Loyola University Chicago
Loyola eCommons

Biology: Faculty Publications and Other Works

Faculty Publications and Other Works by Department

2022

Discordant Population Structure Among Rhizobium Divided Genomes and Their Legume Hosts

Alex B. Riley
University of Illinois at Urbana-Champaign

Michael A. Grillo
Loyola University Chicago, mgrillo1@luc.edu

Brendan Epstein
University of Minnesota, St Paul

Peter Tiffin
University of Minnesota, St Paul

Katy D. Heath
University of Illinois at Urbana-Champaign

Follow this and additional works at: https://ecommons.luc.edu/biology_facpubs

Part of the Biology Commons

Recommended Citation

Riley, Alex B.; Grillo, Michael A.; Epstein, Brendan; Tiffin, Peter; and Heath, Katy D.. Discordant Population Structure Among Rhizobium Divided Genomes and Their Legume Hosts. Molecular Ecology, 31, 24: , 2022. Retrieved from Loyola eCommons, Biology: Faculty Publications and Other Works, http://dx.doi.org/10.1111/mec.16704

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Biology: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
© 2022 The Authors. Molecular Ecology published by John Wiley & Sons Ltd.

Check for updates

Received: 3 September 2021 | Revised: 15 September 2022 | Accepted: 20 September 2022

DOI: 10.1111/mec.16704

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Discordant population structure among rhizobium divided genomes and their legume hosts

Alex B. Riley[1] | Michael A. Grillo[2] | Brendan Epstein[3] | Peter Tiffin[3] | Katy D. Heath[1,4]

[1]Department of Plant Biology, University of Illinois, Urbana, Illinois, USA

[2]Department of Biology, Loyola University Chicago, Chicago, Illinois, USA

[3]Department of Plant and Microbial Biology, University of Minnesota, St. Paul, Minnesota, USA

[4]Carl R. Woese Institute for Genomic Biology, University of Illinois, Urbana, Illinois, USA

Correspondence
Katy D. Heath, Department of Plant Biology, University of Illinois, Urbana, Illinois, USA.
Email: kheath@illinois.edu

Funding information
Directorate for Biological Sciences, Grant/Award Number: 1856744; Division of Integrative Organismal Systems, Grant/Award Number: 1401864 and 1645875; Joint Genome Institute, Grant/Award Number: CSP-503446

Handling Editor: Tatiana Giraud

Abstract

Symbiosis often occurs between partners with distinct life history characteristics and dispersal mechanisms. Many bacterial symbionts have genomes comprising multiple replicons with distinct rates of evolution and horizontal transmission. Such differences might drive differences in population structure between hosts and symbionts and among the elements of the divided genomes of bacterial symbionts. These differences might, in turn, shape the evolution of symbiotic interactions and bacterial evolution. Here we use whole genome resequencing of a hierarchically structured sample of 191 strains of *Sinorhizobium meliloti* collected from 21 locations in southern Europe to characterize population structures of this bacterial symbiont, which forms a root nodule symbiosis with the host plant *Medicago truncatula*. *S. meliloti* genomes showed high local (within-site) variation and little isolation by distance. This was particularly true for the two symbiosis elements, pSymA and pSymB, which have population structures that are similar to each other, but distinct from both the bacterial chromosome and the host plant. Given limited recombination on the chromosome, compared to the symbiosis elements, distinct population structures may result from differences in effective gene flow. Alternatively, positive or purifying selection, with little recombination, may explain distinct geographical patterns at the chromosome. Discordant population structure between hosts and symbionts indicates that geographically and genetically distinct host populations in different parts of the range might interact with genetically similar symbionts, potentially minimizing local specialization.

KEYWORDS
co-evolution, *Ensifer*, horizontal gene transfer, MGE, mutualism, plasmid

1 | INTRODUCTION

Host–microbe symbioses are ubiquitous in nature and underly crucial processes at every level of biological organization (Moran, 2007). Our understanding of the functions of microbial symbionts in natural systems, and how the information encoded by their genomes cascades up through hosts to affect communities and ecosystems, is growing rapidly (Pita et al., 2018; Rosenberg & Zilber-Rosenberg, 2018; Simon et al., 2019). Because many symbioses occur between species with distinct life histories, range sizes, mating systems and modes of dispersal, they are likely to have very different population structures (Pita et al., 2018; Revillini et al., 2016;

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Molecular Ecology* published by John Wiley & Sons Ltd.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Thrall et al., 2007). This is important because how hosts and symbionts influence each others' demography and evolution is likely to vary across the landscape and through time (Carlsson-Granér & Thrall, 2015; Fernandes et al., 2019; Laine, 2005; Tack et al., 2014; Thompson, 2005). The population structures of hosts and symbionts appear to vary widely, from largely congruent (Anderson et al., 2004; Caldera & Currie, 2012; Feurtey et al., 2016; Hartmann et al., 2020; Smith et al., 2008) to largely discordant, where one species exhibits substantially less population structure than its partner (Baums et al., 2014; Dybdahl & Lively, 1996; Harrison et al., 2017; Strobel et al., 2016).

To add to this complexity, bacterial symbionts are often host to mobile genetic elements (MGEs), which drive nonvertical inheritance and highly dynamic genomes of functional importance in symbiosis (Batstone, 2021; Drew et al., 2021; Porter et al., 2019; Wardell et al., 2022). Many bacterial genomes are even "divided," with nonchromosomal replicons, often referred to as megaplasmids, that vary in size and essentiality of the genes they contain (diCenzo & Finan, 2017; Hall et al., 2022; Harrison et al., 2010). Genes on these elements move between individuals independently of genes on other elements, potentially resulting in different parts of the genome having distinct patterns of population structure. Moreover, these genome elements can have widely differing rates of mutation and recombination, leading to higher rates of evolution in plasmids not required for bacterial survival (Cooper et al., 2010; Epstein et al., 2012, 2014; Epstein & Tiffin, 2021). Divided genomes are hypothesized to be an adaptation to distinct environments experienced by bacteria (e.g., rhizosphere vs. bulk soil), by allowing for the functional division of genes across elements and thus independent adaptation (diCenzo & Finan, 2017). This adaptive "divided genome" hypothesis contrasts with the "selfish operon" hypothesis, which proposes that operons (and by extension entire MGEs) act as their own evolutionary agents, with fitness interests separate from their bacterial hosts (Hall et al., 2022; Heath et al., 2022; Lawrence & Roth, 1996; Wardell et al., 2022). In either of these nonmutually exclusive scenarios, we expect that elements within a bacterial genome might differ in population structure from each other.

The legume–rhizobium mutualism is a tractable model system for studying microbial symbiosis across the landscape, due to the relative ease of isolating individuals and the availability of genetic resources for several legume–rhizobium partnerships (Kaneko et al., 2002; Yates et al., 2015). In this symbiosis, rhizobia form nodules on plant roots, wherein they fix atmospheric nitrogen in exchange for carbon derived from plant photosynthesis. However, rhizobia inhabit the soil as saprophytes when not in symbiosis with a host, thereby experiencing profoundly different environmental conditions and probably different selective pressures (Burghardt, 2019; Burghardt et al., 2019). In rhizobia, the canonical symbiosis genes (e.g., *nod* genes for signalling the host, *nif* and *fix* genes for fixing atmospheric nitrogen) are carried by MGEs (Wardell et al., 2022). These can be within integrative and conjugative elements (Greenlon et al., 2019; Hollowell et al., 2016; Porter et al., 2019; Sullivan et al., 1995; Weisberg et al., 2022). In rhizobia with divided genomes, however, symbiosis genes occur on separate, horizontally transmissible megaplasmids (Hall et al., 2022; Wardell et al., 2022). These pSyms have evolutionary histories, diversity levels and recombination rates that differ from the chromosome (Bailly et al., 2011; Cavassim et al., 2020; Epstein et al., 2012, 2014; Epstein & Tiffin, 2021; Wardell et al., 2022). Whether the mobility of these symbiosis plasmids results in distinct population genetic structures, which might affect adaptation to local environments in and out of host plants, is an unaddressed question in microbial evolution.

Metagenomic studies suggest that rhizobial taxa have characteristics associated with large geographical ranges and, by extension, large dispersal ability (Choudoir et al., 2018). However, the extent of geographical structure within these ranges, necessary for understanding local adaptation, co-evolution with hosts and other evolutionary processes, has not been characterized (Hoeksema & Forde, 2008; Hoetzinger et al., 2021; Kraemer & Boynton, 2017; Thompson, 2005; Whitaker & Banfield, 2006). Here we sequenced whole genomes of 191 isolates of the rhizobium *Sinorhizobium meliloti* (Becker et al., 2009; Kuzmanović et al., 2022; Tang et al., 2014) from 21 sites in southern Europe to test whether the chromosome and two symbiosis elements (megaplasmid pSymA and chromid pSymB) in this divided genome have congruent population genetic structures. We also reanalysed RADseq data from 192 *Medicago truncatula* plants (Grillo et al., 2016) to test whether patterns of population structure in the symbiont (at all three genome elements) match that of the host, which past population genetics research indicates has a pattern of isolation by distance (IBD) and genetic structure at both the population and regional levels (Bonhomme et al., 2015; Bonnin et al., 1996; Grillo et al., 2016; Ronfort et al., 2006; Siol et al., 2008).

