Authors' accepted manuscript. Published as
Wang, X.Y., Hua, F.Y., Wang, L., Wilcove, D.W., Yu, D.W. (2019) The biodiversity benefit of native forest over Grain-for-Green plantations. Diversity and Distributions. doi: 10.1111/ddi.12972. bioRXiv: https://doi.org/10.1101/437418

Title: The biodiversity benefit of native forest over Grain-for-Green plantations

Authors: Xiaoyang Wang[1,2], Fangyuan Hua[3,4], Lin Wang[1], David S. Wilcove[5,6], Douglas W Yu[1,7,8*]

Author affiliations:

[1] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[2] Kunming College of Life Sciences, University of Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[3] Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3QZ, U.K.

[4] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, China

[5] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, United States

[6] Program in Science, Technology, and Environmental Policy, Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544, United States

[7] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming Yunnan, 650223 China

[8] School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK

***Correspondence to:** Email: dougwyu@mac.com

Author declaration: XW, FH, and DY conceived the research and designed the experiments. XW and FH carried out the fieldwork, WL created Figure 1, and XW carried out the lab work. XW and DY carried out the bioinformatic and statistical analyses. XW, DY, and FH wrote the manuscript, with comments from all other authors.

Fangyuan Hua: slcyane@gmail.com

Lin Wang: wanglin@xtbg.ac.cn

David S. Wilcove: dwilcove@Princeton.EDU

Biosketch: Prof. Douglas Yu is a Principal Investigator at the Kunming Institute of Zoology, China, and Professor of Ecology at the University of East Anglia, UK. His group is a leading developer of methods for rapid assessment of eukaryotic biodiversity using high-throughput DNA sequencing (metabarcoding, mitogenomics, metagenomics) and participates in multiple collaborations to provide biodiversity information for ecosystem-scale management. Yu is also a co-founder of NatureMetrics (www.naturemetrics.co.uk), which supplies DNA-based services to governments, NGOs, and

businesses so that they can improve policies and processes for protecting and growing biodiversity and natural capital through adaptive management.

ACKNOWLEDGMENTS

We thank Y. Liu, Q. Wei, S.K. Robinson, S. Francis, R. Zeidler, W. Hua, P. Li, M. Xu and Y. Yao for helpful discussions and logistical support. Special thanks go to our field assistants: S. Huang, M. Tong, and K. Du. We thank Ding Zhaoli for sequencing support. We are grateful for the support of Ji Yinqiu, Wu Chunying, Yang Chunyan, and Wang Yunyu. D.W. Yu and X. Y. Wang were supported by the National Natural Science Foundation of China (41661144002, 31670536, 31400470, 31500305), the Key Research Program of Frontier Sciences, CAS (QYZDY-SSW-SMC024), the Bureau of International Cooperation project (GJHZ1754), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA20050202, XDB31000000), the Ministry of Science and Technology of China (2012FY110800), the University of East Anglia, and the State Key Laboratory of Genetic Resources and Evolution at the Kunming Institute of Zoology. F. Hua and D. S. Wilcove were supported by the High Meadows Foundation and Princeton University. L. Wang was supported by the National Nature Science Foundation of China (31872963) and the Informatization Program of the Chinese Academy of Sciences (XXH13505-03-102-2).

Title: The biodiversity benefit of native forests and mixed-species plantations over monoculture plantations

Word count: Abstract = 339; Main text = 5605

Number of references: 92

Number of Figures, Tables: 6 and 0 respectively

Data accessibility statement: Sequence data have been submitted to GenBank under accession number SAMN09981891. Bioinformatic scripts are in Supplementary Information. R scripts and data tables are available at https://github.com/dougwyu/Sichuan2014.

ABSTRACT

Aim: China's Grain for Green Program (GFGP) is the largest reforestation program in the world and has been operating since 1999. The GFGP has promoted the establishment of tree plantations over the restoration of diverse native forests. In a previous study, we showed that native forests support a higher species richness and abundance of birds and bees than do GFGP plantations and that mixed-species GFGP plantations support a higher level of bird (but not bee) diversity than do any individual GFGP monocultures (although still below that of native forests). Here, we use metabarcoding of arthropod diversity to test the generality of these results.

Location: Sichuan, China

Methods: We sampled arthropod communities using pan traps in the land-cover types concerned under the GFGP. These land-use types include croplands (the land cover being reforested under the GFGP), native forests (the reference ecosystem as the benchmark for the GFGP's biodiversity effects), and the dominant GFGP reforestation outcomes: monoculture and mixed-species plantations. We used COI-amplicon sequencing ('metabarcoding') of the arthropod samples to quantify and assess the arthropod community profiles associated with each land-cover type.

Results: Native forests support the highest overall levels of arthropod species diversity, followed by mixed-species plantations, followed by bamboo and other monocultures. Also, the arthropod community in native forests shares more species with mixed-species plantations than it does with any of the monocultures. Together, these results broadly corroborate our previous conclusions on birds and bees but show a higher arthropod biodiversity value of mixed-species plantations than previously indicated by bees alone.

Main conclusion: In our previous study, we recommended that GFGP should prioritize the conservation and restoration of native forests. Also, where plantations are to be used, we recommended that the GFGP should promote mixed-species arrangements over monocultures. Both these recommendations should result in more effective protection of terrestrial biodiversity, which is an important objective of China's land-sustainability spending. The results of this study strengthen these recommendations because our policy prescriptions are now also based on a dataset that includes over 500 species-resolution taxa, ranging across the Arthropoda.

KEYWORDS: Arthropoda, biodiversity, China, forest management, Grain for Green Program, metabarcoding, reforestation

1 INTRODUCTION

An important challenge for conservation science is to quantify the biodiversity impacts of major policy initiatives, especially in regions undergoing large shifts in land-use change. Nowhere is this more true than in China, which combines a high level of native biodiversity (Tao, Huang, Jin, & Guo, 2010) with a large human population that is increasing its ecological footprint (Liu & Diamond, 2005; Pyne, 2013; Sayer & Sun, 2003; Xie et al., 2012). Moreover, for decades, China has had the managerial, political, and financial capacity to implement the largest land-sustainability programs ever seen, from nature-reserve protection to reforestation to de-desertification (Bryan et al., 2018; Liu et al., 2003; Wu et al., 2019; Xu, Wang, & Xue, 1999). These programs have caused major land-use changes and successfully slowed land degradation caused by economic activities (Liu, Li, Ouyang, Tam & Chen, 2008; Ouyang et al., 2016; Ren et al., 2015). For example, China established its first nature reserve in 1956 and reached 2740 reserves at the end of 2015 (Ma, Shen, Grumbine, & Corlett, 2017). Nearly two-thirds of the area of those nature reserves have national-level status, meaning that they receive the highest level of protection and funding, and analysis of Landsat imagery has shown that national-level reserves successfully deter deforestation (Ren et al., 2015).

Two other major land-sustainability programs are the Natural Forest Protection Program (NFPP, also known as Natural Forest Conservation Program) and the Grain for Green Program (GFGP, also known as the Sloping Land Conservation Program and the Farm to Forest Program), which were implemented after widespread flooding in 1998 (Liu et al., 2008; Xu, Yin, Li, & Liu, 2006; Yin, Yin, & Li, 2009). The NFPP aims to reduce soil erosion and flooding by protecting native forests in the upstream watersheds of the Yangtze and Yellow Rivers (Liu et al., 2008; Ren et al., 2015). The GFGP complements the NFPP by controlling soil erosion on sloping land. The government pays cash and grain to farmers in exchange for tree planting on sloping farmland (Delang &Yuan, 2015; Liu et al., 2008; Ma et al., 2017; Xu et al., 2006; Zhai, Xu, Dai, Cannon, & Grumbine, 2014). Having reforested 9.06 million ha of cropland over 16 years (~2014) since its inception in 1999, the GFGP is the world's largest reforestation program.

However, relative to their scale and budgets, little is known about the biodiversity consequences of China's land-sustainability programs, even though an important and expected co-benefit is biodiversity conservation (Wu et al., 2019). In a recent, massive review, Bryan et al. (2018) were able to cite only one study on the consequences of China's large-scale reforestation programs for biodiversity, Hua et al. (2016). This paucity of understanding contrasts starkly with the large volume of information on other consequences of these programs: water and soil maintenance (Deng, Shangguan, & Li, 2012; Long et al., 2006; Wang, Peng, Zhao, Liu, & Chen, 2017; Wang, Jiao, Rayburg, Wang, & Su, 2016), carbon storage (Deng, Liu, & Shangguan, 2014; Wei et al., 2014), vegetation cover (Hua et al., 2018; Zhai et al., 2014; Zhou, Van Rompaey, & Wang, 2009), and socioeconomic

outcomes (Liu & Lan, 2015; Yin, Liu, Zhao, Yao, & Liu, 2014; Yin et al., 2009). A better understanding of the biodiversity implications of reforestation programs is needed to guide these programs for China and the rest of the world (Turner, Lambin, & Reenberg, 2007; United Nations, 2015).

Guided by the goal of soil erosion control, and operating under the implicit assumption that any type of tree cover should achieve this goal, the GFGP has predominantly established tree plantations ('plantations' hereafter) on retired croplands, rather than restoring native forests (Hua et al., 2016; Hua et al., 2018; Zhai et al., 2014). However, compared with native forest ecosystems, plantations are known to support lower levels of biodiversity across the world's forest biomes and across taxa (Barlow, Overal, Araujo, Gardner, & Peres, 2007; Bremer & Farley, 2010; Brockerhoff, Jactel, Parrotta, Quine, & Sayer, 2008; Gardner, Hernandez, Barlow, & Peres, 2008; Lindenmayer & Hobbs, 2004), although certain management regimes, such as maintaining understory structure and mixed cropping, can somewhat increase biodiversity (Hartley, 2002). On the other hand, compared with croplands, plantations are known to support different species assemblages, with potentially higher levels of biodiversity, although there are indications that croplands in low-intensity agricultural systems – which the croplands retired under GFGP tend to be (Hu, Fu, Chen, & Gulinck, 2006) – may support considerable biodiversity which potentially exceeds that associated with plantations (Allan, Harrison, Navarro, Wilgen, & Thompson, 1997; Buscardo et al., 2008; Elsen, Ramesh, & Wilcove, 2018). Together, these insights suggest that plantations should have been expected to support low levels of biodiversity and that the GFGP could support more biodiversity if it restored native forests.

Indeed, this is what Hua et al. (2016) found. They surveyed bird and bee communities in GFGP-related tree covers in south-central Sichuan, comparing native-forest remnants to GFGP-financed tree-cover types, which include monoculture stands of bamboo, Eucalyptus, and Japanese cedar, as well as 'mixed plantations,' which are mostly patchworks (checkerboards) of two to five different monocultures and, to a lesser extent, *bona fide* tree-level mixtures (Hua et al., 2018). Most importantly, this study documented that bird and bee species diversities were higher in native forests than in any of the monocultures. In addition, they found that in mixed plantations, bird diversity for non-breeding species was higher than in any of the individual monocultures, albeit lower than in native forests. In contrast, bee diversity was equally low in mixed plantations and monocultures. The lack of a boost to bee diversity in mixed plantations was not surprising, since as with monocultures, the understory vegetation in mixed plantations was notably lacking in flowering plants (Hua et al., 2016).

The above findings, however, raise the question of why bird diversity was increased just by planting monocultures of different tree species next to each other. One possibility that could not be investigated in Hua et al. (2016) is that general arthropod diversity might also have been boosted in the mixed plantations, since, unlike bees, other arthropods can exploit a range of food resources available even in plantations, via direct consumption of plants and

fungi, and via decomposition, parasitism, and predation of other animals, including other arthropods (Jactel & Brockerhoff, 2007). Increased arthropod diversity might in turn support more bird diversity. In addition, as a large component of biodiversity, how arthropods themselves (and subgroups thereof) are affected by the GFGP is an important part of understanding the GFGP's biodiversity effects. For instance, Barlow et al. (2007) compared primary forest and Eucalyptus plantations in Brazil and found that birds achieve highest diversity in primary forest, while bees have similar levels of species richness in primary forest and Eucalyptus plantations. They also found that butterflies and dung beetles achieve low diversity but that fruit flies and moths achieve high diversity in Eucalyptus plantations.

The purpose of this study is to test the generality of Hua et al.'s (2016) results by interrogating the 'rest of the biodiversity' that was captured in the same sites analyzed by Hua et al. (2016). We employ the technique of metabarcoding, which combines traditional DNA barcoding with high-throughput DNA sequencing to characterize the biodiversity of mixed samples of eukaryotes (Cristescu, 2014; Deiner et al., 2017; Yu et al., 2012), and which has been shown to be a reliable and efficient method for biodiversity characterization (Ji et al., 2013). Through metabarcoding the non-bee arthropods caught in the same pan traps previously used to trap bees in Hua et al. 2016, we hope to answer the following questions: (1) Do native forests support higher levels of arthropod species richness and diversity than all four GFGP plantations? (2) For all GFGP plantations, do mixed plantations support higher levels of arthropod species richness and diversity than do the three individual monocultures? (3) How does community composition compare among these tree covers and what underlies the potential differences?

2 METHODS

2.1 Study location

The study region and locations are as in Hua et al. (2016). In short, our study region was a 7,949 km^2 area in south-central Sichuan province (Figure 1) spanning 315–1,715 m above sea level, historically forested and then deforested starting in the 1950s. The GFGP established ~54,800 ha of new tree cover between 1999 and 2014, dominated by short-rotation (6-20 years) monocultures of bamboo (BB), Eucalyptus (EC), and Japanese cedar (JC), and short-rotation mixed plantations (MP) of two to five tree species (including the three monoculture species). Monocultures are created by households planting the same tree species in neighboring landholdings. Correspondingly, mixed plantations are, in most cases, created by planting different species, resulting in a checkerboard, although about a quarter of mixed plantations consist of tree-level mixtures. In Hua et al. (2016), we used the term 'mixed forests', but in Hua et al. (2018), we switched to 'mixed plantations.'

The two other surveyed land covers were croplands (CL) and native forests (NF). Croplands mostly consist of low-intensity plantings of rice, corn, and vegetables and is the land-cover type that has been reforested by GFGP. Native forests are broadleaf, subtropical, evergreen forest that have been subject to decades of selective logging and other forms of extraction.

Because this region of China has been inhabited for millennia, there are no undisturbed native forests. Croplands are typically located on flatter land than are the tree covers, since GFGP reforestation targeted sloped land, and the native forests are concentrated toward the more hilly, southern end of the study region. For sampling, we chose larger expanses (> 60 ha) of these six land-cover types: BB, EC, JC, MP, NF, and CL.

2.2 Sampling design

Each land-cover type was represented by at least two locations set ≥15 km apart. All tree-cover stands sampled had closed canopy. For each land-cover type, we sampled with at least 10 one-ha quadrats, within each of which we operated 40 fluorescent pan traps for 24 hrs (Bartholomew & Prowell, 2005) (Fig. S1). In total, we sampled 74 quadrats (BB: 10, EC: 10, JC: 12, MP: 10, NF: 16, CL: 16). Different quadrats were separated by ⩾300 m if placed in the same tree-cover stand. Samples were stored in 100% ethanol at ambient temperature until shipment to the lab, where they were stored at -20 ℃ before DNA extraction. The original reason for using pan traps had been to trap bees, which we individually DNA-barcoded in Hua et al. (2016). Here we analyze the bycatch.

2.2 Amplicon preparation

For each of the 74 quadrats, we pooled all 40 pan traps into a single sample. Three quadrats had very few individuals, and we pooled them with their nearest neighbor of the same land-cover type (EC01+EC02+EC03; NF02+NF03), leaving us with 71 samples. Storage ethanol was removed by air drying on single-use filter papers. Our samples were dominated by Diptera and Hymenoptera, as expected. We equalized input DNA across species by using one leg of every individual larger than a mosquito (~5 mm long) and the whole body if smaller (e.g. midges). This was to reduce the effect of large-biomass individuals outcompeting small-biomass individuals during PCR, which improves taxon detection (Elbrecht, Peinert, & Leese, 2017). DNA extraction followed the protocols of Qiagen DNeasy Blood&Tissue Kits (Hilden, Germany), followed by quantification via Nanodrop 2000 (Thermo Fisher Scientific, Wilmington, DE).

