



OPEN ACCESS

EDITED BY

Akiko Sugio,
Institut National de la Recherche Agronomique,
France

REVIEWED BY

Hiroshi Arai,
National Agriculture and Food Research
Organization (NARO), Japan
Emily Hornett,
University of Liverpool, United Kingdom

*CORRESPONDENCE

Logan D. Moore
✉ ldm427@msstate.edu

RECEIVED 19 January 2023

ACCEPTED 19 April 2023

PUBLISHED 18 May 2023

CITATION

Moore LD and Ballinger MJ (2023) The toxins of vertically transmitted *Spiroplasma*.
Front. Microbiol. 14:1148263.
doi: 10.3389/fmicb.2023.1148263

COPYRIGHT

© 2023 Moore and Ballinger. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The toxins of vertically transmitted *Spiroplasma*

Logan D. Moore* and Matthew J. Ballinger

Department of Biological Sciences, Mississippi State University, Mississippi State, MS, United States

Vertically transmitted (VT) microbial symbionts play a vital role in the evolution of their insect hosts. A longstanding question in symbiont research is what genes help promote long-term stability of vertically transmitted lifestyles. Symbiont success in insect hosts is due in part to expression of beneficial or manipulative phenotypes that favor symbiont persistence in host populations. In *Spiroplasma*, these phenotypes have been linked to toxin and virulence domains among a few related strains. However, these domains also appear frequently in phylogenetically distant *Spiroplasma*, and little is known about their distribution across the *Spiroplasma* genus. In this study, we present the complete genome sequence of the *Spiroplasma* symbiont of *Drosophila atripex*, a non-manipulating member of the Ixodetis clade of *Spiroplasma*, for which genomic data are still limited. We perform a genus-wide comparative analysis of toxin domains implicated in defensive and reproductive phenotypes. From 12 VT and 31 non-VT *Spiroplasma* genomes, ribosome-inactivating proteins (RIPs), OTU-like cysteine proteases (OTUs), ankyrins, and ETX/MTX2 domains show high propensity for VT *Spiroplasma* compared to non-VT *Spiroplasma*. Specifically, OTU and ankyrin domains can be found only in VT-*Spiroplasma*, and RIP domains are found in all VT *Spiroplasma* and three non-VT *Spiroplasma*. These domains are frequently associated with *Spiroplasma* plasmids, suggesting a possible mechanism for dispersal and maintenance among heritable strains. Searching insect genome assemblies available on public databases uncovered uncharacterized *Spiroplasma* genomes from which we identified several *spaid*-like genes encoding RIP, OTU, and ankyrin domains, suggesting functional interactions among those domain types. Our results suggest a conserved core of symbiont domains play an important role in the evolution and persistence of VT *Spiroplasma* in insects.

KEYWORDS

Spiroplasma, Ixodetis, ETX/MTX2, OTU-like cysteine protease, ribosome-inactivating proteins, ankyrin, vertical transmission

Introduction

Gene duplications, losses, and horizontal transfers can facilitate dramatic shifts in bacterial lifestyle and capabilities (Romero and Palacios, 1997; Moran, 2002; Arnold et al., 2022). Gene loss is a dominant feature of symbiont evolution due to the selective benefits of removing metabolically costly genes (McCutcheon and Moran, 2012; McCutcheon et al., 2019). Alternatively, gene gain via horizontal transfer and duplication can lead to rapid adaptation across symbiotic species. For example, several mutualistic soil bacteria with nitrogen-fixing capabilities have benefited from receiving “symbiont islands” and plasmids enriched with nitrogen-fixing genes (Sullivan and Ronson, 1998; Andrews et al., 2018). Vertically transmitted symbionts are no exception to this phenomenon and can acquire novel phenotypes that benefit

a heritable lifestyle via the transfer of symbiosis-supporting genes (Ballinger et al., 2019; Manzano-Marin et al., 2020; Waterworth et al., 2020).

Vertically transmitted microbial symbionts can strongly influence the life history traits of their insect hosts to maintain stable vertical transmission (VT) (Moran et al., 2008). For example, the facultative endosymbiont *Wolbachia* can defend against viral infections in their *Drosophila* hosts (Teixeira et al., 2008) but can also be penetrant reproductive parasites of *Drosophila*, using manipulative phenotypes like male-killing and cytoplasmic incompatibility to favor their transmission to subsequent generations. Heritable symbionts are ubiquitous among insects (Moran et al., 2008), and yet, a central question remains for many heritable microbes—what gene effectors promote vertical transmission and how do they arise?

Spiroplasma is a genus of helical, cell wall-less bacteria estimated to infect up to 7% of terrestrial arthropod species (Duron et al., 2008). The genus is phylogenetically structured into three large clades—the Apis clade, the Citri-Chrysopicola-Mirum clade, and the Ixodetis clade (Gasparich et al., 2004). Despite their rich diversity, all *Spiroplasma* are non-free-living bacteria, infecting the gut or hemolymph of invertebrates and/or the phloem of plants. At ecological timescales, *Spiroplasma* infections are transmitted horizontally by ingestion or maternally via the cytoplasm of mature oocytes (Herren et al., 2013). The Apis clade contains several insect pathogens, including *Spiroplasma taiwanense* and *Spiroplasma culicicola* of mosquitoes and *Spiroplasma apis* of European honey-bees (Mouches et al., 1982; Humphery-Smith et al., 1991; Bolaños et al., 2015). No VT *Spiroplasma* have been described in the Apis clade, while both VT and non-VT *Spiroplasma* have been reported in the Citri-Chrysopicola-Mirum clade (hereafter referred to as the Citri clade) and the Ixodetis clade (Jiggins et al., 2000; Haselkorn, 2010; Ballinger et al., 2018; Binetruy et al., 2019; Vera-Ponce León et al., 2021).

Heritable *Spiroplasma* have been described across insect orders, and are well-studied in vinegar flies, butterflies, beetles, aphids and parasitoid wasps. Some members of the Citri clade defend *Drosophila* hosts against parasitic nematodes and parasitoid wasps (Xie et al., 2010; Jaenike et al., 2010b; Haselkorn and Jaenike, 2015). Others in this clade are both male-killers and defensive symbionts of *Drosophila* (Montenegro et al., 2005; Xie et al., 2014). Reproductive manipulation is also common among Ixodetis clade *Spiroplasma*. Ixodetis *Spiroplasma* of the butterfly *Danaus chrysippus* (*sChrys*) and the ladybird beetle *Anisosticta novemdecimpunctata* are effective VT male-killers (Jiggins et al., 2000; Tinsley and Majerus, 2006) and the recently discovered *Spiroplasma* symbiont that causes cytoplasmic incompatibility in the parasitoid wasp *Lariophagus distinguendus* (*sDis*) is also an Ixodetis clade member (Pollmann et al., 2022). Cytoplasmic incompatibility is a form of reproductive manipulation whereby toxic or modified sperm of symbiont-infected fathers kills the offspring of uninfected mothers, thus driving symbiont spread through the population. As defensive symbionts, Ixodetis *Spiroplasma* have been shown to confer resistance to fungal infections in their aphid hosts (Łukasik et al., 2013). Indeed, VT strains are present in divergent *Spiroplasma* clades, yet little is known about the genetic mechanisms supporting their host-associated phenotypes beyond those of the Poulsonii group in the Citri clade. Within the Poulsonii group, symbiont toxins have been implicated in both defensive and reproductive phenotypes (Hamilton et al., 2016; Ballinger and Perlman, 2017; Harumoto and Lemaitre, 2018).

Toxic peptides are known effectors of defensive and reproductive phenotypes across symbiotic bacterial taxa (Gillespie et al., 2018; Oliver and Perlman, 2020; Massey and Newton, 2022). *Hamiltonella defensa* utilizes phage-encoded toxins to defend aphid hosts against parasitoid wasps (Degnan and Moran, 2008). *Wolbachia* symbionts employ a toxin-antitoxin or host modification system to inflict cytoplasmic incompatibility in their hosts (Beckmann et al., 2017; LePage et al., 2017; Horard et al., 2022; Kaur et al., 2022). Within the Citri-based Poulsonii group, including *Spiroplasma* strains infecting *Drosophila hydei*, *Drosophila neotestacea*, and *Drosophila melanogaster* (*sHyd*, *sNeo*, and MSRO respectively), defensive strains utilize ribosome-inactivating protein toxins (RIPs) to contribute to defense against parasitoid wasps and parasitic nematodes (Hamilton et al., 2016; Ballinger and Perlman, 2017). RIPs cause irreversible damage to a key enzymatic residue of ribosomes, resulting in cell death (Stirpe, 2004). RIP gene families are remarkably diverse in Citri clade *Spiroplasma*, ostensibly due to extensive horizontal transmission, duplication, loss, and recombination (Ballinger et al., 2019; Gerth et al., 2021). *Spiroplasma poulsonii* strain MSRO is a reproductive manipulator that encodes a multidomain toxin called Spaid that causes male-killing in *Drosophila* (Harumoto and Lemaitre, 2018). Spaid consists of three integral domains including ankyrin repeat domains, an OTU-like cysteine protease domain (OTU) with predicted function in deubiquitination, and a C-terminal transmembrane domain. OTUs and ankyrin domains are also associated with type V cifB-like proteins of CI-inducing *Wolbachia*, suggesting they can be adapted for different reproductive phenotypes in diverse systems (Martinez et al., 2021). While toxins have been identified to play an essential role in maintaining symbiosis in these heritable Citri clade *Spiroplasma*, far less is known about effector genes supporting the diverse Ixodetis *Spiroplasma* phenotypes. However, recent advances in genomic representation among these strains (Yeoman et al., 2019; Martin et al., 2020; Vera-Ponce León et al., 2021; Pollmann et al., 2022) facilitates an in-depth comparative analysis of toxin gene content in this enigmatic clade.

In this study, we report the genome sequence of a *Drosophila*-associated Ixodetis clade member, the *Spiroplasma* symbiont of *Drosophila atriplex* (*sAtri*). This strain does not induce male-killing or cytoplasmic incompatibility in *Drosophila* (Haselkorn et al., 2009) but does provide protection against parasitoid wasps in the family Figitidae (T. Chris Amuwa, unpublished results). We identify a diverse set of toxin and virulence domains in the genome, including the same domains contributing to defense and reproductive manipulation in the Citri clade. We conducted a genus-wide survey of 43 *Spiroplasma* genomes across all three clades to characterize the distribution of toxin domains between heritable and non-heritable *Spiroplasma*. Our results uncover a striking association of RIPs, OTUs, ankyrin, and ETX/MTX2 domains with heritable strains, suggesting that these domains play an important role in promoting VT symbiosis across *Spiroplasma*.

