LETTER

ECOLOGY LETTERS | cnrs WILEY

Host specificity of plant-associated bacteria is negatively associated with genome size and host abundance along a latitudinal gradient

Zihui Wang[1,2] | Geneviève Lajoie[3,4] | Yuan Jiang[5] | Minhua Zhang[2] | Chengjin Chu[5] | Yongfa Chen[5] | Shuai Fang[6] | Guangze Jin[7] | Mingxi Jiang[8] | Juyu Lian[9] | Yanpeng Li[10] | Yu Liu[2] | Keping Ma[11] | Xiangcheng Mi[11] | Xiujuan Qiao[8] | Xihua Wang[12] | Xugao Wang[6] | Han Xu[10] | Wanhui Ye[9] | Li Zhu[11] | Yan Zhu[11] | Fangliang He[2,13] | Steven W. Kembel[1,2]

[1]Université du Québec à Montréal, Montréal, Québec, Canada

[2]ECNU-Alberta Joint Lab for Biodiversity Study, Tiantong Forest Ecosystem National Observation and Research Station, School of Ecology and Environmental Sciences, East China Normal University, Shanghai, China

[3]Université de Montréal, Montréal, Québec, Canada

[4]Jardin Botanique de Montréal, Montréal, Québec, Canada

[5]State Key Laboratory of Biocontrol, School of Ecology, Sun Yat-Sen University, Guangzhou, China

[6]CAS Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, China

[7]Center for Ecological Research, Northeast Forestry University, Harbin, China

[8]Key Laboratory of Aquatic Botany and Watershed Ecology, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

[9]Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

[10]Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou, China

[11]State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing, China

[12]Tiantong Forest Ecosystem National Observation and Research Station, School of Ecology and Environmental Sciences, East China Normal University, Shanghai, China

[13]Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada

Correspondence
Steven W. Kembel, SB-2130, 141 Ave Président-Kennedy, Montréal, QC H2X1Y4, Canada.
Email: kembel.steven_w@uqam.ca

Fangliang He, 322A, 11328 - 89 Ave NW, Edmonton, AB T6G 2J7, Canada.
Email: fhe@ualberta.ca

Funding information
Natural Sciences and Engineering Research Council of Canada; China Scholarship Council; Canada Research Chairs

Editor: Vojtech Novotny

Abstract

Host specialization plays a critical role in the ecology and evolution of plant–microbe symbiosis. Theory predicts that host specialization is associated with microbial genome streamlining and is influenced by the abundance of host species, both of which can vary across latitudes, leading to a latitudinal gradient in host specificity. Here, we quantified the host specificity and composition of plant–bacteria symbioses on leaves across 329 tree species spanning a latitudinal gradient. Our analysis revealed a predominance of host-specialized leaf bacteria. The degree of host specificity was negatively correlated with bacterial genome size and the local abundance of host plants. Additionally, we found an increased host specificity at lower latitudes, aligning with the high prevalence of small bacterial genomes and rare host species in the tropics. These findings underscore the importance of genome streamlining and host abundance in the evolution of host specificity in plant-associated bacteria along the latitudinal gradient.

INTRODUCTION

Host specialization is a widespread phenomenon in plant–microbe associations (Mariadassou et al., 2015), defined here as the tendency that microbes are associated with phylogenetically related plant species (Gilbert & Webb, 2007). This specialization plays a critical role in shaping the ecology and evolution of both host plants and microbial symbionts. From the microbial perspective, specialization on a certain set of hosts can promote microbial coexistence through resource partitioning, driving genome evolution and, in some cases, fostering microbial diversification (Brochet et al., 2021; Oh et al., 2010; Qin et al., 2019). Conversely, from the perspective of plants, interactions with specialist microbial symbionts facilitate biotic niche partitioning among plant species, thereby influencing plant species diversity and distribution (Benítez et al., 2013; Bever et al., 2010; Semchenko et al., 2022). Despite the significance of host specialization in the ecology and evolution of both partners, our understanding of the origins and maintenance of host specialization in plant-associated microbes remains limited (Cobian et al., 2019; Lajoie & Parfrey, 2022).

Over the course of microbe–host coevolution, plants impose a consistent selection pressure on microbial traits that convey a high fitness in host-dependent habitats, driving the genomic and functional evolution of microbes (Kwong & Moran, 2015; Sieber et al., 2021). The adaptive evolution of microbes can lead to different consequences for host specialization. On the one hand, microbes may gain fitness advantages in a specific host-associated habitat at the cost of reduced fitness in other habitats (i.e. performance trade-offs), leading to host specialization (Caley & Munday, 2003). Specialization provides a stable environment conducive to genome streamlining, where only the most essential functions that fit well in the specialized habitats are retained (Baroncelli et al., 2016; Klasson & Andersson, 2004). This could result in a reduction of genome size and genomic GC content in specialist versus generalist microbes, as commonly observed in both symbiotic and free-living bacteria (McCutcheon & Moran, 2012; Qin et al., 2019; Swan et al., 2013). On the other hand, microbial symbionts can acquire functional genes needed to colonize novel host habitats through horizontal gene transfer, thereby expanding phenotypic and genetic repertoires, potentially promoting host generalization (Arnold et al., 2022; von Meijenfeldt et al., 2023). As a result, large microbial genomes consisting of high functional diversity are more likely adapted to a variety of habitats and should be associated with lower host specificity compared to those

of small genome size and less functional complexity, as illustrated in Figure 1a.

Given their short generation times, large populations and high mutation rates, microbes have the ability to quickly adapt to various environments through adaptive evolution (Shi et al., 2021). The variability of biotic and abiotic environments that microbes are exposed to could influence the dynamics of adaptation and further shape the evolution of host specialization (Lajoie & Parfrey, 2022; Mestre et al., 2020). For example, studies on animal–microbial symbioses have shown that microbes associated with mobile and socially gregarious mammals are broadly shared among host individuals, which fosters host-generalized lifestyles in microbial symbionts (Grieneisen et al., 2017; Perofsky et al., 2021). For plant-associated microbes that mainly disperse among neighbouring host plants, the variability of microbial habitats tends to depend on the diversity of neighbouring host species that they potentially colonize (Meyer et al., 2022, 2023). Common plant species are more widespread and thus neighbour more of the plant species in a community than rare species do, and therefore their associated microbes are expected to have a broader distribution across host species compared to those associated with rare plant species (Figure 1b, Miller et al., 2018). As a result, the adaptive evolution of microbes across varying host-associated habitats impedes the maintenance of host specialization, leading to a decreased host specificity on common versus rare host species (Figure 1b). Nevertheless, empirical evidence regarding the interplay among host species abundance, microbial distribution and host specialization is lacking.