We amplified a 319-bp fragment of COI using forward primer LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and reverse primer mlCOIintR (5'-GGNGGRTANANNGTYCANCCNGYNCC-3') (Leray et al., 2013). All samples were carried out with two rounds of PCR. In the first round, both forward and reverse primers were tailed with tags (12-17 bp) for sample identification. In the second round, we added Illumina adapters to the amplicons from the first PCR, thus avoiding the tag jumping that can arise during library preparation of amplicon mixtures (Schnell, Bohmann, & Gilbert, 2015). A table of tags and primers is in Supplementary Information (Table S1). All PCRs were performed on a Mastercycler Pro (Eppendorf, Germany) in 20-μl reaction volumes, each containing 2 μl 10x buffer (Mg^{2+} plus), 0.2 mM dNTPs, 0.4 μM of each primer, 1 μl DMSO, 0.4 μl BSA (bovine serum albumin) (TaKaRa Biotechnology Co. Ltd, Dalian, China), 0.6 U exTaq DNA polymerase (TaKaRa Biotechnology), and approximately 60 ng genomic DNA. Both rounds of PCR started with an initial denaturation at 94 ℃ for 4 mins, followed by 35 cycles of

94 ℃ for 45s, 45 ℃ for 45s, 72 ℃ for 90s, and finishing at 72 ℃ for 10 mins. PCR products were gel-purified with QIAquick PCR Purification Kit (Qiagen). One sample failed to amplify. We pooled the 70 PCR products into two libraries and sequenced on the Illumina MiSeq (Reagent Kit V3, 300PE) at the Southwest Biodiversity Institute Regional Instrument Center in Kunming. The total number of paired-end reads returned was 13,601,908.

2.3 Data analyses

The bioinformatic script, including parameters, for the analyses below is in Supplementary Information and will be archived in datadryad.org, along with sequence data and metadata. The *R* scripts and data tables are on https://github.com/dougwyu/Sichuan2014. Below, *R* packages are indicated with single quotes, and other software is italicized.

2.3.1 Bioinformatic processing

Initial processing. – We removed remnant Illumina adapter sequences with *AdapterRemoval* 2.2.0 (Schubert, Lindgreen, & Orlando, 2016), followed by Schirmer et al.'s (2015) pipeline to filter, trim, denoise, and merge read pairs. Specifically, we trimmed low-quality ends using *sickle* 1.33 (Joshi & Fass, 2011), corrected sequence errors using *BayesHammer* in *SPAdes* 3.10.1 (Nikolenko, Korobeynikov, & Alekseyev, 2013), and merged reads using *PandaSeq* 2.11 (Masella, Bartram, Truszkowski, Brown, & Neufeld, 2012), all with default parameters.

Demultiplexing and Clustering. – We then used *QIIME* 1.9.1's *split_libraries.py* (Caporaso et al., 2010) to demultiplex reads by sample and used *usearch* 9.2.64 (Edgar, 2010) to retain reads between 300 and 330 bp, inclusive, since our amplicon is 319 bp. We used *vsearch* 2.4.3 (Rognes, Flouri, Nichols, Quince, & Mahé, 2016) for *de-novo* chimera removal and used *CROP* 1.33 (Hao, Jiang, & Chen, 2011) to cluster the remaining reads at 97%-similarity. This step produced 3,507 OTUs. We also tried *swarm* 2.2.2 (Mahé, Rognes, Quince, Vargas, & Dunthorn, 2015), but it returned huge numbers of OTUs that could not be reduced even after running through 'lulu' (see below).

OTU filtration and taxonomic assignment. – From the resulting sample X OTU table, we used 'lulu' 0.1.0 (Frøslev et al., 2017) to combine OTUs that were likely from the same species but which had failed to be clustered by *CROP*. 'lulu' identifies such 'parent-child' sets by calculating pairwise similarities of all OTUs (using *vsearch*) to identify sets of high-similarity OTUs and then combining OTUs within such sets that show nested sample distributions. For example, four OTUs might be highly similar, and within this set of four, one OTU contains the most reads and is observed in ten samples. This OTU is the parent, and daughters are inferred if they are present in a subset of the parent's samples. We ended with 1,506 OTUs.

A common filtering step is to remove OTUs made up of few reads (e.g. 1-read OTUs), as these are more likely to be artefactual (e.g. Yu et al., 2012, Zepeda-Mendoza et al., 2016). For instance, PCR errors can generate clusters of sequences that are sufficiently different from the

parent that they cannot be identified as daughters. Such OTUs are more likely to be small because novel haplotypes typically arise in a later PCR cycle. However, the definition of small is subjective and differs with the size of the sequence dataset. We therefore used 'phyloseq' 1.19.1 (McMurdie & Holmes, 2013) to plot the number of OTUs that would be filtered out at different minimum OTU sizes (see http://evomics.org/wp-content/uploads/2016/01/phyloseq-Lab-01-Answers.html, accessed 19 July 2018), and we chose a minimum OTU size of 44 reads, which was roughly the graph's inflection point and thus filtered out the most OTUs for the lowest minimum size. We ended with 594 OTUs.

We then used *PyNAST* 1.2.2 to align the 594 OTU sequences to a reference alignment of Arthropoda COI sequences (Yu et al., 2012) at a minimum similarity of 60%; one sequence failed to align and was deleted. The remaining sequences were translated to amino acids using the invertebrate mitochondrial codon table, and we removed 32 OTUs with sequences that contained stop codons. We carried out taxonomic assignment of the OTUs using a Naïve Bayesian Classifier (Wang, Garrity, Tiedje, & Cole, 2007) trained on the Midori UNIQUE COI dataset (Machida, Leray, Ho, & Knowlton, 2017). Sixteen OTUs assigned to non-Arthropoda taxa and two OTUs assigned to Collembola were removed. We ended with 543 OTUs.

Finally, we inspected the OTU table and set to zero those cells that had <5 reads representing that OTU in that sample, since these were more likely to be the result of sequencing error (Yu et al., 2012). In addition, we removed two samples (rows) that contained ≤100 reads total (i.e. samples with little data) and removed seven samples (rows) with <5 OTUs because these samples were potentially overly influential in analyses of species richness. These seven samples included two from native forests and five from monocultures (3 BB, 1 EC, 1 JC), meaning that we disproportionately removed monocultures, making our species diversity analyses below more conservative. After these sample removals, seven OTUs were removed because they were left with few (<20) reads. Because we do not consider OTU size to be reliable measures of biomass or abundance (Nichols et al., 2018; Piñol, Mir, Gomez-Polo, & Agustí, 2015; Yu et al., 2012), we converted the OTU table into a presence/absence (0/1) dataset. Throughout, our bias was to remove false-positive detections even at the expense of losing true-positive detections, thereby resulting in a dataset with less, but more reliable (and thus more replicable), data. We ended with 536 OTUs and 61 samples.

2.3.2 Community analysis

OTU richness and diversities. – All community analyses were performed in *R* 3.3.3 (R Core Team, 2017). We estimated species richness and Shannon and Simpson diversities using two sample-based estimators: function *specpool* in 'vegan' 2.4-5 (Chiu, Wang, Walther, & Chao, 2014) and 'iNEXT' 2.0.12 (Hsieh, Ma, & Chao, 2016).

OTU phylogenetic diversities. – Because we used a combination of *CROP*+'lulu' and 'phyloseq' to combine and remove small OTUs that were likely to be artefactual, the

remaining OTUs were more likely to represent true presences. Nonetheless, it remained possible that we had over-split some biological species into multiple OTUs, since there is no single correct similarity threshold for species delimitation, and this oversplitting might have occurred more often for some taxa in some land-cover types, leading to artifactual differences in species richness. However, oversplit OTUs should cluster together in a phylogenetic tree and thus contribute less to estimates of *phylogenetic* diversity than would OTUs from different biological species. Phylogenetic diversity should thus be a robust estimator of alpha diversity (Yu et al., 2012). To estimate sample phylogenetic diversities, we used 'iNextPD' 0.3.2 (Hsieh & Chao, 2017). We built a maximum-likelihood (ML) tree in *RaxML* 8.0.0 (Stamatakis, 2014) with an alignment of the OTU-representative sequences, using a General Time Reversible (GTR) model of nucleotide substitution and a gamma model of rate heterogeneity estimating the proportion of invariable sites (-m GTRGAMMAI). The algorithm used a rapid bootstrap analysis and searched for the best-scoring ML tree (-f a), with -N 1000 times bootstrap and -p 12345 as the parsimony random seed. Three OTU sequences produced very long branches in the ML tree, which would skew estimates of phylogenetic diversity, and we removed them. Two of these OTUs were found in all land-cover types (and thus would not have been informative), and one was only found in some cropland samples (and thus would not have informed analyses of the tree-cover sites).

Beta diversity. – To visualize changes in community composition across land-cover types, we ran a Bayesian ordination with 'boral' 1.6.1 (Hui, 2016), which is more statistically robust than non-metric multidimensional scaling (NMDS) analysis because 'boral' is model-based and thus allows us to apply a suitable error distribution so that fitted-model residuals are properly distributed. We used a binomial error distribution and no row effect since we were using presence/absence data (Figure S5). For the same reasons, we used 'mvabund' 3.12.3 (Wang, Naumann, Wright, & Warton, 2012) to test the hypotheses that native forests and mixed plantations differ compositionally from each other and differ from the monocultures and croplands.

We also visualized changes in community composition with an 'UpSetR' 1.3.3 intersection diagram, an alternative to Venn diagrams (Conway, Lex, & Gehlenborg, 2017), with a heatmap using the *tabasco* function in 'vegan', and with a 'betapart' 1.4-1 (Baselga & Orme, 2012) analysis, which partitions beta diversity into turnover and nestedness components using binary Jaccard dissimilarities, which we visualized with NMDS using the *metaMDS* function in 'vegan'. Finally, we used 'metacoder' 0.2.0 (Foster, Sharpton, & Grunwald, 2017) to generate taxonomic 'heat trees' to pairwise-compare the six land-cover types and identify the taxa most strongly driving compositional differences.

3 RESULTS

3.1 Alpha diversity

Species richness and diversity are highest in native forests and croplands, followed by mixed plantations, which are in turn richer and more diverse than the monoculture plantations, with the possible exception of bamboo.

OTU richness and diversities. – The Chao2 estimator indicates that native forests, mixed plantations, and croplands have the highest estimated species richnesses and do not differ significantly from each other (Figure 2a). Importantly, all three monocultures (bamboo, Eucalyptus, and Japanese cedar) exhibit less than half the species richness of native forests and around half the species richness of mixed plantations (Figure 2a). The pairwise differences between native forests and monocultures are all statistically significant (Table S2), and the pairwise differences between mixed plantations and the three monocultures are marginally or significantly different (Figure 2a, Table S2), all after table-wide correction.

The iNEXT analysis reveals even clearer contrasts: native forests have the highest estimated asymptotic species richnesses and Shannon diversities, followed by croplands and mixed plantations, followed by the three monocultures (Figures 2b, S3). The iNEXT-estimated richness and diversity of mixed plantations are significantly higher than all the monocultures, with the possible exception of bamboo, because the MP and BB confidence intervals touch.

Phylogenetic diversities. – The iNextPD analysis mirrors the iNEXT results (Figures 2b, S4). Using 'iNextPD' to visualize phylogenetic coverage by land-cover type (Figure 3) reveals that native forests and croplands exhibit almost complete coverage of the OTU tree, whereas mixed plantations and bamboo exhibit some coverage deficits, followed by larger coverage deficits in the other two monocultures.

3.2 Beta diversity

Native forests are compositionally most similar to mixed plantations and most dissimilar to croplands. The differences in community composition are driven primarily by species turnover.

Differences in community compositions. Ordination with 'boral' (Figure 4a) shows that the primary separation is between the tree cover types and croplands, with a significantly positive correlation between latent variable 1 and elevation ($r = -0.457$, $df = 59$, $p = 0.0002$). The cropland sites themselves cluster into two groups by elevation. Latent variable 2 largely separates Eucalyptus monoculture from the other tree-cover types, which might reflect its distinct phytochemistry. Importantly, the mixed-plantation and (most of) the native-forest sites overlap and are encircled by the monocultures, indicating that native forests and mixed plantations are compositionally most similar.

The ‘UpSetR’ intersection diagram (Figure 4b) is consistent with the diversity analyses (Figures 2, S3, S4): native forests (110 OTUs) and croplands (130 OTUs) support more than 2.5 times the number of ‘unique species’ (species detected in only one land-cover type) than any of the plantations, and secondly, of the plantations, mixed plantations support the highest number of unique species (44 OTUs). The greater compositional similarity that native forests have with mixed plantations (Figure 4a) is displayed by native forests uniquely sharing more OTUs with mixed plantations (22 OTUs) than with any of the monocultures (13, 9, and 5). However, despite their overlap, ‘mvabund’ analysis shows that the arthropod communities of mixed plantations and native forests are still significantly distinct from each other, and from the three monocultures and croplands (Table S3).

Turnover versus nestedness. – Consistent with the UpSetR result that the mode in each land-cover type is unique species, we found that turnover, not nestedness, dominates compositional differences (Figure 5; see Figure S7 for a heatmap visualisation). In other words, the arthropod communities in the monocultures are not simply subsets of native forests or mixed plantations but contain distinct sets of species.

Taxonomic compositions of and differences between land-cover types. – The 536 arthropod species in our metabarcoding dataset represent a wide range of arachnid and insect orders and thus, represent a wide range of ecological functions (Figure 6), including generalist predators (Araneae, Formicidae) and more specialized parasites and parasitoids (Tachinidae, Phoridae, Braconidae) of other arthropods. We also observe taxa that are noted for pollination (Thysanoptera, Syrphidae), xylophagy (Isoptera), and various modes of detritivory, fungivory, frugivory, herbivory, and animal parasitism (Lepidoptera, Hemiptera, Diptera, Orthoptera, Formicidae, Thysanoptera).

Although the ‘boral’ ordination (Figure 4a) reveals compositional similarity between mixed plantations and native forests, it does not reveal the taxa that are most responsible for this similarity, and for the differences with the other tree-cover types. With ‘metacoder’ heat trees (Figure 6 inset), we can identify the taxa that are driving this similarity and the differences, and what we see is that mixed plantations and native forests ‘differ in the same ways’ from the monocultures. (1) Relative to bamboo, mixed plantations and native forests both have slightly more Lepidoptera OTUs. (2) Relative to Eucalyptus, mixed plantations and native forests both have more Diptera OTUs and fewer of the three OTUs assigned to genera *Mycetophila*, *Sonema*, and *Homaloxestis*, which can be taken as Eucalyptus indicator species. (3) Finally, relative to Japanese cedar, mixed plantations and native forests both have more Araneae and Lepidoptera OTUs, fewer Hemiptera OTUs, and fewer of the OTU assigned to *Mycetophila*. Heat-tree differences at higher taxonomic ranks (e.g. more Araneae-assigned OTUs) mean that the species which separate the two land-cover types differ across samples but nonetheless are in the same higher taxon (e.g. Araneae). Finally, when we include croplands in the heat-tree comparisons (Figure S8), we observe the largest number of heat-tree-tip differences between any two land-cover types. In other words, there are multiple

species-level indicators of croplands (or in the case of the *Mycetophila* OTU, an indicator of Japanese cedar and Eucalyptus).

