Evolution and diversity of inherited Spiroplasma in Myrmica ants

Matthew J. Ballinger, Logan D. Moore and Steve J. Perlman

February 2018

Accepted Manuscript:

Ballinger MJ, Moore LD, Perlman SJ. 2018. Evolution and diversity of inherited *Spiroplasma* in *Myrmica* ants. Applied and Environmental Microbiology. 84 (4): e02299-17.

Evolution and diversity of inherited *Spiroplasma* in *Myrmica* ants

Matthew J. Ballinger[a,1], Logan D. Moore[a], and Steve J. Perlman[a,]

[a]Department of Biology, University of Victoria, Victoria, BC, Canada V8W 2Y2

[1]To whom correspondence should be addressed: mattball@uvic.ca

Abstract (204 words)

Microbial partners play an important role in the biology and ecology of animals. In insects, maternally-transmitted symbionts are especially common and can have host effects ranging from reproductive manipulation to nutrient provisioning and defense against natural enemies. In this study, we report a genus-wide association of *Myrmica* ants with the inherited bacterial symbiont*, Spiroplasma*. We screen *Myrmica* ants collected from the wild, including the invasive European fire ant, *Myrmica rubra*, and find an extraordinarily high prevalence of this symbiont – 8 of 9 species, 42 of 43 colonies, and 250 of 276 individual workers were harboring *Spiroplasma* – only one host species was uninfected. In our screens, each host species carried a distinct *Spiroplasma* strain, and none were infected with more than one strain. All symbionts belong to the citri clade, allied most closely with pathogenic *Spiroplasma* of corn crops and honeybees, and there is strong evidence of host-symbiont persistence across evolutionary timescales. Genome sequencing of two *Spiroplasma* symbionts revealed candidate genes that may play a part in the symbiosis, a nutrient transporter absent from other *Spiroplasma*, and a ribosome-inactivating protein previously implicated in parasite defense. These results together suggest long-term, likely mutualistic relationships atypical of *Spiroplasma*-insect associations with potential significance for broad ecological interactions of *Myrmica*.

Importance (129 words)

Animal-associated microbial symbionts can dramatically affect the biology of their hosts. Identification and characterization of these intimate partnerships remains an essential component of describing and predicting species interactions, especially for invasive host species. Ants perform crucial ecological functions as ecosystem engineers, scavengers, and predators, and ants in the genus *Myrmica* can be aggressive resource competitors and reach high densities in their native and invaded habitats. In this study, a novel symbiosis is identified between *Myrmica* ants and the facultative bacterial symbiont, *Spiroplasma*. Broad host distribution, high frequencies of infection, and host-symbiont codivergence over evolutionary timescales, an uncommon feature of *Spiroplasma* associations, suggest an important, likely mutualistic interaction. Genome sequencing identified highly divergent gene candidates that may contribute to *Spiroplasma*'s role as a possible defensive or nutritional partner in *Myrmica*.

Introduction

It is now well established that most insects harbor maternally inherited bacterial endosymbionts that play critical roles in the ecology and evolution of their hosts (1). Insect lineages that feed exclusively on nutrient-poor diets, such as plant sap or animal blood, typically host obligate nutritional endosymbionts that provide essential vitamins and amino acids. These obligate endosymbionts are often housed in specialized symbiont organs and show patterns of strict and ancient co-diversification with their hosts.

More common still are facultative inherited symbionts, of which the best known is *Wolbachia* (2). While these symbionts are transmitted almost exclusively through females over ecological timescales, they very rarely cospeciate with their hosts, and instead, repeatedly colonize new host lineages over evolutionary timescales via horizontal transmission. Although not essential for host survival or reproduction, many facultative inherited symbionts increase host fitness under certain conditions, for example by protecting their hosts against natural enemies or environmental stresses (3–8). Others manipulate their host's reproduction in order to increase the frequency of symbiont-infected females (9–13). Five bacterial lineages are particularly widespread as facultative symbionts of insects. In addition to *Wolbachia*, these are *Arsenophonus*, *Cardinium*, *Rickettsia*, and *Spiroplasma* (14, 15). Initial surveys have found that these bacteria infect ~5-30% of insect species – this represents millions of infected species. However, most insect lineages have been poorly sampled, and even when infections have been reported, it is often not understood how these facultative symbionts affect host fitness or persist in host populations.

Spiroplasma is an incredibly diverse genus of bacteria that infect arthropods, with a wide range of fitness effects and transmission strategies (16). Many *Spiroplasma* are pathogenic, including pathogens of bees, crayfish, and plants (17–20). Even more prevalent are horizontally transmitted gut commensals that have been isolated from a wide range of insects, including beetles and flies (21, 22). Finally, maternal transmission has evolved independently in a number of *Spiroplasma* lineages. Vertically transmitted *Spiroplasma* can be found both inside and outside cells, often at high densities in insect hemolymph, as well as in ovarian tissues (23). While the effects of most vertically transmitted *Spiroplasma* are not known, a number of strains manipulate their host's reproduction by killing male embryos; male-killing *Spiroplasma* strains have been documented in butterflies, planthoppers, beetles, flies, and lacewings (10, 11, 24–26). Some *Spiroplasma* protect their hosts against natural enemies, with strains that infect aphids providing protection against pathogenic fungi (27), and strains that infect *Drosophila* flies protecting against parasitic wasps and nematodes (4, 8). Recent studies have implicated a diverse arsenal of toxins called ribosome-inactivating proteins (RIPs) in *Drosophila* defense (28, 29). Identification of a *Spiroplasma*-encoded RIP transcript in the publically-available transcriptome of the European invasive fire ant, *Myrmica rubra*, motivated closer examination of the relationship between *Spiroplasma* and this ant genus in the present study.

Although *Spiroplasma* infects a wide range of arthropods, few studies have examined a specific group of hosts in detail. The best studied inherited *Spiroplasma* are those that infect *Drosophila*. At least 18 species have been found to harbor *Spiroplasma* (8, 30, 31), with infection frequencies ranging from less than 5% to greater than 85%

(32, 33). *Drosophila* flies have been independently colonized by five different lineages of inherited *Spiroplasma*, from the citri, poulsonii, ixodetis, and tenebrosa clades (30). In this study, ant species in the genus *Myrmica* were surveyed for *Spiroplasma*. This genus also appears to be a hotspot for *Spiroplasma* infection, with all species except one infected at high frequency. Unlike *Drosophila*, however, *Spiroplasma* infecting *Myrmica* are all members of the citri clade, and there is a strong phylogenetic signal suggesting persistent host-symbiont associations across evolutionary timescales.

Results

Spiroplasma symbionts are widespread in *Myrmica*

Nine species of *Myrmica*, broadly distributed across the genus (Fig 1), were screened for *Spiroplasma* by PCR amplification of the *ftsZ* gene; all but one were positive (Table 1). *Spiroplasma* genes were also detected in all three publicly available *Myrmica* transcriptomes (*M. rubra*, *M. ruginodis*, and *M. sulcinodis;* NCBI BioProject PRJDB4088). Among screened species, the prevalence of infection was high, with 250 out of 276 individuals, and 42 out of 43 colonies, testing positive. *COI* failed to amplify or yield quality sequence from nine of the 284 DNA extractions and these samples were excluded from prevalence calculations. In one case, *ftsZ* amplified from a *Myrmica* sp. sample in which *COI* failed and this sample was conservatively excluded from subsequent analysis. For the two best sampled species, *M. rubra* and *M. scabrinodis* (two mtDNA haplotypes), infection frequencies were 86% and 96%, respectively, and all colonies were infected. Infection frequencies were similarly high in juvenile stages, with 9 of 10 larvae and 8 of 9 pupae from one *M. scabrinodis* colony testing positive. Many of the species in our data set are represented by a single colony or individual, yet in most

cases *Spiroplasma* was consistently detected despite limited sampling; however, high prevalence for these species should not be assumed until it can be demonstrated through a similarly thorough sampling effort.

To determine where *Spiroplasma* infection is localized, DNA extractions from the head, thorax, gaster, and legs of adult *Myrmica vandeli* and *M. scabrinodis* were screened; all tissue types were positive, indicating *Spiroplasma* is present in the hemolymph and is not restricted to the gut.

Spiroplasma-host specificity and evolutionary relationships

Phylogenetic analysis of symbiont *ftsZ* sequences places all of the *Myrmica Spiroplasma* strains in the citri clade (Fig 2A), although they are not monophyletic. No *Spiroplasma* strain was shared between species. Two distinct *ftsZ* sequences were recovered from the published transcriptome of *M. ruginodis*, suggesting a coinfection, although this was not examined in greater detail, as *M. ruginodis* samples were not screened. *Myrmica Spiroplasma* form three clades that appear to correspond with host species groups, one with members of the scabrinodis group, one with members of the fracticornis group and allies, and a third with *M. rubra* and *M. alaskensis*. Unlike the *Spiroplasma* of *Myrmica*, those of other ants are more broadly distributed throughout the genus *Spiroplasma* (Fig S1). ParaFit was used to perform a global test of host-symbiont codivergence among all ten distinct host lineages that were screened plus *M. ruginodis*. The null hypothesis of independent host and symbiont evolution, was not rejected at a significance threshold of .05 (Fig 2B; $p = .07$). Exclusion of *M. ruginodis* and its dual *Spiroplasma* strains resulted in rejection of the null hypothesis, though at a marginally significant $p = .02$.

The most thorough sampling was from Lac Remoray, France, where 22, 2, and 3 colonies of *M. scabrinodis*, *M. vandeli*, and *Formica picea* were collected within meters of each other. *Myrmica scabrinodis* contained two distinct mitochondrial haplotypes (96.2% similar at *COI*), and each of these haplotypes harbored its own *Spiroplasma ftsZ* haplotype (99.2% similar at *ftsZ*). Sixteen colonies had one mitochondrial haplotype, six had the other, and no colony had both. Sanger sequencing of a fragment of the long wavelength rhodopsin gene, as well as Illumina sequencing, confirmed that these two mitochondrial haplotypes are one species (i.e. there were no differences in nuclear genes). The two *M. vandeli* colonies harbored a distinct *Spiroplasma* strain that was 99.0-99.2% similar at *ftsZ* to the *Spiroplasma* in *M. scabrinodis*. *Spiroplasma* was absent from the three *Formica picea* colonies (n=26 individuals), further highlighting the absence of lateral transfer of *Spiroplasma* symbionts among microsympatric hosts.

Lastly, two of the *M. scabrinodis* colonies were initially keyed as *M. martini*, a species that was only recently described based on complex morphometrics (34), with no clear morphological features distinguishing it from *M. scabrinodis* (and with the authors' discriminant function misclassifying 10% of individuals). Molecular data from both colonies – mitochondrial and nuclear loci, as well as the symbiont locus – are identical to those of the other 14 *M. scabrinodis* haplotype B colonies in our study, suggesting *M. martini* is not a valid species.

