Essential Amino Acid Enrichment and Positive Selection Highlight Endosymbiont's Role in a Global Virus-Vectoring Pest

Kaitlyn N. Myers,[a] Daniel Conn,[a] Amanda M. V. Brown[a]

[a]Department of Biological Sciences, Texas Tech University, Lubbock, Texas, USA

ABSTRACT Host-associated microbes display remarkable convergence in genome repertoire resulting from selection to supplement missing host functions. Nutritional supplementation has been proposed in the verrucomicrobial endosymbiont *Xiphinematobacter* sp., which lives within a globally widespread group of plant-parasitic nematodes that vector damaging nepoviruses to plants. Only one genome sequence has been published from this symbiont, leaving unanswered questions about its diversity, host range, role, and selective pressures within its hosts. Because its hosts are exceptionally resistant to culturing, this symbiont is best studied through advanced genomic approaches. To analyze the role of *Xiphinematobacter* sp. in its host, sequencing was performed on nematode communities, and then genomes were extracted for comparative genomics, gene ontology enrichment tests, polymorphism analysis, de Bruijn-based genome-wide association studies, and tests of pathway- and site-specific selection on genes predicted play a role in the symbiosis. Results showed a closely clustered set of *Xiphinematobacter* isolates with reduced genomes of ~917 kbp, for which a new species was proposed. Symbionts shared only 2.3% of genes with outgroup *Verrucomicrobia*, but comparative analyses showed high conservation of all 10 essential amino acid (EAA) biosynthesis pathways plus several vitamin pathways. These findings were supported by gene ontology enrichment tests and high polymorphisms in these pathways compared with background. Genome-wide association analysis confirmed high between-species fixation of alleles with significant functional enrichment for EAA and thiamine synthesis. Strong positive selection was detected on sites within these pathways, despite several being under increased purifying selection. Together, these results suggest that supplementation of EAAs missing in the host diet may drive this widespread symbiosis.

IMPORTANCE *Xiphinematobacter* spp. are distinctly evolved intracellular symbionts in the phylum *Verrucomicrobia*, which includes the important human gut-associated microbe *Akkermansia muciniphila* and many highly abundant free-living soil microbes. Like *Akkermansia* sp., *Xiphinematobacter* sp. is obligately associated with the gut of its hosts, which in this case consists of a group of plant-parasitic nematodes that are among the top 10 most destructive species to global agriculture, by vectoring plant viruses. This study examined the hypothesis that the key to this symbiont's stable evolutionary association with its host is through provisioning nutrients that its host cannot make that may be lacking in the nematode's plant phloem diet, such as essential amino acids and several vitamins. The significance of our research is in demonstrating, using population genomics, the signatures of selective pressure on these hypothesized roles to ultimately learn how this independently evolved symbiont functionally mirrors symbionts of phloem-feeding insects.

KEYWORDS amino acid biosynthesis, comparative genomics, genome-wide association study, mutualism, nutritional symbiont, plant-parasitic nematode, population genomics, symbiosis

Microbes often play critical roles in plant-feeding organisms (1), such as breaking down plant starches or supplementing essential amino acids (EAAs) for phloem-feeding insects (2, 3). Microbes have converged evolutionarily on this nutritional supplementation role many times across the bacterial tree of life, displaying reduction in costly and unnecessary metabolic processes except those necessary to maintain basic cellular processes and key symbiotic contributions (2, 4–6). While endosymbionts of plant-feeding organisms living above ground have been a major research focus, symbionts in plant feeders below ground have been poorly studied. One remarkable underground endosymbiont is *Xiphinematobacter* sp. (*Verrucomicrobia*), which lives within a globally widespread group of plant-parasitic nematodes. While this ancient symbiosis is not well understood, it has been proposed that the endosymbiont may have evolved to supplement essential amino acids missing in its host diet, mirroring symbionts of phloem-feeding Hemiptera (7, 8). This endosymbiont undergoes tight coevolution with its hosts (9–11) and has a degenerate genome with disproportionate conservation of essential amino acid biosynthesis genes (8) suggestive of a mutualistic role. Yet, to date, there is only one published genome study from this globally distributed symbiont. By investigating additional *Xiphinematobacter* genomes using population-level and species-level comparative genomics, the present study seeks to better understand the biological role of this symbiont in the evolutionary context of this symbiosis.

Xiphinematobacter sp. is of interest not only for convergent symbiosis evolution (7) but also as a contributor to the success of one of the most damaging plant-parasitic nematodes in agriculture (10, 12, 13). Broadly, nematodes may be the most numerically abundant animals on Earth (14) with millions of individuals per m^2 of soil (15–17) that perform ecosystem services, such as nutrient cycling (18–20). Plant-parasitic nematodes make up a large portion of the community close to roots, causing up to 25% of global crop yield loss at costs of ~$100 billion annually (21). Dagger nematodes (*Xiphinema* spp.) which host *Xiphinematobacter* sp. are among the top 10 most damaging nematodes as vectors of at least 13 nepoviruses, including tobacco ringspot and tomato ringspot virus (12, 13). *Xiphinematobacter* endosymbionts occur within the *Xiphinema americanum* species-complex, which includes up to 61 described species (22, 23) that infect an enormous range of host plants globally, including grape, almond, apricot, peach, nectarine, plum, prune, walnut, cherry, strawberries, soybeans, perennial orchards, and wild plants, including spruce, pine, and common weeds and grasses (24–26), sometimes at densities of 500 individuals per 250 cm^3 of soil (27). Thus, given the host abundance of *Xiphinematobacter* sp., the symbiont's high titer in its host (28), and its global distribution in soils (29, 30), *Xiphinematobacter* sp. appears to be a globally abundant endosymbiont.

Experimental research on the *Xiphinema-Xiphinematobacter* symbiosis remains exceedingly difficult, due to the difficulty of manipulating and experimenting with these nematodes under laboratory conditions (13, 24, 31, 32). Thus, in the present study, we sought to uncover field-sampled genomes of *Xiphinematobacter* sp. that could be analyzed using comparative genomics, single nucleotide polymorphism (SNP) analysis and population genomics using de Bruijn-based genome-wide association (DBGWAS) (33), and tests of site-wise natural selection in genes or pathways hypothesized to be important in the symbiont's role to better understand the potential nutritional supplementation role of *Xiphinematobacter* sp. While PCR- and Sanger sequencing-based approaches have been helpful for identifying potential diversity and coevolution of *Xiphinematobacter* spp. with their hosts (9, 11, 34, 35), we focus on deeper analyses using comparative genomics (8, 36, 37) and microbial population genomics approaches (38–42) which are increasingly valuable tools to use to discover microbial function (43). We combine these approaches with community genome skimming, which is a promising new approach to simultaneously access plant-parasitic nematode hosts and their microbiomes (28).

Results showed that skimming was highly effective for yielding insights on

TABLE 1 Genome features among *Xiphinematobacter* spp.

Sample name[a]	Genome length (bp)	GC content (%)	No. of genes	No. of proteins	No. of rRNAs	No. of tRNAs	No. of pseudogenes	NCBI accession no.
P15	917,845	49.0	893	844	3	46	3	CP068477
P18	915,441	49.0	896	847	3	46	4	CP068476
P21	926,970	49.0	899	848	3	48	2	CP068475
P22	909,775	49.0	885	836	3	46	3	CP068474
P23	917,278	49.0	888	839	3	46	2	CP068473
P3-11[b]	734,636	49.0	794	752	3	39	30	CP068472
XipG	915,884	47.7	867	818	3	46	1	CP012665.1
Xip2	915,884	47.7	867	818	3	46	1	SRX1527792

[a]Names in bold denote isolates collected in this study.
[b]The genome for this sample was interpreted to be incomplete.

nematode communities while producing full-length *Xiphinematobacter* genomes for which we propose a new species. Detailed comparative genomics, gene ontology (GO) enrichment analysis, population genomics, genome-wide association (GWAS), and tests for selection demonstrate distinct dynamics for essential amino acid and thiamine biosynthesis pathways, supporting them as nutrients that could be provisioned from symbiont to host.

RESULTS

***Xiphinematobacter* and *Xiphinema* spp. occurrence in field sample sites.** *Xiphinematobacter* spp. or *Xiphinema* spp. were detected at 11 sampled sites in Texas and New Mexico, as distant as 1,160 km apart. Six sites were found to have *Xiphinematobacter* spp., including samples from farms and mixed natural grass and shrub communities (see Table S1 in the supplemental material). There were up to 60 nematode species per 100-g sample, based on top cytochrome oxidase I (COI) blastn matches. There were 29 *Xiphinema* species (based on DNA fragment identities) detected overall, across samples based on COI matches. They included up to 9 species of *Xiphinema* per sample (see Table S2 in the supplemental material). Assembled nematode mitochondrial contigs include 16 contigs of >12,000 bp, comprising nearly full-length mitogenomes. Using both lower and higher kmer ranges during assembly yielded several long *Xiphinematobacter* contigs that comprised 50% to 95% of the bacterial genome at coverage depths ranging from 23.45× to 84.22× (Table S2), with assembly N_{50} values ranging from 16,926 to 871,420 bp. Postassembly contig alignment and reassembly produced full-length or nearly full-length genome sequences.

Genome features of new *Xiphinematobacter* isolates. Five of six *Xiphinematobacter* genomes appear complete based on gene repertoire, rRNA and tRNA completeness, and genome length, whereas the sixth genome appeared to be ∼80% complete (Table 1). For the *Xiphinematobacter* isolates sequenced in the current study with complete genomes (samples P15, P18, P21, P22, and P23), genome features such as GC content (49.0%), number of tRNAs (44), and number of pseudogenes (2 to 4) were similar among isolates and the previously sequenced *Xiphinematobacter* genomes (Table 1). Genome length ranged from 909,775 to 926,970 bp among isolates, and the number of predicted proteins ranged from 836 to 848 (Table 1). The average predicted gene length was ∼867 bp, which was somewhat smaller than average gene lengths in outgroup *Verrucomicrobia* (e.g., ∼1,015 to 1,041 bp); however, ortholog comparisons did not suggest an abundance of partial (i.e., recently inactivated) genes in *Xiphinematobacter* genomes.

Phylogenetic analysis of *Xiphinematobacter* isolates. Phylogenetic analysis of full-length 16S rRNA gene sequences for the 6 *Xiphinematobacter* isolates from this study and 38 other *Xiphinematobacter* sequences placed these isolates with high support into a single clade (100% bootstrap support and Bayesian posterior probability of 1), close to *Xiphinematobacter* isolates from *Xiphinema luci* collected in Spain (Fig. 1). This clade, designated "phylotype K," is shown in the context of phylotypes proposed as *Xiphinematobacter* species by Lazarova et al. (10). Notably, the only *Xiphinematobacter* isolate whose genome
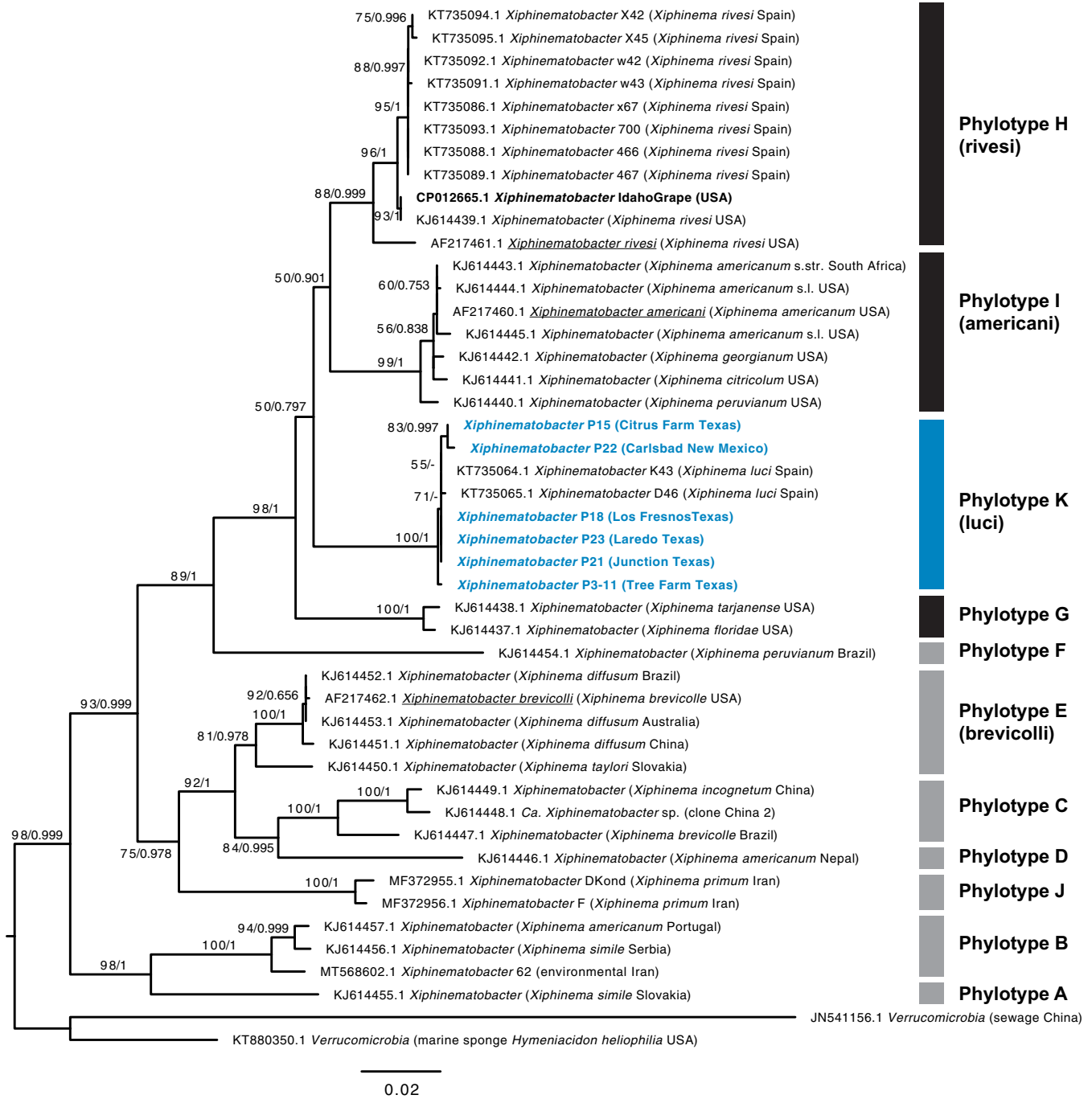
FIG 1 Phylogenetic tree of *Xiphinematobacter* spp. and verrucomicrobial outgroups, based on 1,513 aligned positions of the 16S rRNA gene, generated using RAxML GTR+G with 500 and MrBayes with GTR+G with 4 rate categories model, showing bootstrap replicate and posterior probability values on branches. Taxon names in blue bold font are from sequences obtained in this study. Phylotype designation based on reference 10 is shown alongside taxa, with dominant widespread North American groups indicated in black bars and other more globally diverse groups indicated with gray bars. Underlined taxon names indicate previously named *Xiphinematobacter* species. Black bold font indicates the isolate with previously sequenced genome.