4 DISCUSSION

Improving biodiversity conservation under the GFGP

Our study found that native forests support the highest levels of arthropod species richness, Shannon and Simpson diversity, and Faith's and phylogenetic diversity (Figures 2, 3, S3, S4) and that most of those species are unique to native forests (Figure 4b), consistent with the patterns of bird diversity that were reported in Hua et al. (2016) and other biodiversity studies in plantations (Barlow et al., 2007; Gardner et al., 2008). In addition, our findings pertaining to the higher level of alpha diversity in mixed plantations over monocultures (Figures 2, S3, S4), and their greater degree of compositional similarity to native forests relative to monocultures (Figures 4, 5, 6, S6, S7, S8), corroborate those reported for birds (but not bees) in Hua et al. (2016) and are consistent with other studies of biodiversity in tree plantations. Butterfield and Malvido (1992) showed that mixtures of broadleafs and conifers resulted in a higher species richness of carabid beetles than in conifer monocultures, and Recher et al. (1987) showed that some bird species are present when in Eucalyptus-pine mixtures but absent from pine monocultures. In short, mixed plantations not only support a higher diversity of non-breeding birds but also provide a small but detectable biodiversity boost for arthropods. Finally, we found that compositional differences amongst tree-cover types are almost entirely dominated by species turnover, not nestedness, meaning that some species were only detected in the monocultures. This result is consistent with the pattern of moth communities in primary, secondary and plantation forests studied by Hawes et al. (2009). In their findings, all three of their tree-cover types (primary and secondary forest, Eucalyptus plantation) contained large numbers of unique species in three moth families (Arctiidae, Saturniidae, Sphingidae).

Given the balance of evidence, we re-affirm our previous policy recommendations that the GFGP should prioritize the retention and restoration of native forests, and when restoring native forests is not possible, we secondarily encourage mixed-species plantings over extensive monocultures, at least in western China where we conducted this study. The foundation of these recommendations is now broadened to include 536 species-resolution taxa ranging across the Arthropoda. Given the growing understanding of the biodiversity implications of plantations compared with native forests in different forest biomes across the world (Bremer & Farley, 2010; Fierro, Grez, Vergara, Ramírez-Hernández, & Micó, 2017), these recommendations likely apply to other regions in China where GFGP is relevant, but their applicability will benefit from additional field studies and from anticipated technical advances in DNA-based biodiversity assessment. In the future, it will likely be insightful to carry out time-series biodiversity surveys, since our dataset represents only a single time point, but the temporal turnover of forest arthropod communities is high (Barsoum et al.,

2019). It is possible that the differences in biodiversity levels that we have detected are even stronger when integrated over time. Another important variable that we did not measure is sample biomass, given recent evidence that insect biomass has been dropping around the world (e.g. Hallmann et al., 2017). Because we observed high species richness and diversity in our cropland sampling sites (Figures 2, 3, S3, S4), where agriculture is small-scale in nature, our *a priori* expectation is that biomass has probably not declined here as rapidly as elsewhere, but this clearly needs testing and should of course now be a standard metric in biodiversity surveys.

Greater levels of arthropod biodiversity in native forest is not a surprise, given their more diverse vegetation structures and species compositions, which are well known to be positively correlated with arthropod diversity (Castagneyrol & Jactel, 2012; Haddad et al., 2009; Stork, Mcbroom, Gely, & Hamilton, 2015; Zhang et al., 2016), but the greater diversity and similarity of mixed plantations to native forests is somewhat surprising, especially since they mostly just comprise small-scale monocultures, planted in checkerboard pattern. However, planting different tree species near each other not only provides more diverse vegetation per se but also, because the species vary in height and three-dimensional structure, almost certainly allow greater sunlight penetration to the understory, which in turn should result in greater availability of food and other resources. This mechanism is consistent with our finding that bamboo, which does not create closed canopies, exhibits the highest richness and diversity of the monocultures (Figures 2, S3, S4). We note that 95% confidence-interval overlap is considered an overly conservative test for statistical significance at the p=0.05 level (MacGregor-Fors & Payton, 2013). A more diverse, and presumably higher-biomass, arthropod community in turn could also support a richer bird community, at least for the insectivorous subset of the community. Our results thus point to a plausible mechanism for why bird diversity is boosted in mixed plantations.

In this study, we report evidence for a biodiversity benefit of native forests over GFGP plantations, which we might think trades off against a greater value of timber sales from plantations. However, even excluding biodiversity, which they did not study, Cao et al. (2019) have recently shown that plantations in China also return a lower net value of other ecosystem services relative to native forests, even after counting income from timber sales. Plantations require a high initial outlay for tree planting, some non-native tree species like Eucalyptus require more water input than do native tree species, and more management effort is required to protect plantations from pest attack. In contrast, timber sale values are low. Cao et al.'s findings complement and strengthen our recommendation (Hua et al., 2016) to prioritize native forest recovery and expansion over creating plantations.

Methodological comments on metabarcoding and studies of biodiversity patterns.

Metabarcoding provides an efficient method for interrogating biodiversity samples, but because of its reliance on PCR, metabarcoding datasets tend to contain a non-trivial amount

of noise. This noise manifests as a large number of false-positive OTUs, which are filtered out heuristically. Such false OTUs especially complicate efforts to estimate alpha diversity. Here, we applied several filtering steps to remove false OTUs, and we also used 'iNextPD' to generate robust comparisons of alpha diversity by estimating phylogenetic diversity instead of species richness. This approach has been previously shown to be reliable (Yu et al., 2012). Another approach, which became available only after we had completed the wet-lab portion of our study, is to subject each sample to multiple, independently tagged PCRs (typically three) and to bioinformatically filter out sequences that fail to appear in at least two of the PCRs above some minimum number of reads; such sequences are more likely to be PCR or sequencing errors. This is implemented in the DAMe protocol of Zepeda-Mendoza et al. (2016, also see Alberdi, Aizpurua, Gilbert, & Bohmann, 2018).

With regard to studies of biodiversity patterns, we follow Magurran et al. (2015; Magurran, 2016) in recommending that we should focus less on explaining change in species *richness* and more on explaining change in species *composition* as a function of natural and anthropogenic causes. The argument is that anthropogenically disturbed communities can maintain species richness and even phylogenetic diversity, even as local, or worse still, endemic, species go extinct and are replaced by cosmopolitan species. In our study, croplands support an arthropod community similar in richness and diversity to that of mixed plantations and just below that of native forests (Figures 2, 3, 4b, S3, S4), but the species composition of croplands is distinct from those in native forests (Figures 4, 5, S6, S7, S8). Croplands therefore cannot compensate for the loss of the biodiversity dependent on native forests.

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*, 134–147. https://doi.org/10.1111/2041-210X.12849

Allan, D. G., Harrison, J. A., Navarro, R. A., Wilgen, B. W. Van, & Thompson, M. W. (1997). The impact of commercial afforestation on bird populations in Mpumalanga Province, South Africa—insights from bird-atlas data. *Biological Conservation*, *79*, 173–185.

Barlow, J., Gardner, T. A., Araujo, I. S., Avila-Pires, T. C., Bonaldo, A. B., Costa, J. E., … Peres, C. A. (2007). Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. *PNAS*, *104*, 18555–18560.

Barlow, J., Overal, W. L., Araujo, I. S., Gardner, T. A., & Peres, C. A. (2007). The value of primary, secondary and plantation forests for fruit-feeding butterflies in the Brazilian Amazon. *Journal of Applied Ecology*, *44*, 1001–1012.

Barsoum, N., Bruce, C., Forster, J., Ji, Y., Yu, D. W., Lodge, A. H., … Kingdom, U. (2019). The devil is in the detail: Metabarcoding of arthropods provides a sensitive measure of biodiversity response to forest stand composition compared with surrogate measures of biodiversity. *Ecological Indicators*, *101*, 313–323. https://doi.org/10.1016/j.ecolind.2019.01.023

Bartholomew, C. S., & Prowell, D. (2005). Pan compared to malaise trapping for bees (Hymenoptera: Apoidea) in a longleaf pine savanna. *Journal of the Kansas Entomological Society*, *78*, 390–392.

549 Baselga, A., & Orme, C. D. L. (2012). Betapart: An R package for the study of beta diversity. *Methods in*
550 *Ecology and Evolution*, *3*, 808–812. https://doi.org/10.1111/j.2041-210X.2012.00224.x
551 Bremer, L. L., & Farley, K. A. (2010). Does plantation forestry restore biodiversity or create green deserts? A
552 synthesis of the effects of land-use transitions on plant species richness. *Biodiversity and Conservation*,
553 *19*, 3893–3915. https://doi.org/10.1007/s10531-010-9936-4
554 Brockerhoff, E. G., Jactel, H., Parrotta, J. A., Quine, C. P., & Sayer, J. (2008). Plantation forests and
555 biodiversity: oxymoron or opportunity? *Biodiversity and Conservation*, *17*, 925–951.
556 https://doi.org/10.1007/s10531-008-9380-x
557 Bryan, B. A., Gao, L., Ye, Y., Sun, X., Connor, J. D., Crossman, N. D., … Hou, X. (2018). China's response to
558 a national land-system sustainability emergency. *Nature*, *559*, 193–204. https://doi.org/10.1038/s41586-
559 018-0280-2
560 Buscardo, E., Smith, G. F., Kelly, D. L., Freitas, H., Iremonger, S., Mitchell, F. J. G., … Mckee, A. (2008). The
561 early effects of afforestation on biodiversity of grasslands in Ireland. *Biodiversity and Conservation*, *17*,
562 1057–1072. https://doi.org/10.1007/s10531-007-9275-2
563 Butterfield, J., & Malvido, J. B.. (1992). Effect of mixed-species tree planting on the distribution of soil
564 invertebrates. *Special Publications Series of the British Ecological Society*, *11*, 255-268
565 Cao, S., Zhang, J., & Su, W. (2019). Difference in the net value of ecological services between natural and
566 artificial forests in China. *Conservation Biology*, *00*, 1–8.
567 Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Knight, D.
568 (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*, 335–
569 336. https://doi.org/10.1038/nmeth.f.303.QIIME
570 Castagneyrol, B., & Jactel, H. (2012). Unraveling plant – animal diversity relationships: a meta-regression
571 analysis. *Ecology*, *93*, 2115–2124.
572 Chiu, C. H., Wang, Y. T., Walther, B. A., & Chao, A. (2014). An improved nonparametric lower bound of
573 species richness via a modified good-turing frequency formula. *Biometrics*, *70*, 671–682.
574 https://doi.org/10.1111/biom.12200
575 Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting
576 sets and their properties. *Bioinformatics*, *33*, 2938–2940. https://doi.org/10.1093/bioinformatics/btx364
577 Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: Towards
578 an integrative approach to the study of global biodiversity. *Trends in Ecology and Evolution*, *29*, 566–571.
579 https://doi.org/10.1016/j.tree.2014.08.001
580 Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L.
581 (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities.
582 *Molecular Ecology*, *26*, 5872–5895. https://doi.org/10.1111/mec.14350
583 Delang, C. O., & Yuan, Z. (2015). *China's Grain for Green Program*. Springer US.
584 Deng, L., Liu, G. bin, & Shangguan, Z. ping. (2014). Land-use conversion and changing soil carbon stocks in
585 China's "Grain-for-Green" Program: A synthesis. *Global Change Biology*, *20*, 3544–3556.
586 https://doi.org/10.1111/gcb.12508
587 Deng, L., Shangguan, Z. ping, & Li, R. (2012). Effects of the grain-for-green program on soil erosion in China.
588 *International Journal of Sediment Research*, *27*, 120–127. https://doi.org/10.1016/S1001-6279(12)60021-
589 3
590 Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*, 2460–

591 2461. https://doi.org/10.1093/bioinformatics/btq461
592 Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass
593 on DNA metabarcoding. *Ecology and Evolution*, *7*, 6918–6926. https://doi.org/10.1002/ece3.3192
594 Elsen, P. R., Ramesh, K., & Wilcove, D. S. (2018). Conserving Himalayan birds in highly seasonal forested and
595 agricultural landscapes. *Conservation Biology*, *32*, 1313–1324.
596 Fierro, A., Grez, A. A., Vergara, P. M., Ramírez-Hernández, A., & Micó, E. (2017). How does the replacement
597 of native forest by exotic forest plantations affect the diversity, abundance and trophic structure of
598 saproxylic beetle assemblages? *Forest Ecology and Management*, *405*, 246–256.
599 https://doi.org/10.1016/j.foreco.2017.09.026
600 Foster, Z. S. L., Sharpton, T., & Grunwald, N. J. (2017). MetacodeR: An R package for manipulation and heat
601 tree visualization of community taxonomic data from metabarcoding. *PLoS Computational Biology*,
602 *13*, e1005404. https://doi.org/10.1101/071019
603 Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., & Hansen, A. J. (2017).
604 Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates.
605 *Nature Communications*, *8*, 1-11. https://doi.org/10.1038/s41467-017-01312-x
606 Gardner, T. A., Hernandez, M. I. M., Barlow, J., & Peres, C. A. (2008). Understanding the biodiversity
607 consequences of habitat change: the value of secondary and plantation forests for neotropical dung
608 beetles. *Journal of Applied Ecology*, *45*, 883–893.
609 Haddad, N. M., Crutsinger, G. M., Gross, K., Haarstad, J., Knops, J. M. H., & Tilman, D. (2009). Plant species
610 loss decreases arthropod diversity and shifts trophic structure. *Ecology Letters*, *12*, 1029–1039.
611 https://doi.org/10.1111/j.1461-0248.2009.01356.x
612 Hallmann, C. A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Sumser, H., … Kroon, H. De. (2017). More
613 than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE*, *12*,
614 e0185809.
615 Hao, X., Jiang, R., & Chen, T. (2011). Clustering 16S rRNA for OTU prediction: A method of unsupervised
616 Bayesian clustering. *Bioinformatics*, *27*, 611–618. https://doi.org/10.1093/bioinformatics/btq725
617 Hartley, M. J. (2002). Rationale and methods for conserving biodiversity in plantation forests. *Forest Ecology*
618 *and Management*, *155*, 81–95.
619 Hawes, J., Motta, S., Overal, W. L., Barlow, J., Gardner, T. A., & Peres, C. A. (2009). Diversity and
620 composition of Amazonian moths in primary, secondary and plantation forests. *Journal of Tropical*
621 *Ecology*, *25*, 281–300. https://doi.org/10.1017/S0266467409006038
622 Hsieh, T. C., & Chao, A. (2017). Rarefaction and extrapolation: Making fair comparison of abundance-sensitive
623 phylogenetic diversity among multiple assemblages. *Systematic Biology*, *66*, 100–111.
624 https://doi.org/10.1093/sysbio/syw073
625 Hsieh, T. C., Ma, K. H., & Chao, A. (2016). iNEXT: an package for rarefaction and extrapolation of species
626 diversity (Hill numbers). *Methods in Ecology and Evolution*, *7*, 1451–1456. https://doi.org/10.1111/2041-
627 210X.12613
628 Hu, C., Fu, B., Chen, L., & Gulinck, H. (2006). Farmer's attitudes towards the Grain-for-Green programme in
629 the Loess hilly area, China: A case study in two small catchments. *International Journal of Sustainable*
630 *Development & World Ecology*, *13*, 211–220. https://doi.org/10.1080/13504500609469673
631 Hua, F., Wang, L., Fisher, B., Zheng, X., Wang, X., Yu, D. W., … Wilcove, D. S. (2018). Tree plantations
632 displacing native forests: The nature and drivers of apparent forest recovery on former croplands in