Genome content of *Spiroplasma* symbionts of *Myrmica*

Spiroplasma genomes were sequenced from two host species, *M. vandeli* and *M. scabrinodis*. Eighty-seven and eighty-six million reads, respectively, were generated from the pooled DNA of five ants per species. Preliminary metagenomes were

assembled from low-GC reads (≤ 31%) and consisted of almost exclusively ant, *Wolbachia*, and *Spiroplasma* contigs by blastp (Table 2). These preliminary *Spiroplasma* contigs were used to improve mapping and assembly of *Spiroplasma* reads in each final assembly (see methods). From the final assemblies, 481 contigs encoding 1,019 proteins and 402 contigs encoding 995 proteins were assigned to the *Spiroplasma* symbionts of *M. scabrinodis* and *M. vandeli*, respectively. 98.4% of *M. scabrinodis* proteins were also identified in *M. vandeli*, and 96.8% in the reciprocal comparison, suggesting that the majority of protein coding genes are represented in our *Spiroplasma* assemblies. As expected, the vast majority of these putative genes also yielded blastp hits to the genomes of *S. citri*, *S. kunkelii*, *S. melliferum* and *S. poulsonii* (Table 2). Genome read coverage for *M. scabrinodis* and *M. vandeli* respectively, was 10.4 and 17.2 (median coverage), and 9.1 and 13.6 (mode coverage)(Fig 3A and B). A majority fraction of the top blastp hits for each *Spiroplasma* assembly was to taxa belonging to the citri clade; 70.8% of 1,019 in *M. scabrinodis* and 68.9% of 995 genes in *M. vandeli* (Fig 3C and D). The species receiving the largest fraction of top hits was *Spiroplasma melliferum*, a honey bee pathogen closely allied with the plant pathogens *S. citri* and *S. kunkelii*. Genome sequencing facilitated a more thorough comparison of nucleotide identity between strains than the *ftsZ* locus alone – across 30 kb of syntenous coding and intergenic sequence the two share 95% identity.

Genes that are unique to these strains relative to other *Spiroplasma* taxa may hint toward the biological role of *Spiroplasma* in *Myrmica*. Hypothetical ORFs located on the same contig as a *Spiroplasma* gene were translated and queried by blastp and HMMER against the nr protein database and reference proteomes. Using a

188 conservative minimum of 600 nucleotides for ORF prediction, nine and eight candidates
189 from *M. scabrinodis* and *M. vandeli*, respectively, were identified. All but one returned
190 no significant similarity or domain conservation to known proteins. The exception
191 encodes a nutrient transporter gene that is absent from all other sequenced
192 *Spiroplasma* genomes: the substrate component of an energy-coupling factor (ECF)
193 membrane transporter. No disruption in read coverage was evident between the ECF
194 transporter and the *Spiroplasma* genes flanking it, it uses the *Mycoplasma/Spiroplasma*
195 genetic code, and is also present in the transcriptome of *M. sulcinodis*, suggesting it is
196 not an artifact of contaminating sequence reads in the assembly. PCR screens
197 confirmed its presence in *M. vandeli* and both *M. scabrinodis* strains, but did not yield
198 amplicons from *Myrmica* specimens from outside of the scabrinodis species group.
199 Phylogenetic analysis alongside the most similar blastp matches and ECF transporter
200 gene families of other *Spiroplasma* taxa placed the putative novel transporter on a long
201 branch, distantly related to characterized families (Fig 4).

202 **Ribosome-inactivating proteins in *Myrmica Spiroplasma* symbionts**

203 Ribosome-inactivating protein (RIP) coding regions were identified in each of the
204 *Spiroplasma* genomes and in the transcriptome of *M. rubra*. The RIP of *M. vandeli* is not
205 predicted to encode a functional protein due to reading frame disruptions, while the
206 RIPs of *M. scabrinodis* and *M. rubra* appear to encode intact ORFs with conserved
207 active site residues, though the latter is only partially represented (~60% of the gene).
208 Phylogenies of RIPs are not congruent with hosts (Fig 5), as was found in other
209 *Spiroplasma* (35). The RIP of *M. scabrinodis* assembled into a 14 kb contig with an
210 order of magnitude greater read coverage than that of *M. vandeli*, suggesting copy

number variation between the two (Fig 3A and B). Other genes encoded on this contig include those with strong amino acid similarity to proteins involved in type IV secretion systems used for conjugative DNA transfer and encoded on plasmids of *S. citri* and *S. kunkelii* (36, 37), including *soj*, *mob*, and *traE*.

Presence of RIPs was confirmed by PCR for both *M. vandeli* and *M. scabrinodis*, while reactions targeting the *M. rubra* RIP failed to amplify from our Toronto, Vancouver, and Victoria samples. Detection of RIPs from individual colonies of scabrinodis group hosts varied by *Spiroplasma* strain. The RIP pseudogene was detected in workers from each of the two colonies of *M. vandeli*, while the intact RIP was detected in all of the colonies of *M. scabrinodis* bearing *Spiroplasma* haplotype B but none of the haplotype A colonies.

Discussion

In this study, *Myrmica* ants are shown to be a hotspot for *Spiroplasma* infection. Eight of nine species screened in the current study, as well as all three species with publicly available transcriptomes, harbor *Spiroplasma*. In addition, infection prevalence within species is high, with 42 of 43, and 250 of 276 colonies and individuals respectively, infected. No strains were shared between multiple ant species, while one species, *M. ruginodis*, hosted two different strains. Other broad surveys of symbionts in ants, using universal 16S ribosomal RNA primers, have also reported *Spiroplasma* infections from multiple species groups, including citri, ixodetis and platyhelix (Fig S1), in 27 of 95 species (38) and in 24 of 464 species (39). In this latter study, one half of infections were in the genus *Polyrachis*. Although these screens do not typically distinguish between inherited and horizontally transmitted *Spiroplasma*, or provide

234 information about prevalence within ant species, they suggest that *Spiroplasma*
235 infections may not be uncommon in ants.

236 The perfect association between mitochondrial haplotype and *Spiroplasma*
237 variant, which was found in our detailed screening of microsympatric colonies of *M.*
238 *scabrinodis* and *M. vandeli*, is strong evidence for symbiont vertical transmission. Two
239 mitochondrial haplotypes were found in *M. scabrinodis* (96% similar at *COI*), and these
240 are perfectly associated with *Spiroplasma* haplotypes, as indicated by variation in the
241 *ftsZ* gene (99% similar). Individuals (and colonies) with different mitochondrial
242 haplotypes showed no differences in their nuclear genes, and it is not known how or
243 why this mitochondrial polymorphism persists. This pattern of genetic variation is in
244 contrast with *M. rubra*, where nuclear but *not* mitochondrial differences are associated
245 with a queen reproductive polymorphism (40, 41). The close relative *M. vandeli* harbors
246 a closely related *Spiroplasma* strain (95% similar to *M. scabrinodis* symbionts, 99% at
247 *ftsZ* alone). It is interesting that these ants do not share or appear to exchange
248 *Spiroplasma*, despite being found so close to each other, and is in contrast with other
249 studies finding the exchange of *Wolbachia* between socially parasitic ants and those
250 within the host colonies (42–44). Interestingly, one of these studies found that unlike
251 *Wolbachia*, *Spiroplasma* strains were not exchanged between host and social parasite
252 (44). Finally, the high infection prevalence in *Myrmica* larvae and pupae, and
253 widespread tissue distribution, also suggest vertical transmission.

254 All *Spiroplasma* in the present study were from the citri clade. A 16S rRNA
255 screen also found a strain from the citri clade infecting *Myrmica incompleta* (39). The
256 *Myrmica* symbionts are not monophyletic; plant pathogens *S. kunkellii* and *S. citri*, bee

257 pathogenic *S. melliferum*, and symbionts of *Drosophila wheeleri, D. aldrichi*, and *D.*
258 *mojavensis* are all nested within the group of *Myrmica Spiroplasma*. Although there is
259 not strict cospeciation between *Myrmica* and their symbionts, there is strong
260 phylogenetic signal, with three lineages of *Spiroplasma* each closely associated with a
261 lineage of *Myrmica*, although much more detailed sampling and screening of *Myrmica* is
262 required to determine how many independent acquisitions of *Spiroplasma* have
263 occurred. This will be challenging, as *Myrmica* is a very diverse clade of ants whose
264 taxonomy and evolutionary relationships are unresolved and sometimes controversial,
265 with many cryptic species (40, 45–47). We know of no cases of cospeciation between
266 *Spiroplasma* with their hosts, and in general cospeciation between facultative symbionts
267 and their hosts is very rare (see ref (48) for a *Wolbachia* example). Most facultative
268 symbionts are lost before their hosts speciate, and infect new hosts via horizontal
269 transmission (30, 49, 50).

270 The patterns of association between *Spiroplasma* and *Myrmica* differ greatly from
271 *Drosophila*, the best studied insect lineage that is commonly infected by inherited
272 *Spiroplasma*. There, at least five lineages from four clades have colonized *Drosophila*,
273 and hosts are often infected at low frequencies, largely depending on maternal
274 transmission efficiency, as well as the fitness and phenotypic effects of the symbiont.
275 For example, male-killing strains are found at low frequencies in host populations (51),
276 whereas a strain that protects against a very common virulent nematode parasite
277 occurs at high frequency (33).

278 Of course, the obvious next step is to determine what effects *Spiroplasma* might
279 have on their *Myrmica* hosts. It is unlikely that these microbes are essential, as some

280 individuals (and one species) were uninfected. Obligate inherited symbionts of ants
281 include *Blochmannia* that recycle nitrogen for their carpenter ant hosts (52); *Myrmica*
282 ants, however, are primarily predaceous and not thought to feed on nutrient-limited diets
283 (53). Also, no *Spiroplasma* are known to be obligate symbionts.

284 Perhaps *Spiroplasma* manipulates *Myrmica* reproduction, for example by killing
285 males. It is challenging though to demonstrate sex ratio distortion in ants and other
286 social Hymenoptera, as this would involve isolating symbiont-free queens, rearing
287 colonies to produce reproductives, and comparing their sex ratios with those of infected
288 colonies. As far as we are aware, only one study has shown a convincing link between
289 symbionts and sex ratio distortion in ants (54). In that study, artificial selection on sex
290 ratio in colonies of the pharaoh ant *Monomorium pharaonis* resulted in rapid changes in
291 the frequency of *Wolbachia*.

