LETTER

doi:10.1038/nature22399

Leaf bacterial diversity mediates plant diversity and ecosystem function relationships

Isabelle Laforest-Lapointe[1,2], Alain Paquette[1,2], Christian Messier[1,2,3] & Steven W. Kembel[1,2]

Research on biodiversity and ecosystem functioning has demonstrated links between plant diversity and ecosystem functions such as productivity[1,2]. At other trophic levels, the plant microbiome has been shown to influence host plant fitness and function[3,4], and host-associated microbes have been proposed to influence ecosystem function through their role in defining the extended phenotype of host organisms[5,6] However, the importance of the plant microbiome for ecosystem function has not been quantified in the context of the known importance of plant diversity and traits. Here, using a tree biodiversity–ecosystem functioning experiment, we provide strong support for the hypothesis that leaf bacterial diversity is positively linked to ecosystem productivity, even after accounting for the role of plant diversity. Our results also show that host species identity, functional identity and functional diversity are the main determinants of leaf bacterial community structure and diversity. Our study provides evidence of a positive correlation between plant-associated microbial diversity and terrestrial ecosystem productivity, and a new mechanism by which models of biodiversity–ecosystem functioning relationships can be improved.

The identification of the mechanisms that promote and maintain primary production in terrestrial ecosystems is a central question in ecology, especially in the context of anthropogenic global change[7], and increasing biodiversity loss[8]. After years of research on biodiversity–ecosystem functioning, the importance of diversity in driving ecosystem productivity and services has been demonstrated in many ecosystems[1,2]. Two effects have been proposed to explain the positive effect of diversity on productivity[9]: a complementarity effect, in which more diverse communities make better use of resources through niche partitioning; and a selection effect, in which dominance by species possessing particular traits drives ecosystem properties. Biodiversity–ecosystem functioning studies have shown that plant species richness, functional diversity and functional identity[10] are among the key factors driving terrestrial ecosystem productivity, but recent work suggests that these relationships could differ among trophic levels[11].

The development of high-throughput sequencing technologies has revolutionized our understanding of microbial ecology, and led to calls for the consideration of host-associated microbial communities as part of the host's extended phenotype or 'holobiont'[3] with potential effects on host ecology and evolution. Plant-associated microbial communities have direct roles in ecosystem functioning through effects on carbon[12,13] and nitrogen cycles[12,14]. They also influence ecosystem function indirectly through their effects on host plant health and productivity via numerous mechanisms[4,5] such as modifying plant hormone production[15] and increasing host resistance to abiotic and biotic stress[16]. Healthy hosts have been shown to contain a greater diversity of microorganisms than hosts infected by pathogens in systems including the human gut[17] and plant root[18] and leaf[19] microbiomes. There is accumulating evidence that higher leaf bacterial diversity influences host productivity through a variety of mechanisms, including (1) inducing plant-mediated resistance by improving host resistance to pathogens through increasing competition for niches, depleting nutrient pools and enhancing the production of antibiotic molecules[20]; (2) influencing plant hormone production (that is, auxins[21] and cytokinins[22]); and (3) augmenting nitrogen availability through atmospheric nitrogen

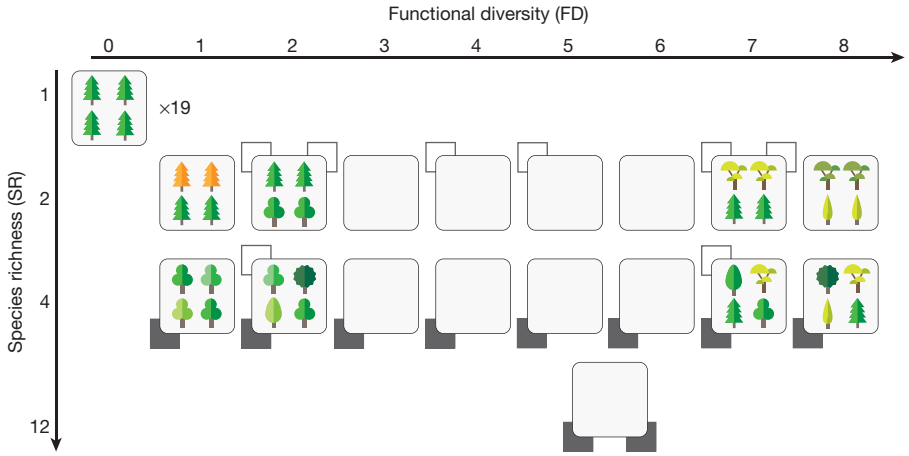


Figure 1 | The IDENT experiment near Montreal, Canada. A total of 54 community mixtures involving 19 tree species replicated four times were established in spring 2009, including gradients of species richness (1, 2, 4 and 12) and functional diversity (8 initial levels). The functional diversity of all possible mixtures was ordered into 8 bins from which communities were chosen. Smaller white squares placed as exponents denote additional plots at some functional diversity levels (different communities producing similar functional diversity values). Exotic species were included as monocultures and in mixtures of 4 and 12 with native species in equal proportion, denoted as subscript black squares. IDENT, International Diversity Experiment Network with Trees.

[1]Département des Sciences Biologiques, Université du Québec à Montréal, Montréal H3C 3P8, Québec, Canada. [2]Centre d'étude de la Forêt, Université du Québec à Montréal, Montréal H2X 3Y7, Québec, Canada. [3]Institut des Sciences de la Forêt Tempérée, Université du Québec en Outaouais, Ripon J0V 1V0, Québec, Canada.

Table 1 | Bacterial community structure explained by various factors

Variables	*F* value	df	R^2	*P*(>*F*)
Host species identity	12.58	18	26.6	0.001
Functional identity	13.95	1	1.6	0.001
Functional diversity	11.15	1	1.3	0.001
Species richness	1.68	1	0.2	0.055
Host species identity × functional identity	1.67	18	3.5	0.001
Host species identity × functional diversity	1.18	18	2.5	0.021

The model explains a total of 36% of the variation in bacterial community structure (Bray–Curtis dissimilarity) in 606 samples of leaf bacterial communities from trees. The full model tested the influence of host species identity, plant functional identity, species richness and functional diversity, as well as the interaction of each of these last three variables with host species identity, on leaf bacterial community structure. The interaction between host species identity and species richness was not significant ($P=0.23$) and was therefore removed from the model. *P* values determined by PERMANOVA. df, degrees of freedom.

Table 2 | Drivers of tree-level leaf bacterial diversity

Variables	*F* value	df_n	df_d	*P*(>*F*)
Host species identity	37.99	18	535	<0.0001
Functional identity	26.16	1	474	<0.0001
Functional diversity	21.90	1	302	<0.0001

Determinants of leaf bacterial diversity were quantified using the Shannon diversity index. Parameters determined by type II ANOVA with Kenward–Rodger approximation of the degrees of freedom in a linear-mixed model. The model explains 53% of the marginal variation (only due to fixed effects) in leaf alpha diversity in 606 samples of leaf bacterial communities from trees. Random effects were the factors plot (54 levels, nested in block) and block (4 levels). Species richness was not significant ($P=0.30$) and was thus removed from the model. df_d, degrees of freedom denominator; df_n, degrees of freedom numerator.

fixation by leaf bacterial communities[14]. Despite their potential importance in mediating plant biodiversity–ecosystem function relationships, the role of microbial communities in driving ecosystem productivity and function has never been evaluated in an experimental context that allows direct quantification of the association between plant-associated microbes and ecosystem function.

