Multiple origins of obligate nematode and insect symbionts by a clade of bacteria closely related to plant pathogens

Vincent G. Martinson, Ryan M. R. Gawryluk, Brent E. Gowen, Caitlin I. Curtis, John Jaenike, & Steve J. Perlman

December 2020

The final publication is available at:

https://doi.org/10.1073/pnas.2000860117

Accepted Manuscript:

Martinson VG, Gawryluk RMR, Gowen BE, Curtis CI, Jaenike J, Perlman SJ. 2020. Multiple origins of obligate nematode and insect symbionts by a clade of bacteria closely related to plant pathogens. *Proceedings of the National Academy of Sciences, USA*. *117, 31979-31986*. *doi/10.1073/pnas.2000860117*

Main Manuscript for

Multiple origins of obligate nematode and insect symbionts by members of a newly characterized bacterial clade

Authors.
Vincent G. Martinson[1,2], Ryan M. R. Gawryluk[3], Brent E. Gowen[3], Caitlin I. Curtis[3], John Jaenike[1], Steve J. Perlman[3]
[1] Department of Biology, University of Rochester, Rochester, NY, USA, 14627
[2] Department of Biology, University of New Mexico, Albuquerque, NM, USA, 87131
[3] Department of Biology, University of Victoria, Victoria, BC, Canada, V8W 3N5

Corresponding author.
Vincent G. Martinson, Department of Biology, MSC03 2020, 1 University of New Mexico, Albuquerque, NM 87131-0001; Tel: 1-406-697-4660; Email: vmartinson@unm.edu

ORCID.
Vincent G. Martinson (ORCID: 0000-0001-5824-3548), Ryan M. R. Gawryluk (0000-0003-1251-882X), Steve J. Perlman (0000-0002-9020-5506)

Classification.
Biological Sciences

Keywords.
Howardula, symbiosis, *Drosophila*, genome reduction, *Sodalis*

Short Title.
Widespread Symbiont of Nematodes and Insects

Author Contributions.
VGM, SJP, RMRG, and JJ designed and performed experiments and sequencing. BEG performed electron microscopy. VGM performed analyses. CIC performed experiments. VGM and SJP wrote the paper with input from all the authors.

Competing Interests Statement.
The authors declare no competing interest.

This PDF file includes:

Main Text	Supporting Information Appendix
Figures 1 to 4	Figures S1 to S7
Tables 1 to 2	Tables S1 to S2

Abstract.
Obligate symbioses involving intracellular bacteria have transformed eukaryotic life, from providing aerobic respiration and photosynthesis, to enabling colonization of previously inaccessible niches, such as feeding on xylem and phloem, and surviving in deep-sea hydrothermal vents. A major challenge in the study of obligate symbioses is to understand how they arise. Because the best studied obligate symbioses are ancient, it is especially challenging to identify early or intermediate stages. Here we report the discovery of a nascent obligate symbiosis in *Howardula aoronymphium*, a well-studied nematode parasite of *Drosophila* flies. We have found that *H. aoronymphium* and its sister species harbor a maternally inherited intracellular bacterial symbiont. We never find the symbiont in nematode-free flies and virtually all nematodes in the field and the lab are infected. Treating nematodes with antibiotics causes a severe reduction in fly infection success. The association is recent, as more distantly related insect-parasitic tylenchid nematodes do not host these endosymbionts. We also report that the *Howardula* nematode symbiont is a member of a widespread monophyletic group of invertebrate host-associated microbes that has independently given rise to at least four obligate symbioses, one in nematodes and three in insects, and that is sister to *Pectobacterium*, a lineage of plant pathogenic bacteria. Comparative genomic analysis of this group, which we name *Candidatus* Symbiopectobacterium, shows signatures of genome erosion characteristic of early stages of symbiosis, with the *Howardula* symbiont's genome containing over a thousand predicted pseudogenes, comprising a third of its genome.

Significance Statement.
Obligate symbioses are intimate associations between species in which neither partner can live without the other. It is challenging to study how obligate symbioses arise because they are often ancient and it is difficult to uncover early or intermediate stages. We have discovered a nascent obligate symbiosis involving *Howardula aoronymphium*, a well-studied nematode parasite of *Drosophila* flies, and a bacterium related to *Pectobacterium*, a lineage of plant pathogens. Moreover, this nematode symbiont is a member of a widespread group of invertebrate host-associated microbes that has independently given rise to at least four obligate symbioses in nematodes and insects, making it an exciting model to study transitions to obligate symbiosis.

Introduction.

Intimate symbioses involving intracellular bacteria have transformed eukaryotic life (1, 2), with mitochondria and chloroplasts as canonical examples. More recent, yet still ancient, acquisitions of obligate bacterial intracellular endosymbionts have enabled colonization and radiation by animals into previously inaccessible niches, such as feeding on plant sap and animal blood (3), and surviving in deep-sea hydrothermal vents (4). Among the most difficult questions to resolve in the study of obligate symbiosis are how do obligate symbioses evolve, and where do obligate symbionts come from? This is particularly challenging because most of the obligate symbioses that have been studied are ancient, making it is extremely difficult to identify early or intermediate stages.

One of the most common ways to acquire an obligate symbiont is via symbiont replacement (5). As a result of a lifestyle shaped by genetic drift, vertically transmitted obligate symbionts follow a syndrome of accumulation of deleterious mutations, leading to genome degradation and reduction (6). A common pattern is that they are replaced by other less-broken symbionts that may then renew the cycle of genomic degradation (7). Here the symbiont, which is often descended from common facultative symbionts or parasites (8, 9), is fitted into an established and well-functioning symbiosis (i.e. with a 'symbiont-experienced' host). For example, the symbiont *Sodalis* has independently given rise to numerous obligate nutritional symbioses in blood-feeding flies and lice, sap-feeding mealybugs, spittlebugs, hoppers, and grain-feeding weevils (9).

Less studied are young obligate symbioses in host lineages that did not already house obligate symbionts (i.e. 'symbiont-naïve' hosts') (10). Some of the best-known examples originate through host manipulation by the symbiont via addiction or reproductive control. Addiction or dependence may be a common route for obligate symbiosis (11), and one of the most famous examples occurred in the laboratory, on the timescale of years, where strains of *Amoeba* evolved to become entirely dependent on intracellular symbionts (12). Many maternally-inherited symbionts of terrestrial arthropods induce parthenogenetic (i.e. all female) reproduction in their hosts (13); accumulation of deleterious mutations in genes required for sexual reproduction will result in hosts that are unable to reproduce if cured of their symbiont (14). However, despite advances in microbial surveys, there are still few examples of young obligate symbioses that result in novel host functions. One intriguing example involves spheroid bodies, nitrogen-fixing organelles found in rhopalodiacean diatoms, that originated from a single acquisition of a cyanobacterial symbiont as recently as ~12 Mya (15, 16).

Here we report the discovery of a nascent obligate symbiosis in *Howardula aoronymphium*, a well-studied nematode parasite of *Drosophila* (17), most recently in the context of a defensive symbiosis. A common host species, *D. neotestacea*, harbors a strain of the facultative inherited symbiont *Spiroplasma* that protects it against nematode-induced sterility (18). The protection provided by *Spiroplasm*a is so strong that symbiont-infected flies are spreading across N. America and replacing their uninfected counterparts (19). Surprisingly, we have found that *H. aoronymphium* itself harbors an intracellular bacterial symbiont that is related to *Pectobacterium*, a well-studied group of plant pathogens, often vectored by insects. We also report that the nematode symbiont, which we name *Candidatus* Symbiopectobacterium (and hereafter *Symbiopectobacterium*), is a member of a widespread lineage of invertebrate symbionts that has independently given rise to at least four obligate symbioses, one in nematodes and three in insects, representing an exciting new model for the study of obligate symbiosis.

Results.

Virtually all H. aoronymphium in the field and lab host Symbiopectobacterium

We surveyed wild-caught *Drosophila* spp. across North America and Europe, and over multiple years, for the presence of *H. aoronymphium* and *Symbiopectobacterium*. Using specific primers for *Symbiopectobacterium*, virtually all *Howardula*-infected flies were also positive for the symbiont (74 of 79, Table 1; the 5 individuals that tested positive for *Howardula* but negative for *Symbiopectobacterium* may have contained dead or dying nematodes). The symbiont was never amplified from *Drosophila* not infected with *Howardula*. Our lab strain of *Howardula aoronymphium*, maintained in the lab for over ten years, is also always infected with *Symbiopectobacterium*, and in a controlled laboratory infection of *D.*

neotestacea we confirmed the perfect association of *H. aoronymphium* and *Symbiopectobacterium*, i.e. only nematode-infected flies are positive for the symbiont (Figure 1a).

Symbiopectobacterium is an intracellular inherited symbiont of *H. aoronymphium*

Within an adult fly host, a single adult parasitic female nematode, commonly referred to as a motherworm at this stage, can produce around 500 larvae that fill her body until they rupture her hypodermis, disseminate throughout the host fly hemocoel, and finally exit the fly through the intestinal or genital tract (20). Using TEM and FISH microscopy, we observed pervasive bacterial infection in nematode motherworms, juveniles developing within motherworms, and juveniles that have been released from motherworms (Figure 1b-g; SI Appendix, Figure S1). In motherworms, this bacterium is situated intracellularly within hypertrophied hypodermal cells, below the outer layer of microvilli, that distinguish the parasitic female (SI Appendix, Figure S1). Within juveniles, bacteria are present in areas consistent with hypodermal cells. 16S rRNA amplicon sequencing of *H. aoronymphium* was dominated by *Symbiopectobacterium*, and there were no other potential symbiont sequences (SI Appendix, Figure S2).

Antibiotic treatment decreases *Howardula* parasitism success

We attempted to clear *Symbiopectobacterium* from *H. aoronymphium* by rearing nematodes on their host *D. neotestacea* in media with the antibiotics ampicillin or rifampin. Adult flies were then screened for nematodes and symbionts. The proportion of parasitized *D. neotestacea* was significantly lower with ampicillin (0.085±0.03) and rifampin (no infection) compared to the control *H. aoronymphium* exposure (0.25±0.9) (Control-Amp $\chi^2(1, N = 687) = 23.48$, $p < 0.0001$; Control-Rif $\chi^2(1, N = 469) = 45.74$, $p < 0.0001$) (Figure 2a). Further, in a subset of flies that we dissected, *Symbiopectobacterium* was found in all flies that contained motherworms (25/25), while visually non-parasitized flies were almost all negative (41/45) (Figure 2b). Some visually non-parasitized flies were PCR positive for *H. aoronymphium* but lacked *Symbiopectobacterium* (10/14), possibly indicating unsuccessful parasitism. Thus, we are at present unable to generate symbiont-free nematodes.

Symbiopectobacterium is a common associate of a clade of *Drosophila*-parasitic nematodes

In order to determine if other insect-parasitic nematodes in the suborder Hexatylina are ancestrally associated with *Symbiopectobacterium*, we screened nematode species with a range of primers designed to amplify *Symbiopectobacterium*. We detected *Symbiopectobacterium* in two close relatives of *H. aoronymphium* that also parasitize *Drosophila* (21) - *H. neocosmis* and an unnamed Japanese *Howardula* sp. (12/12 and 1/1 respectively) (Table 2). The *Symbiopectobacterium* gyrA gene sequences amplified from these nematodes formed a strongly supported (100% bootstrap support) monophyletic clade, sharing >99% nucleotide sequence identity (SI Appendix, Figure S3abc). In contrast, *Symbiopectobacterium* was not found in two species of *Fergusobia* (0/6 individuals), the lineage that is sister to *Drosophila*-parasitic *Howardula* (22), in *Parasitylenchus nearcticus*, a more distantly related tylenchid parasite of *Drosophila* flies (21) (0/3 individuals), or in three unnamed nematodes that infect sphaerocerid flies (0/13, 0/3, and 0/9), using a number of primers (gyrA, groEL, purK spacer, and 16S rRNA genes) designed to target *Symbiopectobacterium*, as well as universal 16S rRNA primers (Table 2, SI Appendix, Table S1, Figure S2, Figure S3c).

Symbiopectobacterium is an invertebrate-associated lineage allied with the plant-associated genus *Pectobacterium*

A large number of gene sequences that had >97% sequence identity to the *Howardula* symbiont 16S rRNA gene were recovered from GenBank. Phylogenetic reconstruction revealed a monophyletic clade closely related to the genera *Pectobacterium*, *Dickeya*, and *Brenneria* (SI Appendix, Figure S4). This clade included several known intracellular symbionts, including in *Cimex* bed bugs (23), *Euscelidius* leafhoppers (24), and *Chilacis* bulrush bugs (25); the remaining sequences were almost completely observed in association with invertebrates (SI Appendix, Table S2). *Symbiopectobacterium* is most commonly associated with insects from the order Hemiptera; however, these insects are taxonomically diverse, sharing a most recent common ancestor nearly 300 Mya, and ecologically diverse with disparate life history traits (e.g. blood-feeding, phloem-feeding, seed-feeding) (26).