The pace of microbial genome evolution, the abundance of host species and the degree of host specialization can differ between tropical and temperate regions, potentially exhibiting latitudinal gradients. For example, molecular evolution rates tend to increase with temperature, resulting in a rapid genome streamlining, and thus smaller bacterial genomes in warmer climates (Sabath et al., 2013; Sorensen et al., 2019; Swan et al., 2013). Moreover, tropical forests harbour a higher proportion of rare plant species than temperate forests, providing more isolated microbial habitats that favour host specialization (Enquist et al., 2019). Many broad-scale empirical studies have indicated that trophic interactions are more intense in lower latitudes, and that tropical tree species may be more chemically distinct than temperate tree species (Baskett & Schemske, 2018; Schemske et al., 2009). These differences between tropical and temperate communities could lead to more diverse microbial niches and stronger selection pressures from host plants in lower latitudes, therefore fostering a higher host specificity
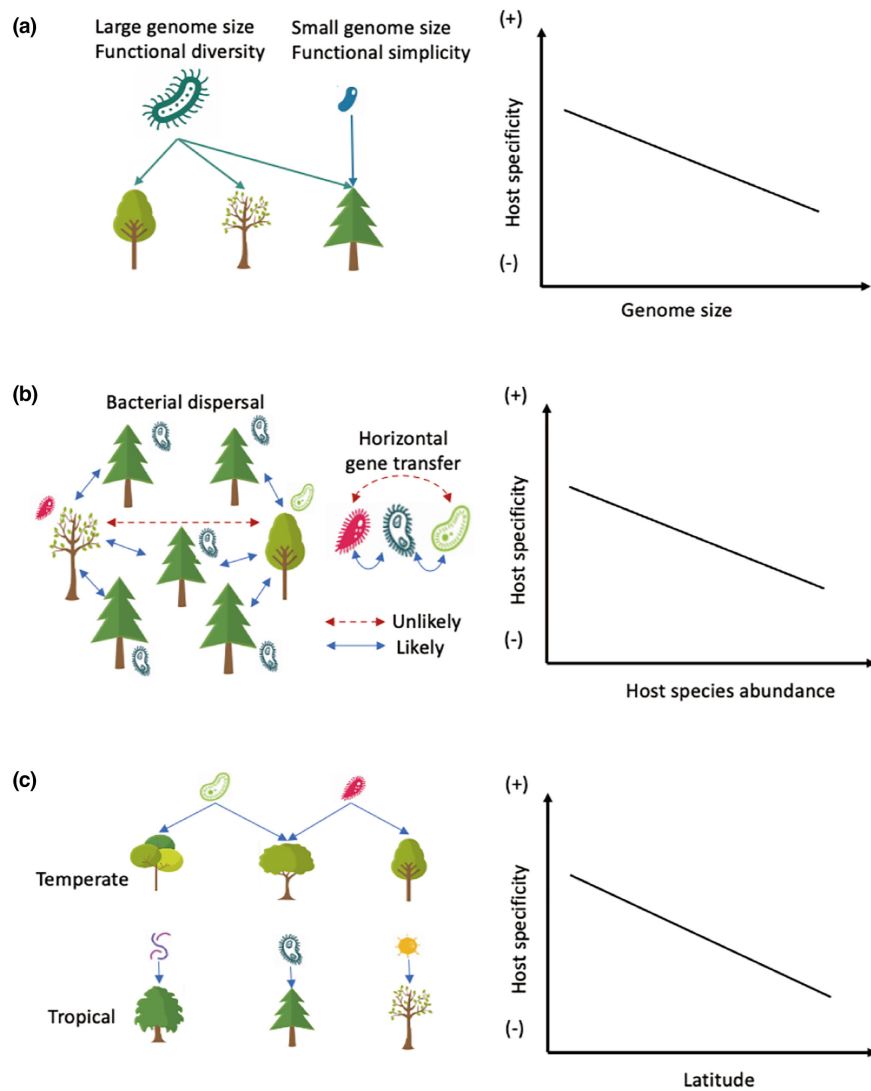
FIGURE 1 Hypotheses on the factors that influence host specialization of plant–bacteria symbioses. (a) Bacteria with larger genome size and higher functional complexity are associated with lower host specificity; (b) host plant abundance facilitates the dispersal and horizontal gene transfer of bacteria, reducing host specialization; and (c) host specificity is higher at lower latitudes as a result of stronger selection pressure and higher diversity of host plants in tropical than temperate regions.

in the tropics (Figure 1c). While latitudinal gradients of host specificity have been widely studied in plant–herbivore and plant-pollinator interactions (Dyer et al., 2007; Novotny et al., 2006; Ollerton & Cranmer, 2002), our understanding of the latitudinal pattern of host specificity in plant–microbe symbiosis remains limited, despite the crucial role of microbial symbionts in maintaining plant diversity and supporting ecosystem functioning (Benítez et al., 2013; Bever et al., 2010; Semchenko et al., 2022).

Phyllosphere bacteria, the bacterial colonizers of the aerial surfaces of plants, represent a major group of plant–microbial symbioses (Lindow & Leveau, 2002; Vorholt, 2012). Numerous studies have shown that phyllosphere bacteria play a critical role in promoting host plant health, productivity, and ecosystem functioning (Fürnkranz et al., 2008; Laforest-Lapointe et al., 2017; Liu et al., 2020; Vacher et al., 2016). The phyllosphere habitat is exposed to surrounding environments including air, soil and neighbouring plants, and thus phyllosphere bacterial assemblages are shaped by both host selection and bacterial dispersal among habitats (Vorholt, 2012; Wang et al., 2023). Host selection pressure can facilitate the host specialization of phyllosphere bacteria, while bacterial dispersal reduces host specialization, leading to conflicting pressures in generating specialists versus generalists (Lajoie & Parfrey, 2022). Previous studies have shown that phyllosphere bacteria ranged from specialists to generalists, making it an excellent model system for investigating the origin and maintenance of host specialization (Lajoie & Kembel, 2021). In this study, we assess the host specificity of phyllosphere epiphyte bacteria across 329 plant species along a broad latitudinal gradient to test the following predictions: (1) host specialization should be correlated with a reduced genome

size, genomic GC content and cell size in bacterial symbionts; (2) bacteria on common host species should (a) be more widespread and have a general composition within the local community, and (b) exhibit lower host specificity relative to bacteria on rare plant species; and (3) host specialization in plant–bacteria symbioses should be more prevalent in tropical than temperate regions (Figure 1).

MATERIALS AND METHODS

Study sites

We conducted the study at 10 forest plots spanning a latitude range of 18.7° N to 51.8° N in China (Wang et al., 2023). These plots varied in size from 10 to 60 ha and provided habitats for 44 to 289 tree species. Plant individuals within each forest plot were identified and mapped according to the protocol of the Forest Global Earth Observatory Networks (ForestGEO http://forestgeo.si.edu/). Approximately one-third of tree species within each plot were selected based on a species abundance gradient, ranging from rarer to more common plant species in the forest plot. In total, we selected 329 tree species from 148 genera and 59 families, among which 225 species were collected at a single site and the remaining 104 species were collected at more than one site. We randomly chose approximately three individuals per plant species per site, resulting in a total of 1453 tree individuals (Table S1).

Sampling and processing

We collected 50–100 g of leaves from the sub-canopy (2–10 m above ground) of each tree using sterile plastic bags from July to August 2019. Within 24 h of field sampling, we washed epiphytic microbes from the leaves using a 1:50 dilution of Redford buffer (1 M Tris, 500 mM EDTA, 1.2% CTAB) by agitating leaves for 5 min (Kembel et al., 2014). The suspension was centrifuged at 3300 *g* for 25 min, and the resulting pellet was collected and stored at −20°C until DNA extraction. Microbial DNA was extracted following the manufacturer's protocol (QIAGEN PowerSoil DNA extraction kit, single-tube kit). We amplified V5–V6 region of bacteria 16S rRNA genes using primers 799F and 1115R to determine the composition of the phyllosphere bacterial community (Kembel et al., 2014). These primers exclude plant chloroplasts and cyanobacterial sequences to prevent PCR contamination by plant DNA amplification (Rastogi et al., 2010). PCR products were normalized using a Just-a-plate 96 PCR Purification and Normalization kit (Charm Biotech, Cape Girardeau, MO, USA), pooled and then purified using AMPure XP beads (Beckman Coulter, Mississauga, Canada) to remove contaminants.

