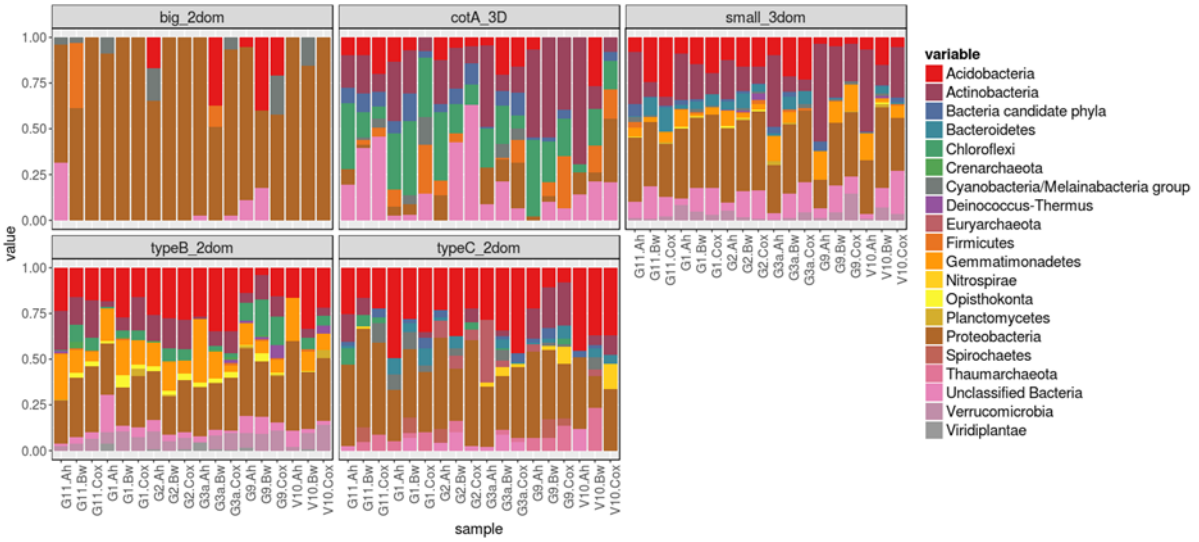
MSNRRGFLRNILAGAGAMASAKVLSAQEMDMSHGMQGMKMKGVKKSEHAHASRMLVETPD
VAQLPWRMEGNVKEFHLIAEPVKQEIFPGRVVDLWGYNGSAPGPTIQVNQGDRVRIIVEN
RLPEATSMHWHGFEVPNEMDGAPGSSQDPIPPGGRFVYEFTLHQEGTYFYHSHMAMQEMM
GMIGAFVMRPKQAYRPRVDHDFAIILQEYAILPNISVPNSMNMEFNWLTFNGKAGPATTP
LIVRQGERVRIRLINLGMDHHPIHLHGHQFVVTGTEGGRQPESTWGPGNTVFVGVAQSRD
IEFVASNPGDWMLHCHLPHHMMNQMSSNVGPMTRRNGIPAGLDMERGMGMLRQGSATSEE
HGPSLGRGMGVGSTFEQRMSNSPLKAGSPMQHQGMQMGKPDVSKDANSVPGFPQDAFMEG
PMMAMDQRVDKPENFGLRPGWSGFMAGMMTFVRVLPQDKYDHIMELRKKQEGNKPMKMDM
PEMEHKHE

>LacC15

MSNRRGFLRNILAGAGAMASAKVLSAQEMEMSHGMQGMKMKGVKKSERDHASRMLVETPD
VAQLPWRMDGSVKEFHLIAEPVKQEIFPGRVVDLWGYNGSVPGPTIQVNQGDRVRIIVDN
HLPEATSMHWHGFEIPNEMDGAPGSSQDPIPPGGRFVYEFTLHQEGTYFYHSHMAMQEMM
GMIGAFIMHPKEPYKPRSDKDFAIIMQEYAILPNIKVPNSMNMEFNWLTFNGKAGPATTP
LIVRHGDRVRVRLINLGMDHHPIHLHGHQFVVTGTEGGRQPQSTWGPGNTVLVGVAQSRD
VEFVASNPGDWMLHCHLPHHMMNQMSSMVGPMSRRNGMPAGLDMERGMGMLRQGSAMSEE
NGPSLGRGMGVGSTAEQTMSNSPLKAGNPMQHQAMPNMQQQGMKMGKPDVSKDANSVPGF