Results

Genome of *sAtri*

Using a hybrid long and short read sequencing strategy, we produced a closed genome assembly of 1.27 Mb for *sAtri*. The

genome is covered to 29X depth by short reads (Supplementary Figure S1). Phylogenomic analysis nests *sAtri* with strong support among other Ixodetis clade endosymbiont members that infect a wide range of insect hosts (Figure 1). Genome completeness based on single copy orthologs was estimated by BUSCO at 98% (Table 1). Missing orthologs from the genome completeness analysis include the genes *rplL* and *rsmL*. We identified *rplL* in the *sAtri* genome with a manual BLAST search but were unable to identify *rsmL*. While *rsmL* (rRNA ribose-2'-O-methyltransferase) has a conserved role in bacterial translation, it is absent from all currently available Ixodetis clade genomes (Pollmann et al., 2022). The hybrid genome assembly includes eight putative plasmids, which were assigned as such based on gene content, coverage depth, and circularity (Supplementary Tables S1, S2). Repeat content is 10.73%, but only nine transposable elements could be annotated. Of the 1,313 putative protein coding genes, 639 could be assigned a functional annotation by prokka. *sAtri* encodes complete ATP synthase and glycolysis pathways, as well as transport components for glucose, fructose, GlcNAc, glycerol, and possibly mannose. Summaries of functional predictions and carbon metabolism pathway components are shown in Supplementary Figure S2A and Supplementary Table S3. Amino acid and lipid biosynthesis pathways are incomplete or absent, as reported previously for hemolymph-dwelling *Spiroplasma* species, in both the Citri and Ixodetis clades (Paredes et al., 2015; Yeoman et al., 2019; Vera-Ponce León et al., 2021). In addition to the PTS genes for carbon import (Supplementary Figure S2B), ABC membrane proteins for transport of phosphate (*ptsABCs*), nucleosides (*bmpA*, *nupABC*), and micronutrients (*ecfA1*, *ecfA2*) could be identified. A comparative summary of membrane transport, biosynthesis, and metabolism genes in *sAtri* and two other Ixodetis clade genomes is present in Supplementary Table S3. *sAtri* also encodes a Type II-C CRISPR/Cas9 system which has not been previously reported in a heritable *Spiroplasma*. No matches between *sAtri* spacers and known phages or phage regions were identified. However, several of the spacers in *sAtri* are exact matches to sequences present on *sAtri* plasmids, suggesting a possible interaction with CRISPR/Cas9 (e.g.,

plasmid copy number suppression). *sAtri* spacers also match to non-phage genomic sequences of other Ixodetis *Spiroplasma*, though it is unclear if those sequences are associated with plasmids.

sAtri encodes a diverse set of toxins and virulence genes

The *sAtri* genome assembly reveals similarities to and departures from other *Spiroplasma* with regard to toxin and virulence gene content. Three RIP domain-containing proteins are present, two of which encode a signal peptide for secretion. RIP toxins contribute to defense against natural enemies in some Citri clade members (Hamilton et al., 2016; Ballinger and Perlman, 2017). The *sAtri* RIPs are phylogenetically distant from one another (Figure 2) and from other non-Ixodetis *Spiroplasma* RIPs with less than 30% similarity at the amino acid level. We also identified RIPs in the genome assemblies of the *Spiroplasma* symbiont of *Danaus chrysippus* (*sChrys*) and the *Spiroplasma* symbiont of *Cephus cinctus* (*sCinc*), and retrieved the RIPs reported in the *Spiroplasma* symbiont of *Lariophagus distinguendus* (*sDis*) and the *Spiroplasma* symbiont of *Dactylopius coccus* (*sCoccus*). Examining the phylogenetic relationships of Ixodetis clade RIPs reveals them to be polyphyletic, grouping separately with distinct clades of Citri RIPs. This is consistent with ancient ancestry or horizontal transfer within the genus (Ballinger et al., 2019). *sAtri* encodes two ETX/MTX2 toxins, both of which possess signal peptides for secretion. ETX/MTX2 toxins are β pore-forming toxins that inflict cytotoxic activity against their target cells and are noted for their insecticidal activity (Palma et al., 2014; Moar et al., 2017). Like *sAtri* RIPs, these toxins are diverse, grouping separately with other Ixodetis clade-encoded ETX/MTX2 toxins (Supplementary Figure S3). ETX/MTX2 toxins are numerous and widespread in *Spiroplasma* but their potential role in host-symbiont relationships is unknown. *sAtri* encodes 11 total genes containing ankyrin domains (Supplementary Figure S4) and a separate gene with an OTU-like cysteine protease. While not inherently toxic, ankyrin and OTU domains are associated with genes involved in host manipulation, and work in conjunction to cause male-killing by MSRO and possibly cytoplasmic incompatibility by *Wolbachia* (Pan et al., 2008; Nguyen et al., 2014; Harumoto and Lemaitre, 2018; Martinez et al., 2021). The OTU of *sAtri* is highly conserved among other heritable Ixodetis members (Supplementary Figure S5) and does not share an open reading frame with ankyrin domains. Despite the presence of these domains, no sex-ratio distortion has been observed in previous studies (Haselkorn et al., 2009) or in our lab stock (T. Chris Amuwa, unpublished results).

sAtri also encodes virulence genes that are scarcely found or not yet described in the *Spiroplasma* genus. These include an adenylate cyclase domain annotated by HMMER as anthrax toxin LF subunit. This domain is also found on the edema factor of *Bacillus anthracis* anthrax toxin. The edema factor is itself an adenylate cyclase toxin that causes large increases in the ubiquitous signaling molecule cAMP, resulting in severe disruptions to cellular processes (Tang and Guo, 2009).

sAtri encodes two metalloproteases of the M60 family which are implicated in virulence and infection outcomes in some bacteria and viruses (Nakjang et al., 2012; Belousov et al., 2018). M60-like

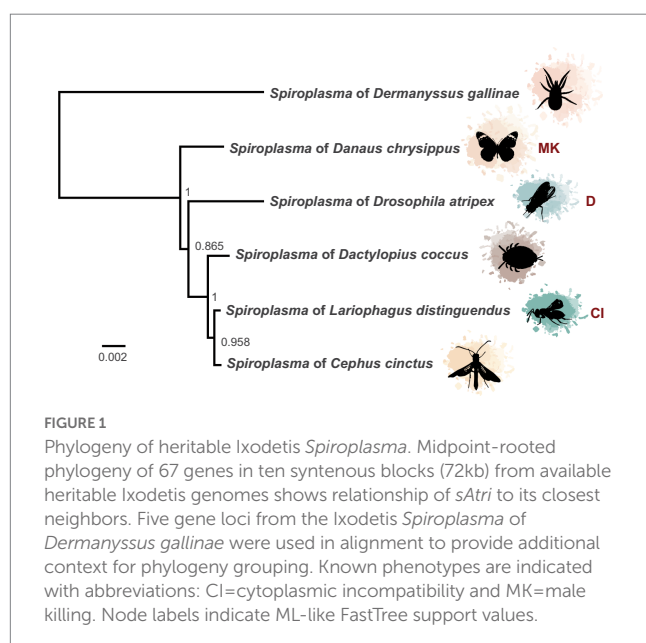
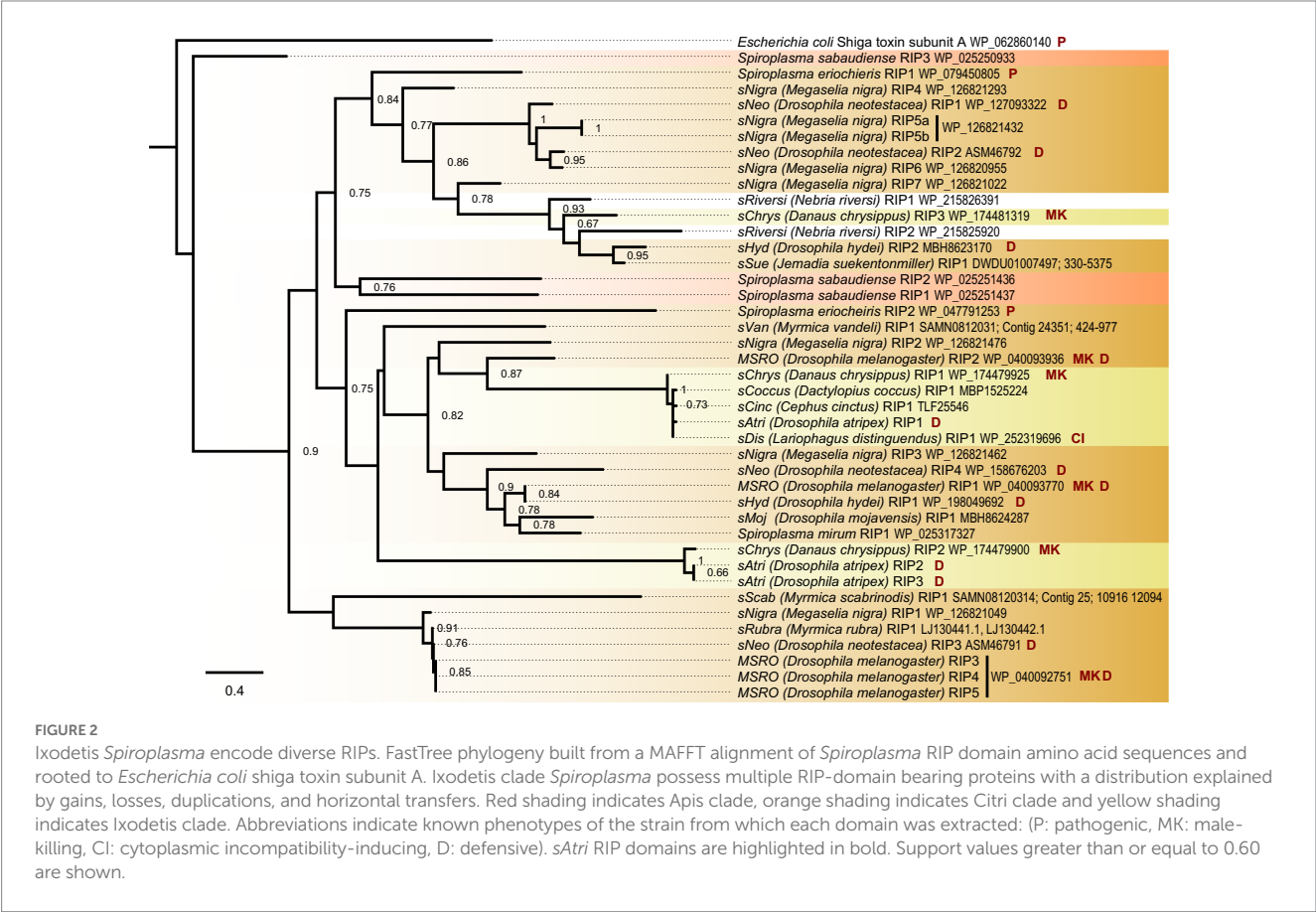


TABLE 1 Genome assembly statistics of *Spiroplasma*.