Signatures of independent genome erosion across the *Symbiopectobacterium* clade

We sequenced the genomes of *Symbiopectobacterium* in nematodes and bulrush bugs, and compared them with publicly available sequences of related symbionts in mealybugs (7), leafhoppers (27), bed bugs (28), and parasitoid wasps (29), as well as free-living *Pectobacterium* (30). Regardless of the three-fold size difference between the genomes of *Symbiopectobacterium* in nematodes (4.5 Mb) and in bulrush bugs (1.5 Mb), both genomes harbored >90% of 203 single copy ortholog (SiCO) gammaproteobacterial genes (31), indicating we captured nearly full chromosomes. Similarly, the *Symbiopectobacterium* genomes recovered from the mealybug and parasitoid wasp genome projects were nearly complete, containing 98% and 86% of the SiCO genes. Genomic fragments of the bed bug and leafhopper symbionts were incomplete and lacked most of the SiCO genes, but could be positively identified as members of the *Symbiopectobacterium* clade and were included in the genome tree (Figure 3, 4a).

Phylogenetic analysis of 203 single copy genes resulted in an overall branching pattern similar to previous publications of Gammaproteobacteria (27, 32) and highly supported the *Symbiopectobacterium* lineage within the plant-pathogenic 'soft rot Enterobacteriaceae' as the sister clade to the genera *Pectobacterium* and *Brenneria* (Figure 3; SI Appendix, Figure S5). Regardless of sequence divergence, species within the *Symbiopectobacterium* group shared large areas of gene order synteny with the closely related *Pectobacterium carotovorum* genome. Alignment within syntenic regions revealed that, although certain genomic rearrangements are shared, each genome has independent insertions, deletions and mutations that have resulted in differential pseudogene formation (Figure 4bc).

Compared to their closest relative, *Pectobacterium carotovorum*, the *Symbiopectobacterium* clade members had signatures of genome erosion often associated with the early stages of symbiosis, including more pseudogenes, fewer tRNAs, shorter average CDS size, lower coding density, and decreased genome size (Figure 4a). Despite their similarities, there is large variation among members of *Symbiopectobacterium* in these same genomic traits, suggesting that they are at different stages in the process of genomic reduction, which might be connected with the age of their association with their host invertebrate. Using sequence divergence measurements (33), we estimate that *Symbiopectobacterium* split from *Pectobacterium* 400-500 ky and the symbiosis events among *Symbiopectobacterium* species occurred independently <100 ky (SI Appendix, Figure S6).

Along with genomic erosion, the metabolic potential of *Symbiopectobacterium* members has changed greatly. Pseudogene formation or deletion has interrupted many of the amino acid and vitamin synthesis pathways. While *Symbiopectobacterium* in nematodes has maintained many genes associated with synthesis of amino acids and vitamins, chemotaxis, motility, and secretion systems, the bulrush bug symbiont has lost the majority of functions, except for basic DNA replication and repair, and biosynthesis of lysine and several vitamins.

Proposal of "*Candidatus* Symbiopectobacterium"

We propose the genus name "*Candidatus* Symbiopectobacterium" for the lineage of Enterobacteriaceae that forms a monophyletic clade sister to *Pectobacterium* (SI Appendix, Figure S4), and whose members are commonly found in association with diverse invertebrates, including intracellular symbionts that are vertically transmitted within their host (i.e. symbionts of *Howardula aoronymphium, Euscelidius variegatus*, *Cimex lectularius*, *Chilacis typhae*, and *Pseudococcus longispinus*). While the 16S rRNA gene in *Symbiopectobacterium* lacks sufficient genetic diversity to differentiate it from sister genera at the historically used 95% similarity level, this lineage is clearly divergent in the phylogeny produced with the genome-wide single-copy ortholog genes, which is mirrored in their ecological associations. The symbionts of the bulrush bug and sycamore seed bug, *Chilacis typhae* and *Belonochilus numenius*, respectively, have previously been named *Candidatus* Rohrkolberia, which refers to the German word for bulrush (25, 34). It was recently pointed out that this name does not conform to the current naming standards laid out in the International Code of Nomenclature of Prokaryotes (35, 36), because the German word for bulrush is used instead of the Latin (Rule and Recommendation 6). Our proposal of *Candidatus* Symbiopectobacterium would supersede the genus *Ca*. Rohrkolberia; this name indicates both the widespread symbiotic nature of this lineage and its relatedness to *Pectobacterium*.

Discussion

We have discovered a nascent obligate symbiosis involving *Howardula aoronymphium*, a nematode parasite of *Drosophila*, and a bacterium from a cryptic but widespread lineage of endosymbionts, allied

with *Pectobacterium*, and which we here name *Candidatus* Symbiopectobacterium. We present four different lines of evidence supporting obligate symbiosis. First, there is almost perfect concordance between *Symbiopectobacterium* and *H. aoronymphium* presence. We never detect symbionts in nematode-free flies, and virtually all wild-caught flies that test positive for *H. aoronymphium* DNA are also positive for *Symbiopectobacterium*. The very rare mismatches could be due to false negative PCR amplification, perhaps due to low quality and/or low titer DNA, for example due to dead or dying nematodes within a fly. Second, microscopy revealed intracellular bacterial infection, including inside motherworms, juveniles developing inside motherworms, and shed juveniles. Third, closely related symbionts were also detected in the sister species *H. neocosmis* and a Japanese *Howardula* sp. Finally, treating nematodes with antibiotics caused a severe reduction in fly infection success, although we cannot rule out the possibility that the antibiotic directly affected the nematode.

Only a handful of obligate symbioses involving nematodes and bacteria has been described, and these include *Wolbachia* in some filarial nematodes (37), *Xiphinematobacter* and *Burkholderia* in some plant-parasitic dagger nematodes (38, 39), *Photorhabdus* and *Xenorhabdus* in entomopathogenic nematodes (40), and *Thiosymbion* sulfur-oxidizing ectosymbionts of marine stilbonematines (41). In contrast to many nematode-bacterium associations, this symbiosis appears to be very young. We detected *Symbiopectobacterium* only in the closely related *H. aoronymphium*, *H. neocosmis* and Japanese *Howardula* sp., which represents just a sliver of the insect-parasitic, and *Drosophila*-parasitic, diversity within the Hexatylina (21, 42, 43). We were unable to amplify *Symbiopectobacterium* DNA in the related nematodes *Fergusobia* sp., *Parasitylenchus nearcticus*, or three unnamed nematodes that infect sphaerocerid flies. Nor did we find any evidence of bacterial endosymbiont infection in electron microscopy studies of allied Hexatylina nematodes, including *Deladenus*, *Thripinema*, and *Contortylenchus* (44-46), and a very detailed and extensive study of *Howardula husseyi* (47) [note that the genus *Howardula* is not monophyletic] (SI Appendix, Figure S3bc). As there is very little available insect-parasitic nematode DNA sequence, it is difficult to provide an estimate for the age of the *Symbiopectobacterium*-*Howardula* association. *Fergusobia* nematodes, the sister group of the *Howardula* clade that hosts *Symbiopectobacterium*, are obligate mutualists of fergusoninid gall-making flies, and the age of this family of flies has been estimated to be not more than 42 My (48, 49), providing a conservative upper limit to the age of the symbiosis. Our sequence divergence measurements, however, suggest that the symbiosis is much more recent, with *Symbiopectobacterium* splitting from *Pectobacterium* about a half a million years ago.

Symbiopectobacterium is a surprisingly widespread lineage of symbionts, closely related to *Pectobacterium, Dickeya,* and *Brenneria*, often referred to as 'soft rot Enterobacteriaceae' (SRE). These SREs are common pathogens of plants that are often vectored by insects (50). The high sequence similarity shared between *Symbiopectobacterium* and *Pectobacterium* (e.g. 16S rRNA gene) may explain why this lineage has remained largely undetected despite the increase in amplicon surveys of invertebrates. *Symbiopectobacterium* symbionts are diverse and include at least four independently evolved obligate mutualistic symbioses – once in *Drosophila*-parasitic *Howardula*, and three times as a nutritional symbiont in sap-feeding hemipteran insects, including a young symbiont replacement in *Pseudococcus longispinus,* the long-tailed mealybug (7). They have also independently colonized the lygaeoid seed bugs *Chilacis typhae*, the bulrush bug, and *Belonochilus numenius*, the sycamore seed bug (25, 34). They are also common facultative symbionts of insects, including in *Cimex lectularius* bed bugs (23), *Dipetalogaster maximus* and possibly related kissing bugs (51, 52), and *Euscelidius variegatus* leafhoppers (24, 53). This latter symbiont, called BEV ('B'acterium of 'E'uscelidius 'v'ariegatus), was an early model in insect symbiosis (53) that is noteworthy because it exhibits both transovarial and horizontal transmission, and can be easily cultured on media outside the host. We also found some *Symbiopectobacterium* sequences in GenBank that are not associated with invertebrate hosts, but instead were identified in association with plants. These sequences may represent free-living *Symbiopectobacterium* strains that have repeatedly forged symbioses with invertebrate hosts; alternatively, new host-symbiont combinations may have become established via horizontal transmission of facultative symbionts. *Symbiopectobacterium* is the fourth example of a lineage of symbionts that is shared between insects and nematodes, joining *Wolbachia* (54), *Cardinium* (55), and *Burkholderia* (38, 56).

The evolution and distribution of *Symbiopectobacterium* is reminiscent of *Sodalis*, another widespread lineage of mostly facultative insect symbionts that has repeatedly given rise to obligate nutritional symbioses in sap and blood-feeding hemipterans and flies (9); acquisition of some *Sodalis* has been estimated to have occurred very recently, ~30 kya (33). Like *Sodalis*, *Symbiopectobacterium* is an exciting model for studying the evolution and dynamics of symbiosis, because it and its close relatives run the gamut, from free-living, well-studied, genetically tractable pathogens (57, 58) to cultivable facultative inherited symbionts, all the way to obligate symbionts. This diversity is also reflected in the dynamic genome evolution of this lineage. Similar to *Sodalis* (33), the *Howardula* symbiont genome exhibits the hallmark of recent symbiosis, as it is very large and full of pseudogenes.

So what is the role of *Symbiopectobacterium* in *Howardula*? There are a number of not mutually exclusive possibilities. The symbiont may be boosting nematode fitness, for example, by aiding in evading the *Drosophila* immune response, or providing a novel function, such as supplementing nutrition or metabolism. One possibility is that the nematode symbiont supplements heme. While nematodes are the only animals known to have lost the ability to synthesize heme (59), two lineages of animal-parasitic nematodes have independently acquired the gene for ferrochelatase, the last step in the synthesis of heme, via horizontal gene transfer from bacteria (60, 61); this enzyme, including conserved active sites, is also encoded in the nematode *Symbiopectobacterium* genome. Alternatively, the symbiont may persist as a result of addiction by the host (11), for example by producing a persistent toxin and its antidote, such that symbiont removal is deleterious. The symbiont may also be providing an essential function that was lost by the host. Comparing genomes of related nematodes with and without *Symbiopectobacterium*, as well as transcriptome studies to uncover highly expressed symbiont genes, may provide some useful clues. Interestingly, the role of the obligate *Wolbachia* symbiont of filarial nematodes is still unclear and metabolic, defensive, and addictive processes have all been invoked (11, 62-64) particularly because a number of species have lost *Wolbachia* without having gained new symbionts or genes (65, 66).

Methods.

PCR survey of *Symbiopectobacterium* in insect-parasitic nematodes

Drosophila flies collected at mushroom baits at sites in N. America, Europe, and Asia, were individually subjected to DNA extraction and screened with specific primers for *Howardula* parasitism and the presence of *Symbiopectobacterium*; additional samples were dissected to determine nematode infection, followed by screening, which was done blind. Additionally, *Fergusonina* flies, that are obligately associated with *Fergusobia* nematodes, were collected at eucalyptus trees and screened for *Symbiopectobacterium*. Sphaerocerid flies were collected at mushroom baits and dissected for nematodes, which were subsequently genotyped and screened for *Symbiopectobacterium*. 16S rRNA amplicon sequencing was performed on *H. aoronymphium* and sphaerocerid nematode motherworms, along with their uninfected fly hosts, *D. falleni* and *Spelobia* sp. Specific information about species sampled, including dates, localities, primers and PCR conditions are available in the SI Appendix, SI Materials and Methods.

Controlled *Howardula* and *Symbiopectobacterium* lab infection and qPCR

Drosophila neotestacea larvae were infected with *H. aoronymphium*. Upon emergence, *Symbiopectobacterium* titer was quantified in two-day-old adult flies, using quantitative PCR. Specific information about lab infection, and primers and qPCR conditions are available in the SI Appendix, SI Materials and Methods.