2 | METHODS

2.1 | Study system

Sinorhizobium (formerly *Ensifer*) *meliloti* is a rhizobium species in the Alphaproteobacteria (Kuzmanović et al., 2022; Young & Haukka, 1996) that forms N-fixing nodules on the roots of multiple species in the genus *Medicago*, and is one of two *Sinorhizobium* species that forms root nodules on *Medicago truncatula* (Zribi et al., 2004). *M. truncatula* is a self-fertilizing annual native to the Mediterranean region of Europe (Bonnin et al., 1996; Siol et al., 2008). The genome of *S. meliloti* is ~6.79 Mb, divided into three major genomic elements: the chromosome (3.69 Mb), megaplasmid pSymA (1.41 Mb) and chromid pSymB (1.69 Mb), as well as smaller plasmids in some strains (Galibert et al., 2001; Nelson et al., 2018). Metabolic modelling, along with genetic manipulation of the *S. meliloti* genome, has shown that gene content differs functionally between the three genomic elements (diCenzo et al., 2014; Galibert et al., 2001). The chromosome carries primarily genes related to core metabolic function in soil, pSymA carries the majority of genes required for symbiotic N fixation and pSymB carries primarily genes important for life in rhizosphere environments (diCenzo et al., 2014).

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2 | Sample collection

We isolated *S. meliloti* strains from 21 sites where *M. truncatula* host plants are found, using a hierarchical sampling design: three populations were sampled within 337 km of each other in southern Spain, 14 populations were sampled from within 328 km of each other in southern France and four populations were sampled from within 91 km of each other in Corsica. Sampling locations in Spain and southern France were separated by a minimum of 619 km, between southern France and Corsica 457 km, and Spain and Corsica 968 km (Table S1). We also sampled *M. truncatula* plants from eight of these 21 sites (Grillo et al., 2016). At each site we collected soil surrounding the top six inches of the roots of multiple host plants. To avoid cross-contamination within each site, the sampling shovel was wiped clean of excess soil between samples and was pierced into the ground adjacent to a plant numerous times before sampling soil. Between sampling locations, the shovel was sterilized with dilute bleach. Soil samples were kept at 4°C prior to isolating cultures.

Sinorhizobium strains were isolated or "trapped" in the laboratory from the field samples following standard protocols (Heath, 2010; Vincent, 1970). In brief, *M. truncatula* seeds were nicked with a razor blade, surface sterilized with 30% bleach, rinsed with sterile water and imbibed in sterile water for ~30 min. Seeds were then directly sown into a given soil sample housed in a sterilized, fully self-contained Magenta box (Brown et al., 2020). Magenta boxes were randomly placed in a temperature-controlled growth room (23°C) with light set to 12-h days. After 4 weeks, plants were harvested, and soil was washed from the roots. Individual nodules were removed with forceps, surface sterilized by soaking in 30% bleach for 10 min and then rinsed with sterilized water. Surface-sterilized nodules were crushed with sterilized forceps and streaked on tryptone-yeast (TY) media plates. Plates were incubated at 30°C for 48 h, then colonies were streaked onto TY plates and again incubated at 30°C. Strains were replated until individual colonies could be isolated; isolates were grown in liquid TY media and these pure cultures were stored in 50% TY 50% glycerol at −80°C. Given variation among host genotypes in rhizobium infection rates (Batstone et al., 2017; Heath & Tiffin, 2009), we used 10 host genotypes planted in each soil to trap rhizobia in a common garden experiment and maximize strain variation from all populations (strains per host genotype in this analysis varied from four to 38). While host genotypes undoubtedly differ in the frequency of strains sampled from a mixed rhizobium population (and we cannot rigorously assess this here), our host genotypes were found with a diversity of strain lineages. Moreover, closely related pSymA clusters (see Section 3) were sampled by by six or seven host genotypes each (see Table S1 for details on final host–strain composition). Because both *S. meliloti* and *S. medicae* infect the roots of *M. truncatula*, we used a post-PCR (polymerase chain reaction) restriction enzyme (*Rsa*I) digestion of the 16S gene to assign strains to species (following Biondi et al., 2003), resulting in 199 putative *S. meliloti* strains sequenced.

2.3 | Sequencing

We extracted DNA from cultures of *S. meliloti* grown in liquid TY media using Qiagen DNeasy and sent samples to the DOE Joint Genome Institute (JGI) for sequencing. JGI prepared a paired-end library for each strain, and sequenced samples on an Illumina HiSeq-2500 1 TB platform (101-nt read length; Illumina). Of the 199 strains submitted to JGI, we received high-quality whole genome sequences for 166. We regrew the remaining 33 strains from frozen cultures (as above), and extracted DNA using the Zymo Quick-DNA kit for Fungi or Bacteria. These samples were sequenced (2×150 or paired-end 150 nt read length) on the Novaseq 6000 platform (Illumina) by the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign. We successfully recovered quality *S. meliloti* genome sequences from 25 of these 33 isolates, for a total of 191 *S. meliloti* strains analysed in this study.

2.4 | Genome assembly, annotation and SNP calling

To ensure high-quality single nucleotide polymorphism (SNP) calling and genome assembly, we trimmed PCR adaptors and removed PCR duplicates and PhiX contamination using HT-STREAM (github.com/s4hts/HTStream), followed by further adaptor removal, removal of bases with quality scores <30 from the ends of reads, and removal of reads <80 bp long with TRIMGALORE! (github.com/FelixKrueger/TrimGalore). We aligned reads to the *S. meliloti* reference genome USDA1106 using BWA with default settings (Li & Durbin, 2009). We then used FREEBAYES (Garrison & Marth, 2012) to identify haplotype variants, which we split into SNPs using VCFTOOLS (Danecek et al., 2011). We filtered SNPs to retain only those with depth values between 20 and 230; the minimum was set to ensure enough coverage to correctly call the allele, and the maximum was set to avoid inaccurate SNP calls due to duplications and other misalignments. We retained sites with minor allele frequencies ≥0.009 to eliminate singletons (except for diversity analyses, see below). Finally, we retained SNPs present in >80% of strains to capture loci present in most strains while allowing some presence–absence variation (Epstein et al., 2012). After applying these filters, 72,311 SNPs remained (chromosome: 34,689 SNPs, pSymA: 15,162 SNPs, pSymB: 22,460 SNPs; see Table S2).

For variable gene content, we assembled genomes de novo using SPADES (Bankevich et al., 2012) with default parameters, followed by annotation using PROKKA (Seemann, 2014) and scanned these annotations for presence–absence variants using default settings in ROARY (Page et al., 2015). ROARY indicated that the variable gene content did not saturate even in our sample of 191 strains, but that the core gene content stabilized quickly as strains were added to the analysis (Figure S1). Because our data came from Illumina sequencing runs from two sequencing facilities that differed in read length, we down-sampled the larger set (JGI) and reran ROARY on the two data sets separately; we found little evidence that the size of the core and

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

variable genomes were influenced by this difference in read length (Table S3).

2.5 | Population genomic analyses

We used POPGENOME (Pfeifer et al., 2014) implemented in R 4.0.2 with the full SNP data set (without minor allele frequency filter; see Table S2) to calculate nucleotide diversity (π) and Tajima's *D* (D_T) for our entire sample as well as within individual sampling sites, and site frequency spectra (SFS). To estimate π, we divided the total diversity by the length of the appropriate element after removing sites with FREEBAYES variant calls that had quality score <30, indel sites and sites with coverage from <80% of the strains (chromosome: 3,379,577, pSymA: 839,391, pSymB: 1,534,650). To explore the extent of linkage disequilibrium across the three elements, we used PLINK (Purcell et al., 2007) to calculate linkage disequilibrium (r^2) for all pairwise comparisons between sites on each element, then used a random subsample of 10,000 pairwise comparisons for visualization using GGPLOT2 (Wickham, 2016). To compare the distance between linked SNPs among elements, we visualized the distributions of inter-SNP distances for the subset of SNPs with $r^2 > .5$. To explore the population structure of the three genomic elements of *S. meliloti*, we first used principal components analysis (PCA). We used the glPca function in the ADEGENET (Jombart & Ahmed, 2011) library in R on a random subsample of 15,000 SNPs (to equilibrate data set size) for each element of the *S. meliloti* genome to naively cluster individuals by genome-wide similarity, and GGPLOT2 (Wickham, 2016) to plot the positions of individuals along the first three axes of variation.

We used the LEA package (Frichot et al., 2014; Frichot & François, 2015) implemented in R to further characterize population structure, using the SNP data for each element. We used non-negative matrix factorization algorithms via the sNMF function to naively cluster individuals into populations (Frichot et al., 2014), performing 10 repetitions for each number of clusters (*K*) ranging from 1 to 20. We calculated the cross-entropy criterion and performed visual inspection of plots to determine the value of *K* where the criterion began levelling off and after which additional groups were composed of admixed individuals.

To quantify genetic differences, we calculated individual and population-level D_{XY} (Nei, 1972) using the stamppNeisD function from the package STAMPP (Pembleton et al., 2013) on all variants for each element. To test whether genetic distances among *S. meliloti* strains were congruent across the three genome elements, we used these matrices of individual-based distance metrics in pairwise Mantel tests comparing the three *S. meliloti* genome elements (chromosome vs. pSymA, chromosome vs. pSymB, pSymA vs. pSymB). Mantel tests were implemented in the R package ADE4 (Dray & Dufour, 2007).

We used AMOVA (poppr.amova) with clone correction and 1000 random permutations (Kamvar et al., 2014) to partition the genetic variance among individuals, sites and regions, as well as to assess the probability of obtaining these levels of variance by chance. To test the hypothesis of IBD, we used Mantel tests of correlations between individual-level D_{XY} and geographical distances between sampling sites. For each of the three elements, we built phylogenies based on the same random subsample of 15,000 core genome SNPs used for PCA. We used the neighbour-joining (nj) function in the R package APE (Paradis & Schliep, 2018) with 1000 bootstrap replicates and visualized trees in FIGTREE (Rambaut, 2018).