had been sequenced previously, *Xiphinematobacter* sp. from Idaho Grape (GenBank accession number CP012665.1) falls into a distinct, strongly supported clade (phylotype H) that includes *Xiphinematobacter rivesi*. As shown in Fig. 1, the isolates from the current study and the Idaho Grape isolate clustered within a larger clade dominated by these symbionts from *Xiphinema* species that are widespread in North America (phylotypes G, H, I, and K), whereas other phylotypes occur in *Xiphinema* spp. that are more broadly distributed globally.
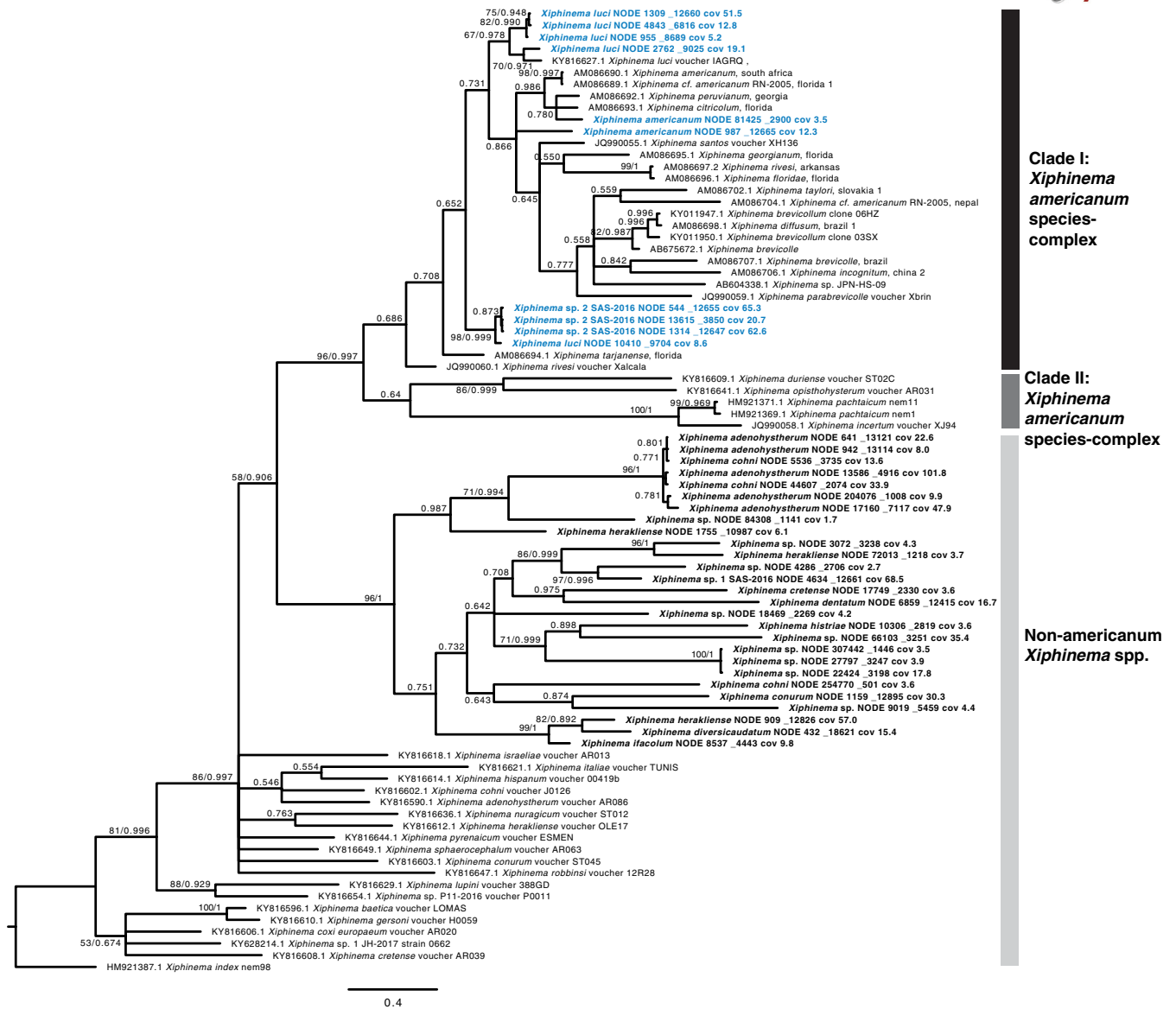
FIG 2 Phylogenetic tree of *Xiphinema americanum* species complex and non-*X. americanum Xiphinema* spp. showing the groups found in this study, based on 372 aligned positions of the partial cytochrome oxidase I (COI) gene, generated using Bayesian 50% majority rule in MrBayes with GTR+G with 4 rate categories model, and showing posterior probabilities and maximum likelihood bootstrap values from RAxML GTR+G with 500 and bootstrap replicates on the branches. Taxon names in bold are from sequences obtained in this study with *X. americanum* group representatives indicated in blue.

An analysis of a subsection of the 16S rRNA gene for which additional North American isolates have been sequenced (see Fig. S1 in the supplemental material) showed results consistent with full-length 16S rRNA analyses, supporting phylotype K. This clade included the *Xiphinematobacter* isolates from this study and the *Xiphinematobacter* isolates from *X. luci* collected in Spain as a sister clade to clade mt-C (11), which is notable as a distinct clade that is apparently not associated with nepoviruses.

Concatenated protein-coding gene-based phylogenetic analyses showed strong support for a close monophyletic group of *Xiphinematobacter* spp. with long branch distances to the closest outgroup *Verrucomicrobia* (see Fig. S2 in the supplemental material).

Phylogenetic analysis and community profiles of *Xiphinema* nematodes. Phylogenetic analysis of the sequences with highest blastn similarity to the *Xiphinema* spp. COI gene from the 11 sampled locations produced a distinct pattern with 3 major clades (Fig. 2), representing previously characterized clades I and II of the *Xiphinema americanum sensu lato* species complex and a non-*X. americanum Xiphinema* spp. clade. Sequences from the current study fell into three subgroups within clade I *X. americanum* and one large and highly diverged group of non-*X. americanum*
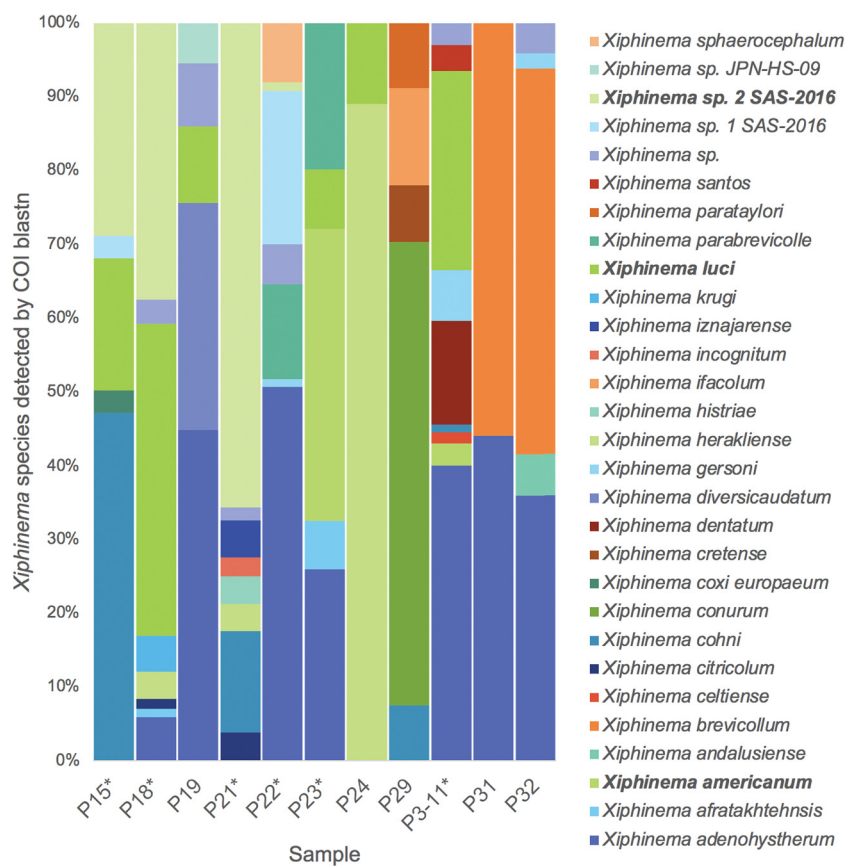
FIG 3 Relative percentage of each *Xiphinema* species in each sample, based on partial cytochrome oxidase I (COI) gene blast searches.

Xiphinema spp. (Fig. 2). The clade I *X. americanum* isolates were close matches to *X. luci*, *X. americanum*, and *Xiphinema* sp. 2 SAS-2016 from reference 45. There were no COI sequences from the current study with any similarity to clade II *X. americanum*.

The relative proportions of COI matches in the 11 samples (considering sequencing coverage) are shown in Fig. 3 and indicate that most samples had 1 or 2 dominant *Xiphinema* species.

Absolute coverage comparisons (Fig. 4) among samples indicate that *Xiphinematobacter*-positive samples (P15, P18, P21, P22, P23, and P3-11) generally had high coverage of *X. americanum* species-complex nematodes relative to *X. americanum Xiphinema* nematodes compared with *Xiphinematobacter*-negative samples (especially P19, P24, and P29) except for P32, which was dominated by the of *X. americanum* species complex nematode *Xiphinema brevicollum*, but did not carry matches close to *X. luci*.

Genome repertoire and enrichment analysis of Xiphinematobacter isolates and outgroups. Genome-wide sequence comparisons, including gene order and gene presence-absence comparisons, showed conserved synteny and similar gene compaction patterns and core genome repertoire among all eight *Xiphinematobacter* genomes, which differed distinctly from genomes of outgroup Verrucomicrobia (see Fig. S3 in the supplemental material).

Pangenome analyses (Fig. 5) revealed only 349 conserved genes out of 14,893 (2.3%) shared between *Xiphinematobacter* isolates and *Verrucomicrobia* outgroups. There were 747 genes unique to the *Xiphinematobacter* pangenome (i.e., not occurring in outgroups), of which a majority (470/747, 62.9%) were not annotated to function (i.e., predicted "hypothetical protein"). In contrast, the majority of genes (716 of 1,096, 65.3%) were conserved across the 7 fully sequenced *Xiphinematobacter* isolates, and
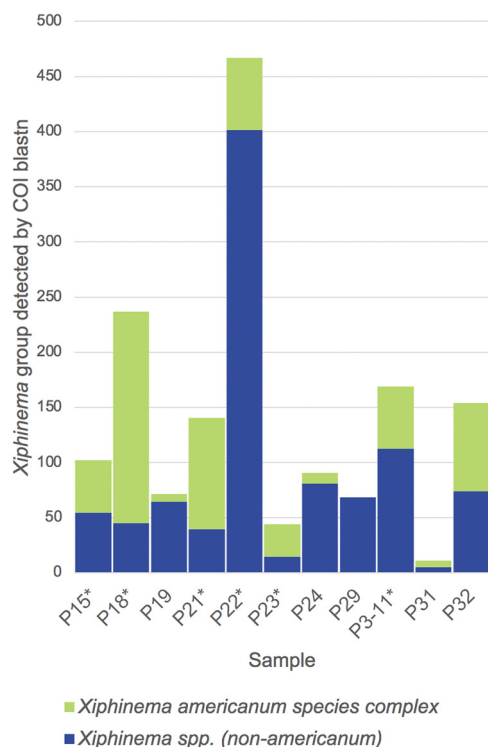
FIG 4 Total coverage of each *Xiphinema* type (*X. americanum* group versus non-*X. americanum* group) in each sample, based on partial cytochrome oxidase I (COI) gene blast searches.

among the 5 new genomes sequenced here, 75.8% (739 of 975) of genes were shared across isolates (Fig. 5). There were 379 core genes shared across all *Xiphinematobacter* species isolates but not present in outgroups, 254 genes unique to the pangenome of our 5 new *Xiphinematobacter* isolates, and 115 genes unique to the pangenome of the 2 previously sequenced *Xiphinematobacter* isolates (XipG and Xip2). However, from the 254 genes unique to the new isolates, only 10 of could be annotated with known gene function, while the others were annotated as hypothetical protein. Similarly, from the 115 genes unique to the previously sequenced isolates, only 5 genes could be annotated to a specific function.