Antibiotic exposure bioassay

To eliminate *Symbiopectobacterium* from *H. aoronymphium* and measure the subsequent parasitism success, we exposed nematodes to ampicillin or rifampin and compared them to a control group. Exposure lasted the entire parasitic life cycle. Briefly, free-living juvenile nematodes were allowed to parasitize larval *D. neotestacea* and carry out development within fly pupae and adults as internal parasites in vials with antibiotic impregnated food. One week post eclosion, adult flies were dissected to ascertain nematode parasitism, and a subset were screened for *Symbiopectobacterium*. Each treatment was repeated 4 times and the proportion of parasitism was assessed with a chi-squared test. Detailed

procedures for laboratory rearing and antibiotic treatment of *H. aoronymphium* are found in the SI Appendix, SI Materials and Methods.

Microscopy to localize the symbiont of *Howardula*

Laboratory-reared *H. aoronymphium* were dissected from adult *Drosophila*, nematode samples for TEM were prepared with Karnovsky's fixative before being embedded in Epon, whereas FISH samples were fixed in Carnoy's solution for either TEM or FISH microscopy and probed with the general bacterial target, Eub339. Detailed procedures for TEM and FISH are available in the SI Appendix, SI Materials and Methods.

Sequencing, genome binning, and annotation of nematode and bulrush bug *Symbiopectobacterium*

DNA obtained from *H. aoronymphium*, dissected from lab-raised *Drosophila putrida*, was sequenced with Illumina HiSeq technology and genome assembly was performed in UniCycler (67). The *Symbiopectobacterium* genome was separated from the nematode and fly genomes using Blobtools (68) and the de Bruijn graph visualization tool Bandage (69) (SI Appendix, Figure S7). Similar sequencing and assembly methods were employed for the symbiont genome of the bulrush bug (*Chilacis typhae*). Gene and metabolic pathways were annotated with Prokaryotic Genomes Annotation Pipeline (70, 71) and KEGG Automatic Annotation Server (72) for each genome and gene synteny comparisons were made in Geneious 11.1.5 (https://www.geneious.com) with Mauve (73). Detailed description of tissues, extraction, sequence data, and assembly parameters used for genomic analyses are available in the SI Appendix, SI Materials and Methods.

Acknowledgements.
We thank Sonja Scheffer for providing *Fergusobia* nematodes / *Fergusonina* flies; Evan Tandy for maintaining *Drosophila* stocks; Alexandria Marshall and Finn Hamilton for help in the early stages of the project; and Ben Parker and Ellen Martinson for constructive conversations. This study was supported by National Science Foundation grant 1144581 to JJ and grants from the Natural Sciences and Engineering Research Council of Canada (Discovery Grant Program) and the Swiss National Science Foundation (Sinergia grant number CRSII3_154396) to SJP.

Data Accessibility.
Genomes are deposited in GenBank (SyHa, PRJNA415854; SyCt, PRJNA521717; SyPl, PRJNA510127; SyDa, PRJNA510132; SyCl, PRJNA510131) and raw reads for SyHa and SyCt are available at the SRA under the same bioprojects. Bacterial 16S rRNA (MK943676, MT859673, MT859692, MT859693), gyrA (MN175990-MN175995, MT860065), and groEL (MT860063, MT860064) genes. Nematode 18S (MN175314-MN175319, MT863735-MT863741) and CO1 (MN167829-MN167835) genes. Fly hosts CO1 (MT863696-MT863702). 16S rRNA amplicon dataset (PRJNA655365).

References.

1. N. A. Moran, Symbiosis. *Curr Biol* **16**, R866-R871 (2006).
2. N. A. Moran, Symbiosis as an adaptive process and source of phenotypic complexity. *Proc Natl Acad Sci USA* **104**, 8627-8633 (2007).
3. N. Moran, J. P. McCutcheon, A. Nakabachi, Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet* **42**, 165-190 (2008).
4. N. Dubilier, C. Bergin, C. Lott, Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Microbiol* **6**, 725-740 (2008).
5. S. Sudakaran, C. Kost, M. Kaltenpoth, Symbiont acquisition and replacement as a source of ecological innovation. *Trends Microbiol* **25**, 375-390 (2017).
6. G. M. Bennett, N. A. Moran, Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proc Natl Acad Sci USA* **112**, 10169-10176 (2015).
7. F. Husnik, J. P. McCutcheon, Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc Natl Acad Sci USA* **113**, E5416-E5424 (2016).
8. Y. Matsuura *et al.*, Recurrent symbiont recruitment from fungal parasites in cicadas. *Proc Natl Acad Sci USA* **115**, E5970-E5979 (2018).

9. J. P. McCutcheon, B. M. Boyd, C. Dale, The life of an insect endosymbiont from the cradle to the grave. *Curr Biol* **29**, R485-R495 (2019).
10. A. Y. Kostygov *et al.*, Novel trypanosomatid-bacterium association: evolution of endosymbiosis in action. *mBio* **7** (2016).
11. W. Sullivan, *Wolbachia*, bottled water, and the dark side of symbiosis. *MBoC* **28**, 2343-2346 (2017).
12. K. W. Jeon, Development of cellular dependence on infective organisms: micrurgical studies in amoebas. *Science* **176**, 1122-1123 (1972).
13. R. Stouthamer, "*Wolbachia*-induced parthenogenesis" in Influential Passengers: Inherited Microorganisms and Arthropod Reproduction, S. L. O''Neill, A. A. Hoffmann, J. H. Werren, Eds. (Oxford University Press, New York, 1997), pp. 102-124.
14. W. Ma, B. Pannebakker, L. Beukeboom, T. Schwander, L. Van de Zande, Genetics of decayed sexual traits in a parasitoid wasp with endosymbiont-induced asexuality. *Heredity* **113**, 424-431 (2014).
15. T. Nakayama *et al.*, Spheroid bodies in rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. *J Plant Res* **124**, 93-97 (2011).
16. T. Nakayama *et al.*, Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc Natl Acad Sci USA* **111**, 11407-11412 (2014).
17. J. Jaenike, S. J. Perlman, Ecology and evolution of host-parasite associations: mycophagous *Drosophila* and their parasitic nematodes. *Am Nat* **160**, S23-S39 (2002).
18. J. Jaenike, R. Unckless, S. N. Cockburn, L. M. Boelio, S. J. Perlman, Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science* **329**, 212-215 (2010).
19. S. N. Cockburn *et al.*, Dynamics of the continent-wide spread of a *Drosophila* defensive symbiont. *Ecol Lett* **16**, 609-616 (2013).
20. H. E. Welch, Taxonomy, life cycle, development, and habits of two new species of Allantonematidae (Nematoda) parasitic in drosophilid flies. *Parasitology* **49**, 83-103 (1959).
21. S. J. Perlman, G. S. Spicer, D. D. Shoemaker, J. Jaenike, Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Mol Ecol* **12**, 237-249 (2003).
22. W. Ye *et al.*, Molecular phylogenetics and the evolution of host plant associations in the nematode genus *Fergusobia* (Tylenchida: Fergusobiinae). *Mol Phylogenet Evol* **45**, 123-141 (2007).
23. V. Hypša, S. Aksoy, Phylogenetic characterization of two transovarially transmitted endosymbionts of the bedbug *Cimex lectularius* (Heteroptera: Cimicidae). *Insect Mol Biol* **6**, 301-304 (1997).
24. B. C. Campbell, A. H. Purcell, Phylogenetic affiliation of BEV, a bacterial parasite of the leafhopper *Euscelidius variegatus*, on the basis of 16S rDNA sequences. *Curr Microbiol* **26**, 37-41 (1993).
25. S. M. Kuechler, K. Dettner, S. Kehl, Characterization of an obligate intracellular bacterium in the midgut epithelium of the bulrush bug *Chilacis typhae* (Heteroptera, Lygaeidae, Artheneinae). *Appl Environ Microbiol* **77**, 2869-2876 (2011).
26. B. Misof *et al.*, Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763-767 (2014).
27. P. H. Degnan *et al.*, Origin and examination of a leafhopper facultative endosymbiont. *Curr Microbiol* **62**, 1565-1572 (2011).
28. J. B. Benoit *et al.*, Unique features of a global human ectoparasite identified through sequencing of the bed bug genome. *Nat Comm* **7**, 10165 (2016).
29. E. S. Tvedte *et al.*, Genome of the parasitoid wasp *Diachasma alloeum*, an emerging model for ecological speciation and transitions to asexual reproduction. *Genome Biol Evol* **11**, 2767-2773 (2019).
30. X. Li *et al.*, Comparative genomics of 84 *Pectobacterium* genomes reveals the variations related to a pathogenic lifestyle. *BMC Genomics* **19**, 889 (2018).
31. E. Lerat, V. Daubin, N. A. Moran, From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-proteobacteria. *PLoS Biol* **1**, e19 (2003).

32. S. Denman *et al.*, *Brenneria goodwinii* sp. nov., associated with acute oak decline in the UK. *Int J Syst Evol Microbiol* **62**, 2451-2456 (2012).
33. A. L. Clayton *et al.*, A novel human-infection-derived bacterium provides insights into the evolutionary origins of mutualistic insect–bacterial symbioses. *PLoS Genet* **8**, e1002990 (2012).
34. S. M. Kuechler, P. Renz, K. Dettner, S. Kehl, Diversity of symbiotic organs and bacterial endosymbionts of lygaeoid bugs of the families Blissidae and Lygaeidae (Hemiptera: Heteroptera: Lygaeoidea). *Appl Environ Microbiol* **78**, 2648-2659 (2012).
35. A. Oren, A plea for linguistic accuracy–also for *Candidatus* taxa. *Int J Syst Evol Microbiol* **67**, 1085-1094 (2017).
36. C. T. Parker, B. J. Tindall, G. M. Garrity, International code of nomenclature of prokaryotes. Prokaryotic code (2008 revision). *Int J Syst Evol Microbiol* **69**, S7-S111 (2015).
37. M. Sironi *et al.*, Molecular evidence for a close relative of the arthropod endosymbiont *Wolbachia* in a filarial worm. *Mol Biochem Parasitol* **74**, 223-227 (1995).
38. J. E. Palomares-Rius, A. Archidona-Yuste, C. Cantalapiedra-Navarrete, P. Prieto, P. Castillo, Molecular diversity of bacterial endosymbionts associated with dagger nematodes of the genus *Xiphinema* (Nematoda: Longidoridae) reveals a high degree of phylogenetic congruence with their host. *Mol Ecol* **25**, 6225-6247 (2016).
39. T. Vandekerckhove, A. Willems, M. Gillis, A. Coomans, Occurrence of novel verrucomicrobial species, endosymbiotic and associated with parthenogenesis in *Xiphinema americanum*-group species (Nematoda, Longidoridae). *Int J Syst Evol Microbiol* **50**, 2197-2205 (2000).
40. G. O. Poinar, G. M. Thomas, Significance of *Achromobacter nematophilus* Poinar and Thomas (Achromobacteraceae: Eubacteriales) in the development of the nematode, DD-136 (*Neoaplectana* sp. Steinernematidae). *Parasitology* **56**, 385-390 (1966).
41. M. F. Polz *et al.*, Phylogenetic analysis of a highly specific association between ectosymbiotic, sulfur-oxidizing bacteria and a marine nematode. *Appl Environ Microbiol* **60**, 4461-4467 (1994).
42. V. N. Chizhov, N. N. Butorina, S. A. Subbotin, Entomoparasitic nematodes of the genus *Skarbilovinema*: *S. laumondi* and *S. lyoni* (Nematoda: Tylenchida), parasites of the flies of the family Syrphidae (Diptera), with phylogeny of the suborder Hexatylina. *Russ J Nematol* **20**, 141-155 (2012).
43. E. Koshel, V. Aleshin, G. Eroshenko, V. Kutyrev, Phylogenetic analysis of entomoparasitic nematodes, potential control agents of flea populations in natural foci of plague. *BioMed Res Int* http://dx.doi.org/10.1155/2014/13521, article 135218 (2014).
44. K. Sims, J. J. Becnel, J. Funderburk, The morphology and biology of the entomophilic *Thripinema fuscum* (Tylenchida: Allantonematidae), and the histopathological effects of parasitism on the host *Frankliniella fusca* (Thysanoptera: Thripidae). *J Nat Hist* **46**, 1111-1128 (2012).
45. S. A. Subbotin, V. N. Chizhov, Ultrastructure of the integument of parasitic females in entomogenous tylenchids. II. *Howardula phyllotretae*, *Parasitylenchus dispar*, *Contortylenchus* sp., and two allantonematid species. *Russ J Nematol* **4**, 131-138 (1996).
46. V. V. Yushin, H. Kosaka, M. Kusunoki, Spermatozoon ultrastructure in the sphaerularioidid nematode *Contortylenchus genitalicola* (Tylenchomorpha: Sphaerularioidea). *Nematology* **8**, 191-196 (2006).
47. I. L. Riding (1971) Studies on the biology and ultrastructure of nematodes parasitic in insects. in *Department of Zoology & Applied Entomology* (Imperial College London, London, UK).
48. L. A. Nelson *et al.*, An emerging example of tritrophic coevolution between flies (Diptera: Fergusoninidae) and nematodes (Nematoda: Neotylenchidae) on Myrtaceae host plants. *Biol J Linnean Soc* **111**, 699-718 (2014).
49. B. M. Wiegmann *et al.*, Episodic radiations in the fly tree of life. *Proc Natl Acad Sci USA* **108**, 5690-5695 (2011).
50. S. Rossmann, M. W. Dees, J. Perminow, R. Meadow, M. B. Brurberg, Soft rot Enterobacteriaceae are carried by a large range of insect species in potato fields. *Appl Environ Microbiol* **84**, e00281-00218 (2018).
51. F. F. da Mota *et al.*, Cultivation-independent methods reveal differences among bacterial gut microbiota in triatomine vectors of Chagas disease. *PLoS Neg Trop Dis* **6**, e1631 (2012).
52. S. Díaz, B. Villavicencio, N. Correia, J. Costa, K. L. Haag, Triatomine bugs, their microbiota and *Trypanosoma cruzi*: asymmetric responses of bacteria to an infected blood meal. *Parasit Vectors* **9**, 636 (2016).