To compare the spatial genetic structures of rhizobia and host plants, we re-analysed RAD-seq data from the 192 *M. truncatula* genotypes studied in Grillo et al. (2016). We first called SNPs using STACKS with default parameters (Catchen et al., 2013), then filtered the resulting variants using VCFTOOLS (Danecek et al., 2011) to ensure that all variants were present in at least 80% of genotypes, had minor allele frequencies >0.05 to minimize the effects of rare variants and were >5 kb apart given previous studies of linkage disequilibrium in *M. truncatula* (Branca et al., 2011; Grillo et al., 2016); 10,814 SNPs remained after these filters. We performed AMOVA, calculated individual and population-based D_{XY}, used LEA to perform population structure analyses and tested for IBD in *M. truncatula* as detailed above for *S. meliloti*. For the subset of eight sites for which we had both hosts and symbionts, we used Mantel tests based on the Pearson correlation coefficient (mantel.randtest in ADEGENET; Jombart & Ahmed, 2011) of whether the three matrices of pairwise population D_{XY} values from symbionts (chromosome, pSymA and pSymB) were correlated with that of the host, to evaluate congruent population genetic structure between hosts and symbionts.

To explore geographical structure in the variable genome (i.e., genes present in only some strains), we performed an additional PCA, as above, using the matrix of *S. meliloti* gene presence–absence variants (91,840 genes) and plotted the first three principal components for visual inspection. We also used AMOVA (as described above) on the matrix of presence–absence variants to quantify the spatial variation in the pangenome. To test for IBD in the pangenome, we used a Mantel test of the correlation between PC-based distances (PC1) and geographical distances between sampling sites.

3 | RESULTS

To characterize population structure, we sequenced the full genomes of 191 *Sinorhizobium meliloti* strains sampled from 21 locations (two to 21 strains per population; Tables S1 and S4) in Spain and France (mainland and Corsica). Across the entire sample, we detected 287,978 segregating sites: 149,061 on the chromosome, 49,861 on pSymA and 89,056 on pSymB. Pairwise diversity of segregating sites was higher on pSymA (π = 0.0027) than pSymB (π = 0.0021) or the chromosome (π = 0.0020). Sample wide, all three elements (but particularly the pSyms) harboured an excess of rare variants, relative to expectations under a standard neutral model (Tajima's *D* or D_T = −0.18, −0.97 and −1.18 for the chromosome, pSymA and pSymB respectively). This excess is also seen in SFS plots (Figure S2A), though the chromosomal distribution is bimodal and includes a high proportion of intermediate-frequency alleles.

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 AMOVA partitioning the genome-wide genetic variation for each of the three elements of the *Sinorhizobium meliloti* genome for 191 strains sampled from 21 populations from three regions (Spain, France or Corsica) in the native range of the symbiosis, compared to host plant *Medicago truncatula*

	Chromosome		pSymA		pSymB		Pangenome		*M. truncatula*	
	% variance	Phi	% variance	Phi	% variance	Phi	% variance	Phi	% variance	Phi
Among region	24.5	0.34**	7.7	0.27**	10.1	0.24***	7.2	0.20***	19.6	0.20***
Among population within region	9.7	0.13***	19.2	0.21***	13.8	0.15***	13.2	0.14***	36.5	0.45***
Within population	65.8	0.25***	73.1	0.08***	76.1	0.10***	79.6	0.07***	44	0.56***

Note: For each level of spatial division, the percentage variance explained, phi statistic and probability of that level of variance occurring by chance are given.

$^{**}p < .01$; $^{***}p < .001$.

This chromosomal pattern probably results from groups of closely related chromosomal genomes (see results below) and did not depend on the inclusion of strains from Corsica (Figure S2B). Individual populations (sampling sites) varied considerably in diversity, with even some lesser-sampled populations harbouring more diversity than the entire sample (e.g., chromosomal $\pi = 0.0025$ in site 17 with $N = 5$ strains; Table S4), and some well-sampled populations possessing little diversity (e.g., sites 3 and C3; Table S4). Average linkage disequilibrium among SNPs (Figure S3) was much higher overall and decayed less with distance on the chromosome (mean $r^2 = .14$, $SD = .27$, Figures S2 and S3A) than pSymA (mean $r^2 = .03$, $SD = .09$) or pSymB (mean $r^2 = .02$, $SD = .08$). The proportion of linked SNPs ($r^2 > .5$) was much higher on the chromosome (12.7%) compared to the plasmids (0.9% and 0.8% for pSymA and pSymB, respectively; Figure S3B), and the distance between those SNPs was larger on the chromosome (Figure S3C). These patterns are consistent with higher rates of recombination on the pSyms than the chromosome.

Our sample of 191 *S. meliloti* genomes revealed the chromosome and symbiosis elements of *S. meliloti* to have distinct population structures. For each of the elements, and gene content in the entire pangenome, the majority (66%–80%) of genetic variation was found within populations (Table 1). Consistent with the extensive diversity within each sampling location, PCA did not reveal strong among-region differentiation (core-genome, Figure 1a; pangenome, Figure S4), although differentiation was stronger for the chromosome than for symbiosis plasmids (Table 1; Figure 1). Despite the stronger differentiation of the chromosome, there were multiple instances of individuals sampled from separate locations having nearly identical chromosomal sequences (points overlap in the chromosome PCA; Figures 1a and S3).

The chromosome was also the only genomic element to exhibit significant IBD (Figure 1b), although there was extremely large variation in the genetic distances between chromosomes isolated from even the most distant locations (D_{XY} distances ranged from 0 to 0.5 even when strains were sampled from ~1350 km apart; Table 1; Figure 1b). The two symbiotic elements in *S. meliloti* (pSymA and pSymB) were less structured at the among-region and among-population scales (Table 1), and neither exhibited significant IBD (Figure 1b). Thus, although populations differed in genetic composition (significant structure in Table 1), *S. meliloti* strains from distant populations were often as closely related to each other as strains from the same sampling site. We found no evidence for IBD in the pangenome ($r = -.01$; $p = .54$).

The topology of the chromosomal tree showed five tightly clustered groups of sequences, with almost all strains from Corsica closely related and forming a distinct group (orange cluster; Figure 2a), and strains from Spain somewhat interspersed with those from France, mostly in the blue and yellow clusters. The diversity of strains from mainland France (the best sampled region) included representatives from each major chromosomal lineage, with two lineages (pink and purple clusters) found only in mainland France (except strain 710A from Corsica; Figure 2a; Table S1). Differentiation of the chromosome at the regional scale (i.e., among Spain, France

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

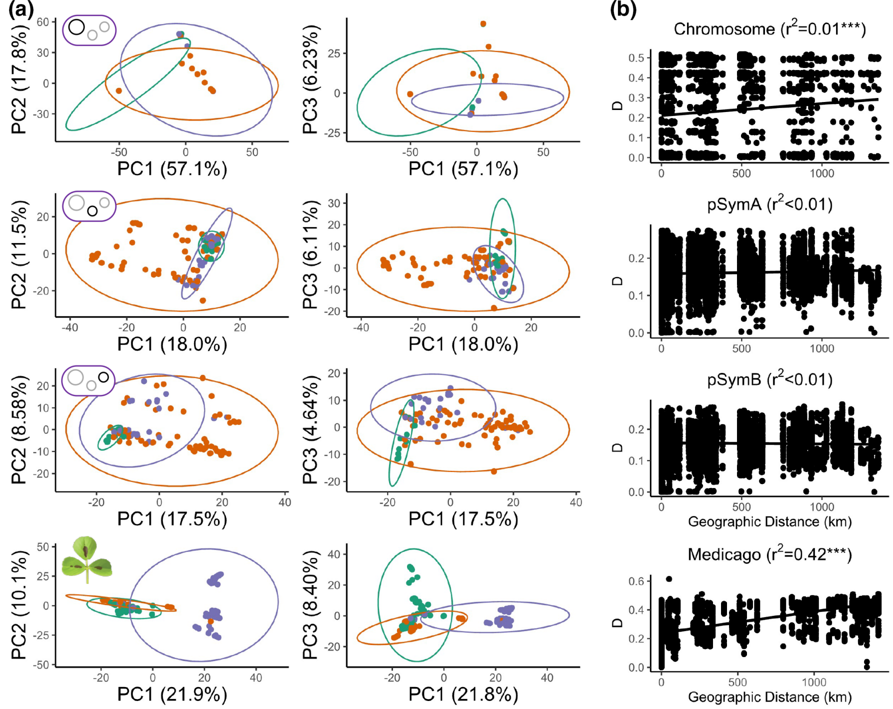


FIGURE 1 Genetic PCAs and isolation by distance in the tripartite symbiont genome and its host plant. (a) Principal component (PA) axis plots of genome-wide similarity and percentage variance explained by PCs 1–3 for 191 individuals of *Sinorhizobium meliloti* for the chromosome, pSymA and pSymB as well as for 192 *Medicago truncatula* host plants. The three ellipses and points are coloured by region: Spain (purple), mainland France (orange) and Corsica (green); many individual strains have nearly identical PCA scores and thus are stacked under a single point, particularly for the chromosome (top) PCA plot. (b) the relationship between geographical distance and individual genetic distance (D_{XY}) for each of the three genome elements in *S. meliloti* and the host plant (from top: Chromosome, pSymA, pSymB, *M. truncatula*). Shown are r^2 values from mantel tests for each comparison (***$p < .001$)

and Corsica; Table 1; Figure 1) is probably driven in large part by the strains sampled from Corsica, as short branch lengths were found among most Corsica strains on the chromosomal tree (within the orange cluster; Figure 2a), and genetic distances were high between Corsica populations and the rest of the range (Table S5).

We next tested whether the patterns of geographical structure were congruent across the three *S. meliloti* elements. Mantel tests of among-population genetic distances indicated tight congruence between pSymA and pSymB (Figure 3), but weak congruence between the chromosome and either pSymA or pSymB (Figure 3). Analyses in LEA indicated much shallower entropy plots (i.e., much less information gained with each additional cluster) and fewer clusters (best estimate of $K = 3$; Figures S5 and S6) for both pSymA and pSymB than for the chromosome ($K = 5$; Figures S5 and S6). Mapping these genetic groups onto sampling locations illustrates the incongruence of population genetic structure across the genomic elements; the chromosome (Figure 4) shows strains from Corsica falling into distinct clusters, whereas the largest genetic clusters for both pSymA and pSymB (green and purple, respectively) contained strains sampled from throughout the range. These results are largely consistent with the PCA results (Figure 1).