633 Southwestern China from 2000 to 2015. *Biological Conservation*, *222*, 113–124.
634 https://doi.org/10.1016/j.biocon.2018.03.034
635 Hua, F., Wang, X., Zheng, X., Fisher, B., Wang, L., Zhu, J., … Wilcove, D. S. (2016). Opportunities for
636 biodiversity gains under the world's largest reforestation programme. *Nature Communications*, *7*, 1–11.
637 https://doi.org/10.1038/ncomms12717
638 Hui, F. K. C. (2016). boral – Bayesian ordination and regression analysis of multivariate abundance data in R.
639 *Methods in Ecology and Evolution*, *7*, 744–750. https://doi.org/10.1111/2041-210X.12514
640 Jactel, H., & Brockerhoff, E. G. (2007). Tree diversity reduces herbivory by forest insects. *Ecology Letters*, *10*,
641 835–848. https://doi.org/10.1111/j.1461-0248.2007.01073.x
642 Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable,
643 verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*, 1245–1257.
644 https://doi.org/10.1111/ele.12162
645 Joshi NA, & Fass JN. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files
646 (Version 1.33) *[Software]*. Available at https://github.com/najoshi/sickle.
647 Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., … Machida, R. J. (2013). A new
648 versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding
649 metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, *10*,
650 1–14. https://doi.org/10.1186/1742-9994-10-34
651 Lindenmayer, D. B., & Hobbs, R. J. (2004). Fauna conservation in Australian plantation forests – a review.
652 *Biological Conservation*, *119*, 151–168. https://doi.org/10.1016/j.biocon.2003.10.028
653 Liu, J., & Diamond, J. (2005). China's environment in a globalizing world. *Nature*, *435*, 1179–1186.
654 https://doi.org/10.1038/4351179a
655 Liu, J., Li, S., Ouyang, Z., Tam, C., Chen, X., Liu, J., … Chen, X. (2008). Ecological and socioeconomic effects
656 of China's policies for ecosystem services. *PNAS*, *105*, 9477–9482.
657 https://doi.org/10.1073/pnas.0706436105
658 Liu, J., Ouyang, Z., Pimm, S. L., Raven, P. H., Wang, X., Miao, H., & Han, N. (2003). Protecting China's
659 biodiversity. *Science*, *300*, 1240–1241. https://doi.org/10.1126/science.1078868
660 Liu, Z., & Lan, J. (2015). The Sloping Land Conversion Program in China: Effect on the livelihood
661 diversification of rural households. *World Development*, *70*, 147–161.
662 https://doi.org/10.1016/j.worlddev.2015.01.004
663 Long, H. L., Heilig, G. K., Wang, J., Li, X. B., Luo, M., Wu, X. Q., & Zhang, M. (2006). Land use and soil
664 erosion in the upper reaches of the Yangtze River: Some socio-economic considerations on China's Grain-
665 for-Green Programme. *Land Degradation & Development*, *17*, 589–603. https://doi.org/10.1002/ldr.736
666 Ma, K., Shen, X., Grumbine, R. E., & Corlett, R. (2017). China's biodiversity conservation research in progress.
667 *Biological Conservation*, *210*, 1–2. https://doi.org/10.1016/j.biocon.2017.05.029
668 MacGregor-Fors, I., & Payton, M. E. (2013). Contrasting diversity values: Statistical inferences based on
669 overlapping confidence intervals. *PLoS ONE*, *8*, e56794. https://doi.org/10.1371/journal.pone.0056794
670 Machida, R. J., Leray, M., Ho, S. L., & Knowlton, N. (2017). Data descriptor: Metazoan mitochondrial gene
671 sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data*, *4*, 1–7.
672 https://doi.org/10.1038/sdata.2017.27
673 Magurran, A. E. (2016). How ecosystems change. *Science*, *351*, 448–449.
674 https://doi.org/10.1126/science.aad6758

675 Magurran, A. E., Dornelas, M., Moyes, F., Gotelli, N. J., & McGill, B. (2015). Rapid biotic homogenization of
676 marine fish assemblages. *Nature Communications*, *6*, 1–5. https://doi.org/10.1038/ncomms9405
677 Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2015). Swarm v2: highly-scalable and high-
678 resolution amplicon clustering. *PeerJ*, *3*, e1420. https://doi.org/10.7717/peerj.1420
679 Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G., & Neufeld, J. D. (2012). PANDAseq: Paired-
680 end assembler for illumina sequences. *BMC Bioinformatics*, *13*, 31-37. https://doi.org/10.1186/1471-
681 2105-13-31
682 McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and
683 graphics of microbiome census data. *PLoS ONE*, *8*, e61217. https://doi.org/10.1371/journal.pone.0061217
684 Nichols, R. V, Vollmers, C., Newsom, L. A., Wang, Y., Heintzman, P. D., Leighton, M., … Shapiro, B. (2018).
685 Minimizing polymerase biases in metabarcoding. *Molecular Ecology Resources*, *18*, 927–939.
686 https://doi.org/10.1111/1755-0998.12895
687 Nikolenko, S. I., Korobeynikov, A. I., & Alekseyev, M. A. (2013). BayesHammer: Bayesian clustering for error
688 correction in single-cell sequencing. *BMC Genomics*, *14*, 1-11. https://doi.org/10.1186/1471-2164-14-S1-
689 S7
690 Ouyang, Z., Zheng, H., Xiao, Y., Polasky, S., Liu, J., Xu, W., … Daily, G. C. (2016). Improvements in
691 ecosystem servcies from investments in natural capital. *Science*, *352*, 1455–1460. Retrieved from
692 http://www.gdrc.org/sustdev/concepts/26-nat-capital.html
693 Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use
694 of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology*
695 *Resources*, *15*, 819–830. https://doi.org/10.1111/1755-0998.12355
696 Pyne, K. (2013). Conserving China's biodiversity. *Earth Common Journal*, *3*, 44-53.
697 Recher, H. F., Davis, W. E., & Holmes, R. T. (1987). Ecology of brown and striated thornhills in forests of
698 south-eastern new south wales, with comments on forest management. *Emu - Austral Ornithology*, *87*, 1–
699 13. https://doi.org/10.1071/MU9870001
700 Ren, G., Young, S. S., Wang, L., Wang, W., Long, Y., Wu, R., … Yu, D. W. (2015). Effectiveness of China's
701 National Forest Protection Program and nature reserves. *Conservation Biology*, *29*, 1368–1377.
702 https://doi.org/10.1111/cobi.12561
703 Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for
704 metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584
705 Sayer, J. A., & Sun, C. (2003). The impacts of policy reforms on forest environments and biodiversity. In
706 *China's Forests: Global Lessons from Market Reforms* (pp. 177–194).
707 https://doi.org/10.4324/9781936331239
708 Schirmer, M., Ijaz, U. Z., D'Amore, R., Hall, N., Sloan, W. T., & Quince, C. (2015). Insight into biases and
709 sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Research*, *43*,
710 e37. https://doi.org/10.1093/nar/gku1341
711 Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated - reducing sequence-to-sample
712 misidentifications in metabarcoding studies. *Molecular Ecology Resources*, *15*, 1289–1303.
713 https://doi.org/10.1111/1755-0998.12402
714 Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: Rapid adapter trimming, identification,
715 and read merging. *BMC Research Notes*, *9*, 1–7. https://doi.org/10.1186/s13104-016-1900-2
716 Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large

717 phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033
718 Stork, N. E., Mcbroom, J., Gely, C., & Hamilton, A. J. (2015). New approaches narrow global species estimates
719 for beetles, insects, and terrestrial arthropods. *PNAS*, *112*, 7519–7523.
720 https://doi.org/10.1073/pnas.1502408112
721 Tao, Y., Huang, D., Jin, X., & Guo, L. (2010). Make an inventory of China's biodiversity. *Bulletin of the*
722 *Chinese Academy of Sciences*, *24*, 212–217.
723 Turner, B. L., Lambin, E. F., & Reenberg, A. (2007). The emergence of land change science for global
724 environmental change and sustainability. *PNAS*, *104*, 20666–20671.
725 United Nations. (2015). *Transforming our world: the 2030 agenda for sustainable development*. Retrieved from
726 https://sustainabledevelopment.un.org/post2015
727 Wang, J., Peng, J., Zhao, M., Liu, Y., & Chen, Y. (2017). Significant trade-off for the impact of Grain-for-Green
728 Programme on ecosystem services in North-western Yunnan, China. *Science of the Total Environment*,
729 *574*, 57–64. https://doi.org/10.1016/j.scitotenv.2016.09.026
730 Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naïve Bayesian classifier for rapid assignment of
731 rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*, 5261–
732 5267. https://doi.org/10.1128/AEM.00062-07
733 Wang, Y., Naumann, U., Wright, S. T., & Warton, D. I. (2012). Mvabund - an R package for model-based
734 analysis of multivariate abundance data. *Methods in Ecology and Evolution*, *3*, 471–474.
735 https://doi.org/10.1111/j.2041-210X.2012.00190.x
736 Wang, Z. J., Jiao, J. Y., Rayburg, S., Wang, Q. L., & Su, Y. (2016). Soil erosion resistance of "Grain for Green"
737 vegetation types under extreme rainfall conditions on the Loess Plateau, China. *Catena*, *141*, 109–116.
738 https://doi.org/10.1016/j.catena.2016.02.025
739 Wei, Y., Yu, D., Lewis, B. J., Zhou, L., Zhou, W., Fang, X., … Dai, L. (2014). Forest carbon storage and tree
740 carbon pool dynamics under natural forest protection program in northeastern China. *Chinese*
741 *Geographical Science*, *24*, 397–405. https://doi.org/10.1007/s11769-014-0703-4
742 Wu, R., Possingham, H. P. ., Yu, G., Jin, T., Wang, J., Yang, F., … Zhao, H. (2019). Strengthening China's
743 national biodiversity strategy to attain an ecological civilization. *Conservation Letters*, 1–7.
744 https://doi.org/10.1111/conl.12660
745 Xie, G., Cao, S., Yang, Q., Xia, L., Fan, Z., Chen, B., … Sarah, H. (2012). *China Ecological Footprint Report*
746 *2012*.
747 Xu, H., Wang, S., & Xue, D. (1999). Biodiversity conservation in China: legislation, plans and measures.
748 *Biodiversity and Conservation*, *8*, 819–837.
749 Xu, J., Yin, R., Li, Z., & Liu, C. (2006). China's ecological rehabilitation: Unprecedented efforts, dramatic
750 impacts, and requisite policies. *Ecological Economics*, *57*, 595–607.
751 https://doi.org/10.1016/j.ecolecon.2005.05.008
752 Yin, R., Liu, C., Zhao, M., Yao, S., & Liu, H. (2014). The implementation and impacts of China's largest
753 payment for ecosystem services program as revealed by longitudinal household data. *Land Use Policy*, *40*,
754 45–55. https://doi.org/10.1016/j.landusepol.2014.03.002
755 Yin, R., Yin, G., & Li, L. (2009). Assessing China's ecological restoration programs: What's been done and
756 what remains to be done? *Environmental Management*, *45*, 442–453. https://doi.org/10.1007/978-90-481-
757 2655-2_2
758 Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup:

759 Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology*
760 *and Evolution*, *3*, 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x
761 Zepeda-Mendoza, M. L., Bohmann, K., Carmona Baez, A., & Gilbert, M. T. P. (2016). DAMe: A toolkit for the
762 initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding
763 analyses. *BMC Research Notes*, *9*, 1–13. https://doi.org/10.1186/s13104-016-2064-9
764 Zhai, D. L., Xu, J. C., Dai, Z. C., Cannon, C. H., & Grumbine, R. E. (2014). Increasing tree cover while losing
765 diverse natural forests in tropical Hainan, China. *Regional Environmental Change*, *14*, 611–621.
766 https://doi.org/10.1007/s10113-013-0512-9
767 Zhang, K., Lin, S., Ji, Y., Yang, C., Wang, X., Yang, C., … Yu, D. W. (2016). Plant diversity accurately
768 predicts insect diversity in two tropical landscapes. *Molecular Ecology*, *25*, 4407–4419.
769 https://doi.org/10.1111/mec.13770
770 Zhou, H., Van Rompaey, A., & Wang, J. (2009). Detecting the impact of the “Grain for Green” program on the
771 mean annual vegetation cover in the Shaanxi province, China using SPOT-VGT NDVI data. *Land Use*
772 *Policy*, *26*, 954–960. https://doi.org/10.1016/j.landusepol.2008.11.006
773

774 Data Accessibility
775 Sequence data have been submitted to GenBank under accession number SAMN09981891.
776 Bioinformatic scripts are in Supplementary Information. R scripts and data tables are
777 available at https://github.com/dougwyu/Sichuan2014.
778

Figure legends

Figure 1. Study area in south-central Sichuan province, subdivided into counties and shaded by elevation. Each cross represents a pan-trap sampling location, color-coded by land-cover type: BB = bamboo monoculture, blue; EC = Eucalyptus monoculture, light green; CL = croplands, orange; JC = Japanese cedar monoculture, red; MP = mixed plantations, purple; NF = native forests, dark green.

Figure 2. Species richness estimates across land-cover type. (a) Comparisons of Chao2 species richness estimates. Land-cover types sharing the same superscript are not significantly different at the p=0.05 level (Welch's t-test) after table-wide correction for multiple tests (Bonferroni). (b) 'iNEXT' estimates of species richness, Shannon diversity, and 'iNextPD' estimates of phylogenetic diversity by land-cover type, using sample-based rarefaction and extrapolation. Native forests (NF) have the highest species richness and diversities, followed by croplands (CL) and mixed plantations (MP), followed by the three monoculture plantations (BB, EC, and JC). Codes for land-cover types as in Figure 1. Symbols on each curve indicate the number of sampled locations per land-cover type, solid lines represent interpolations, and dashed lines represent extrapolations, with 95% confidence intervals. Statistically significant pairwise differences are detected visually by non-overlapping confidence intervals and are considered conservative (MacGregor-Fors & Payton, 2013). Full iNEXT and iNextPD figures are in S3 and S4.

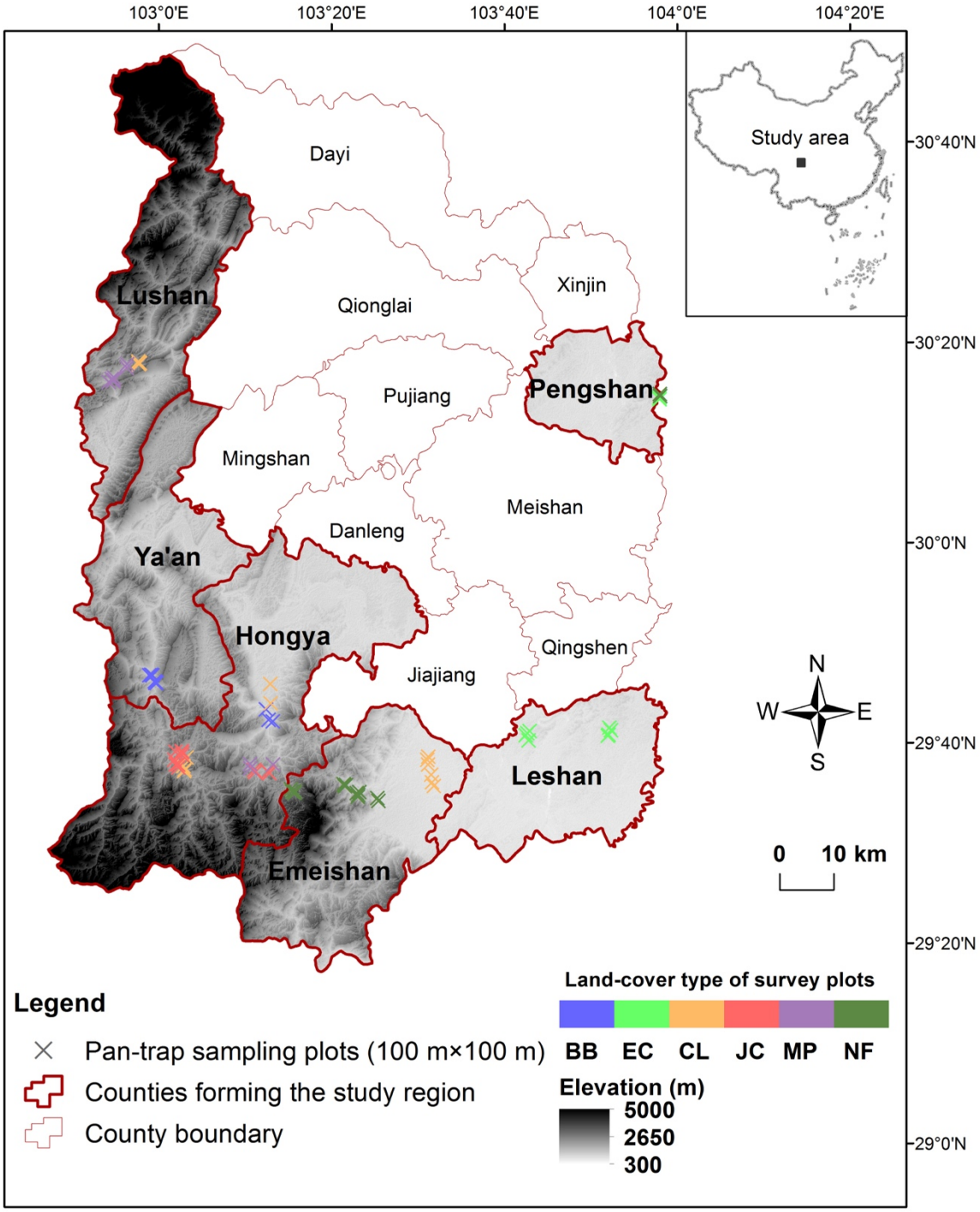
Figure 3. Phylogenetic distribution of OTUs by land-cover type, created using 'iNextPD'. Terminal nodes are black and represent the OTUs. Internal nodes are white. Sizes of the squares on the right indicate each OTU's incidence frequency (number of samples in which the OTU is observed). Phylogenetic coverage is most complete in native forests (NF) and croplands (CL), followed by mixed plantations (MP), followed by the three monocultures (BB, EC, JC). Codes for land-cover types as in Figure 1.