The DNA library was then sequenced using Illumina MiSeq Reagent Kit for 300 bp paired-end sequencing.

Bioinformatics

We used the DADA2 package in R (package version 1.24.0, R version 4.2) to identify the amplicon sequence variants (ASV) present in each sample (Callahan et al., 2016). The nucleotides at 19–240 and 16–240 bp positions were kept for forward and reverse DNA sequences, respectively. We estimated error rates and inferred ASVs using DADA2 default parameters (Callahan et al., 2016). Nonbacterial sequences and chimeric sequences were removed. Taxonomy was assigned to each ASV by comparison to the SILVA ribosomal RNA gene database (v.138.1; Quast et al., 2013). We obtained approximately 116 million high-quality 16S rRNA gene sequences and identified 55,693 ASVs from 1453 samples. To account for variation in sequencing depth among samples, we randomly rarefied the ASV table to 15,000 sequences per sample using the R package 'vegan' (v.2.6.2), and 49,254 ASVs were retained after rarefaction. Most ASVs (82%) were taxonomically classified to the family level, 58% were classified to genus and only 4% were identified at the species level.

Metadata

We collected three main categories of metadata related to host plants, bacterial ASVs and abiotic environments. Host plant-related metadata included host species identity, species relative abundance within the forest plot, and plant leaf traits such as specific leaf area, leaf phosphorus and nitrogen concentrations. Leaf traits were measured following the procedures of Cornelissen et al. (2003): 5–6 of the largest and smallest individuals of each tree species were selected for sampling, and two to five leaves were measured for each individual. Fresh mass of the sampled leaves was recorded upon leaf removal and dry mass was also recorded after drying at 60°C for 72 h. Additionally, we constructed a plant species-level phylogenetic tree using the R package 'V.phyloMaker' (v. 0.1.0) based on a comprehensive phylogeny database for vascular plants (Jin & Qian, 2019).

For bacterial attributes, we collected bacterial phylogenies and genomic traits. A bacterial phylogenetic tree was reconstructed using 16S rRNA gene sequence of ASVs. Multiple sequence alignment was performed using the *AlignSeqs* function in R package 'DECIPHER' (v. 2.26) with default parameters (Wright, 2016). The aligned sequences were used to build a phylogenetic tree with FastTree 2.0 using the GTRCAT model of evolution (Price et al., 2009). Bacterial traits were obtained from a compilation of microbial traits from published data repositories (Madin et al., 2020). While bacterial

functional traits may exhibit plasticity within species or even strains, certain traits such as genome size and GC content display robust phylogenetic conservatism and can be predicted by taxonomic identity (Martinez-Gutierrez & Aylward, 2022). Six traits that could be reliably predicted by bacterial taxonomic identity were selected, including categorical variables (Gram-positive/negative, motility) and numerical variables (cell length, genome size, genomic GC content, number of tRNA genes) (Madin et al., 2020). Trait values were assigned to each ASV based on taxonomic identity. ASVs with species identity were assigned the mean trait value of the species (for numerical traits) or the predominant status of categorical traits. ASVs with only genus identity were assigned the mean trait value of the genus, while those with only family identity were assigned the mean trait value of the family. ASVs lacking identification beyond the family level were excluded from trait assignment and subsequent analysis. 71% of bacterial ASVs were assigned at least one bacterial trait.

Abiotic environmental data include two climatic variables, annual temperature and precipitation, collected from the WorldClim database for each study site. We also collected local environmental variables for each sample, which includes elevation, slope, soil pH, soil total concentration of nitrogen, phosphorus, organic matter, water, and soil available concentration of nitrogen, phosphorus, and potassium. These variables have been measured for all forest plots following the ForestGEO protocol (https://forestgeo.si.edu/protocols). We used Kriging to map each of these variables and estimate their values in our sampling locations, implemented with the 'automap' package (v. 1.19, Hiemstra et al., 2009).

Statistical analysis

To evaluate the host specificity of phyllosphere bacteria, we calculated the rescaled distance-based specialization index (DSI*) for each bacterial ASV using *dsi* function in 'dizzy' package (Jorge et al., 2017). The DSI* measures the phylogenetic relatedness of host species, scaled by the null expectation of random host species. The index ranges between −1 and 1 with negative values indicating generalists associated with phylogenetically dispersed plant species and positive values representing specialists associated with phylogenetically clustered plant species (Jorge et al., 2014, 2017). The null model was defined at both global scale (i.e. including plant species from all study sites) and plot scale (i.e. including the host plant species that cooccur with the ASV in the same plot), therefore generating global and plot-level DSI*, respectively. ASVs with a total abundance of less than 20 were excluded due to the potential bias associated with extreme rarity.

To explore the potential link between host specialization and bacterial traits (hypothesis 1), we employed phylogenetic generalized least squares (PGLS) to model DSI* as a function of bacterial traits. We included bacterial phylogenetic distance as a random term in the model to control for phylogenetic conservatism of host specificity (Martinez-Gutierrez & Aylward, 2022). Model fitting was conducted using the *gls* function in 'nlme' package (v. 3.1.155), and variables were selected through a stepwise algorithm implemented with the *stepAIC* function in MASS package (v. 7.3.58). Additionally, we ran a univariate linear mixed-effect model to predict DSI* as a function of genome size or GC content with bacterial family identity taken as a random term, to visualize the relationships between bacterial genomics and host specificity for different bacterial taxa.

To investigate the influence of host abundance on the composition and specialization of bacterial symbionts (hypothesis 2a and b, respectively), we calculated two indices, the local contribution to beta diversity (LCBD) and the averaged DSI*, for each sample. The LCBD, calculated with *beta.div* function in 'adespatial' package (v. 0.3.14), indicates the uniqueness of bacterial community composition, with larger values representing more distinctive bacterial community composition and greater contribution to bacterial beta diversity (Legendre, 2014; Legendre & Cáceres, 2013). The averaged DSI* was calculated as the abundance-weighted mean DSI* across ASVs to indicate the overall host specificity of the bacterial community. We then used linear mixed-effects models to model these two indices as a function of host plant abundance (i.e. the species relative abundance in the forest plot), leaf traits and abiotic environmental variables, with study site and plant species identity treated as random effects. We conducted backward selection using *step* function in 'lmerTest' package (v. 2.0.36) to remove those explanatory variables having no significant effects (at *p*-value=0.05) from the final model (Kuznetsova et al., 2017). The marginal and conditional R squared of the model were calculated using *r.squaredGLMM* function in 'MuMIn' package (v. 1.45.7). We also ran a univariate linear mixed-effect model with host abundance as the explanatory variable, study site and plant species as random effects to visualize how the two indices changed along host abundance gradients across plant species and study sites.