Annotating pSymA and pSymB trees with colours corresponding to the chromosomal clusters helps visualize how pSym lineages are interspersed across the chromosomal tree (coloured outer rings in Figure 2b,c), indicative of plasmid transfer (whole or partial) across chromosomal lineages. Some strains were closely related at all three elements of their genomes; for example, the chromosomal lineages in pink from the western coast of France appear together in all three trees (with the exception of strain 78), indicating little transfer of pSym genes among these and other lineages. On the other hand, the other chromosomal clusters (purple, orange and yellow clusters in Figure 2a) appear throughout the pSymA and pSymB tree (see same colours interspersed; Figure 2b,c), indicating that these chromosomal lineages are found with diverse pSym genotypes. For example, the tightly clustered group of closely related chromosomal lineages in orange (Figure 2a) were found with lineages from across both the pSymA and pSymB trees (Figure 2b,c).

Host plants were more strongly structured than any element of the symbiont genome, and their population structure was incongruent with the structure in *S. meliloti*. This was particularly true at the among-population and among-region scale, with only 44% of the variation found within populations (compared to 66%–80% in the symbiont; Table 1). PCA revealed regional differentiation, particularly between individuals from Spain (purple ellipse in Figure 1a) along PC1 (21.9% of genome-wide variance) as well as differentiation between individuals sampled from Corsica (green ellipse) along PC3 (8.4% of genome-wide variance). As in previous studies (Bonhomme et al., 2015; Grillo et al., 2016; Ronfort et al., 2006), we

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

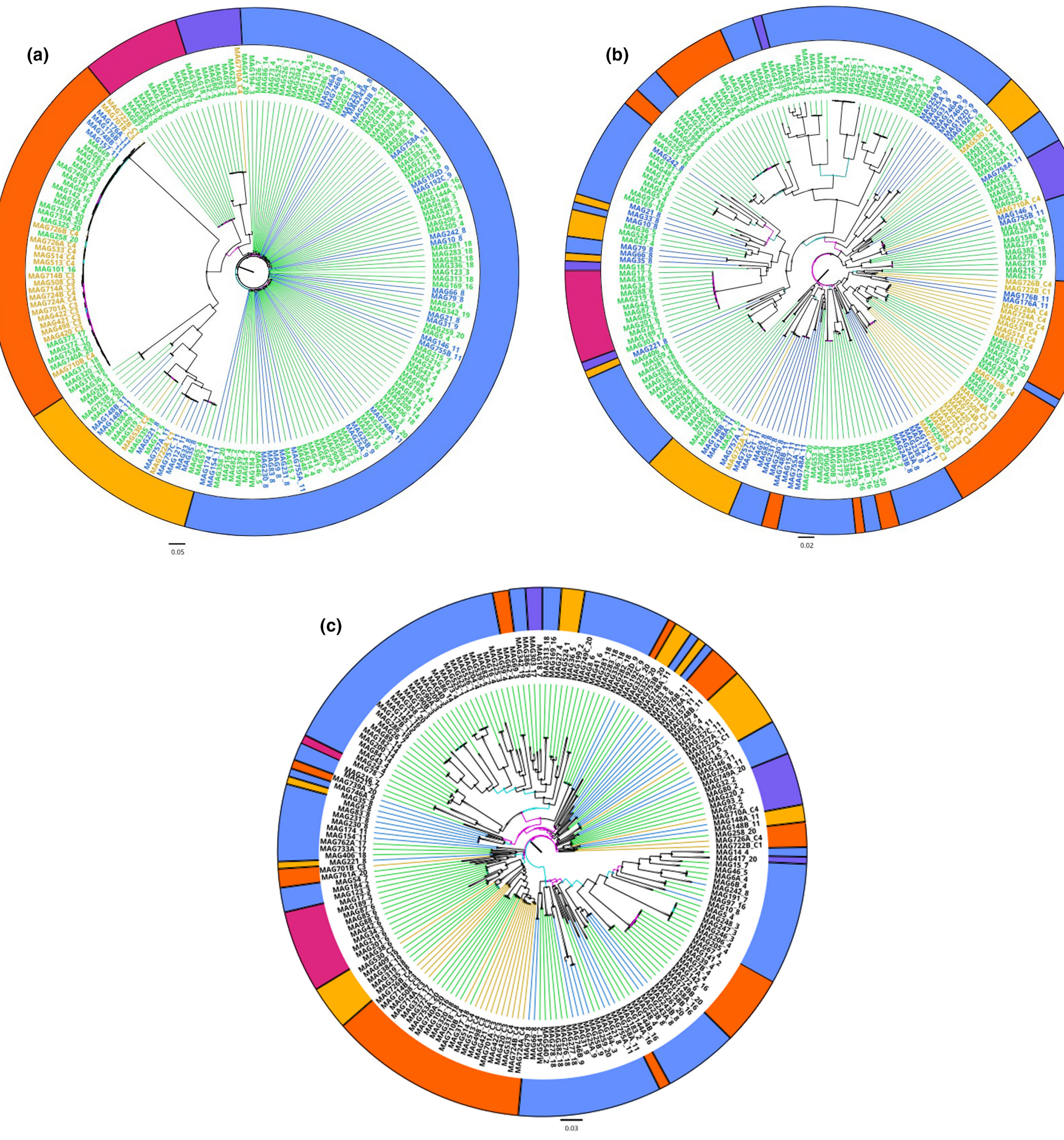
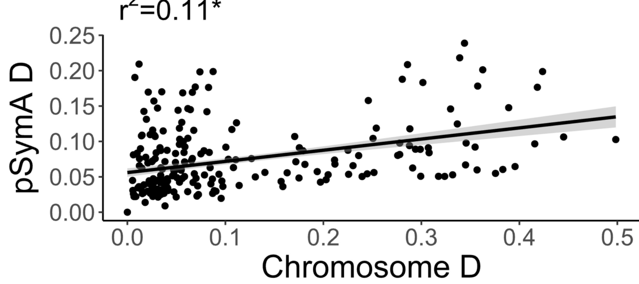
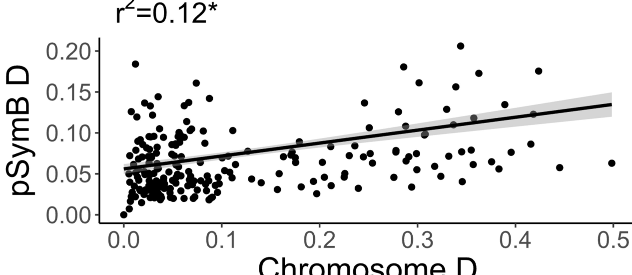


FIGURE 2 Neighbour-joining trees of all *Sinorhizobium meliloti* strains based on (a) chromosomal, (b) pSymA and (c) pSymB variant data. For all trees individual tip labels (strain ID followed by soil population) are coloured based on region of origin (orange from Corsica, green from mainland France, blue from Spain). The outer ring coloration in all trees represents the five major clusters of the chromosome (tree a), allowing for comparison across elements. Individual tip labels (strain ID followed by soil population) are coloured based on region of origin (orange from Corsica, green from mainland France, blue from Spain). Branch support is indicated in teal (<70) or pink (<50)

found a strong pattern of IBD for *Medicago truncatula* (Figure 1b), much stronger than for any element of the symbiont genome. While both the host plant and the symbiont chromosome exhibit significant IBD (though again the latter was weak), their spatial patterns of genetic variation were distinct. For the eight overlapping populations, we found no significant correlations between the among-population differentiation (D_{XY}) between *M. truncatula* hosts and any element in the *S. meliloti* genome (all $p > .05$; Figure S7). Host plants primarily clustered along axes that separated Spain from France and Corsica (Figures 1a and 4; Grillo et al., 2016, and references therein). By contrast, the *S. meliloti* chromosome was most clearly differentiated into mainland Europe versus Corsica (PC1, Figures 1a and 4; Table S5). The lack of IBD and little regional structure in pSymA and pSymB indicate that host plants at a given site, from all five *M. truncatula*

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License





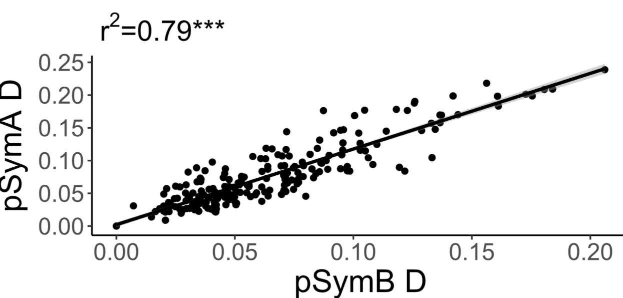


FIGURE 3 Pairwise comparisons of the matrices of among-population genetic distances, between the three genomic elements (chromosome, pSymA and pSymB) in 191 strains of the bacterial symbiont *Sinorhizobium meliloti* sampled from 21 populations in the native range. Each point represents a pairwise population mean D_{XY} value (genetic distance between two populations), with the mantel test of their correlation (*$p < .05$; ***$p < .001$)

genetic clusters, have the potential to interact with the diversity of symbiosis plasmids from across this part of the species range.