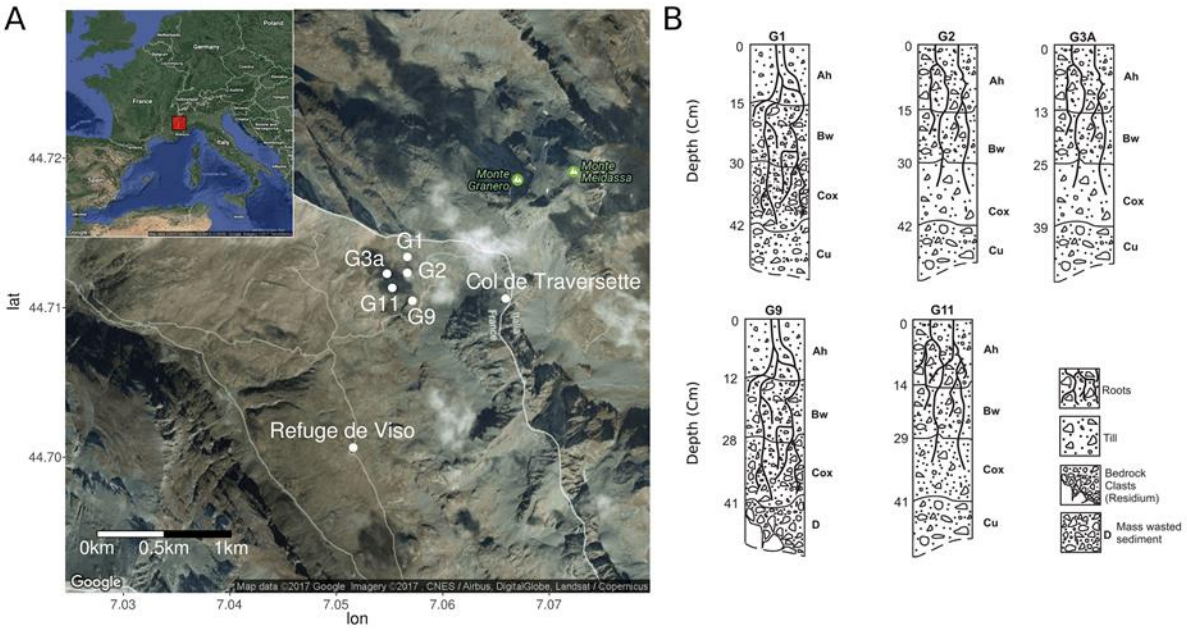
PQDAFMEGPMMAMDQMVDKPENFGLRRGWSGFMAGMMTFVRVLPQDKYDHIMEPRKKQEG
NKPMKMDMPEMEHKHEY

S4: Laccase taxonomic distribution by type Phylum level



Taxonomic profile for each class of laccase genes Isolated. Laccase gene sequences were taxonomically annotated using Kaiju (Menzel et al., 2016), as described in the main methods section of the paper. See figure 4 for origin of the five laccase type designations.

S.5. Map of sampling sites used in the study (from Young et al., 2019).



Overview of Sampling sites and soil stratigraphy. Panel A: Satellite map of sampling area. Panel B: Stratigraphy of soil profiles for each sample (from Mahaney et al., 2016). Sampling sites are shown at

467 the top of each profiles, depth in centimetres is displayed on the left and soil horizon classification is
468 displayed on the right of each profile.

469

470 7. REFERENCES

471

472

473 Ausec, L., Zakrzewski, M., Goesmann, A., Schlüter, A., Mandic-Mulec, I., 2011. Bioinformatic analysis
474 reveals high diversity of bacterial genes for laccase-like enzymes. PLoS ONE. 6(10), e25724

475 Ausec, L., Berini, F., Casciello, C., Cretoiu, M.S., Van Elsas, J.D., Marinelli, F. and Mandic-Mulec, I., 2017.
476 The first acidobacterial laccase-like multicopper oxidase revealed by metagenomics shows high salt
477 and thermo-tolerance. Applied Microbiology and Biotechnology, 101(15), pp. 6261-6276.

478 Bertani, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli.
479 J. Bacteriol. 62, 293-300

480 Birkland, P.W., Colman, S,M., Burke, R.M, Shroba, R.R., Meierding, T.C., 1979. Nomenclature of alpine
481 glacial deposits, or, What's in a name? Geology, 7 ,pp. 532-536.

482 Bugg, T.D.H., Ahmad, M., Hardiman, E.M., Singh, R., 2011. The emerging role for bacteria in lignin
483 degradation and bio-product formation. Current Opinion in Biotechnology, 22(3), pp. 394-400.

484 Bushnell B. (2015). BBMap short-read aligner, and other bioinformatics tools. Available from
485 sourceforge net/projects/bbmap.

486 Call, H.P., Mucke, I., 1997. History, overview and applications of mediated lignolytic systems, especially
487 laccase-mediator-systems (Lignozym®-process). Journal of biotechnology. 53 (2–3), pp. 163-202

488 Chauhan, P.S., Goradia, B. And Jha, B., 2018. Optimization and up scaling of ionic liquid tolerant and
489 thermo-alkali stable laccase from a marine Staphylococcus arlettae S1-20 using tea waste. Journal of
490 the Taiwan Institute of Chemical Engineers, 86, pp. 1-8.

491 Chauhan, P.S. And Jha, B., 2018. Pilot scale production of extracellular thermo-alkali stable laccase
492 from Pseudomonas sp. S2 using agro waste and its application in organophosphorous pesticides
493 degradation. Journal of Chemical Technology and Biotechnology, 93(4), pp. 1022-1030.

494 Couto, S. R., Herrera, J.L.T., 2006. Industrial and biotechnological applications of laccases: a review.
495 Biotechnology advances 24(5), 500-513.