292 Another possibility is that *Spiroplasma* persist in *Myrmica* by providing protection
293 against natural enemies. *Myrmica* are commonly infected with parasites (55). In fact,
294 new species of *Myrmica* have been erroneously described due to parasitic nematodes,
295 because infected ants often look different from uninfected ones, with distended
296 abdomens (56, 57). To explore the potential for protection, we sequenced *Spiroplasma*
297 genomes and surveyed for ribosome-inactivating proteins (RIPs), toxins that are
298 widespread and diverse in *Spiroplasma* and that have been implicated in defense
299 against parasitic nematodes and wasps (28, 29). These toxins appear to evolve rapidly
300 and exhibit elevated rates of gains and losses in *Spiroplasma*, making initial detection
301 by PCR difficult, even with degenerate primers. However, our genome surveys
302 uncovered two RIPs. One of these, in the *M. vandeli* symbiont, is a pseudogene, while

303 the other, in the *M. scabrinodis* symbiont, appears to lie on a plasmid that is present in
304 the strain associated with haplotype B, but not haplotype A. That the RIP toxins are
305 pseudogenized or found on plasmids suggests that their host associations are dynamic,
306 perhaps evolving in concert with changing pressures from natural enemies. In support
307 of this, we also uncovered a *Spiroplasma* RIP from the transcriptome of *M. rubra*
308 collected from its native range in Europe, but we could not detect it in N. American
309 colonies, suggesting that it also occurs on a plasmid and was lost when *M. rubra*
310 invaded N. America, perhaps due to enemy release. Of course, much work remains in
311 order to determine whether these RIPs might be protective and against what.
312 Genomes can also provide useful clues for understanding symbiont biology (1).
313 We searched the *Myrmica*-*Spiroplasma* metagenome for novel *Spiroplasma* genes, i.e.
314 genes that do not occur in any sequenced *Spiroplasma* genomes, of which 21 are
315 currently available, from four clades, including *S. melliferum*, *S. citri* and *S. kunkellii*
316 from the citri clade. All but one of the new genes were of unknown function, and there
317 were no new metabolic pathways uncovered, further suggesting that the symbiont does
318 not fill an obligate nutritional or metabolic role for its host. We identified a divergent ECF
319 S (substrate) component gene responsible for conferring substrate specificity to a
320 transport complex. In shared energy-coupling transport systems, an ATPase binding
321 cassette and transmembrane protein, the so-called AT module, is a universal
322 component of each transporter, with substrate specificity being conferred by S
323 component genes (58). Substrates include vitamins and transition metals, therefore, it is
324 possible that *Spiroplasma* symbionts of *Myrmica* are supplementing or siphoning
325 nutrients. Phylogenetic analysis can help to identify candidate protein functions among

members of specialized protein classes; however, this gene could not be conclusively attributed to a characterized substrate family, thus its role and importance in this symbiosis remain open questions.

Finding such prevalent inherited *Spiroplasma* in *Myrmica* opens up many interesting questions. The next step is to perform experiments comparing symbiont-infected and uninfected ants, following antibiotic treatment of lab colonies. A promising model would be the European fire ant, *Myrmica rubra,* one of the most invasive ant species globally. Interestingly, a little cited study from twenty years ago treated *M. rubra* lab colonies with antibiotics, and found an effect on ant growth and queen production (59).

Materials and Methods

Sample collection, DNA extraction, and *Spiroplasma* screening

Individuals from twenty-nine ant colonies from two sites that were 35 m apart (site 1: 46.7594 °N, 6.2527 °E, and site 2: 46.7595 °N, 6.2573 °E) in Réserve Naturelle du Lac de Remoray, France, were collected in August 2016. This area was already known to harbor a number of different *Myrmica* species. Species determinations were made by Mesut Koken, a local ant expert. The number of ants sampled from each colony ranged from 6-20; for one *M. scabrinodis* colony, seven larvae and ten pupae were also collected. Colony samples were stored separately in 95% ethanol. In addition, *Myrmica* samples were received from colleagues as whole ants in ethanol or as DNA extractions (see Acknowledgements and Table S1).

In preparation for DNA extraction, ants were removed from ethanol and air dried for 10 minutes. DNA was extracted from individual ants using the PrepMan Ultra

(Applied Biosystems) sample preparation reagent. To rule out the possibility that *Spiroplasma* was an ectosymbiont harbored on the ant cuticle, all sampled individuals from one *M. scabrinodis* colony were surface sterilized by submerging ant specimens in 2.5% bleach for two minutes, then submerging in 70% ethanol for four minutes, and then rinsing in distilled water twice for three minutes. To test for systemic infection, the head, thorax, gaster and legs of eight ants from one *M. scabrinodis* and one *M. vandeli* colony were carefully dissected, and DNA was extracted and screened separately.

DNA extractions were screened for *Spiroplasma* by PCR using primers targeting a 780 bp fragment of the single copy cell division protein gene *ftsZ* (F2: 5' TGAACAAGTCGCGTCAATAAA and R2: 5' CCACCAGTAACATTAATAATAGCATCA (30)). Some initial screens were also performed using primers targeting 300 bp of *Spiroplasma* 16S rRNA (Spi16S-F: 5' CCTGAGTAGTATGCTCGCAAGAG and Spi16S-R: 5' CCCACCTTCCTCTAGCTTAC). Primers targeting the *Myrmica* host gene, mitochondrially-encoded cytochrome oxidase subunit I (*COI*) were used as a positive control for DNA quality and to sequence a region of *Myrmica* DNA for molecular identification and analysis. Primer sequences are MyrmCOI F: 5' TCGTTTAGAATTAGGATCTTGT and R: 5' ATGAGAAATTAATCCAAATCCAG for species in the scabrinodis group taxa, uMyrmicaCOI F: 5' TAATTAATAATGAYCAAATTTATAATAC and R: 5' GTRGGRATTGCAATAATTATAGTTGC for all other Myrmica taxa, as well as LCO1490: 5' TAAACTTCAGGGTGACCAAAAAATCA and HCO2198: 5' GGTCAACAAATCATAAAGATATTGG (60) for *Formica*. Primers targeting 500 bp of a variable portion of the nuclear encoded long wavelength rhodopsin gene (LR143F: 5'

CACTGGTATCARTTCGCACCSAT LR672R and LR672R: 5' CCRCAMGCWGTCATGTTRCCTTC (47)) were used to confirm that divergent *M. scabrinodis* mitochondrial haplotypes corresponded to the same species. RIP primers were scabRIP F: 5' GAGGAACTAAAATTGAAGTAGTTCT and R: 5' AATCTTCATCTTGATACTTGACCAC and vanRIP F: 5' TCCTTGGTTAGATACTATTTCTGCTC and R: 5' ATTATTGAGTTTGAGGTATCGC. ECF transporter primers were Sp-ECF F: 5' CTTAGCAGCTGTAATGTTAGCATTAAC and R: 5' CTAATTCCACAGCCATAAATAAAGTAG. Thermal cycling programs for ftsZ, MyrmCOI, Spi16S, HCO/LCO, LR, RIPs, and ECFs were 35 cycles of 95° C for :30, 54° C for :30 72° C for 1:15 and for uMyrmicaCOI 35 cycles of 95° C for :30, 49° C for :30 72° C for 1:15. PCR products were assessed by DNA gel electrophoresis on a 1% agarose gel stained with ethidium bromide and visualized under UV light. One or more *ftsZ* and *COI* amplicons per *Myrmica* species was sequenced with the Sanger method by Sequetech (California, USA) using forward and reverse primers. At least one *ftsZ* amplicon for each host taxon in our sample set was sequenced. Intrahost symbiont diversity was examined by sequencing *ftsZ* and *COI* from all individuals from one *M. vandeli* and two *M. scabrinodis* colonies (n=10, 9, 8 individuals, repectively).

Sequence processing and phylogenetic analysis

Primer regions were trimmed from the Sanger sequences by hand, yielding a final product of 648 bases for *ftsZ* and 690 bases for *COI*. *ftsZ* sequences were used to query nucleotide sequences deposited in the NCBI transcriptome shotgun assembly (TSA) database. For *Myrmica* taxa with TSA hits to *Spiroplasma ftsZ* sequences, *COI* sequences were also recovered from the transcriptome by blastn. Nucleotide and amino

395 acid sequences were aligned with MAFFT 7.309 (61). For each nucleotide alignment,
396 the best substitution model for phylogenetic analysis was determined in jModelTest
397 2.1.7 (62). For *ftsZ*, the best model was HKY+I+G and for *COI* it was TPM2uf+G. RIP
398 and ECF transporter phylogenies were built from amino acid alignments and used the
399 LG substitution model. Maximum likelihood phylograms were constructed using PhyML
400 2.2.0 (63) implemented in Geneious R10. SH-like approximate likelihood ratio test
401 scores were calculated for each branch. Alignments and phylogenies used to test for
402 evidence of cophylogeny were generated as described above. Tests of codivergence
403 were carried out with ParaFit (64) implemented in the R package ape (65). Briefly,
404 patristic distance matrices were calculated for the host *COI* and symbiont *ftsZ* loci using
405 the same models of nucleotide substitution as above. Included in these alignments were
406 all the hosts that were positive for *Spiroplasma* by our PCR screens, ten distinct host
407 mitochondrial lineages and their symbionts, plus *M. ruginodis* host and *Spiroplasma*
408 sequences from the public transcriptome. Distance matrices were permuted randomly to
409 create a distribution of data against which to test the null hypothesis of independent
410 host and symbiont evolutionary histories. Matrices were permuted 999 times. Because
411 infection frequency in *M. ruginodis* could not be determined nor could read
412 contamination in the sequence read archive be ruled out, the codivergence analysis
413 was carried out once with and once without this host. The tanglegram depicted in Fig 2
414 was visualized with TreeMap 3 (66).

415 ***Spiroplasma* genome DNA extractions, sequencing, and assembly**

416 To prepare samples for Illumina sequencing, genomic DNA was extracted from
417 pools of five ants for each of two species in the *scabrinodis* clade, *Myrmica scabrinodis*

418 and *M. vandeli*, using the phenol-chloroform method. Short insert shotgun libraries were
419 prepared and 125 bp paired-end reads were sequenced by Genome Québec (Montréal,
420 Québec, Canada) on a HiSeq 2500 v4 system.