4 | DISCUSSION

The evolution of multipartite bacterial genomes will depend, at least in part, on whether the replicons share similar population structure. Similarly, the population structures of bacterial symbionts and their eukaryotic hosts have the potential to alter the outcomes of co-evolution by determining the spatial scale at which interactions occur, from local to global (Fernandes et al., 2019). Here we studied population structure in a hierarchically structured sample of rhizobial symbionts and found that: (i) the elements of the tripartite bacterial genome (chromosome and symbiosis elements pSymA and pSymB) have distinct population genetic structures at the local and regional scales, with more structure and IBD at the chromosome compared to the symbiotic elements; (ii) *Medicago truncatula* host plants are more strongly structured than *Sinorhizobium meliloti* symbionts, particularly when compared to the two symbiotic elements (pSymA and pSymB); and (iii) population structure between hosts and symbionts is not congruent across this part of the range.

4.1 | Spatial structure of the bacterial chromosome and symbiotic elements

Despite the importance of population structure in local adaptation and co-evolution (Hoeksema & Forde, 2008; Kraemer & Boynton, 2017; Thompson, 2005), the population structures of most microbes, including important symbiotic species, remain poorly understood (Chase et al., 2019; Hanson et al., 2012; Martiny et al., 2006; VanInsberghe et al., 2020). Genomic sequences of 191 strains of *S. meliloti* isolated from 21 geographical locations in Spain and France revealed that, like other microbial systems (Chase et al., 2019; Hoetzinger et al., 2021; Vos & Velicer, 2008), *S. meliloti* displays some population genetic structure; moreover, our element-specific analyses allow us to show how the extent and patterns of this structure vary among the replicons. The chromosome showed the most regional structure and was the only element showing evidence of IBD – at least at the scale at which we sampled. Despite the IBD, we found nearly identical chromosomes in strains sampled from up to 1350 km apart (e.g., strains in chromosomal clade 1, which are 99.85% identical and overlapping in PC space in Figure 1a), suggesting the possibility of long-distance dispersal. Previous work has found either abundant phenotypic variation within local populations (Heath, 2010; Heath & Stinchcombe, 2014; Heath & Tiffin, 2009; Pahua et al., 2018; Stoy et al., 2020) or abundant genomic diversity from a single soil sample (Bailly et al., 2006, 2011). Here we partition the genome-wide variation in space using a large sample of strains to demonstrate that the high proportion of local genomic diversity, relative to population structure, is a general characteristic of this symbiont.

We show that the patterns of population structure differed among the three elements of the *S. meliloti* genome; pSymA and pSymB had similar population structures, whereas both differed from the structure of the chromosome. Previous work lacking an explicit spatial component has shown that the three elements can have distinct evolutionary histories (Bailly et al., 2006; Epstein et al., 2014; Galardini et al., 2013; Nelson et al., 2018; Toro et al., 2016). Similarly, chromosomes and symbiosis genes often have differing evolutionary histories in other rhizobial taxa with symbiosis plasmids (Carrascal et al., 2016; Cavassim et al., 2020; Klinger et al., 2016; Koppell & Parker, 2012; Kumar et al., 2015; Young et al., 2006), as well as those with symbiosis integrative and conjugative elements (Hollowell et al., 2016; Porter et al., 2019; Weisberg et al., 2022).

What drives these patterns in time and space, and what are the possible implications of this incongruence? All three elements are components of the same organismal genome and thus disperse together and share vertical transmission, but pSymA and pSymB have much higher rates of recombination presumably due to plasmid

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

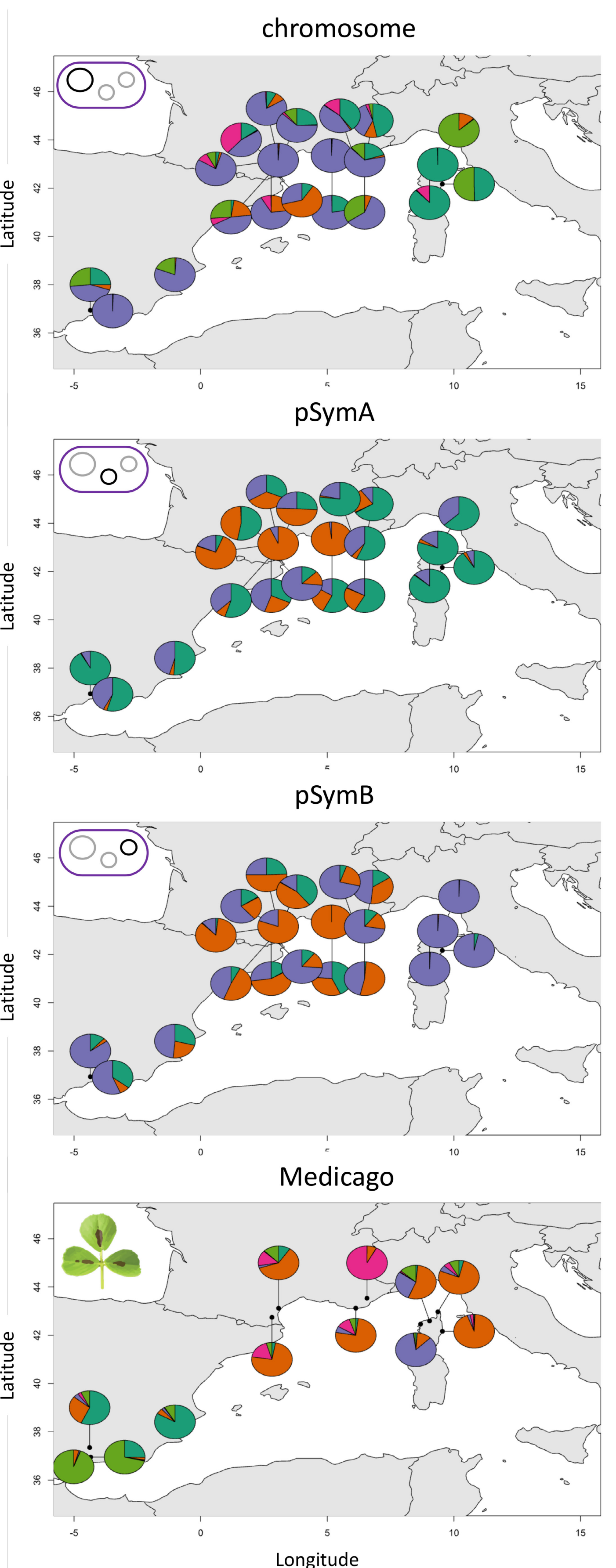


FIGURE 4 Population genetic structure for each of the three genomic elements (chromosome, pSymA and pSymB) in 21 populations of the bacterial symbiont *Sinorhizobium meliloti* and 12 populations of the host plant *Medicago truncatula*. Shown is the proportion of each population assigned to each genetic cluster ($K = 5$ for the bacterial chromosome and the host plant; $K = 3$ for the bacterial pSymA and pSymB)

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

conjugation (Blanca-Ordóñez et al., 2010; Epstein & Tiffin, 2021; Nelson et al., 2018). Thus, for a given rate of bacterial dispersal, effective gene flow might be higher on the plasmids than on the chromosome, resulting in even less population structure in pSymA and pSymB. The "divided genome" hypothesis posits that strong within-chromosome linkage favours independent transmission of genes involved in different phases of complex bacterial life cycles such as those of facultatively symbiotic rhizobia (diCenzo & Finan, 2017). Indeed, functional genetic annotations, metabolic models and gene expression data predict that the three major elements in the *S. meliloti* genome play distinct roles in its life history, with the chromosome for basic metabolism, pSymA for nodulation and N fixation, and pSymB for rhizosphere interactions (Barnett et al., 2004; diCenzo et al., 2014; diCenzo & Finan, 2017; Galibert et al., 2001). Our results add distinct population genetic structure across the landscape to the list of differences among elements.

Interestingly, chromosomes found in Corsica are all nearly identical, and these chromosome lineages were found elsewhere at low frequency. This striking result might be due to a founder event on this island, or selection having swept this chromosomal lineage to high frequency. Given the high linkage disequilibrium across the chromosome compared to the pSyms, targets of selection on this element are difficult to distinguish, a common problem in bacterial elements (Sheppard et al., 2018). Given these limitations, using genetic manipulations to uncouple the three elements by curing and exchanging individual replicons (e.g., Checcucci et al., 2018) would enable strong tests of how the elements interact with each other and whether each of the three genome elements are locally adapted.

4.2 | Mismatched host–symbiont population structure

Our results indicate discordance in population structure between hosts and symbionts. Although both plants and symbionts are diverse within sites, hosts show stronger IBD, and much (56%) of the genome-wide variance in the host is structured among populations and regions. By contrast, the majority (65%–80%) of genome-wide variation in *S. meliloti* is found within individual sites. More interestingly, our analyses indicate that regional differentiation in the symbiont is organized quite differently from that of host plants, since chromosomes from Corsica were distinct from most mainland populations (see discussion above), whereas hosts were most differentiated between Spain and elsewhere. Put another way, though most *M. truncatula* hosts from Spain are genetically distinct from those living elsewhere, they are likely to encounter similar *S. meliloti* symbionts. These incongruent population structures may limit advantages that might come from specialization between host genotypes and rhizobium strains (Fernandes et al., 2019; Lion & Gandon, 2015; Nuismer et al., 1999). However, it is important to note that many *S. meliloti* strains can form nodules with other *Medicago* species, and multiple *Medicago* species are sympatric with *M. truncatula* (Bailly et al., 2006; Bena et al., 2005). A full understanding of the forces shaping the population structure in *S. meliloti* would require not only characterizing the population structure of those other species, but also evaluating the role of selection, either to the soil or host environments. Identifying the genes responsible for fitness differences, and characterizing their evolutionary history, would be necessary to differentiate the role of gene flow, drift and selection in shaping the population structure of this species.

5 | CONCLUSIONS

Here we use a hierarchically structured sample of 191 strains of symbiotic nitrogen-fixing bacteria to show that the three elements in a divided bacterial genome can have distinct population genetic structure across the landscape—distinct from each other and also distinct from that of their eukaryotic host plant. Our results highlight the importance of considering spatial genetic processes in the co-evolutionary dynamics of host–microbe symbiosis, particularly given the widespread action of plasmids and other MGEs in mobilizing interaction genes within and among natural populations of microbial symbionts.