496 Cragg, S.M., Beckham, G.T., Bruce, N.C., Bugg, T.D., Distel, D.L., Dupree, P., Etxabe, A.G., Goodell, B.S.,
497 Jellison, J., McGeehan, J.E., McQueen-Mason, S.J., 2015. Lignocellulose degradation mechanisms
498 across the Tree of Life. Current Opinion in Chemical Biology. 1(29), pp.108-119.

499 Dixon, P., 2003. VEGAN, a package of R functions for community ecology. Journal of Vegetation
500 Science, 14(6), pp. 927-930.

501 Durão, P., Chen, Z., Fernandes, A.T., Hildebrandt, P., Murgida, D.H., Todorovic, S., Pereira, M.M., Melo,
502 E.P. and Martins, L.O., 2008. Copper incorporation into recombinant CotA laccase from Bacillus

503 subtilis: Characterization of fully copper loaded enzymes. Journal of Biological Inorganic Chemistry,
504 13(2), pp. 183-193.

505 Fang, Z., Li, T., Wang, Q., Zhang, X., Peng, H., Fang, W., Hong, Y., Ge, H., Xiao, Y., 2011, A bacterial
506 laccase from marine microbial metagenome exhibiting chloride tolerance and dye decolorization
507 ability. Applied microbiology and Biotechnology. 89(4), pp1103-1110.

508 Francisco, J.E., Marçal, D., Martins, L.O., Grenha, R., Henriques, A.O., Lindley, P.F., Carrondo, M.A.,
509 2004. Substrate and Dioxygen Binding to the Endospore Coat Laccase from Bacillus subtilis. Journal of
510 Biological Chemistry, 279 (22), pp. 23472–23476.

511 Freedman, Z.B. and Zak, B.R., 2015. Atmospheric N deposition alters connectance, but not functional
512 potential among saprotrophic bacterial communities. Molecular Ecology. 24(12), pp3170

513 Godornes, C., Leader, B.T., Molini, B.J., Centurion-Lara, A., Lukehart, S.A. 2007, Quantitation of rabbit
514 cytokine mRNA by real-time PT-PCR. Cytokine 38(1), 1-7

515 Griffin, A.S., West, S.A., Buckling, A. 2004, Cooperation and competition in pathogenic bacteria. Nature
516 430, 1024-1027

517 Guan, Z., Song, C., Zhang, N., Zhou, W., Xu, C., Zhou, L., Zhao, H., Cai, Y. and Liao, X., 2014.
518 Overexpression, characterisation, and dye-decolorizing ability of a thermostable, pH-stable, and
519 organic solvent-tolerant laccase from Bacillus pumilus W3. Journal of Molecular Catalysis B: Enzymatic,
520 101, pp. 1-6

521 Harris, E.D., 2000. Cellular copper transport and metabolism. Annual Review of Nutrition, 20, pp. 291-
522 310

523 Hyatt, D., Chen, G-L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic
524 gene recognition and translation initiation site iden0tification. BMC Bioinformatics 11, pp119.

525 Ihssen, J., Reiss, R., Luchsinger, R., Thöny-Meyer, L. and Richter, M., 2015. Biochemical properties and
526 yields of diverse bacterial laccase-like multicopper oxidases expressed in Escherichia coli. Scientific
527 Reports, 5.

528 Jacquiod, S., Demaneche, S., Franqueville, L., Ausec, L., Xu, Z., Delmont, T.O. et al., 2014.
529 Characterization of new bacterial catabolic genes and mobile genetic elements by high throughput
530 genetic screening of a soil metagenomic library. Journal of Biotechnology, 190(20), pp.18-29.

531

532 Jones, R.T., Robeson, M.S., Lauber, C.L., Hamady, M., Knight, R., Fierer, N., 2009. A comprehensive
533 survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. The ISME
534 Journal. 3, pp. 442–453

535 Koroleva, O.V., Stepanova, E.V., Binukov, V.I., Timofeev, V.P. and Pfeil, W., 2001. Temperature-
536 induced changes in copper centers and protein conformation of two fungal laccases from Coriolus
537 hirsutus and Coriolus zonatus. Biochimica et Biophysica Acta - Protein Structure and Molecular
538 Enzymology, 1547(2), pp. 397-407.

539 Kruskal, W.H., Wallis, W.A., 1952. Use of ranks in one-criterion variance analysis. Journal of the
540 Americam Statistical Association. 47: 583–621

Kumar,S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 33 (7), pp 1870–1874

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H., and Lam, T.W., 2016. MEGAHIT v1. 0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods, 102, pp3-11

Lu, L., Wang, T.-., Xu, T.-., Wang, J.-., Wang, C.-. and Zhao, M., 2013. Cloning and expression of thermo-alkali-stable laccase of Bacillus licheniformis in Pichia pastoris and its characterization. Bioresource technology, 134, pp. 81-86.