We employed linear regression to depict how host specificity, genome size and host abundance of bacterial ASVs changed with latitudes (hypothesis 3). The host abundance of a bacterial ASV was quantified by abundance-weighted mean value of host species' relative abundance within the plot. To assess the influence of genome size and host abundance on the latitudinal gradients of host specificity, we modelled DSI* as a function of latitude, genome size and averaged host abundance in a linear mixed-effect model, where study plot and bacterial family identity were incorporated as random terms to control for the variation of host specificity among study sites and bacterial taxa.

RESULTS

After excluding rare bacterial ASVs (those with fewer than 20 sequences), a total of 14,774 bacterial ASVs were retained, and the host specialization index DSI* was calculated for these bacterial ASVs. Our analysis showed that 89% and 73% of phyllosphere bacterial ASVs exhibited positive values of global and plot-level DSI*, respectively, indicating that the majority of bacteria residing on plant leaves were host-specialized, forming associations with phylogenetically related host plants. There was a significant phylogenetic signal in DSI* (Pagel's Lambda = 0.922, $p = 0.001$), suggesting that closely related bacteria tend to exhibit similar degrees of host specificity. We demonstrated a negative correlation between DSI* and bacterial traits including genome size and genomic GC content (Figure 2; Figure S1), consistent with the prediction of hypothesis 1. Additionally, Gram-positive bacteria displayed higher DSI* values than Gram-negative bacteria, and bacteria with motility ability exhibited greater host specificity (Table 1 and Table S2 for global and plot-level DSI*, respectively).

We found a negative relationship between host species abundance and the averaged DSI*, indicating that phyllosphere bacteria associated with common host species exhibit a lower host specificity (Table 2 and Table S3 for global and plot-level DSI*, respectively). The negative correlation between host specificity and host abundance held both within and across sites, that is, common plant species were associated with less host-specialized bacteria than rarer plant species within a study site, and higher host relative abundance was associated with lower host specificity of bacteria for the same plant species across different study sites (Figure 3), consistent with the expectation of hypothesis 2b. Apart from host abundance, the averaged DSI* of bacterial communities was influenced by factors including the specific leaf area of host plants, annual temperature and soil nutrients (Table 2 and Table S3 for global and plot-level DSI*, respectively). Overall, the model explained a total of 43% of the variation in averaged DSI*, of which 23% was explained by the fix effects of variables.

The analysis of bacterial community composition revealed that bacteria on more common host species contributed less to bacterial beta diversity (indicated by lower LCBD), therefore exhibiting a more general composition compared to bacteria on rarer plant species (Figure 3c). This result suggested that common plant species-associated bacteria are more widespread within the local forest plot, as expected in hypothesis 2a. Additionally, community uniqueness index LCBD was also affected by abiotic factors such as annual temperature, elevation and soil nutrient conditions (Table 2). Together, our model explained 72% of the variation in LCBD, with the fixed effects of variables explaining 34%.

We revealed a decreasing host specificity in leaf-associated bacteria with increasing latitudes (Figure 4a
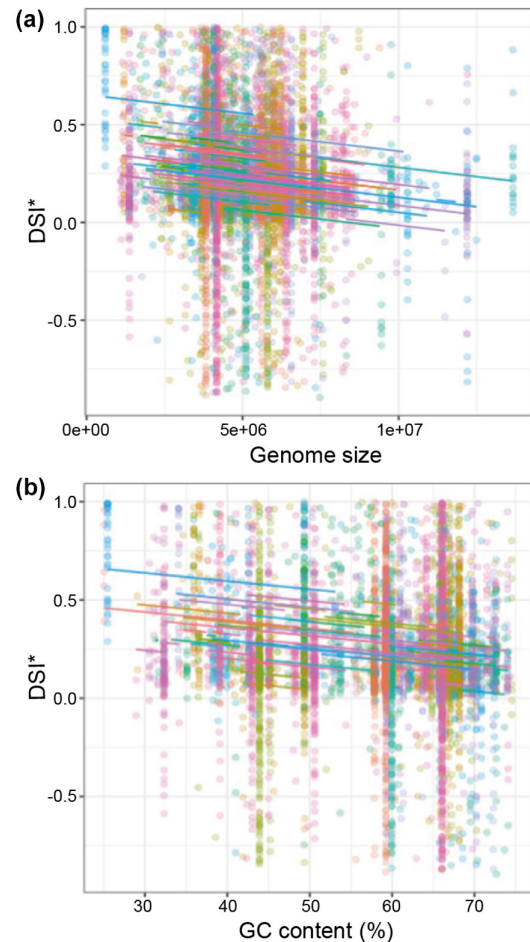
FIGURE 2 The correlations between (a) bacterial genome size, (b) GC content and the host specificity index DSI*. These relationships were fitted using a linear mixed-effect model with bacterial family taken as a random term (bacterial families were indicated by different colours and lines). 8822 ASVs from 140 bacterial families and 7652 ASVs from 126 bacterial families were included in (a) and (b), respectively. Both bacterial genome size and GC content were negatively correlated with DSI* ($p < 0.001$, see Figure S1 for the effect size). The family annotations were not shown in the legend due to the large number of bacteria families.

and Figure S2 for global and plot-level DSI*, respectively), consistent with the expectation of hypothesis 3. Conversely, bacterial genome size and host abundance increased with latitude (Figure 4b,c), in line with the prevalence of small-genome bacteria and rare plant species in tropical forests. Notably, we found negative effect of bacterial genome size and host abundance on the host specificity of leaf bacteria, even after accounting for the effect of latitude, study site and bacterial taxa (Table S4).

DISCUSSION

Our study addressed fundamental hypotheses regarding the mechanisms underlying the host specialization of plant–bacteria symbiosis, by a comprehensive assessment of the host specificity in leaf-associated bacteria along a

TABLE 1 Coefficients of bacterial traits related to the host specificity (global DSI*) in a phylogenetic generalized least squares model.

	Estimate	Std.Error	T-value	p-Value
Intercept	0.113	0.031	3.703	<0.001***
Cell length	−0.014	0.005	−2.517	0.012*
GC content	−0.014	0.005	−2.759	0.006**
Genome size	−0.039	0.006	−6.073	<0.001***
Gram-positive (vs. negative)	0.046	0.011	4.047	<0.001***
Motility-Flagella (vs. non)	0.0001	0.012	0.008	0.994
Motility (vs. non)	0.066	0.009	7.109	<0.001***
tRNA genes	0.048	0.005	10.300	<0.001***

Note: Bacterial phylogenies were incorporated as a random effect to account for the phylogenetic conservatism of host specificity. 6567 ASVs with all the traits available were included in the model. Significance levels: *$p<0.05$, **$p<0.01$, ***$p<0.001$.

TABLE 2 Coefficients of factors that influence the host specificity (global DSI*) and community uniqueness (LCBD index) of bacteria in linear mixed-effect models.

|	Estimate	Std.Error	df	t_value	Pr(>|t|)
Modelling averaged DSI*					
Intercept	−0.013	0.019	103.6	−0.678	0.499
Host species abundance	−0.002	0.001	744.3	−2.477	0.0146*
SLA	0.055	0.015	354.7	3.645	<0.001***
Annual temperature	0.093	0.019	23.76	4.952	<0.001***
SoilTN	0.082	0.027	111.4	3.104	0.002**
SoilAN	−0.071	0.026	16.32	−2.656	0.017*
SoilOrgan	−0.066	0.021	350.1	−3.068	0.002**
Modelling LCBD					
Intercept	1.418	0.509	26.62	2.785	0.010**
Host species abundance	−0.082	0.012	784.0	−6.617	<0.001***
Elevation	−1.204	0.352	48.64	−3.451	0.001**
SoilTP	−0.914	0.313	1016	−2.770	0.006**
SoilAP	0.511	0.248	1138	2.021	0.044*
Annual temperature	−1.388	0.572	11.69	−2.430	0.032*
SoilAK	0.804	0.233	1124	3.454	0.001***
Soilwater	0.743	0.235	831.6	3.007	0.003**
SoilpH	0.972	0.391	255.0	2.455	0.015*

Note: Study site and plant species identity were treated as random terms. 1253 samples with available SLA data were included in the models.