In this study, we quantified the relationships among leaf bacterial diversity, plant species richness, plant functional diversity and identity, and plant community productivity in a biodiversity–ecosystem function experiment with trees. We first compared the relative influence of host species identity and diversity on host-level leaf bacterial community structure and diversity. We then evaluated the hypothesis that effects mediated by leaf bacterial diversity explain an important part of the influence of plant diversity and identity on productivity. We hypothesized that (1) host species identity and functional diversity will be the strongest driver of leaf bacterial community structure and diversity on individual trees; and (2) a higher leaf bacterial diversity will be positively linked to plant community productivity. We tested these hypotheses by measuring leaf bacterial community structure on 620 trees from 19 species in a common field garden biodiversity experiment near Montreal, Canada, where tree species richness and functional diversity were manipulated in a replicated design with 1–12 tree species grown together for 5 years in 4 × 4 m experimental plots (Fig. 1).

The strongest driver of leaf bacterial community structure at the tree level was host species identity (permutational multivariate analysis of variance (PERMANOVA); $F=12.68$, $R^2=26.6\%$, $P=0.001$; Table 1), in accordance with previous studies[23]. Although their relative influence was much smaller, plant functional identity ($R^2=1.7\%$) and diversity ($R^2=1.3\%$) were also significant drivers of leaf bacterial community structure and interacted with host species identity to shape bacterial communities on leaves. Likewise, host species identity ($F_{18,535}=38.0$, $P<0.0001$) was the strongest determinant of leaf bacterial diversity (linear mixed model on leaf bacterial diversity; marginal $R^2=53\%$; Table 2), followed by functional identity ($F_{1,474}=26.2$, $P<0.0001$) and functional diversity ($F_{1,302}=21.9$, $P<0.0001$). These results suggest that host species identity has a dominant role in determining leaf microbial community structure even after accounting for changes in plot-level plant functional diversity, identity and species richness. In addition, our results support the idea that plant-associated microbial communities vary predictably with host plant ecological strategy[24], and thus potentially affect host growth and ecosystem productivity.

The diversity of bacterial communities on tree leaves explained significant amounts of variation in plant community productivity ($r=0.12$; $P=0.002$; Fig. 2), even when accounting for the effects of all other variables (structural equation model; $\chi^2=1.451$, $P=0.484$; Fig. 2). Removing the link between leaf bacterial diversity and community productivity in the structural equation model yielded an unstable model ($\chi^2=11.906$, $P=0.008$), providing further evidence for the importance of leaf bacterial diversity for plant community productivity. Furthermore, we tested for a possible influence of leaf bacterial composition on plant community productivity and found no significant effect (Extended Data Figs 5, 6). At the plot level, plant species richness ($r=0.61$; $P<0.001$), functional identity ($r=-0.28$; $P<0.001$) and functional diversity ($r=0.26$; $P<0.001$) had a strong effect on productivity in the model, with plant species richness being the strongest determinant of plant community productivity ($R^2=85\%$; Fig. 2). In addition, plant species richness ($r=0.29$; $P=0.003$), functional identity ($r=0.44$; $P<0.001$) and functional diversity ($r=0.18$; $P=0.08$) also drove leaf bacterial diversity, explaining 41% of the variance in bacterial diversity between plots. These results provide empirical evidence that leaf bacterial diversity is positively related to terrestrial ecosystem productivity, even after accounting for other explanatory factors, and that biodiversity–ecosystem functioning relationships in plant communities could in part be driven by positive interactions involving other trophic levels. Here we show that plant-associated microbial diversity is positively related to plant community productivity, explaining a portion of the variation in productivity that would otherwise have been attributed to plant diversity and functional traits, both adding to the explanatory power of the model of plot productivity and mediating the tree diversity-identity effect on productivity.

Our finding that leaf bacterial diversity but not community composition is related to plant community productivity suggests that the effect of bacterial diversity on plant productivity is more likely to be a complementarity effect rather than a selection effect. Many studies have proposed that niche complementarity is one of the principal mechanisms that explain positive biodiversity–ecosystem function relationships[25], through more efficient capture of resources with increasing species diversity and complementarity[26]. Recent food-web studies have

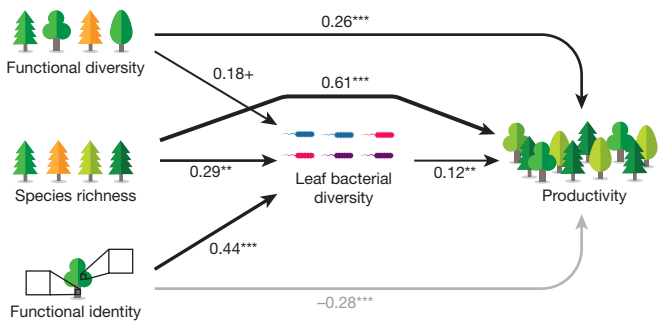


Figure 2 | Structural equation model of plant diversity and identity explaining leaf bacterial diversity and plant community productivity. The path analysis ($n=216$, $\chi^2=1.451$, $P=0.484$, degrees of freedom (df) = 2; root mean square error of approximation (RMSEA) = 0.00, $P=0.644$) explains 41% of the variance in leaf bacterial diversity and 85% of the variance in plot productivity. Numbers adjacent to arrows and arrow width indicate the relationship's effect size and the associated bootstrap *P* value. $+P<0.1$; $*P<0.05$; $**P<0.01$; $***P<0.001$. Black and grey arrows indicate positive and negative relationships, respectively.

introduced the idea of trophic complementarity, a concept based on complementarity occurring through differential resource use, predation by distinct predators, or both[27]. Here we provide unprecedented evidence that leaf bacterial diversity could have a role in stimulating plant community productivity. Our work concurs with previous studies that demonstrated the influence of leaf bacterial diversity on plant community productivity through mechanisms such as fixation of atmospheric nitrogen[14] or protection from pathogen infection[20,28]. The demonstration of causality between diversity and productivity is a common concern raised in biodiversity–ecosystem functioning studies, and since we did not manipulate leaf bacterial diversity experimentally, it is not possible to state definitively that it caused the observed increase in plant community productivity. However, our findings do support the hypothesis that adding a multi-trophic component to studies of biodiversity–ecosystem functioning in plant communities is a promising route to understand diversity mechanisms better by improving models of plant ecosystem productivity. Our work indicates the need for future research oriented towards system-level and multi-trophic experiments to confirm this hypothesis and to identify the underlying mechanisms driving the influence of the leaf microbiota on plant community productivity. Given the capacity of microbes to respond rapidly to environmental changes[29], studying how the effect of microbial communities on plant productivity interacts with global change and intensified anthropogenic pressures will be crucial to optimize or maintain primary production. Using one of the most extensive studies of tree leaf bacterial communities so far, our results suggest that considering plant-associated microbial diversity can improve models of biodiversity–ecosystem functioning.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 July 2016; accepted 21 April 2017.
Published online 24 May 2017.