Machczynski , M.C., Vijgenboom, E., Samyn, B., Canters, G.W., 2009. Characterization of SLAC: A small laccase from Streptomyces coelicolor with unprecedented activity. Protein Science. 13(9), pp 2388-2397

Mahaney WC, Somelar P, Dirszowsky RW, Kelleher B, Pentlavalli P, McLaughlin S, et al. (2016). A Microbial Link to Weathering of Postglacial Rocks and Sediments, Mount Viso Area, Western Alps, Demonstrated through Analysis of a Soil/Paleosol Bio/Chronosequence. Journal of Geology, 124, 149–169.¶

Martins, L.O., Durão, P., Brissos, V., Lindley , P.F., 2015. Laccases of prokaryotic origin: enzymes at the interface of protein science and protein technology. Cellular and molecular life sciences. 72(5) pp.911-22.

McKnight, C.J., Briggs, M.S. and Gierasch, L.M., 1989. Functional and nonfunctional LamB signal sequences can be distinguished by their biophysical properties. Journal of Biological Chemistry, 264(29), pp. 17293-17297.

Menzel, P., Ng, K.L., Krogh, A., Marth, G., Lipman, D., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature Communications. 7: pp11257.

Mohammadian, M., Fathi-Roudsari, M., Mollania, N., Badoei-Dalfard, A. and Khajeh, K., 2010. Enhanced expression of a recombinant bacterial laccase at low temperature and microaerobic conditions: Purification and biochemical characterization. Journal of Industrial Microbiology and Biotechnology, 37(8), pp. 863-869.

Mollania, N., Heidari, M. and Khajeh, K., 2018. Catalytic activation of Bacillus laccase after temperature treatment: Structural & biochemical characterization. International journal of biological macromolecules, 109, pp. 49-56.

Muyzer, G., De Waal, E.C. & Uitterlinden, A.G. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology, 59(3), pp. 695-700.

Nakamura, K., Kawabata, T., Nobuhiro Go, K.Y., 2003. Novel types of two-domain multi-copper oxidases: possible missing links in the evolution. FEBS Letters, 553, pp1873-3468

NCBI Resource Coordinators, 2017. Database Resources of the National Center for Biotechnology Information. Nucleic Acids Research. 45: D12–D17

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MH, Wagner H (2011). vegan: Community ecology package. R package version. 2011, pp.117-118.

580 Papinutti, L., Dimitriu, P. and Forchiassin, F., 2008. Stabilization studies of Fomes sclerodermeus
581 laccases. Bioresource technology, 99(2), pp. 419-424.

582 R Development CoreTeam, 2011. A language and environment for statistical computing. R Foundation
583 for Statistical Computing, Vienna. http://www.R-project.org

584 Suzuki, M.T., Taylor, L.T. & DeLong, E.F. 2000, "Quantitative analysis of small-subunit rRNA genes in
585 mixed microbial populations via 5'-nuclease assays", Applied and Environmental Microbiology, 66(11),
586 pp. 4605-4614.

587 Torsvik, V., Jostein, G., Daae. F.L., 1990. "High diversity in DNA of soil bacteria." Applied and
588 environmental microbiology 56(3), pp. 782-787.

589 Vulpe, C.D. and Packman, S., 1995. Cellular copper transport. Annual Review of Nutrition, 15, pp. 293-
590 322

591 Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM
592 measure is inconsistent among samples. Theory Biosciences. 131, pp.281–285.

593 West, S.A., Griffin, A.S., Gardner, A., Diggle, S.P., 2006. Social evolution theory for microorganisms.
594 Nature Reviews Microbiology, 4, pp. 597-607.

595 Yang, Q., Zhang, M., Zhang, M., Wang, C., Liu, Y., Fan, X., Li, H., 2018. Characterization of a Novel, Cold-
596 Adapted, and Thermostable Laccase-Like Enzyme With High Tolerance for Organic Solvents and Salt
597 and Potent Dye Decolorization Ability, Derived From a Marine Metagenomic Library. Frontiers in
598 Microbiology. 9, pp.2998

599

600 Ye, M., Li, G., Liang, W.Q. and Liu, Y.H., 2010. Molecular cloning and characterization of a novel
601 metagenome-derived multicopper oxidase with alkaline laccase activity and highly soluble expression.
602 Applied Microbiology and Biotechnology, 87(3), pp. 1023-1031.

603 Young, J.M., Skvortsov, T,. Kelleher, B.P., Mahaney, W.C., Somelar, P., Allen, C,C,R, 2019. Effect of soil
604 horizon stratigraphy on the microbial ecology of alpine paleosols. Science of The Total Environment.
605 657, pp. 1183-1193

