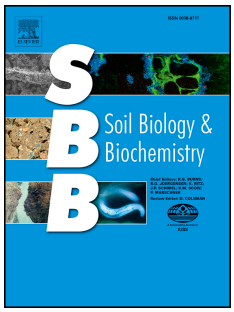
Journal Pre-proof



Altered microbial CAZyme families indicated dead biomass decomposition following afforestation

Chengjie Ren, Xinyi Zhang, Shuohong Zhang, Jieying Wang, Miaoping Xu, Yaoxin Guo, Jun Wang, Xinhui Han, Fazhu Zhao, Gaihe Yang, Russell Doughty

PII: S0038-0717(21)00235-2

DOI: https://doi.org/10.1016/j.soilbio.2021.108362

Reference: SBB 108362

To appear in: *Soil Biology and Biochemistry*

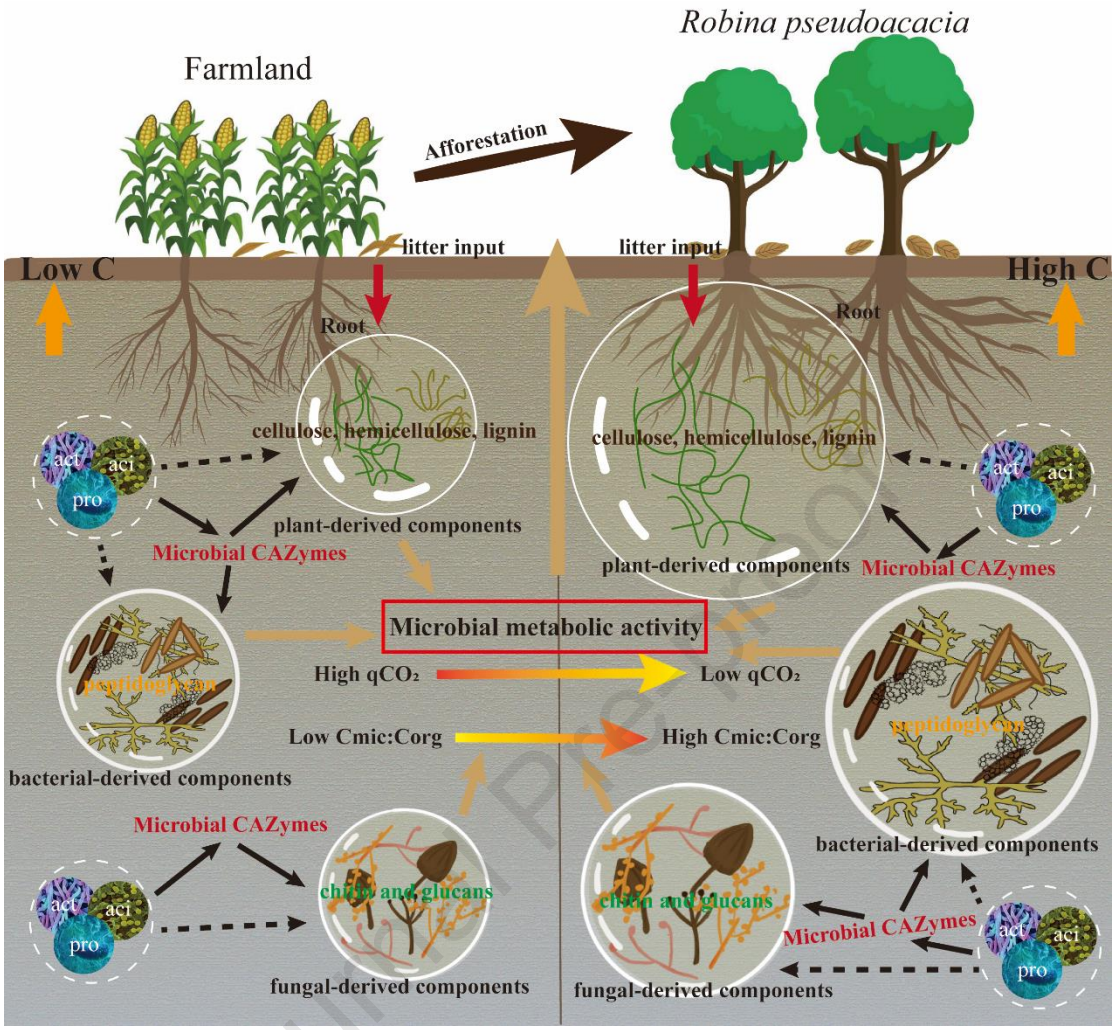
Received Date: 21 February 2021

Revised Date: 15 July 2021

Accepted Date: 16 July 2021

Please cite this article as: Ren, C., Zhang, X., Zhang, S., Wang, J., Xu, M., Guo, Y., Wang, J., Han, X., Zhao, F., Yang, G., Doughty, R., Altered microbial CAZyme families indicated dead biomass decomposition following afforestation, *Soil Biology and Biochemistry*, https://doi.org/10.1016/j.soilbio.2021.108362.

Graphic abstract



Altered microbial CAZyme families indicated dead biomass decomposition following afforestation

Chengjie Ren[1, 2], Xinyi Zhang[1, 2], Shuohong Zhang[1, 2], Jieying Wang[3], Miaoping Xu[1, 2], Yaoxin Guo[4], Jun Wang[3], Xinhui Han[1, 2], Fazhu Zhao[3*], Gaihe Yang[1, 2*], Russell Doughty[5]

[1]College of Agronomy, Northwest A&F University, Yangling, 712100 Shaanxi, China; [2]The Research Center of Recycle Agricultural Engineering and Technology of Shaanxi Province, Yangling 712100 Shaanxi, China; [3]Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, Northwest University, Xi'an, 710072, Shaanxi, China; [4]College of Life Sciences, Northwest University, Xi'an, 710072, Shaanxi, China; [5]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, 91125, USA

***** Corresponding author: Fazhu Zhao, Gaihe Yang

Email: zhaofazhu@nwu.edu.cn; ygh@nwsuaf.edu.cn;

Abstract: Afforestation can modify terrestrial carbon (C) pools, some of which are primarily fixed in the plant dead biomass and then incorporated into the microbial dead biomass. Soil microorganisms exert a critical role in C flow and potentially influence C balance through the degradation of plant and microbial dead biomass. Here, we compared sites along a 45-year *Robinia pseudoacacia* (RP) afforestation chronosequence on the Loess Plateau of China. Subsequently, the trends of microbial carbohydrate-active enzymes (CAZyme) and their responses to the decomposition of dead biomass of different origins were studied using metagenomics. The results show that soil microbial CAZyme families, which degrade the plant- and microbial-derived components, significantly increased after afforestation, with a significant peak at the 20-year site. The dominant bacterial phyla (i.e., *Actinobacteria*, *Proteobacteria*, and *Acidobacteria*) mineralized C sources from plant and microbial biomass components through their corresponding CAZyme families. Moreover, the increased abundance of CAZymes involved in the decomposition of plant-derived components (e.g., cellulose, hemicellulose, and lignin) contributed to the formation of C pools. In the case of microbial-derived components, the abundance of CAZymes encoding the bacterial-derived components (peptidoglycan) was larger than that encoding fungal-derived components (chitin and glucans) and was more associated with microbial metabolic activity (qCO_2 and Cmic: Corg ratio), indicating a higher investment of bacterial-derived components for microbial carbon turnover following afforestation. Overall, our study compares plant- and microbial-derived biomass to illustrate the differential contributions of dead biomass to C accumulation and confirms the importance of the

bacterial community and derived biomass for C turnover following afforestation.

Keywords: Metagenomics; Microbial CAZyme families; plant-derived components; microbial-derived components; Microbial metabolic activity; Afforestation.