Specific gene presence-absence patterns for amino acid and vitamin/cofactor biosynthetic gene pathways among *Xiphinematobacter* isolates and *Verrucomicrobia* outgroups are depicted in Fig. 6. These patterns show a high level of conservation of genes for the 10 essential amino acid (EAA) biosynthesis pathways (arginine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), with lower conservation of genes in some of the nonessential amino acid (non-EAA) pathways and most of the vitamin/cofactor pathways. Among the essential amino acid pathway genes, methionine, histidine, and leucine appear to be missing several genes in *Xiphinematobacter* isolates, although the missing genes are conserved among the isolates. Among nonessential amino acids (alanine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, selenocysteine, serine, and tyrosine), only genes for tyrosine appear to be well conserved; however, most of this pathway is shared with phenylalanine. Among the vitamins/cofactors, riboflavin and lipoate biosynthesis stand out as having largely intact genes in *Xiphinematobacter* sp. (Fig. 6).

From the 747 genes unique to the *Xiphinematobacter* pangenome (i.e., not occurring in outgroups) (Fig. 5), 262 protein-coding genes could be annotated beyond hypothetical protein. These genes were analyzed to assess functional enrichment compared with the background or "universe" genes for the combined
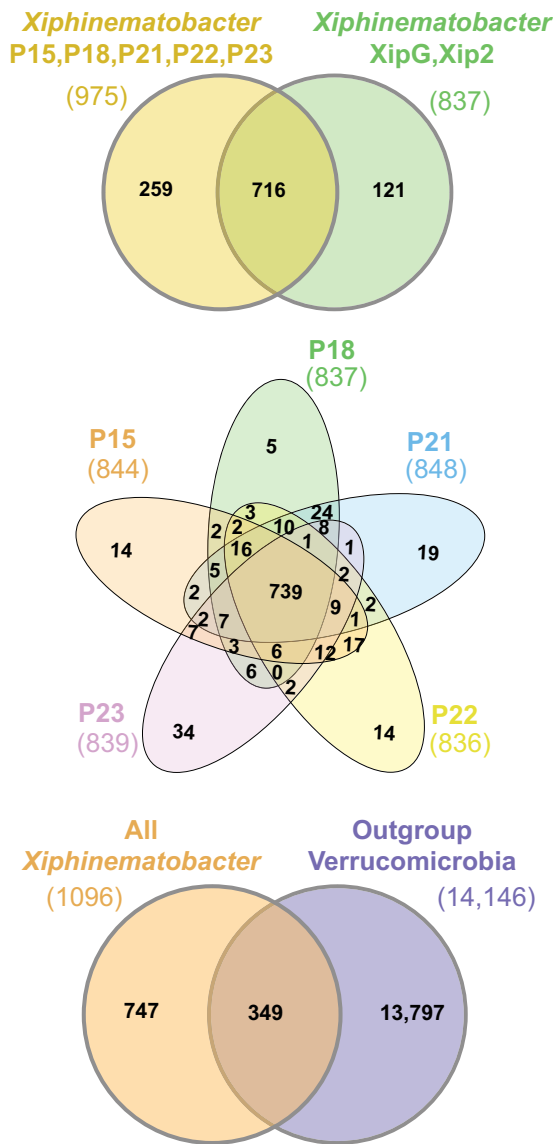
FIG 5 Pangenomes for *Xiphinematobacter* isolate from this study, previously sequenced *Xiphinematobacter*, and outgroup Verrucomicrobia, including *Spartobacteria bacterium* AMD-G4 (NEUK01000001.1), *Chthoniobacter flavus* Ellin428 (ABVL01000062.1), *Terrimicrobium sacchariphilum* NM-5 (NZ_BDCO01000003), and "*Candidatus* Udaeobacter sp. AEW3 Udaeo2_1" (JAALOD010000001.1).

Xiphinematobacter and *Verrucomicrobia* pangenome using topGO. The results (Table 2) showed significantly enriched GO terms (*P* values of <0.05) for 17 biological processes, 2 cellular compartments, and 6 metabolic functions. Most of these significantly enriched terms include housekeeping functions (e.g., cell division, regulation of cell shape, and cell cycle), but among these terms were also enriched an biosynthesis of several EAAs and vitamins, including aromatic amino acids, tryptophan, riboflavin, thiamine, and folic acid-containing compound biosynthesis.

Polymorphism analysis within and between *Xiphinematobacter* isolates. Single nucleotide polymorphism (SNP) density analysis across 441 conserved genes from 8 *Xiphinematobacter* genomes (Fig. 7), calculated using snp-sites software, showed several genes with exceptionally high SNP density, including some amino acid and vitamin/cofactor biosynthesis genes. SNP density was slightly higher for amino acid biosynthesis genes than for vitamin/cofactor genes and background genes ("other") not falling within these categories; Mann-Whitney U test of these categorized SNP
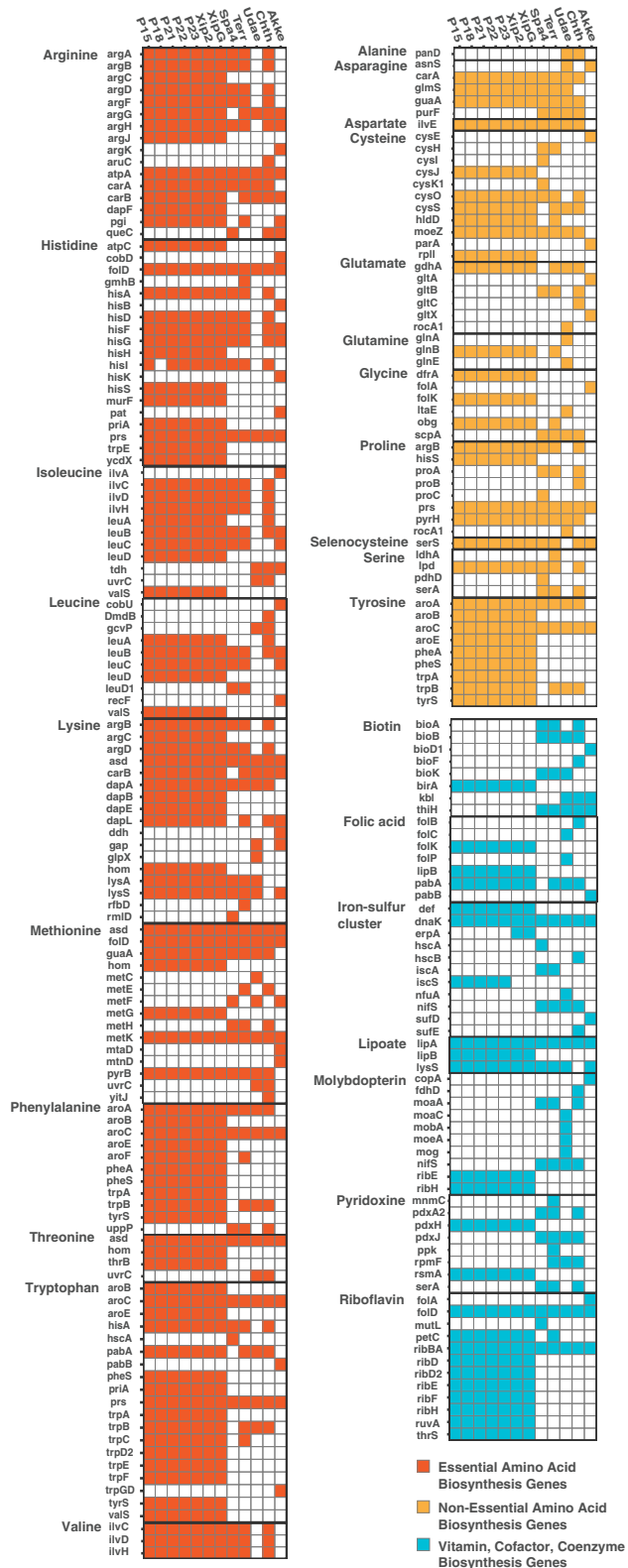
FIG 6 Amino acid and vitamin/cofactor biosynthesis gene presence and absence patterns for *Xiphinematobacter* isolates from this study (P15, P18, P21, P22, and P23), previously sequenced *Xiphinematobacter* spp. (XipG and Xip2), and outgroup *Verrucomicrobia* spp., including *Spartobacteria bacterium* AMD-G4 (NEUK01000001.1; Spa4), *Chthoniobacter flavus* Ellin428 (ABVL01000062.1; Chth), *Terrimicrobium sacchariphilum* NM-5 (NZ_BDCO01000003; Terr), "*Candidatus* Udaeobacter sp. AEW3 Udaeo2_1" (JAALOD010000001.1; Udae), and *Akkermansia muciniphila* YL44 (NZ_CP021420.1; Akke).

TABLE 2 Significantly enriched gene ontology categories among *Xiphinematobacter* species[a]

GO_iD by category	Term	No. of annotated genes	No. of significant genes	No. of genes expected	*P* value
Biological processes					
GO:0051301	Cell division	33	14	3.95	7.8e-06
GO:0008360	Regulation of cell shape	36	12	4.31	0.00025
GO:0007049	Cell cycle	33	12	3.95	0.00043
GO:0015937	Coenzyme A biosynthetic process	5	4	0.6	0.00064
GO:0009073	Aromatic amino acid family biosynthetic process	30	12	3.59	0.00213
GO:0009423	Chorismate biosynthetic process	10	5	1.2	0.00237
GO:0009252	Peptidoglycan biosynthetic process	34	10	4.07	0.00243
GO:0015986	ATP synthesis coupled proton transport	7	4	0.84	0.00372
GO:0009231	Riboflavin biosynthetic process	9	4	1.08	0.01123
GO:0009396	Folic acid-containing compound biosynthetic process	10	4	1.2	0.01177
GO:0009245	Lipid A biosynthetic process	17	5	2.04	0.03069
GO:0000162	Tryptophan biosynthetic process	7	3	0.84	0.03230
GO:0043165	Gram-negative-bacterium-type cell outer membrane assembly	7	3	0.84	0.03230
GO:0008299	Isoprenoid biosynthetic process	16	4	1.92	0.03294
GO:0009228	Thiamine biosynthetic process	13	4	1.56	0.04488
GO:0006364	rRNA processing	22	7	2.63	0.04615
GO:0043093	Ftsz-dependent cytokinesis	8	3	0.96	0.04759
Cellular components					
GO:0016021	Integral component of membrane	441	61	51.26	0.00095
GO:0032153	Cell division site	7	3	0.81	0.03042
Metabolic functions					
GO:0003723	RNA binding	121	27	13.61	0.0020
GO:0046933	Proton-transporting ATP synthase activity, rotational mechanism	7	4	0.79	0.0031
GO:0005524	ATP binding	362	50	40.71	0.0145
GO:0003887	DNA-directed DNA polymerase activity	8	3	0.9	0.0418
GO:0000049	tRNA binding	26	6	2.92	0.0457
GO:0016757	Transferase activity, transferring glycosyl groups	85	17	9.56	0.0491

[a]Based on topGO gene ontology enrichment analysis among *Xiphinematobacter* species, based on genes unique to the *Xiphinematobacter* pangenome that do not occur in outgroups.

densities showed significant differences, including EAA versus vitamin/cofactor (Z-score, 2.03461; *P* = 0.02118) and EAA versus other (Z-score, −2.06105; *P* = 0.0197). Notably high SNP density genes included the EAA genes *aroA*, *aroC*, *argB*, *dapA*, *hisB1*, *hisS*, *metG*, and *ycdX*; the non-EAAs *aroC*, *hisS*, and *serS*; and the vitamin gene *iscC*. Notably low SNP density genes included *argA*, *gdhA*, *pdhC*, *petC*, *ribAB*, and *thrB*.