Abbreviations: SLA, specific leaf area; SoilTN, AN, soil total and available nitrogen; SoilTP, AP, soil total and available phosphorus; soil AK, soil available potassium; SoilOrgan, soil organic matters. Significance levels: *$p<0.05$, **$p<0.01$, ***$p<0.001$.

latitudinal gradient. We revealed negative effects of genome size and host species abundance on the host specificity of leaf bacteria, consistent with the expectation that host specialization is accompanied by genome streamlining and is influenced by host abundance (Klasson & Andersson, 2004; Lajoie & Parfrey, 2022). Additionally, we demonstrated an increased host specificity at lower latitudes, in line with the decreased bacterial genome size and increased prevalence of rare host species in the tropics (Hutchinson, 1959; Ngugi et al., 2023; Sorensen et al., 2019). These findings highlighted the interplay among host specificity, genome streamlining and host plant abundance in driving the latitudinal diversity gradient of plant–bacteria symbioses.

Genome streamlining in host specialization

Our findings of reduced genome size and genomic GC content with increased host specificity provided empirical evidence for the hypothesis that host specialization is associated with microbial genome streamlining (Klasson & Andersson, 2004; Lajoie & Parfrey, 2022). While much evidence for genome streamlining has been derived from experimental studies with cultured cells, our finding extended this concept to many unculturable bacteria living in natural environments (Giovannoni et al., 2014; McCutcheon & Moran, 2012). The negative correlations between genome size and host specificity may reflect a bidirectional influence between
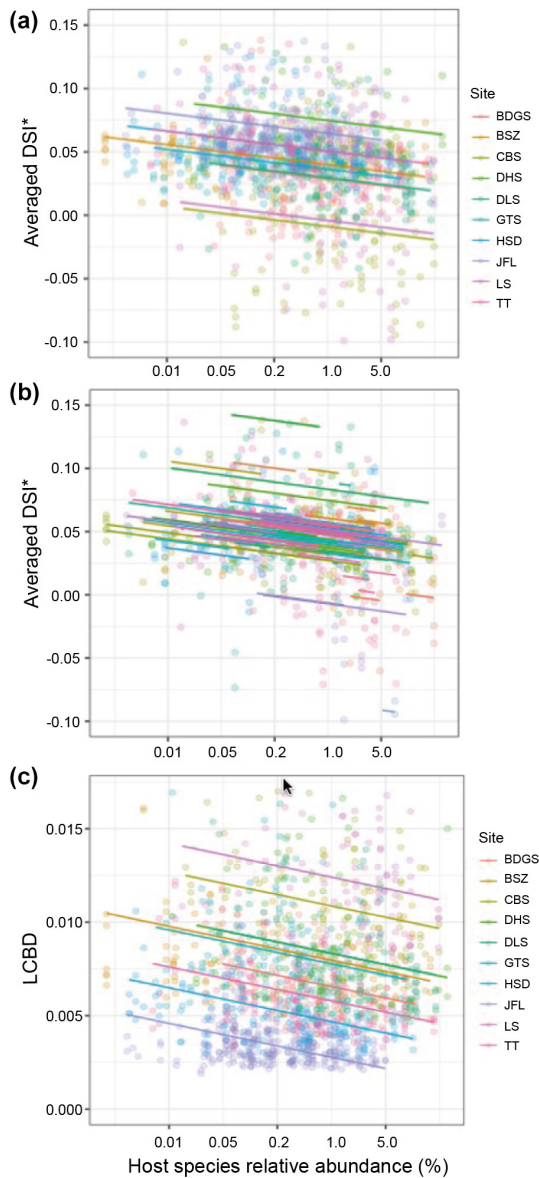
FIGURE 3 The correlations (a) between host plant abundance and host specificity (averaged DSI*) of bacteria within study plots, (b) across study plots and (c) between host plant abundance and community uniqueness (LCBD) of bacteria within study plots. Host plant abundance was calculated as the percentage of plant individuals in the forest plot. In (a), the relationship was fitted using a linear mixed-effect model with study plots as a random term, where each line indicated the change of host specificity in bacteria from rarer to common host species within each plot. In (b), the relationship was fitted with plant species identity as a random term, and each line represented a plant species, indicating the change of host specificity of bacteria when the same host species showed different abundance across study plots. Plant species names are not shown in the legend of (b) due to the large number of species. In (c), the relationships were fitted with study plot as a random term, indicating how the uniqueness of bacterial composition changed with host abundance. All these correlations were significantly negative ($p < 0.05$, also see Table 2).

genome evolution and specialization. For example, specialization could provide a stable environment conducive to genome streamlining, favouring the essential
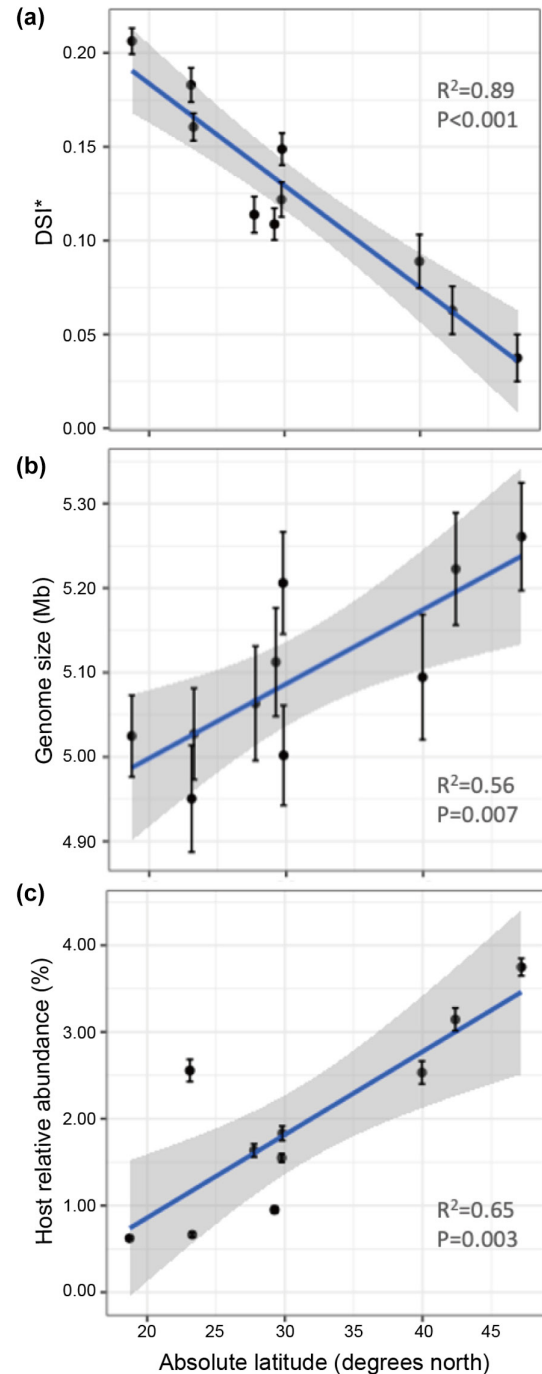
FIGURE 4 The latitudinal gradients of averaged (a) host specificity (global DSI*), (b) genome size and (c) host species abundance of phyllosphere bacterial ASVs. Mean value of the index and the 95% confidence interval of mean were shown for each study site. R^2 and p-value of the linear regression were reported.

functions that suit a specific habitat but imposing little selection pressure on the non-essential ones (Baroncelli et al., 2016; Klasson & Andersson, 2004). In turn, genome reduction enhanced bacterial fitness within their specialized environment while inhibiting colonization of alternative habitats, thereby reinforcing host specialization (Caley & Munday, 2003). The observed relationship between genome size and host specificity established a

foundation for future studies to test the causation between genome evolution and host specialization in laboratory experiments.