1. Introduction

Afforestation has been widely implemented in several countries, increasing the area of planted forests globally by approximately 1.05×10^8 ha, it has undoubtedly increased C stored in biomass and mitigated climate change (Lu et al., 2018; Hong et al., 2020). Bastin et al. (2019) revealed that afforestation has the potential to sequester 205 Gt C in ecosystems, offsetting 68% of global CO_2 emissions. In terrestrial ecosystems, the C in soils originates predominantly through decomposing plant litter; however, almost all plant-derived C is transformed by soil microbes (Lopez-Mondejar et al., 2018; van der Heijden et al., 2008). Hence, soil microorganisms greatly contribute to terrestrial C flow and play critical roles in the C balance through the decomposition of dead biomass of different origins (Lopez-Mondejar et al., 2018). Recent evidence suggests that the dead biomass turnover is primarily tracked by analyzing the microbial enzymes that mediate them (Žifčáková et al., 2017; Lopez-Mondejar et al., 2018; Lopez-Mondejar et al., 2020). Thus, a better understanding of their role in C cycling in afforested soils is essential for predicting and quantifying C pools.

When considering C cycling in forest soils, the dead biomass from plants and microbes represents different compartments (Schnecker et al., 2015; Žifčáková et al., 2017). Plant biomass is composed of cellulose, hemicelluloses, and lignin, forming a complex and recalcitrant matrix, whereas the biomass derived from the living microbiota (chitin and peptidoglycan) represents an important pool of organic matter (Kallenbach et al., 2016; Žifčáková et al., 2017). The carbohydrate-active enzyme (CAZyme) have recently been used to explore the microbial functional responses for C

turnover, and a few studies have identified the specific glycoside hydrolases (GH) and auxiliary activity (AA) enzymes driving C dynamics, which are associated with the decomposition of polysaccharides and lignin, respectively (Llado et al., 2019; Lopez-Mondejar et al., 2020). For example, the cellulases or hemicellulases from several GHs are the primary enzymes degrading plant biomass, whereas lysozymes and chitinases from GHs are involved in the degradation of dead biomass from bacterial and fungal communities, respectively (Lombard et al., 2014; Žifčáková et al., 2017). Recent studies have also shown that soil substrate and environmental parameters are the dominant factors that control microbial CAZyme activities due to their control of soil microbial communities. For instance, soil substrates, such as C sources, could be considered as drivers for the transcription of genes encoding CAZymes for the utilization of plant C compounds (Hu et al., 2020; Lopez-Mondejar et al., 2020), and the changes in the soil temperature between seasons may also cause changes in CAZymes activities, thereby affecting C cycling (Žifčáková et al., 2017). Afforestation results in extensive changes in soil substrate and environmental parameters that strongly affect the demand of microorganisms for C sources (Zhong et al., 2020), and these responses may change the functional genes encoding CAZymes for the degradation of plant and microbial biomass. However, the trends of microbial CAZyme families and their influence on C turnover following afforestation have to be elucidated.

In this study, we characterize the enzymatic toolbox used for the decomposition of various biomass types for elucidating the microbial mechanisms underlying the C turnover of different origins in afforested soils from the perspective of functional genes

encoding CAZymes. We selected the Loess Plateau, one of the most fragile environments worldwide, as our study site and compared sites along a 45 year *Robinia pseudoacacia* afforestation chronosequence. We predicted an increased abundance of microbial CAZyme genes for C metabolic activity with increases in soil carbon and *in situ* soil respiration following afforestation (Ren et al., 2018). Moreover, we hypothesized that microbes can encode different CAZyme families involved in biomass degradation, and bacterial community will have a bigger role in the degradation of plant-and microbial-derived components. This hypothesis is developed from our previous findings that bacterial dominant phyla responded sensitively to afforestation with *Robinia pseudoacacia* plantations in this region (Ren et al., 2018; Zhong et al., 2020). Therefore, the objectives of this study are to: (i) determine the trends of the microbial CAZyme pool; (ii) characterize the potential of microbial taxa for their dead biomass decomposition by analyzing specific CAZyme families; (iii) evaluate the different contributions of microbial-derived decomposition components to C accumulation following afforestation.

2. Methods and materials

2.1 Study area description

Our study was carried out in the Wuliwan catchment, located in Ansai County on the Loess Plateau, China (latitude 36◦ 51'N; longitude 109◦ 20′ E; altitude 1061–1371 m above sea level). The study area is characterized by the temperate semi-arid climate. The annual mean temperature (MAT) is 8.8 □ and annual mean precipitation (MAP) is 505 mm, 70% of which falls between July and September. Over the past decades, vegetation restoration has been implemented to prevent land degradation and soil

erosion. Since the 1970s, *Robinia pseudoacacia* has been extensively planted for the degraded ecosystem, which was widely implemented by the Institute of Soil and Water Conservation of the Chinese Academy of Science (Ren et al., 2017; Zhong et al., 2020).

2.2 Experimental design and soil sampling

Samples were collected between July and August 2019 from four *Robinia pseudoacacia* stands of different ages: 14, 20, 30, and 45 years, as of 2019. The age of each *Robinia pseudoacacia* plantation was determined using an increment borer (Fig. S1). Each of the four age classes was replicated three times, representing a total of 12 stands. Neither thinning, pruning, nor nutrient amendment manipulation was performed during the development of these plantations. Three active sloped farmlands growing millet using traditional cultivation techniques were used as references. Three independent replicate stands (20 m x 30 m) were established for each land-use type, which had similar geographical characteristics within the same area, including elevation, slope, and parent soil material. Detailed information on these sampling sites is shown in Table S1 and Table S2.

After excluding aboveground plants without any disturbance, ten soil cores (5.0 cm inner diameter) were randomly obtained from the topsoil (0–10 cm) of each plot in an "S" shape. The soil samples were mixed to form a composite sample. A total of 45 soil samples (4 age classes × 3 replicate stands × 3 replicate plots + 3 farmlands × 3 replicate plots) were collected. For each variable, the average value of the three replicate plots at each stand was used to represent the stand. These samples were sieved (2 mm) and homogenized, and the roots and other debris were removed. A portion of each soil sample was stored at 4 °C for microbial biomass analyses. Subsamples for molecular analysis were stored on ice and transported to the laboratory, where they were stored at −80 °C. The other soil subsamples were air-dried and stored at room

temperature prior to chemical analysis.

2.3 Measurement of soil microbial respiration and chemical properties

Microbial respiration (MR: CO_2 efflux) was estimated using the alkali absorption method, which was described in our previous study (Zhong et al., 2020). Soil moisture (SM, %) was determined by oven drying to constant mass at 105 °C, soil pH was estimated with a soil/water ratio (w/v) of 1:2.5 using a glass electrode meter, and soil bulk density (BD) was determined from the volume of the core sampler before and after oven drying at 105°C for 24h to calculate the volume of each core. Soil clay was measured using the hydrometer method (Bouyoucos, 1962). The soil organic C (SOC) content was determined using the Walkley-Black method, and soil microbial biomass carbon (MBC) was analyzed using a chloroform fumigation-extraction method as previously described (Vance et al., 1987).

2.4 Calculation of microbial metabolic coe□cients

The metabolic quotient (qCO_2) was calculated as the ratio of soil MR to MBC (BR/MBC), and the microbial quotient was calculated as the ratio of microbial biomass C to soil organic carbon (C_{mic}: C_{org}). These two indices were used to describe the microbial metabolic activity (Bhople et al., 2019; Deng et al., 2019; Bene et al., 2013)

2.5 DNA extraction and sequencing

Soil DNA was extracted in triplicate from 0.5 g of fresh soil sample using the FastDNA spin kit for soil (MP Biomedicals, Cleveland, United States), following the manufacturer's instructions. The quality and integrity of the DNA extracts were assessed using a NanoDrop 2000 spectrophotometer. To obtain sufficient DNA for shotgun metagenomic sequencing, six replicates from each soil sample were analyzed. The extracted microbial DNA was processed to construct metagenome shotgun sequencing libraries with insert sizes of 400 bp using the Illumina TruSeq Nano DNA

LT Library Preparation Kit. Each library was sequenced by Illumina HiSeq X-ten platform (2 × 150; Illumina, USA) at Personal Biotechnology Co., Ltd. (Shanghai, China). The sequences can be found on the National Center for Biotechnology Information (NCBI), with the accession number SRP319300.