Figure 4. Community composition differences in all land-cover types. (a) 'Boral' ordination. Colors represent land-cover types, and numbers represent individual samples. Cropland (CL) sites separate into two clusters by elevation. Overlap of native forests (NF) and mixed-plantations (MP) points indicates greater compositional similarity between these two land-cover types. Ovals manually added to visualize community groupings. Residuals of the 'boral' fit in Fig. S5. (b) UpSetR intersection map of OTUs unique to and shared between and among land-cover types. Croplands and native forests support the highest numbers of unique OTUs (CL=130, NF=110), followed by the four plantations (MP=44, BB=37, EC=31, JC=27). Native forests uniquely share almost as many OTUs with mixed plantations (22 OTUs) as native forests share with the three monocultures combined (27 OTUs, =13+9+5). Horizontal bars on the left indicate the total number of OTUs in each land-cover class. Codes for land-cover types as in Figure 1. For clarity, only pairwise comparisons are shown. A non-truncated version is presented in Fig. S6.

Figure 5. NMDS (non-metric multidimensional scaling) ordination of beta diversity by land-cover type (binary Jaccard dissimilarities), partitioned with 'betapart'. (a) Total beta diversity. (b) Beta diversity based on species turnover only. (c) Beta diversity based on species nestedness only. Turnover accounts for most the observed beta diversity across land-cover types, which is visualized as greater distances between points in the turnover figure (b) and almost no distances between points in the nestedness figure (c). Codes for land-cover types as in Figure 1.

Figure 6. Pairwise taxonomic comparisons of all land-cover types. Upper right triangle: greener branches indicate taxa that are relatively more abundant (in numbers of OTUs) in the land-cover types along the right column, and browner branches indicate taxa that are relatively more abundant in the land-cover types along the top row. Lower left: taxonomic identities of the branches. Note that this is a taxonomic tree, not a phylogenetic tree. Legend: width indicates number of OTUs at a given taxonomic rank, and color indicates relative differences in $\log_2$(number of OTUs). Codes for land-cover types as in Figure 1. A figure including croplands and a zoomable taxonomic tree is in supplementary information (Figure S8, S9).

840 Figure 1.



841

842 Figure 1. Study area in south-central Sichuan province, subdivided into counties and shaded
843 by elevation. Each cross represents a pan-trap sampling location, color-coded by land-cover
844 type: BB = bamboo monoculture, blue; EC = Eucalyptus monoculture, light green; CL =
845 croplands, orange; JC = Japanese cedar monoculture, red; MP = mixed plantations, purple;
846 NF = native forests, dark green.

Figure 2

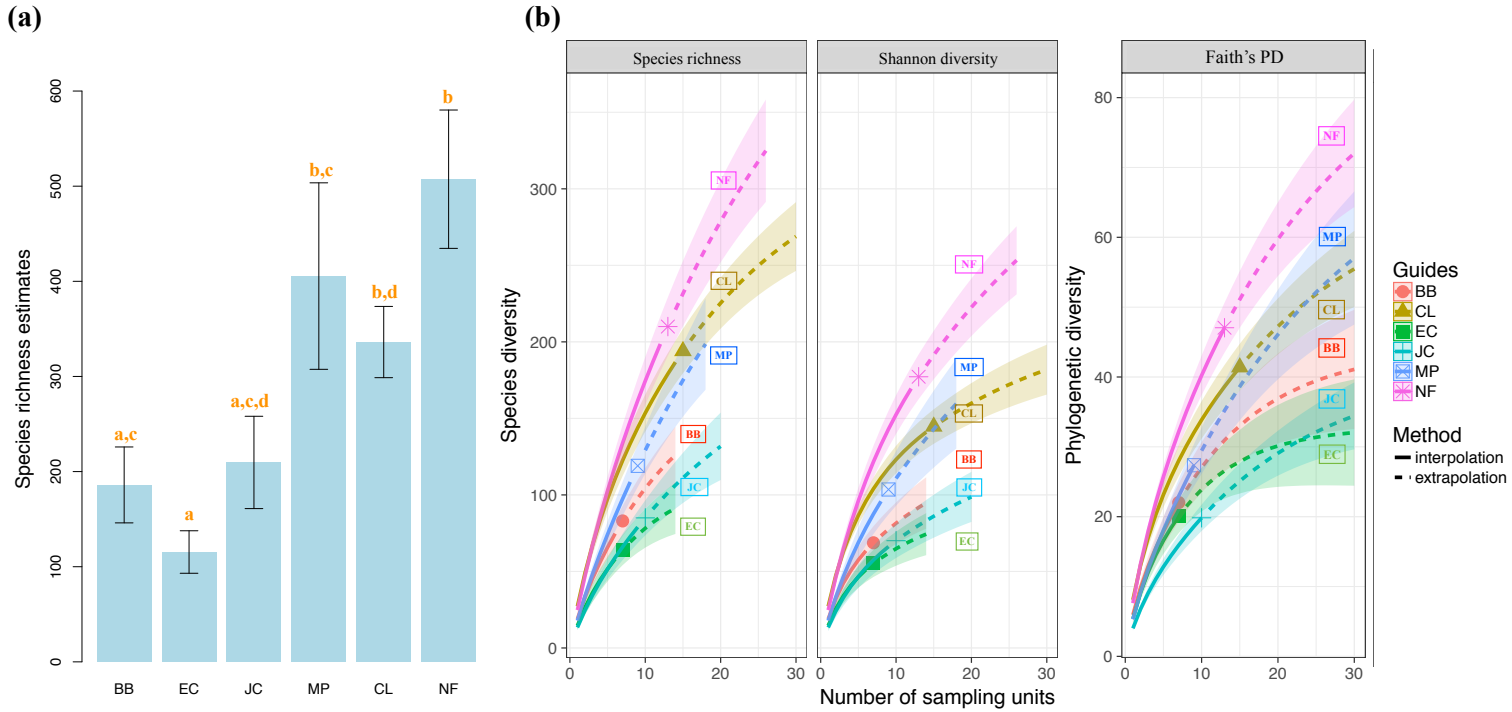


Figure 2. Species richness estimates across land-cover type. (a) Comparisons of Chao2 species richness estimates. Land-cover types sharing the same superscript are not significantly different at the p=0.05 level (Welch's t-test) after table-wide correction for multiple tests (Bonferroni). (b) 'iNEXT' estimates of species richness, Shannon diversity, and 'iNextPD' estimates of phylogenetic diversity by land-cover type, using sample-based rarefaction and extrapolation. Native forests (NF) have the highest species richness and diversities, followed by croplands (CL) and mixed plantations (MP), followed by the three monoculture plantations (BB, EC, and JC). Codes for land-cover types as in Figure 1. Symbols on each curve indicate the number of sampled locations per land-cover type, solid lines represent interpolations, and dashed lines represent extrapolations, with 95% confidence intervals. Statistically significant pairwise differences are detected visually by non-overlapping confidence intervals and are considered conservative (MacGregor-Fors & Payton, 2013). Full iNEXT and iNextPD figures are in S3 and S4.

Figure 3

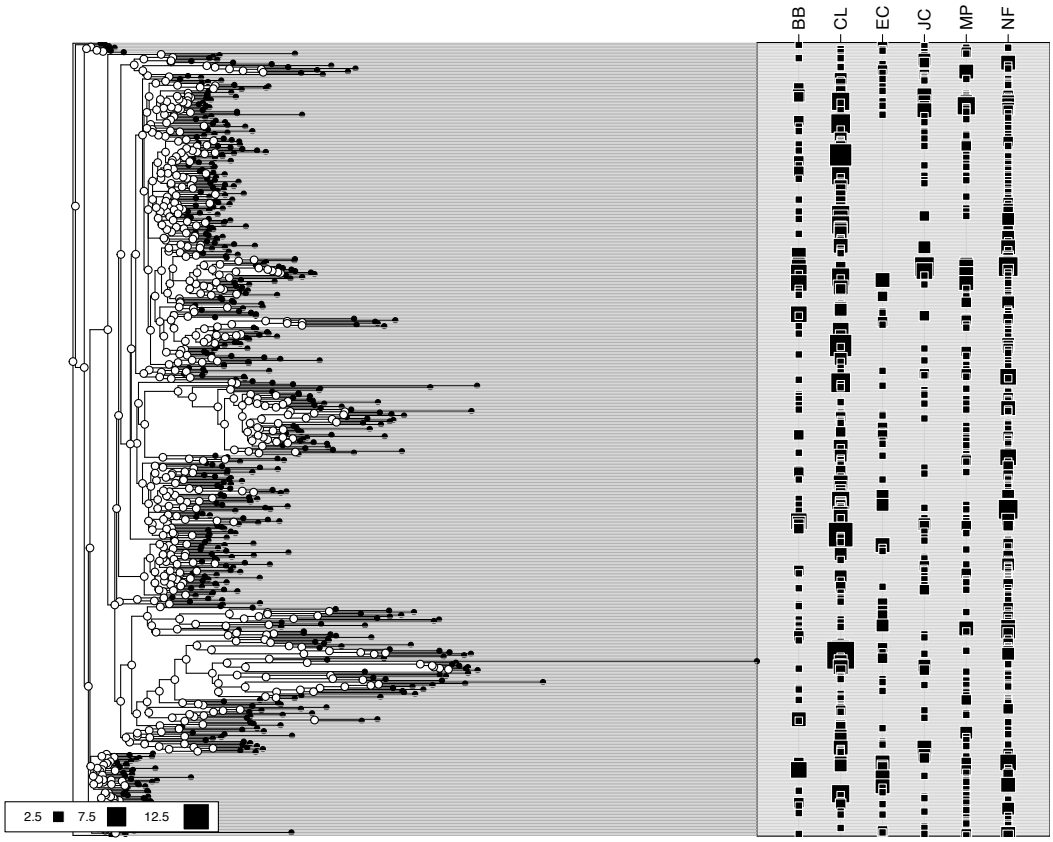


Figure 3. Phylogenetic distribution of OTUs by land-cover type, created using 'iNextPD'. Terminal nodes are black and represent the OTUs. Internal nodes are white. Sizes of the squares on the right indicate each OTU's incidence frequency (number of samples in which the OTU is observed). Phylogenetic coverage is most complete in native forests (NF) and croplands (CL), followed by mixed plantations (MP), followed by the three monocultures (BB, EC, JC). Codes for land-cover types as in Figure 1.

Figure 4

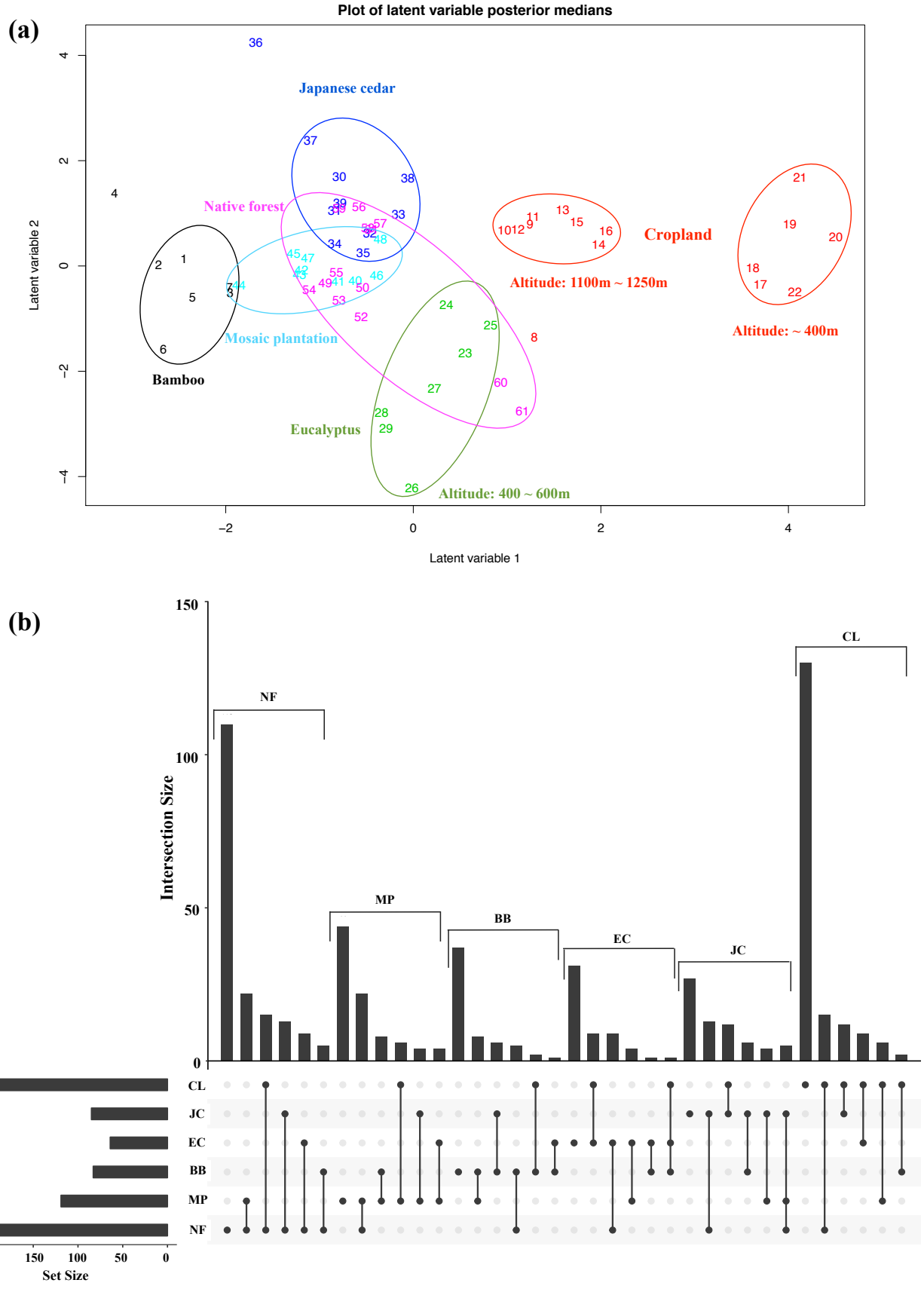

(a)
Plot of latent variable posterior medians
Latent variable 1
Latent variable 2
Japanese cedar
Native forest
Mosaic plantation
Bamboo
Eucalyptus
Cropland
Altitude: 1100m ~ 1250m
Altitude: ~ 400m
Altitude: 400 ~ 600m
(b)
Intersection Size
Set Size
NF
MP
BB
EC
JC
CL

Figure 4. Community composition differences in all land-cover types. (a) ‘Boral’ ordination. Colors represent land-cover types, and numbers represent individual samples. Cropland (CL) sites separate into two clusters by elevation. Overlap of native forests (NF) and mixed-plantations (MP) points indicates greater compositional similarity between these two land-cover types. Ovals manually added to visualize community groupings. Residuals of the ‘boral’ fit in Fig. S5. (b) UpSetR intersection map of OTUs unique to and shared between and among land-cover types. Croplands and native forests support the highest numbers of unique OTUs (CL=130, NF=110), followed by the four plantations (MP=44, BB=37, EC=31, JC=27). Native forests uniquely share almost as many OTUs with mixed plantations (22 OTUs) as native forests share with the three monocultures combined (27 OTUs, =13+9+5). Horizontal bars on the left indicate the total number of OTUs in each land-cover class. Codes for land-cover types as in Figure 1. For clarity, only pairwise comparisons are shown. A non-truncated version is presented in Fig. S6.