421 Preliminary metagenomes were assembled for each of the two *Myrmica* host
422 species, *M. vandeli* and *M. scabrinodis* after filtering out reads with GC content greater
423 than 31% and their pairs. The low GC data set was mapped to the mitochondrial
424 genome sequence of *Myrmica scabrinodis* (NCBI Reference Sequence NC_026133)
425 and removed. Reads were trimmed, filtered, and mapped with BBMap 37.36 (by Brian
426 Bushnell, sourceforge.net/projects/bbmap/). Filtered reads were assembled de novo
427 using SPAdes 3.10.1 (67). Open reading frames with a minimum length of 300
428 nucleotides were predicted and translated with Geneious R10 and compared against
429 the nr protein database available on NCBI using blastp. All contigs encoding
430 *Spiroplasma* genes, i.e. sequences for which the top blastp hit was to *Spiroplasma*,
431 were retained as a preliminary genome assembly. *Spiroplasma* genes with higher GC
432 content, though rare, were absent from these assemblies. To produce a more complete
433 assembly for each symbiont strain, a second iteration of mapping and assembly was
434 carried out. Using the original read sets, reads were quality trimmed, and GC filtered at
435 45%. All remaining reads that mapped to contigs in the preliminary assemblies, as well
436 as to the sequenced genome of *Spiroplasma citri* and its plasmids were again
437 assembled de novo using SPAdes. The blastp search was repeated to identify *Myrmica*
438 and bacterial contigs that had not been removed through the mapping procedure. A
439 collection of genes with top hits to either non-*Spiroplasma* or unclassified
440 Entomoplasmatales taxa, 56 in *M. vandeli* and 53 in *M. scabrinodis*, were interpreted as

Spiroplasma genes nonetheless if they were encoded on *Spiroplasma* contigs and showed strong blastp hits to *Spiroplasma* taxa as well.

To identify genes unique to the *Myrmica Spiroplasma*, proteins with top blastp matches to *Spiroplasma* were annotated back onto the assembly contigs and compared against all predicted ORFs on each contig, i.e. most ORFs had two annotations, one as a predicted protein coding gene and one as a *Spiroplasma* blastp match. ORFs with only the former annotation were investigated individually with blastp, blastn, and HMMER to identify putative functions. RIPs were identified by tblastn to the final assemblies and to the three *Myrmica* transcriptomes and sequence read archives available as part of NCBI BioProject PRJDB4088.

Acknowledgments

We thank Megan Frederickson, Rob Higgins, Danielle Hoefele, Caroline Cameron, Corrie Moreau, and Jake Russell for donating *Myrmica* specimens, and Jocelyn Claude and Mesut Koken for assisting in collecting and identifying specimens. We also thank Ryan Gawryluk and Cuong Le for technical advice. This work was funded by a Sinergia grant from the Swiss National Science Foundation awarded to SJP and a Jamie Cassels Undergraduate Research Award from the University of Victoria to LDM.

Data Availability

Genomic DNA sequence reads and PCR amplicon sequences generated during this study (68) have been submitted to GenBank under BioProject PRJNA419549 and accession numbers MG558353-MG558456.

References

1. Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. Ann Rev Genet 42:165–190.
2. Zug R, Hammerstein P. 2012. Still a host of hosts for Wolbachia: Analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PLoS One 7.
3. Oliver KM, Russell JA, Moran NA, Hunter MS. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. Proc Natl Acad Sci U S A 100:1803–7.
4. Xie J, Vilchez I, Mateos M. 2010. Spiroplasma bacteria enhance survival of Drosophila hydei attacked by the parasitic wasp Leptopilina heterotoma. PLoS One 5.
5. Kellner RLL, Dettner K. 1996. Differential efficacy of toxic pederin in deterring potential arthropod predators of Paederus (Coleoptera: Staphylinidae) offspring. Oecologia 107:293–300.
6. Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. 2008. Wolbachia and virus protection in insects. Science (80-) 322:702.
7. Teixeira L, Ferreira Á, Ashburner M. 2008. The bacterial symbiont Wolbachia induces resistance to RNA viral infections in Drosophila melanogaster. PLoS Biol 6:2753–2763.
8. Jaenike J, Unckless R, Cockburn SN, Boelio LM, Perlman SJ. 2010. Adaptation via symbiosis: recent spread of a Drosophila defensive symbiont. Science 329:212–5.
9. Hurst GDD, Jiggins FM, Hinrich Graf von der Schulenburg J, Bertrand D, West

SA, Goriacheva II, Zakharov IA, Werren JH, Stouthamer R, Majerus MEN. 1999. Male-killing Wolbachia in two species of insect. Proc R Soc B Biol Sci 266:735–740.

10. Jiggins FM, Hurst GD, Jiggins CD, von der Schulenburg JH, Majerus ME. 2000. The butterfly Danaus chrysippus is infected by a male-killing Spiroplasma bacterium. Parasitology 120:439–46.

11. Majerus TMO, Graf Von Der Schulenburg JH, Majerus MEN, Hurst GDD. 1999. Molecular identification of a male-killing agent in the ladybird Harmonia axyridis (Pallas) (Coleoptera: Coccinellidae). Insect Mol Biol 8:551–555.

12. Stouthamer R, Breeuwert JA, Luck RF, Werren JH. 1993. Molecular identification of microorganisms associated with parthenogenesis. Nature 361:66–8.

13. Rousset F, Bouchon D, Pintureau B, Juchault P, Solignac M. 1992. Wolbachia Endosymbionts Responsible for Various Alterations of Sexuality in Arthropods. Proc R Soc B Biol Sci 250:91–98.

14. Duron O, Bouchon D, Boutin SSS, Bellamy L, Zhou L, Engelstadter J, Hurst GD, Engelstädter J. 2008. The diversity of reproductive parasites among arthropods: Wolbachia do not walk alone. BMC Biol 6:27.

15. Perlman SJ, Hunter MS, Zchori-Fein E. 2006. The emerging diversity of Rickettsia. Proc R Soc B Biol Sci 273:2097–2106.

16. Gasparich GE, Whitcomb RF, Dodge D, French FE, Glass J, Williamson DL. 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.

509 17. Ding Z, Tang J, Xue H, Li J, Ren Q, Gu W, Meng Q, Wang W. 2014. Quantitative
510 detection and proliferation dynamics of a novel Spiroplasma eriocheiris pathogen
511 in the freshwater crayfish, Procambarus clarkii. J Invertebr Pathol 115:51–54.
512 18. Clark TB, Whitcomb RF, Tully JG, Mouches C, Saillard C, Bove JM, Wroblewski
513 H, Carle P, Rose DL, Henegar RB, Williamson DL. 1985. Spiroplasma melliferum,
514 a New Species from the Honeybee (Apis mellifera). Int J Syst Bacteriol 35:296–
515 308.
516 19. Whitcomb RF, Chen TA, Williamson DL, Liao C, Tully JG, Bové JM, Mouches C,
517 Rose DL, Coan ME, Clark TB. 1986. Spiroplasma kunkelii sp. nov.:
518 Characterization of the Etiological Agent of Corn Stunt Disease. Int J Syst Evol
519 Microbiol 36:170–178.
520 20. Saglio P, Lhospital M, Lafleche D, Dupont G, Bove JM, Tully JG, Freundt E a.
521 1973. Spiroplasma citri gen. and sp. n.: A Mycoplasma-Like Organism Associated
522 with “Stubborn” Disease of Citrus. Int J Syst Bacteriol 23:191–204.
523 21. Hackett KJ, Whitcomb RF, French FE, Tully JG, Gasparich GE, Rose DL, Carle
524 P, Bové JM, Henegar RB, Clark TB, Konai M, Clark E a, Williamson DL. 1996.
525 Spiroplasma corruscae sp. nov., from a firefly beetle (Coleoptera: Lampyridae)
526 and tabanid flies (Diptera: Tabanidae). Int J Syst Bacteriol 46:947–50.
527 22. Whitcomb RF, French FE, Joseph G, Gasparich GE, Rose DL, Carle P, Bove JM,
528 Henegar RB, Konai M, Hackett KJ, Adams JR, Williamson DL, Truman B. 1997.
529 Spiroplasma chrysopicola sp. nov., Spiroplasma gladiatoris sp. nov., Spiroplasma
530 helicoides sp . nov ., and Spiroplasma tabanidicola sp . nov ., from Tabanid (
531 Diptera : Tabanidae) Flies. Int J Syst Bacteriol 47:713–719.

532 23. Sakaguchi B, Poulson DF. 1961. Distribution of “sex-ratio” agent in tissues of
533 Drosophila willistoni. Genetics 46:1665–1676.
534 24. Hayashi M, Watanabe M, Yukuhiro F, Nomura M, Kageyama D. 2016. A
535 nightmare for males? A maternally transmitted male-killing bacterium and strong
536 female bias in a green lacewing population. PLoS One 11.
537 25. Sanada-Morimura S, Matsumura M, Noda H. 2013. Male killing caused by a
538 Spiroplasma symbiont in the small brown planthopper, Laodelphax striatellus. J
539 Hered 104:821–829.
540 26. Poulson DF, Sakaguchi B. 1961. Nature of “sex-ratio” agent in Drosophila.
541 Science 133:1489–90.
542 27. Łukasik P, van Asch M, Guo H, Ferrari J, Godfray HCJ. 2013. Unrelated
543 facultative endosymbionts protect aphids against a fungal pathogen. Ecol Lett
544 16:214–218.
545 28. Hamilton PT, Peng F, Boulanger MJ, Perlman SJ. 2015. A ribosome-inactivating
546 protein in a Drosophila defensive symbiont. Proc Natl Acad Sci U S A
547 113:1518648113-.
548 29. Ballinger MJ, Perlman SJ. 2017. Generality of toxins in defensive symbiosis:
549 Ribosome-inactivating proteins and defense against parasitic wasps in
550 Drosophila. PLoS Pathog 13.
551 30. Haselkorn TS, Markow TA, Moran NA. 2009. Multiple introductions of the
552 Spiroplasma bacterial endosymbiont into Drosophila. Mol Ecol 18:1294–1305.
553 31. Chandler JA, Lang J, Bhatnagar S, Eisen JA, Kopp A. 2011. Bacterial
554 communities of diverse Drosophila species: Ecological context of a host-microbe