Genome	Size (bp)	Contigs	%GC	N50	CDS	tRNAs
<i>Spiroplasma</i> of <i>Drosophila atripex</i>	1,271,056	1	23.7	Closed	1,506	27
<i>Spiroplasma</i> of <i>Lariophagus distinguendus</i> ^{††}	1,163,832	198	24.3	14,219	1,175	27
<i>Spiroplasma</i> of <i>Cephus cinctus</i>	713,566	145	24.9	5,160	754	23
<i>Spiroplasma</i> of <i>Danaus chrysippus</i> [†]	1,745,430	12	23.7	215,399	1,782	27
<i>Spiroplasma</i> of <i>Dactylopius coccus</i> (DCF)	1,195,508	286	23.7	6,014	1,253	27
<i>Spiroplasma poulsonii</i> strain MSRO ^{†,‡}	1,883,005	1	26.4	Closed	2,217	31

[†]Male-killing strain.
^{††}CI-inducing strain.
[‡]Citri clade.



peptidases, also commonly referred to as enhancins, are well studied in, but not restricted to, insect pathogenic dsDNA viruses in the family Baculoviridae and function by binding to and degrading mucin layers that protect epithelial tissues (Ishimwe et al., 2015).

A conserved core of toxin domains in heritable *Spiroplasma*

The discovery that the *Drosophila* symbiont *sAtri* harbors multiple functional domains previously reported in distantly related heritable Citri clade symbionts motivated a more detailed analysis of domain distributions across the genus *Spiroplasma*. We focused the

scope of this analysis on RIP, OTU and ankyrin domains due to their role in defensive and reproductive manipulation phenotypes. We also analyzed ETX/MTX2 domains across *Spiroplasma* which are found widespread across this genus with little understanding of their role. VT-capable strains were determined based on a number of qualifying characteristics including evidence of transovarial transmission, PCR detection in hemolymph, ovaries and/or eggs, and presence of reproductive manipulation phenotypes (Table 2). VT *Spiroplasma* strains investigated in this study and their associated hosts, taxonomy and phenotypes are listed in Table 3. OTU and ankyrin domains can be found only in VT *Spiroplasma* across the Citri and Ixodetis clade (Figure 3). RIP domains are found in all VT *Spiroplasma* and in three non-VT *Spiroplasma*. ETX/MTX2 domains are more widely

TABLE 2 Evidence of vertical transmission.

Strain (host)	Evidence of vertical transmission	References
<i>sAtri</i> (<i>Drosophila atripex</i>)	Transgenerational transmission PCR detection in hemolymph Genome sequenced from host ovaries	T. Chris Amuwa (unpublished results) and this study
<i>MSRO</i> (<i>Drosophila melanogaster</i>)	Transgenerational transmission Detection in hemolymph via transinfection experiments FISH visualization in eggs FISH visualization of ovaries Reproductive manipulation	Pool et al. (2006) and Harumoto and Lemaitre (2018) Herren et al. (2013)
<i>sNeo</i> (<i>Drosophila neotestacea</i>)	Transgenerational transmission Detection in hemolymph via transinfection experiments	Jaenike et al. (2010a) and Haselkorn and Jaenike (2015)
<i>sHyd</i> (<i>Drosophila hydei</i>)	Transgenerational transmission Detection in hemolymph via transinfection PCR detection in eggs (unpublished)	Mateos et al. (2006) and Xie et al. (2010)
<i>sMoj</i> (<i>Drosophila mojavensis</i>)	Transgenerational transmission	Haselkorn et al. (2013)
<i>sVan</i> (<i>Myrmica vandeli</i>)	PCR detection in hemolymph High infection frequency (92.3%) Host specificity in sympatry	Ballinger et al. (2018)
<i>sScab</i> (<i>Myrmica scabrinodis</i>)	PCR detection in hemolymph High infection frequency (91.9–97.6%) Host specificity in sympatry	Ballinger et al. (2018)
<i>sNigra</i> (<i>Megaselia nigra</i>)	Transgenerational transmission (unpublished) PCR detection in hemolymph (unpublished) High infection frequency (73.3–80.2%)	Ballinger et al. (2019)
<i>sRiversi</i> (<i>Nebria riversi</i>)	PCR detection in egg, larval and adult stages	Weng et al. (2021)
<i>sChrys</i> (<i>Danaus chrysippus</i>)	Transgenerational transmission Reproductive manipulation	Martin et al. (2020)
<i>sCoccus</i> (<i>Dactylopius coccus</i>)	PCR detection in ovaries	Vera-Ponce León et al. (2021)
<i>sDis</i> (<i>Lariophagus distinguendus</i>)	Transgenerational transmission FISH visualization in ovaries Reproductive manipulation	Pollmann et al. (2022)

distributed among members of the Citri and Ixodetis clades, including most VT *Spiroplasma* and a few non-VT *Spiroplasma*. All four domains are almost entirely absent from the Apis clade except for three RIPs in *Spiroplasma sabaudiense* and a single ETX/MTX2 copy in *Spiroplasma culicicola*. Interestingly, no VT strains have been described in the Apis clade (Bolaños et al., 2015). We used BayesTraits software to identify correlations between distributions of toxin domains and heritability under two different models—one that assumes independent evolution and one that assumes dependent evolution. We find that all four domains are significantly better represented under a dependent model of evolution compared to an independent model (Table 4). For comparison, we used the same approach to determine if a correlation existed between heritability and other putative virulent *Spiroplasma* genes including *glpO* (glycerol 3 phosphate oxidase), *chiA* (chitinase), and *spi* (spiralin) (Alexeev et al., 2012; Chang et al., 2014; Duret et al., 2014). We find little to no support that the distribution of these domains across *Spiroplasma* are associated with heritability (Table 4; Supplementary Figure S6).

VT-specific domains vary in copy number and origin

While RIP, OTU, ankyrin, and ETX/MTX2 domains are distributed across diverse heritable *Spiroplasma*, the number of genes possessing these domains can vary drastically, even between closely related strains (Figure 4). For example, the number of genes possessing RIP domains varies both within and between clades, likely due to gene duplications and horizontal gene transfers. The number of genes possessing ankyrin domains are consistent within clade but show extreme copy number variation between clades characterized by ankyrin enrichment in Ixodetis *Spiroplasma*. Conversely, the number of OTU and ETX/MTX2 copies is relatively consistent within and between clades. All four of these domains can be found on plasmids (Harumoto and Lemaitre, 2018; Ballinger et al., 2019). In some cases, a given strain's entire domain repertoire can only be found on plasmids such as ankyrins and OTUs in MSRO, and ETX/MTX2 in *Spiroplasma citri*. In contrast, only one domain could be found within an endogenous phage region (ETX/MTX2 of MSRO), suggesting

TABLE 3 Vertically transmitted strains of *Spiroplasma* in study.

Host	<i>Spiroplasma</i> strain	Strain taxonomy	Phenotype
<i>Drosophila melanogaster</i>	MSRO	Citri-Poulsonii	Male-killing; defensive
<i>Drosophila neotestacea</i>	sNeo	Citri-Poulsonii	Defensive
<i>Drosophila hydei</i>	sHyd	Citri-Poulsonii	Defensive
<i>Myrmica vandeli</i>	sVan	Citri	–
<i>Myrmica scabrinodis</i>	sScab	Citri	–
<i>Drosophila mojavensis</i>	sMoj	Citri	–
<i>Megaselia nigra</i>	sNigra	Citri	–
<i>Nebria riversi</i>	sRiversi	Unplaced	–
<i>Dactylopius coccus</i>	sCoccus	Ixodetis	–
<i>Danaus chrysippus</i>	sChrys	Ixodetis	Male-killing
<i>Lariophagus distinguendus</i>	sDis	Ixodetis	Cytoplasmic incompatibility
<i>Drosophila atripex</i>	sAtri	Ixodetis	Defensive

Dash (–) indicates a phenotype has not yet been published.

Spiroplasma phages do not explain the distribution of RIP, OTU, ankryrin, or ETX/MTX2 domains across the genus.

VT-specific domains source diverse recombinant genes

Across Citri and Ixodetis-clade *Spiroplasma*, we observe multiple structural variants of RIPs, OTUs, ankryrins, and ETX/MTXs within the same open reading frame (Figure 5). In a particularly interesting case, we uncovered two Spaid-like toxins from the genome of the *Spiroplasma* endosymbiont of *Drosophila hydei* (strain sHyd1, NCBI BioProject PRJNA274591) and a *Spiroplasma* genome extracted from whole-genome sequence data of the South American butterfly *Jemadia suekentonmilleri* (sSue). These open reading frames include ankryrin domains, an OTU domain, and the Spaid C-terminal transmembrane domain. Toward the N-terminus, these proteins are also equipped with two RIP domains which together possess all the conserved active site residues of a single RIP domain. These Spaid-like toxins are present across several *Spiroplasma* genomes extracted from hesperid butterfly assemblies (Supplementary Figure S7).

Spiroplasma with VT-associated domains are widespread across Hesperidae butterflies and other insects

We further investigated the presence of *Spiroplasma* in insect genome assemblies using MSRO Spaid, and FtsZ and rpoB proteins from Apis, Citri, and Ixodetis clade taxa as queries against whole genome shotgun (WGS) sequence databases on NCBI and identified Citri clade *Spiroplasma* genomic sequences in several South American members of the Hesperidae butterfly family. A few of these *Spiroplasma* infections are also reported in an unpublished preprint (Medina et al., 2020; Manzano-Marin et al., 2020). These Hesperidae-infecting *Spiroplasma* strains group closely to the heritable *Spiroplasma* strains infecting *Drosophila mojavensis* and *Megaselia nigra* (sMoj and sNigra, respectively) (Figure 6).