Figure 5

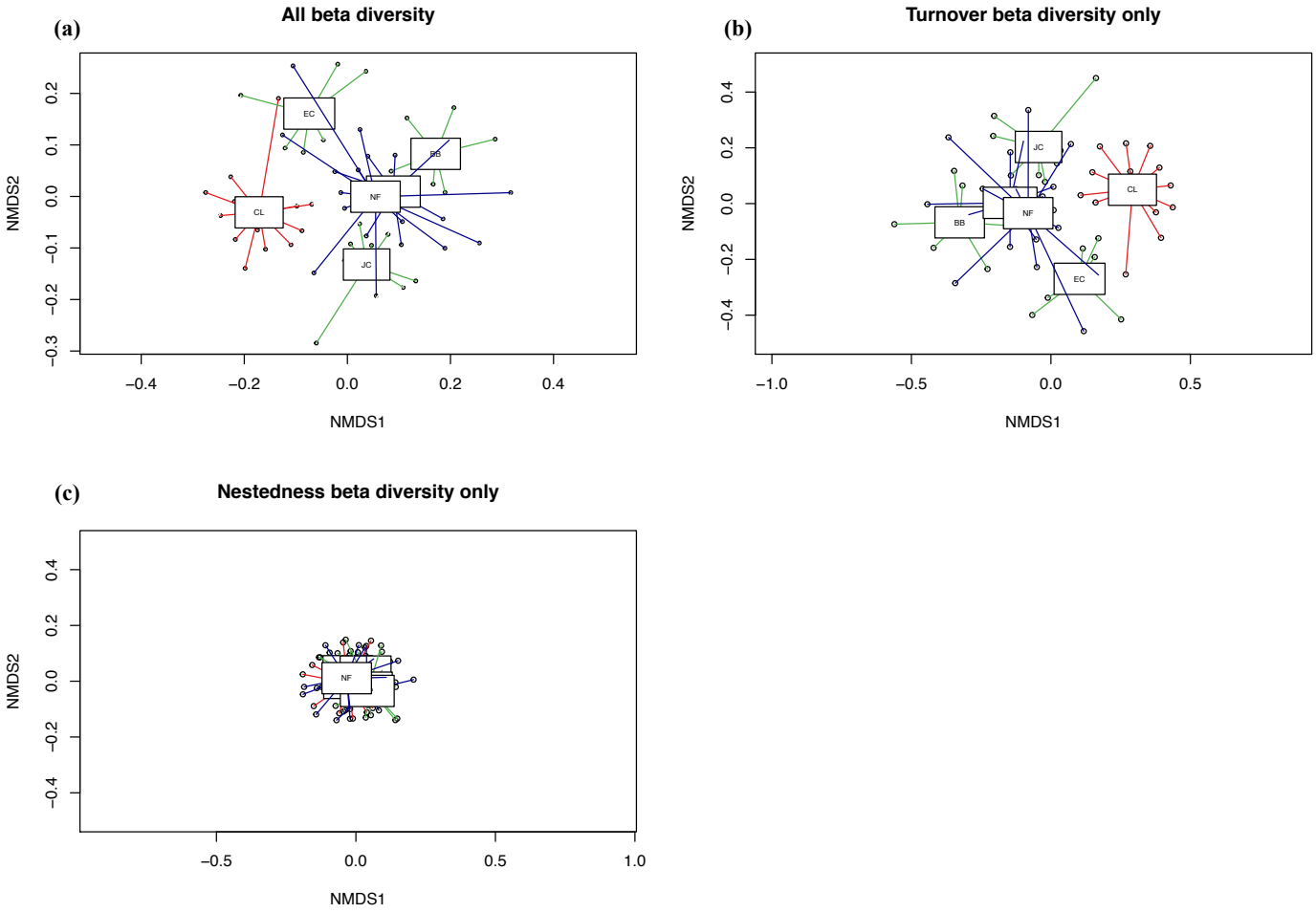


Figure 5. NMDS (non-metric multidimensional scaling) ordination of beta diversity by land-cover type (binary Jaccard dissimilarities), partitioned with 'betapart'. (a) Total beta diversity. (b) Beta diversity based on species turnover only. (c) Beta diversity based on species nestedness only. Turnover accounts for most the observed beta diversity across land-cover types, which is visualized as greater distances between points in the turnover figure (b) and almost no distances between points in the nestedness figure (c). Codes for land-cover types as in Figure 1.

Figure 6

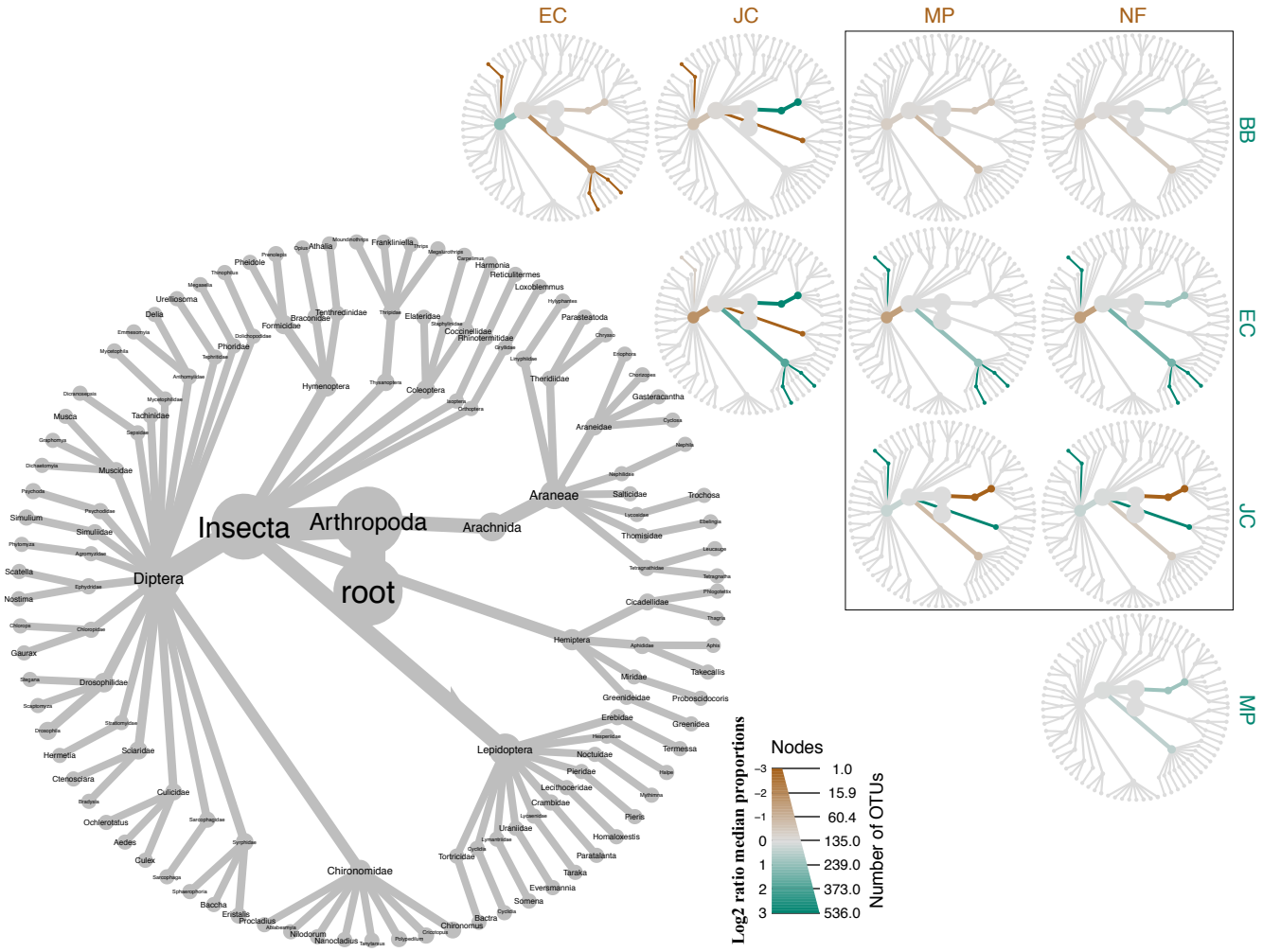


Figure 6. Pairwise taxonomic comparisons of all land-cover types. Upper right triangle: greener branches indicate taxa that are relatively more abundant (in numbers of OTUs) in the land-cover types along the right column, and browner branches indicate taxa that are relatively more abundant in the land-cover types along the top row. Lower left: taxonomic identities of the branches. Note that this is a taxonomic tree, not a phylogenetic tree. Legend: width indicates number of OTUs at a given taxonomic rank, and color indicates relative differences in $\log_2$(number of OTUs). Codes for land-cover types as in Figure 1. A figure including croplands and a zoomable taxonomic tree is in supplementary information (Figure S8, S9).

Title: The biodiversity benefit of native forests and mixed-species plantations over monoculture plantations

Appendix

Authors: Xiaoyang Wang[1,2], Fangyuan Hua[3,4], Lin Wang[1], David S. Wilcove[5,6], Douglas W Yu[1,7,8*]

Author affiliations:

[1] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[2] Kunming College of Life Sciences, University of Chinese Academy of Sciences, Kunming, Yunnan 650223, China

[3] Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3QZ, U.K.

[4] Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, China

[5] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, United States

[6] Program in Science, Technology, and Environmental Policy, Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544, United States

[7] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming Yunnan, 650223 China

[8] School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK

Appendix S1: Figures S1 – S9, Table S1 – S3

Figure S1. Spatial arrangement of pan traps in each one-hectare quadrat (= 1 sampling site). Each quadrat was subdivided into four subquadrats to balance pan colors. Each dot's color represents that pan-trap's color (white, yellow, blue, red, purple), which within each subquadrat were arranged randomly.

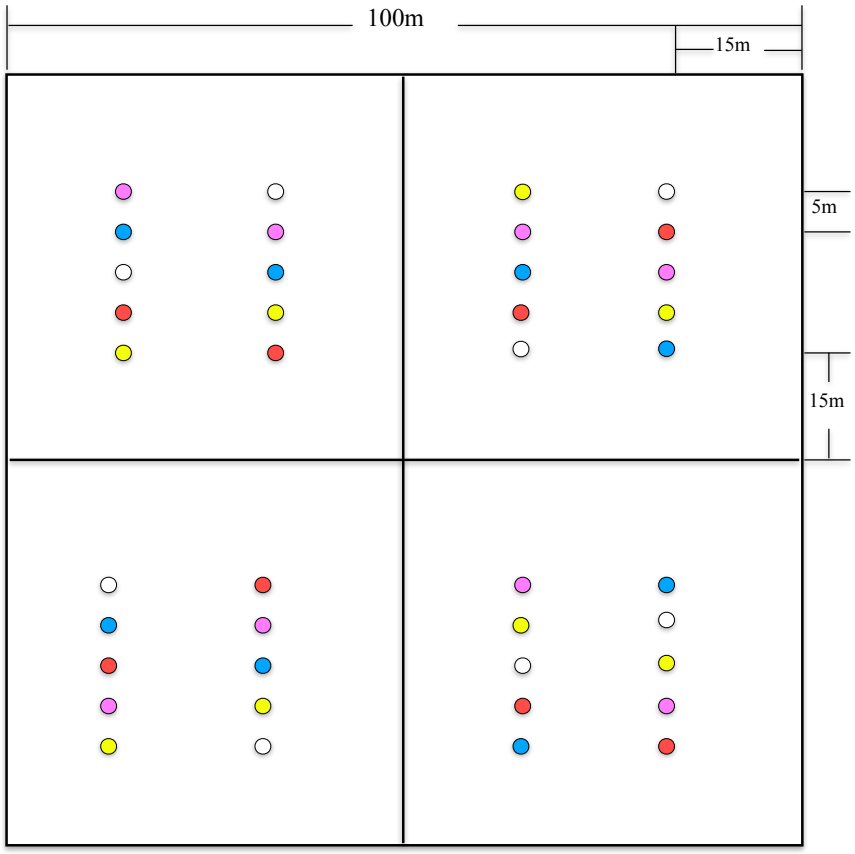


Figure S2. Each land-cover type's observed species richness, visualized using 'beanplot' 1.2 (Kampstra, 2008). White lines are observed values at each sampling site, black lines are the mean per land-cover type, and the dashed line is the grand mean. Codes for land-cover types as in Figure 1.

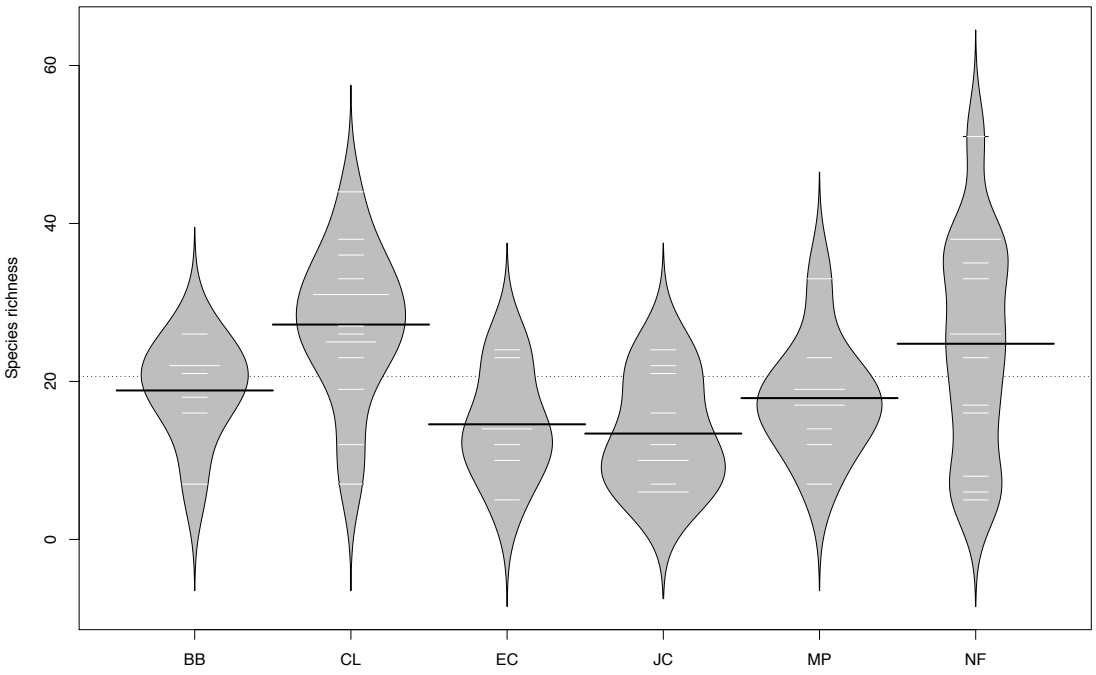


Figure S3. ‘iNEXT’ estimates of species richness, Shannon diversity, and Simpson diversity by land-cover type, using sample-based rarefaction and extrapolation. Native forests (NF) have the highest species richness and diversities, followed by croplands (CL) and mixed plantations (MP), followed by the three monoculture plantations (BB, EC, and JC). Codes for land-cover types as in Figure 1. Symbols on each curve indicate the number of sampled locations per land-cover type, solid lines represent ‘iNEXT’ interpolations, and dashed lines represent ‘iNEXT’ extrapolations, with 95% confidence intervals. Statistically significant pairwise differences are detected visually by non-overlapping confidence intervals and are somewhat conservative (MacGregor-Fors & Payton, 2013).