Gene ontology enrichment analysis of phenotype-associated loci from de Bruijn GWAS. Using DBGWAS for 7 *Xiphinematobacter* genomes, 51,449 unique SNPs were detected in de Bruijn graphs, from which 4,193 SNPs (de Bruijn kmers) showed a statistically significant association with the host *Xiphinema* group (specifically, *X. rivesi* or *X. luci*) (see Table S3 in the supplemental material). Functional enrichment analysis was performed on the set of 3,440 of these SNPs for which the gene could be annotated (i.e., removing hypothetical proteins) and a gene ontology (GO) category could be assigned. GO enrichment tests were performed on the 525 genes containing these SNPs by comparing the associated gene set to the background or universe genes for *Xiphinematobacter* and *Verrucomicrobia* sp. using topGO. The results (Table 3) showed 25 significantly enriched GO biological processes (*P* < 0.05), which included several housekeeping functions (e.g., translation, *de novo* UMP biosynthesis, and cell division). Among the 25 significantly enriched processes, 7 were for EAA (histidine, arginine, iso-leucine, valine, threonine, leucine, and tryptophan) biosynthesis, 1 was aromatic amino
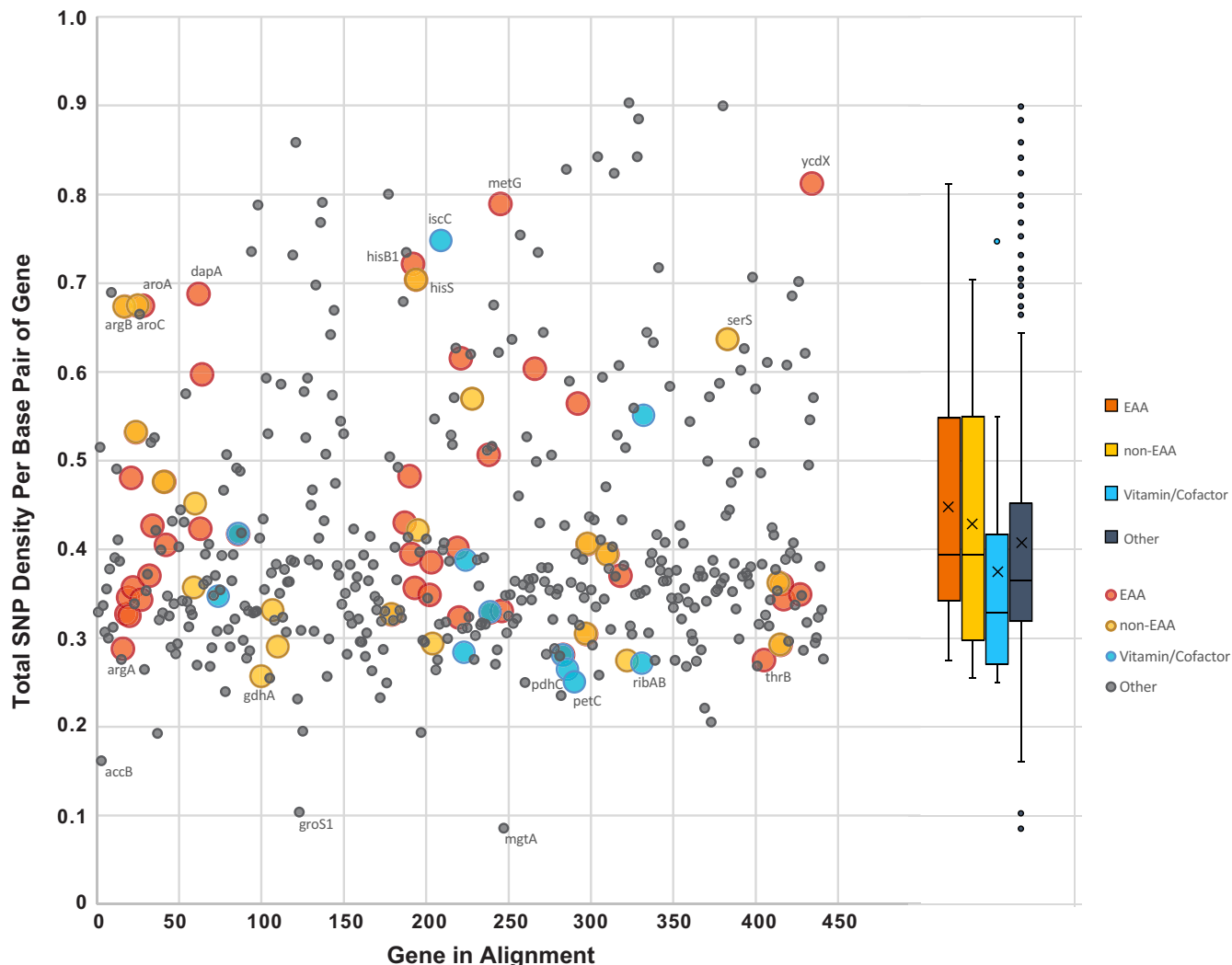
FIG 7 Single nucleotide polymorphism (SNP) density across conserved genes from eight *Xiphinematobacter* genomes compared in this study, across 441 genes, showing amino acid and vitamin/cofactor biosynthesis gene SNP density patterns. Mann-Whitney U test showed significant difference between essential amino acid (EAA) and vitamin/cofactor groups and between EAA and background (other) genes. Boxplot (right) shows 1st to 3rd quartile range, median (line), mean (x), and outliers.

acid biosynthesis (tyrosine, phenylalanine, and tryptophan), and 1 was for biosynthesis of the vitamin thiamine.

Polymorphism density for associated loci from de Bruijn GWAS. Our higher-resolution analysis of these data showed the density of the significantly associated DBGWAS SNPs per gene (Fig. 8). These host-phenotype-associated SNP sites showed a trend in contrast to the total SNP density (Fig. 7), with essential amino acid gene DBGWAS SNPs having lower density, although this was not significant and this density matched that of the background genes (Fig. 8). Mann-Whitney U test of these categorized DBGWAS SNP densities showed a significant difference between non-EAA versus vitamin/cofactor (Z-score, 1.66675; $P = 0.04746$). Notably high DBGWAS SNP density genes included the non-EAA genes *pheS*, *glgC*, and *gltX*. Notably low-SNP genes include the vitamin genes *rnc*, *ruvA*, and *thiH*; and amino acid genes *grpE*, *ilvH*, and *iolG*.

Analysis of signatures of selection across gene groups. The ratio of nonsynonymous and synonymous substitutions (dN/dS) between *Xiphinematobacter* clades for genes categorized into individual "nutritional" biosynthesis pathways identified in Table 2 are shown in Fig. 9. Compared with typical background genome-wide values,

TABLE 3 Significantly enriched gene ontology categories from DBGWAS analysis[a]

GO_iD	Term	No. of annotated genes	No. of signficant genes	No. of genes expected	P value
GO:0006412	Translation	100	71	23.89	8.0e-19
GO:0000105	Histidine biosynthetic process	10	9	2.39	1.0e-05
GO:0044205	"De novo" UMP biosynthetic process	7	7	1.67	2.7e-05
GO:0051301	Cell division	33	17	7.88	6.8e-05
GO:0006526	Arginine biosynthetic process	11	9	2.63	0.00017
GO:0015937	Coenzyme A biosynthetic process	5	5	1.19	0.00055
GO:0008360	Regulation of cell shape	36	17	8.6	0.00076
GO:0007049	Cell cycle	33	14	7.88	0.00079
GO:0009073	Aromatic amino acid family biosynthetic process	30	22	7.17	0.00144
GO:0009423	Chorismate biosynthetic process	10	7	2.39	0.00169
GO:0042254	Ribosome biogenesis	36	14	8.6	0.00239
GO:0015986	ATP synthesis coupled proton transport	7	6	1.67	0.00265
GO:0009097	Isoleucine biosynthetic process	11	7	2.63	0.00375
GO:0009245	Lipid A biosynthetic process	17	9	4.06	0.00548
GO:0009252	Peptidoglycan biosynthetic process	34	14	8.12	0.01007
GO:0019288	Isopentenyl diphosphate biosynthetic process, methylerythritol 4-phosphate pathway	5	4	1.19	0.01012
GO:0009099	Valine biosynthetic process	5	4	1.19	0.01012
GO:0009088	Threonine biosynthetic process	5	4	1.19	0.01012
GO:0006351	Transcription, DNA-templated	148	18	35.36	0.01196
GO:0006096	Glycolytic process	18	9	4.3	0.02044
GO:0009098	Leucine biosynthetic process	9	5	2.15	0.03067
GO:0006396	RNA processing	63	25	15.05	0.04495
GO:0000162	Tryptophan biosynthetic process	7	4	1.67	0.04821
GO:0009228	Thiamine biosynthetic process	13	6	3.11	0.04879
GO:0006353	DNA-templated transcription, termination	6	5	1.43	0.04929

[a]Based on topGO gene ontology enrichment analysis for *Xiphinematobacter*, based on genes that are significantly associated with host phenotype (*Xiphinema* species clade) in de Bruijn genome wide association (DBGWAS) SNP analysis.

dN/dS was lower in genes for arginine, isoleucine, leucine, valine, and tryptophan biosynthesis. In contrast, dN/dS was higher than background levels in genes for threonine and thiamine biosynthesis and aromatic amino acids.

Within species of *Xiphinematobacter*, isoleucine and valine again showed lower dN/dS than background, but arginine showed a higher dN/dS, while other trends in Fig. 9 were not observed, with most genes having values close to that of the background (see Fig. S4 in the supplemental material).

Analysis of site-specific positive selection across gene groups. Site-specific tests of positive selection in the same groups of genes were calculated in codeML using the likelihood ratio test (LRT) for nested models of positive selection (M7 = beta −; 10 categories versus M8 beta ω of >1; 11 categories), with results shown in Table S4 in the supplemental material, and Fig. 10 P values showed significance at <0.05 for ω of >1 (i.e., positively selected sites) for all biosynthesis pathways except valine biosynthesis. Histidine was the second lowest in ω and second highest in P value, whereas the most outstanding high ω pathways were, in descending order, leucine, thiamine, isoleucine, aromatic amino acids, threonine, and tryptophan.

DISCUSSION

A wide range of endosymbionts play critical roles in their hosts (2, 4, 46). Familiar examples include *Buchnera* sp. (*Enterobacteriaceae*) which supplement missing essential amino acids in their aphid hosts and *Wolbachia* (*Alphaproteobacteria*) which are either reproductive parasites or mutualists in insects and nematodes. However, for some unculturable obligate endosymbionts living within nonmodel hosts, such as *Xiphinematobacter* sp., ecological and physiological experiments are extremely challenging; therefore, studies of the endosymbiont role depend critically on genomics. Here, we present novel insights into the role of *Xiphinematobacter* sp. by focusing on genomics-based approaches to understand the selective pressure on nutritional supplementation pathways.
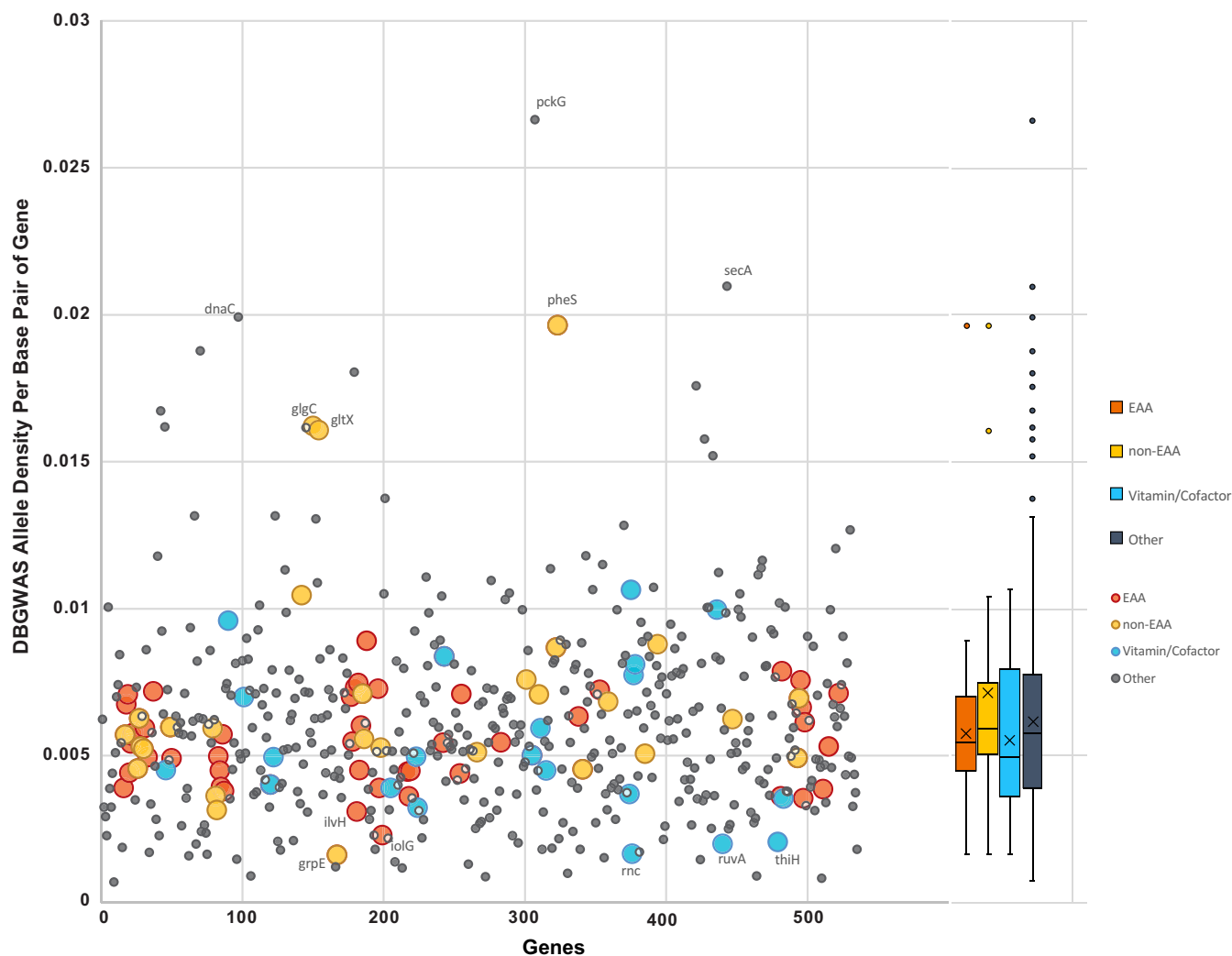
FIG 8 Density of de Bruijn genome-wide association (DBGWAS) SNPs that are significantly associated with different host phenotype (*Xiphinema* species clade) plotted for 536 genes, showing amino acid and vitamin/cofactor biosynthesis gene SNP density patterns. Mann-Whitney U test showed significant differences between non-EAA and vitamin/cofactor genes. Boxplot (right) shows 1st to 3rd quartile range, median (line), mean (x), and outliers.