1. Tilman, D., Reich, P. B. & Isbell, F. Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. *Proc. Natl Acad. Sci. USA* **109,** 10394–10397 (2012).
2. Liang, J. *et al.* Positive biodiversity-productivity relationship predominant in global forests. *Science* **354,** aaf8957 (2016).
3. Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A. & Dufresne, A. The importance of the microbiome of the plant holobiont. *New Phytol.* **206,** 1196–1206 (2015).
4. Vorholt, J. A. Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* **10,** 828–840 (2012).
5. Bringel, F. & Couée, I. Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Front. Microbiol.* **6,** 486 (2015).
6. Müller, D. B., Vogel, C., Bai, Y. & Vorholt, J. A. The plant microbiota: systems biology insights and perspectives. *Annu. Rev. Genet.* **50,** 211–234 (2016).
7. Hautier, Y. *et al.* Plant ecology. Anthropogenic environmental changes affect ecosystem stability via biodiversity. *Science* **348,** 336–340 (2015).
8. Allan, E. *et al.* Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. *Ecol. Lett.* **18,** 834–843 (2015).
9. Loreau, M. & Hector, A. Partitioning selection and complementarity in biodiversity experiments. *Nature* **412,** 72–76 (2001).
10. Flynn, D. F., Mirotchnick, N., Jain, M., Palmer, M. I. & Naeem, S. Functional and phylogenetic diversity as predictors of biodiversity–ecosystem-function relationships. *Ecology* **92,** 1573–1581 (2011).
11. O'Connor, M. I. *et al.* A general biodiversity–function relationship is mediated by trophic level. *Oikos* **126,** 18–31 2016).
12. Knief, C. *et al.* Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J.* **6,** 1378–1390 (2012).
13. Jo, Y. *et al.* Bacterial communities in the phylloplane of *Prunus* species. *J. Basic Microbiol.* **55,** 504–508 (2015).
14. Moyes, A. B. *et al.* Evidence for foliar endophytic nitrogen fixation in a widely distributed subalpine conifer. *New Phytol.* **210,** 657–668 (2016).
15. Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M. & Vorholt, J. A. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet.* **10,** e1004283 (2014).
16. Zamioudis, C. & Pieterse, C. M. J. Modulation of host immunity by beneficial microbes. *Mol. Plant Microbe Interact.* **25,** 139–150 (2012).
17. Khanna, S. *et al.* A novel microbiome therapeutic increases gut microbial diversity and prevents recurrent *Clostridium difficile* infection. *J. Infect. Dis.* **214,** 173–181 (2016).
18. Berendsen, R. L., Pieterse, C. M. & Bakker, P. A. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17,** 478–486 (2012).
19. Agler, M. T. *et al.* Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biol.* **14,** e1002352 (2016).
20. Ritpitakphong, U. *et al.* The microbiome of the leaf surface of *Arabidopsis* protects against a fungal pathogen. *New Phytol.* **210,** 1033–1043 (2016).
21. Brandl, M. T., Quiñones, B. & Lindow, S. E. Heterogeneous transcription of an indoleacetic acid biosynthetic gene in *Erwinia herbicola* on plant surfaces. *Proc. Natl Acad. Sci. USA* **98,** 3454–3459 (2001).
22. Brandl, M. T. & Lindow, S. E. Contribution of indole-3-acetic acid production to the epiphytic fitness of *Erwinia herbicola*. *Appl. Environ. Microbiol.* **64,** 3256–3263 (1998).
23. Laforest-Lapointe, I., Messier, C. & Kembel, S. W. Host species identity, site and time drive temperate tree phyllosphere bacterial community structure. *Microbiome* **4,** 27 (2016).
24. Kembel, S. W. *et al.* Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proc. Natl Acad. Sci. USA* **111,** 13715–13720 (2014).
25. Sapijanskas, J., Paquette, A., Potvin, C., Kunert, N. & Loreau, M. Tropical tree diversity enhances light capture through crown plasticity and spatial and temporal niche differences. *Ecology* **95,** 2479–2492 (2014).
26. Fargione, J. *et al.* From selection to complementarity: shifts in the causes of biodiversity–productivity relationships in a long-term biodiversity experiment. *Proc. R. Soc. B* **274,** 871–876 (2007).
27. Poisot, T., Mouquet, N. & Gravel, D. Trophic complementarity drives the biodiversity-ecosystem functioning relationship in food webs. *Ecol. Lett.* **16,** 853–861 (2013).
28. Wei, F., Hu, X. & Xu, X. Dispersal of *Bacillus subtilis* and its effect on strawberry phyllosphere microbiota under open field and protection conditions. *Sci. Rep.* **6,** 22611 (2016).
29. Lau, J. A. & Lennon, J. T. Rapid responses of soil microorganisms improve plant fitness in novel environments. *Proc. Natl Acad. Sci. USA* **109,** 14058–14062 (2012).

Acknowledgements We are grateful to R. Fréchon, S. Guérard and M. A. Chadid Hernandez for their help in the field, and to J. Shapiro and his laboratory for technical support. We also thank B. Shipley for his help with structural equation modelling techniques used in the manuscript. C. M. Tobner, P. B. Reich and D. Gravel helped in designing the original experiment (IDENT-Montréal) together with A.P. and C.M. The study site is part of McGill University and we much appreciate their support.

Author Contributions I.L.-L., C.M. and S.W.K. designed the study; C.M. and A.P. established the field experiment; I.L.-L. collected the data; I.L.-L. analysed the data with support from A.P. and S.W.K.; all authors contributed to the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to I.L.L. (isabelle.laforest.lapointe@gmail.com).

Reviewer Information *Nature* thanks D. Wardle, S. Lindow, J. Grace and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

No statistical methods were used to predetermine sample size, and investigators were not blinded to allocation during experiments and outcome assessment.

Experiment description. The common garden experiment is located at Ste-Anne-de-Bellevue, near Montreal, Quebec, Canada (45° 26′ N, 73° 56′W, 39 metres above sea level (MASL)) where the mean annual temperature and mean annual precipitation are 6.2 °C and 963 mm, respectively (http://climate.weatheroffice.gc.ca). This experiment was established in 2009 as part of the International Diversity Experiment Network with Trees (IDENT) present in North America and Europe[30], and the TreeDivNet[31]. The experiment is organized in a randomized block (4 levels) design that includes densely planted (50-cm spacing) trees in 8 × 8 plots (16 m^2) in monocultures of 19 temperate and boreal tree species, 14 two-species mixtures, 18 four-species mixtures and three 12-species mixtures of a set of 12 native and 7 exotic species (Fig. 1). The study site is a flat agricultural field intensively managed for decades. The soil consists of a 20–70-cm deep sandy layer overtopping clay. Microtopography (the difference in elevation between plot centres) was measured to account for slight differences in drainage[32]. The experiment is surrounded by a buffer of random tree species from the same pool. At the end of the 2014 season, tree height ranged from 1.3 to 5.7 m with a mean of 3.2 m, while diameter at 5 cm from ground ranged between 20 and 60 mm with a mean of 38 mm. At the beginning of that season, tree mortality since establishment was below 4%. Species mixtures were established to create functional diversity gradients over each of the fixed and independent species richness levels. Additional species combinations were also established at some functional diversity levels to increase resolution (see Fig. 1, Extended Data Fig. 1, and refs 30, 32 for added details on the experimental design).