Metagenomic binning of *Spiroplasma* contigs reveals this novel group of *Spiroplasma* is also equipped with RIPs, OTUs, ankryrins, and ETX/MTX2 domains. Phylogenetic analysis of Ixodetis clade *Spiroplasma* extracted from WGS sequences reveals an association with diverse insects including two ants, a butterfly, a springtail and a drosophilid (Figure 6). The presence of RIPs, OTUs, ankryrins, and ETX/MTX2 domains is variable among these novel Ixodetis strains (Figure 6). For example, the *Spiroplasma* genome extracted from a *Monomorium pharaonic* ant assembly (sPharaoh) is estimated to be 98% complete (Supplementary Table S4) and only encodes a single ETX/MTX2 domain. Alternatively, Ixodetis clade *Spiroplasma* genomes extracted from the assemblies of the butterfly *Colias croceus* (sCroceus) and the drosophilid *Zaprionus kolodkinae* (sKolod) both encode RIPs, OTUs, ankryrins, and ETX/MTX2 domains.

Discussion

In this study, we report a conserved core of toxin and virulence domains associated with heritable *Spiroplasma*. As one of the most successful and diverse arthropod-associated heritable symbionts on Earth (Duron et al., 2008), the mechanisms that support *Spiroplasma* persistence across ecological timescales are of great interest and yet largely unknown. Previous work has demonstrated that *Spiroplasma* exploit conserved yolk protein import machinery to invade maturing oocytes, providing a possible explanation for their effective vertical transmission in diverse arthropods (Herren et al. 2013). But how *Spiroplasma* navigates numerous other obstacles required for growth and persistence in host hemolymph throughout host development is not well understood. While toxins are often viewed through the lens of pathogenesis, they can also mediate and manipulate host processes to maintain symbiont persistence. For instance, RIPs and an OTU-ankryrin-bearing Spaid toxin promote *Spiroplasma* persistence by supporting defensive and reproductive phenotypes, respectively (Hamilton et al., 2016; Ballinger and Perlman, 2017; Harumoto and Lemaitre, 2018). Our assembly of the Ixodetis *Spiroplasma* genome from *Drosophila atripex* and accompanying comparative analyses reveal that the presence of these domains is highly conserved across

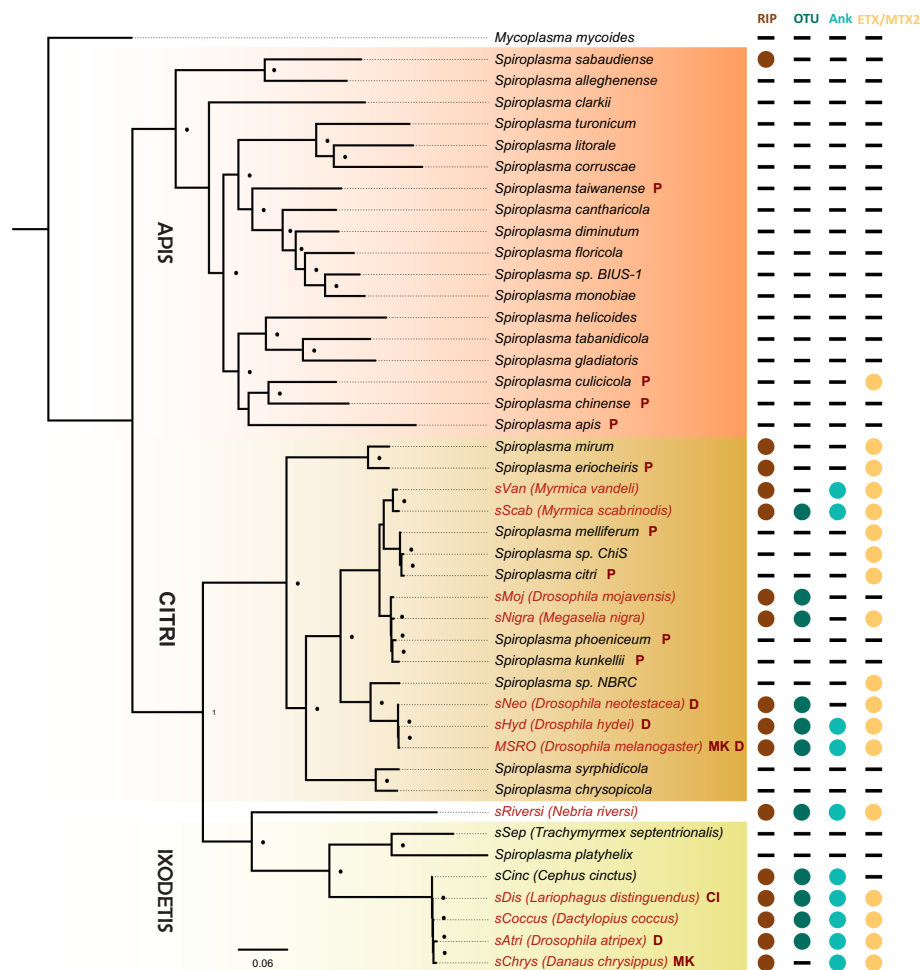


FIGURE 3

Phylogenetic distribution of select toxins and virulence domains in the genus *Spiroplasma*. FastTree phylogeny built from MAFFT nucleotide alignments of concatenated *Spiroplasma* *ftsZ*, *rpoB* and *gyrB* and rooted to *M. mycoides*. RIP, OTU, ankyrin, and ETX/MTX2 domains are preferentially distributed among VT *Spiroplasma*. Small black dots next to nodes indicate FastTree support values greater than 75%. Large colored circles to the right of branch labels indicate at least one domain copy is present in the genome. Black dash indicates no domain copies present in the genome. Red text is used for VT *Spiroplasma* and black text is used for non-VT *Spiroplasma*. Abbreviations indicate known phenotypes of the strain from which each domain was extracted: (P: pathogenic, MK: male-killing, CI: cytoplasmic incompatibility-inducing, D: defensive).

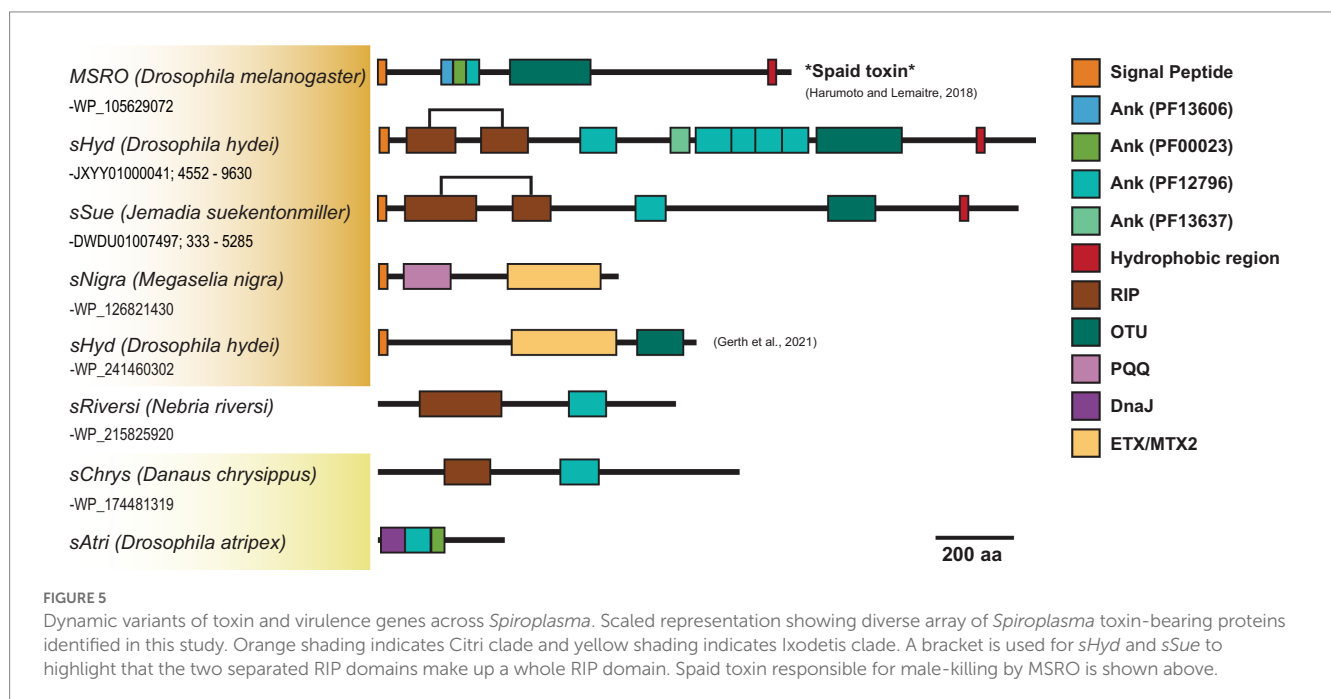
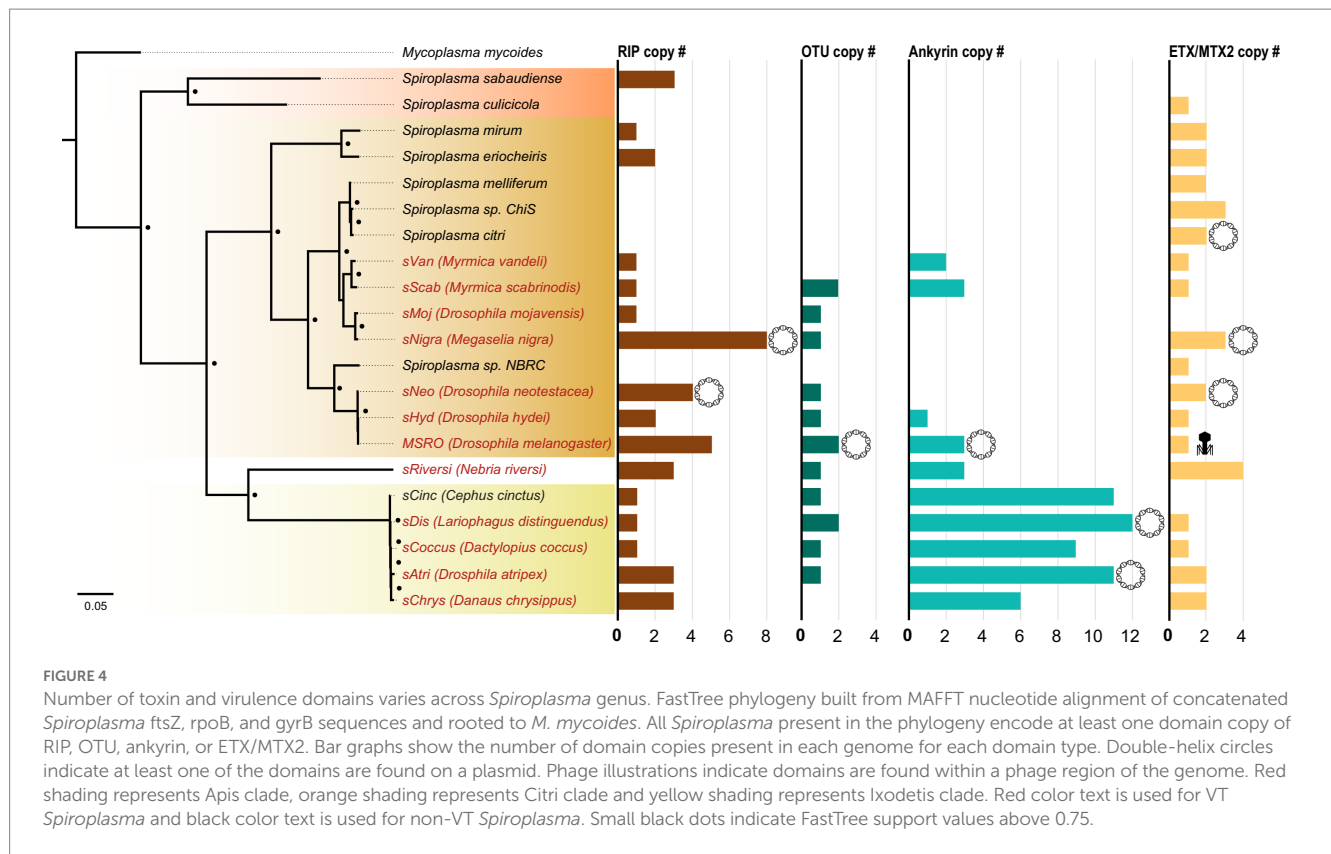
TABLE 4 Model comparisons for describing toxin and virulence domains across heritable *Spiroplasma*.