606

607

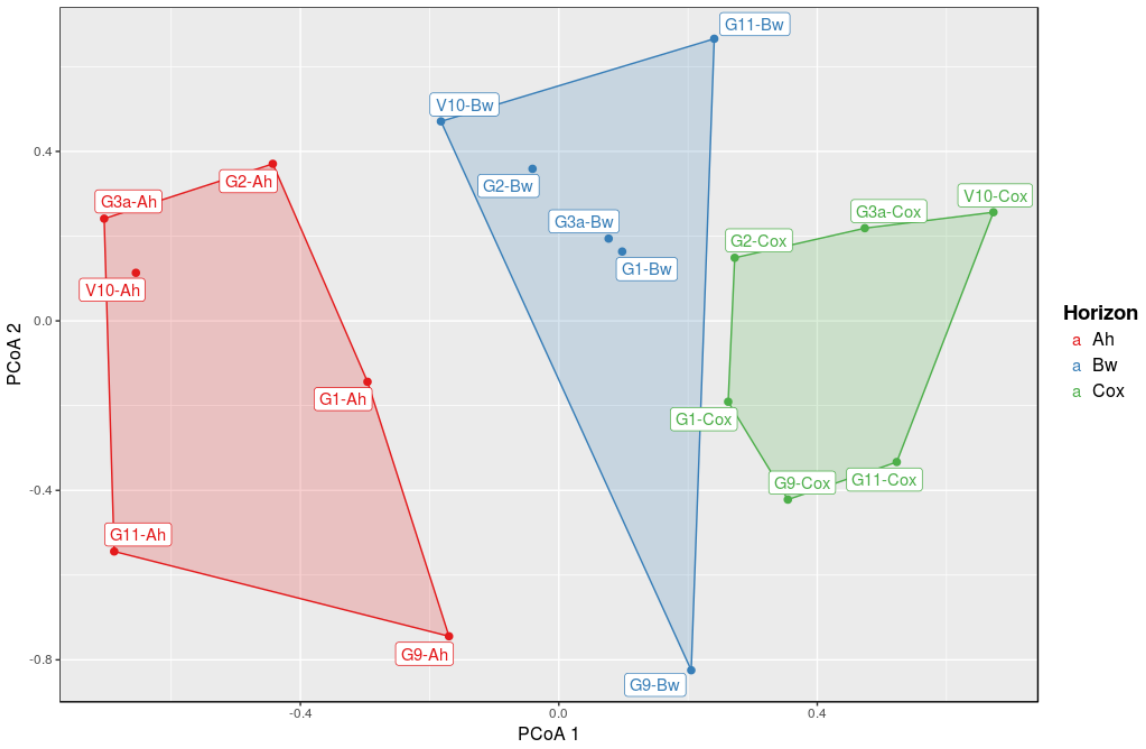
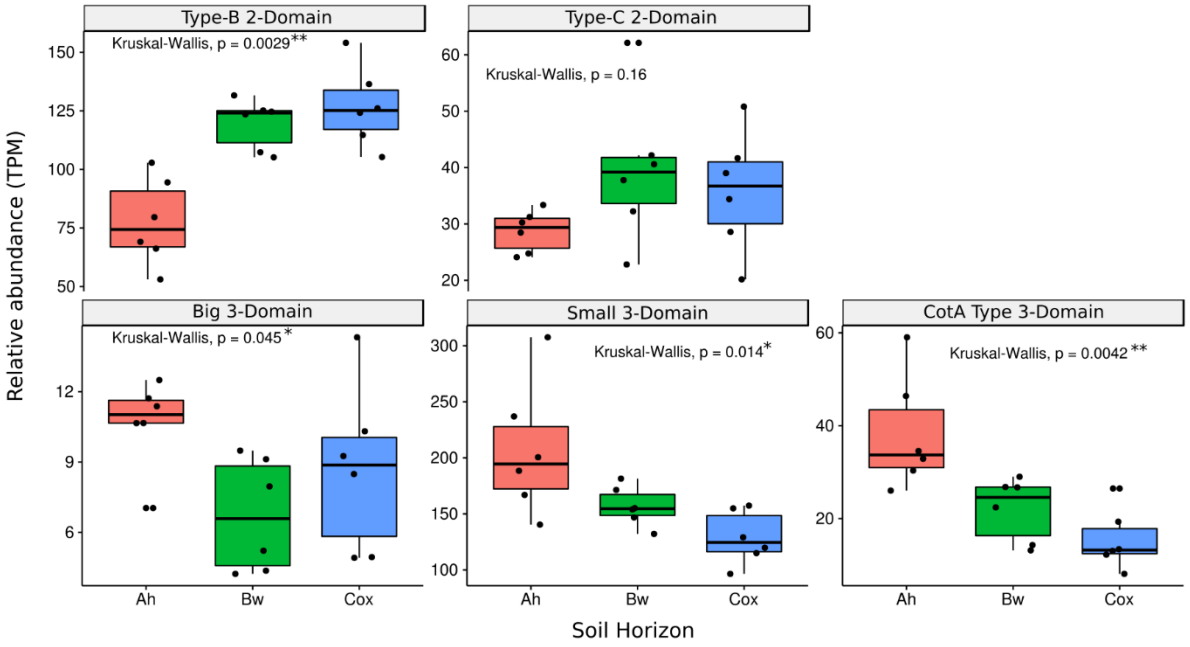


Figure 1. Principal coordinates analysis, of Alpine soil physical-chemical characteristics. A total of 26 Soil Abiotic variables were compared, (supplementary table 1.). Variables were log transformed and normalised such that the sum of squares for each sample = 1. Sample dissimilarities were calculated by their Euclidean distances. PcoA was performed using the capscale function in the package vegan (Oksanen et al., 2016) in R version 3.4.1 (R Core Team, 2011), The first two principle co-ordinate axes were used to ordinate the samples in 2d space. Samples are coloured by their soil horizon classification and convex hulls are displayed for each horizon.