Based on our comparative genomic analyses of genome repertoire and enrichment, a major pattern that emerged was the conservation of several essential nutrient biosynthetic pathway genes across *Xiphinematobacter* isolates compared with the outgroup *Verrucomicrobia*, and we also found evidence of directional selection in these pathways using population genomics. Gene repertoire data showed these highly reduced *Xiphinematobacter* endosymbiont genomes have universally conserved biosynthesis pathways for most essential amino acids, tyrosine, and several vitamins/ cofactors, with statistically overrepresented GO terms for tryptophan, aromatic amino acid, riboflavin, folate, and thiamine biosynthesis, despite an endosymbiosis-driven pattern of gene loss, expanding previous results based on a single *Xiphinematobacter* genome (8). While endosymbionts often evolve reduced genome repertoires in the absence of purifying selection on genes for factors supplied by the host (5, 6), it is exceptional for the metabolically expensive essential amino acid and vitamin synthesis pathways to be conserved unless the symbiont is provisioning these nutrients to supplement its host diet.

Many examples of highly conserved pathways for *de novo* biosynthesis of essential amino acids and vitamins/cofactors in endosymbionts occur in insects feeding strictly on low amino acid or low vitamin diets, such as phloem (47) or blood, in which the
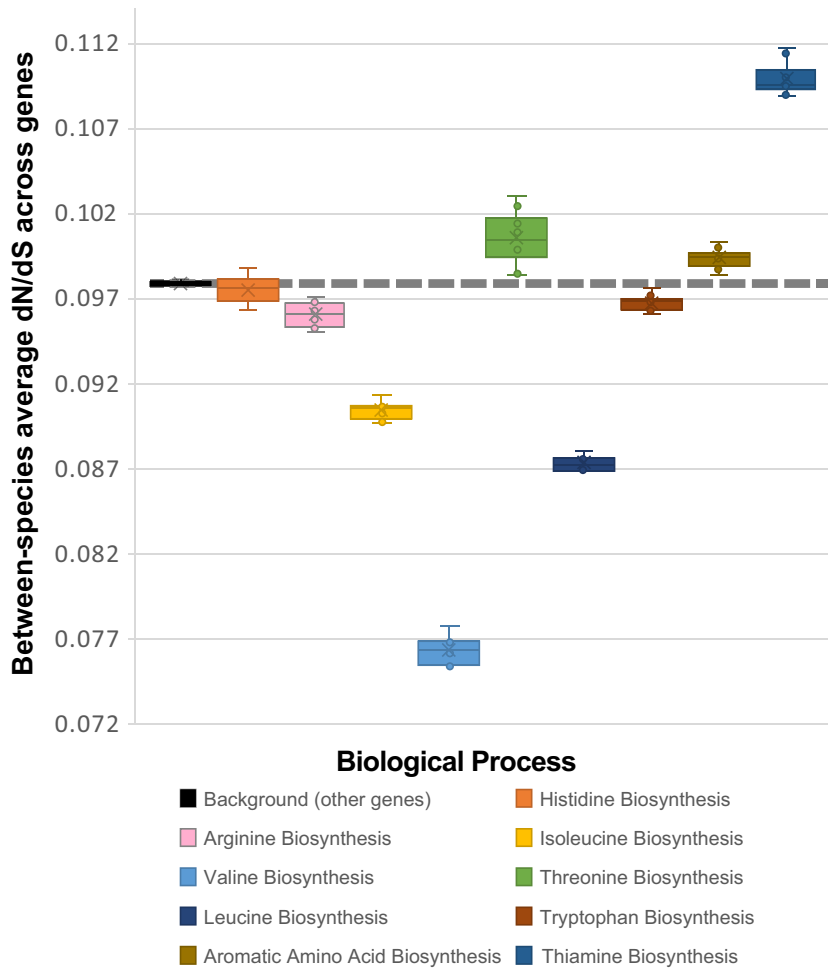
FIG 9 Between-species dN/dS in *Xiphinematobacter* sp. for select amino acid and vitamin/cofactor biosynthesis pathways, identified as enriched from de Bruijn genome-wide association (DBGWAS SNP) gene ontology analysis, compared to remainder of the genome (i.e., background), and calculated using the Nei and Gojobori 1986 method for dN/dS in codeML.

symbionts match the host nutritional demands (48) and facilitate host plant use (1). Our findings of a disproportionate conservation of these pathways hint at a similar role of *Xiphinematobacter* sp. in supplementing these essential nutrients in its nematode diet. These results are supported by the gut wall tissue distribution of this symbiont (8, 49). While nematodes, like other animals, are unable to synthesize these essential molecules *de novo*, nutritional mutualism in plant-feeding nematodes seems to be rare. This is likely because most plant-feeding and plant-parasitic nematodes have sophisticated adaptations to secrete plant cell wall-digesting enzymes that release plant cell contents, which are rich in nutrients (50), explaining the observation that most *Xiphinema* species do not appear to require symbiotic microbes (51–53). Yet, based on strong coevolutionary patterns and 100% prevalence globally in individuals and nominal species in the *X. americanum* complex clade I and most of clade II (9, 11, 34, 35), *Xiphinematobacter* sp. appears to be a long-standing resident mutualist (7, 8). Details of *Xiphinema* feeding patterns suggest some species may specialize on drinking from vascular bundle and phloem (54), potentially explaining nutrient deficiencies that would demand supplementation by a symbiont.

Polymorphism analyses among isolates of *Xiphinematobacter* sp., including gene- and pathway-based SNP density, showed significantly increased standing genetic variation and allele fixation in EAAs compared with vitamin/cofactor and background genes, suggesting a potentially enhanced adaptive potential in EAA biosynthesis
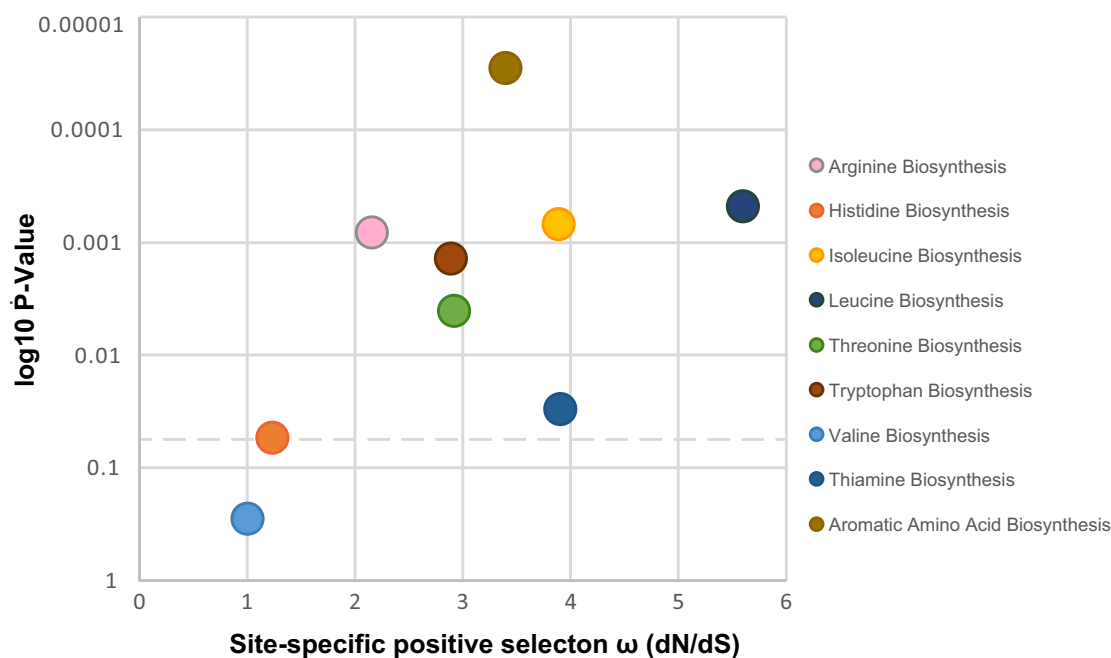
FIG 10 Strength of positive selection and positively selected sites on amino acid and vitamin/cofactor biosynthesis pathway genes among *Xiphinematobacter* species based on the likelihood ratio test (LRT) for nested models of positive selection (M7 = beta -; 10 categories versus M8 beta $\omega > 1$; 11 categories) calculated in codeML. Dotted line indicates a *P* value of 0.05.

pathways. Specific genes with higher SNP density included three genes for histidine (none known to be rate limiting), two critical genes for phenylalanine and tryptophan (*aroA* and *aroC*; encoding 3-phosphoshikimate 1-carboxyvinyltransferase and chorismite synthase, respectively), two for lysine (*argB* encoding acetylglutamate kinase, and *dapA* encoding the lysine biosynthesis rate-limiting product 4-hydroxy-tetrahydrodipicolinate synthase) (55), and one for methionine (*metG* encoding methionyl-tRNA synthase that may affect methionine biosynthesis through altering exogenous methionine levels which generally repress *met* synthesis genes; however, such feedback inhibition likely varies between species and is still poorly understood) (56). The less critical genes in these pathways may accumulate SNPs due to relative relaxation of purifying selection; however, it is unclear why the critical or rate-limiting genes should accumulate more SNPs, unless these mutations accumulate through positive selection. Lower SNP-density genes included two for riboflavin, including the critical riboflavin biosynthesis protein RibAB, which is consistent with riboflavin synthesis being critical, as has been proposed for other symbionts (57).

However, the findings above include both synonymous and nonsynonymous sites, whereas gene ontology (GO) enrichment analyses supported this pattern and focused on the set of variants that were fixed between groups. For example, GO analysis of de Bruijn-based genome-wide association (GWAS) SNPs showed 7 out of 25 (32%) significant terms were for essential amino acid biosynthesis (histidine, arginine, isoleucine, valine, threonine, leucine, and tryptophan). Because this association analysis for DBGWAS SNPs focused on those significantly associated with host species clade as the phenotype (i.e., *X. rivesi* clade versus *X. luci* clade), the focus emphasized fixed SNPs in these symbionts within the *X. americanum* species complex. Thus, the EAA biosynthesis overrepresentation in DBGWAS SNPs in *Xiphinematobacter* sp. is indicative of positive selection driving the fixation of alleles for these functions. These findings add to a growing list of studies demonstrating adaptation through population genomics of host-associated microbes (38, 40, 41, 43, 58–60). Densities of DGBWAS SNPs suggested that fixation in non-EAA biosynthetic pathways was greater, overall, than fixation in vitamin/cofactor pathways, perhaps due to a lower evolutionary rate or higher

purifying selection in the latter. Together, overrepresentation of DBGWAS SNPs in the biosynthesis of EAAs and several vitamin/cofactor processes suggests more rapidly evolving fixation at the species level—a finding consistent with higher overall SNP density, high retention of these genes in gene presence/absence data, and selection, which is discussed below.

Selection tests (e.g., dN/dS, LRT test) showed differential signatures of positive and purifying selection both across biosynthetic pathways and at specifically selected sites. For example, some pathways appeared to be under lower selective constraints (notably thiamine, threonine, and aromatic amino acid biosynthesis) than background dN/dS, while other pathways appeared to be under higher constraint (i.e., purifying selection), such as valine, leucine, and isoleucine, and to a lesser extent, arginine and tryptophan. This finding suggests there are dynamic evolutionary effects on these conserved nutritional pathways—noting that background genome dN/dS was substantially lower than 1, indicating the *Xiphinematobacter* symbiont genome is under strong purifying selection. Mirroring this result, within-species dN/dS was also low for valine and isoleucine, suggesting more constrained evolution than background on these biosynthetic pathways and implying their potential importance across *Xiphinematobacter* sp. In comparative LRT tests of site-specific positive selection, assuming selection (ω) is constant across branches in the phylogeny of *Xiphinematobacter* isolates, we found strong statistical support for positive selection in numerous sites in all the pathways above, except for valine, which is consistent with valine's overall stronger purifying selection. Together, these results suggest higher positive selection, ω, in pathways for the biosynthesis of leucine and isoleucine, with stronger purifying selection between species, but high site-specific positive selection, in contrast to biosynthesis pathways for thiamine and aromatic amino acids, which showed overall higher dN/dS as well as high (i.e., $\omega > 1$) site-specific positive selection, emphasizing likely directional selection on these fixed divergences in these pathways. Between-isolate positive selection in this case may indicate that different hosts may place different selection pressures on these genes (e.g., due to hosts feeding on different plants or living in different habitats). Such analyses of selection on candidate microbial mutualism functions are rare (43, 61) but can be powerful when combined with manipulative experiments (60) and studies of molecular regulation between host and mutualist (62–64) and could form targets for exploration of new methods to control these pests.