Functional diversity, functional identity and productivity. We measured functional diversity using the functional dispersion index[33] calculated as the mean distance of each species to the centre of mass of all species in a multidimensional trait space. We quantified functional identity using the first axis of a principal components analysis on community weighted mean traits[34] based on planted relative abundances, explaining 80% of variation in traits among species (Extended Data Fig. 2). We obtained data on host plant functional traits including maximum photosynthetic capacity (A_{mass}), leaf longevity (LL), leaf mass per area (LMA), leaf nitrogen content (N_{mass}), and wood density (WD) from global databases (Extended Data Table 1). To estimate total plant community productivity, we measured the diameter and height of each 13,824 trees at the end of the sixth growth year (2014) since planting and then estimated the aboveground stem volume (V_{plot}) with the following formula:

$$V_{plot} = \sum_{i}^{n} (D_i^2 \times H_i)$$

in which D_i represents tree i diameter and H_i denotes tree i height. Plot volume was calculated only for the inner 36 trees, leaving out trees from the outer rows of each plot, to minimize edge effects.

Bacterial community sampling and DNA extraction. On 2 July 2014, we collected one sample of leaves per host species per plot for a total of 620 samples. For bacterial community collection and amplification, we used previously described protocols[24]. We took one 50–100 g sample of leaves per host species per plot summing up to 620 samples. In the laboratory, all samples were uniformly trimmed to 50 g mass. We collected microbial communities from leaf surfaces by agitating the samples in 100 ml of diluted Redford buffer solution for 5 min. We re-suspended cells in 500 µl of PowerSoil bead solution (MoBio). We extracted DNA from isolated cells using the PowerSoil kit according to the manufacturer's instructions and stored at −80 °C. All samples were amplified using the same one-step PCR step and normalized with primers designed to attach a 12-base-pair barcode and Illumina adaptor sequence to the fragments during PCR[35]. We used chloroplast-excluding primers targeting the V5–V6 region (799F and 1115R; ref. 36) of the 16S rRNA gene. These primers contained a heterogeneity spacer along with the Illumina linker sequence (forward (799F): 5′-CAAGCAGAAGACGGCATACGAGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTxxxxxxxxxxxxHS-AACMGGATTAGATACCCKG-3′, reverse (1115R): 5′-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTxxxxxxxxxxxxHS-AGGGTTGCGCTCGTTG-3′) where x represents barcode nucleotides and HS represents a 0–7-base-pair heterogeneity spacer. Each sample was submitted to a single 25 µl PCR reaction containing 5 µl 5×HF buffer (Thermo Scientific), 0.5 µl dNTPs (10 µM), 0.5 µl forward primer (10 µM), 0.5 µl reverse primer (10 µM), 0.75 µl DMSO, 0.25 µl Phusion HotStart II polymerase (Thermo Scientific), 1 µl DNA, and 16.5 µl molecular-grade water. The reaction was performed using: 30 s initial denaturation at 98 °C, 35 cycles of 15 s at 98 °C, 30 s at 60 °C, and 30 s at 72 °C, with a final 10-min elongation at 72 °C. The samples were processed with an Invitrogen Sequalprep PCR Cleanup and Normalization Kit to be then pooled with equal concentration and then sequenced.

Sequence processing. Samples were sequenced on the Illumina MiSeq platform. We processed the raw sequence data with PEAR[37] and QIIME[38] pipelines to merge paired-end sequences to a single sequence of length of approximately 350 bp, eliminate low quality sequences (mean quality score < 30 or with any series of five bases with a quality score < 30), and de-multiplex sequences into samples. We eliminated chimaeric sequences using the Uclust and Usearch algorithms[39]. Then, we binned the remaining sequences into operational taxonomic units (OTUs) at a 97% sequence similarity cut-off. We determined the taxonomic identity of each OTU using the BLAST algorithm and Greengenes database[40] as implemented in QIIME[38]. After filtering OTUs that were represented by less than 20 sequences, our database contained 6,834 OTUs. The number of sequences per sample ranged from 4,006 to 40,900. From a database of 8,965,472 quality sequences, we rarefied each sample to 3,500 sequences, with 14 samples excluded from subsequent analyses due to insufficient sequence reads as a result of extraction or sequencing errors, totalling 2,121,000 sequences from 606 samples. We performed analyses with the ape[41], picante[42], and vegan[43] packages in R[44].

Statistical analyses. At the tree level, we used a PERMANOVA (Bray–Curtis dissimilarities) to identify the main drivers of leaf bacterial community structure, and a linear mixed-model to test the effect of the same drivers on bacterial diversity (Shannon index). Plant species richness, functional diversity and functional identity are continuous variables describing the characteristics of the plot where the tree was sampled, whereas host species identity is a categorical variable with 19 levels. The leaf bacterial community structure is the OTU community matrix of the 606 samples. Only the interaction between plant species richness and host species identity was removed from the initial model fitted. The linear mixed-model is also tested at the tree level, there is a sample representing the alpha-diversity of the leaf bacterial community of each tree sampled. We compared the strength of the variables in the linear mixed model by an ANOVA type II test and computed a marginal pseudo-R^2 for the model[45,46]. The model formula is:

$$\text{leaf diversity} \approx \text{host species identity} + \text{functional diversity} + \text{functional identity} + (1|\text{block/plot})$$

where fixed effects included leaf bacterial diversity, functional diversity and functional identity as continuous variables and host species identity as a 19 levels factor.

At the plot level, we built a structural equation model to test for the direct and indirect effects of host tree identity and diversity on leaf bacterial diversity and plant community productivity (Fig. 2 and Extended Data Figs 3–6). We quantified plot alpha-bacterial diversity using the Shannon diversity index on all samples from each plot combined. Before fitting the structural equation model, variables were transformed to achieve normality. Productivity and species richness were log-transformed while functional diversity and leaf bacterial diversity were both rank transformed. The following two covariances were removed from the a priori model (Extended Data Fig. 3): the covariances between plant functional identity and both plant species richness ($P = 0.45$) and plant functional diversity ($P = 0.90$). We also tested the influence of soil microtopography on plant community productivity and leaf bacterial diversity (Extended Data Fig. 4). The correlations between microtopography with both plant community productivity ($P = 0.27$) and leaf bacterial diversity ($P = 0.99$) were not significant and therefore were excluded from the final model. To address the challenge of controlling for changes in leaf bacterial community composition at different bacterial community diversities, we quantified leaf bacterial community structure by using the two axes of a non-metric multi-dimensional scaling (NMDS) analysis on Bray–Curtis dissimilarities of the OTU community matrix. We added the plot scores on each axis as a variable representing leaf bacterial community composition in the structural equation modelling (SEM) model to account for their potential influence on plant community diversity (see Extended Data Fig. 5). However, none of the components of leaf bacterial community composition had a significant effect on productivity ($P = 0.11$ and $P = 0.17$) and the effect of leaf bacterial diversity remained unchanged (0.11, $P = 0.02$; see Extended Data Fig. 6). The influence of variation in leaf bacterial community composition was not significant and did not change the correlation between bacterial diversity and plant community productivity, and thus was removed from the final model (Extended Data Figs 5, 6).