608

609



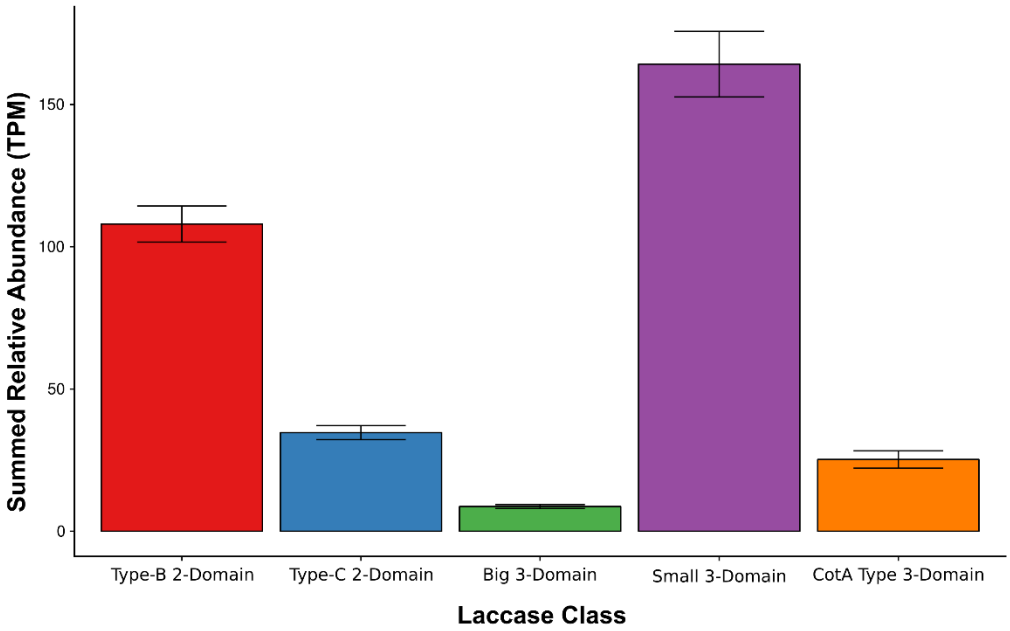
610

611

612 **Figure 2.** Laccase genes by Type Comparison of relative abundances of laccases between stratified soil
613 horizons. Relative abundances of metagenome derived type B 2 domain, type C 2 domain, Big 3
614 domain, small 3 domain, cot A 3 domain respectively. Metagenomic proteins were classified using
615 previously published Hidden Markov Models (HMMs) (Ausec et al., 2011). Abundances are expressed
616 as TPMs. The statistical significance of differences in mean abundances between soil horizon was
617 assessed by Kruskal-Wallis one way ANOVA (Kruskal & Wallis, 1952), the resulting p values are
618 displayed on each plot.

619

620



621

622

623 **Figure 3.** Total relative abundance for each of the 323 putative identified laccase genes was performed
624 across all sample microbiomes. Error bars represent standard error of the mean abundance across all
625 samples. The laccase genes identified in each sample were annotated using the package Kaiju (Menzel
626 et al., 2016). Statistical analysis of each enzyme subgroup across the metagenome datasets was
627 performed using Kruskal-Wallis one way ANOVA (Kruskal & Wallis, 1952). The distinction between
628 type B and C 2-domain laccase is from Nakuamura et al. (2003). The designation of small and large 3-
629 domain laccases comes from Machczynski et al. (2009). The CotA 3-Domain laccases were described
630 by Martin et al. (2015).