AUTHOR CONTRIBUTIONS

KDH and MAG conceived of and designed the work. KDH, MAG and PT acquired funding. MAG collected soil, performed trapping experiments and isolated strains, and MAG and ABR extracted and submitted DNA for sequencing. ABR and BE performed bioinformatic and evolutionary analyses. ABR, KDH and MAG drafted the article, and all authors participated in critical revisions and approved the final version for submission.

ACKNOWLEDGEMENTS

We thank Laurène Gay and Joelle Ronfort for advice and contributions to field collection, and Amy Marshall-Colón, Julian Catchen and Rachel Whitaker for feedback on an early manuscript draft. Tatiana Giraud and four anonymous reviewers provided constructive reviews that led to substantial improvements to the manuscript. We thank various sources of funding including NSF IOS-1401864 to MAG, NSF IOS-1645875 to KDH, NSF PGRP-1856744 to PT and KDH, JGI CSP-503446 to PT and KDH, and the Department of Plant Biology at the University of Illinois Urbana-Champaign.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at 10.5061/dryad.wpzgmsbqj. (Riley et al., 2021).

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DATA AVAILABILITY STATEMENT

Raw sequence reads and assemblies are archived at NCBI (see Table S1 for accessions). VCF files and R code are archived at DRYAD (doi:10.5061/dryad.wpzgmsbqj).

ORCID

Brendan Epstein https://orcid.org/0000-0001-7083-1588
Peter Tiffin https://orcid.org/0000-0003-1975-610X
Katy D. Heath https://orcid.org/0000-0002-6368-744X

REFERENCES

Anderson, B., Olivieri, I., Lourmas, M., & Stewart, B. A. (2004). Comparative population genetic structures of local adaptation of two mutualists. *Evolution*, *58*(8), 1730–1747. https://doi.org/10.1554/03-317

Bailly, X., Giuntini, E., Sexton, M. C., Lower, R. P., Harrison, P. W., Kumar, N., & Young, J. P. W. (2011). Population genomics of Sinorhizobium medicae based on low-coverage sequencing of sympatric isolates. *The ISME Journal*, *5*(11), 1722–1734. https://doi.org/10.1038/ismej.2011.55

Bailly, X., Olivieri, I., Mita, S. D., Cleyet-Marel, J.-C., & Bena, G. (2006). Recombination and selection shape the molecular diversity pattern of nitrogen-fixing Sinorhizobium sp. associated to Medicago. *Molecular Ecology*, *15*(10), 2719–2734. https://doi.org/10.1111/j.1365-294x.2006.02969.x

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Barnett, M. J., Tolman, C., Fisher, R., & Long, S. R. (2004). A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(47), 16636–16641. https://doi.org/10.1073/pnas.0407269101

Batstone, R. T. (2021). Genomes within genomes: Nested symbiosis and its implications for plant evolution. *New Phytologist*, *234*, 28–34. https://doi.org/10.1111/nph.17847

Batstone, R. T., Dutton, E. M., Wang, D., Yang, M., & Frederickson, M. E. (2017). The evolution of symbiont preference traits in the model legume *Medicago truncatula*. *New Phytologist*, *213*(4), 1850–1861. https://doi.org/10.1111/nph.14308

Baums, I. B., Devlin-Durante, M. K., & LaJeunesse, T. C. (2014). New insights into the dynamics between reef corals and their associated dinoflagellate endosymbionts from population genetic studies. *Molecular Ecology*, *23*(17), 4203–4215. https://doi.org/10.1111/mec.12788

Becker, A., Barnett, M. J., Capela, D., Dondrup, M., Kamp, P.-B., Krol, E., Linke, B., Rüberg, S., Runte, K., Schroeder, B. K., Weidner, S., Yurgel, S. N., Batut, J., Long, S. R., Pühler, A., & Goesmann, A. (2009). A portal for rhizobial genomes: RhizoGATE integrates a *Sinorhizobium meliloti* genome annotation update with postgenome data. *Journal of Biotechnology*, *140*(1–2), 45–50. https://doi.org/10.1016/j.jbiotec.2008.11.006

Bena, G., Lyet, A., Huguet, T., & Olivieri, I. (2005). Medicago-Sinorhizobium symbiotic specificity evolution and the geographic expansion of Medicago. *Journal of Evolutionary Biology*, *18*(6), 1547–1558. https://doi.org/10.1111/j.1420-9101.2005.00952.x

Biondi, E. G., Pilli, E., Giuntini, E., Roumiantseva, M. L., Andronov, E. E., Onichtchouk, O. P., Kurchak, O. N., Simarov, B. V., Dzyubenko, N. I., Mengoni, A., & Bazzicalupo, M. (2003). Genetic relationship of Sinorhizobium meliloti and Sinorhizobium medicae strains isolated from Caucasian region. *FEMS microbiology letters*, *220*(2), 207–213. https://doi.org/10.1016/S0378-1097(03)00098-3

Blanca-Ordóñez, H., Oliva-García, J. J., Pérez-Mendoza, D., Soto, M. J., Olivares, J., Sanjuán, J., & Nogales, J. (2010). pSymA-dependent mobilization of the *Sinorhizobium meliloti* pSymB Megaplasmid. *Journal of Bacteriology*, *192*(23), 6309–6312. https://doi.org/10.1128/jb.00549-10

Bonhomme, M., Boitard, S., Clemente, H. S., Dumas, B., Young, N. D., & Jacquet, C. (2015). Genomic signature of selective sweeps illuminates adaptation of Medicago truncatula to root-associated microorganisms. *Molecular Biology and Evolution*, *msv092*, 2097–2110. https://doi.org/10.1093/molbev/msv092

Bonnin, I., Huguet, T., Gherardi, M., Prosperi, J.-M., & Olivieri, I. (1996). High level of polymorphism and spatial structure in a selfing plant species, *Medicago truncatula* (Leguminosae), shown using RAPD markers. *American Journal of Botany*, *83*(7), 843–855.

Branca, A., Paape, T. D., Zhou, P., Briskine, R., Farmer, A. D., Mudge, J., Bharti, A. K., Woodward, J. E., May, G. D., Gentzbittel, L., Ben, C., Denny, R., Sadowsky, M. J., Ronfort, J., Bataillon, T., Young, N. D., & Tiffin, P. (2011). Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume *Medicago truncatula*. *Proceedings of the National Academy of Sciences*, *108*(42), E864–E870. https://doi.org/10.1073/pnas.1104032108

Brown, S. P., Grillo, M. A., Podowski, J. C., & Heath, K. D. (2020). Soil origin and plant genotype structure distinct microbiome compartments in the model legume *Medicago truncatula*. *Microbiome*, *8*(1), 139. https://doi.org/10.1186/s40168-020-00915-9

Burghardt, L. T. (2019). Evolving together, evolving apart: Measuring the fitness of rhizobial bacteria in and out of symbiosis with leguminous plants. *New Phytologist*, *228*, 28–34. https://doi.org/10.1111/nph.16045

Burghardt, L. T., Epstein, B., & Tiffin, P. (2019). Legacy of prior host and soil selection on rhizobial fitness in planta. *Evolution*, *73*(9), 2013–2023. https://doi.org/10.1111/evo.13807

Caldera, E. J., & Currie, C. R. (2012). The population structure of antibiotic-producing bacterial symbionts of *Apterostigma dentigerum* ants: Impacts of coevolution and multipartite Symbiosis. *The American Naturalist*, *180*(5), 604–617. https://doi.org/10.1086/667886

Carlsson-Granér, U., & Thrall, P. H. (2015). Host resistance and pathogen infectivity in host populations with varying connectivity. *Evolution*, *69*(4), 926–938. https://doi.org/10.1111/evo.12631

Carrascal, O. M. P., VanInsberghe, D., Juárez, S., Polz, M. F., Vinuesa, P., & González, V. (2016). Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing rhizobium species associated with *Phaseolus vulgaris*. *Environmental Microbiology*, *18*(8), 2660–2676. https://doi.org/10.1111/1462-2920.13415

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Cavassim, M. I. A., Moeskjær, S., Moslemi, C., Fields, B., Bachmann, A., Vilhjálmsson, B. J., Schierup, M. H., Young, J. P., & Andersen, S. U. (2020). Symbiosis genes show a unique pattern of introgression and selection within a *Rhizobium leguminosarum* species complex. *Microbial Genomics*, *6*(4), 1–15. https://doi.org/10.1099/mgen.0.000351

Chase, A. B., Arevalo, P., Brodie, E. L., Polz, M. F., Karaoz, U., & Martiny, J. B. H. (2019). Maintenance of sympatric and allopatric populations in free-living terrestrial bacteria. *MBio*, *10*(5), e02361-19. https://doi.org/10.1128/mbio.02361-19

Checcucci, A., diCenzo, G. C., Ghini, V., Bazzicalupo, M., Becker, A., Decorosi, F., Döhlemann, J., Fagorzi, C., Finan, T. M., Fondi, M., Luchinat, C., Turano, P., Vignolini, T., Viti, C., & Mengoni, A. (2018). Creation and characterization of a genomically hybrid strain in the nitrogen-fixing symbiotic bacterium *Sinorhizobium meliloti*. *ACS*

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Synthetic Biology, *7*, 2365–2378. https://doi.org/10.1021/acssynbio.8b00158

Choudoir, M. J., Barberán, A., Menninger, H. L., Dunn, R. R., & Fierer, N. (2018). Variation in range size and dispersal capabilities of microbial taxa. *Ecology*, *99*(2), 322–334. https://doi.org/10.1002/ecy.2094