53. A. H. Purcell, T. Steiner, F. Mégraud, J. Bové, In vitro isolation of a transovarially transmitted bacterium from the leafhopper *Euscelidius variegatus* (Hemiptera: Cicadellidae). *J Invertebr Pathol* **48**, 66-73 (1986).
54. M. Casiraghi *et al.*, Phylogeny of *Wolbachia pipientis* based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. *Microbiology* **151**, 4015-4022 (2005).
55. Y. Nakamura *et al.*, Prevalence of *Cardinium* bacteria in planthoppers and spider mites and taxonomic revision of "*Candidatus* Cardinium hertigii" based on detection of a new *Cardinium* group from biting midges. *Appl Environ Microbiol* **75**, 6757-6763 (2009).
56. K. Takeshita, Y. Kikuchi, *Riptortus pedestris* and *Burkholderia* symbiont: an ideal model system for insect–microbe symbiotic associations. *Res Microbiol* **168**, 175-187 (2017).
57. J. Mansfield *et al.*, Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol* **13**, 614-629 (2012).
58. I. K. Toth, K. S. Bell, M. C. Holeva, P. R. Birch, Soft rot erwiniae: from genes to genomes. *Mol Plant Pathol* **4**, 17-30 (2003).
59. A. U. Rao, L. K. Carta, E. Lesuisse, I. Hamza, Lack of heme synthesis in a free-living eukaryote. *Proc Natl Acad Sci USA* **102**, 4270-4275 (2005).
60. E. Nagayasu *et al.*, Identification of a bacteria-like ferrochelatase in *Strongyloides venezuelensis*, an animal parasitic nematode. *PLoS ONE* **8**, e58458 (2013).
61. B. Wu *et al.*, Interdomain lateral gene transfer of an essential ferrochelatase gene in human parasitic nematodes. *Proc Natl Acad Sci USA* **110**, 7748-7753 (2013).
62. A. C. Darby *et al.*, Analysis of gene expression from the *Wolbachia* genome of a filarial nematode supports both metabolic and defensive roles within the symbiosis. *Genome Res* **22**, 2467-2477 (2012).
63. V. Foray, M. M. Pérez-Jiménez, N. Fattouh, F. Landmann, *Wolbachia* control stem cell behavior and stimulate germline proliferation in filarial nematodes. *Dev Cell* **45**, 198-211.e193 (2018).
64. E. Lefoulon *et al.*, Breakdown of coevolution between symbiotic bacteria *Wolbachia* and their filarial hosts. *PeerJ* **4**, e1840 (2016).
65. C. A. Desjardins *et al.*, Genomics of *Loa loa*, a *Wolbachia*-free filarial parasite of humans. *Nature Genetics* **45**, 495-500 (2013).
66. E. Ferri *et al.*, New insights into the evolution of *Wolbachia* infections in filarial nematodes inferred from a large range of screened species. *PLoS ONE* **6**, e20843 (2011).
67. R. R. Wick, L. M. Judd, C. L. Gorrie, K. E. Holt, Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**, e1005595 (2017).
68. D. R. Laetsch, M. L. Blaxter, BlobTools: interrogation of genome assemblies. *F1000Res* **6**, 1287 (2017).
69. R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350-3352 (2015).
70. D. H. Haft *et al.*, RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* **46**, D851-D860 (2018).
71. T. Tatusova *et al.*, NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* **44**, 6614-6624 (2016).
72. Y. Moriya, M. Itoh, S. Okuda, A. C. Yoshizawa, M. Kanehisa, KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**, W182-W185 (2007).
73. A. E. Darling, B. Mau, N. T. Perna, progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **5**, e11147 (2010).

Figures Legends.

Figure 1. *Symbiopectobacterium* is pervasive in *Howardula aoronymphium*. A. *Symbiopectobacterium* relative abundance in *Drosophila neotestacea* not-parasitized (HA-) or parasitized (HA+) with *Howardula aoronymphium* (qPCR measurements). B-G. Localization of bacteria within larval *Howardula* using FISH microscopy; B. background autofluorescence, C. excitation of probe labeled bacteria. Localization of *Symbiopectobacterium* to the *Howardula* hypodermis. Transmission electron microscopy image showing bacterial cells visible in a *Howardula* young motherworm (D & E) and F1 instar larva (F & G) developing within a *Drosophila neotestacea* individual.

Figure 2. Experimental antibiotic exposure reduces or eliminates successful parasitism in *Howardula aoronymphium*. A. Percent of adult *Drosophila* that were visually parasitized by *Howardula* with the addition of different antibiotics (Control n=268, Ampicillin n=399, Rifampin n=181). B. PCR survey for *Howardula* and *Symbiopectobacterium* in a subset of visually parasitized or non-parasitized *Drosophila* individuals.

Figure 3. Phylogenetic tree of Gammaproteobacteria using the conserved set of 203 single copy orthologous genes. The clade containing the *Howardula* symbiont, *Symbiopectobacterium* is sister to the genera *Pectobacterium, Dickeya,* and *Brenneria* and other soft-rot Enterobacteriaceae. Taxonomic classification of the host organism of each member of the putative symbiont clade. Phylogeny was constructed with RAxML using 100 bootstrap replicates. The full phylogeny is available in SI Appendix, Figure S5.

Figure 4. *Symbiopectobacterium* genomes compared to their closest relative, *Pectobacterium carotovorum*. A. Comparison of genomic features. B. Example of conserved gene order across *Symbiopectobacterium* species and *P. carotovorum*, highlighting the differential pseudogene events across the symbiont clade. Red lines indicate contig boundaries. C. Analysis for central metabolic pathways across *P. carotovorum* and the symbiont clade.

Table 1. PCR survey for *Symbiopectobacterium* in wild-caught *Drosophila neotestacea*.

Table 2. PCR survey for *Symbiopectobacterium* in wild-caught flies and parasitic nematodes.

Figures and Tables

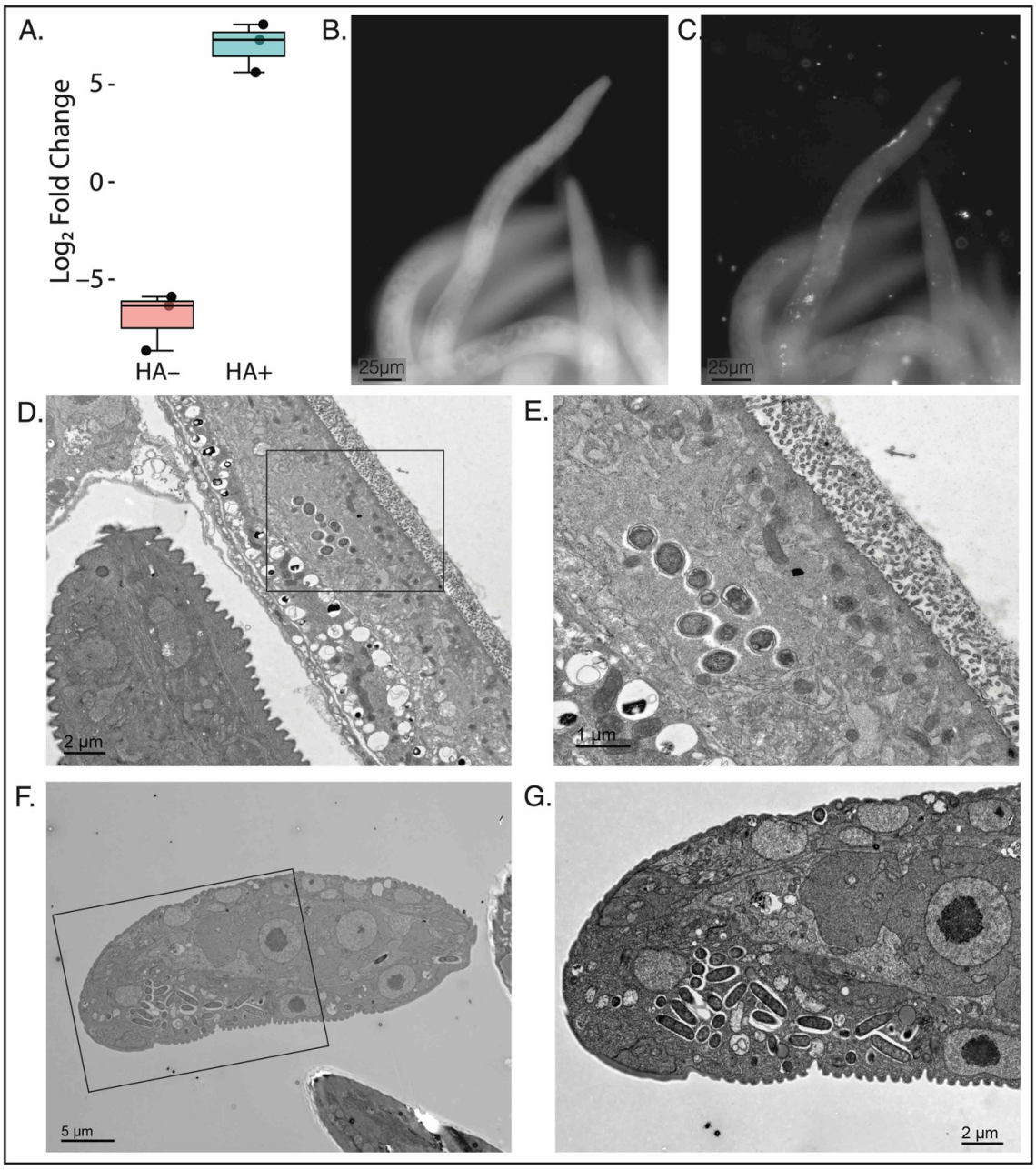

Figure 1. *Symbiopectobacterium* is pervasive in *Howardula aoronymphium*. A. *Symbiopectobacterium* relative abundance in *Drosophila neotestacea* not-parasitized (HA-) or parasitized (HA+) with *Howardula aoronymphium* (qPCR measurements). B-G. Localization of bacteria within larval *Howardula* using FISH microscopy; B. background autofluorescence, C. excitation of probe labeled bacteria. Localization of *Symbiopectobacterium* to the *Howardula* hypodermis. Transmission electron microscopy image showing bacterial cells visible in a *Howardula* young motherworm (D & E) and F1 instar larva (F & G) developing within a *Drosophila neotestacea* individual.

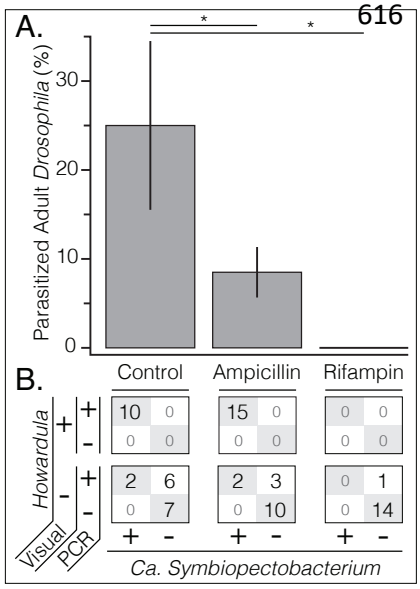

Figure 2. Experimental antibiotic exposure reduces or eliminates successful parasitism in *Howardula aoronymphium*. A. Percent of adult *Drosophila* that were visually parasitized by *Howardula* with the addition of different antibiotics (Control n=268, Ampicillin n=399, Rifampin n=181). B. PCR survey for *Howardula* and *Symbiopectobacterium* in a subset of visually parasitized or non-parasitized *Drosophila* individuals.