2.6 Metagenome assembly

Raw sequencing reads were processed to obtain quality-filtered reads for further analysis. Sequencing adapters were removed from sequencing reads using Cutadapt (v1.2.1) (Martin, 2011). Low quality reads were trimmed using a sliding-window algorithm in fastp (Chen et al., 2018). Taxonomic classifications of metagenomic sequencing reads from each sample were performed using Kaiju (Menzel et al., 2016) in greedy-5 mode against an nr-derived database, which included proteins from archaea, bacteria, viruses, fungi, and microbial eukaryotes. Megahit (v1.1.2) (Li et al., 2015a) was used to assemble for each samples using the meta-large presented parameters (Steinegger and Söding, 2017). The lowest common ancestor taxonomy of the non-redundant contigs was obtained by aligning them against the NCBI nucleotide database by BLASTN, and contigs assigned to Viridiplantae or Metazoa were removed from the following analysis. The annotation of contigs (longer than 200bp) was performed using both MetaGeneMark (Zhu et al., 2010) and MetaEuk (Karin et al., 2020) to predict genes. In particular, MetaEuk considered both prokaryotic and eukaryotic exons.

2.7 CAZyme annotation

The annotation of CAZymes in the metagenome contigs was performed after gene calling using the pipeline dbCAN2 with non-redundant protein (Yin et al., 2012), followed by manual curation considering the alignments to the sequences in the CAZy database (July 2020). The taxonomical annotation of non-redundant proteins was performed using MMseqs2 with easy-taxonomy mode and nr database (Steinegger and

Söding, 2017). To assess the abundance of these genes, high-quality sequences from each sample were mapped onto the predicted gene sequences using Salmon (Patro et al., 2017) and the TPM (transcripts per kilobase per million mapped reads) was used to normalize the abundance values in metagenomes.

2.8 Statistical analyses

Before the analysis, all data were tested for normal distribution. One-way analysis of variance (ANOVA) was performed to assess the effect of afforestation on the soil properties, microbial metabolic coe□cients (MBC, SOC, MR, qCO_2, C_{mic}: C_{org}), microbial phyla, and specific CAZyme families. Pearson analysis and Mantel test were conducted to identify the relationship between the abundance of the CAZymes involved in the degradation of plant- and microbial-derived components and microbial metabolic coe□cients. All analyses were conducted using R statistical software (v.4.0.3).

3. Results

3.1 Trends of Glycoside hydrolases and auxiliary enzyme families following afforestation

A total of 192,604 CAZymes were identified from the 19,336,492 predicted proteins of the whole metagenome (1.00%, average value). Among these, 75.54% and 0.10% were assigned to bacteria and fungi, respectively, and the remaining 24.36% were unassigned. Moreover, among the total CAZyme pools, the GH and AA were identified to represent average value of 36.29% and 2.05%, respectively (Table S3), and most GH and AA CAZyme were largely explained by the bacterial community, which accounted for 63.95% of the GH reads (*Actinobacteria:* 27.49%, *Proteobacteria:* 21.20%, *Chloroflexi*: *Acidobacteria* 4.54%, and 2.15%) and maximum of 63.65% of the AA reads (*Actinobacteria:* 26.00%, *Proteobacteria:* 24.43%, and *Acidobacteria*

206 3.66%). In contrast, the fungal community exhibited low variations in both the GH and
207 AA CAZyme pools (Fig. 1). After afforestation, the abundance of GHs and AA
208 significantly increased, peaking at the 20-year chronosequence. The NMDS further
209 illustrated the significant differences in GHs (ANOSIM: r=0.72, P=0.001) and AA
210 (ANOSIM: r=0.19, P=0.045) families among different afforestation stages, particularly
211 for GHs (Fig. 1c and Fig. 1d).

212 **3.2 Variations in specific CAZyme families involved in the degradation of plant-**
213 **and microbial- derived components following afforestation**

214 Among the functional groups of CAZymes, as presented in Table 1, the GH and
215 AA that targeted the plant cell wall, fungal dead biomass, and bacterial dead biomass
216 significantly increased after afforestation, peaking at 20-year afforestation (Fig. 2).
217 These trends were largely assigned to the dominant bacterial phyla, including
218 *Actinobacteria*, *Proteobacteria*, and *Acidobacteria* (Fig. 3 and Table S4). In particular,
219 *Actinobacteria* exhibited a higher number of TMP, which was higher in afforested soil
220 than in farmland. *Proteobacteria* also significantly increased as the stand age increased.

221 The specific families that degrade plant- and microbial-derived components also
222 changed with afforestation (Table S5). In detail, both GH3 (β-glucosidase: 431.31 TMP)
223 and GH2 (β-galactosidase: 179.63 TMP) were the most abundant families involved in
224 plant-derived component decomposition. In the case of microbial-derived components,
225 the most abundant families in fungal CAZyme were GH16 (chitinase; 200.04 TMP) and
226 GH18 (chitinase: 77.87 TMP), whereas the most abundant bacterial CAZyme was the
227 GH23 family (lysozyme: 505.59), all of which significantly increased following

afforestation and peaked at 20-year afforestation (Table S5). Hence, the trends of dominant GHs were also largely explained by the dominant bacterial phyla, despite both fungal and bacterial communities being rich in numerous CAZymes.

3.3 Specific CAZyme family responses for microbial metabolic activity following afforestation

We found that afforestation had a significant effect on microbial carbon metabolic activity (Table 2, $p<0.05$). MR first increased and significantly decreased after 14 years, qCO_2 decreased significantly with increasing stand age ($p<0.05$), and the C_{mic}: C_{org} ratio increased significantly with stand age. Moreover, the microbial carbon metabolic activity greatly varied with the CAZymes involved in the decomposition of both plant- and microbial-derived components. In detail, SOC and MBC were positively and significantly related with the CAZymes involved in the degradation of plant and microbial-derived components, with the exception of fungal-derived glucans ($p<0.05$). The qCO_2 exhibited a negative and strong association with the decomposition of plant-derived cellulose (R=-0.82, $p<0.001$), plant-derived hemicellulose (R=-0.72, $p<0.01$), plant-derived lignin (R=-0.57, $p<0.05$), fungal-derived chitin (R=0.62 $p<0.05$), and bacterial-derived peptidoglycan (R=-0.80, $p<0.001$). The Cmic: Corg ratio is positively and significantly correlated with bacterial-derived peptidoglycan (R=0.69, $p<0.01$), plant-derived cellulose (R=0.65, $p<0.01$), and plant-derived hemicellulose (R=0.66 $p<0.01$), but show insignificant with fungal-derived chitin (Fig.4).

4. Discussion

4.1 Occurrence of microbial CAZyme genes after afforestation

Our study revealed that soil microbes that use C sources from plant and microbial biomass possess a diverse set of microbial CAZymes that are involved in the degradation of both types of biomass (Fig. 1 and Fig. 2). This finding supports the previous viewpoint that soil microbes are the primary consumers of simple and recalcitrant substrates (Kramer et al., 2016; Lopez-Mondejar et al., 2018; Rousk and Frey, 2015). In addition, soil microbial CAZyme families significantly increased following afforestation (p <0.05), with a significant peak at the 20-year site (Fig. 2). Such trends were largely attributed to the effects of the soil environment (e.g., soil temperature and soil moisture) (Žifčáková et al., 2017), plant community, and soil substrates (Žifčáková et al., 2017; Lopez-Mondejar et al., 2018). In this regard, previous studies have showed that higher plant input (litter and root) can result in higher soil substrates and improve the demands of microorganisms for C sources (Ren et al., 2017; Xu et al., 2020), subsequently leading to the synthesis of soil enzymes. This corresponds with previous observations showing that high-quality litter was enriched in transcripts associated with cellulases and lignin-targeting enzymes (Lopez-Mondejar et al., 2016; Margida et al., 2020). Moreover, the differences in microbial nutrient limitations are also an important reason. The C and P limitation of microbial growth occurred in the farmland and late stage of afforestation (Ren et al., 2017; Zhong et al., 2020), and corresponded to the lower value of microbial CAZymes, indicating that higher microbial nutrient limitation may restrict the release of enzymes by microbes. This is also consistent with the fact that resource nutrient limitation may affect the CAZymes associated with mobilizing plant, and microbial dead biomass under such

circumstances may need to be maintained at the cost of metabolic reserves (Žifčáková et al., 2017).