The phylogenetic results were consistent with those of previous studies suggesting *Xiphinematobacter* clades codiversify with their hosts (9–11, 34, 35, 45) due to high fidelity vertical transmission of this obligate mutualist. Furthermore, phylogenies showed our new *Xiphinematobacter* isolates formed a well-supported monophyletic clade in all analyses, supporting the distinct species of *Xiphinematobacter*. These isolates had similar 16S rRNA sequences (<0.7% divergence) and similar genomic features (similar gene content, high synteny, and identical GC content). Comparing 16S rRNA divergences between close relatives to this clade (e.g., 4.0% to *Xiphinematobacter rivesi*, 4.1% to *Xiphinematobacter* from Idaho Grape, and 4.6% to *Xiphinematobacter americani*) and using fossil-based calibration from a *Wolbachia* sp. from apid bees (65) that estimates 216 million years divergence per ~2.8% 16S rRNA divergence, we estimate this clade may be ~308 million years old, perhaps diverging in the Carboniferous Period. A fascinating aspect of the *Xiphinema-Xiphinematobacter* symbiosis is the possibility of a replacement symbiont from among the *Burkholderiaceae* which was discovered in *Xiphinema pachtaicum*, *Xiphinema incertum*, *Xiphinema astaregiense*, *Xiphinema parapachydermum*, *Xiphinema vallense*, and *Xiphinema madeirense* in Spain and Iran (9, 34) but is not yet found in North America. However, we did not find reads or contigs with blastn similarity to *Burkholderia*-like symbionts; although, our study did not recover any matches to this symbiont.

To date, there have been few field surveys of nematodes and their microbes (66, 67), but the success of *Xiphinematobacter* and *Xiphinema* spp. recovery from field sites in this study demonstrates the effectiveness of community genome skimming for

survey purposes while simultaneously yielding quality data for deeper analyses, such as phylogenomics and comparative and population genomics. Our census result showing *Xiphinematobacter* sp. in half of our sampled sites is consistent with this endosymbiont being common in soils (29, 30). Our results emphasize an advantage of genome skimming for surveys compared with PCR for long-evolved soil-dwelling endosymbionts which may have priming site differences and inhibitors from soil contaminants (9, 68–70). Although the skimming method does not directly assess host-symbiont combinations, they may be analyzed in complementary studies, for example using Hi-C or other technologies (71–73).

In conclusion, this study presents novel insight into the *Xiphinema-Xiphinematobacter* symbiosis, which closely mirrors nutrient provisioning symbioses in hemipteran insects and highlights convergent roles in the evolution of microbiota associated with animals. We showed isolates of *Xiphinematobacter* conserve essential amino acid and vitamin biosynthetic pathways and yet hold increased standing genetic variation and allele fixation in some of the genes in these pathways, suggesting enhanced adaptive potential in nutrient provisioning by this symbiont. Finally, sequence divergence and phylogenetic evidence provide support for the focal isolates as a distinct species of *Xiphinematobacter* described below. Broadly, these analyses support a genomics-based approach to species designation that could be useful for other experimentally intractable insect and nematode symbiosis.

Description of "*Candidatus* Xiphinematobacter luci" n. sp. *Xiphinematobacter luci* (lu'ci. N.L. gen. masc. n. luci of Luc, i.e., lives in the species *Xiphinema luci*, which was named by Lamberti and Bleve-Zacheo [74] from specimens collected by Michel Luc).

Characteristics are as defined by the DNA sequence analysis description for the genus *Xiphinematobacter*, including 16S rRNA gene similarity among species (53) compared with outgroup *Verrucomicrobia*. Characteristics distinguishing this species are phylogenomic and 16S rRNA phylogeny placement in a supported monophyletic clade (phylotype K) with high 16S rRNA gene sequence similarity, 99.3%, between isolates (this study) and two related isolates of *Xiphinematobacter* sp. from *Xiphinema luci* in Spain (GenBank accession KT735064 and KT735065). This monophyletic clade is diverged in 16S rRNA gene sequences from all three formally described *Xiphinematobacter* species (53), including 4% divergence from *Xiphinematobacter rivesi* (GenBank accession AF217461), 4.6% divergence from *Xiphinematobacter americani* (AF217460), and 7.9% divergence from *Xiphinematobacter brevicolli* (GenBank accession AF217462). Other distinguishing characters include a distinct genomic repertoire from other *Xiphinematobacter* isolates (strains XipG or *Xiphinematobacter* sp. Idaho Grape, GenBank accession CP012665.1; and Xip2 assembled from GenBank BioSample SRX1527792) which also show 16S rRNA gene sequence divergence of 4.1%.

MATERIALS AND METHODS

Field sample collection and nematode isolation from soil. To obtain new *Xiphinematobacter* isolates, approximately 100 to 500 g of soil and roots were collected from locations in Texas and New Mexico (see Table S1), using a double-serrated shovel (Root Assassin, USA) or nonspiral soil auger. Collection and transport were performed in compliance with USDA APHIS permits. Soil and roots were kept in coolers (<10°C) and processed within a few days. Nematodes were isolated by Baermann funnel and, if needed, also using sucrose flotation. Briefly, soil from each sample was divided into ~100-g aliquots and placed in the top of glass funnels lined with two layers of Kimwipe (Kimtech, USA). Distilled water was added to saturate soil, and nematodes were collected over 3 to 5 days in clamped tubing at the bottom. Nematodes were further cleaned by three repetitions of adding 50 ml sterile distilled water, gentle centrifugation (400 × *g* for 3 min), and removal of the supernatant. After the rinse steps, nematodes were inspected in water using an inverted microscope. If soil was present, nematodes were further cleaned by mixing with 1:1 80% sucrose solution, gently centrifuging, and rinsing on a 20-μm mesh sieve (75).

DNA extraction, Illumina library preparation, and sequencing. To obtain and sequence symbiont genomes, DNA was isolated from bulk nematode communities (~500 to 2,000 nematodes) by using five cycles of freeze-thaw to break cuticles, followed by DNA isolation using the Qiagen DNeasy blood and tissue kit (Valencia, CA) following the manufacturer's directions. DNA was checked for quantity and quality on the Nanodrop spectrophotometer, and then ~0.5 to 1 μg of DNA was used for shotgun metagenomic library preparation with the QIAseq FX 96 DNA library kit (Valencia, CA) with enzymatic fragmentation and AMPure bead size selection targets of 450 to 550 bp. Libraries were checked for quality and

quantified on the TapeStation 2200 system (Agilent, USDA) before normalizing and pooling multiplex barcoded libraries. Illumina HiSeq 150-bp paired-end sequencing was performed at Genewiz, Inc. (NJ).

Sequence assembly and assessment. To process and assemble reads into full *Xiphinematobacter* genomes, read overlaps were merged using PEAR v0.9.11 (76). We find this step improves read assembly into contigs if performed before downstream steps. Next, merged and unmerged paired reads were filtered and trimmed using Trimmomatic v0.38 (44). Reads (trimmed merged and paired) were then assembled with metaSPAdes v3.13.0 (77, 78) using lower kmers (25, 33, and 45) and higher kmers (65, 87, 101, and 115). Initial assembly quality assessment was performed using QUAST v5.0.1 (79). To isolate all contigs matching *Xiphinematobacter* sp., blastn in BLAST+ v2.10.1 (80) (E value, 10) was used to search for matches to a custom database from the available genome for *Xiphinematobacter* sp. Idaho Grape (GenBank accession CP012665.1). The resulting contigs were extracted and then subjected to a second blastn against the nucleotide database, and all top hit contigs to *Xiphinematobacter* sp. were extracted using custom scripts. Resulting low and high kmer contigs were imported into Geneious Prime v2020.0.4 (Biomatters, Ltd.) and aligned using the ProgressiveMauve v1.1.1 (81) plugin and LASTZ alignment plugin v7.0.2 (Biomatters, Ltd.) to order and orient contigs to one another within each sample and against samples with high sequence identity, and then contigs were inspected for coverage discrepancies (if any) between contigs within a sample. Finally, similar-coverage contigs were laced together with or without regions of Ns relative to high-similarity reference contigs using the consensus among LASTZ-oriented contigs to consolidate regions of overlap between assemblies within each sample to generate long contigs (genomes). Final genome completeness was assessed based on an evaluation of annotated genomes (described below), evaluating the relative presence of housekeeping genes and tRNAs.

Annotation, ortholog detection, pangenomes, and gene repertoire analysis. To compare genome content functionally between *Xiphinematobacter* isolates and related *Verrucomicrobia* sp., genomes were first annotated using Prokka v1.13.3 (82) (E value, 0.001) which uses Prodigal for *ab initio* gene prediction, HMMER3 for protein family profiles, BLAST+ for comparative annotation, Barrnap for rRNAs, and Aragorn for tRNAs. For consistency among gene annotation calls, all genomes (complete or incomplete) were annotated using the same Prokka parameters. In addition to genomes from *Xiphinematobacter* isolates from Texas and New Mexico, we downloaded and annotated the genome from *Xiphinematobacter* sp. Idaho Grape (GenBank accession CP012665.1), a second variant draft genome sequence obtained from the same study, and five outgroup *Verrucomicrobia* genomes from *Spartobacteria bacterium* AMD-G4 (NEUK01000001.1), *Chthoniobacter flavus* Ellin428 (ABVL01000062.1), *Terrimicrobium sacchariphilum* NM-5 (NZ_BDCO01000003), "*Candidatus* Udaeobacter sp. AEW3 Udaeo2_1" (JAALOD010000001.1), and *Akkermansia muciniphila* YL44 (NZ_CP021420.1). Orthologs and pangenomes were obtained using Roary v3.13.0 (83), which uses blastp on the gff files produced by Prokka, using parameters -e for codon-aware alignment using PRANK (84), using -i 60 to allow for distant matches with outgroups, and using the subpackage FastTree v2.1.12 (85) with -gtr -nt < core_gene_alignment.aln > file.new to generate a preliminary tree. Using Roary's gene_presence_absence.csv output, subsets of present/absent genes corresponding to different isolates and outgroups were extracted for the whole genome and metabolic pathway comparisons. Whole-genome presence-absence visualization was performed using Phandango (86). Pangenome overlap was visualized using InteractiVenn (87). Gene assignment to pathways was performed using a custom database based on function annotations primarily from UniProtKB downloads for available specific genomes and supplemented with pathway data from MetaCyc, KEGG pathways, and EMBL-EBI InterPro. Visualization of presence-absence within pathways was performed in R with ggplot2 "melt" and reshape2 packages.

Phylogenetic and phylogenomic analysis. To understand the evolutionary relationships for this symbiosis, phylogenetic analysis was performed on *Xiphinematobacter* sp. and outgroups for nearly full-length 16S rRNA genes and partial 16S rRNA genes and for *Xiphinema* spp. on partial cytochrome oxidase I (COI) genes. Sequences for these analyses were either obtained in this study or downloaded from GenBank (National Center for Biotechnology Information [NCBI]). Genes were aligned with MAFFT v1.0.4 (88). Maximum likelihood phylogenetic analysis was performed using RAxML v4.0 (89) with the GTR Gamma nucleotide model, with rate heterogeneity alpha estimated, and with rapid bootstrapping and search for the best-scoring ML tree (-f a -x 1) with 500 replicates. Bayesian inference phylogenetic analysis was performed using MrBayes v2.2.4 (90, 91) with substitution model GTR+G with 4 categories and with Markov chain Monte Carlo settings as follows: chain length 1,100,000; 4 heated chains; heated chain temperature, 0.2; subsampling frequency, 200; Burn-in length, 100,000; with random seed, 31,569; and priors with unconstrained branch lengths GammaDir (1,0.1,1,1), checking for convergence with minESS of >200. Phylogenomic analysis was performed using the set of core conserved genes from Roary core_gene_alignment.aln (nucleotides aligned by codons) for *Xiphinematobacter* sp. and outgroups. Prior to phylogenetic analysis, alignment positions with gaps in >5% of sequences were removed. Phylogenomic analyses were performed with RAxML MrBayes as described above. Final phylogenetic trees were visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/).

Analysis of nematode communities. To characterize nematode communities in our genome skimming samples, blastn matches between contigs from each of the 11 sample sites to our custom nematode COI database were filtered for matches of >75% identity and 100 bp long, and they were assessed using the kmer coverage conversion to coverage equation $C = (C_K s^2 R)/(R - K + 1)$, where C is total coverage, C_K is kmer coverage, K is the length of kmers, and R is read length. For hits to the same species within a sample, coverages were added (i.e., combining variants) to plot abundances of each species in a proportional bar plot. For classification of hits into *Xiphinema americanum* species complex or non-*X.*

americanum Xiphinema spp., designations were gathered from published classification studies (23, 25, 35, 92, 93), and absolute coverage levels were calculated as above and plotted in a bar plot.

SNP analysis and DBGWAS. To determine if the gene- and pathway-specific mutations in nutritional pathways were in these symbionts, we used the output of *Xiphinematobacter* sp.-only ortholog analysis in Roary, and the core_gene_alignment.aln alignment file was analyzed to extract a variant call format (vcf) of polymorphisms (fixed SNPs within genome isolates) using snp-sites v2.5.1 (94). The number of SNPs per gene was calculated and then normalized for the gene length to obtain relative SNP densities. Within these data, gene assignment to pathways (described above), such as essential amino acid, nonessential amino acid, and vitamin/cofactor or coenzyme biosynthesis, were distinguished and plotted. Box and whisker plots were drawn, and Mann-Whitney U tests were performed in excel.