Domain	Dependent model likelihood score	Independent model likelihood score	Chi score	df	value of p^*
RIP	-29.855283	-40.389039	21.067512	4	0.000307
OTU	-26.546300	-37.764235	22.43587	4	0.000164
ankyrin	-26.036137	-35.577156	19.082038	4	0.000757
ETX/MTX2	-33.930099	-44.915624	21.97105	4	0.000203
glpO	-44.902777	-46.963108	4.056606	4	0.398347
chiA	-37.674589	-40.678499	6.00782	4	0.198552
spiralin	-32.183710	-37.138576	9.91	4	0.041971

*Value of p threshold has been set at 0.001.

diverse heritable *Spiroplasma*. RIPs, OTUs and ankyrin domains are specific to VT *Spiroplasma*, and are rare (RIPs) or entirely absent (OTUs and ankyrins) in non-VT *Spiroplasma* genomes. These toxin and virulence domains may play a role in facilitating and maintaining

a vertically transmissible lifestyle across heritable *Spiroplasma* as has been demonstrated for a small number of Citri clade members. Finally, ETX/MTX2 toxins are more widespread in non-VT Citri and Ixodetis *Spiroplasma* compared to RIPs, OTUs, and ankyrins but are



still highly correlated with a VT lifestyle based on BayesTraits analysis. As previously stated, the role of ETX/MTX2 toxins in host-*Spiroplasma* interactions is unknown but their propensity for VT-*Spiroplasma* may suggest a role in maintaining a VT lifestyle similar to RIPs, OTUs and ankyrin domains (i.e., defense or host manipulation).

We find that similar toxin evolutionary processes previously reported in Poulsonii clade strains (Ballinger et al., 2019; Gerth et al., 2021) are mirrored in toxin domain-containing proteins of the Ixodetis clade, suggesting conservation of dominant toxin evolution mechanisms across heritable strains throughout the genus. Citri clade RIPs experience duplications, losses, and domain

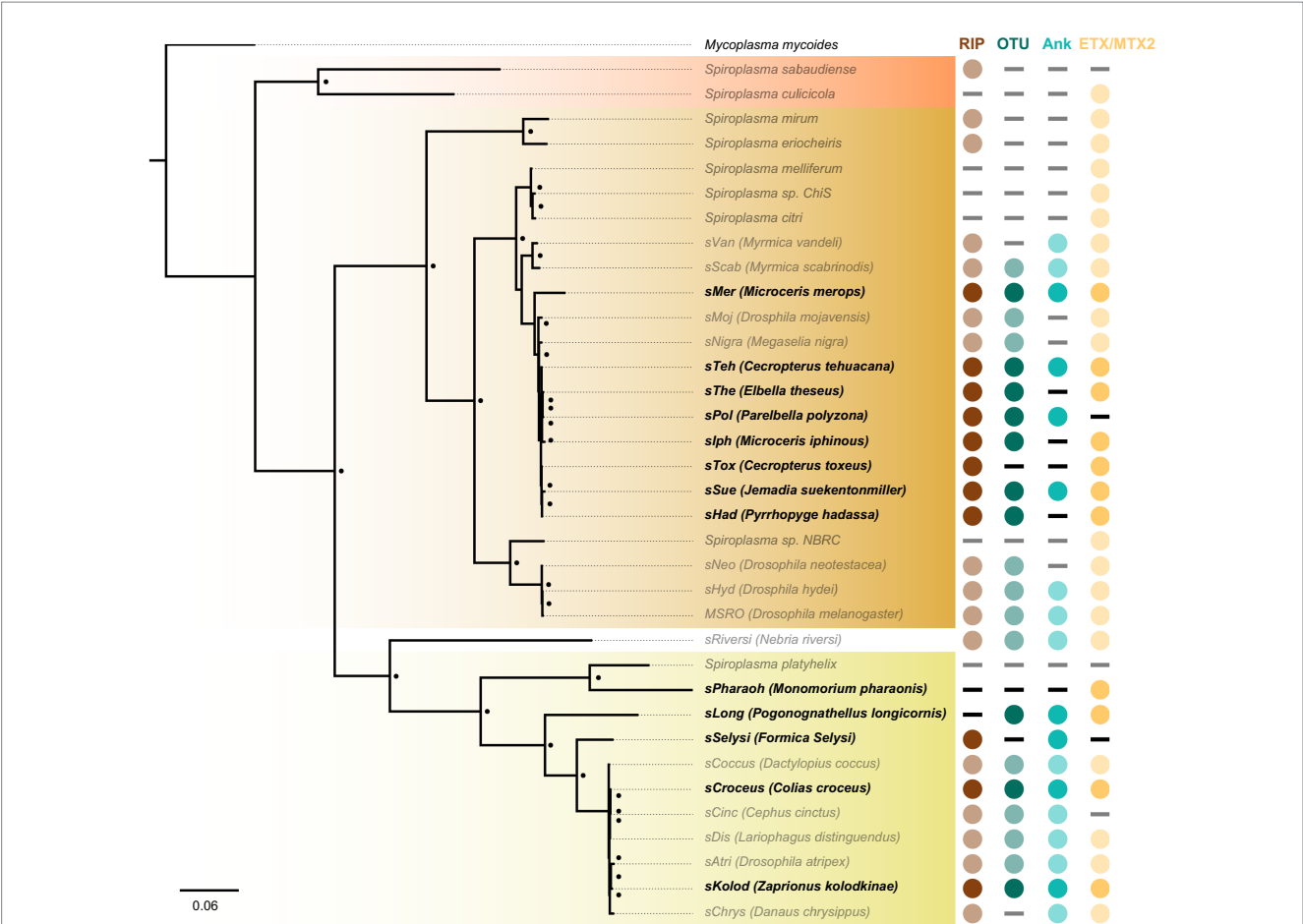


FIGURE 6 Cryptic *Spiroplasma* carry domains associated with a heritable lifestyle. FastTree phylogeny built from MAFFT nucleotide alignment of concatenated *Spiroplasma* *ftsZ*, *rpoB*, and *gyrB* and rooted to *M. mycoides*. This phylogram shows the placement of novel *Spiroplasma* strains (black bold) extracted from insect WGS sequences. Many of these *Spiroplasma* strains carry RIPs, OTUs, and/or ankyrin domains that are strongly associated with a heritable lifestyle. Circles indicate the presence of a given domain and dashes indicate their absence. Red shading indicates Apis clade, orange shading indicates Citri clade and yellow shading indicates Ixodetis clade. WGS-extracted taxa are shown in bold typeface. Black dots adjacent to nodes indicate a FastTree support value greater than 0.75.

swapping events that have accompanied an explosion in RIP diversity. Similarly, Ixodetis RIPs are polyphyletic and RIP representation differs between strains. For example, a RIP from the Ixodetis endosymbiont of the butterfly *Danaus chrysippus* groups closely with the enigmatic double RIP domain-possessing proteins recently uncovered from *sHyd1* and *sSue* (Figure 5; Gerth et al., 2021). Ankyrin domain-containing proteins are especially numerous in members of the Ixodetis clade, ranging from six to fourteen copies across members of this clade. Believed to originate in eukaryotes, ankyrin domains are more commonly found in heritable microbes than other bacterial taxa (Jernigan and Bordenstein, 2014), and may facilitate a variety of microbe-host interactions. This makes the expansion of ankyrin domains in Ixodetis particularly interesting, especially given the broad range of insect hosts they infect (Figure 1).

Dynamic evolution is also demonstrated in the varied number of genes that possess these domains (i.e., RIP, OTU, ankyrin, and ETX/MTX2)—even among closely related strains—indicating that duplications, gains, and losses are common generators of their diversity in *Spiroplasma*. RIP, OTU, ankyrin, and ETX/MTX2

domains are also present on plasmids, possibly contributing functional roles that favor their dispersal and persistence through the genus. Plasmids often confer adaptive phenotypes to their bacterial hosts to help drive their own spread among bacterial populations (Dimitriu et al., 2021). Interestingly, no RIP, OTU, or ankyrin domains are present within the diverse plasmids of non-VT Citri clade *Spiroplasma*. For example, some *Spiroplasma citri* strains can have upwards of nine unique plasmids (Rattner et al., 2021). Given the promiscuity of plasmids across *Spiroplasma*, the lack of RIP, OTU, and ankyrin domains on plasmids in non-heritable strains suggests that they may have little adaptive function in that lifestyle.