555 model system. PLoS Genet 7.
556 32. Watts T, Haselkorn TS, Moran NA, Markow TA. 2009. Variable incidence of
557 Spiroplasma infections in natural populations of Drosophila species. PLoS One 4.
558 33. Cockburn SN, Haselkorn TS, Hamilton PT, Landzberg E, Jaenike J, Perlman SJ.
559 2013. Dynamics of the continent-wide spread of a Drosophila defensive symbiont.
560 Ecol Lett 16:609–616.
561 34. Seifert B, Yazdi AB, Schultz R. 2014. Myrmica martini sp.n. - A cryptic species of
562 the myrmica scabrinodis species complex (Hymenoptera: Formicidae) revealed
563 by geometric morphometrics and nest-centroid clustering. Myrmecological News
564 19:171–183.
565 35. Ballinger MJ, Perlman SJ. 2016. Generality of toxins in defensive symbiosis:
566 Ribosome-inactivating proteins and parasitoid wasp defense in Drosophila. prep.
567 36. Bai X, Fazzolari T, Hogenhout SA. 2004. Identification and characterization of
568 traE genes of Spiroplasma kunkelii. Gene 336:81–91.
569 37. Joshi BD, Berg M, Rogers J, Fletcher J, Melcher U. 2005. Sequence comparisons
570 of plasmids pBJS-O of Spiroplasma citri and pSKU146 of S. kunkelii: implications
571 for plasmid evolution. BMC Genomics 6:175.
572 38. Kautz S, Rubin BER, Moreau CS. 2013. Bacterial infections across the ants:
573 Frequency and prevalence of Wolbachia, Spiroplasma, and Asaia. Psyche (New
574 York).
575 39. Russell JA, Funaro CF, Giraldo YM, Goldman-Huertas B, Suh D, Kronauer DJC,
576 Moreau CS, Pierce NE. 2012. A Veritable Menagerie of Heritable Bacteria from
577 Ants, Butterflies, and Beyond: Broad Molecular Surveys and a Systematic

578 Review. PLoS One 7.
579 40. Steiner FM, Schlick-Steiner BC, Konrad H, Moder K, Christian E, Seifert B,
580 Crozier RH, Stauffer C, Buschinger A. 2006. No sympatric speciation here:
581 Multiple data sources show that the ant Myrmica microrubra is not a separate
582 species but an alternate reproductive morph of Myrmica rubra. J Evol Biol
583 19:777–787.
584 41. Leppänen J, Seppä P, Vepsäläinen K, Savolainen R. 2015. Genetic divergence
585 between the sympatric queen morphs of the ant Myrmica rubra. Mol Ecol
586 24:2463–2476.
587 42. Dedeine F, Ahrens M, Calcaterra L, Shoemaker DD. 2005. Social parasitism in
588 fire ants (Solenopsis spp.): A potential mechanism for interspecies transfer of
589 Wolbachia. Mol Ecol.
590 43. Liberti J, Sapountzis P, Hansen LH, Sørensen SJ, Adams RMM, Boomsma JJ.
591 2015. Bacterial symbiont sharing in Megalomyrmex social parasites and their
592 fungus-growing ant hosts. Mol Ecol 24:3151–3169.
593 44. Haapaniemi K, Pamilo P. 2015. Social parasitism and transfer of symbiotic
594 bacteria in ants (Hymenoptera: Formicidae). Myrmecological News 21:49–57.
595 45. Radchenko A, Elmes GW. 2003. A taxonomic revision of the socially parasitic
596 Myrmica ants (Hymentoptera: Formicidae) of the Palaearctic region. Ann Zool
597 53:217–243.
598 46. Savolainen R, Vepsalainen K. 2003. Sympatric speciation through intraspecific
599 social parasitism. Proc Natl Acad Sci U S A 100:7169–74.
600 47. Jansen G, Savolainen R, Vepsäläinen K. 2010. Phylogeny, divergence-time

estimation, biogeography and social parasite-host relationships of the Holarctic ant genus Myrmica (Hymenoptera: Formicidae). Mol Phylogenet Evol 56:294–304.

48. Jaenike J, Stahlhut JK, Boelio LM, Unckless RL. 2010. Association between Wolbachia and Spiroplasma within Drosophila neotestacea: An emerging symbiotic mutualism? Mol Ecol 19:414–425.

49. Werren JH, Zhang W, Guo LR. 1995. Evolution and Phylogeny of Wolbachia: Reproductive Parasites of Arthropods. Proc R Soc B Biol Sci 261:55–63.

50. Jaenike J. 2012. Population genetics of beneficial heritable symbionts. Trends Ecol Evol.

51. Montenegro H, Solferini VN, Klaczko LB, Hurst GDD. 2005. Male-killing Spiroplasma naturally infecting Drosophila melanogaster. Insect Mol Biol 14:281–287.

52. Feldhaar H, Straka J, Krischke M, Berthold K, Stoll S, Mueller MJ, Gross R. 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont Blochmannia. BMC Biol 5:48.

53. Fiedler K, Kuhlmann F, Schlick-Steiner BC, Steiner FM, Gebauer G. 2007. Stable N-isotope signatures of central European ants - Assessing positions in a trophic gradient. Insectes Soc 54:393–402.

54. Pontieri L, Schmidt AM, Singh R, Pedersen JS, Linksvayer TA. 2017. Artificial selection on ant female caste ratio uncovers a link between female-biased sex ratios and infection by Wolbachia endosymbionts. J Evol Biol 30:225–234.

55. Witek M, Barbero F, Markó B. 2014. Myrmica ants host highly diverse parasitic

communities: from social parasites to microbes. Insectes Soc.

56. Czechowski W, Radchenko A, Czechowska W. 2007. Mermithid infestation strikingly alters the morphology of myrmica rubba (L.) (Hymenoptera: Formicidae): Possible Taxonomical Involvements. Ann Zool 57:325–330.

57. Csosz S. 2012. Nematode infection as significant source of unjustified taxonomic descriptions in ants (Hymenoptera: Formicidae). Myrmecological News.

58. Rodionov DA, Hebbeln P, Eudes A, Ter Beek J, Rodionova IA, Erkens GB, Slotboom DJ, Gelfand MS, Osterman AL, Hanson AD, Eitinger T. 2009. A novel class of modular transporters for vitamins in prokaryotes. J Bacteriol 91:42–51.

59. Pearson B, Raybould AF. 1998. The effects of antibiotics on the development of larvae and the possible role of bacterial load in caste determination and diapause in Myrmica rubra (Hymenoptera: Formicidae). Sociobiology 31:77–90.

60. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3:294–299.

61. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780.

62. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772–772.

63. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. Syst Biol 59:307–321.

64. Legendre P, Desdevises Y, Bazin E. 2002. A statistical test for host-parasite

coevolution. Syst Biol 51:217–234.

65. Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

66. Charleston MA, Robertson DL. 2002. Preferential host switching by primate lentiviruses can account for phylogenetic similarity with the primate phylogeny. Syst Biol 51:528–535.

67. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., 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.

68. Ballinger MJ, Moore LD, Perlman SJ. 2017. Metagenomic sequence reads of Myrmica scabrinodis and Myrmica vandeli.

662 **Table 1.** *Myrmica* species or species groups screened for *Spiroplasma*

***Myrmica* species**	**Screened**	**Positive**	**Frequency**
	individuals/colonies	individuals/colonies	%
alaskensis	16/1	16/1	100.0
fracticornis	8/1	0/0	0.0
rubra	68/12	58/12	85.2
scabrinodis mtDNA type A	37/6	34/6	91.9
scabrinodis mtDNA type B	126/16	123/16	97.6
sp. 1 (group A; CSM0598 W1)	1/1	1/1	100.0
sp. 2 (group A; CSM0202, CSM0296)	4/2	4/2	100.0
sp. 3 (group B; CSM1794c)	1/1	1/1	100.0
sp. 4 (group B; CSM1798)	2/1	1/1	50.0
vandeli	13/2	12/2	92.3

663

664 **Table 2. Genome assembly statistics of novel *Spiroplasma* genomes**

	Myrmica scabrinodis	*Myrmica vandeli*
Preliminary metagenomes		
Contigs (> 300 nt)	113,714	107,962
N50	593	605
Wolbachia genes	338	574
Final *Spiroplasma* genomes		
Spiroplasma contigs (>300 nt)	481	402
Spiroplasma N50	3477	4471
Spiroplasma nucleotides	1,150,673	1,206,483
Spiroplasma ORFs (> 300 nt)	1,019	995
with matches to:		
S. citri genome	957 (93.9%)	916 (92.1%)
S. kunkelli genome	928 (91.1%)	891 (89.5%)
S. melliferum genome	895 (87.8%)	881 (88.5%)
S. poulsonii genome	936 (91.8%)	907 (91.2%)

665

666 **Figure 1. Diverse *Myrmica* ants harbor *Spiroplasma***

667 Maximum-likelihood phylogram of mitochondrially-encoded cytochrome oxidase subunit

668 I (*COI*) nucleotide sequences of ants in the genus *Myrmica*. Gold type indicates taxa

669 screened for *Spiroplasma*. Species groups are designated by alternately shaded boxes
670 and labeled at the right side of the phylogram. Bold type marks a species group
671 containing one or more screened taxa. Branches are labeled with SH-like approximate
672 likelihood ratio test scores greater than 0.75.

673 **Figure 2. *Spiroplasma* in *Myrmica* belong to the citri clade**

674 (*A*) Maximum-likelihood phylogram of cell division protein *ftsZ* nucleotide sequences
675 from *Spiroplasma* bacteria. Gold type indicates a taxon amplified from *Myrmica* ants in
676 this study. Branches are labeled with SH-like approximate likelihood ratio test scores of
677 0.65 or higher. (*B*) Tanglegram depicting the pattern of cophylogeny between host and
678 symbiont gene trees. Horizontal black lines between trees connect host taxa to
679 symbiont taxa. Codivergence is not statistically significant as tested by ParaFit (p = .07).