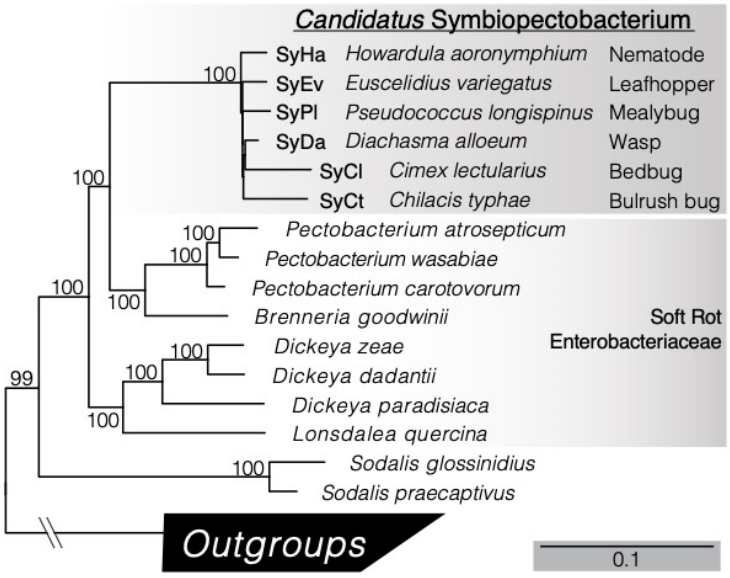

Figure 3. Phylogenetic tree of Gammaproteobacteria using the conserved set of 203 single copy orthologous genes. The clade containing the *Howardula* symbiont, *Symbiopectobacterium* is sister to the genera *Pectobacterium, Dickeya,* and *Brenneria* and other soft-rot Enterobacteriaceae. Taxonomic classification of the host organism of each member of the putative symbiont clade. Phylogeny was constructed with RAxML using 100 bootstrap replicates. The full phylogeny is available in SI Appendix, Figure S5.

Figure 4. *Symbiopectobacterium* genomes compared to their closest relative, *Pectobacterium carotovorum*. A. Comparison of genomic features. B. Example of conserved gene order across *Symbiopectobacterium* species and *P. carotovorum*, highlighting the differential pseudogene events across the symbiont clade. Red lines indicate contig boundaries. C. Analysis for central metabolic pathways across *P. carotovorum* and the symbiont clade.

645

Table 1. PCR survey for *Symbiopectobacterium* in wild-caught *Drosophila neotestacea*.

	Howardula +		*Howardula* –	
	Sym. +	*Sym.* –	*Sym.* +	*Sym.* –
2013	20	0	0	20
2014	8	5	0	27
2015	17	0	0	23
2020	29	0	0	55

Sym., *Symbiopectobacterium;* +, detected by PCR; –, not detected by PCR.

646

Table 2. PCR survey for *Symbiopectobacterium* in wild-caught flies and parasitic nematodes

	Nematode +		*Howardula* –	
	Sym. +	*Sym.* –	*Sym.* +	*Sym.* –
Howardula aoronymphium (England) (Host: *Drosophila*)	1	0	0	19
Howardula aoronymphium (Germany) (Host: *Drosophila*)	4	0	0	16
Howardula sp. B (Japan) (Host: *Drosophila*)	1	0	0	19
Howardula neocosmis (Host: *Drosophila*)	11	0	–	–
Fergusobia sp. (Host: *Fergusonina*)	0	6	–	–
Parasitylenchus nearcticus (Host: *Drosophila*)	0	3	–	–
Spelobia sphaerocerid parasite #1 (Host: *Spelobia*)	0	13	–	–
Spelobia sphaerocerid parasite #2 (Host: *Spelobia*)	0	3	–	–
Spelobia sphaerocerid parasite #3 (Host: *Spelobia*)	0	9	–	–

Sym., *Symbiopectobacterium;* +, detected by PCR; –, not detected by PCR.

647
648
649

Supporting Information for

Multiple origins of obligate nematode and insect symbionts by members of a newly characterized bacterial clade

Vincent G. Martinson, Ryan M. R. Gawryluk, Brent E. Gowen, Caitlin I. Curtis, John Jaenike, Steve J. Perlman

Corresponding author.
Vincent G. Martinson, Department of Biology, MSC03 2020, 1 University of New Mexico, Albuquerque, NM 87131-0001; Tel: 1-406-697-4660; Email: vmartinson@unm.edu

SI Materials and Methods.

Survey of *Candidatus* Symbiopectobacterium (hereafter *Symbiopectobacterium*) in insect-parasitic nematodes.

To determine how taxonomically and geographically widespread the association between *Symbiopectobacterium* and insect-parasitic nematodes extends, we surveyed through a time series, distant localities, and diverse nematode taxa.

Wild-collected *Drosophila neotestacea* were examined for *Howardula aoronymphium* parasitism and the presence of *Symbiopectobacterium*. As part of a long-term survey of nematode prevalence, *D. neotestacea* were collected monthly near Rochester, NY, DNA was extracted (Puregene kit, Qiagen, Valencia, CA, USA), and individuals were screened with nematode-specific primers. We screened 20 putative *Howardula*-parasitized flies and 20 putative *Howardula*-unparasitized flies with nematode-specific and *Symbiopectobacterium*-specific primers (Table S1), that had been collected consecutive Augusts between 2013 and 2015. European flies are parasitized by *H. aoronymphium*, and therefore we screened *Drosophila* collected at mushroom baits in 2009 from Southampton, United Kingdom and Munich, Germany. In 2020, we collected wild *D. neotestacea* at mushroom baits in Victoria, BC, dissected them to determine *Howardula aoronymphium* infection, and screened them, along with uninfected flies, for *Symbiopectobacterium;* infection status was not known to the person performing the PCR screening.

We also surveyed the closest relatives of *H. aoronymphium,* two species that also parasitize *Drosophila*: *H. neocosmis*, which is native to Florida and western N. America, and an unnamed species of *Howardula* found in E. Asia. We screened 11 *D. munda* parasitized with *H. neocosmis* (Chiricahua mountains, AZ, USA, 2008) and 20 *Drosophila* sp. collected in Sapporo, Japan (2009) for the closely related unnamed *Howardula* species.

Outside of the *aoronymphium/neocosmis/*unnamed clade, the closest related genus is *Fergusobia*. These nematodes are obligate symbionts with herbivorous flies (*Fergusonina* sp.), where they are parasites of adults and live mutualistically in plant galls with fly larvae. We screened for *Symbiopectobacterium* in three fly individuals of two species of *Fergusonina*. These specimens were collected Australia from their native host trees: 1) Falls Creek Village, Victoria, *Eucalyptus pauciflora*, 24-Sep-2007; 2) Kangaroo Island, South Australia, *Eucalyptus cosmophylla*, 19-Nov-2007. We also screened the more distantly related *Parasitylenchus nearcticus*, a nematode that is native to N. America and can also parasitize *D. neotestacea* (Vancouver, BC, 2002). Finally, we collected and dissected wild sphaerocerid flies (mostly *Spelobia* sp.) at mushroom baits in Victoria, BC, in 2019 and 2020. We amplified and sequenced nematode rRNA and attempted to amplify *Symbiopectobacterium* with a range of primers. There were three tylenchid nematode species in these samples. *Fergusobia* and *Parasitylenchus* were screened with purK spacer and gyrA primers for *Symbiopectobacterium* (Table S1). Sphaerocerid parasites were screened with gyrA, groEL, 16S, and universal 16S primers.

Nematode infection and *Symbiopectobacterium* titer.

Howardula aoronymphium, *D. falleni*, and *D. neotestacea* were originally collected from mushroom baits in West Hartford, CT, USA in 2006, and maintained in the laboratory on vials containing mushrooms (*Agaricus bisporus*) and Carolina instant fly food; nematodes are maintained in *D. falleni* in the lab. *Drosophila neotestacea* previously cured of *Wolbachia* and *Spiroplasma* infections were incubated overnight on mushroom agar at 21°C. Eggs were collected in groups of 20 the subsequent morning, and placed on mushroom wedges that were treated with either *Drosophila falleni* homogenate containing approximately 400 *H. aoronymphium* infective juveniles in Ringer's solution (HA^+ condition), or an equivalent volume of homogenate from uninfected *D. falleni* (HA^- condition). A single mushroom wedge was placed onto moistened cheesecloth in each vial.

Adult flies were collected immediately after eclosion and placed in solitary mushroom vials for two days. Individual flies were homogenized in 50 µl of lysis buffer containing 100 mM Tris, pH 7.6, 100 mM sodium chloride, 50 mM disodium EDTA, 1% w:v sodium dodecyl sulfate, and 0.1 mg/ml proteinase K with 10-15 1 mm silica-zirconium beads in a Mini Beadbeater (BioSpec Products) for 10 s. Homogenates were incubated at 65°C for one hour and extracted twice with 50 µl buffer-saturated phenol and 25 µl chloroform. The final aqueous phase was combined with 5 µl of 3.3 M sodium acetate and precipitated with 110 µl of ethanol. The pellet was washed twice with 200 µl of 75% ethanol and dissolved in 20 µl of water.

Titer measurements for *Symbiopectobacterium* were carried out in a CFX96 Real-Time System interfaced with CFX Manager 3.0 software using Applied Biosystems SYBR Select Master Mix (Applied Biosystems, Carlsbad, CA, USA). The *D. neotestacea rpl32 gene* was used as a positive control, and the *rplQ* gene was used to measure *Symbiopectobacterium* titer. Three biological replicates were measured for each condition (*i.e.*, HA^- or HA^+), and each primer set was measured in triplicate for each sample. Fold changes in *Symbiopectobacterium* titer were calculated according to the Pfaffl method (1). Because rplQ failed to amplify in all HA^- samples, we treated the Ct as 40. To compare HA^+ and HA^- samples side-by-side, we subtracted the Ct of each sample from that of the global mean. Primer efficiency (E) was incorporated into the calculations as $1+E^{\Delta Ct}$.

Effect of antibiotic on nematode parasitism.

A population of *H. aoronymphium*, initially isolated from *D. neotestacea* collected at mushroom baits (*Agaricus bisporus*) near Rochester, NY, was maintained in the laboratory using the host fly *Drosophila putrida*. Briefly, parasitized adult *D. putrida* were homogenized in sterile H_2O to make a suspension of free-living nematodes which were used to infect the next generation of *D. putrida* larvae. Adult flies were allowed to oviposit for two days in standard *Drosophila* vials that contained Formula 4-24 Instant *Drosophila* Medium (Carolina Biological Supply Co, Burlington, NC, USA) hydrated with water and a small mushroom slice (*Agaricus bisporus*). After oviposition, adults were removed from the vial and the slurry containing juvenile nematodes was added. Nematodes were allowed to parasitize the *Drosophila* larvae and resulting adult *D. putrida* were screened for nematode infection. *Howardula* used in this study had been maintained in the laboratory for greater than five generations.

To determine the effect of antibiotics on nematode parasitism, lab-raised *D. neotestacea* were infected with *H. aoronymphium* in the absence of antibiotics, with ampicillin, or with rifampicin. Nematodes were exposed in vials that were prepared as above except the 4-24 media was hydrated with antibiotic stocks to a final concentration of (100 μg/ml). Larvae were allowed to fully develop and adult *D. neotestacea* were collected 6-7 days post eclosion to ensure that infecting-nematodes had the chance to produce juveniles. Adult flies were dissected to observe nematodes and the proportion parasitized was recorded. The parasitism experiment was performed in four replicate vials. To independently confirm the presence of *H. aoronymphium* and *Symbiopectobacterium*, a subset of 10-15 flies that were visually positive or negative were used for DNA extraction and screened with specific primers. Variation between replicates was checked with a normal logistic fit test and significance was assessed in JMP using the chi-square test.

Microscopy to localize the symbiont of *Howardula.*

Samples for TEM (a lab culture of *H. aoronymphium* adults and infectives, dissected from adult *Drosophila falleni* and *D. neotestacea*) were double fixed and Epon embedded using standard TEM methodology (2). After initial fixation in Karnovsky's fixative, the small samples were embedded within low melt agarose (Sigma A9045) and the resulting blocks osmicated, dehydrated, infiltrated with Epon, and the Epon polymerized. TEM sections were stained in uranyl acetate and lead citrate and viewed in a Jeol JEM 1400 TEM at 80 kV. Images were captured using a Gatan SC-1000 digital camera.

The legs and head were removed from *H. aoronymphium*-parasitized *D. putrida* (from a lab culture) to allow better penetration of the fixative and probes. Specimens were fixed with Carnoy's solution overnight and rinsed with 100% EtOH using previously published methods (3). Fluorescent in situ hybridization (FISH) microscopy was performed on as outlined in Martinson et al. 2012, on larval nematodes that were released from the *D. putrida* body cavity following the final rinse step. The general bacterial probe Eub339 was used to label all bacteria in larval nematodes (4).