Not only do RIPs, OTUs, and ankyrin domains exist within diverse proteins across *Spiroplasma*, but many VT *Spiroplasma* species are equipped with several structural variants of these domains on the same protein. In one notable example, our analysis revealed a *Spiroplasma* genome within the WGS assembly of the South American butterfly *J. suekentonmilleri* that encodes a protein with RIP domains, an OTU domain, an ankyrin domain and a C-terminal transmembrane domain, i.e., a RIP-Spaid fusion protein.

These VT-associated domains appear to frequently recombine to create novel protein configurations. For instance, Spaid and Spaid-like proteins encode multiple ankyrin domains each and these ankyrin domains vary greatly in copy number and protein family (Figure 5; Supplementary Figure S7B), suggesting that ankyrin-spanning region is especially prone to domain losses and gains. The close association of RIPs, OTUs, and ankyrins on the same proteins across Citri and Ixodetis *Spiroplasma* suggests functional evolutionary ties between these domains.

Extracting symbiont genomes from publicly available nucleotide databases is a powerful and convenient approach to studying symbiont evolution (Salzberg et al., 2005; Gerth and Hurst, 2017; Pascari and Chandler, 2018; Scholz et al., 2020; Pilgrim et al., 2021). We explored the presence of RIP, OTU, ankyrin, and ETX/MTX2 domains in other insect assemblies by mining partial to near complete genomes of both Ixodetis and Citri clade *Spiroplasma* from a diverse array of insect WGS assemblies. Within Citri, there is a large group of *Spiroplasma* strains infecting various members of the South American butterfly family Hesperidae, and are equipped with RIP, OTU, and ankyrin domain-possessing proteins. Within Ixodetis, we identified five novel *Spiroplasma* genomes extracted from a diverse variety of insect hosts with a varied presence of RIP, OTU, ankyrin, and ETX/MTX2 domains. This analysis provides promising candidates for future studies on heritable *Spiroplasma* and the identification of RIP, OTU, ankyrin, and ETX/MTX2 domains within insect WGS assemblies may be useful for developing early hypotheses on *Spiroplasma*-host interactions.

Methods

Genome sequencing, assembly, and annotation

Nucleic acids were extracted from ovaries of eight adult female *Drosophila atriplex* using the phenol-chloroform method. Paired end 150 bp Illumina genomic DNA reads were sequenced by Novogene (CA, United States). Long reads were sequenced on an Oxford Nanopore MinION device and R9.4.1 flow cell following library preparation using the Ligation Sequencing Kit (SQK-LSK109). Guppy 6.1.5 implemented in MinKNOW 22.05.5 was used to generate high accuracy basecalls. Only long reads above 10 kb were used in hybrid assembly. Illumina reads were trimmed for adapter sequences and quality (qtrim = r trimq = 10) with BBMap 38.35 (Bushnell, 2014), and a metagenome was hybrid assembled using Unicycler 0.5.0 (Wick et al., 2017). *Spiroplasma*-derived contigs were retained through a combination of metagenomic binning (Laczny et al., 2017) and manual filtering. Manual filtering was performed through targeted searches of sequences likely to contaminate the bin based on similar nucleotide content and was guided by kmer abundance analyses performed with KAT (Mapleson et al., 2016). The draft genome was annotated through Prokka 1.14.5 (Seemann, 2014) and KEGG numbers were assigned with BLASTKoala (Kanehisa et al., 2016). Genome coverage and plasmid coverage were determined using BBMap (Bushnell, 2014). Genome completeness was estimated with BUSCO v5 (Simão et al., 2015) on the gVolante webserver (Nishimura et al., 2017). *Spiroplasma* homologs of interest, including those used for phylogenetics, were identified using HMMER (Finn et al., 2011)

and PfamScan (Mistry et al., 2021) with a threshold of E-value = 0.01. Proteins of interest that could not be annotated by the previously mentioned approaches were investigated further with HHpred (Zimmermann et al., 2018), BLASTp and alignments. Protein and nucleotide sequences were aligned with MAFFT 7.388 (Katoh and Standley, 2013) and phylogenies were built using FastTree 2.1.11 (Price et al., 2010) unless stated otherwise. BLASTp searches with relaxed significance thresholds (expect values ≤ 0.01) were also performed to ensure detection of more distantly related protein domains if present.

Characterizing distribution of heritability, toxin, and virulence domains across *Spiroplasma* genus

All *Spiroplasma* sp. genomes, plasmids, and reads were downloaded from NCBI, and their accession numbers are listed in Supplementary Table S5. Heritable *Spiroplasma* were identified based on evidence provided from transgenerational screenings, ovarian tissue and egg screens, high infection frequencies among populations, multi-year screenings and systemic infections (Table 2). We searched for RIP, OTU, ankyrin, and ETX/MTX2 domains in all *Spiroplasma* genomes using tBLASTn (threshold E-value = 0.01) with a curated list of *Spiroplasma* domains that have been identified in this study and extracted using HMMER domain annotations as a guide (Supplementary Table S6). Protein domains were confirmed with HMMER 3.3 and pfamscan with a 0.01 e-value threshold, and through alignments to confirm presence of conserved residues. Domains present on open reading frames and on pseudogenes are both included in the total count for domains of interest. Phage regions were determined using Phaster webserver (Zhou et al., 2011; Arndt et al., 2016). Annotated phage regions were extracted from the genomes of *Spiroplasma* strains known to contain RIP, ankyrin, OTU, or ETX/MTX2 domains. These phage regions were investigated to determine if they included any of these domains.

We determined whether a correlation existed between heritability and specific encoded domains using the software BayesTraits V4.0.0 (Pagel and Meade, 2006). Heritability was treated as a discrete binary trait (heritable or non-heritable) and domains were also treated as a binary trait (domain present domain absent in genome). Likelihood scores were calculated for both a dependent model (i.e., assumes heritability and domain acquisition/retention evolved dependently) and an independent model (i.e., assumes heritability and domain acquisition/retention evolved independently). Likelihood scores were compared with a chi-square analysis as recommended by BayesTraits documentation. We rejected the null hypothesis of independent evolution at $p < 0.001$.

Identification and extraction of *Spiroplasma* genomes from insect whole genome shotgun assemblies

Using Spaid toxin from MSRO, and ftsZ and rpoB from Ixodetis, Citri, and Apis clade taxa as queries, we performed tBLASTn searches of whole genome shotgun contigs (WGS) located on the NCBI database. Default search parameters were used, and we limited

our organism search to Insecta (taxid:50557). Assemblies with matches to these genes were submitted to BusyBee web server for genomic binning (Laczny et al., 2017). Bins outputted by BusyBee were manually inspected for the presence of core *Spiroplasma* genes including *ftsZ*, *rpoB*, and *gyrB*. These core genes were also used to construct the phylogeny in Figure 6. We performed tBLASTn against the extracted *Spiroplasma* genomes using a curated list of *Spiroplasma* RIP, OTU and ankyrin peptide sequences. If a domain was determined to be missing from an extracted *Spiroplasma* genome, we searched for it in the original insect assembly to ensure it wasn't present on a *Spiroplasma* contig missed in the binning process. Genome completeness was determined using BUSCO through gVolante webserver, and coding DNA sequences and tRNA content was determined with Prokka 1.14.5. Genome statistics and insect assembly accession numbers are available in Supplementary Table S4.

Constructing phylogenies

MAFFT 7.388 was used to create all alignments and FastTree 2.1.11 was used to construct *Spiroplasma* phylogenies. Jukes-Cantor model was used to build phylogenies from nucleotide sequence alignments. RIP and OTU proteins are often flanked with diverse, nonhomologous accessory domains. For this reason, RIP and OTU regions immediately outside of the conserved active site residues were trimmed out of alignments manually; conserved active site residues span the majority of both domains. *sSue* RIP, *sHyd* RIP2, and *sChrys* RIP3 all possess peptide insertion sequences ranging from 54 to 157 aa that split their RIP domain. This appears to be a conserved feature among this clade of RIPs. Due to their large size and highly divergent sequence variation, we identified these insertion sequence regions using alignments to other RIPs and removed them manually. *sHyd* RIP2 and *sRiversi* RIP1 also group among these RIPs, however they are small proteins encoding only a partial RIP domain and therefore did not require additional trimming. The trimmed RIP and OTU sequences were then aligned with MAFFT 7.388 and submitted to MEGA software (Kumar et al., 2018) to determine the best substitution model for PhyML tree building. WAG+G+I substitution model was used to construct a RIP phylogeny and cpREV+G substitution model was used to construct an OTU phylogeny. ETX/MTX2 are divergent toxins that lack conserved active site residues to help guide trimming. We created an alignment of ETX/MTX2-possessing proteins with MAFFT 7.388 and confirmed that the annotated ETX/MTX2 domains were aligned. The alignment was then trimmed using ClipKIT software (Steenwyk et al., 2020) in kpi-gappy mode. The trimmed alignment output was submitted to MEGA software to determine the best substitution model for PhyML tree building. WAG+G+F substitution model was used to construct an ETX/MTX2 phylogeny.

References

Alexeev, D., Kostrjukova, E., Aliper, A., Popenko, A., Bazaleev, N., Tyakht, A., et al. (2012). Application of *Spiroplasma melliferum* proteogenomic profiling for the discovery of virulence factors and pathogenicity mechanisms in host-associated *Spiroplasmas*. *J. Proteome Res.* 11, 224–236. doi: 10.1021/pr2008626

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: The Bioproject accession is “PRJNA928682” and the Biosample accession is “SAMN32939082” and the SRA accessions are “SRX19199603” and “SRX19199602.”

Author contributions

LM and MB conceived and designed the study, collected data and performed analyses, and revised the manuscript and approved the final submission. LM prepared figures and wrote the first draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by National Science Foundation award 2144270 to MB.