Cooper, V. S., Vohr, S. H., Wrocklage, S. C., & Hatcher, P. J. (2010). Why genes evolve faster on secondary chromosomes in bacteria. *PLoS Computational Biology*, *6*(4), e1000732. https://doi.org/10.1371/journal.pcbi.1000732

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

diCenzo, G. C., & Finan, T. M. (2017). The divided bacterial genome: Structure, function, and evolution. *Microbiology and Molecular Biology Reviews*, *81*(3), 208–237. https://doi.org/10.1128/mmbr.00019-17

diCenzo, G. C., MacLean, A. M., Milunovic, B., Golding, G. B., & Finan, T. M. (2014). Examination of prokaryotic multipartite genome evolution through experimental genome reduction. *PLoS Genetics*, *10*(10), e1004742. https://doi.org/10.1371/journal.pgen.1004742

Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20. https://doi.org/10.18637/jss.v022.i04

Drew, G. C., Stevens, E. J., & King, K. C. (2021). Microbial evolution and transitions along the parasite–mutualist continuum. *Nature Reviews Microbiology*, *1–16*, 623–638. https://doi.org/10.1038/s41579-021-00550-7

Dybdahl, M. F., & Lively, C. M. (1996). The geography of coevolution: Comparative population structures for a snail and its trematode parasite. *Evolution*, *50*(6), 2264–2275. https://doi.org/10.2307/2410696

Epstein, B., Branca, A., Mudge, J., Bharti, A. K., Briskine, R., Farmer, A. D., Sugawara, M., Young, N. D., Sadowsky, M. J., & Tiffin, P. (2012). Population genomics of the facultatively mutualistic bacteria *Sinorhizobium meliloti* and *S. medicae*. *PLoS Genetics*, *8*(8), e1002868. https://doi.org/10.1371/journal.pgen.1002868

Epstein, B., Sadowsky, M. J., & Tiffin, P. (2014). Selection on horizontally transferred and duplicated genes in Sinorhizobium (Ensifer), the root-nodule symbionts of Medicago. *Genome Biology and Evolution*, *6*(5), 1199–1209. https://doi.org/10.1093/gbe/evu090

Epstein, B., & Tiffin, P. (2021). Comparative genomics reveals high rates of horizontal transfer and strong purifying selection on rhizobial symbiosis genes. *Proceedings of the Royal Society B*, *288*(1942), 20201804. https://doi.org/10.1098/rspb.2020.1804

Fernandes, L. D., Lemos-Costa, P., Guimarães, P. R., Thompson, J. N., & Aguiar, M. A. M. (2019). Coevolution creates complex mosaics across large landscapes. *The American Naturalist*, *194*(2), 217–229. https://doi.org/10.1086/704157

Feurtey, A., Gladieux, P., Hood, M. E., Snirc, A., Cornille, A., Rosenthal, L., & Giraud, T. (2016). Strong phylogeographic co-structure between the anther-smut fungus and its white campion host. *New Phytologist*, *212*(3), 668–679. https://doi.org/10.1111/nph.14125

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210x.12382

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*(4), 973–983. https://doi.org/10.1534/genetics.113.160572

Galardini, M., Pini, F., Bazzicalupo, M., Biondi, E. G., & Mengoni, A. (2013). Replicon-dependent bacterial genome evolution: The case of *Sinorhizobium meliloti*. *Genome Biology and Evolution*, *5*(3), 542–558. https://doi.org/10.1093/gbe/evt027

Galibert, F., Finan, T. M., Long, S. R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M. J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R. W., Dreano, S., … Batut, J. (2001). The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, *293*(5530), 668–672. https://doi.org/10.1126/science.1060966

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing.

Greenlon, A., Chang, P. L., Damtew, Z. M., Muleta, A., Carrasquilla-Garcia, N., Kim, D., Nguyen, H. P., Suryawanshi, V., Krieg, C. P., Yadav, S. K., Patel, J. S., Mukherjee, A., Udupa, S., Benjelloun, I., Thami-Alami, I., Yasin, M., Patil, B., Singh, S., Sarma, B. K., von Wettberg, E., … Cook, D. R. (2019). Global-level population genomics reveals differential effects of geography and phylogeny on horizontal gene transfer in soil bacteria. *Proceedings of the National Academy of Sciences*, *116*(30), 15200–15209. https://doi.org/10.1073/pnas.1900056116

Grillo, M. A., Mita, S., Burke, P. V., Solórzano-Lowell, K. L. S., & Heath, K. D. (2016). Intrapopulation genomics in a model mutualist: Population structure and candidate symbiosis genes under selection in *Medicago truncatula*. *Evolution*, *70*(12), 2704–2717. https://doi.org/10.1111/evo.13095

Hall, J. P. J., Botelho, J., Cazares, A., & Baltrus, D. A. (2022). What makes a megaplasmid? *Philosophical Transactions of the Royal Society B*, *377*(1842), 20200472. https://doi.org/10.1098/rstb.2020.0472

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506. https://doi.org/10.1038/nrmicro2795

Harrison, P. W., Lower, R. P. J., Kim, N. K. D., & Young, J. P. W. (2010). Introducing the bacterial 'chromid': Not a chromosome, not a plasmid. *Trends in Microbiology*, *18*(4), 141–148. https://doi.org/10.1016/j.tim.2009.12.010

Harrison, T. L., Wood, C. W., Heath, K. D., & Stinchcombe, J. R. (2017). Geographically structured genetic variation in the *Medicago lupulina*-Ensifermutualism. *Evolution*, *62*, 220–215. https://doi.org/10.1111/evo.13268

Hartmann, F. E., Snirc, A., Cornille, A., Godé, C., Touzet, P., Van Rossum, F., Fournier, E., Le Prieur, S., Shykoff, J., & Giraud, T. (2020). Congruent population genetic structures and divergence histories in anther-smut fungi and their host plants Silene italica and the Silene nutans species complex. *Molecular Ecology*, *29*(6), 1154–1172. https://doi.org/10.1111/mec.15387

Heath, K. D. (2010). Intergenomic epistasis and coevolutionary constraint in plants and rhizobia. *Evolution*, *64*(5), 1446–1458. https://doi.org/10.1111/j.1558-5646.2009.00913.x

Heath, K. D., Batstone, R. T., Romero, M. C., & McMullen, J. G. (2022). MGEs as the MVPs of partner quality variation in legume-rhizobium Symbiosis. *MBio*, *13*, e00888-22. https://doi.org/10.1128/mbio.00888-22

Heath, K. D., & Stinchcombe, J. R. (2014). Explaining mutualism variation: A new evolutionary paradox? *Evolution*, *68*(2), 309–317. https://doi.org/10.1111/evo.12292

Heath, K. D., & Tiffin, P. (2009). Stabilizing mechanisms in a legume-rhizobium mutualism. *Evolution*, *63*(3), 652–662. https://doi.org/10.1111/j.1558-5646.2008.00582.x

Hoeksema, J. D., & Forde, S. E. (2008). A meta-analysis of factors affecting local adaptation between interacting species. *American Naturalist*, *171*(3), 275–290. https://doi.org/10.1086/527496

Hoetzinger, M., Pitt, A., Huemer, A., & Hahn, M. W. (2021). Continental-scale gene flow prevents allopatric divergence of pelagic freshwater bacteria. *Genome Biology and Evolution*, *13*(3), evab019. https://doi.org/10.1093/gbe/evab019

Hollowell, A. C., Regus, J. U., Turissini, D., Gano-Cohen, K. A., Bantay, R., Bernardo, A., Moore, D., Pham, J., & Sachs, J. L. (2016). Metapopulation dominance and genomic-Island acquisition of Bradyrhizobium with superior catabolic capabilities. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1829), 20160496. https://doi.org/10.1098/rspb.2016.0496

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Jombart, T., & Ahmed, I. (2011). Adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M., & Tabata, S. (2002). Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Research*, *9*(6), 189–197. https://doi.org/10.1093/dnares/9.6.189

Klinger, C. R., Lau, J. A., & Heath, K. D. (2016). Ecological genomics of mutualism decline in nitrogen-fixing bacteria. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1826), 20152563. https://doi.org/10.1098/rspb.2015.2563

Koppell, J. H., & Parker, M. A. (2012). Phylogenetic clustering of Bradyrhizobium symbionts on legumes indigenous to North America. *Microbiology*, *158*(Pt_8), 2050–2059. https://doi.org/10.1099/mic.0.059238-0

Kraemer, S. A., & Boynton, P. J. (2017). Evidence for microbial local adaptation in nature. *Molecular Ecology*, *26*(7), 1860–1876. https://doi.org/10.1111/mec.13958

Kumar, N., Lad, G., Giuntini, E., Kaye, M. E., Udomwong, P., Shamsani, N. J., Young, J. P., & Bailly, X. (2015). Bacterial genospecies that are not ecologically coherent: Population genomics of *Rhizobium leguminosarum*. *Open Biology*, *5*(1), 140133. https://doi.org/10.1098/rsob.140133

Kuzmanović, N., Fagorzi, C., Mengoni, A., Lassalle, F., & diCenzo, G. C. (2022). Taxonomy of Rhizobiaceae revisited: Proposal of a new framework for genus delimitation. *International Journal of Systematic and Evolutionary Microbiology*, *72*(3), 1–17. https://doi.org/10.1099/ijsem.0.005243

Laine, A. (2005). Spatial scale of local adaptation in a plant-pathogen metapopulation. *Journal of Evolutionary Biology*, *18*(4), 930–938. https://doi.org/10.1111/j.1420-9101.2005.00933.x

Lawrence, J. G., & Roth, J. R. (1996). Selfish operons: Horizontal transfer may drive the evolution of gene clusters. *Genetics*, *143*(4), 1843–1860. https://doi.org/10.1093/genetics/143.4.1843

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lion, S., & Gandon, S. (2015). Evolution of spatially structured host–parasite interactions. *Journal of Evolutionary Biology*, *28*(1), 10–28. https://doi.org/10.1111/jeb.12551