Genome of *Symbiopectobacterium* associated with the nematode, *Howardula aoronymphium.*

A mix of juvenile and motherworm *H. aoronymphium* were dissected from infected *D. putrida* in 1x PBS buffer. The resulting nematodes were further rinsed twice with 1x PBS, for 30 sec in 2% bleach, and again twice with 1x PBS. DNA was extracted with the DNeasy blood & tissue kit (Qiagen, Valencia, CA, USA) and sequenced on the Illumina HiSeq2500 using 250 bp paired-end reads. Initial cleaning and trimming of the data was performed with Trimmomatic-0.32 (5) using the commands "SLIDINGWINDOW:4:20 TRAILING:13 LEADING:13 ILLUMINACLIP:adapters.fasta:2:30:10 MINLEN:15". Because the sample contained a mix of DNA from the host (*Drosophila*), the parasite nematode (*Howardula*), and microorganisms (including *Symbiopectobacterium* and any gut-associated microbiota),

an initial assembly was performed to partition contigs by taxonomic identity. Contigs assembled with Metaspades (6) were classified by top blast hits to the GenBank nr database (accessed Feb 2016) and visualized in plots based on length, GC content, and coverage using Blobtools v1.0 (7). The fly and nematode genomes were highly fragmented and could not be improved; however, we identified a putative *Symbiopectobacterium* genome that comprised the long, high-coverage contigs and had top blast hits to the genera *Pectobacterium* and *Brenneria* (Figure S7ab).

To improve the assembly and isolate this bacterial genome, a subset of the cleaned reads (24 million paired reads) were used for assembly with Unicycler v0.2.0 with the default settings (8). This assembler utilizes SpADES (9) with multiple steps to iteratively improve assembly and to close circular DNA (e.g. plasmids, bacterial chromosomes). Additionally, Unicycler outputs the de Bruijn graphs used to create the assembly which preserve the tentative links between contigs that could not be integrated into the final assembly because repeat regions create conflicting signal. After assembly, contigs were again assigned a taxonomic identity by blast and Blobtools. By loading each contigs taxonomic identity into the genome assembly visualization tool Bandage (10) with the "csv" option and "Node labels: CSV data", all linked sets of contigs could be classified, binned as individual genomes, and analyzed separately (Figure S7cd). The contigs identified as Proteobacteria were further error corrected and checked for overlaps with the polish command in Unicycler. Contigs with fewer than 200 bp (256) were removed from the assembly, after they were shown to not have sufficient sequence similarity to identify a top hit in GenBank or they matched to insertion sequence (IS) families that were abundant in the genome. The final genomic contigs were ordered relative to the closest genome sequence, *Pectobacterium carotovorum*, with Mauve (11) using the progressiveMauve feature. Gene prediction and annotation (including pseudogene prediction) was performed by the Prokaryotic Genomes Annotation Pipeline (PGAP) of NCBI (12, 13). To ensure that the removal of short contigs (<200bp) had not artifactually led to the removal of genes in metabolic pathway annotations, the top blast hit for each contig was identified.

Genome of *Symbiopectobacterium* associated with the bulrush bug, *Chilacis typhae.*

Chilacis typhae were collected from bulrushes (*Typha*) at Rithet's Bog, Victoria, British Columbia, Canada. Individual insects (~40 in total) were dissected with sterile pins, and the red 'symbiont belt' was excised with a sterile razor blade. Symbiont belts were combined in 100 µl of cell lysis buffer containing 100 mM Tris pH 7.6, 100 mM sodium chloride, 50 mM disodium EDTA, 1% sodium dodecyl sulfate, and 0.1 mg/ml proteinase K, and manually disrupted with a plastic pestle. The homogenate was incubated overnight at 65°C, and extracted twice with 100 µl buffer-saturated phenol and 50 µl chloroform, precipitated with 100 µl of isopropanol, washed twice in 200 µl of 75% ethanol, and dissolved in TE buffer (10 mM Tris, 1 mM disodium EDTA, pH 7.6). 0.3 µl of 100 mg/ml RNAse A (Qiagen) was added and the solution was incubated for 20 minutes at room temperature. DNA was extracted, precipitated, washed, and dissolved as described above. The presence of symbiont DNA was confirmed via Sanger sequencing after PCR amplification of the 16S rRNA gene with primer pair Symbiopecto_16S F and R (Table S1).

Genomic DNA libraries were generated at the Genome Québec Innovation Centre using the NEBNext Ultra II kit (New England Biolabs), and 125 bp paired-end reads were generated on an Illumina HiSeq 2500 instrument from one third of a flow cell. The BBTools suite was used for quality control of reads (https://sourceforge.net/projects/bbmap/). Raw sequence reads were trimmed for quality and adaptor sequence with BBDuk, and reads ≥ 40 bp were retained. To remove reads from less abundant symbiont belt bacteria prior to assembly, BBNorm was used to filter out kmers represented at a depth of less than 200. SPAdes 3.10.1 (9) was used to assemble the genome, employing kmer values of 21, 33, 55, 77, and 99. Megablast (cite) was used to further decontaminate the genome assembly. Genomic contigs for the symbiont chromosome were isolated with Bandage, <200 bp contigs removed, and annotated with PGAP as performed for the *H. aoronymphium* symbiont.

Genome comparisons.

While querying GenBank with the *Howardula*-associated *Symbiopectobacterium*, several highly-similar bacterial matches were identified within insect genome sequencing projects, including: the bed bug (*Cimex lectularius*), the long-tailed mealybug (*Pseudococcus longispinus*), and a parasitoid wasp (*Diachasma alloem*). Additionally, the incomplete genomic assembly of the symbiont "BEV" from a leafhopper (*Euscelidius variegatus*) contained a few orthologous gene fragments that were found to have high sequence similarity to *Symbiopectobacterium* (14). Sets of contigs, tentatively representing partial bacterial genomes, were extracted from each of the insect genome projects, re-ordered, and annotated

with previously mentioned methods. The KEGG Automatic Annotation Server was used to find orthologous genes and to evaluate the completeness of central metabolic pathways across the genus *Symbiopectobacterium* and *P. carotovorum* (15). Gene order synteny was compared in Geneious 11.1.5 (https://www.geneious.com) using the multiple-genome alignment feature of the Mauve plugin (11).

Phylogenetic reconstruction.

The gyrase A gene (gyrA) of *Symbiopectobacterium* and CO1 and 18S rRNA genes from the nematode and fly host were amplified from a subset of the *aoronymphium*/*neocosmis*/unnamed specimens. Sanger sequencing was performed at Genewiz, Inc (South Plainfield, NJ, USA) or Sequetech (Mountain View, CA, USA) and sequences were cleaned using Geneious. For each sequenced gene (18S rRNA, CO1, gyrA) and for the bacterial 16S rRNA, closely related sequences were identified via blast to GenBank and searching the Ribosomal Database Project (16). Sequences were aligned with MUSCLE (17) in Geneious and phylogenies were built with RAxML (18) using 100 bootstrap replicates, except for the 16S rRNA phylogeny that was built with Fasttree (19).

The genome phylogeny of Gammaproteobacteria was created with the 203 single-copy orthologs (20). The gene set from *Escherichia coli K-12 substr. MG1655* was used to query each genome to find top blast hits, each gene was individually aligned with MAFFT (21), and trimmed with gblocks (22, 23). The resulting alignments were concatenated, and the full genome phylogeny was built with RAxML using the GAMMA WAG model with 100 bootstrap replicates.

Dating the *Symbiopectobacterium-Howardula* symbiosis.

The age of the *Symbiopectobacterium* lineage was estimated by calculating the divergence time of from *Pectobacterium* using sequence-based substitution rate measurements (24). Pairwise SiCO genes were identified for each genome pair with the online tool OrthoVenn2 (25). Genes identified as pseudogenes in the PGAP annotation pipeline (12) were excluded from further analysis to constrain the analysis to only the most conserved regions of the genome. Amino acid sequence pairs for each SiCO gene were individually aligned with MUSCLE (17) and codon-aligned nucleotide sequences were created with pal2nal (26). Sequence changes were analyzed at only the 2nd codon positions because these are the most functionally constrained sites. Substitutions at 2nd codon positions were calculated with custom perl scripts incorporating codeml in PAML (27) and divergence times were estimated using the measurement of 2.2x10^{-7} substitutions/site/year for the insect endosymbiont *Buchnera aphidicola* (28).

16S rDNA amplicon sequencing.

DNA was extracted and 16S rDNA amplicon sequencing was performed, for the following samples: a) *Howardula aoronymphium* motherworms dissected out of *D. falleni*, from a nematode culture maintained in the Perlman lab, and originally collected in West Hartford, CT, USA in 2006 (n=6), b) uninfected *D. falleni* (n=4), c) sphaerocerid parasite species #3 motherworms, dissected from *Spelobia* sp. flies collected at mushrooms baits in Victoria, BC, in 2020 (n=6), and d) uninfected *Spelobia* sp. collected at the same baits (n=4).

Tissue was homogenized in 100 uL of lysis buffer (100mM Tris-HCL, 100mM NaCl, 50mM Na2EDTA and 1%SDS) and 0.5mg/mL Proteinase K (New England Biolabs Ltd.) in a Mini-Beadbeater. Samples were incubated at 37˚C for 20 minutes, followed by a de-activation of Proteinase K at 95˚C for 5 minutes. A 1/10 volume of 3.3 M sodium acetate and 1 uL of Glycogen (Thermoscientific) were added, and samples were incubated at 65˚C for one hour. DNA was isolated using the phenol-chloroform method and precipitated with ethanol.

All DNA samples were normalized to a concentration of 20 ng/µl and submitted to GeneWiz (South Plainfield, New Jersey) for construction of V3V4 hypervariable amplicon libraries, and 2 x 250 bp paired-end sequencing using an Illumina MiSeq platform (*i.e.*, 16S-EZ service). Cutadapt v2.6 (50) was used to remove 5' primers, and low-quality bases (< 20) from the 3' end of reads. All subsequent analyses were carried out with R v.4.0.2 (R Core Team) within RStudio v1.2.5033 (RStudio Team). DADA2 v1.16 (51) was used for denoising, dereplication, merging, chimera detection, and inference of amplicon sequence variants (ASVs). Briefly, forward and reverse reads were discarded if they were likely to contain more than two erroneous base calls or were shorter than 200 bp; pseudo-pooling was used in ASV inference; and read merging required ≥ 20 bp of overlapping sequence. Taxonomic assignment of ASVs was performed using Decipher v2.16.1 (52) based on the SILVA SSU r138 2019 dataset. Sampling depth was normalized according to the variance stabilizing transformation using DESeq2 (53). Phyloseq v1.32.0

(54) was used to generate a proportion table containing sums of all ASVs within each bacterial family contained in the dataset, and to subsequently filter the table to include only ASVs that make up > 5% of any individual sample. A taxonomic boxplot was generated with ggplot2 v3.3.2 (55) to depict the proportion of ASV counts from each bacterial family as a function of host (Figure S2).

SI References

1. M. W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45 (2001).
2. M. Hayat, *Principles and techniques of electron microscopy: biological applications* (CRC Press, Boca Raton, FL, USA, ed. 3rd, 1989).
3. R. Koga, T. Tsuchida, T. Fukatsu, Quenching autofluorescence of insect tissues for in situ detection of endosymbionts. *Appl Entomol Zool* **44**, 281-291 (2009).
4. V. G. Martinson, J. Moy, N. A. Moran, Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl Environ Microbiol* **78**, 2830-2840 (2012).
5. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
6. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**, 824-834 (2017).
7. D. R. Laetsch, M. L. Blaxter, BlobTools: interrogation of genome assemblies. *F1000Res* **6**, 1287 (2017).
8. R. R. Wick, L. M. Judd, C. L. Gorrie, K. E. Holt, Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**, e1005595 (2017).
9. A. Bankevich *et al.*, SPAdes: a new genome assembly algorithm and Its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477 (2012).
10. R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350-3352 (2015).
11. A. E. Darling, B. Mau, N. T. Perna, progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **5**, e11147 (2010).
12. D. H. Haft *et al.*, RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* **46**, D851-D860 (2018).
13. T. Tatusova *et al.*, NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* **44**, 6614-6624 (2016).
14. P. H. Degnan *et al.*, Origin and examination of a leafhopper facultative endosymbiont. *Curr Microbiol* **62**, 1565-1572 (2011).
15. Y. Moriya, M. Itoh, S. Okuda, A. C. Yoshizawa, M. Kanehisa, KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**, W182-W185 (2007).
16. J. R. Cole *et al.*, Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**, D633-D642 (2014).
17. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).
18. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690 (2006).
19. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
20. E. Lerat, V. Daubin, N. A. Moran, From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-proteobacteria. *PLoS Biol* **1**, e19 (2003).
21. K. Katoh, D. M. Standley, MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**, 772-780 (2013).
22. J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**, 540-552 (2000).
23. G. Talavera, J. Castresana, Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56**, 564-577 (2007).
24. A. L. Clayton *et al.*, A novel human-infection-derived bacterium provides insights into the evolutionary origins of mutualistic insect–bacterial symbioses. *PLoS Genet* **8**, e1002990 (2012).