To determine if host species place differential selection on nutritional pathways in these symbionts, we used de Bruijn-based genome-wide association analysis in DBGWAS (33, 95) for whole annotated nucleotide fasta (genome) files from *Xiphinematobacter* sp. with outgroup *Verrucomicrobia* used as the input along with a phenotype file designating phenotypes as 0 = host clade *Xiphinema luci*, 1 = host clade *Xiphinema rivesi*, and 2 = host clade outgroups. Resulting DBGWAS analyses produced variant clusters or nodes (denoted DBGWAS SNPs) at various levels of significant association with phenotypes. Significant DBGWAS SNP sites distinguishing phenotypes 0 and 1, enriched for fixed variants, were extracted for gene ontology (GO) enrichment analysis (described below) and plotted as described above for total SNPs.

GO enrichment analysis. To determine functional overrepresentation in gene content and SNPs among *Xiphinematobacter* isolates, gene ontology (GO) enrichment analysis was performed on the *Xiphinematobacter* sp.-only pangenome not overlapping with outgroup *Verrucomicrobia*, designated "diff," and the universe gene set included all genes in both sets. Full hierarchical GO annotations for the universe gene set were obtained primarily from UniProtKB and as needed from other databases (MetaCyc and KEGG). Enrichment analyses were performed in topGO v2.4.0 (96) which accounts for GO-term graph topology. TopGO was implemented in R using the script aip_topgo_usage.consider_universe.R (https://github.com/lyijin/topGO_pipeline/). Test statistics were assigned by the "weight01.fisher" algorithm for which returned *P* values are regarded as corrected (or not affected) by multiple testing. GO enrichment analysis in topGO was also performed on the genes with statistically significant DBGWAS SNPs associated with host *Xiphinema* phenotype as the filter on the complete *Xiphinematobacter* diff gene set in topGO.

Analysis of dN/dS and site-specific positive selection. To assess differential selection pressures on proposed symbiont-supplementation pathways, we analyzed within and between *Xiphinematobacter* species signatures of selection based on statistically significant GO biological process terms for nutritional pathways, beginning with the Roary core_gene_alignment.aln alignment file with the corresponding embl headers imported into Geneious. Pathway-associated gene alignment blocks were extracted and concatenated from the full *Xiphinematobacter* alignment. From each pathway subalignment block (e.g., for each set of genes associated with an amino acid or vitamin/cofactor biosynthetic pathway identified in DBGWAS) as well as for the remaining genome, dN/dS was calculated using the Nei and Gojobori 1986 estimator within PAML 4 CodeML (97), implemented through EasyCodeML v1.21 (98). For the same alignment blocks, we used CodeML with the Site Model analysis variance across the sequence blocks by applying the likelihood ratio test (LRT) for positive selection (ω), in which multiple nested models were evaluated. In particular, the LRT tests for selecting M8 over M7 [M7 beta; $\omega \sim B(p,q)$ with 2 free parameters, and 10 categories or equal classes; versus M8 beta& ω; with proportion p0 of sites $\sim B(p,q)$, p1 of sites from discrete ω class with 4 free parameters, and 11 categories] which has been shown to be a very stringent test of positive selection (97–99).

Data availability. The raw SRA data are deposited in GenBank under BioProject accession number PRJNA687334, and the genomes are in GenBank under BioSample accession numbers SAMN17141338, SAMN17141340, SAMN17141343, SAMN17141344, SAMN17141345, and SAMN17141353.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
FIG S1, EPS file, 1.6 MB.
FIG S2, EPS file, 0.7 MB.
FIG S3, EPS file, 0.8 MB.
FIG S4, EPS file, 0.6 MB.
TABLE S1, DOCX file, 0.01 MB.
TABLE S2, DOCX file, 0.01 MB.
TABLE S3, XLSX file, 1.1 MB.
TABLE S4, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Hansen AK, Moran NA. 2014. The impact of microbial symbionts on host plant utilization by herbivorous insects. Mol Ecol 23:1473–1496. https://doi.org/10.1111/mec.12421.
2. Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165–190. https://doi.org/10.1146/annurev.genet.41.110306.130119.
3. Bennett GM, Moran NA. 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. Genome Biol Evol 5:1675–1688. https://doi.org/10.1093/gbe/evt118.
4. Toft C, Andersson SGE. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. Nat Rev Genet 11:465–475. https://doi.org/10.1038/nrg2798.
5. McCutcheon JP, Moran NA. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. Genome Biol Evol 2:708–718. https://doi.org/10.1093/gbe/evq055.
6. Moran NA, Bennett GM. 2014. The tiniest tiny genomes. Annu Rev Microbiol 68:195–215. https://doi.org/10.1146/annurev-micro-091213-112901.
7. Brown AV. 2018. Endosymbionts of plant-parasitic nematodes. Annu Rev Phytopathol 56:225–242. https://doi.org/10.1146/annurev-phyto-080417-045824.
8. Brown AMV, Howe DK, Wasala SK, Peetz AB, Zasada IA, Denver DR. 2015. Comparative genomics of a plant-parasitic nematode endosymbiont suggest a role in nutritional symbiosis. Genome Biol Evol 7:2727–2746. https://doi.org/10.1093/gbe/evv176.
9. Palomares-Rius JE, Archidona-Yuste A, Cantalapiedra-Navarrete C, Prieto P, Castillo P. 2016. Molecular diversity of bacterial endosymbionts associated with dagger nematodes of the genus Xiphinema (Nematoda: Longidoridae) reveals a high degree of phylogenetic congruence with their host. Mol Ecol 25:6225–6247. https://doi.org/10.1111/mec.13904.
10. Lazarova SS, Brown DJF, Oliveira CMG, Fenton B, MacKenzie K, Wright F, Malloch G, Neilson R. 2016. Diversity of endosymbiont bacteria associated with a non-filarial nematode group. Nematol 18:615–623. https://doi.org/10.1163/15685411-00002982.
11. Howe DK, Smith M, Tom DM, Brown AMV, Peetz AB, Zasada IA, Denver DR. 2019. Analysis of nematode-endosymbiont coevolution in the Xiphinema americanum species complex using molecular markers of variable evolutionary rates. Nematol 21:533–546. https://doi.org/10.1163/15685411-00003233.
12. Jones JT, Haegeman A, Danchin EGJ, Gaur HS, Helder J, Jones MGK, Kikuchi T, Manzanilla-López R, Palomares-Rius JE, Wesemael WML, Perry RN. 2013. Top 10 plant-parasitic nematodes in molecular plant pathology. Mol Plant Pathol 14:946–961. https://doi.org/10.1111/mpp.12057.
13. Jeger M, Bragard C, Caffier D, Candresse T, Chatzivassiliou E, Dehnen-Schmutz K, Gilioli G, Grégoire JC, Miret JA, MacLeod A, Navarro M, Parnell S, Potting R, Rafoss T, Rossi V, Urek G, Van Bruggen A, Van der Werf W, West J, Winter S, Kaluski T, Niere B. 2018. Pest categorisation of Xiphinema americanum sensu lato. EFSA J 16:5298. https://doi.org/10.2903/j.efsa.2018.5298.
14. Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. Proc Natl Acad Sci U S A 115:6506–6511. https://doi.org/10.1073/pnas.1711842115.
15. Ferris H, Griffiths BS, Porazinska DL, Powers TO, Wang K-H, Tenuta M. 2012. Reflections on plant and soil nematode ecology: past, present and future. J Nematol 44:115–126.
16. Ferris H, Tuomisto H. 2015. Unearthing the role of biological diversity in soil health. Soil Biol Biochem 85:101–109. https://doi.org/10.1016/j.soilbio.2015.02.037.
17. Bongers T, Ferris H. 1999. Nematode community structure as a bioindicator in environmental monitoring. Trends Ecol Evol 14:224–228. https://doi.org/10.1016/s0169-5347(98)01583-3.
18. Beare MH. 1997. Fungal and bacterial pathways of organic matter decomposition and nitrogen mineralization in arable soils, p 37–70. In Lijbert Brussaard RF-C (ed), Soil ecology in sustainable agricultural systems. CRC Press, Boca Raton, FL.
19. Neher DA. 2001. Role of nematodes in soil health and their use as indicators. J Nematol 33:161–168.
20. van den Hoogen J, Geisen S, Routh D, Ferris H, Traunspurger W, Wardle DA, de Goede RGM, Adams BJ, Ahmad W, Andriuzzi WS, Bardgett RD, Bonkowski M, Campos-Herrera R, Cares JE, Caruso T, de Brito Caixeta L, Chen X, Costa SR, Creamer R, Mauro da Cunha Castro J, Dam M, Djigal D, Escuer M, Griffiths BS, Gutiérrez C, Hohberg K, Kalinkina D, Kardol P, Kergunteuil A, Korthals G, Krashevska V, Kudrin AA, Li Q, Liang W, Magilton M, Marais M, Martín JAR, Matveeva E, Mayad EH, Mulder C, Mullin P, Neilson R, Nguyen TAD, Nielsen UN, Okada H, Rius JEP, Pan K, Peneva V, Pellissier L, Carlos Pereira da Silva J, et al. 2019. Soil nematode abundance and functional group composition at a global scale. Nature 572:194–198. https://doi.org/10.1038/s41586-019-1418-6.
21. Nicol JM, Turner SJ, Coyne DL, den Nijs L, Hockland S, Tahna Maafi Z. 2011. Current nematode threats to world agriculture, p 21–44. In Jones JT, Gheysen G, Fenoll C (ed), Genomics and molecular genetics of plant-nematode interactions. Springer, Heidelberg, Germany.
22. Lazarova SS, Malloch G, Oliveira CMG, Hübschen J, Neilson R. 2006. Ribosomal and mitochondrial DNA analyses of Xiphinema americanum-group populations. J Nematol 38:404–410.
23. Zasada IA, Peetz A, Howe DK, Wilhelm LJ, Cheam D, Denver DR, Smythe AB. 2014. Using mitogenomic and nuclear ribosomal sequence data to investigate the phylogeny of the Xiphinema americanum species complex. PLoS One 9:e90035. https://doi.org/10.1371/journal.pone.0090035.
24. Koenning SR, Overstreet C, Noling JW, Donald PA, Becker JO, Fortnum BA. 1999. Survey of crop losses in response to phytoparasitic nematodes in the United States for 1994. J Nematol 31:587–618.
25. Robbins RT. 1993. Distribution of Xiphinema americanum and related species in North America. J Nematol 25:344–348.
26. Lamberti F, Ciancio A. 1993. Diversity of Xiphinema americanum-group species and hierarchical cluster analysis of morphometrics. J Nematol 25:332–343.
27. McKenry M. 1992. Nematodes, p 281–285. In Flaherty DL, Christensen LP, Lanini WT, Marois JJ, Phillips P, Wilson LT (ed), Grape pest management. University of California Press, Oakland, CA.
28. Denver DR, Brown AMV, Howe DK, Peetz AB, Zasada IA. 2016. Genome skimming: a rapid approach to gaining diverse biological insights into multicellular pathogens. PLoS Pathog 12:e1005713. https://doi.org/10.1371/journal.ppat.1005713.
29. de Araujo ASF, Bezerra WM, dos Santos VM, Rocha SMB, Carvalho NS, de Lyra MCC, Figueiredo MVB, de Almeida Lopes ÂC, Melo VMM. 2017. Distinct bacterial communities across a gradient of vegetation from a preserved Brazilian Cerrado. Antonie Van Leeuwenhoek 110:457–469. https://doi.org/10.1007/s10482-016-0815-1.
30. Bergmann GT, Bates ST, Eilers KG, Lauber CL, Caporaso JG, Walters WA, Knight R, Fierer N. 2011. The under-recognized dominance of Verrucomicrobia in soil bacterial communities. Soil Biol Biochem 43:1450–1455. https://doi.org/10.1016/j.soilbio.2011.03.012.
31. Mordechai M, Cohn E. 1969. Investigations on the life cycles and host preference of some species of Xiphinema and Longidorus under controlled conditions. Nematol 15:295–302. https://doi.org/10.1163/187529269X00335.
32. Lownsbery BF, Mitchell JT. 1965. Some effects of chemical amendments

and cultural conditions on population levels of *Xiphinema americanum*. Plant Dis Rep 49:994–998.

33. Jaillard M, Lima L, Tournoud M, Mahé P, van Belkum A, Lacroix V, Jacob L. 2018. A fast and agnostic method for bacterial genome-wide association studies: bridging the gap between k-mers and genetic events. PLoS Genet 14:e1007758. https://doi.org/10.1371/journal.pgen.1007758.

34. Mobasseri M, Hutchinson MC, Afshar FJ, Pedram M. 2019. New evidence of nematode-endosymbiont bacteria coevolution based on one new and one known dagger nematode species of *Xiphinema americanum*-group (Nematoda, Longidoridae). PLoS One 14:e0217506. https://doi.org/10.1371/journal.pone.0217506.

35. Archidona-Yuste A, Navas-Cortés JA, Cantalapiedra-Navarrete C, Palomares-Rius JE, Castillo P. 2016. Cryptic diversity and species delimitation in the *Xiphinema americanum*-group complex (Nematoda: Longidoridae) as inferred from morphometrics and molecular markers. Zool J Linn Soc 176:231–265. https://doi.org/10.1111/zoj.12316.