However, the broad applicability of our findings was partly limited by the precision and completeness of bacteria trait data. While the phylogenetic conservatism of bacterial traits such as genomic attributes allowed for the prediction of bacterial traits based on their taxonomic classification, there may be intraspecific trait plasticity and genetic variation at lower taxonomic levels that we were unable to capture in our data. Additionally, most bacterial ASVs were taxonomically identified and assigned trait values; however, there remains a proportion of ASVs that were either unidentified or lacked available trait data. Furthermore, our findings demonstrated correlations between bacterial genomic traits and host specialization at a coarse scale, while understanding how these correlations reflect genome evolution in host specialization at finer scales, that is, within species, will require future investigation (Li et al., 2021; Vega, 2019). The increasing availability of bacteria genomic data makes it possible to examine the genomic change of bacterial populations during their adaptation to different host ranges, for example, via comparative genomic analysis (Sheppard et al., 2018). This will provide an opportunity to unravel the genetic mechanisms underlying the evolution of host specialization (Ellegaard et al., 2019; Kwong et al., 2014).

Effect of host abundance and abiotic environment on the host specificity

Understanding the evolution of specialization on rare versus common resources is a key question in ecology. A prevailing theory in explaining the specialization of trophic interactions suggested that the cost for specialization is high, and therefore specialists should favour common and 'reliable' host resources in order to evolutionarily persist (Levins, 1968). However, the cost–benefit theory does not necessarily apply to bacteria which have high abundance, short generation time and rapid evolution. This is particularly so given that bacteria can quickly adapt to variable environments by mutation and horizontal gene transfer, expanding their functional and genetic repertoires (Kwong & Moran, 2015). As a result, the environmental variability that bacteria are exposed to may play a key role in driving the evolution of specialization. For plant-associated bacteria, neighbouring plant species provide new habitats for bacteria to colonize, which widens the habitat in which bacterial adaptation occurs (Meyer et al., 2022, 2023). Common plant species are more widespread and can neighbour with many other plant species. This provides ample opportunities for the bacteria associated with common host species to interact with diverse potential hosts (Miller et al., 2018). The evolutionary adaptation of bacteria to various hosts could weaken host specialization, leading to a decreased host specificity on common host species.

In addition to the influence of host abundance, the degree of host specificity increased with plant-specific leaf area and annual temperature. One of the explanations for this correlation is that warm temperatures and higher specific leaf area select for host-specialized bacteria (Cobian et al., 2019; Lajoie & Kembel, 2021). However, it could also be caused by other latitude-related factors that were not included in our study. For example, plant chemical traits such as leaf-released phenolics could play a critical role in influencing the composition and specialization of plant-associated bacteria (Bhattacharya et al., 2010). We suggest plant chemicals be worthwhile to consider in future study to understand how plant traits drive the host specificity of microbial symbionts. Additionally, there is significant effect of abiotic environments on the compositional uniqueness of phyllosphere bacteria, which is consistent with the established role of selection in shaping the community assembly of leaf bacteria (Vorholt, 2012). Incorporation of these effects in the model did not change the relationship between host abundance and host specificity, suggesting that the effect of host abundance on the host specificity of bacterial symbiosis is independent of the effect of plant traits, abiotic environments and study sites.

Latitudinal gradient of host specificity in plant leaf-associated bacteria

Our finding of decreasing latitudinal gradient of host specificity in leaf-associated bacteria supports the long-standing hypothesis of ecological specialization, which predicts a higher degree of host specialization in lower latitudes as an explanation of the remarkable biodiversity seen in the tropics (Hutchinson, 1959). While this hypothesis has been studied extensively in plant–symbiont interactions such as insect herbivores and pollinators (Dyer et al., 2007; Novotny et al., 2006), our study is among the first test of this hypothesis in plant–microbial interactions. It is well established that host specialization plays a key role in driving microbial diversity and also in maintaining plant species diversity (Klironomos, 2002; Schroeder et al., 2020). Therefore, the greater prevalence of host-specialized plant-associated bacteria at lower latitudes may contribute to the high diversity of plants and bacteria in tropical forests.

We demonstrated that the genome size of leaf bacteria increased with latitude, aligning with previously documented latitudinal gradients of genome size in soil and ocean microbes (Ngugi et al., 2023; Sorensen et al., 2019). Additionally, the averaged host abundance of leaf bacteria increased with latitude, which can be attributed to the larger proportion of rare plant species in tropical than temperate forests (Enquist et al., 2019). Intriguingly, the reduced genome size and host abundance in tropics were

associated with lower host specificity, even after considering the influence of latitude, study sites and bacterial taxa. This result implies there are intrinsic correlations among genome size, host abundance and host specificity across latitudinal gradients.

The increased host specificity of phyllosphere bacteria in lower latitudes can be partly attributed to the fact that many rare bacterial ASVs in tropical regions are only present in one or a few sites while ASVs in temperate regions are less unique to sites. For instance, 3.3%–19.9% of the ASVs in tropical regions occurred in only one site, compared to 1.3%–5.9% of ASVs in temperate regions. As a result, the global DSI* index, which consider both geographic distribution and host range of bacterial ASVs, exhibited a decreasing latitudinal gradient of host specificity. However, even after controlling the influence of geographic distribution (i.e. the plot-level DSI* index), there was still a decreasing latitudinal gradient of host specificity, suggesting that bacterial ASVs are more host-specialized in lower latitudes at both t local and larger spatial scales.

Conclusion

Our study presented evidence supporting three fundamental predictions regarding the mechanisms of host specialization in plant–bacteria symbiosis. The reduced genome size and GC content in host-specialized leaf bacteria aligned with the prediction of genomic streamlining theory, suggesting that genome size reduction is crucial in the evolution of host specialization (Klasson & Andersson, 2004; McCutcheon & Moran, 2012). Moreover, bacteria associated with common host species display greater ubiquity and lower level of host specialization than those associated with rarer host species. This implies that host abundance can influence host specialization, possibly through the facilitation of bacterial dispersal among diverse plant species (Lajoie & Parfrey, 2022). Remarkably, we revealed that the degree of host specificity in leaf bacteria decreased with increasing latitude, congruent with the observed rise in bacterial genome size and host abundance at higher latitudes (Hutchinson, 1959). These findings suggest that genome size and host abundance play a key role in the evolution of host specialization in plant-associated bacteria along a latitudinal gradient.