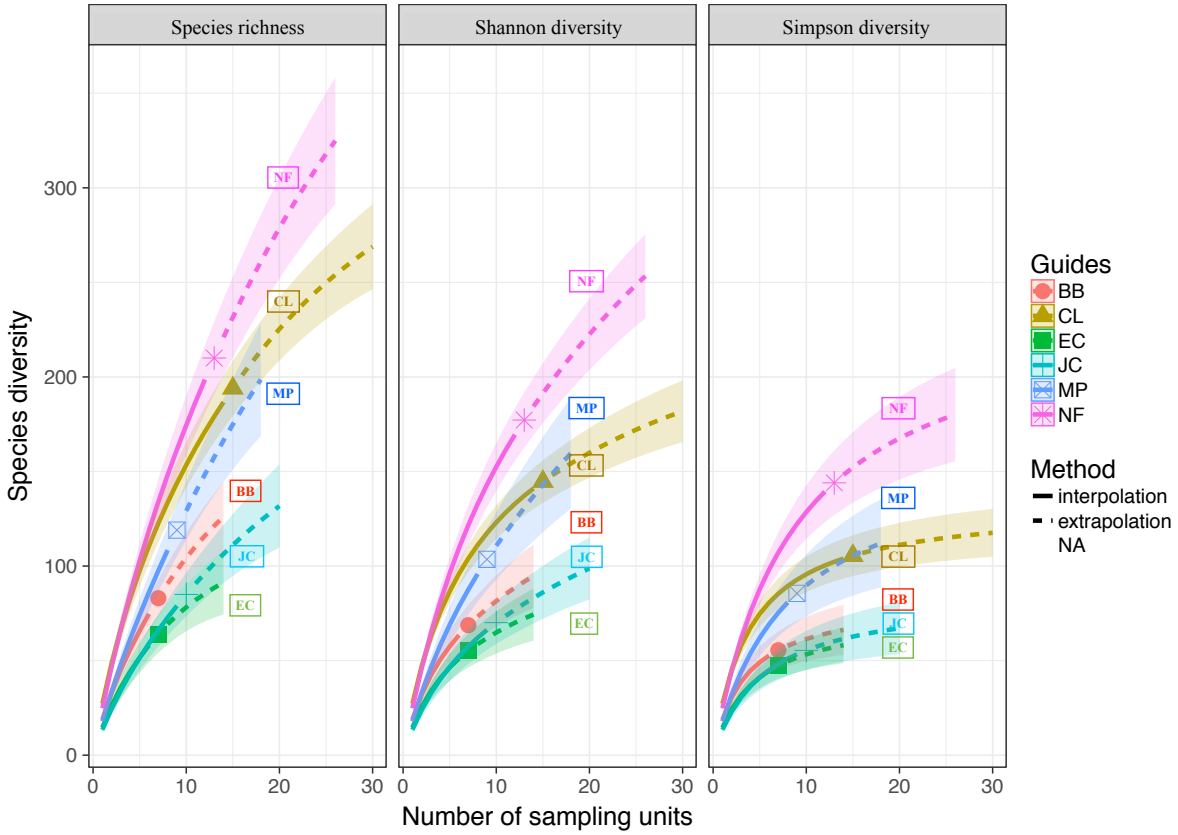


Figure S4. ‘iNextPD’ estimates of phylogenetic diversity by land-cover type, using sample-based rarefaction and extrapolation. Similar to the results in Figure S3, two of the three estimators of phylogenetic diversity are higher in native forests (NF), followed by croplands (CL) and mixed plantations (MP), followed by the three monocultures (BB, EC, and JC). Codes for land-cover types as in Figure 1). Symbols indicate sample sizes per land-cover type, solid lines represent ‘iNextPD’ interpolations, and dashed lines represent ‘iNextPD’ extrapolations, with 95% confidence intervals. Statistically significant pairwise differences are detected visually by non-overlapping confidence intervals and are somewhat conservative (MacGregor-Fors & Payton, 2013).

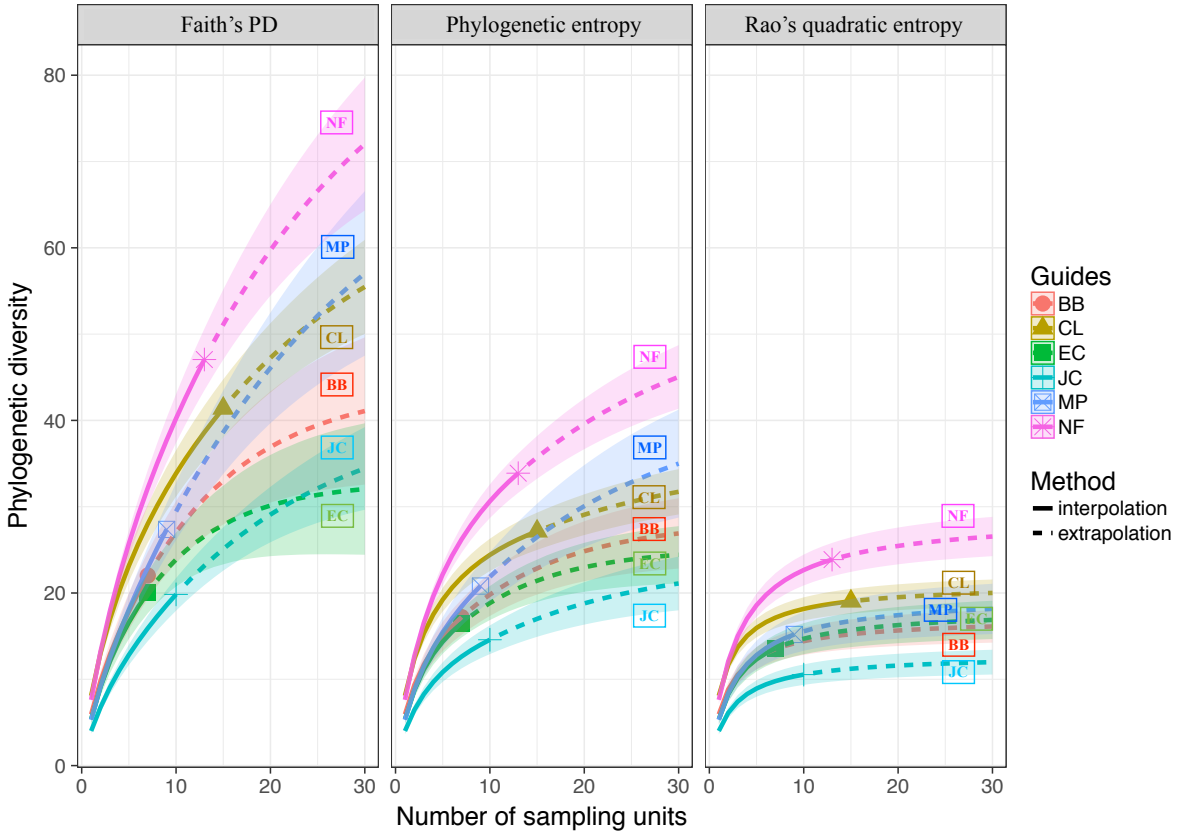


Figure S5. Residual plots of the *boral* model fit in Fig. 4a.

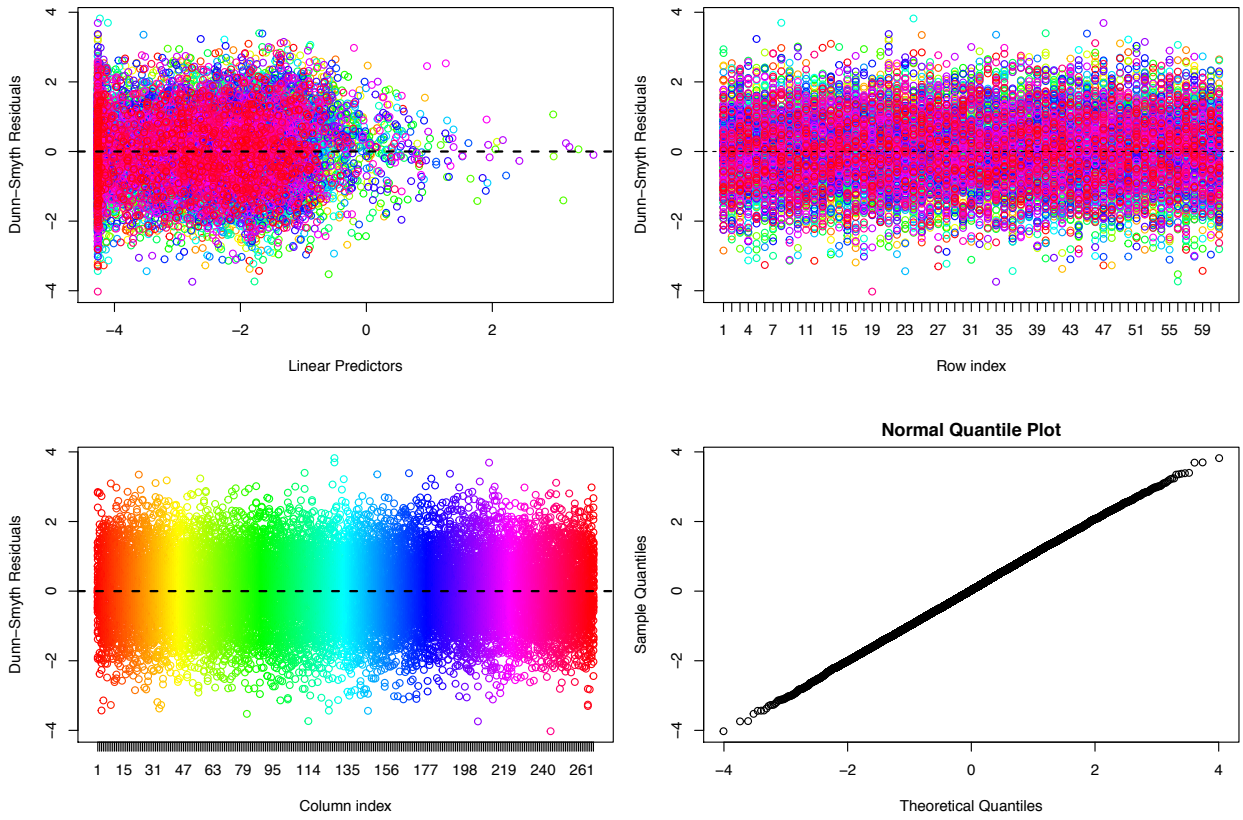


Figure S6. UpSetR intersection map of OTU distribution by land-cover type. Number of comparisons not truncated. Horizontal bars on the left bottom indicate the number of OTUs in each land-cover type, and vertical bars indicate the number of unique or shared OTUs. Codes for land-cover types as in Figure 1.

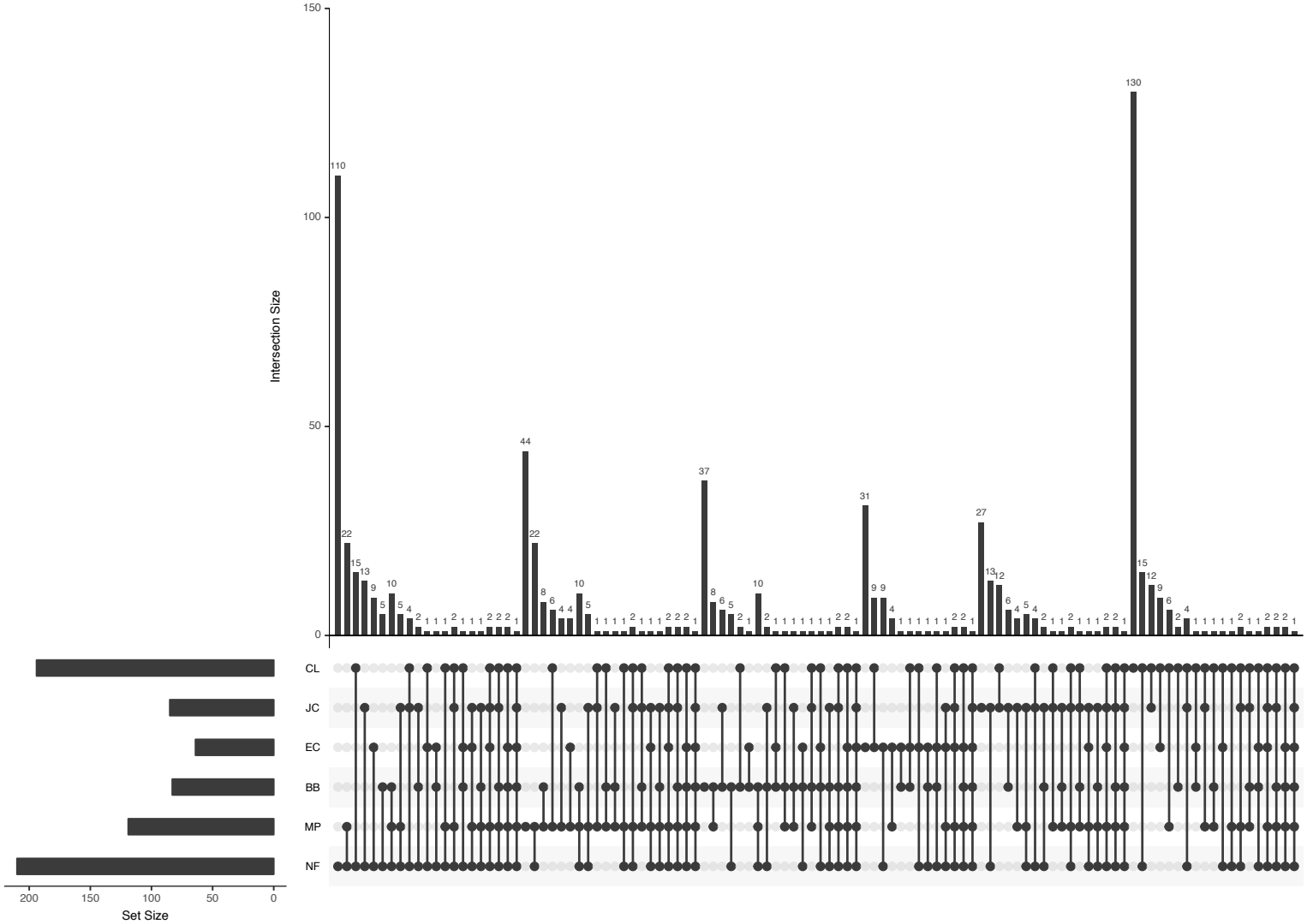


Figure S7. Heat map of OTU distribution by land-cover type, showing that beta diversity is dominated by species turnover rather than by nestedness. The vertical line separates two compartments of communities, one dominated by croplands and one dominated by forests and plantations. Each column is a sample site, and rows are OTUs. Codes for land-cover types as in Figure 1.