25. L. Xu *et al.*, OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res* **47**, W52-W58 (2019).
26. M. Suyama, D. Torrents, P. Bork, PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* **34**, W609-W612 (2006).
27. Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**, 1586-1591 (2007).
28. N. A. Moran, H. J. McLaughlin, R. Sorek, The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. *Science* **323**, 379-382 (2009).
29. B. C. Campbell, A. H. Purcell, Phylogenetic affiliation of BEV, a bacterial parasite of the leafhopper *Euscelidius variegatus*, on the basis of 16S rDNA sequences. *Curr Microbiol* **26**, 37-41 (1993).
30. F. F. da Mota *et al.*, Cultivation-independent methods reveal differences among bacterial gut microbiota in triatomine vectors of Chagas disease. *PLoS Neg Trop Dis* **6**, e1631 (2012).
31. S. Díaz, B. Villavicencio, N. Correia, J. Costa, K. L. Haag, Triatomine bugs, their microbiota and *Trypanosoma cruzi*: asymmetric responses of bacteria to an infected blood meal. *Parasit Vectors* **9**, 636 (2016).
32. O. Duron *et al.*, The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol.* **6**, 12 (2008).
33. T. Hosokawa, R. Koga, Y. Kikuchi, X.-Y. Meng, T. Fukatsu, *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc Natl Acad Sci USA* **107**, 769-774 (2010).
34. F. Husnik, J. P. McCutcheon, Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc Natl Acad Sci USA* **113**, E5416-E5424 (2016).
35. R. Koga, N. A. Moran, Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. *ISME J* **8**, 1237-1246 (2014).
36. S. M. Kuechler, P. Renz, K. Dettner, S. Kehl, Diversity of symbiotic organs and bacterial endosymbionts of lygaeoid bugs of the families Blissidae and Lygaeidae (Hemiptera: Heteroptera: Lygaeoidea). *Appl Environ Microbiol* **78**, 2648-2659 (2012).
37. Y. Matsuura *et al.*, Evolution of symbiotic organs and endosymbionts in lygaeid stinkbugs. *ISME J* **6**, 397-409 (2012).
38. A. R. Pitman, P. J. Wright, M. D. Galbraith, S. A. Harrow, Biochemical and genetic diversity of pectolytic enterobacteria causing soft rot disease of potatoes in New Zealand. *Australasian Plant Pathology* **37**, 559-568 (2008).
39. H. Salmonová, J. Killer, V. Bunešová, M. Geigerová, E. Vlková, Cultivable bacteria from *Pectinatella magnifica* and the surrounding water in South Bohemia indicate potential new Gammaproteobacterial, Betaproteobacterial and Firmicutes taxa. *FEMS Microbiol Lett* **365** (2018).
40. N. Someya, Y. Ohdaira Kobayashi, S. Tsuda, S. Ikeda, Molecular characterization of the bacterial community in a potato phytosphere. *Microbes and Environments* **28**, 295-305 (2013).
41. E. Yashiro, P. S. McManus, Effect of streptomycin treatment on bacterial community structure in the apple phyllosphere. *PLoS ONE* **7**, e37131 (2012).
42. M. J. Ballinger, R. M. R. Gawryluk, S. J. Perlman, Toxin and genome evolution in a *Drosophila* defensive symbiosis. *Genome Biol Evol* **11**, 253-262 (2018).
43. R. M. Floyd, A. D. Rogers, P. J. D. Lambshead, C. R. Smith, Nematode-specific PCR primers for the 18S small subunit rRNA gene. *Molecular Ecology Notes* **5**, 611-612 (2005).
44. O. Folmer, M. Black, W. Hoeh, R. Lutz, R. Vrijenhoek, DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3**, 294-299 (1994).
45. J. Jaenike, T. D. Brekke, Defensive endosymbionts: a cryptic trophic level in community ecology. *Ecol Lett* **14**, 150-155 (2011).
46. J. R. Marchesi *et al.*, Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl Environ Microbiol* **64**, 795-799 (1998).
47. V. G. Martinson *et al.*, A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol* **20**, 619-628 (2011).
48. J. P. Sandström, J. A. Russell, J. P. White, N. A. Moran, Independent origins and horizontal transfer of bacterial symbionts of aphids. *Mol Ecol* **10**, 217-228 (2001).

49. L. Iasur-Kruh *et al.*, Bacterial associates of *Hyalesthes obsoletus* (Hemiptera: Cixiidae), the insect vector of bois noir disease, with a focus on cultivable bacteria. *Res Microbiol* **168**, 94-101 (2017).

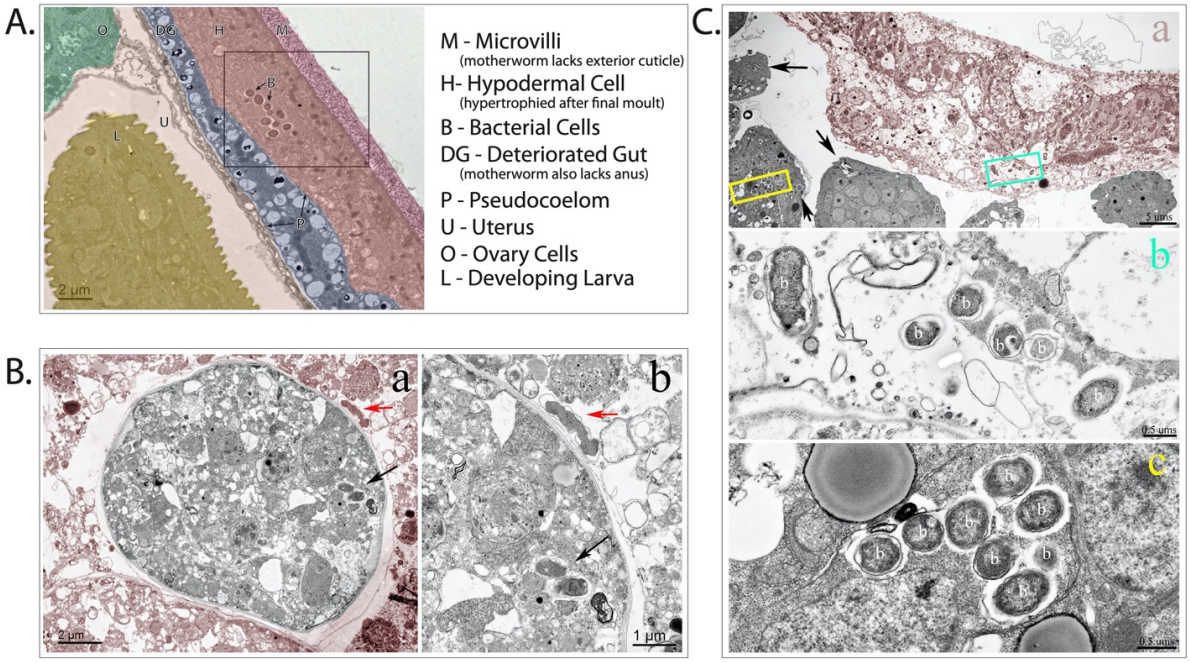

Figure S1. Tissue localization of bacteria within *Howardula aoronymphium* adults and juveniles.

Plate A. Localization of symbionts within a motherworm.

Plate B. Low magnification of sepia colored mother worm material with a grey scaled developing embryos/daughter worm within (a), higher magnification in (b) of the plate. The bacteria within the mother worm (red arrow for both figures) is external to the developing embryo/daughter worm. The bacteria within the developing embryo/daughter worm is shown by the black arrows.

Plate C. Low magnification of a sepia colored mother worm cell wall and grey scaled developing embryos/daughter worms within the lumen of the mother worm (a). The teal colored box encloses bacteria within the motherworm and is the magnified in (b) of the plate (no longer sepia colored). The yellow colored box encloses bacteria within a developing embryo/daughter worm and is the magnified in (c) of the plate. The arrows within *figure a* show other groups of bacteria. The easily recognized bacteria in the magnified figures are labeled with a ‘b’.

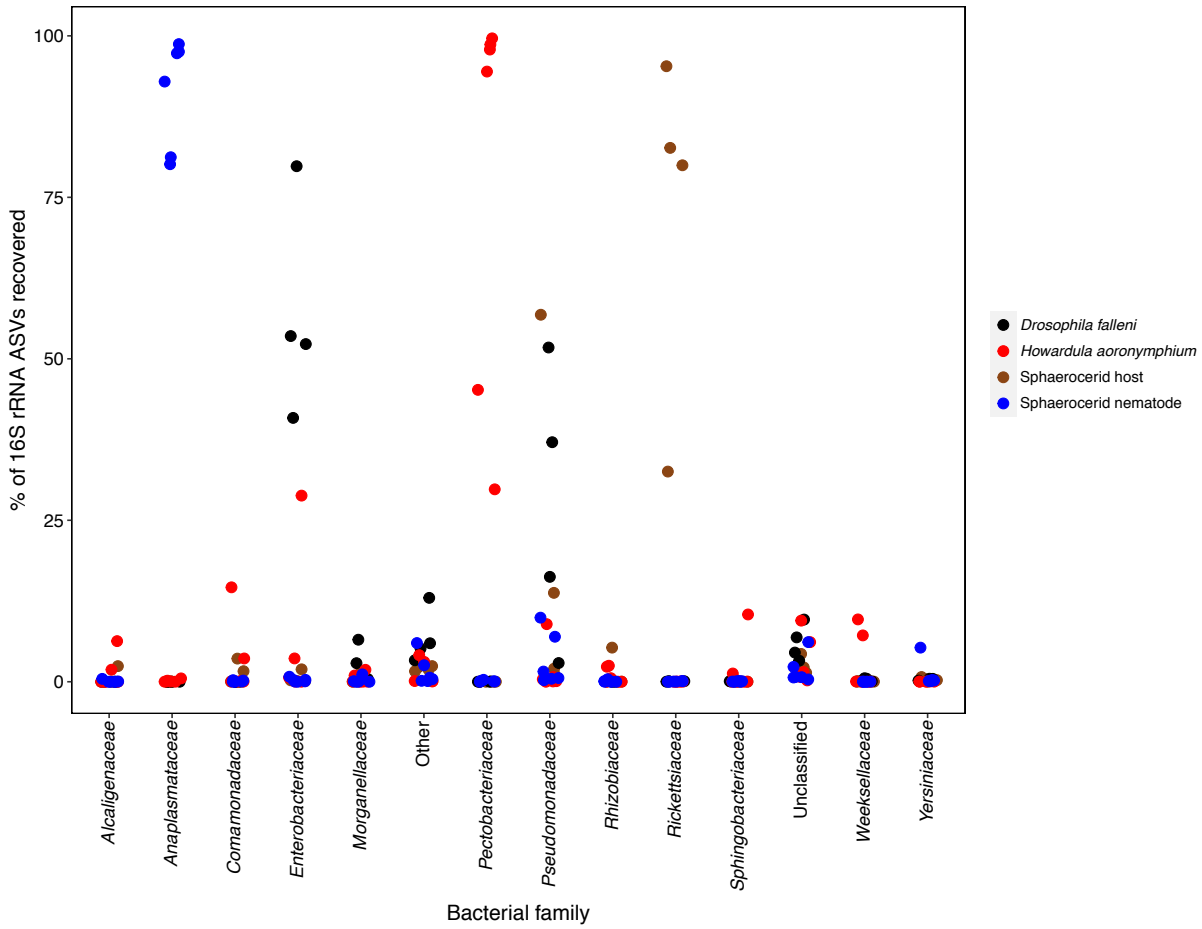

Figure S2. Relative abundance of V3V4 amplicon sequence variants according to bacterial family. The proportion of ASV counts corresponding to each bacterial family present across metagenomic datasets of *Drosophila falleni*, *Howardula aoronymphium*, *Spelobia* sp. (Sphaerocerid host) and its nematode (Sphaerocerid nematode species #3). ASVs that could not be resolved to the level of family are reported as unclassified. In order to avoid displaying excessive numbers of low-abundance bacterial families, we show only those that constitute > 5% of any individual sample. *Symbiopectobacterium* (=Pectobacteriaceae) is only present in *H. aoronymphium*, and is the only microbe found exclusively in all *H. aoronymphium* samples.