680 **Figure 3. Genome coverage and *Spiroplasma* gene assignment of *Myrmica***
681 **symbionts**

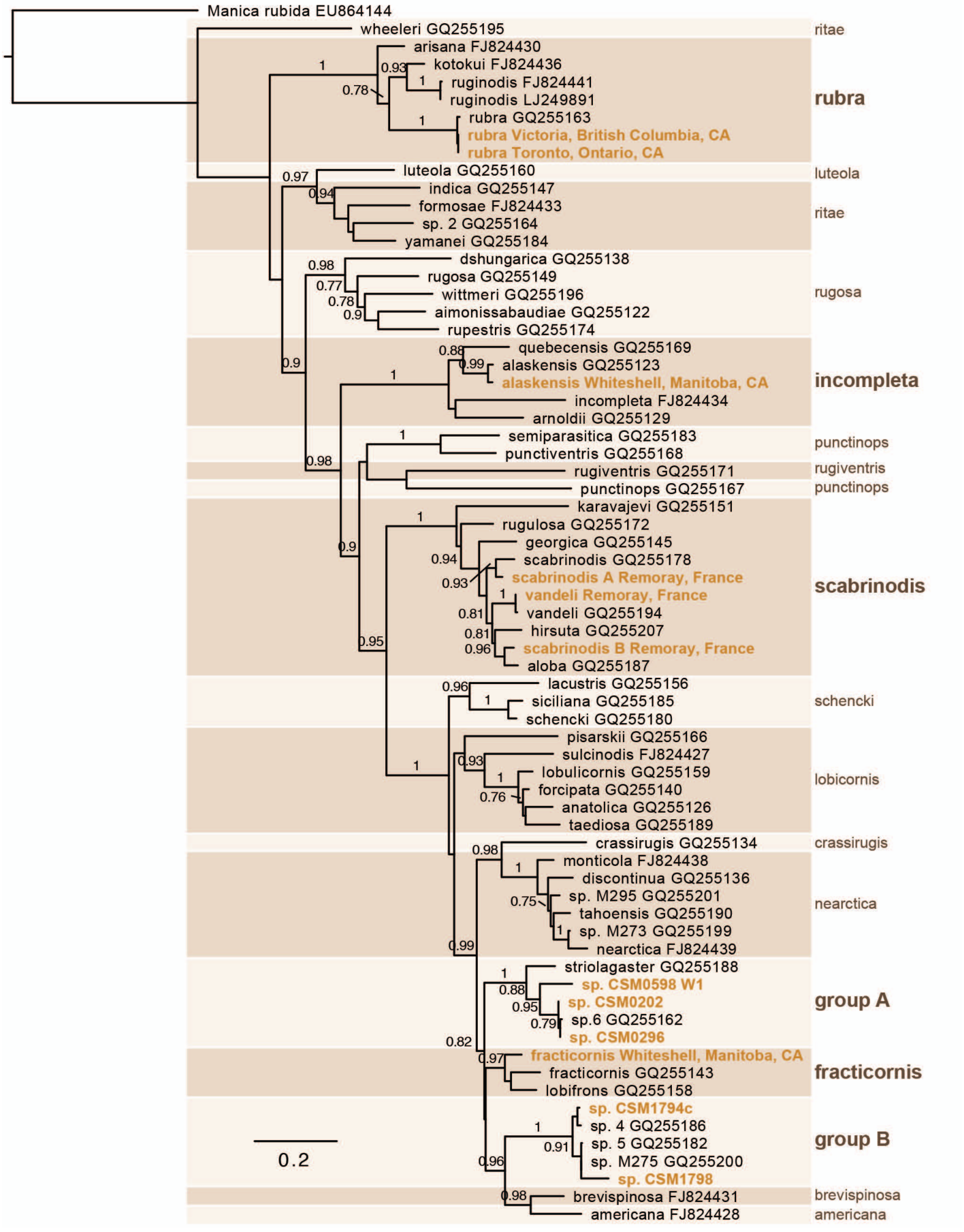
682 Read coverage distribution graphed in a 0-50 range and 0-300 range for (*A*) *M.*
683 *scabrinodis* and (*B*) *M. vandeli Spiroplasma* symbionts. Read coverage of the RIP-
684 encoding contig for each symbiont is indicated with an arrow. Pie charts summarize
685 gene assignments within each symbiont's genome. The majority fraction of *Spiroplasma*
686 genes in (*C*) *M. scabrinodis* and (*D*) *M. vandeli Spiroplasma* symbionts match best by
687 blastp to the honeybee pathogen, *Spiroplasma melliferum*, and overall, most genes
688 match to members of the citri clade: *Spiroplasma citri*, *S. kunkelii*, *S. melliferum*, and *S.*
689 *poulsonii*. In the legend, right, each taxon label is accompanied by the number of top
690 blastp matches from *M. scabrinodis* and *M. vandeli* symbionts, respectively.

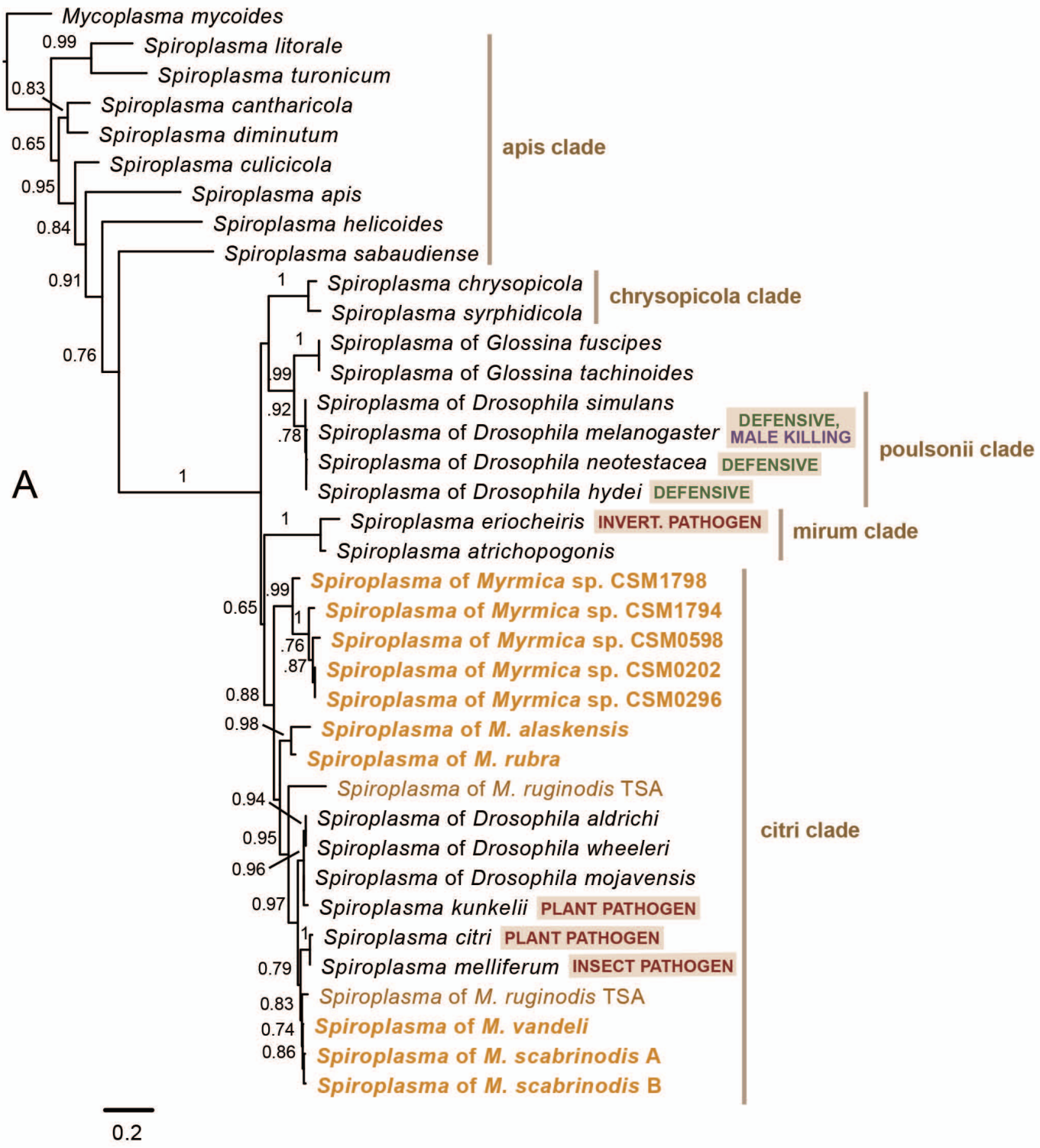
Figure 4. A divergent ECF transporter in the genomes of *Myrmica Spiroplasma* symbionts

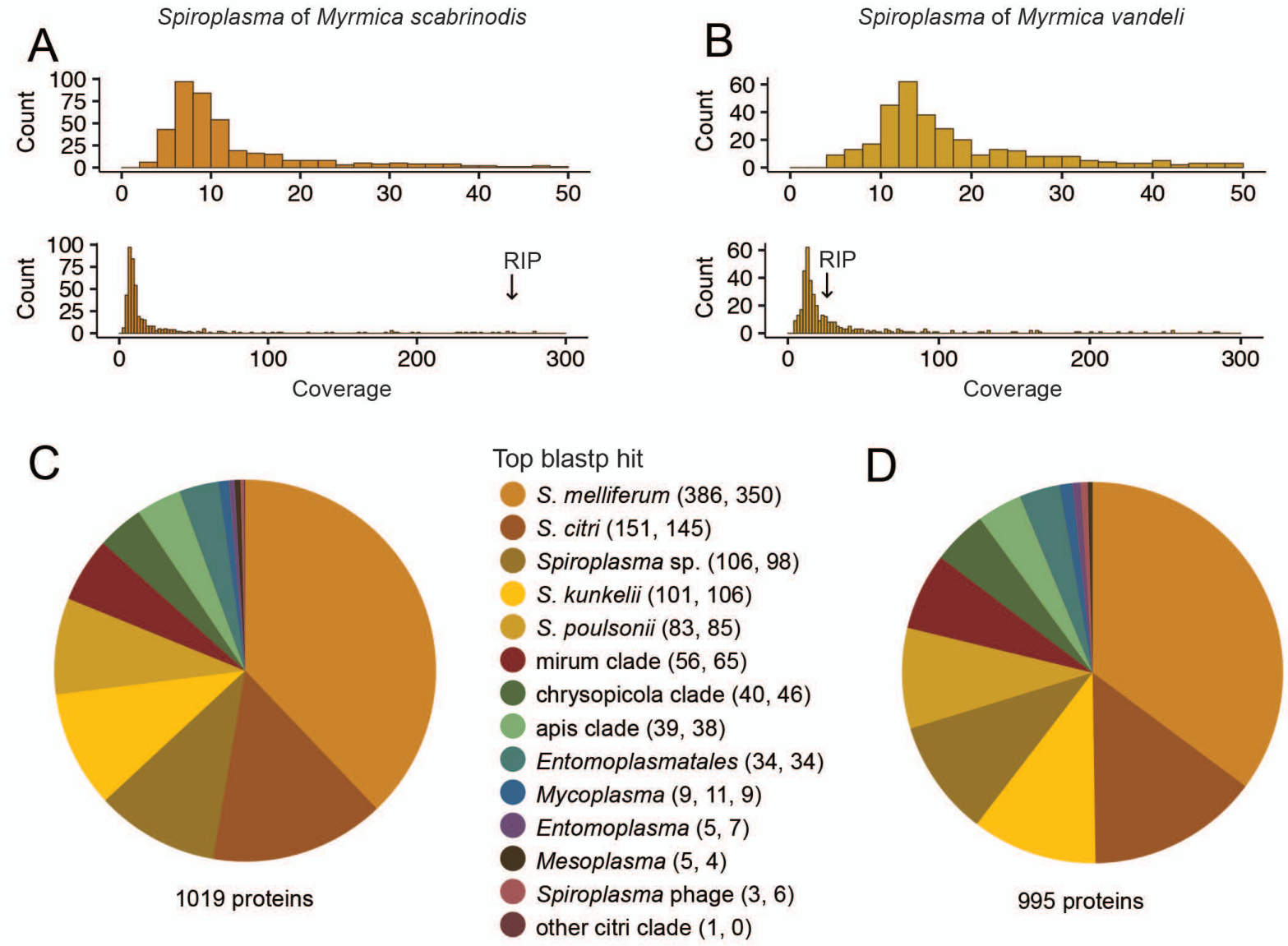
Maximum-likelihood phylogram of energy-coupling factor (ECF) transporter substrate component amino acid sequences. Tips are labeled with taxonomic identifiers and clades with protein family information, if available. Gold type indicates the novel ECF transporter of *Myrmica*-associated *Spiroplasma* symbionts, and red type indicates other ECF transporters identified in the genomes of these symbionts. Branches are labeled with SH-like approximate likelihood ratio test scores of 0.75 or higher.