Acknowledgments

We are grateful to Luciano Matzkin for sharing *Drosophila atripex* lines.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1148263/full#supplementary-material>

Andrews, M., De Meyer, S., James, E., Stępkowski, T., Hodge, S., Simon, M., et al. (2018). Horizontal transfer of symbiosis genes within and between Rhizobial genera: occurrence and importance. *Genes* 9:321. doi: 10.3390/genes9070321

- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Arnold, B. J., Huang, L.-T., and Hanage, W. P. (2022). Horizontal gene transfer and adaptive evolution in bacteria. *Nat. Rev. Microbiol.* 20, 206–218. doi: 10.1038/s41579-021-00650-4
- Ballinger, M. J., Gawryluk, R. M. R., and Perlman, S. J. (2019). Toxin and genome evolution in a *Drosophila* defensive symbiosis. *Genome Biol. Evol.* 11, 253–262. doi: 10.1093/gbe/evy272
- Ballinger, M. J., Moore, L. D., and Perlman, S. J. (2018). Evolution and diversity of inherited *Spiroplasma* symbionts in myrmica ants. *Appl. Environ. Microbiol.* 84:e02299–17. doi: 10.1128/AEM.02299-17
- Ballinger, M. J., and Perlman, S. J. (2017). Generality of toxins in defensive symbiosis: ribosome-inactivating proteins and defense against parasitic wasps in *Drosophila*. *PLoS Pathog.* 13:e1006431. doi: 10.1371/journal.ppat.1006431
- Beckmann, J. E., Ronau, J. A., and Hochstrasser, M. (2017). A Wolbachia deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* 2:17007. doi: 10.1038/nmicrobiol.2017.7
- Belousov, M. V., Bondarev, S. A., Kosolapova, A. O., Antonets, K. S., Sulatskaya, A. I., Sulatsky, M. I., et al. (2018). M60-like metalloprotease domain of the *Escherichia coli* YghJ protein forms amyloid fibrils. *PLoS ONE* 13, e0191317. doi: 10.1371/journal.pone.0191317
- Binetruy, F., Bailly, X., Chevillona, C., Martin, O. Y., Bernasconi, M. V., and Duron, O. (2019). Phylogenetics of the *Spiroplasma ixodetis* endosymbiont reveals past transfers between ticks and other arthropods. *Ticks Tick-Borne Dis.* 10, 575–584. doi: 10.1016/j.ttbdis.2019.02.001
- Bolaños, L. M., Servín-Garcidueñas, L. E., and Martínez-Romero, E. (2015). Arthropod–*Spiroplasma* relationship in the genomic era. *FEMS Microbiol. Ecol.* 91, 1–8. doi: 10.1093/femsec/fiu008
- Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. Lawrence Berkeley National Laboratory. LBNL Report #: LBNL–7065E.
- Chang, T.-H., Lo, W.-S., Ku, C., Chen, L.-L., and Kuo, C.-H. (2014). Molecular evolution of the substrate utilization strategies and putative virulence factors in mosquito-associated *Spiroplasma* species. *Genome Biol. Evol.* 6, 500–509. doi: 10.1093/gbe/evu033
- Degnan, P. H., and Moran, N. A. (2008). Diverse phage-encoded toxins in a protective insect Endosymbiont. *Appl. Environ. Microbiol.* 74, 6782–6791. doi: 10.1128/AEM.01285-08
- Dimitriu, T., Matthews, A. C., and Buckling, A. (2021). Increased copy number couples the evolution of plasmid horizontal transmission and plasmid-encoded antibiotic resistance. *Proc. Natl. Acad. Sci. U. S. A.* 118:e2107818118. doi: 10.1073/pnas.2107818118
- Duret, S., Batailler, B., Dubrana, M.-P., Saillard, C., Renaudin, J., Béven, L., et al. (2014). Invasion of insect cells by *Spiroplasma citri* involves spiralin relocalization and lectin/glycoconjugate-type interactions: Spiralin relocalization and insect cell adhesion. *Cell. Microbiol.* 16, 1119–1132. doi: 10.1111/cmi.12265
- Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L., Engelstädter, J., et al. (2008). The diversity of reproductive parasites among arthropods: Wolbachia do not walk alone. *BMC Biol.* 6:27. doi: 10.1186/1741-7007-6-27
- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367
- Gasparich, G. E., Whitcomb, R. F., Dodge, D., French, F. E., Glass, J., and Williamson, D. L. (2004). The genus *Spiroplasma* and its non-helical descendants: phylogenetic classification, correlation with phenotype and roots of the *Mycoplasma mycoides* clade. *Int. J. Syst. Evol. Microbiol.* 54, 893–918. doi: 10.1099/ijs.0.02688-0
- Gerth, M., and Hurst, G. D. D. (2017). Short reads from honey bee (*Apis* sp.) sequencing projects reflect microbial associate diversity. *Peer J* 5:e3529. doi: 10.7717/peerj.3529
- Gerth, M., Martínez-Montoya, H., Ramirez, P., Masson, F., Griffin, J. S., Aramayo, R., et al. (2021). Rapid molecular evolution of *Spiroplasma* symbionts of *Drosophila*. *Microbial Genomics* 7, 1–15. doi: 10.1099/mgen.0.000503
- Gillespie, J. J., Driscoll, T. P., Verhoeve, V. I., Rahman, M. S., Macaluso, K. R., and Azad, A. F. (2018). A tangled web: origins of reproductive parasitism. *Genome Biol. Evol.* 10, 2292–2309. doi: 10.1093/gbe/evy159
- Hamilton, P. T., Peng, F., Boulanger, M. J., and Perlman, S. J. (2016). A ribosome-inactivating protein in a *Drosophila* defensive symbiont. *Proc. Natl. Acad. Sci. U. S. A.* 113, 350–355. doi: 10.1073/pnas.1518648113
- Harumoto, T., and Lemaître, B. (2018). Male-killing toxin in a bacterial symbiont of *Drosophila*. *Nature* 557, 252–255. doi: 10.1038/s41586-018-0086-2
- Haselkorn, T. S. (2010). The *Spiroplasma* heritable bacterial endosymbiont of *Drosophila*. *Fly* 4, 80–87. doi: 10.4161/fly.4.1.10883
- Haselkorn, T. S., and Jaenike, J. (2015). Macroevolutionary persistence of heritable endosymbionts: acquisition, retention and expression of adaptive phenotypes in *Spiroplasma*. *Mol. Ecol.* 24, 3752–3765. doi: 10.1111/mec.13261
- Haselkorn, T. S., Markow, T. A., and Moran, N. A. (2009). Multiple introductions of the *Spiroplasma* bacterial endosymbiont into *Drosophila*. *Mol. Ecol.* 18, 1294–1305. doi: 10.1111/j.1365-294X.2009.04085.x
- Haselkorn, T. S., Watts, T. D., and Markow, T. A. (2013). Density dynamics of diverse *Spiroplasma* strains naturally infecting different species of *Drosophila*. *Fly* 7, 204–210. doi: 10.4161/fly.25469
- Herren, J. K., Paredes, J. C., Schüpfer, F., and Lemaître, B. (2013). Vertical transmission of a *Drosophila* Endosymbiont via cooption of the yolk transport and internalization machinery. *MBio* 4, e00532–e00512. doi: 10.1128/mBio.00532-12
- Horard, B., Terretaz, K., Gosselin-Grenet, A.-S., Sobry, H., Sicar, M., Landmann, F., et al. (2022). Paternal transmission of the Wolbachia CidB toxin underlies cytoplasmic incompatibility. *Curr. Biol.* 32, 1319–1331.e5. doi: 10.1016/j.cub.2022.01.052
- Humphery-Smith, I., Grulet, O., Le Goff, F., and Chastel, C. (1991). *Spiroplasma* (mollicutes: spiroplasmataceae) pathogenic for *Aedes aegypti* and *Anopheles stephensi* (diptera: culicidae). *J. Med. Entomol.* 28, 219–222. doi: 10.1093/jmedent/28.2.219
- Ishimwe, E., Hodgson, J. J., Clem, R. J., and Passarelli, A. L. (2015). Reaching the melting point: degradative enzymes and protease inhibitors involved in baculovirus infection and dissemination. *Virology* 479–480, 637–649. doi: 10.1016/j.virol.2015.01.027
- Jaenike, J., Stahlhut, J. K., Boelio, L. M., and Unckless, R. L. (2010a). Association between Wolbachia and *Spiroplasma* within *Drosophila* neotestacea: an emerging symbiotic mutualism? wolbachia-*Spiroplasma* mutualism. *Mol. Ecol.* 19, 414–425. doi: 10.1111/j.1365-294X.2009.04448.x
- Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M., and Perlman, S. J. (2010b). Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science* 329, 212–215. doi: 10.1126/science.1188235
- Jernigan, K. K., and Bordenstein, S. R. (2014). Ankyrin domains across the tree of life. *PeerJ* 2:e264. doi: 10.7717/peerj.264
- Jiggins, F. M., Hurst, G. D. D., Jiggins, C. D. V. D., Schulenburg, J. H. G., and Majerus, M. E. N. (2000). The butterfly *Danaus chrysippus* is infected by a male-killing *Spiroplasma* bacterium. *Parasitology* 120, 439–446. doi: 10.1017/S0031182099005867
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and Metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kaur, R., Leigh, B. A., Ritchie, I. T., and Bordenstein, S. R. (2022). The Cif proteins from Wolbachia prophage WO modify sperm genome integrity to establish cytoplasmic incompatibility. *PLoS Biol.* 20:e3001584. doi: 10.1371/journal.pbio.3001584
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Lacny, C. C., Kiefer, C., Galata, V., Fehlmann, T., Backes, C., and Keller, A. (2017). BusyBee web: metagenomic data analysis by bootstrapped supervised binning and annotation. *Nucleic Acids Res.* 45, W171–W179. doi: 10.1093/nar/gkx348
- LePage, D. P., Metcalf, J. A., Bordenstein, S. R., On, J., Perlmuter, J. I., Shropshire, J. D., et al. (2017). Prophage WO genes recapitulate and enhance Wolbachia-induced cytoplasmic incompatibility. *Nature* 543, 243–247. doi: 10.1038/nature21391
- Lukasik, P., van Asch, M., Guo, H., Ferrari, J., Charles, J., and Godfray, H. (2013). Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecol. Lett.* 16, 214–218. doi: 10.1111/ele.12031
- Manzano-Marin, A., Coeur Dacier, A., Clamens, A. L., Orvain, C., Cruaud, C., Barbe, V., et al. (2020). Serial horizontal transfer of vitamin-biosynthetic genes enables the establishment of new nutritional symbionts in aphids' di-symbiotic systems. *ISME J.* 14, 259–273. doi: 10.1038/s41396-019-0533-6
- Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., and Clavijo, B. J. (2016). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics* 33, 574–576. doi: 10.1093/bioinformatics/btw663
- Martin, S. H., Singh, K. S., Gordon, I. J., Omufwoko, K. S., Collins, S., Warren, I. A., et al. (2020). Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLoS Biol.* 18:e3000610. doi: 10.1371/journal.pbio.3000610
- Martinez, J., Klasson, L., Welch, J. J., and Jiggins, F. M. (2021). Life and death of selfish genes: comparative genomics reveals the dynamic evolution of cytoplasmic incompatibility. *Mol. Biol. Evol.* 38, 2–15. doi: 10.1093/molbev/msaa209
- Massey, J. H., and Newton, I. L. G. (2022). Diversity and function of arthropod endosymbiont toxins. *Trends Microbiol.* 30, 185–198. doi: 10.1016/j.tim.2021.06.008
- Mateos, M., Castrezana, S. J., Nankivell, B. J., Estes, A. M., Markow, T. A., and Moran, N. A. (2006). Heritable endosymbionts of *Drosophila*. *Genetics* 174, 363–376. doi: 10.1534/genetics.106.058818
- McCutcheon, J. P., Boyd, B. M., and Dale, C. (2019). The life of an insect endosymbiont from the cradle to the grave. *Curr. Biol.* 29, R485–R495. doi: 10.1016/j.cub.2019.03.032
- McCutcheon, J. P., and Moran, N. A. (2012). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26. doi: 10.1038/nrmicro2670