AUTHOR CONTRIBUTIONS
ZW, FH and SWK conceived the project; ZW conducted field sampling and phyllosphere data collection; YJ, MZ, CC, YC, SF, GJ, MJ, J-YL, YL, YL, KM, XM, XQ, XW, XW, HX, WY, LZ and YZ collected and collated the metadata and assisted with fieldwork; ZW analysed data with support from SWK and GL; ZW drafted the manuscript with input from SWK, GL and FH. All authors contributed to editing of the manuscript.

CONFLICT OF INTEREST STATEMENT
The authors declare that they have no conflict of interest.

PEER REVIEW
The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ele.14447.

DATA AVAILABILITY STATEMENT
Sequence data of this study were deposited in the Sequence Read Archive at https://www.ncbi.nlm.nih.gov/sra/PRJNA1001021, with accession number: PRJNA1001021. All data, code and metadata supporting the finding of this study are publicly available at https://doi.org/10.6084/m9.figshare.25727091.v1.

ORCID
Zihui Wang https://orcid.org/0000-0002-2512-7361
Yuan Jiang https://orcid.org/0000-0001-8704-3676
Minhua Zhang https://orcid.org/0000-0001-7184-7952
Chengjin Chu https://orcid.org/0000-0002-0606-449X
Guangze Jin https://orcid.org/0000-0002-9852-0965
Yu Liu https://orcid.org/0000-0001-9869-2735
Keping Ma https://orcid.org/0000-0001-9112-5340
Xiujuan Qiao https://orcid.org/0000-0003-4647-399X
Xugao Wang https://orcid.org/0000-0003-1207-8852
Fangliang He https://orcid.org/0000-0003-0774-4849
Steven W. Kembel https://orcid.org/0000-0001-5224-0952

REFERENCES
Arnold, B.J., Huang, I.-T. & Hanage, W.P. (2022) Horizontal gene transfer and adaptive evolution in bacteria. *Nature Reviews Microbiology*, 20, 206–218.

Baroncelli, R., Amby, D.B., Zapparata, A., Sarrocco, S., Vannacci, G., Floch, G.L. et al. (2016) Gene family expansions and contractions are associated with host range in plant pathogens of the genus Colletotrichum. *BMC Genomics*, 17, 555.

Baskett, C.A. & Schemske, D.W. (2018) Latitudinal patterns of herbivore pressure in a temperate herb support the biotic interactions hypothesis. *Ecology Letters*, 21, 578–587.

Benítez, M.-S., Hersh, M.H., Vilgalys, R. & Clark, J.S. (2013) Pathogen regulation of plant diversity via effective specialization. *Trends in Ecology & Evolution*, 28, 705–711.

Bever, J.D., Dickie, I.A., Facelli, E., Facelli, J.M., Klironomos, J., Moora, M. et al. (2010) Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology & Evolution*, 25, 468–478.

Bhattacharya, A., Sood, P. & Citovsky, V. (2010) The roles of plant phenolics in defence and communication during Agrobacterium and Rhizobium infection. *Molecular Plant Pathology*, 11, 705–719.

Brochet, S., Quinn, A., Mars, R.A., Neuschwander, N., Sauer, U. & Engel, P. (2021) Niche partitioning facilitates coexistence of closely related honey bee gut bacteria. *eLife*, 10, e68583.

Caley, M.J. & Munday, P.L. (2003) Growth trades off with habitat specialization. *Proceedings of the Royal Society B: Biological Sciences*, 270, S175–S177.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13, 581–583.

Cobian, G.M., Egan, C.P. & Amend, A.S. (2019) Plant–microbe specificity varies as a function of elevation. *The ISME Journal*, 13, 2778–2788.

Cornelissen, J.H.C., Lavorel, S., Garnier, E., Díaz, S., Buchmann, N., Gurvich, D.E. et al. (2003) A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 51, 335–380.

Dyer, L.A., Singer, M.S., Lill, J.T., Stireman, J.O., Gentry, G.L., Marquis, R.J. et al. (2007) Host specificity of Lepidoptera in tropical and temperate forests. *Nature*, 448, 696–699.

Ellegaard, K.M., Brochet, S., Bonilla-Rosso, G., Emery, O., Glover, N., Hadadi, N. et al. (2019) Genomic changes underlying host specialization in the bee gut symbiont Lactobacillus Firm5. *Molecular Ecology*, 28, 2224–2237.

Enquist, B.J., Feng, X., Boyle, B., Maitner, B., Newman, E.A., Jørgensen, P.M. et al. (2019) The commonness of rarity: global and future distribution of rarity across land plants. *Science Advances*, 5, eaaz0414.

Fürnkranz, M., Wanek, W., Richter, A., Abell, G., Rasche, F. & Sessitsch, A. (2008) Nitrogen fixation by phyllosphere bacteria associated with higher plants and their colonizing epiphytes of a tropical lowland rainforest of Costa Rica. *The ISME Journal*, 2, 561–570.

Gilbert, G.S. & Webb, C.O. (2007) Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 4979–4983.

Giovannoni, S.J., Thrash, J.C. & Temperton, B. (2014) Implications of streamlining theory for microbial ecology. *The ISME Journal*, 8, 1553–1565.

Grieneisen, L.E., Livermore, J., Alberts, S., Tung, J. & Archie, E.A. (2017) Group living and male dispersal predict the core gut microbiome in wild baboons. *Integrative and Comparative Biology*, 57, 770–785.

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, C.J.W. & Heuvelink, G.B.M. (2009) Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. *Computational Geosciences*, 35, 1711–1721.

Hutchinson, G.E. (1959) Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist*, 93, 145–159.

Jin, Y. & Qian, H. (2019) V.PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography*, 42, 1353–1359.

Jorge, L.R., Novotny, V., Segar, S.T., Weiblen, G.D., Miller, S.E., Basset, Y. et al. (2017) Phylogenetic trophic specialization: a robust comparison of herbivorous guilds. *Oecologia*, 185, 551–559.

Jorge, L.R., Prado, P.I., Almeida-Neto, M. & Lewinsohn, T.M. (2014) An integrated framework to improve the concept of resource specialisation. *Ecology Letters*, 17, 1341–1350.

Kembel, S.W., O'Connor, T.K., Arnold, H.K., Hubbell, S.P., Wright, S.J. & Green, J.L. (2014) Relationships between phyllosphere bacterial communities and plant functional traits in a neotropical forest. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 13715–13720.

Klasson, L. & Andersson, S.G.E. (2004) Evolution of minimal-genesets in host-dependent bacteria. *Trends in Microbiology*, 12, 37–43.

Klironomos, J.N. (2002) Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature*, 417, 67–70.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2017) lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82, 1–26.

Kwong, W.K., Engel, P., Koch, H. & Moran, N.A. (2014) Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proceedings of the National Academy of Sciences*, 111, 11509–11514.

Kwong, W.K. & Moran, N.A. (2015) Evolution of host specialization in gut microbes: the bee gut as a model. *Gut Microbes*, 6, 214–220.

Laforest-Lapointe, I., Paquette, A., Messier, C. & Kembel, S.W. (2017) Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. *Nature*, 546, 145–147.