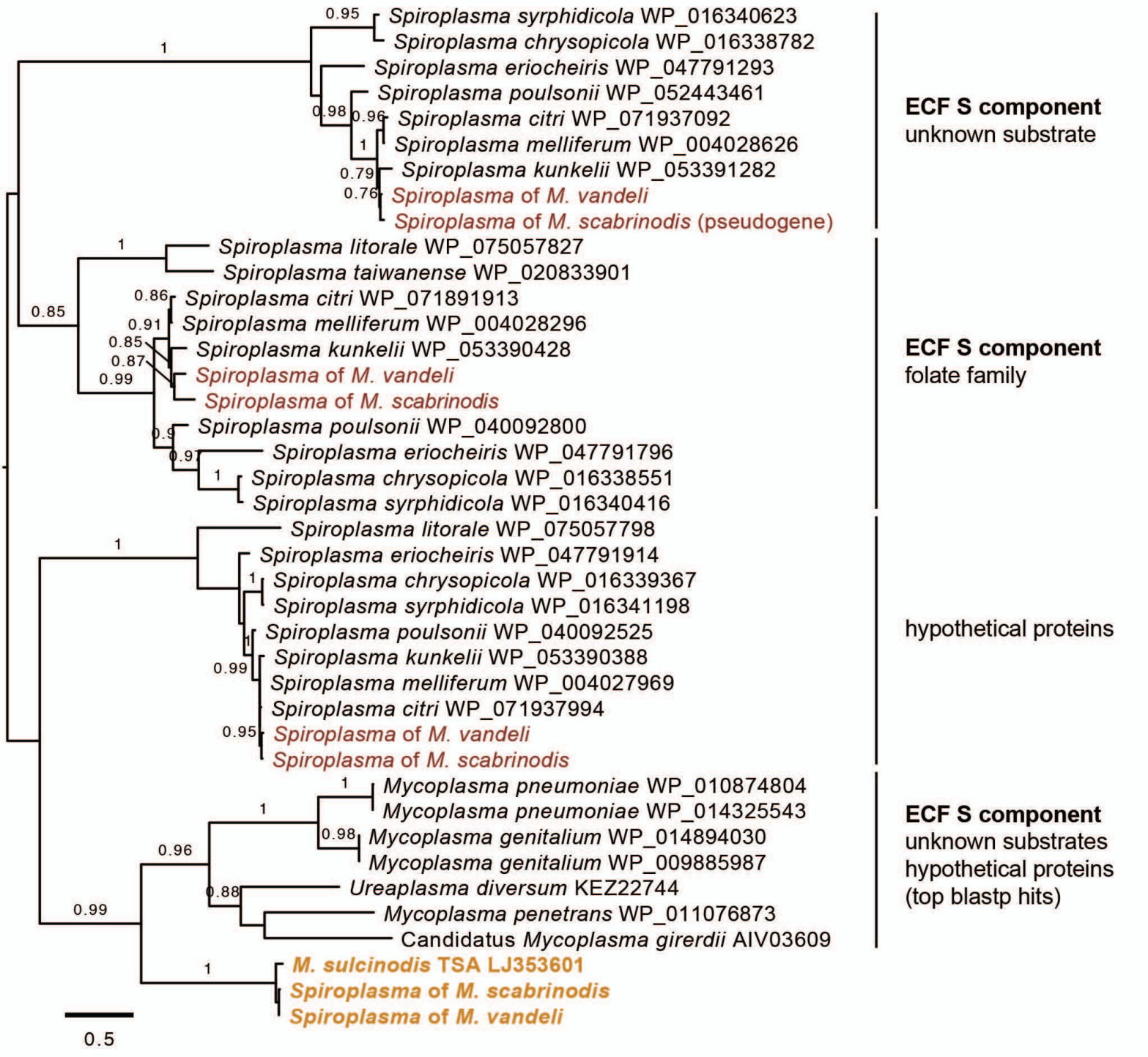
Figure 5. Diversity of ribosome-inactivating proteins in *Myrmica Spiroplasma*

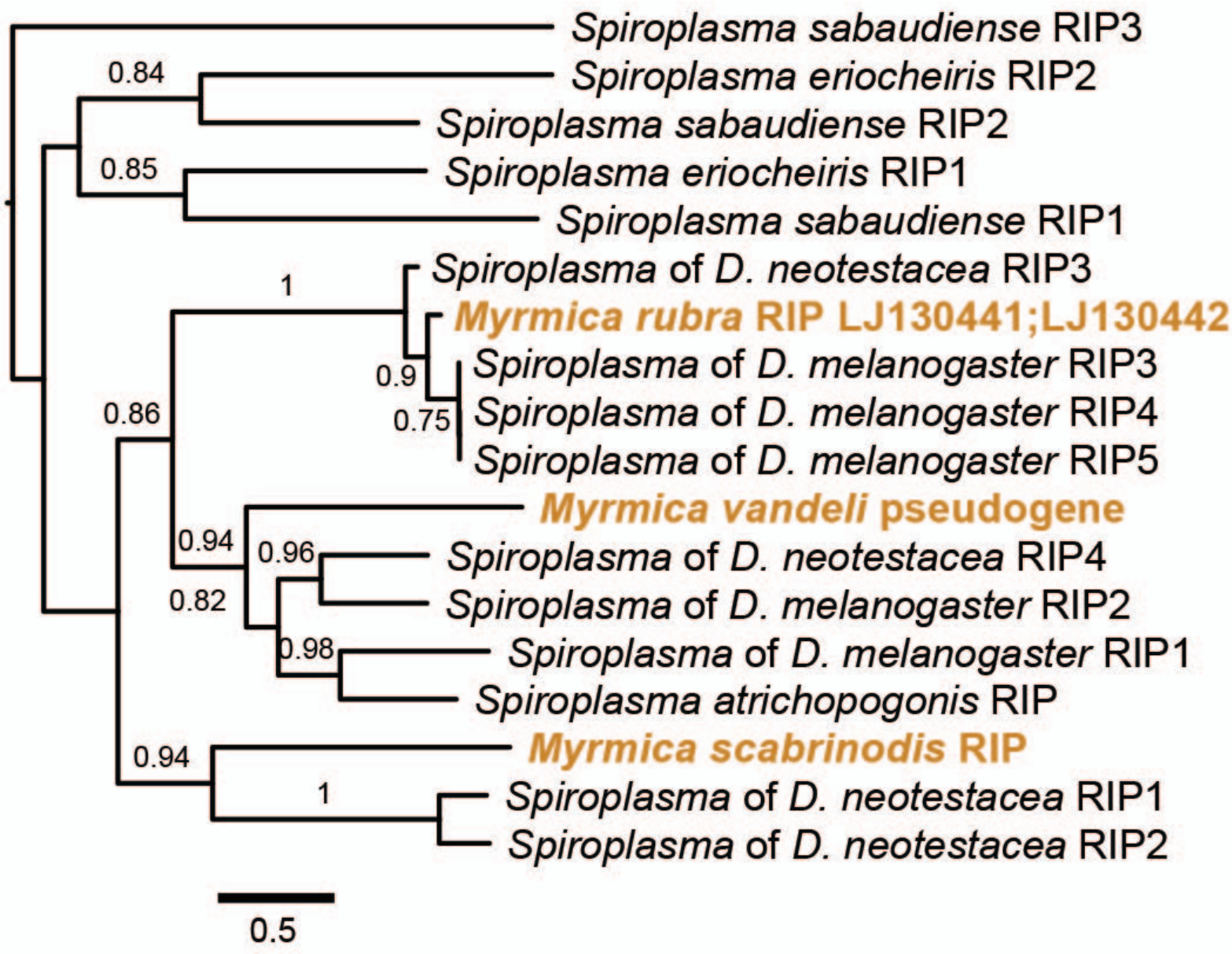
Maximum-likelihood phylogram of *Spiroplasma*-encoded ribosome inactivating protein (RIP) amino acid sequences. Tips are labeled with *Spiroplasma* species names or references to the host species harboring a *Spiroplasma* symbiont. Gold type indicates a RIP sequence identified from a *Myrmica*-associated *Spiroplasma* symbiont.


Manica rubida EU864144
wheeleri GQ255195
ritae
arisana FJ824430
kotokui FJ824436
ruginodis FJ824441
ruginodis LJ249891
rubra GQ255163
rubra Victoria, British Columbia, CA
rubra Toronto, Ontario, CA
rubra
luteola GQ255160
luteola
indica GQ255147
formosae FJ824433
sp. 2 GQ255164
yamanei GQ255184
ritae
dshungarica GQ255138
rugosa GQ255149
wittmeri GQ255196
aimonissabaudiae GQ255122
rupestris GQ255174
rugosa
quebecensis GQ255169
alaskensis GQ255123
alaskensis Whiteshell, Manitoba, CA
incompleta FJ824434
arnoldii GQ255129
incompleta
semiparasitica GQ255183
punctiventris GQ255168
punctinops
rugiventris GQ255171
rugiventris
punctinops GQ255167
punctinops
karavajevi GQ255151
rugulosa GQ255172
georgica GQ255145
scabrinodis GQ255178
scabrinodis A Remoray, France
vandeli Remoray, France
vandeli GQ255194
hirsuta GQ255207
scabrinodis B Remoray, France
aloba GQ255187
scabrinodis
lacustris GQ255156
siciliana GQ255185
schencki GQ255180
schencki
pisarskii GQ255166
sulcinodis FJ824427
lobulicornis GQ255159
forcipata GQ255140
anatolica GQ255126
taediosa GQ255189
lobicornis
crassirugis GQ255134
crassirugis
monticola FJ824438
discontinua GQ255136
sp. M295 GQ255201
tahoensis GQ255190
sp. M273 GQ255199
nearctica FJ824439
nearctica
striolagaster GQ255188
sp. CSM0598 W1
sp. CSM0202
sp.6 GQ255162
sp. CSM0296
group A
fracticornis Whiteshell, Manitoba, CA
fracticornis GQ255143
lobifrons GQ255158
fracticornis
sp. CSM1794c
sp. 4 GQ255186
sp. 5 GQ255182
sp. M275 GQ255200
sp. CSM1798
group B
brevispinosa FJ824431
brevispinosa
americana FJ824428
americana
0.2