Several abundant GHs encoding plant- and microbial-derived components were observed to change along the afforestation chronosequence (Fig. 2, Table S5). The families of GH2 (β-galactosidase) and GH3 (β-glucosidase), encoding hemicellulose and cellulose, initially increased and then decreased following 20 years of afforestation, suggesting a higher degradation of plant-derived components at the middle stage. The presence of numerous established families encoding chitin and peptidoglycan, such as GH16 and GH23, also confirmed their role in the degradation of microbial-derived biomass (Žifčáková et al., 2017). An increase in these two families following afforestation may reflect a higher C decomposition. These results further provide evidence for the importance of microbial-specific CAZymes for C source turnover in forest soil (Stursova et al., 2012; Žifčáková et al., 2017; Lopez-Mondejar et al., 2018; Lopez-Mondejar et al., 2020) and predict that afforestation can stimulate the decomposition of dead biomass through increasing their corresponding CAZyme families, further contributing to C accumulation.

Because of the difference in compound turnover (Lopez-Mondejar et al., 2018; Žifčáková et al., 2017), microbial communities differed in dead biomass decomposition following afforestation. In detail, most microbial CAZyme families in afforested soils were assigned to the dominant bacterial phyla (i.e., *Actinobacteria*, *Proteobacteria*, and *Acidobacteria*) (Fig. 1, Fig. 3, and Table S5), indicating the importance of the dominant bacterial community in the degradation of dead biomass after afforestation. Two

possible reasons can be provided. First, bacteria can produce CAZymes that allow them to access C in cellulose, hemicelluloses, and chitin (Eichorst and Kuske, 2012; Lopez-Mondejar et al., 2016; Stursova et al., 2012). Previous studies have confirmed the production of a variety of extracellular enzymes by the three phyla, particularly by *Acidobacteria* (Ivanova et al., 2016; Zifcakova et al., 2016). Second, bacteria have been confirmed to dominate and prefer these C biomass sources (Brabcová et al., 2018); high percentage of bacteria that potentially decompose cellulose found in forest soil and the high frequency of genes involved in the degradation of structural plant polysaccharides have been observed in bacterial genomes (Lopez-Mondejar et al., 2016). For instance, *Proteobacteria* are fast-growing copiotrophs that thrive in environments with high carbon availability (Fierer et al., 2007). Thus, an increase in *Proteobacteria* following afforestation signals a higher microbial biomass belowground. *Actinobacteria* have a filamentous growth form similar to that of fungi and are presumed to have the ability to effectively mobilize carbon (Zechmeister-Boltenstern et al., 2015). Brabcová et al. (2016) revealed that fungal biomass added to soil represents a hotspot of bacterial and not fungal abundance. Overall, these results indicate the presence of an overlap in substrate utilization by bacterial and fungal decomposers in afforested ecosystems, further confirming the key role played by the dominant bacterial phyla in biomass decomposition following afforestation.

4.2 Microbial CAZymes for metabolic activity after afforestation

Due to the different structures and functions of microbial CAZymes, the pool of microbial CAZyme genes for degrading diverse types of biomass were distinct for

316 plants and microbes. The number of CAZymes that decompose plant-derived
317 components was larger than that of microbial-derived components in the afforested
318 ecosystem (Fig. 2), indicating a greater investment in plant dead biomass for the C pool
319 following afforestation. This result corresponds with the general notion that plant tissue
320 is a C-rich material that is added to the soil in the form of both aboveground and
321 belowground litter and is the material that is most likely to become a large part of the
322 C pool (Fry et al., 2018).

323 In addition, in the case of microbial-derived components, the abundance of genes
324 involved in the decomposition of bacterial dead biomass is greater than that of fungal
325 dead biomass. These findings suggest the importance of bacterial-derived biomass
326 decomposition for C cycling in afforested soils. Similar to our results, Gunina et al.
327 (2017) indicated that dead bacterial biomass is equally abundant in forest soils, with
328 higher turnover rates than fungal biomass. Additionally, a previous study reported that
329 the peptidoglycan, derived from the bacterial community, is a major and universal
330 component of cell walls that changes rapidly (Scheffers and Pinho, 2005; Egan et al.,
331 2017). However, in contrast to our findings, some studies have indicated that fungal
332 biomass fractions are highly recalcitrant and are likely to be a major source of
333 recalcitrant soil organic matter (Clemmensen et al., 2013; Li et al., 2015b). We propose
334 two possible explanations for these different responses. First, bacteria contributed to a
335 larger proportion of the microbial community (more than 95%) in our study (Fig. S2);
336 thus the derivative biomass from the bacterial community may be higher than that from
337 the fungal community. Liang et al. (2017) revealed that the higher abundance of

338 microbial species would produce more microbial residues in forest soils. Second, the
339 variations in the soil environment and substrates could play a role in the differences in
340 the decomposition of microbial-derived components in forest soils. Hu et al. (2020)
341 conducted an *in situ* decomposition experiment and revealed that bacterial
342 peptidoglycan decomposition was correlated with microbial biomass, and fungal chitin
343 decomposition was driven by soil texture. Consistent with these deductions, our results
344 revealed that the microbial metabolic activity (MBC, qCO_2 and C_{mic}: C_{org}) was more
345 associated with the shift in CAZyme genes encoding the decomposition of bacterial
346 biomass (Fig. 4), further highlighting the importance of the bacterial dead biomass
347 (indicated by peptidoglycan) decomposition in regulating carbon metabolic activity in
348 afforested ecosystems.

349 **5. Conclusion**

350 Specific microbial CAZyme families with plant and microbial dead biomass
351 decomposition significantly increased after afforestation, particularly at the middle
352 stage. This suggests that afforestation can stimulate the decomposition of dead biomass
353 through an increase in their corresponding CAZyme families, further contributing to C
354 accumulation. Such increased trends were largely explained by the dominant bacterial
355 phyla members, despite the overlap of substrate utilization by bacteria and fungi.
356 Moreover, except for the pool of microbial CAZyme genes for degrading plant biomass,
357 the abundance of CAZymes targeting bacterial dead biomass was higher than that of
358 fungal dead biomass, which was more associated with microbial metabolic activity,
359 highlighting the importance of bacterial peptidoglycan decomposition in regulating

carbon metabolic activity after afforestation. Overall, our findings provide direct evidence for the relationship between microbial CAZyme families and C decomposition and confirm the importance of bacteria for C cycling in the context of afforestation.

Declaration of competing interest

The authors declare no competing financial interest.

Acknowledgments

This work were financially supported by the National Natural Science Foundation of China (No. 41907031), the Natural Science Basic Research Plan in Shaanxi Province of China (No. 2020JQ-272; 2019JM-211), the Young Talents Support Program of the Science and Technology Association of Shaanxi Provincial Universities (No. 20200701), the Chinese Academy of Sciences "Light of West China" Program for Introduced Talent in the West (No. 31570440) and the China Postdoctoral Science Foundation (No. 2019M650276). The authors are also grateful to anonymous reviewers whose comments and suggestions helped us to enhance the quality of this paper.

References

Bastin, J.F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C.M., Crowther, T.W., 2019. The global tree restoration potential. Science 365, 76-+.

Bene, C.D., Pellegrino, E., Debolini, M., Silvestri, N., Bonari, E., 2013. Short- and long-term effects of olive mill wastewater land spreading on soil chemical and biological properties. Soil Biol. Biochem. 56, 21-30.

Bhople, P., Djukic, I., Keiblinger, K., Zehetnere, F., Liu, D., Bierbaumer, M., Zechmeister-Boltenstern, S., Joergensen, R.G., Murugan, R., 2019. Variations in soil and microbial biomass C, N and fungal biomass ergosterol along elevation and depth gradients in Alpine ecosystems. Geoderma 345: 93-103.