36. Brown AMV, Wasala SK, Howe DK, Peetz AB, Zasada IA, Denver DR. 2016. Genomic evidence for plant-parasitic nematodes as the earliest *Wolbachia* hosts. Sci Rep 6:34955. https://doi.org/10.1038/srep34955.

37. Brown AMV, Wasala SK, Howe DK, Peetz AB, Zasada IA, Denver DR. 2018. Comparative genomics of *Wolbachia*–*Cardinium* dual endosymbiosis in a plant-parasitic nematode. Front Microbiol 9:2482. https://doi.org/10.3389/fmicb.2018.02482.

38. Brown AMV, Huynh LY, Bolender CM, Nelson KG, McCutcheon JP. 2014. Population genomics of a symbiont in the early stages of a pest invasion. Mol Ecol 23:1516–1530. https://doi.org/10.1111/mec.12366.

39. Guo X, Li S, Zhang J, Wu F, Li X, Wu D, Zhang M, Ou Z, Jie Z, Yan Q, Li P, Yi J, Peng Y. 2017. Genome sequencing of 39 *Akkermansia muciniphila* isolates reveals its population structure, genomic and functional diversity, and global distribution in mammalian gut microbiotas. BMC Genomics 18:800. https://doi.org/10.1186/s12864-017-4195-3.

40. Li Y, Pinto-Tomás AA, Rong X, Cheng K, Liu M, Huang Y. 2019. Population genomics insights into adaptive evolution and ecological differentiation in Streptomycetes. Appl Environ Microbiol 85:e02555-18. https://doi.org/10.1128/AEM.02555-18.

41. Choi YJ, Tyagi R, McNulty SN, Rosa BA, Ozersky P, Martin J, Hallsworth-Pepin K, Unnasch TR, Norice CT, Nutman TB, Weil GJ, Fischer PU, Mitreva M. 2016. Genomic diversity in *Onchocerca volvulus* and its *Wolbachia* endosymbiont. Nat Microbiol 2:16207. https://doi.org/10.1038/nmicrobiol.2016.207.

42. Tsementzi D, Castro Gordillo J, Mahagna M, Gottlieb Y, Konstantinidis KT. 2018. Comparison of closely related, uncultivated *Coxiella* tick endosymbiont population genomes reveals clues about the mechanisms of symbiosis. Environ Microbiol 20:1751–1764. https://doi.org/10.1111/1462-2920.14104.

43. Sheppard SK, Guttman DS, Fitzgerald JR. 2018. Population genomics of bacterial host adaptation. Nat Rev Genet 19:549–565. https://doi.org/10.1038/s41576-018-0032-z.

44. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

45. Orlando V, Chitambar JJ, Dong K, Chizhov VN, Mollov D, Bert W, Subbotin SA. 2016. Molecular and morphological characterisation of *Xiphinema americanum*-group species (Nematoda: Dorylaimida) from California, USA, and other regions, and co-evolution of bacteria from the genus *Candidatus* Xiphinematobacter with nematodes. Nematol 18:1015–1043. https://doi.org/10.1163/15685411-00003012.

46. Bell-Roberts L, Douglas AE, Werner GDA. 2019. Match and mismatch between dietary switches and microbial partners in plant sap-feeding insects. Proc Biol Sci 286:20190065. https://doi.org/10.1098/rspb.2019.0065.

47. Sandström J, Moran N. 2001. Amino acid budgets in three aphid species using the same host plant. Physiol Entomol 26:202–211. https://doi.org/10.1046/j.0307-6962.2001.00235.x.

48. Russell CW, Poliakov A, Haribal M, Jander G, van Wijk KJ, Douglas AE. 2014. Matching the supply of bacterial nutrients to the nutritional demand of the animal host. Proc Biol Sci 281:20141163. https://doi.org/10.1098/rspb.2014.1163.

49. Vandekerckhove TTM, Coomans A, Cornelis K, Baert P, Gillis M. 2002. Use of the Verrucomicrobia-specific probe EUB338-III and fluorescent in situ hybridization for fetection of "*Candidatus* Xiphinematobacter" cells in nematode hosts. Appl Environ Microbiol 68:3121–3125. https://doi.org/10.1128/aem.68.6.3121-3125.2002.

50. Blaxter M, Koutsovoulos G. 2015. The evolution of parasitism in Nematoda. Parasitology 142:S26–S39. https://doi.org/10.1017/S0031182014000791.

51. Luc M, Coomans A, P A A L, Baujard P. 1998. The *Xiphinema americanum*-group (Nematoda: Longidoridae). 2. Observations on *Xiphinema brevicollum* Lordello & da Costa, 1961 and comments on the group. Fundam Appl Nematol 21:475–490.

52. Coomans A, Claeys M. 1998. Structure of the female reproductive system of *Xiphinema americanum* (Nematoda: Longidoridae. Fundam Appl Nematol 21:569–580.

53. Vandekerckhove TTM, Willems A, Gillis M, Coomans A. 2000. Occurrence of novel verrucomicrobial species, endosymbiotic and associated with parthenogenesis in *Xiphinema americanum*-group species (Nematoda, Longidoridae). Int J Syst Evol Microbiol 50:2197–2205. https://doi.org/10.1099/00207713-50-6-2197.

54. Cohn E. 1970. Observations on the feeding and symptomatology of *Xiphinema* and *Longidorus* on selected host roots. J Nematol 2:167–173.

55. Laber B, Gomis-Ruth FX, Romao MJ, Huber R. 1992. Escherichia coli dihydrodipicolinate synthase. Biochem J 288:691–695. https://doi.org/10.1042/bj2880691.

56. Lee H-S, Hwang B-J. 2003. Methionine biosynthesis and its regulation in Corynebacterium glutamicum: parallel pathways of transsulfuration and direct sulfhydrylation. Appl Microbiol Biotechnol 62:459–467. https://doi.org/10.1007/s00253-003-1306-7.

57. Russell JA, Oliver KM, Hansen AK. 2017. Band-aids for Buchnera and B vitamins for all. Mol Ecol 26:2199–2203. https://doi.org/10.1111/mec.14047.

58. Mobegi VA, Duffy CW, Amambua-Ngwa A, Loua KM, Laman E, Nwakanma DC, MacInnis B, Aspeling-Jones H, Murray L, Clark TG, Kwiatkowski DP, Conway DJ. 2014. Genome-wide analysis of selection on the malaria parasite *Plasmodium falciparum* in West African populations of differing infection endemicity. Mol Biol Evol 31:1490–1499. https://doi.org/10.1093/molbev/msu106.

59. Castillo JA, Agathos SN. 2019. A genome-wide scan for genes under balancing selection in the plant pathogen *Ralstonia solanacearum*. BMC Evol Biol 19:123. https://doi.org/10.1186/s12862-019-1456-6.

60. Chung SH, Parker BJ, Blow F, Brisson JA, Douglas AE. 2020. Host and symbiont genetic determinants of nutritional phenotype in a natural population of the pea aphid. Mol Ecol 29:848–858. https://doi.org/10.1111/mec.15355.

61. Yoder JB. 2016. Understanding the coevolutionary dynamics of mutualism with population genomics. Am J Bot 103:1742–1752. https://doi.org/10.3732/ajb.1600154.

62. Kim D, Thairu MW, Hansen AK. 2016. Novel insights into insect-microbe interactions—role of epigenomics and small RNAs. Front Plant Sci 7:1164. https://doi.org/10.3389/fpls.2016.01164.

63. Price DRG, Feng H, Baker JD, Bavan S, Luetje CW, Wilson ACC. 2014. Aphid amino acid transporter regulates glutamine supply to intracellular bacterial symbionts. Proc Natl Acad Sci U S A 111:320–325. https://doi.org/10.1073/pnas.1306068111.

64. Duncan RP, Husnik F, Van Leuven JT, Gilbert DG, Dávalos LM, McCutcheon JP, Wilson ACC. 2014. Dynamic recruitment of amino acid transporters to the insect/symbiont interface. Mol Ecol 23:1608–1623. https://doi.org/10.1111/mec.12627.

65. Gerth M, Bleidorn C. 2016. Comparative genomics provides a timeframe for *Wolbachia* evolution and exposes a recent biotin synthesis operon transfer. Nat Microbiol 2:16241. https://doi.org/10.1038/nmicrobiol.2016.241.

66. Treonis AM, Unangst SK, Kepler RM, Buyer JS, Cavigelli MA, Mirsky SB, Maul JE. 2018. Characterization of soil nematode communities in three cropping systems through morphological and DNA metabarcoding approaches. Sci Rep 8:2004. https://doi.org/10.1038/s41598-018-20366-5.

67. Zheng F, Zhu D, Chen Q-L, Bi Q-F, Yang X-R, O'Connor P, Zhu Y-G. 2020. The driving factors of nematode gut microbiota under long-term fertilization. FEMS Microbiol Ecol 96:fiaa037. https://doi.org/10.1093/femsec/fiaa037.

68. Augustinos AA, Santos-Garcia D, Dionyssopoulou E, Moreira M, Papapanagiotou A, Scarvelakis M, Doudoumis V, Ramos S, Aguiar AF, Borges PV, Khadem M, Latorre A, Tsiamis G, Bourtzis K. 2011. Detection and characterization of *Wolbachia* infections in natural populations of aphids: is the hidden diversity fully unraveled? PLoS One 6:e28695. https://doi.org/10.1371/journal.pone.0028695.

69. Takano S, Tuda M, Takasu K, Furuya N, Imamura Y, Kim S, Tashiro K, Iiyama K, Tavares M, Amaral AC. 2017. Unique clade of alphaproteobacterial endosymbionts induces complete cytoplasmic incompatibility in the coconut beetle. Proc Natl Acad Sci U S A 114:6110–6115. https://doi.org/10.1073/pnas.1618094114.

70. Cao Y, Tian B, Ji X, Shang S, Lu C, Zhang K. 2015. Associated bacteria of

different life stages of Meloidogyne incognita using pyrosequencing-based analysis. J Basic Microbiol 55:950–960. https://doi.org/10.1002/jobm.201400816.

71. Stalder T, Press MO, Sullivan S, Liachko I, Top EM. 2019. Linking the resistome and plasmidome to the microbiome. ISME J 13:2437–2446. https://doi.org/10.1038/s41396-019-0446-4.

72. Marbouty M, Cournac A, Flot JF, Marie-Nelly H, Mozziconacci J, Koszul R. 2014. Metagenomic chromosome conformation capture (meta3C) unveils the diversity of chromosome organization in microorganisms. Elife 3:e03318. https://doi.org/10.7554/eLife.03318.

73. Kang JB, Siranosian BA, Moss EL, Banaei N, Andermann TM, Bhatt AS. 2019. Intestinal microbiota domination under extreme selective pressures characterized by metagenomic read cloud sequencing and assembly. BMC Bioinformatics 20:585. https://doi.org/10.1186/s12859-019-3073-1.

74. Lamberti F, Bleve-Zacheo T. 1979. Studies on Xiphinema americanum sensu lato with descriptions of fifteen new species (Nematoda, Longidoridae). Nematol Mediterr 7:51–106.

75. Jenkins WR. 1964. A rapid centrifugal-flotation technique for separating nematodes from soil. Plant Dis Report 48:692.

76. Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30:614–620. https://doi.org/10.1093/bioinformatics/btt593.

77. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. MetaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116.

78. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

79. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

80. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.

81. Darling AE, Mau B, Perna NT. 2010. Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.

82. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

83. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

84. Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. Methods Mol Biol 1079:155–170. https://doi.org/10.1007/978-1-62703-646-7_10.

85. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. https://doi.org/10.1371/journal.pone.0009490.

86. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. 2018. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics 34:292–293. https://doi.org/10.1093/bioinformatics/btx610.

87. Heberle H, Meirelles VG, da Silva FR, Telles GP, Minghim R. 2015. InteractiVenn: a Web-based tool for the analysis of sets through Venn diagrams. BMC Bioinformatics 16:169. https://doi.org/10.1186/s12859-015-0611-3.

88. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010.

89. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

90. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029.

91. Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

92. Cai R, Archidona-Yuste A, Cantalapiedra-Navarrete C, Palomares-Rius JE, Zheng J, Castillo P. 2019. Integrative taxonomy of Xiphinema histriae and Xiphinema lapidosum from Spain. J Nematol 51:1–21. https://doi.org/10.21307/jofnem-2019-037.

93. Oliveira CMG, Hübschen J, Brown DJF, Ferraz LCCB, Wright F, Neilson R. 2004. Phylogenetic relationships among Xiphinema and Xiphidorus nematode species from Brazil inferred from 18S rDNA sequences. J Nematol 36:153–159.

94. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb Genom 2:e000056. https://doi.org/10.1099/mgen.0.000056.

95. Leiserson MDM, Eldridge JV, Ramachandran S, Raphael BJ. 2013. Network analysis of GWAS data. Curr Opin Genet Dev 23:602–610. https://doi.org/10.1016/j.gde.2013.09.003.

96. Alexa A, Rahnenfuhrer J. 2020. topGO: enrichment analysis for Gene Ontology R package version 2.40.0.

97. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591. https://doi.org/10.1093/molbev/msm088.

98. Gao F, Chen C, Arab DA, Du Z, He Y, Ho SYW. 2019. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol Evol 9:3891–3898. https://doi.org/10.1002/ece3.5015.

99. Anisimova M, Bielawski JP, Yang Z. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol Biol Evol 18:1585–1592. https://doi.org/10.1093/oxfordjournals.molbev.a003945.