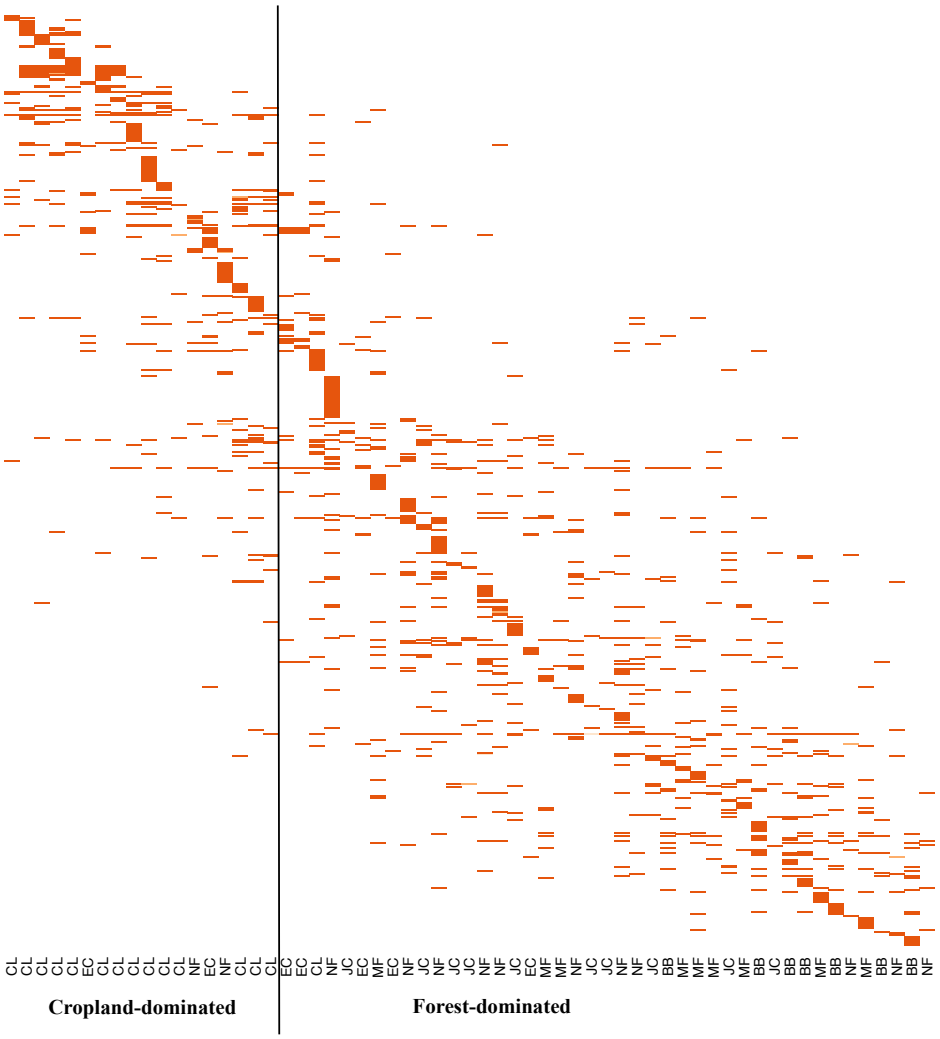


Figure S8. Pairwise taxonomic comparisons of all six land-cover types. Interpretation the same as in Figure 6 except that croplands is included in this version of the figure (boxes). Upper right triangle: greener branches indicate taxa that are relatively more abundant (in terms of numbers of OTUs) in the land-cover types along the right column, and browner branches indicate taxa that are relatively more abundant in the land-cover types along the top row. Lower left: taxonomic identities of the branches. Note that this is a taxonomic tree, not a phylogenetic tree. Legend: width indicates number of OTUs at a given taxonomic rank, and color indicates relative differences in $\log_2$(number of OTUs). Codes for land-cover types as in Figure 1.

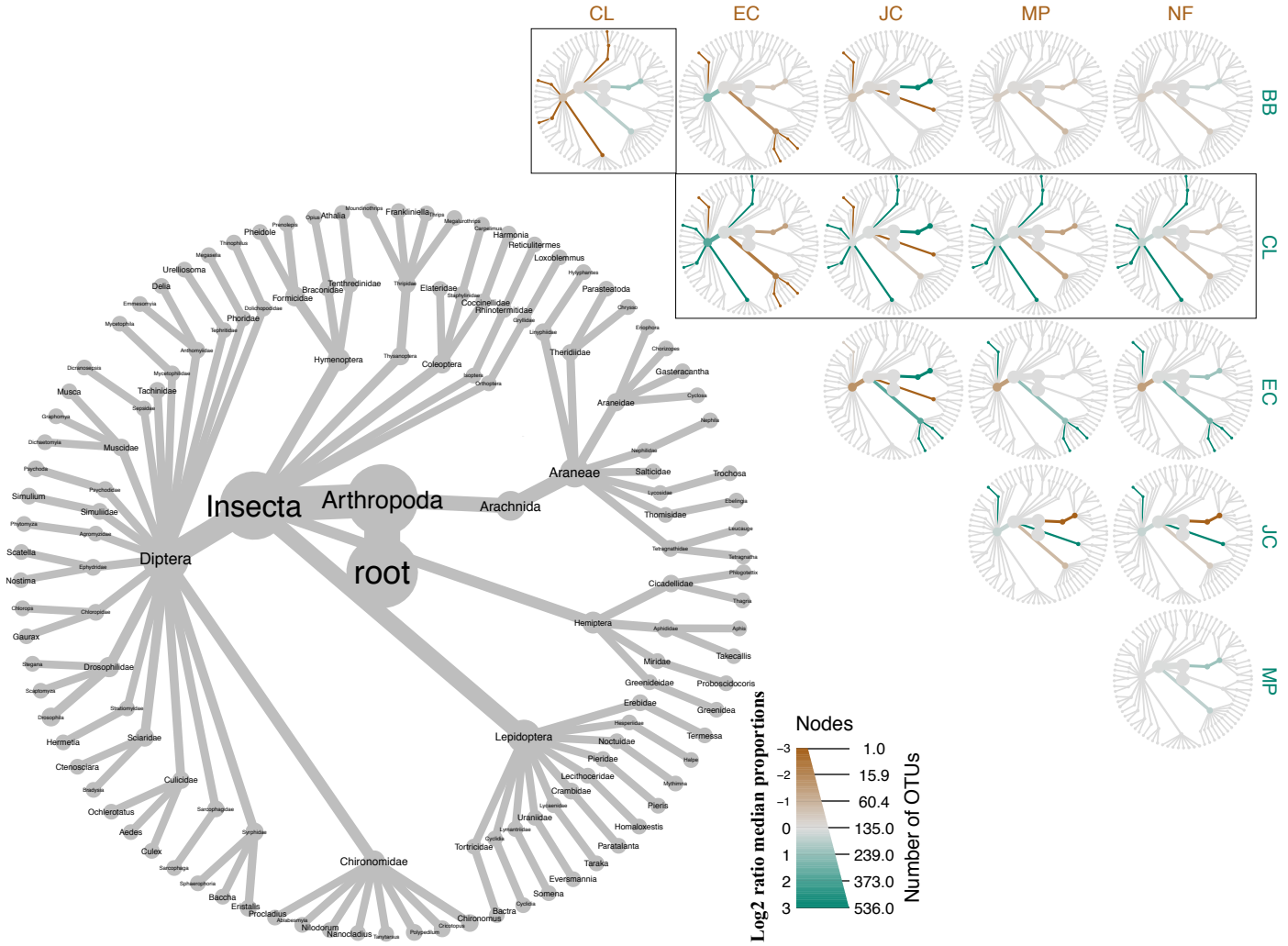


Figure S9. Taxonomic tree of all OTUs in figure 6.

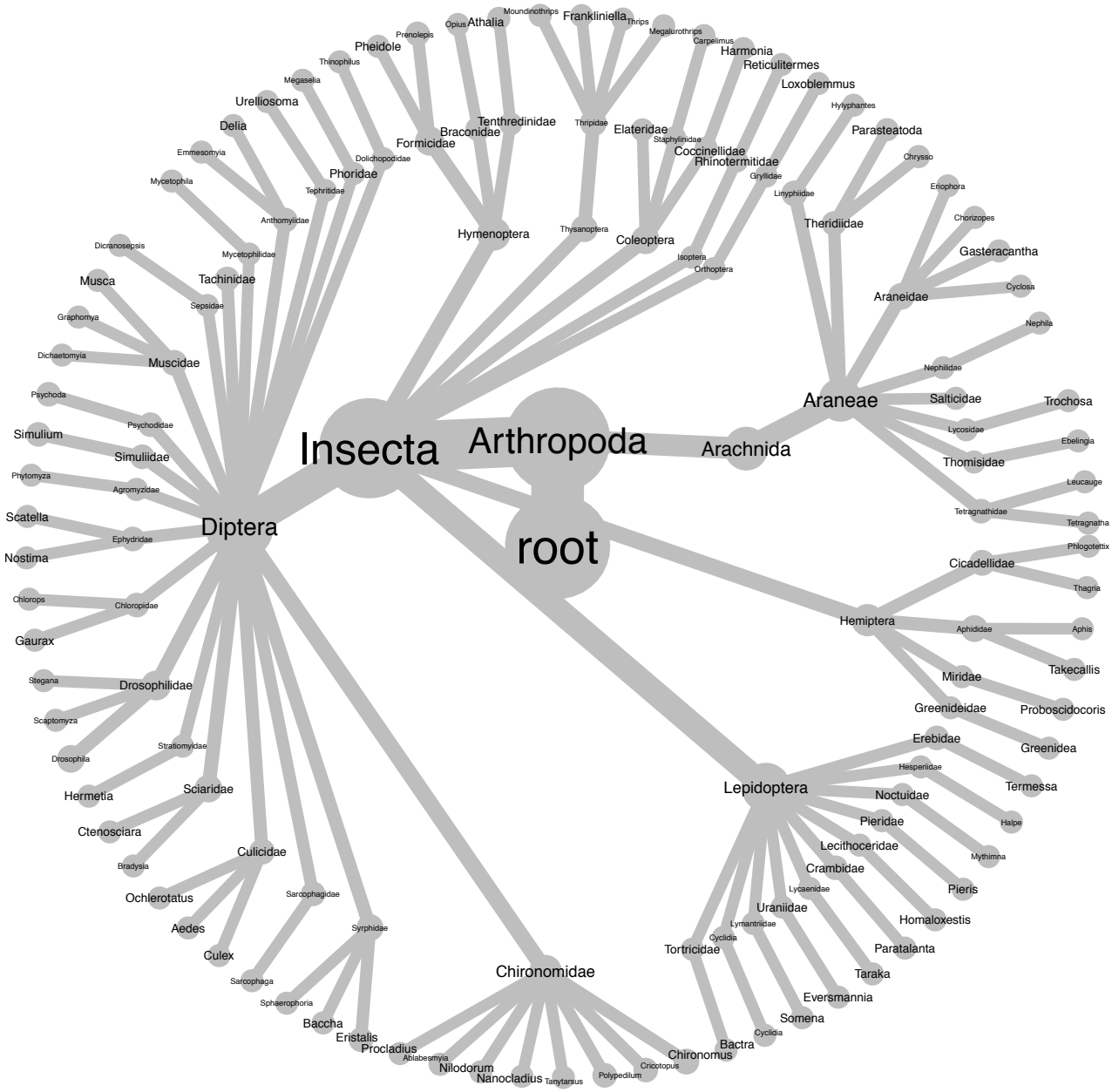


Table S1. Tags and primers used, and a table of tag combinations used for each sample (underlined), spread over two Illumina libraries. Both forward and reverse primers were tagged with sample-identifying tags.

Lib1	Lib2	primer
CL01	JC01	F1-R1
CL02	JC02	F1-R2
CL03	JC03	F1-R3
CL04	JC04	F1-R4
CL05	JC05	F1-R5
CL06	JC06	F1-R6
CL07	JC07	F2-R1
CL08	JC08	F2-R2
CL09	JC09	F2-R3
CL10	JC10	F2-R4
CL11	JC11	F2-R5
CL12	JC12	F2-R6
CL13	EC01-2-3	F3-R1
CL14	EC04	F3-R2
CL15	EC05	F3-R3
CL16	EC06	F3-R4
BB01	EC07	F3-R5
BB02	EC08	F3-R6
BB03	EC09	F4-R1
BB04	EC10	F4-R2
BB05	NF01	F4-R3
BB06	NF02-3	F4-R4
BB07	NF04	F4-R5
BB08	NF05	F4-R6
BB09	NF06	F5-R1
BB10	NF07	F5-R2
MF01	NF08	F5-R3
MF02	NF09	F5-R4
MF03	NF10	F5-R5
MF04	NF11	F5-R6
MF05	NF12	F6-R1
MF06	NF13	F6-R2
MF07	NF14	F6-R3
MF08	NF15	F6-R4
MF09	NF16	F6-R5
MF10	-	F6-R6

Tagged_primer	Forward
Tag1	CCTAAACTACGGGGTCAACAAATCATAAAGATATTGG
Tag2	GTGGTATGGGAGTGGTCAACAAATCATAAAGATATTGG
Tag3	TGTTGCGTTTCTGTGGTCAACAAATCATAAAGATATTGG
Tag4	ACAGCCACCCATCGAGGTCAACAAATCATAAAGATATTGG
Tag5	GTTACGTGGTTGATGAGGTCAACAAATCATAAAGATATTGG
Tag6	TACCGGCTTGCATGCGAGGTCAACAAATCATAAAGATATTGG

Tagged_primer	Reverse
Tag1	CCTAAACTACGGGGNGGRTANANNGTYCANCCNGYNCC
Tag2	GTGGTATGGGAGTGGNGGRTANANNGTYCANCCNGYNCC
Tag3	TGTTGCGTTTCTGTGGNGGRTANANNGTYCANCCNGYNCC
Tag4	ACAGCCACCCATCGAGGNGGRTANANNGTYCANCCNGYNCC
Tag5	GTTACGTGGTTGATGAGGNGGRTANANNGTYCANCCNGYNCC
Tag6	TACCGGCTTGCATGCGAGGNGGRTANANNGTYCANCCNGYNCC

Adapter_link_tag	Forward
Tag1	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCTAAACTACGG
Tag2	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTGGTATGGGAG
Tag3	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGTTGCGTTTCT
Tag4	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACAGCCACCCAT
Tag5	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTTACGTGGTTGATGA
Tag6	CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTACCGGCTTGCATGCGA

Adapter_link_tag	Reverse
Tag1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTAAACTACGG
Tag2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGGTATGGGAG
Tag3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTGTTGCGTTTCT
Tag4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACAGCCACCCAT
Tag5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTTACGTGGTTGATGA
Tag6	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTACCGGCTTGCATGCGA

Table S2. Multiple pairwise Welch's t tests for Chao2 estimates. P values adjusted by Bonferroni. Codes for land-cover types as in Figure 1.

	CL	EC	JC	MP	NF
BB	0.0432	0.1913	0.7236	0.1022	0.006*
CL		0.00075*	0.0973	0.5307	0.0973
EC			0.1420	0.0455*	0.00075*
JC				0.1400	0.0105*
MP					0.5242

Table S3. *mvabund* compositional comparisons. We used *mvabund* to test whether arthropod species compositions in native forests and mixed plantations are significantly different from each other and from the other land-cover types in the study region. After Bonferroni correction, all comparisons were significantly different at $p < 0.01$. Codes for land-cover types as in Figure 1.

	MP	BB	CL	EC	JC
NF	0.00125	0.00125	0.00125	0.00125	0.003
MP		0.001	0.001	0.001	0.001

REFERENCE

Kampstra, P. (2008). Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software*, *28*, 1–9. https://doi.org/10.18637/jss.v028.c01

MacGregor-Fors, I., & Payton, M. E. (2013). Contrasting Diversity Values: Statistical Inferences Based on Overlapping Confidence Intervals. *PLoS ONE*, *8*, 8–11. https://doi.org/10.1371/journal.pone.0056794