Martiny, J. B., Bohannan, B. J., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Ovreås, L., Reysenbach, A. L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102–112. https://doi.org/10.1038/nrmicro1341

Moran, N. A. (2007). Symbiosis as an adaptive process and source of phenotypic complexity. *Proceedings of the National Academy of Sciences*, *104*(suppl 1), 8627–8633. https://doi.org/10.1073/pnas.0611659104

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*(949), 283–292. https://doi.org/10.1086/282771

Nelson, M., Guhlin, J., Epstein, B., Tiffin, P., & Sadowsky, M. J. (2018). The complete replicons of 16 *Ensifer meliloti* strains offer insights into intra- and inter-replicon gene transfer, transposon-associated loci, and repeat elements. *Microbial Genomics*, *4*(5), 37–11. https://doi.org/10.1099/mgen.0.000174

Nuismer, S. L., Thompson, J. N., & Gomulkiewicz, R. (1999). Gene flow and geographically structured coevolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1419), 605–609. https://doi.org/10.1098/rspb.1999.0679

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., Fookes, M., Falush, D., Keane, J. A., & Parkhill, J. (2015). Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, *31*(22), 3691–3693. https://doi.org/10.1093/bioinformatics/btv421

Pahua, V. J., Stokes, P., Hollowell, A. C., Regus, J. U., Gano-Cohen, K. A., Wendlandt, C. E., Quides, K. W., Lyu, J. Y., & Sachs, J. L. (2018). Fitness variation among host species and the paradox of ineffective rhizobia. *Journal of Evolutionary Biology*, *31*(4), 599–610. https://doi.org/10.1111/jeb.13249

Paradis, E., & Schliep, K. (2018). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*(5), 946–952. https://doi.org/10.1111/1755-0998.12129

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient swiss Army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*(7), 1929–1936. https://doi.org/10.1093/molbev/msu136

Pita, L., Rix, L., Slaby, B. M., Franke, A., & Hentschel, U. (2018). The sponge holobiont in a changing ocean: From microbes to ecosystems. *Microbiome*, *6*(1), 46. https://doi.org/10.1186/s40168-018-0428-1

Porter, S. S., Faber-Hammond, J., Montoya, A. P., Friesen, M. L., & Sackos, C. (2019). Dynamic genomic architecture of mutualistic cooperation in a wild population of Mesorhizobium. *The ISME Journal*, *13*(2), 301–315. https://doi.org/10.1038/s41396-018-0266-y

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rambaut, A. (2018). *FigTree*. Edinburgh. Retrieved from. http://tree.bio.ed.ac.uk/software/figtree/

Revillini, D., Gehring, C. A., & Johnson, N. C. (2016). The role of locally adapted mycorrhizas and rhizobacteria in plant–soil feedback systems. *Functional Ecology*, *30*(7), 1086–1098. https://doi.org/10.1111/1365-2435.12668

Riley, A.B., Grillo, M.A., Epstein, B., Tiffin, P., & Heath, K.D. (2021). Discordant population structure among rhizobium divided genomes and their legume hosts. NCBI (see accessions in Table S1) and DRYAD (https://doi.org/10.5061/dryad.wpzgmsbqj).

Ronfort, J., Bataillon, T., Santoni, S., Delalande, M., David, J. L., & Prosperi, J.-M. (2006). Microsatellite diversity and broad scale geographic structure in a model legume: Building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *BMC Plant Biology*, *6*, 28. https://doi.org/10.1186/1471-2229-6-28

Rosenberg, E., & Zilber-Rosenberg, I. (2018). The hologenome concept of evolution after 10 years. *Microbiome*, *6*(1), 78. https://doi.org/10.1186/s40168-018-0457-9

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sheppard, S. K., Guttman, D. S., & Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, *19*(9), 549–565. https://doi.org/10.1038/s41576-018-0032-z

Simon, J.-C., Marchesi, J. R., Mougel, C., & Selosse, M.-A. (2019). Host-microbiota interactions: From holobiont theory to analysis. *Microbiome*, *7*(1), 5. https://doi.org/10.1186/s40168-019-0619-4

Siol, M., Prosperi, J. M., Bonnin, I., & Ronfort, J. (2008). How multilocus genotypic pattern helps to understand the history of selfing populations: A case study in *Medicago truncatula*. *Heredity*, *100*(5), 517–525. https://doi.org/10.1038/hdy.2008.5

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Smith, C. I., Godsoe, W. K. W., Tank, S., Yoder, J. B., & Pellmyr, O. (2008). Distinguishing coevolution from covicariance in an obligate pollination mutualism: Asynchronous divergence in Joshua tree and its pollinators. *Evolution*, *62*(10), 2676–2687. https://doi.org/10.1111/j.1558-5646.2008.00500.x

Stoy, K. S., Gibson, A. K., Gerardo, N. M., & Morran, L. T. (2020). A need to consider the evolutionary genetics of host–symbiont mutualisms. *Journal of Evolutionary Biology*, *33*(12), 1656–1668. https://doi.org/10.1111/jeb.13715

Strobel, H. M., Alda, F., Sprehn, C. G., Blum, M. J., & Heins, D. C. (2016). Geographic and host-mediated population genetic structure in a cestode parasite of the three-spined stickleback. *Biological Journal of the Linnean Society*, *119*(2), 381–396. https://doi.org/10.1111/bij.12826

Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., & Ronson, C. W. (1995). Nodulating strains of rhizobium loti arise through chromosomal symbiotic gene transfer in the environment. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(19), 8985–8989.

Tack, A. J. M., Horns, F., & Laine, A.-L. (2014). The impact of spatial scale and habitat configuration on patterns of trait variation and local adaptation in a wild plant parasite. *Evolution*, *68*(1), 176–189. https://doi.org/10.1111/evo.12239

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K. L., Yandell, M., Gundlach, H., Mayer, K. F., Schwartz, D. C., & Town, C. D. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics*, *15*(1), 312. https://doi.org/10.1186/1471-2164-15-312

Thompson, J. N. (2005). *The geographic mosaic of coevolution (interspecific interactions)* (1st ed.). University Of Chicago Press.

Thrall, P. H., Hochberg, M. E., Burdon, J., & Bever, J. D. (2007). Coevolution of symbiotic mutualists and parasites in a community context. *Trends in Ecology & Evolution*, *22*(3), 120–126. https://doi.org/10.1016/j.tree.2006.11.007

Toro, N., Martínez-Abarca, F., & Fernández-López, M. (2016). The early events underlying genome evolution in a localized *Sinorhizobium meliloti* population. *BMC Genomics*, *17*(1), 556. https://doi.org/10.1186/s12864-016-2878-9

VanInsberghe, D., Arevalo, P., Chien, D., & Polz, M. F. (2020). How can microbial population genomics inform community ecology? *Philosophical Transactions of the Royal Society B*, *375*(1798), 20190253. https://doi.org/10.1098/rstb.2019.0253

Vincent, J. (1970). *A manual for the practical study of root-nodule bacteria*. Blackwell Scientific Publications.

Vos, M., & Velicer, G. J. (2008). Isolation by distance in the spore-forming soil bacterium *Myxococcus xanthus*. *Current Biology*, *18*(5), 386–391. https://doi.org/10.1016/j.cub.2008.02.050

Wardell, G. E., Hynes, M. F., Young, P. J., & Harrison, E. (2022). Why are rhizobial symbiosis genes mobile? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1842), 20200471. https://doi.org/10.1098/rstb.2020.0471

Weisberg, A. J., Rahman, A., Backus, D., Tyavanagimatt, P., Chang, J. H., & Sachs, J. L. (2022). Pangenome evolution reconciles robustness and instability of Rhizobial Symbiosis. *MBio*, *13*, e00074-22. https://doi.org/10.1128/mbio.00074-22

Whitaker, R. J., & Banfield, J. F. (2006). Population genomics in natural microbial communities. *Trends in Ecology & Evolution*, *21*(9), 508–516. https://doi.org/10.1016/j.tree.2006.07.001

Wickham, H. (2016). Positioning. In *ggplot2, elegant graphics for data analysis*. Springer Nature. https://doi.org/10.1007/978-3-319-24277-4_7

Yates, R., Howieson, J., De Meyer, S. E., Tian, R., Seshadri, R., Pati, A., Woyke, T., Markowitz, V., Ivanova, N., Kyrpides, N., Loi, A., Nutt, B., Garau, G., Sulas, L., & Reeve, W. (2015). High-quality permanent draft genome sequence of *Rhizobium sullae* strain WSM1592; a Hedysarum coronarium microsymbiont from Sassari, Italy. *Standards in Genomic Sciences*, *10*(1), 44. https://doi.org/10.1186/s40793-015-0020-2

Young, J. P., Crossman, L. C., Johnston, A. W., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., Wexler, M., Curson, A. R., Todd, J. D., Poole, P. S., Mauchline, T. H., East, A. K., Quail, M. A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., … Parkhill, J. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biology*, *7*(4), R34. https://doi.org/10.1186/gb-2006-7-4-r34

Young, J. P. W., & Haukka, K. (1996). Diversity and phylogeny of rhizobia. *New Phytologist*, *133*, 87–94.

Zribi, K., Mhamdi, R., Huguet, T., & Aouani, M. (2004). Distribution and genetic diversity of rhizobia nodulating natural populations of *Medicago truncatula* in tunisian soils. *Soil Biology & Biochemistry*, *36*, 903–908.

SUPPORTING INFORMATION

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

How to cite this article: Riley, A. B., Grillo, M. A., Epstein, B., Tiffin, P., & Heath, K. D. (2022). Discordant population structure among rhizobium divided genomes and their legume hosts. *Molecular Ecology*, *00*, 1–14. https://doi.org/10.1111/mec.16704

1365294x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/mec.16704 by Loyola University Chicago, Wiley Online Library on [05/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License