Bouyoucos, G.J., 1962. Hydrometer Method Improved for Making Particle Size Analyses of Soils1. Agron. J. 54, 464-465.

Brabcová, V., Nováková, M., Davidová, A., Baldrian, P., 2016. Dead fungal mycelium in forest soil represents a decomposition hotspot and a habitat for a specific microbial community. New Phytol. 210, 1369-1381.

Brabcová, V., Štursová, M., Baldrian, P., 2018. Nutrient content affects the turnover of fungal biomass in forest topsoil and the composition of associated microbial communities. Soil Biol. Biochem. 118,187-198.

Chen, S.F., Zhou, Y.Q., Chen, Y.R., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, 884-890.

Clemmensen, K.E., Bahr, A., Ovaskainen, O., Dahlberg, A., Ekblad, A., Wallander, H., Stenlid, J., Finlay, R.D., Wardle, D.A., Lindahl, B.D., 2013. Roots and Associated Fungi Drive Long-Term Carbon Sequestration in Boreal Forest. Science 339, 1615-1618.

Deng, L., Peng, C., Huang, C., Wang, K., Shangguan, Z., 2019. Drivers of soil microbial metabolic limitation changes along a vegetation restoration gradient on the loess plateau, China. Geoderma, 353, 188-200.

Egan, A.J.F., Cleverley, R.M., Peters, K., Lewis, R.J., Vollmer, W., 2017. Regulation of bacterial cell wall growth. The FEBS Journal 284, 851-867.

Eichorst, S.A., Kuske, C.R., 2012. Identification of Cellulose-Responsive Bacterial and

406 Fungal Communities in Geographically and Edaphically Different Soils by Using
407 Stable Isotope Probing. Appl. Environ. Microbiol. 78, 2316-2327.
408 Fierer, N., Bradford, M.A., Jackson, R.B., 2007. Toward an ecological classification of
409 soil bacteria. Ecology 88, 1354-1364.
410 Fry, E.L., De Long, J.R., Bardgett, R.D., 2018. Chapter 2 - Plant Communities as
411 Modulators of Soil Carbon Storage, In: Singh, B.K. (Ed.), Soil Carbon Storage.
412 Academic Press, pp. 29-71.
413 Gunina, A., Dippold, M., Glaser, B., Kuzyakov, Y., 2017. Turnover of microbial groups
414 and cell components in soil: 13C analysis of cellular biomarkers. Biogeosciences
415 14, 271-283.
416 Hong, S.B., Yin, G.D., Piao, S.L., Dybzinski, R., Cong, N., Li, X.Y., Wang, K., Penuelas,
417 J., Zeng, H., Chen, A.P., 2020. Divergent responses of soil organic carbon to
418 afforestation. Nat. Sustain. 3, 694-+.
419 Hu, Y.T., Zheng, Q., Noll, L., Zhang, S.S., Wanek, W., 2020. Direct measurement of the
420 in situ decomposition of microbial-derived soil organic matter. Soil Biol. Biochem.
421 141, 107660.
422 Ivanova, A.A., Wegner, C.E., Kim, Y., Liesack, W., Dedysh, S.N., 2016. Identification
423 of microbial populations driving biopolymer degradation in acidic peatlands by
424 metatranscriptomic analysis. Mol. Ecol. 25, 4818-4835.
425 Kallenbach, C.M., Frey, S.D., Grandy, A.S., 2016. Direct evidence for microbial-
426 derived soil organic matter formation and its ecophysiological controls. Nat.
427 Commun. 7, 13630.
428 Karin, E.L., Mirdita, M., Soding, J., 2020. MetaEuk-sensitive, high-throughput gene
429 discovery, and annotation for large-scale eukaryotic metagenomics. Microbiome
430 8, 48.

Kramer, S., Dibbern, D., Moll, J., Huenninghaus, M., Koller, R., Krueger, D., Marhan, S., Urich, T., Wubet, T., Bonkowski, M., 2016. Resource Partitioning between Bacteria, Fungi, and Protists in the Detritusphere of an Agricultural Soil. Front. Microbiol. 7, 1524.

Li, D.H., Liu, C.M., Luo, R.B., Sadakane, K., Lam, T.W., 2015a. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31, 1674-1676.

Li, N., Xu, Y. Z., Han, X. Z., He, H. B., Zhang, X. D., Zhang, B., 2015b. Fungi contribute more than bacteria to soil organic matter through necromass accumulation under different agricultural practices during the early pedogenesis of a mollisol. Eur. J. Soil Biol. 67, 51-58.

Liang, C., Schimel, J.P., Jastrow, J.D., 2017. The importance of anabolism in microbial control over soil carbon storage. Nat. Microbiol. 2, 17105.

Llado, S., Vetrovsky, T., Baldrian, P., 2019. Tracking of the activity of individual bacteria in temperate forest soils shows guild-specific responses to seasonality. Soil Biol. Biochem. 135, 275-282.

Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490-D495.

Lopez-Mondejar, R., Brabcova, V., Stursova, M., Davidova, A., Jansa, J., Cajthaml, T., Baldrian, P., 2018. Decomposer food web in a deciduous forest shows high share of generalist microorganisms and importance of microbial biomass recycling. The ISME Journal 12, 1768-1778.

Lopez-Mondejar, R., Tlaskal, V., Vetrovsky, T., Stursova, M., Toscan, R., da Rocha, U.N., Baldrian, P., 2020. Metagenomics and stable isotope probing reveal the

complementary contribution of fungal and bacterial communities in the recycling of dead biomass in forest soil. Soil Biol. Biochem. 148, 107875.

Lopez-Mondejar, R., Zuhlke, D., Becher, D., Riedel, K. and Baldrian, P., 2016. Cellulose and hemicellulose decomposition by forest soil bacteria proceeds by the action of structurally variable enzymatic systems. Sci. Rep. 6, 25279.

Lu, F., Hu, H.F., Sun, W.J., Zhu, J.J., Liu, G.B., Zhou, W.M., Zhang, Q.F., Shi, P.L., Liu, X.P., Wu, X., Zhang, L., Wei, X.H., Dai, L.M., Zhang, K.R., Sun, Y.R., Xue, S., Zhang, W.J., Xiong, D.P., Deng, L., Liu, B.J., Zhou, L., Zhang, C., Zheng, X., Cao, J.S., Huang, Y., He, N.P., Zhou, G.Y., Bai, Y.F., Xie, Z.Q., Tang, Z.Y., Wu, B.F., Fang, J.Y., Liu, G.H., Yu, G.R., 2018. Effects of national ecological restoration projects on carbon sequestration in China from 2001 to 2010. Proc. Natl. Acad. Sci. U. S. A. 115, 4039-4044.

Margida, M.G., Lashermes, G., Moorhead, D.L., 2020. Estimating relative cellulolytic and ligninolytic enzyme activities as functions of lignin and cellulose content in decomposing plant litter. Soil Biol. Biochem. 141, 107689.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data Analysis.

Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 11257.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417-

+.

Ren, C.J., Chen, J., Deng, J., Zhao, F.Z., Han, X.H., Yang, G.H., Tong, X.G., Feng, Y.Z., Shelton, S., Ren, G.X., 2017. Response of microbial diversity to C:N:P stoichiometry in fine root and microbial biomass following afforestation. Biol. Fertil. Soils 53, 457-468.

Ren, C.J., Wang, T., Xu, Y.D., Deng, J., Zhao, F.Z., Yang, G.H., Han, X.H., Feng, Y.Z., Ren, G.X., 2018. Differential soil microbial community responses to the linkage of soil organic carbon fractions with respiration across land-use changes. For. Ecol. Manag. 409,170-178.

Rousk, J., Frey, S.D., 2015. Revisiting the hypothesis that fungal-to-bacterial dominance characterizes turnover of soil organic matter and nutrients. Ecol. Monogr. 85, 457-472.

Scheffers, D.J., Pinho, M.G., 2005. Bacterial cell wall synthesis: New insights from localization studies. Microbiol. Mol. Biol. Rev. 69, 585-+.