631

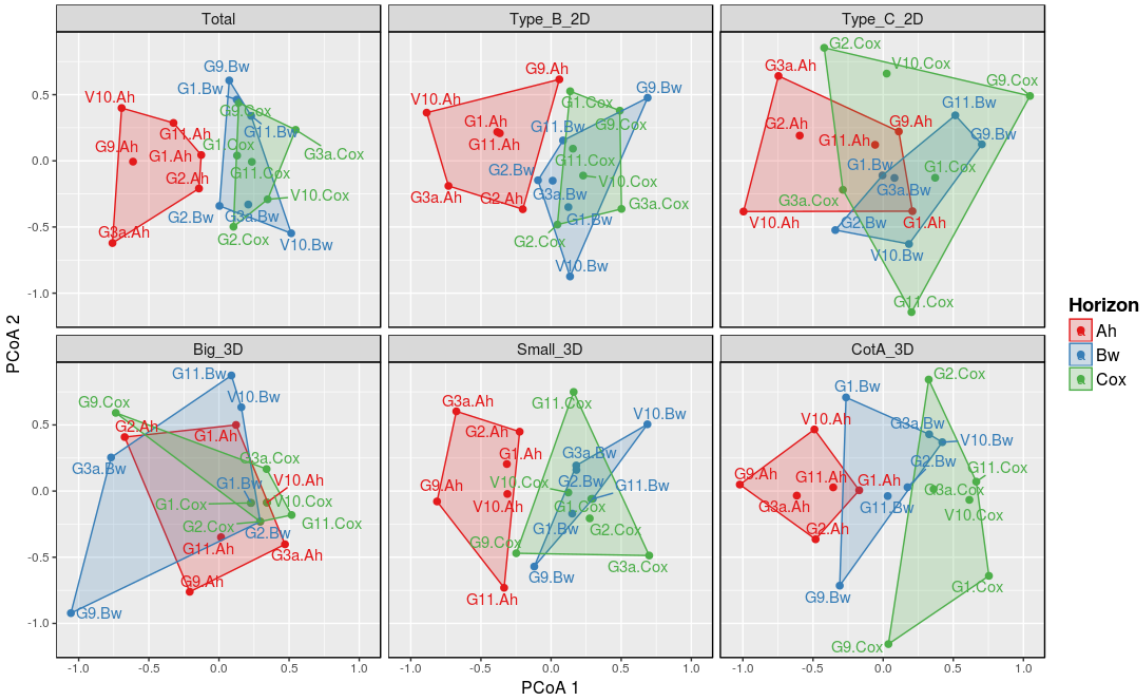


Figure 4. PCA plots for each laccase class abundance plus summed abundances (Phylum level): Laccase genes were taxonomically classified at the Phylum level using the lowest common ancestor algorithm. For each sub-class, Percentage abundances were normalized by Hellinger transformation, sample dissimilarities were calculated by their Euclidean distances. PcoA was performed using the capscale function in the package vegan (Oksanen etal., 2016) in R version 3.4.1 (R Core Team, 2011), The first two principle co-ordinate axes were used to ordinate the samples in 2d space. Samples are coloured by their soil horizon classification and convex hulls are displayed for each horizon.

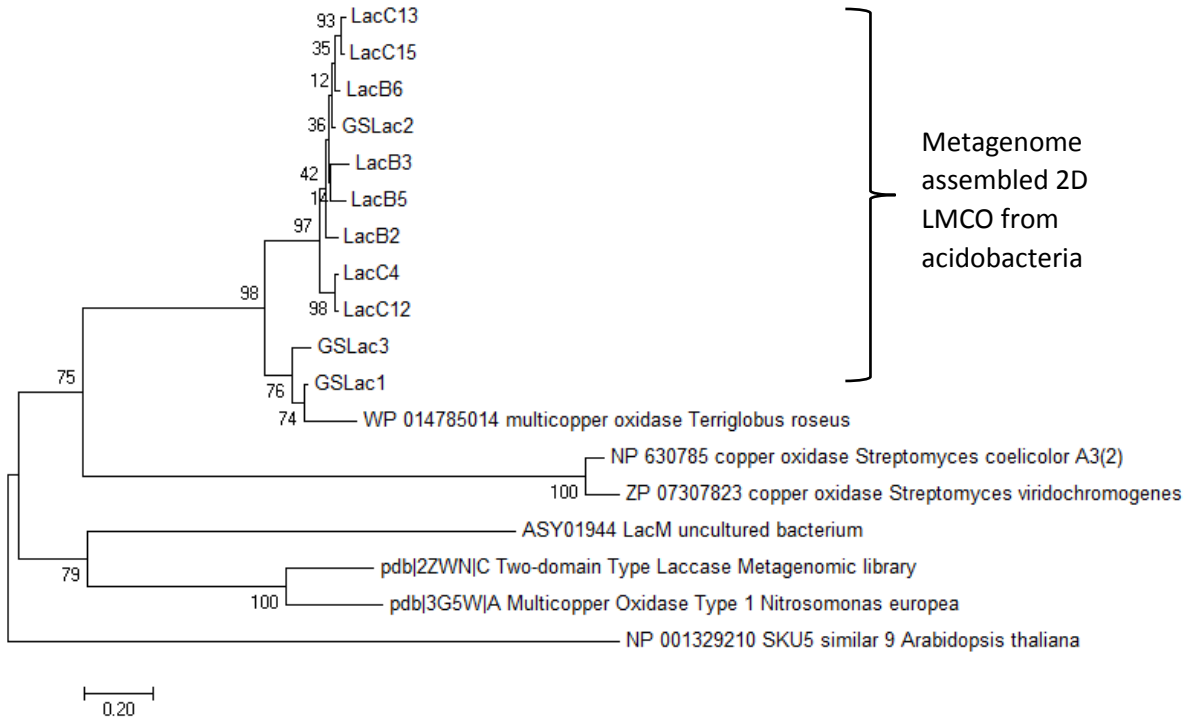


Figure 5. Maximum likelihood phylogenetic tree of alpine metagenome-derived LMCOs with reference sequences of bacterial two domain laccases that demonstrated activity. The sequences are type B 2 domain laccases amplified using the V10_lac PCR primers as described in the text, against DNA exracted from separate V10 Bw and Cox samples (hence designation B or C). Sample LacM (Ausec et al., 2017) was also included as the first acidobacterial metagenome-derived LMCO, though it is a three domain laccase. The tree was drawn using Mega 7.0 (Kumar et al., 2016). The reference sequence shown were downloaded from the NCBI database. The GSLac sequences are representatives of the metagenome-assembled putative laccase genes referred to in 3.2, while the LacB and C samples are sequences obtained after PCR amplification from Bw and Cox horizons I section 3.3.