In all analyses, we started with the fully specified model and eliminated the least non-significant relationship until none remained. Variables' explanatory power is inferred from their respective R^2 value (PERMANOVA, Table 1), F values (ANOVA on linear mixed-model, Table 2) or standardized regression coefficient (structural equation model, Fig. 2) rather than P values. For the PERMANOVA and linear-mixed model, we blocked by block and plot identity to account for any non-random difference in local conditions.

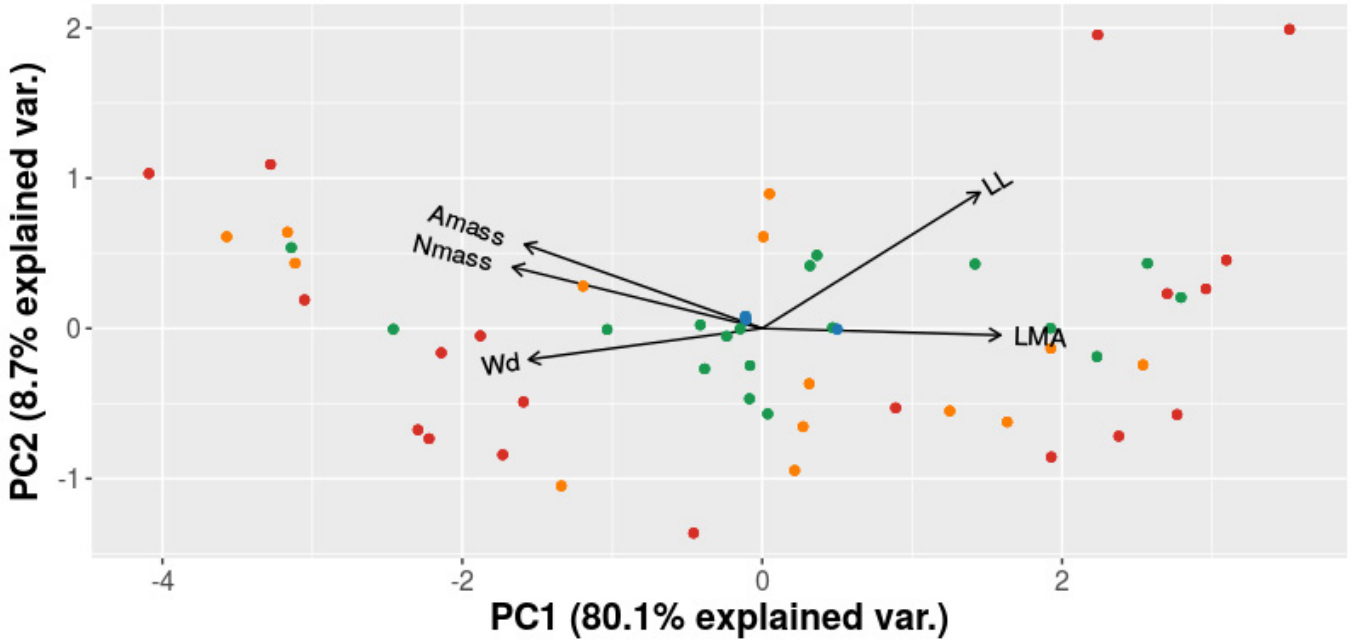
Data availability. The sequence data, metadata and code are deposited at: https://figshare.com/s/2caf0782858e7e706783 (community matrix, all metadata and

community taxonomy), https://figshare.com/s/150f044f16a7714cc395 (sequences) and https://figshare.com/s/3642e0960b600bb0f40c (R markdown code). All other data are available from the corresponding author upon reasonable request.

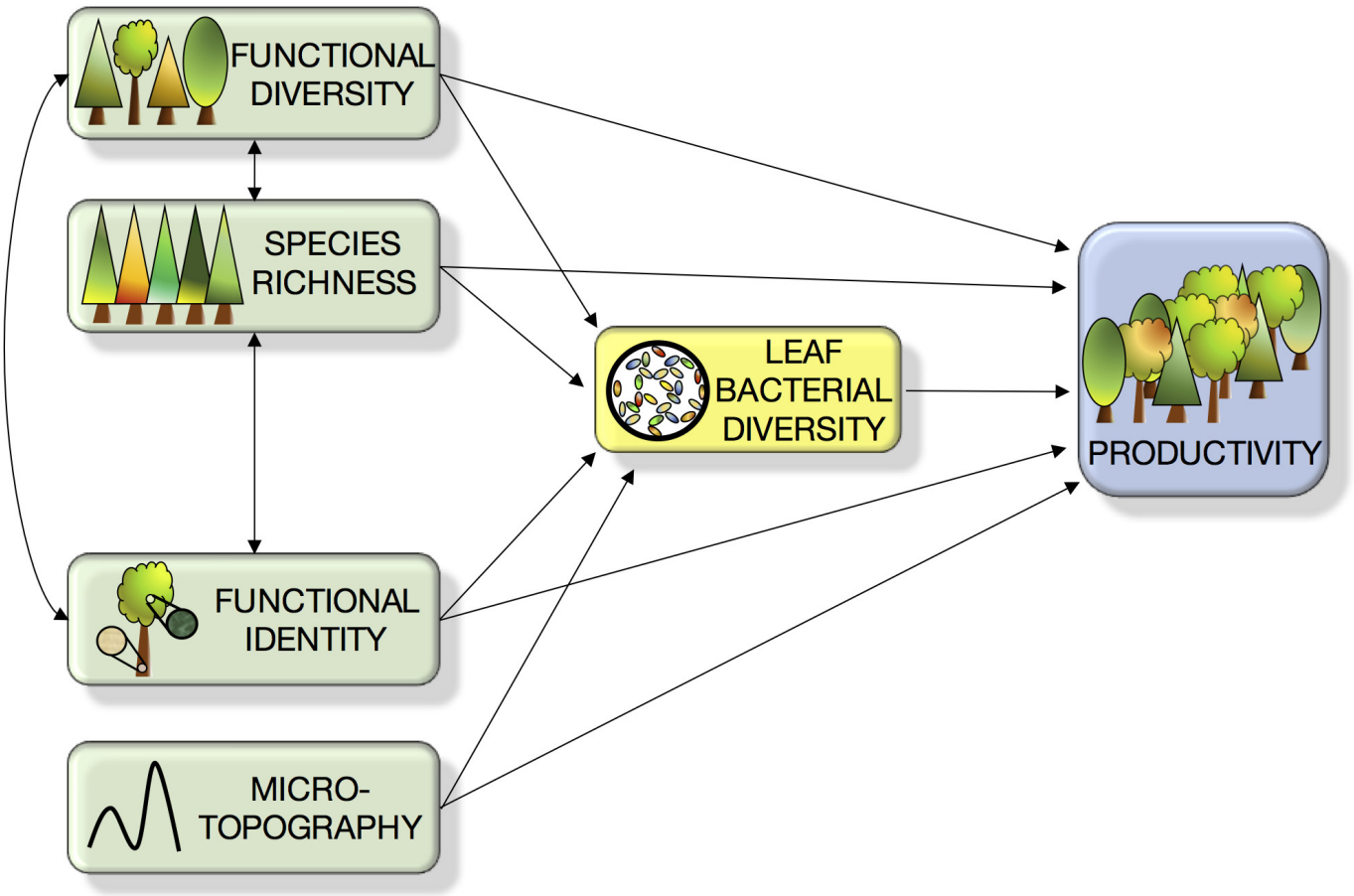
30. Tobner, C. M., Paquette, A., Reich, P. B., Gravel, D. & Messier, C. Advancing biodiversity-ecosystem functioning science using high-density tree-based experiments over functional diversity gradients. *Oecologia* **174,** 609–621 (2014).
31. Verheyen, K. *et al.* Contributions of a global network of tree diversity experiments to sustainable forest plantations. *Ambio* **45,** 29–41 (2016).
32. Tobner, C. M. *et al.* Functional identity is the main driver of diversity effects in young tree communities. *Ecol. Lett.* **19,** 638–647 (2016).
33. Laliberté, E. & Legendre, P. A distance-based framework for measuring functional diversity from multiple traits. *Ecology* **91,** 299–305 (2010).
34. Lavorel, S. *et al.* Assessing functional diversity in the field–methodology matters! *Funct. Ecol.* **22,** 134–147 (2008).
35. Fadrosh, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2,** 6 (2014).
36. Redford, A. J., Bowers, R. M., Knight, R., Linhart, Y. & Fierer, N. The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* **12,** 2885–2893 (2010).
37. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30,** 614–620 (2014).
38. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Met.* **7,** 335–336 (2010).
39. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26,** 2460–2461 (2010).
40. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72,** 5069–5072 (2006).
41. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20,** 289–290 (2004).
42. Kembel, S. W. *et al.* Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26,** 1463–1464 (2010).
43. Oksanen, J. *et al.* The vegan package. *Community Ecology Package* 10, 631–637 (2007).
44. R Development Core Team. *R: A Language and Environment for Statistical Computing*; http://www.R-project.org/ (Vienna, Austria, 2013).
45. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4,** 133–142 (2013).
46. Johnson, P. C. Extension of Nakagawa & Schielzeth's R(2)GLMM to random slopes models. *Methods Ecol. Evol.* **5,** 944–946 (2014).
47. Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12,** 351–366 (2009).
48. Niinemets, Ü. & Valladares, F. Tolerance to shade, drought and waterlogging of temperate, northern hemisphere trees and shrubs. *Ecol. Monogr.* **76,** 521–547 (2006).
49. Royal Botanic Gardens Kew. *Seed Information Database (SID)* version 7.1; http://data.kew.org/sid/ (2008).
50. Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428,** 821–827 (2004).