Schnecker, J., Wild, B., Takriti, M., Alves, R.J.E., Gentsch, N., Gittel, A., Hofer, A., Klaus, K., Knoltsch, A., Lashchinskiy, N., Mikutta, R., Richter, A., 2015. Microbial community composition shapes enzyme patterns in topsoil and subsoil horizons along a latitudinal transect in Western Siberia. Soil Biol. Biochem. 83, 106-115.

Steinegger, M., Söding, J., 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. 35, 1026-1028.

Stursova, M., Zifcakova, L., Leigh, M.B., Burgess, R., Baldrian, P., 2012. Cellulose utilization in forest litter and soil: identification of bacterial and fungal decomposers. Fems Microbiol. Ecol. 80, 735-746.

van der Heijden, M.G.A., Bardgett, R.D., van Straalen, N.M., 2008. The unseen

majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. Ecol. Lett. 11, 296-310.

Vance, E.D., Brookes, P.C. and Jenkinson, D.S., 1987. An extraction method for measuring soil microbial biomass C. Soil Biol. Biochem. 19, 703-707.

Xu, M., Gao, D., Fu, S., Lu, X., Wu, S., Han, X., Yang, G., Feng, Y., 2020. Long-term effects of vegetation and soil on the microbial communities following afforestation of farmland with Robinia pseudoacacia plantations. Geoderma 367, 114263.

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 40, W445-W451.

Zhong, Z.K., Li, W.J., Lu, X.Q., Gu, Y.Q., Wu, S.J., Shen, Z.Y., Han, X.H., Yang, G.H., Ren, C.J., 2020. Adaptive pathways of soil microorganisms to stoichiometric imbalances regulate microbial respiration following afforestation in the Loess Plateau, China. Soil Biol. Biochem. 151, 108048.

Zhu, W.H., Lomsadze, A., Borodovsky, M., 2010. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res. 38, e132.

Žifčáková, L., Větrovský, T., Lombard, V., Henrissat, B., Howe, A., Baldrian, P., 2017. Feed in summer, rest in winter: microbial carbon utilization in forest topsoil. Microbiome. 5, 122.

Žifčáková, L., Vetrovsky, T., Howe, A., Baldrian, P., 2016. Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. Environ. Microbiol. 18, 288-301.

525 **Table 1 Functional classification of glycosyl hydrolases (GH) and auxilliary (AA) encoding the enzymatic activities involved in the**

526 **plant-and microbial compounds degradation according to CAZy (http://www.CAZy.org)**

Group	Compound	CAZy families (GH and AA)
Plant biomass	cellulose	GH1 (β-glucosidase), GH3 (β-glucosidase), GH5 (β-glucosidase/endoglucanase), GH6 (cellobiohydrolase), GH7 (reducing end-acting cellobiohydrolase), GH8 (endoglucanase/endoxylanase), GH9 (endoglucanase), GH12 (endoglucanase), GH45 (endoglucanase), GH48 (reducing end-acting cellobiohydrolase/ endoglucanase), GH116 (β-glucosidase), AA9 (lytic polysaccharide monooxygenase), and AA10 (lytic polysaccharide monooxygenase)
	hemicellulose	GH2 (β-galactosidase/β-glucuronidase), GH10 (endoxylanase), GH11 (endoxylanase), GH26 (endomannanase), GH30 (endoxylanase/β-1,6- glucanase/β-xylosidase), GH36 (α-galactosidase), GH39 (β-xylosidase/α-L-arabinofuranosidase), GH43 (β-xylosidase/endoxylanase), GH44 (xyloglucanase/endoglucanase), GH51 (α-L_x0002_arabinofuranosidase), GH52 (β-xylosidase), GH54 (α-L-arabinofuranosidase), GH62 (α-Larabinofuranosidase) GH67 (xylan α-1,2- glucuronidase), GH74 (xyloglucanase), GH95 (α-L-fucosidase/α-L-galactosidase), GH115 (xylan α-1,2-glucuronidase), GH120 (β-xylosidase)
	lignin	AA1(laccase), AA2(peroxidase), AA3 (oxidase), AA4 (oxidase), AA5(oxidase), AA6 (1,4- benzoquinone reductase)
Fungal biomass	chitin	GH16 (xyloglucanase/endoglucanase), GH18 (chitinase), GH19 (chitinase), GH20 (N-acetyl β-glucosaminidase) , GH72 (β-1,3-glucanosyltransglycosylase)
	glucans	GH17 (endo-1,3-β-glucanase), GH55 (exo-β-1,3- glucanase/endo-1,3-β-glucanase), GH64 (endo- 1,3-β-glucanase), GH81 (endo-1,3-β-glucanase), and GH128 (endo-1,3-β-glucanase)
Bacterial biomass	peptidoglycan	GH22 (lysozyme), GH23 (lysozyme/ peptidoglycan lytic transglycosylase), GH24 (lysozyme), GH25 (lysozyme), GH73 (peptidoglycan hydrolase with endo-β-N-acetylglucosaminidase specificity), GH102 (peptidoglycan lytic transglycosylase), GH103 (peptidoglycan lytic transglycosylase), GH104 (peptidoglycan lytic transglycosylase) and GH108 (lysozyme)

527

528 **Table 2 Changes in soil microbial metabolic coe□cients following afforestation.** Values are represented as the mean value followed by standard error in
529 parentheses (n = 3). Different lowercase letters indicate a significant difference (P <0.05) between different age classes, based on a one-way ANOVA followed by an
530 LSD test. FL: farmland; RP14yr, RP 20yr, RP 30yr, and RP 45yr represent that the Robinia pseudoacacia (RP) plantation. SOC: soil organic carbon, MBC: microbial
531 biomass carbon; qCO_2 was estimated by MR: MBC.

Parameters	FL	RP14yr	RP20 yr	RP30 yr	RP45 yr	$F_{(4,15)}$	*p*
SOC (g/kg)	2.98±0.07d	4.3±0.19c	4.74±0.3c	5.84±0.4b	7.73±0.32a	40.88	<0.001
MBC (mg/kg)	60.27±4.62d	99.26±8.82d	240.53±31.67c	367.73±33.41b	521.54±8.9a	79.49	<0.001
Microbial respiration (MR) (mg/kg)	76.37±2.2a	88.37±3.41a	73.82±9.83a	67.56±6.12a	57.07±6.3a	3.49	0.052
qCO_2	1.28±0.1a	0.9±0.05b	0.32±0.07c	0.18±0.01cd	0.11±0.01d	72.77	<0.001
Cmic: Corg	20.27±1.78c	23.15±2.15c	50.68±5.95b	63.87±8.21ab	67.79±4.05a	19.67	<0.001

532

Fig. 1 The abundance (TPM: transcripts per kilobase per million mapped reads) of microbial glycoside hydrolases (GHs) and auxiliary activities (AAs) following afforestation. (a,b) contribution of microbial (bacterial and fungal) phyla to the GHs and AAs families following afforestation; (c, d) Nonmetric multidimensional scaling of of the GH and AA families following afforestation

Fig. 2 The abundance (TPM: transcripts per kilobase per million mapped reads) of selected GHs and AAs encoding the decomposition of the plant-and microbial-derived components following afforestation. (a) plant-derived cellulose decomposotion; (b) plant-derived hemicellulose decomposition; (c) plant-derived lignin decomposition; (d) fungi-derived chitin decomposition; (e) fungi-derived glucans decomposition; (f) bacteria-derived peptidoglycan decomposition.

Fig. 3 Contribution of microbial (bacterial and fungal) phyla to microbial CAZyme genes for plant-and microbial-derived components decomposition following afforestation. (a) contribution of microbial (bacterial and fungal) phyla to plant-derived cellulose decomposotion. (b) contribution of microbial (bacterial and fungal) phyla to plant-derived hemicellulose decomposition. (c) contribution of microbial (bacterial and fungal) phyla to plant-derived lignin decomposition. (d) contribution of microbial (bacterial and fungal) phyla to fungi-derived chitin decomposition. (e) contribution of microbial (bacterial and fungal) phyla to fungi-derived glucans decomposition. (f) contribution of microbial (bacterial and fungal) phyla to bacteria-derived peptidoglycan decomposition.