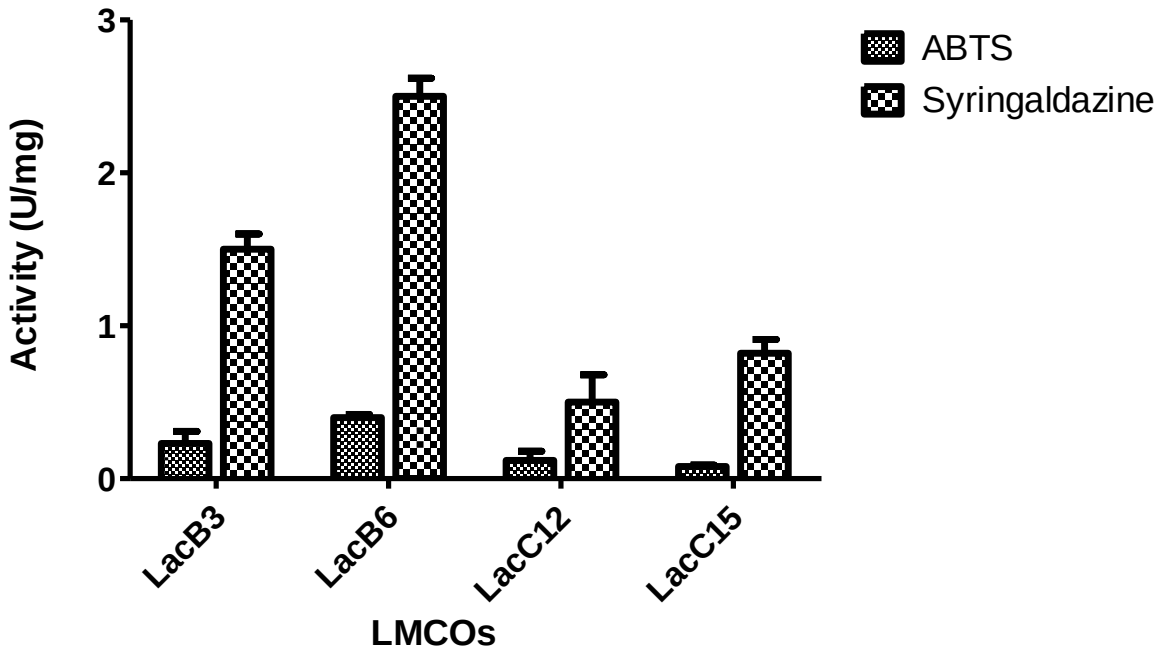


Figure 6: Specific activities of heat treated partially purified metagenome-derived laccase-like multi copper oxidases (LMCOs) from *E.coli* clone cell free extracts. The enzyme activity against these substrates was also compared with a vector control for the *E.coli* expression system (data not shown) where little activity was detectable (<0.01 U/mg for both substrates). Activity measurements are measurements of triplicate analysis (error bars show standard deviation). See methods section for assay details.

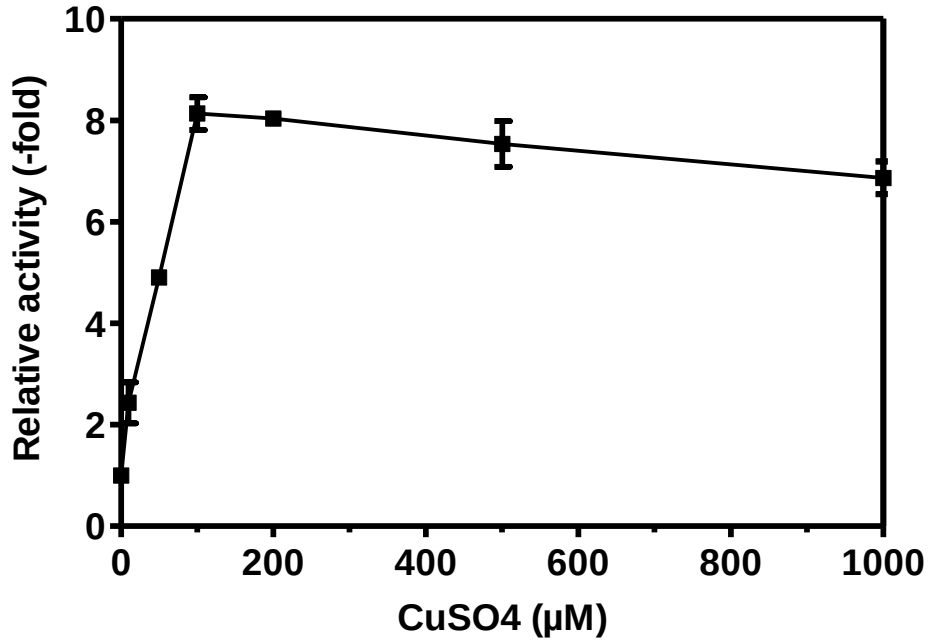


Figure 7: Effect of increasing concentrations of $CuSO_4$ on the cloned laccase LacB6. Values are fold increase at different concentrations relative to the specific activity of 2.5 U/mg recorded in the absence of copper, where relative activity is set as 1. Activity measurements are measurements of triplicate analysis (error bars show standard deviation). See methods section for assay details.