SPECIES RICHNESS	FUNCTIONAL DIVERSITY								
	0	1	2	3	4	5	6	7	8
1	19 host species								
2		Pru*Pre	Ll*Pst Ar*Ba Pre*Pst	Ba*Qru	Bp*Qru Ll*Pg	Pg*Pst Bp*Pst	Ab*Ar	As*Ll Ab*As Ar*To	As*To
4		Ab*Pg*Pru*Pre Ar*Ba*Ap*Tc	Ar*Ba*Bp*Qru Pg*Pru*Pre*Pst Ll*Pru*Pa*Po	Ba*Pru*Pre*Pst Pre*Qru*Ld*Psy	Ab*Bp*Ll*Pg Pg*To*Pa*Psy	Ab*As*Pg*Pru Pg*Pst*Qro*Tc	Ll*Pst*Qru*To Bp*Pru*Qro*Psy	Ar*Ps*Qru*To As*Bp*Pg*Pst Bp*As*Ap*Po	As*Ba*Pg*To Bp*To*Ap*Pa
12						Ab*Ar*As*Ba*Bp*Ll* Pg*Pru*Pre*Pst*Qru*To To*Qru*Pst*Pre*Pg*Bp* Ap*Pa*Po*Psy*Ld*Tc	As*Ba*Pru*Ar*Ab*Ll* Ap*Pa*Po*Psy*Qro*Tc		

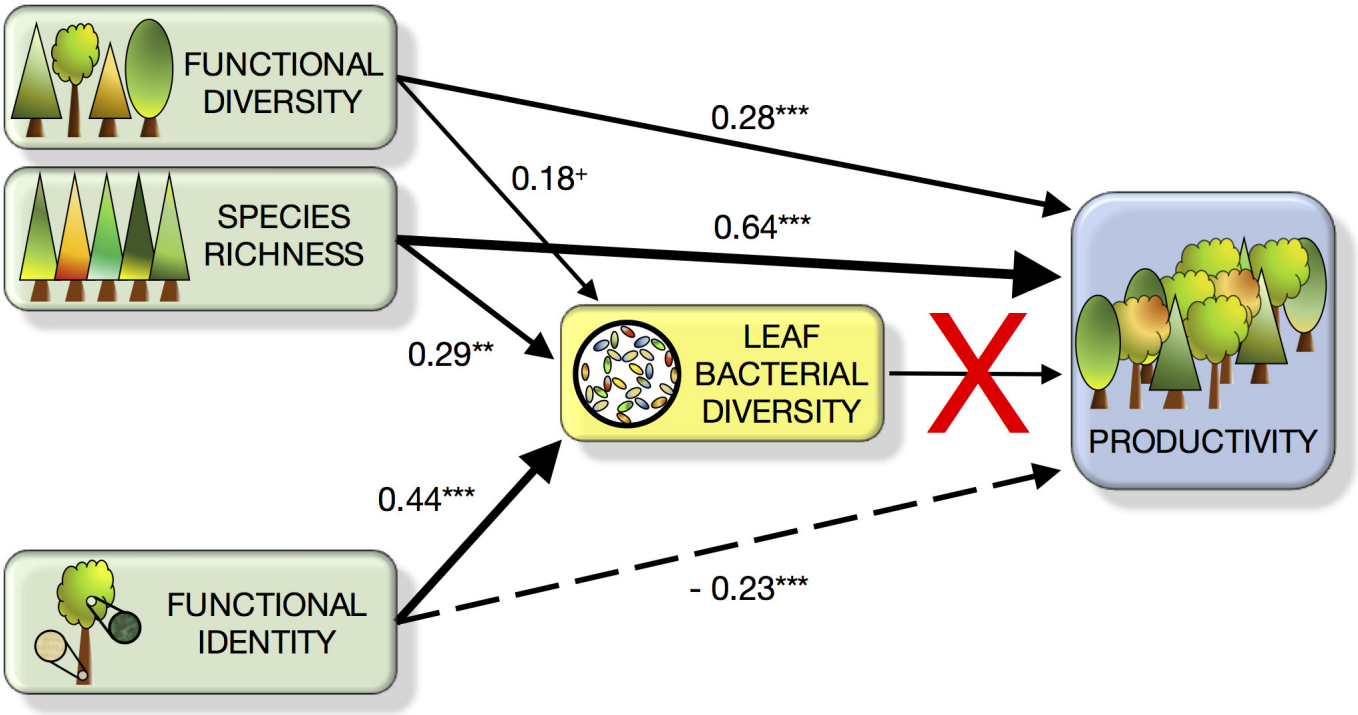
Extended Data Figure 1 | Identity of the tree host species in each of the 54 combinations at the IDENT experiment in Montreal. Ab, *Abies balsamea*; Ap, *Acer platanoides*; Ar, *Acer rubrum*; As, *Acer saccharum*; Ba, *Betula alleghaniensis*; Bp, *Betula papyrifera*; Ld, *Larix decidua*; Ll, *Larix laricina*; Pa, *Picea abies*; Pg, *Picea glauca*; Po, *Picea omorika*; Pre, *Pinus resinosa*; Pru, *Picea rubens*; Pst, *Pinus strobus*; Psy, *Pinus sylvestris*; Qro, *Quercus robur*; Qru, *Quercus rubra*; Tc, *Tilia cordata*; To, *Thuja occidentalis*.