Fig. 4 Relationship between the abundance of the CAZymes involved in the degradation of plant- and microbial-derived components and microbial metabolic activity following afforestation.

Supplementary Information

Fig. S1 An increment borer was used to determine the age of the Robinia pseudoacacia plantations.

Fig. S2 Distribution of microbial domain in the whole metagenome

Table S1. Geographical features and plant characteristics along the afforestation chronosequence

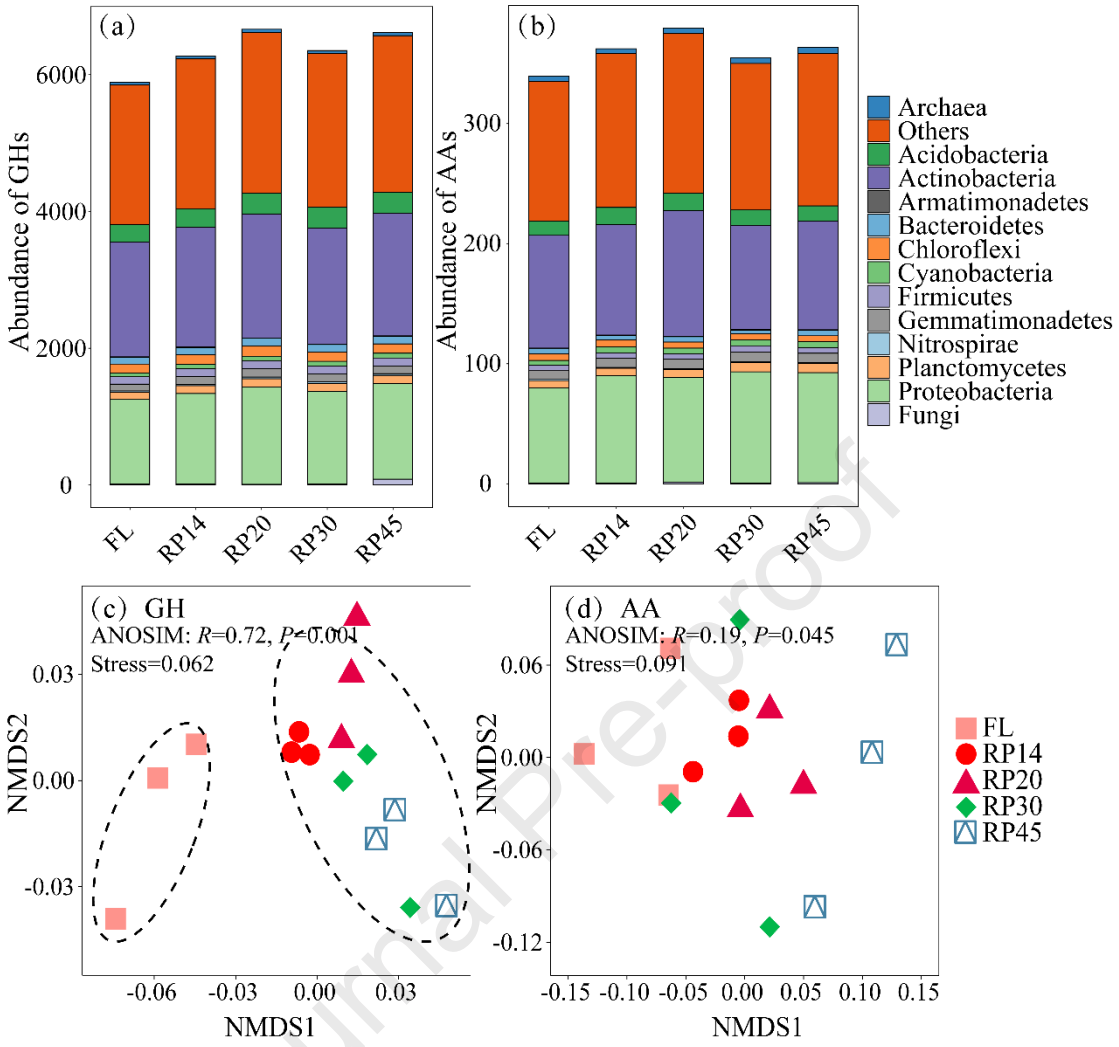
Table S2 Soil nutrients along the afforestation chronosequence

Table S3 Abundance of genes (TPM: transcripts per kilobase per million mapped reads) in the whole metagenome encoding CAZymes. Values are represented as the mean value followed by standard error in parentheses (n = 3). Different lowercase letters indicate a significant difference ($P < 0.05$) between different age classes, based on a one-way ANOVA followed by an LSD test. FL: farmland; RP14yr, RP 20yr, RP 30yr, and RP 45yr represent that the Robinia pseudoacacia (RP) plantation. GH: Glycoside Hydrolases, AA: Auxiliary Activities, CBM: Carbohydrate-Binding Modules, CE: Carbohydrate Esterases, PL: Polysaccharide Lyases, GT: GlycosylTransferases.

Table S4 Abundance of microbial phyla (average values and standard error) encoding the decomposition of the plant-and microbial-derived components after afforestation. Different letters indicate significant differences (ANOVA, $P < 0.05$, Tukey's HSD post-hoc analysis) among different land use types. *, $P<0.05$; **, $P<0.01$.

Table S5 Abundance of selected microbial CAZymes (average values and standard error) encoding the decomposition of the plant-and microbial-derived components after afforestation. Different letters indicate significant differences (ANOVA, $P < 0.05$, Tukey's HSD post-hoc analysis) among different land use types. *, $P<0.05$; **, $P<0.01$.