A
Mycoplasma mycoides
Spiroplasma litorale
Spiroplasma turonicum
Spiroplasma cantharicola
Spiroplasma diminutum
Spiroplasma culicicola
Spiroplasma apis
Spiroplasma helicoides
Spiroplasma sabaudiense
apis clade
Spiroplasma chrysopicola
Spiroplasma syrphidicola
chrysopicola clade
Spiroplasma of Glossina fuscipes
Spiroplasma of Glossina tachinoides
Spiroplasma of Drosophila simulans
Spiroplasma of Drosophila melanogaster
DEFENSIVE, MALE KILLING
Spiroplasma of Drosophila neotestacea
DEFENSIVE
Spiroplasma of Drosophila hydei
DEFENSIVE
poulsonii clade
Spiroplasma eriocheiris
INVERT. PATHOGEN
Spiroplasma atrichopogonis
mirum clade
Spiroplasma of Myrmica sp. CSM1798
Spiroplasma of Myrmica sp. CSM1794
Spiroplasma of Myrmica sp. CSM0598
Spiroplasma of Myrmica sp. CSM0202
Spiroplasma of Myrmica sp. CSM0296
Spiroplasma of M. alaskensis
Spiroplasma of M. rubra
Spiroplasma of M. ruginodis TSA
Spiroplasma of Drosophila aldrichi
Spiroplasma of Drosophila wheeleri
Spiroplasma of Drosophila mojavensis
Spiroplasma kunkelii
PLANT PATHOGEN
Spiroplasma citri
PLANT PATHOGEN
Spiroplasma melliferum
INSECT PATHOGEN
Spiroplasma of M. ruginodis TSA
Spiroplasma of M. vandeli
Spiroplasma of M. scabrinodis A
Spiroplasma of M. scabrinodis B
citri clade
0.2
B
Myrmica COI
Spiroplasma ftsZ
CSM0202
CSM0296
CSM0598 W1
CSM1794c
CSM1798
vandeli
scabrinodis A
scabrinodis B
alaskensis
ruginodis
rubra
CSM0202
CSM0296
CSM0598 W1
CSM1794c
CSM1798
vandeli
ruginodis
scabrinodis A
scabrinodis B
ruginodis
alaskensis
rubra


A
Spiroplasma of Myrmica scabrinodis
Count
100
75
50
25
0
0
10
20
30
40
50
Count
100
75
50
25
0
RIP
0
100
200
300
Coverage
B
Spiroplasma of Myrmica vandeli
Count
60
40
20
0
0
10
20
30
40
50
Count
60
40
20
0
RIP
0
100
200
300
Coverage
C
1019 proteins
Top blastp hit
S. melliferum (386, 350)
S. citri (151, 145)
Spiroplasma sp. (106, 98)
S. kunkelii (101, 106)
S. poulsonii (83, 85)
mirum clade (56, 65)
chrysopicola clade (40, 46)
apis clade (39, 38)
Entomoplasmatales (34, 34)
Mycoplasma (9, 11, 9)
Entomoplasma (5, 7)
Mesoplasma (5, 4)
Spiroplasma phage (3, 6)
other citri clade (1, 0)
D
995 proteins


Spiroplasma syrphidicola WP_016340623
Spiroplasma chrysopicola WP_016338782
Spiroplasma eriocheiris WP_047791293
Spiroplasma poulsonii WP_052443461
Spiroplasma citri WP_071937092
Spiroplasma melliferum WP_004028626
Spiroplasma kunkelii WP_053391282
Spiroplasma of M. vandeli
Spiroplasma of M. scabrinodis (pseudogene)
ECF S component
unknown substrate
Spiroplasma litorale WP_075057827
Spiroplasma taiwanense WP_020833901
Spiroplasma citri WP_071891913
Spiroplasma melliferum WP_004028296
Spiroplasma kunkelii WP_053390428
Spiroplasma of M. vandeli
Spiroplasma of M. scabrinodis
Spiroplasma poulsonii WP_040092800
Spiroplasma eriocheiris WP_047791796
Spiroplasma chrysopicola WP_016338551
Spiroplasma syrphidicola WP_016340416
ECF S component
folate family
Spiroplasma litorale WP_075057798
Spiroplasma eriocheiris WP_047791914
Spiroplasma chrysopicola WP_016339367
Spiroplasma syrphidicola WP_016341198
Spiroplasma poulsonii WP_040092525
Spiroplasma kunkelii WP_053390388
Spiroplasma melliferum WP_004027969
Spiroplasma citri WP_071937994
Spiroplasma of M. vandeli
Spiroplasma of M. scabrinodis
hypothetical proteins
Mycoplasma pneumoniae WP_010874804
Mycoplasma pneumoniae WP_014325543
Mycoplasma genitalium WP_014894030
Mycoplasma genitalium WP_009885987
Ureaplasma diversum KEZ22744
Mycoplasma penetrans WP_011076873
Candidatus Mycoplasma girerdii AIV03609
ECF S component
unknown substrates
hypothetical proteins
(top blastp hits)
M. sulcinodis TSA LJ353601
Spiroplasma of M. scabrinodis
Spiroplasma of M. vandeli
0.5


Spiroplasma sabaudiense RIP3
0.84
Spiroplasma eriocheiris RIP2
Spiroplasma sabaudiense RIP2
0.85
Spiroplasma eriocheiris RIP1
Spiroplasma sabaudiense RIP1
Spiroplasma of D. neotestacea RIP3
1
Myrmica rubra RIP LJ130441;LJ130442
0.9
Spiroplasma of D. melanogaster RIP3
0.86
0.75
Spiroplasma of D. melanogaster RIP4
Spiroplasma of D. melanogaster RIP5
Myrmica vandeli pseudogene
0.94
0.96
Spiroplasma of D. neotestacea RIP4
0.82
Spiroplasma of D. melanogaster RIP2
0.98
Spiroplasma of D. melanogaster RIP1
Spiroplasma atrichopogonis RIP
Myrmica scabrinodis RIP
0.94
1
Spiroplasma of D. neotestacea RIP1
Spiroplasma of D. neotestacea RIP2
0.5

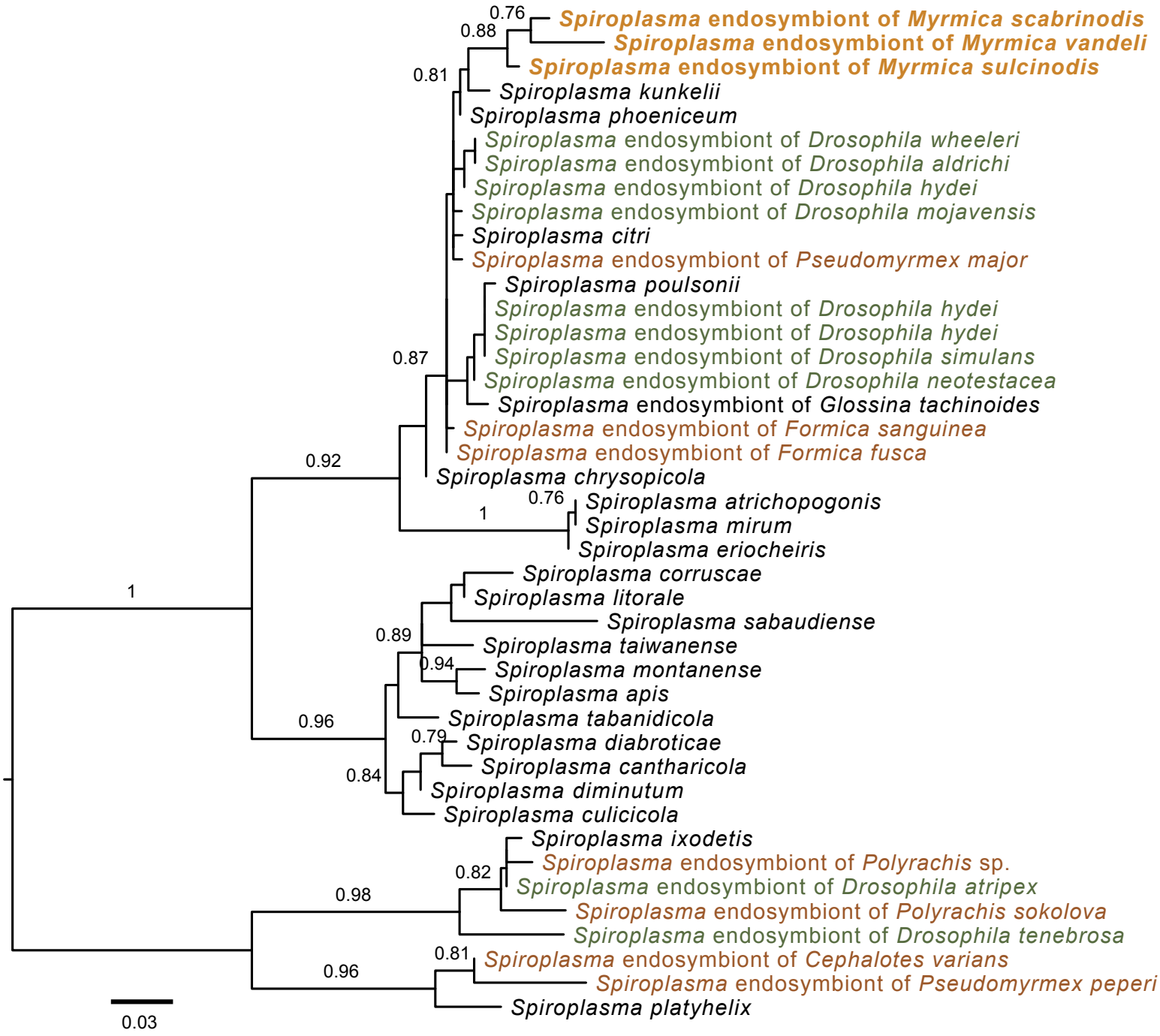


Fig S1. Diversity of *Spiroplasma* in *Myrmica* and other ants
A maximum-likelihood phylogram of *Spiroplasma* 16S ribosomal DNA constructed from an alignment of approximately 400 nucleotide positions. *Spiroplasma* from *Myrmica* are shown in gold, those detected from other ant genera are shown in brown, and those of *Drosophila* are shown in dark green. Branches are labeled with SH-like approximate likelihood ratio test scores of 0.75 or greater.