- Medina, P., Russell, S. L., Aswadhati, K., and Corbett-Detig, R. (2020). Deep data mining reveals variable abundance and distribution of microbial reproductive manipulators within and among diverse host species. *bioRxiv*. doi: 10.1101/679837 [Pre-print].
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: the protein families database in 2021. *Nucleic Acids Res.* 49, D412–D419. doi: 10.1093/nar/gkaa913
- Moar, W. J., Evans, A. J., Kessenich, C. R., Baum, J. A., Bowen, D. J., Edrington, T. C., et al. (2017). The sequence, structural, and functional diversity within a protein family and implications for specificity and safety: the case for ETX_MTX2 insecticidal proteins. *J. Invertebr. Pathol.* 142, 50–59. doi: 10.1016/j.jip.2016.05.007
- Montenegro, H., Solferini, V. N., Klaczko, L. B., and Hurst, G. D. D. (2005). Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*: Male-killing *Spiroplasma* in *D. melanogaster*. *Insect. Mol. Biol.* 14, 281–287. doi: 10.1111/j.1365-2583.2005.00558.x
- Moran, N. A. (2002). Microbial Minimalism. *Cells* 108, 583–586. doi: 10.1016/S0092-8674(02)00665-7
- Moran, N. A., McCutcheon, J. P., and Nakabachi, A. (2008). Genomics and evolution of heritable bacterial Symbionts. *Annu. Rev. Genet.* 42, 165–190. doi: 10.1146/annurev.genet.41.110306.130119
- Mouches, C., Bové, J. M., Albisetti, J., Clark, T. B., and Tully, J. G. (1982). A *Spiroplasma* of serogroup IV causes a may-disease-like disorder of honeybees in southwestern France. *Microb. Ecol.* 8, 387–399. doi: 10.1007/BF02010677
- Nakjang, S., Ndeh, D. A., Wipat, A., Bolam, D. N., and Hirt, R. P. (2012). A Novel Extracellular Metalloprotease Domain Shared by Animal Host-Associated Mutualistic and Pathogenic Microbes. *PLoS ONE* 7, e30287. doi: 10.1371/journal.pone.0030287
- Nguyen, M. T. H. D., Liu, M., and Thomas, T. (2014). Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol. Ecol.* 23, 1635–1645. doi: 10.1111/mec.12384
- Nishimura, O., Hara, Y., and Kuraku, S. (2017). gVolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics* 33, 3635–3637. doi: 10.1093/bioinformatics/btx445
- Oliver, K. M., and Perlman, S. J. (2020). Toxin-mediated protection against natural enemies by insect defensive symbionts. *Adv. Insect Physiol.* 58, 277–316. doi: 10.1016/b.s.aip.2020.03.005
- Pagel, M., and Meade, A. (2006). Bayesian Analysis of Correlated Evolution of Discrete Characters by Reversible-Jump Markov Chain Monte Carlo. *The American Naturalist* 167, 808–825. doi: 10.1086/503444
- Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). *Bacillus thuringiensis* toxins: an overview of their Biocidal activity. *Toxins* 6, 3296–3325. doi: 10.3390/toxins6123296
- Pan, X., Lührmann, A., Satoh, A., Laskowski-Arce, M. A., and Roy, C. R. (2008). Ankyrin repeat proteins comprise a diverse family of bacterial type IV effectors. *Science* 320, 1651–1654. doi: 10.1126/science.1158160
- Paredes, J. C., Herren, J. K., Schüpfer, F., Marin, R., Claverol, S., Kuo, C.-H., et al. (2015). Genome sequence of the *Drosophila melanogaster* male-killing *Spiroplasma* strain MSRO endosymbiont. *mBio* 6, e02437–e02414. doi: 10.1128/mBio.02437-14
- Pascari, J., and Chandler, C. H. (2018). A bioinformatics approach to identifying Wolbachia infections in arthropods. *PeerJ* 6:e5486. doi: 10.7717/peerj.5486
- Pilgrim, J., Thongpreem, P., Davison, H. R., Siozios, S., Baylis, M., Zakharov, E. V., et al. (2021). Torix Rickettsia are widespread in arthropods and reflect a neglected symbiosis. *GigaScience* 10:giab021. doi: 10.1093/gigascience/giab021
- Pollmann, M., Moore, L. D., Krimmer, E., D'Alvise, P., Hasselmann, M., Perlman, S. J., et al. (2022). Highly transmissible cytoplasmic incompatibility by the extracellular insect symbiont *Spiroplasma*. *iScience* 25:104335. doi: 10.1016/j.isci.2022.104335
- Pool, J. E., Wong, A., and Aquadro, C. F. (2006). Finding of male-killing *Spiroplasma* infecting *Drosophila melanogaster* in Africa implies transatlantic migration of this endosymbiont. *Heredity* 97, 27–32. doi: 10.1038/sj.hdy.6800830
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Rattner, R., Thapa, S. P., Dang, T., Osman, F., Selvaraj, V., Maheshwari, Y., et al. (2021). Genome analysis of *Spiroplasma citri* strains from different host plants and its leafhopper vectors. *BMC Genomics* 22:373. doi: 10.1186/s12864-021-07637-8
- Romero, D., and Palacios, R. (1997). Gene amplification and genomic plasticity in prokaryotes. *Annu. Rev. Genet.* 31, 91–111. doi: 10.1146/annurev.genet.31.1.91
- Salzberg, S. L., Hotopp, J., Delcher, A. L., Pop, M., Smith, D. R., Eisen, M. B., et al. (2005). Serendipitous discovery of Wolbachia genomes in multiple *Drosophila* species. *Genome Biol.* 6:R23. doi: 10.1186/gb-2005-6-3-r23
- Scholz, M., Albanese, D., Tuohy, K., Donati, C., Segata, N., and Rota-Stabelli, O. (2020). Large scale genome reconstructions illuminate Wolbachia evolution. *Nat. Commun.* 11:5235. doi: 10.1038/s41467-020-19016-0
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Steenwyk, J. L., Buida, T. J., Li, Y., Shen, X.-X., and Rokas, A. (2020). ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol.* 18:e3001007. doi: 10.1371/journal.pbio.3001007
- Stirpe, F. (2004). Ribosome-inactivating proteins. *Toxicon* 44, 371–383. doi: 10.1016/j.toxicon.2004.05.004
- Sullivan, J. T., and Ronson, C. W. (1998). Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5145–5149. doi: 10.1073/pnas.95.9.5145
- Tang, W.-J., and Guo, Q. (2009). The adenylyl cyclase activity of anthrax edema factor. *Mol. Asp. Med.* 30, 423–430. doi: 10.1016/j.mam.2009.06.001
- Teixeira, L., Ferreira, Á., and Ashburner, M. (2008). The bacterial Symbiont Wolbachia induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol.* 6:e1000002.e2. doi: 10.1371/journal.pbio.1000002
- Tinsley, M. C., and Majerus, M. E. N. (2006). A new male-killing parasitism: *Spiroplasma* bacteria infect the ladybird beetle *Anisosticta novemdecimpunctata* (Coleoptera: Coccinellidae). *Parasitology* 132, 757–765. doi: 10.1017/S0031182005009789
- Vera-Ponce León, A., Dominguez-Mirazo, M., Bustamante-Brito, R., Higuera-Alvarez, V., Rosenbluth, M., and Martínez-Romero, E. (2021). Functional genomics of a *Spiroplasma* associated with the carmine cochineals *Dactylopius coccus* and *Dactylopius opuntiae*. *BMC Genomics* 22:240. doi: 10.1186/s12864-021-07540-2
- Waterworth, S. C., Flórez, L. V., Rees, E. R., Hertweck, C., Kaltenpoth, M., and Kwan, J. C. (2020). Horizontal gene transfer to a defensive Symbiont with a reduced genome in a multipartite beetle microbiome. *mBio* 11:e02430-19. doi: 10.1128/mBio.02430-19
- Weng, Y., Francoeur, C. B., Currie, C. R., Kavanaugh, D. H., and Schoville, S. D. (2021). A high-quality carabid genome assembly provides insights into beetle genome evolution and cold adaptation. *Mol. Ecol. Resour.* 21, 2145–2165. doi: 10.1111/1755-0998.13409
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Xie, J., Butler, S., Sanchez, G., and Mateos, M. (2014). Male killing *Spiroplasma* protects *Drosophila melanogaster* against two parasitoid wasps. *Heredity* 112, 399–408. doi: 10.1038/hdy.2013.118
- Xie, J., Vilchez, I., and Mateos, M. (2010). *Spiroplasma* bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. *PLoS One* 5:e12149. doi: 10.1371/journal.pone.0012149
- Yeoman, C. J., Brutscher, L. M., Esen, Ö. C., Ibaoglu, F., Fowler, C., Eren, A. M., et al. (2019). Genome-resolved insights into a novel *Spiroplasma* symbiont of the wheat stem sawfly (*Cephus cinctus*). *PeerJ* 7:e7548. doi: 10.7717/peerj.7548
- Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res.* 39, W347–W352. doi: 10.1093/nar/gkr485
- Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., et al. (2018). A completely Reimplemented MPI bioinformatics toolkit with a new HHpred server at its Core. *J. Mol. Biol.* 430, 2237–2243. doi: 10.1016/j.jmb.2017.12.007