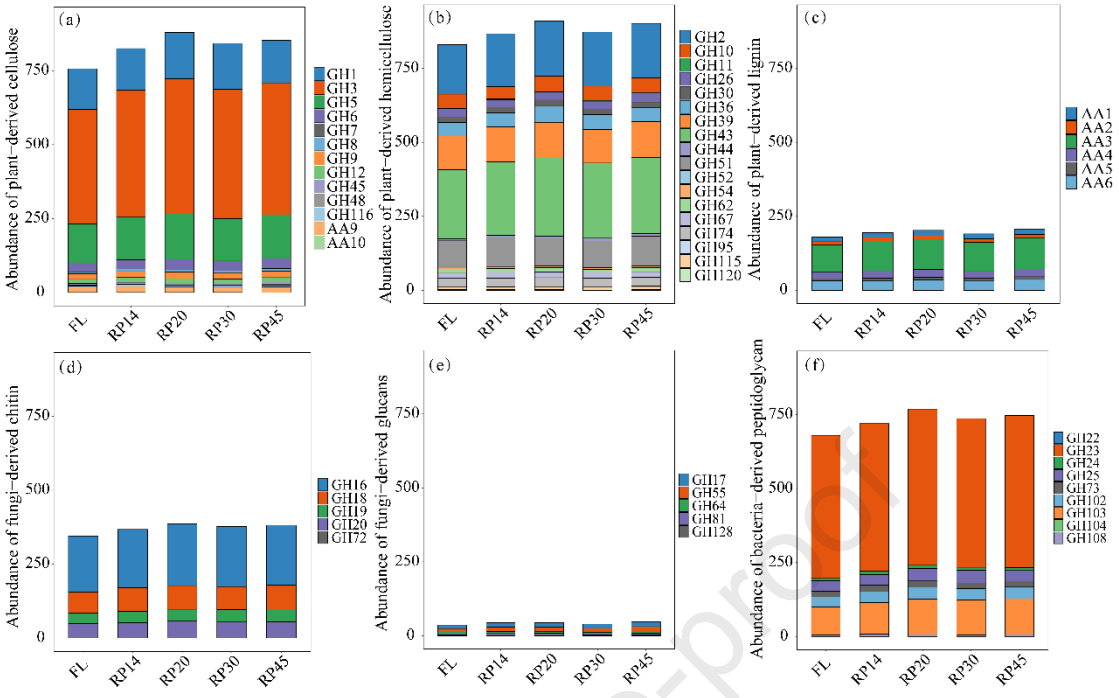
581 **Fig. 1**



582

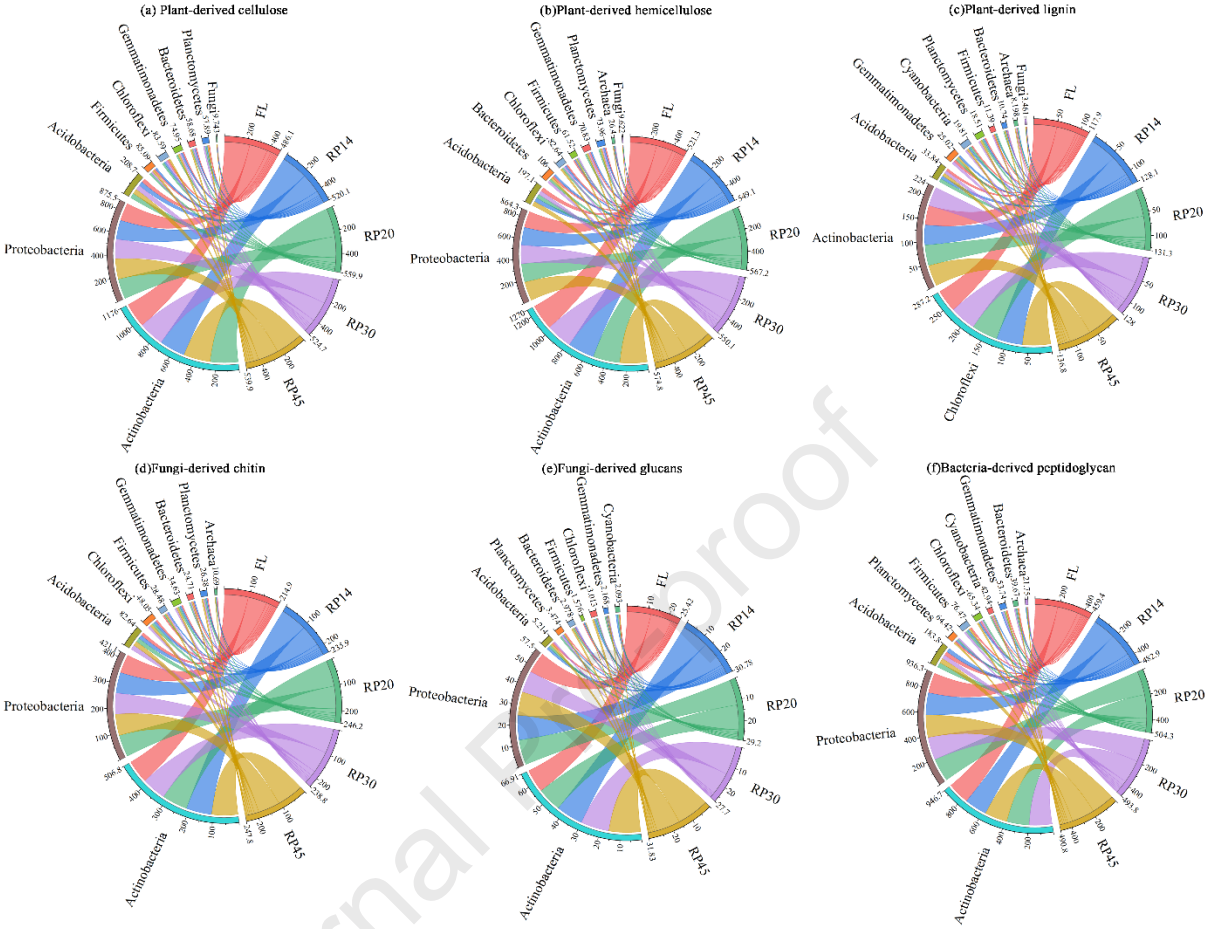
583

584 **Fig. 2**



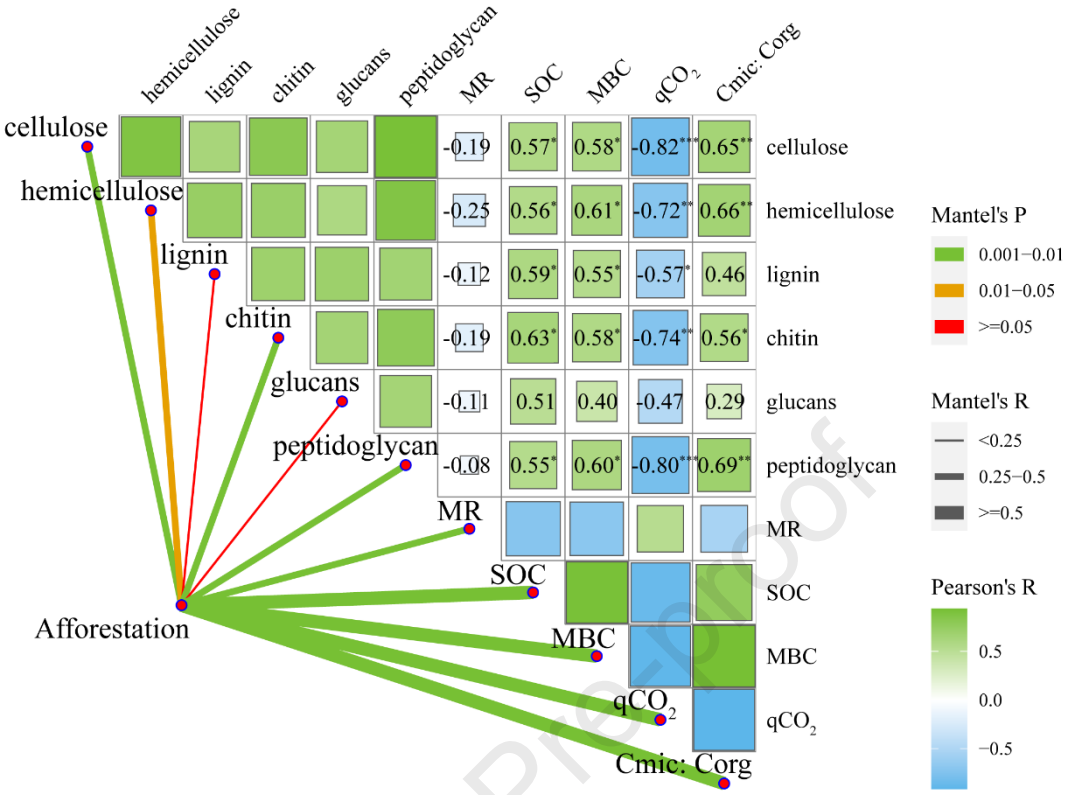
585

586 **Fig. 3**



587

588 **Fig. 4**



589

590

1. Microbial CAZymes increase after a□orestation, peaking at 20-year site

2. Dominant bacterial phyla encode more CAZymes that degrade plant and microbial biomass

3. The CAZymes encoding bacterial biomass are higher than that encoding fungal biomass

4. Bacterial biomass decomposition is more associated with microbial metabolic activity

Declaration of competing interest

All authors declare no competing financial interest.

Chengjie Ren[1, 2], Xinyi Zhang[1, 2], Shuohong Zhang[1, 2], Jieying Wang[3], Miaoping Xu[1, 2], Yaoxin Guo[4], Jun Wang[3], Xinhui Han[1, 2], Fazhu Zhao[3*], Gaihe Yang[1, 2*], Russell Doughty[5]

[1]*College of Agronomy, Northwest A&F University, Yangling, 712100 Shaanxi, China;* [2]*The Research Center of Recycle Agricultural Engineering and Technology of Shaanxi Province, Yangling 712100 Shaanxi, China;* [3]*Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, Northwest University, Xi'an, 710072, Shaanxi, China;* [4]*College of Life Sciences, Northwest University, Xi'an, 710072, Shaanxi, China;* [5]*Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, 91125, USA*

***** Corresponding author: Fazhu Zhao, Gaihe Yang

Email: zhaofazhu@nwu.edu.cn; ygh@nwsuaf.edu.cn;