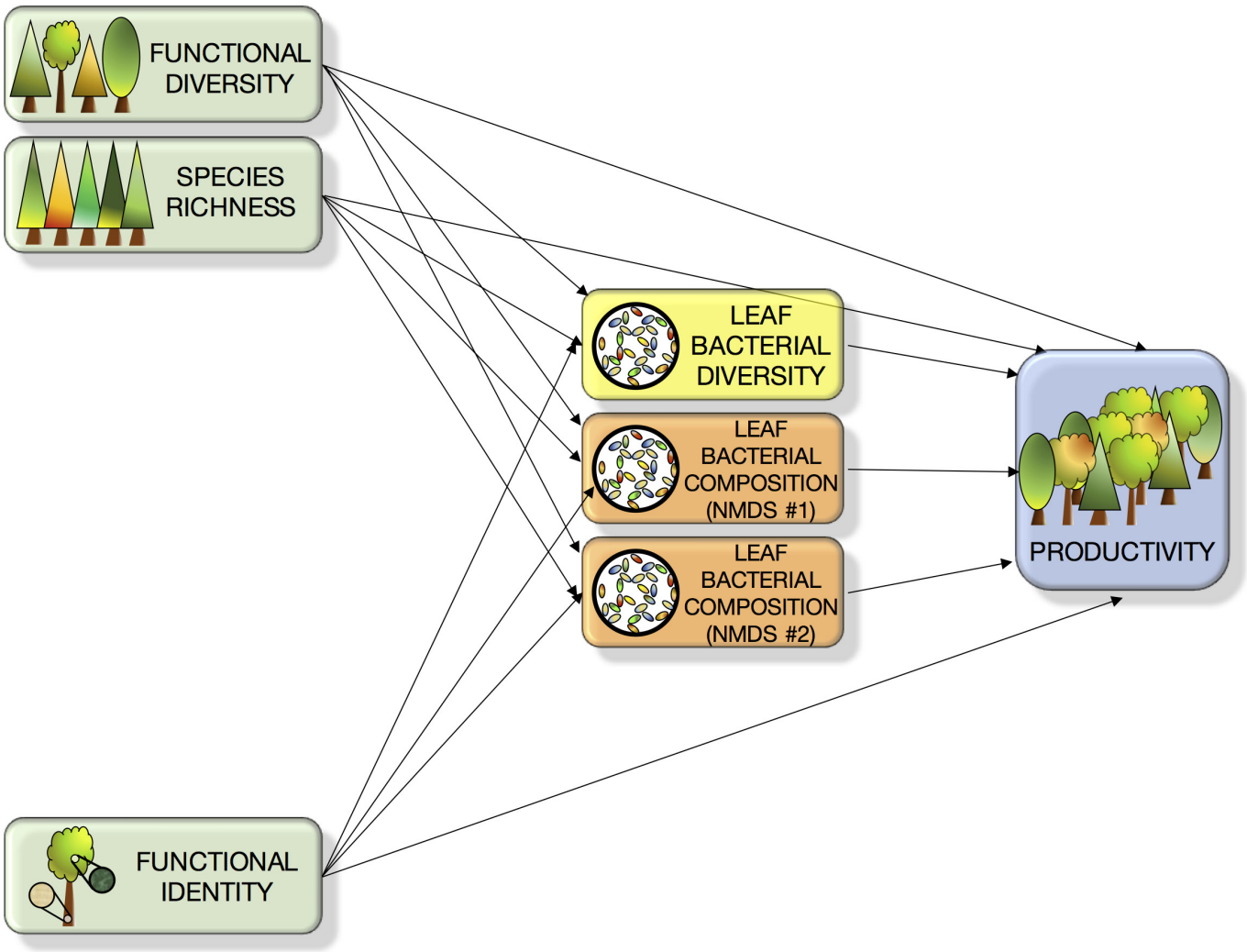
Extended Data Figure 2 | Principal component analysis on functional traits community weighted means. Traits are: maximum photosynthetic capacity (A_{mass}), nitrogen content of leaves (N_{mass}), leaf longevity (LL), wood density (WD) and leaf mass per area (LMA). Colours represent plot species richness levels (red for one species, orange for two, green for four, and blue for twelve).