Lajoie, G. & Kembel, S.W. (2021) Host neighborhood shapes bacterial community assembly and specialization on tree species across a latitudinal gradient. *Ecological Monographs*, 91, e01443.

Lajoie, G. & Parfrey, L.W. (2022) Beyond specialization: re-examining routes of host influence on symbiont evolution. *Trends in Ecology & Evolution*, 37, 590–598.

Legendre, P. (2014) Replacement and richness difference components. *Global Ecology and Biogeography*, 23, 1324–1334.

Legendre, P. & Cáceres, M. (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, 16, 951–963.

Levins, R. (1968) *Evolution in changing environments*. Princeton: Princeton University Press.

Li, E., Ryo, M., Kowalchuk, G.A., Bakker, P.A.H.M. & Jousset, A. (2021) Rapid evolution of trait correlation networks during bacterial adaptation to the rhizosphere. *Evolution: International Journal of Organic Evolution*, 75, 1218–1229.

Lindow, S.E. & Leveau, J.H.J. (2002) Phyllosphere microbiology. *Current Opinion in Biotechnology*, 13, 238–243.

Liu, H., Brettell, L.E. & Singh, B. (2020) Linking the phyllosphere microbiome to plant health. *Trends in Plant Science*, 25, 841–844.

Madin, J.S., Nielsen, D.A., Brbic, M., Corkrey, R., Danko, D., Edwards, K. et al. (2020) A synthesis of bacterial and archaeal phenotypic trait data. *Scientific Data*, 7, 170.

Mariadassou, M., Pichon, S. & Ebert, D. (2015) Microbial ecosystems are dominated by specialist taxa. *Ecology Letters*, 18, 974–982.

Martinez-Gutierrez, C.A. & Aylward, F.O. (2022) Genome size distributions in bacteria and archaea are strongly linked to evolutionary history at broad phylogenetic scales. *PLoS Genetics*, 18, e1010220.

McCutcheon, J.P. & Moran, N.A. (2012) Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology*, 10, 13–26.

Mestre, A., Poulin, R. & Hortal, J. (2020) A niche perspective on the range expansion of symbionts. *Biological Reviews*, 95, 491–516.

Meyer, K.M., Muscettola, I.E., Vasconcelos, A.L.S., Sherman, J.K., Metcalf, C.J.E., Lindow, S.E. et al. (2023) Conspecific versus heterospecific transmission shapes host specialization of the phyllosphere microbiome. *Cell Host & Microbe*, 31, 2067–2079.e5.

Meyer, K.M., Porch, R., Muscettola, I.E., Vasconcelos, A.L.S., Sherman, J.K., Metcalf, C.J.E. et al. (2022) Plant neighborhood shapes diversity and reduces interspecific variation of the phyllosphere microbiome. *The ISME Journal*, 16, 1376–1387.

Miller, E.T., Svanbäck, R. & Bohannan, B.J.M. (2018) Microbiomes as metacommunities: understanding host-associated microbes through metacommunity ecology. *Trends in Ecology & Evolution*, 33, 926–935.

Ngugi, D.K., Acinas, S.G., Sánchez, P., Gasol, J.M., Agusti, S., Karl, D.M. et al. (2023) Abiotic selection of microbial genome size in the global ocean. *Nature Communications*, 14, 1384.

Novotny, V., Drozd, P., Miller, S.E., Kulfan, M., Janda, M., Basset, Y. et al. (2006) Why are there so many species of herbivorous insects in tropical rainforests? *Science*, 313, 1115–1118.

Oh, P.L., Benson, A.K., Peterson, D.A., Patil, P.B., Moriyama, E.N., Roos, S. et al. (2010) Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *The ISME Journal*, 4, 377–387.

Ollerton, J. & Cranmer, L. (2002) Latitudinal trends in plant–pollinator interactions: are tropical plants more specialised? *Oikos*, 98, 340–350.

Perofsky, A.C., Meyers, L.A., Abondano, L.A., Fiore, A.D. & Lewis, R.J. (2021) Social groups constrain the spatiotemporal dynamics of wild sifaka gut microbiomes. *Molecular Ecology*, 30, 6759–6775.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2009) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*, 5, e9490.

Qin, Q.-L., Li, Y., Sun, L.-L., Wang, Z.-B., Wang, S., Chen, X.-L. et al. (2019) Trophic specialization results in genomic reduction in free-living marine Idiomarina bacteria. *MBio*, 10, e02545-18.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41, D590–D596.

Rastogi, G., Tech, J.J., Coaker, G.L. & Leveau, J.H.J. (2010) A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *Journal of Microbiological Methods*, 83, 127–132.

Sabath, N., Ferrada, E., Barve, A. & Wagner, A. (2013) Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic streamlining during thermal adaptation. *Genome Biology and Evolution*, 5, 966–977.

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009) Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology, Evolution, and Systematics*, 40, 245–269.

Schroeder, J.W., Dobson, A., Mangan, S.A., Petticord, D.F. & Herre, E.A. (2020) Mutualist and pathogen traits interact to affect plant community structure in a spatially explicit model. *Nature Communications*, 11, 2204.

Semchenko, M., Barry, K.E., Vries, F.T., Mommer, L., Moora, M. & Maciá-Vicente, J.G. (2022) Deciphering the role of specialist and generalist plant–microbial interactions as drivers of plant–soil feedback. *The New Phytologist*, 234, 1929–1944.

Sheppard, S.K., Guttman, D.S. & Fitzgerald, J.R. (2018) Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, 19, 549–565.

Shi, A., Fan, F. & Broach, J.R. (2021) Microbial adaptive evolution. *Journal of Industrial Microbiology & Biotechnology*, 49, kuab076.

Sieber, M., Traulsen, A., Schulenburg, H. & Douglas, A.E. (2021) On the evolutionary origins of host–microbe associations. *Proceedings of the National Academy of Sciences of the United States of America*, 118, e2016487118.

Sorensen, J.W., Dunivin, T.K., Tobin, T.C. & Shade, A. (2019) Ecological selection for small microbial genomes along a temperate-to-thermal soil gradient. *Nature Microbiology*, 4, 55–61.

Swan, B.K., Tupper, B., Sczyrba, A., Lauro, F.M., Martinez-Garcia, M., González, J.M. et al. (2013) Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences*, 110, 11463–11468.

Vacher, C., Hampe, A., Porté, A.J., Sauer, U., Compant, S. & Morris, C.E. (2016) The phyllosphere: microbial jungle at the plant–climate interface. *Annual Review of Ecology, Evolution, and Systematics*, 47, 1–24.

Vega, N.M. (2019) Experimental evolution reveals microbial traits for association with the host gut. *PLoS Biology*, 17, e3000129.

von Meijenfeldt, F.A.B., Hogeweg, P. & Dutilh, B.E. (2023) A social niche breadth score reveals niche range strategies of generalists and specialists. *Nature Ecology & Evolution*, 7, 768–781.

Vorholt, J.A. (2012) Microbial life in the phyllosphere. *Nature Reviews Microbiology*, 10, 828–840.

Wang, Z., Jiang, Y., Zhang, M., Chu, C., Chen, Y., Fang, S. et al. (2023) Diversity and biogeography of plant phyllosphere bacteria are governed by latitude-dependent mechanisms. *New Phytologist*, 240(4), 1534–1547.

Wright, E.S. (2016) Using DECIPHER v2.0 to analyze big biological sequence data in R. *R Journal*, 8, 352.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.