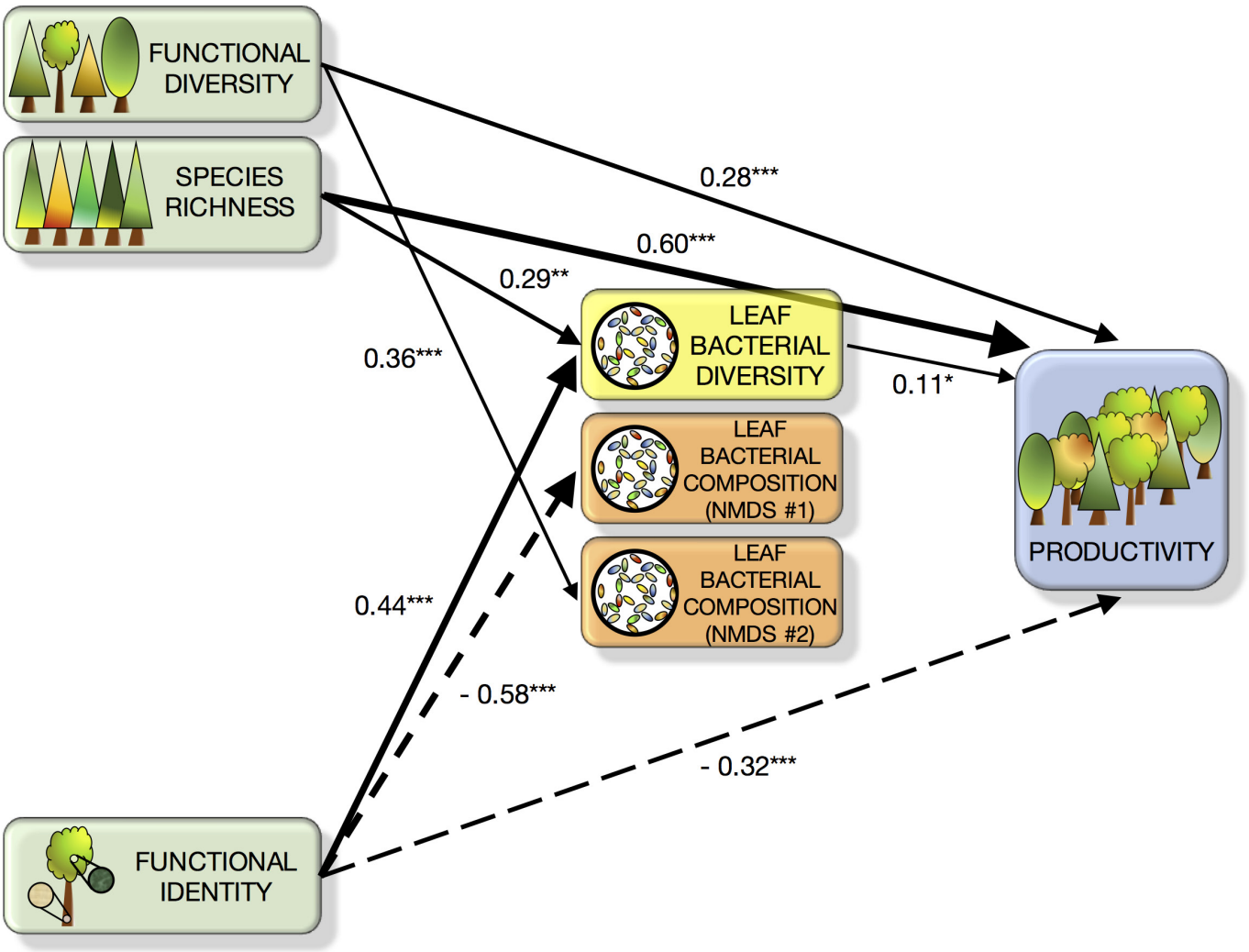
Extended Data Figure 3 | A priori structural equation model. Factors are species richness, functional identity, functional diversity and plot microtopography (elevation at plot centre, cm) as determinants of leaf bacterial diversity and plant community productivity. Green boxes indicate exogenous variables (diversity indices and plot microtopography), whereas responses are in yellow for plot-level leaf bacterial diversity and blue for plant community productivity.



Extended Data Figure 4 | Alternative structural equation model excluding the link between leaf bacterial diversity and plant community productivity. After deletion of this link, the path analysis ($n = 216$, $\chi^2 = 11.906$, $P = 0.008$, df = 3; RMSEA = 0.00, $P = 0.044$) is unstable and inferior to the model with the leaf bacterial diversity–plant community productivity link included. Green boxes indicate plot-level plant diversity indices; yellow denotes plot-level leaf bacterial diversity; blue denotes plant community productivity. Numbers adjacent to arrows and arrow width indicate the effect size of the relationships and the associated bootstrap P value. $+P < 0.1$; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. Continuous and dashed arrows indicate positive and negative relationships, respectively.



Extended Data Figure 5 | Structural equation model of plant diversity and identity explaining leaf bacterial diversity and community composition, as well as plant community productivity (full model tested). Green boxes indicate plot-level plant diversity indices; yellow denotes plot-level leaf bacterial diversity; orange indicates plot-level leaf bacterial identity; blue denotes for plant community productivity. The covariances between leaf bacterial diversity and the two variables of leaf bacterial community composition were also included in the model.



Extended Data Figure 6 | Structural equation model of plant diversity and identity explaining leaf bacterial diversity and community composition, as well as plant community productivity. The path analysis ($n = 216$, $\chi^2 = 1.522$, $P = 0.677$, df = 3; RMSEA = 0.00, $P = 0.821$) explains 38% of the variance in leaf bacterial diversity, 34% and 11% of each components of bacterial identity, and 85% of the variance in plot productivity. Green boxes indicate plot-level plant diversity indices; yellow denotes plot-level leaf bacterial diversity; orange indicates plot-level leaf bacterial identity; blue denotes plant community productivity. Numbers adjacent to arrows and arrow width indicate the effect size of the relationships and the associated bootstrap P value. $+P < 0.1$; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. Continuous and dashed arrows indicate positive and negative relationships, respectively. The covariances between leaf bacterial diversity and the two variables of leaf bacterial community composition were also included in the model.

Extended Data Table 1 | Host species functional traits

Host Species	Functional Traits								
	A_{mass} (umol* $m^{-2}*s^{-1}$)	D_{tol} (scale 0-5)	LL (month)	LMA ($g*m^{-2}$)	N_{mass} (%)	$Seed_{mass}$ (g/1000)	$Shade_{tol}$ (scale 0-5)	$Water_{tol}$ (scale 0-5)	WD ($g*cm^{-3}$)
Abies balsamea	12.0	1.0	110	151.0	1.66	7.6	5.0	2.0	0.33
Acer platanoides	83.1	2.7	6	50.6	1.99	139.0	4.2	1.5	0.52
Acer rubrum	111.2	1.8	5.6	71.1	1.91	26.5	3.4	3.1	0.49
Acer saccharum	84.6	2.3	5.5	70.6	1.83	55.2	4.8	1.1	0.56
Betula alleghaniensis	206.0	3	5.5	46.1	2.20	0.9	3.2	2.0	0.55
Betula papyrifera	195.0	2	3.6	77.9	2.31	0.4	1.5	1.3	0.48
Larix decidua	71.6	2.3	6	93.9	2.05	7.1	1.5	1.1	0.47
Larix laricina	59.4	2	6	120	1.36	2.0	1.0	3.0	0.49
Picea abies	29.3	1.8	103.2	235.2	1.19	7.0	4.5	1.2	0.37
Picea glauca	35.6	2.9	50	302.9	1.28	2.4	4.2	1.0	0.33
Picea omorika	NA	2.8	NA	NA	1.03	2.9	4.7	1.0	0.36
Picea resinosa	24.0	3	36	294.1	1.17	8.0	1.9	1.0	0.41
Picea rubens	NA	2.5	103.2	304.7	1.15	3.3	4.4	2.0	0.37
Pinus strobus	43.8	2.3	20	121.9	1.42	17.0	3.2	1.0	0.34
Pinus sylvestris	36.9	4.3	27.9	254.6	1.33	6.0	1.7	2.6	0.42
Quercus robur	85.1	3	6	68.5	2.37	3378.0	2.5	1.9	0.56
Quercus rubra	148.6	2.9	6	84.2	2.06	3143.0	2.8	1.1	0.56
Thuya occidentalis	32.2	2.7	33	223	1.02	1.4	3.5	1.5	0.30
Tilia cordata	NA	2.8	4.8	49.1	2.13	50.9	4.2	1.83	0.42

Traits are maximum photosynthetic capacity (A_{mass}), drought tolerance (DT), leaf longevity (LL), leaf mass per area (LMA), leaf nitrogen content (N_{mass}), seed mass (SM), shade tolerance (ST), water tolerance (WT), and wood density (WD) from global databases[47–50]. Tolerance indices are based on a 0 (no tolerance) to 5 (maximal tolerance) scale.