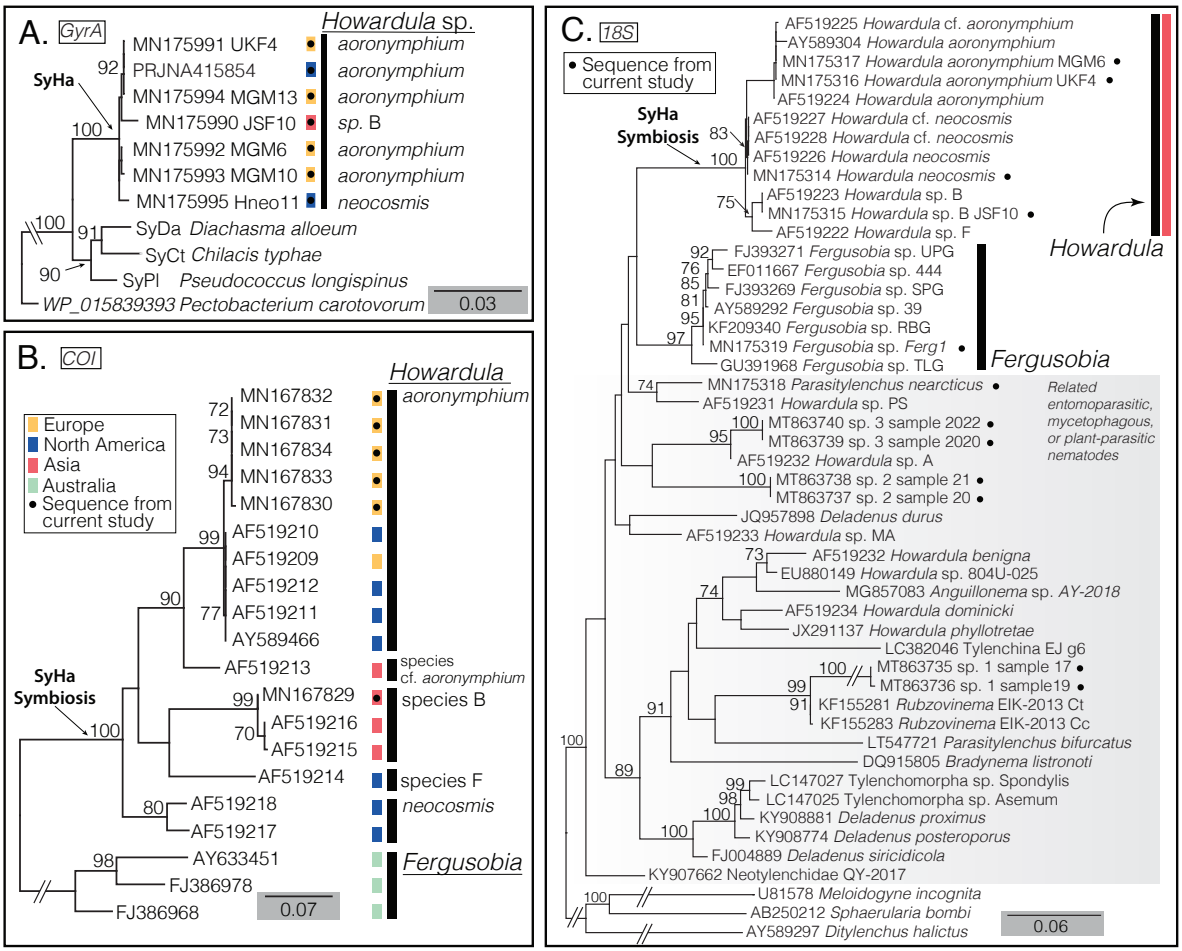

Figure S3. Phylogeny of *Symbiopectobacterium* using the gyrase A gene (gyrA) (A.). Phylogeny of nematode genes COI (B.) and 18S rRNA (C.). The *Howardula* clade harboring *Symbiopectobacterium* is highlighted with a red bar in the 18S rRNA phylogeny of nematode diversity (C). Sequences from the current study are marked with a black circle.

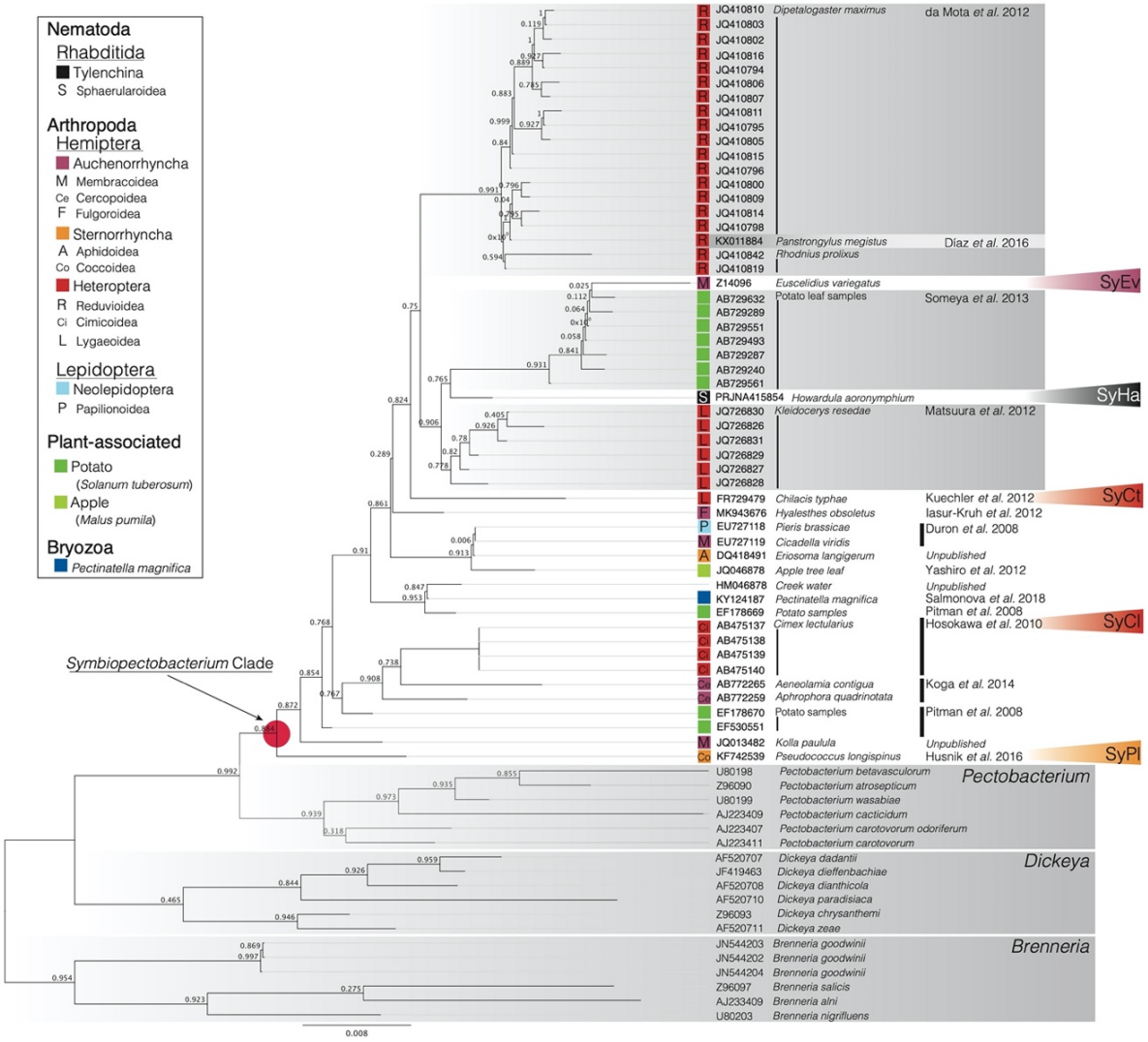

Figure S4. 16S rRNA phylogeny of *Symbiopectobacterium* clade members identified in GenBank. Host taxonomy and sequence citations are listed on the tree (29-41).

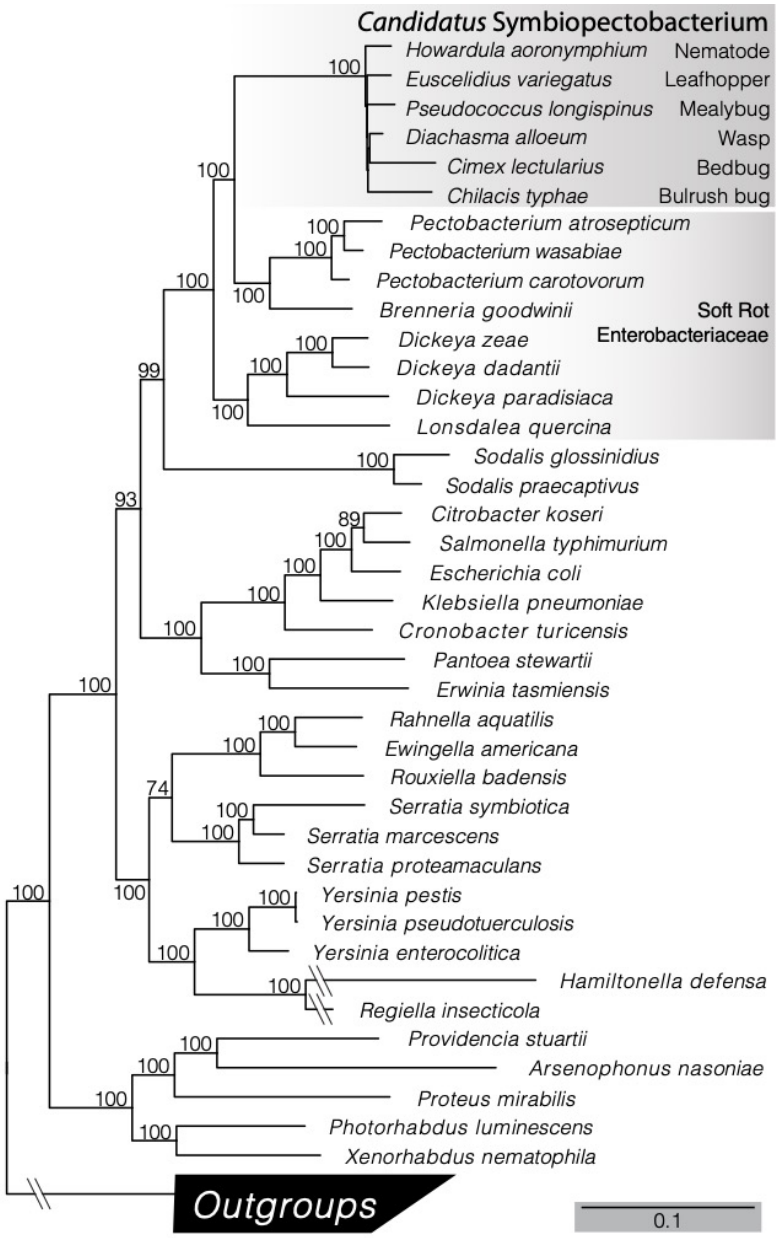

Figure S5. Phylogenetic tree of Gammaproteobacteria using the conserved set of 203 single copy orthologous genes. Phylogeny was constructed with RAxML using 100 bootstrap replicates.

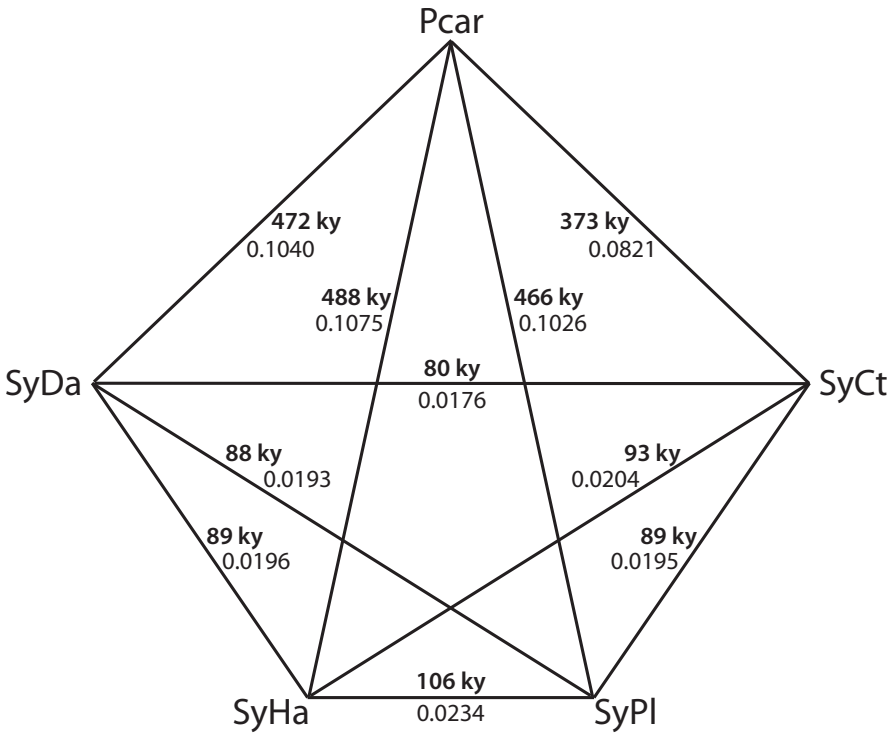

Figure S6. Divergence time estimates among *Symbiopectobacterium* species and their closest relative, *Pectobacterium carotovorum*. Substitution/site rate was calculated at the 2nd codon position for intact single copy ortholog (SiCO) genes (excluding pseudogenes). Divergence time was calculated using the measurement of 2.2×10^{-7} substitutions/site/year for the insect endosymbiont *Buchnera aphidicola* (28). Estimated divergence time (ky = thousand years) is in bold, substitution/site at 2nd codon position for intact genes is below. The number of SiCO genes used for comparisons differed for each pair (Pcar–SyCt 779; Pcar–SyPl 1692; Pcar–SyDa 1992; Pcar–SyHa 2001; SyCt–SyPl 707; SyCt–SyDa 754; SyCt–SyHa 741; SyPl–SyDa 1573; SyPl–SyHa 1786; SyDa–SyHa 1763).

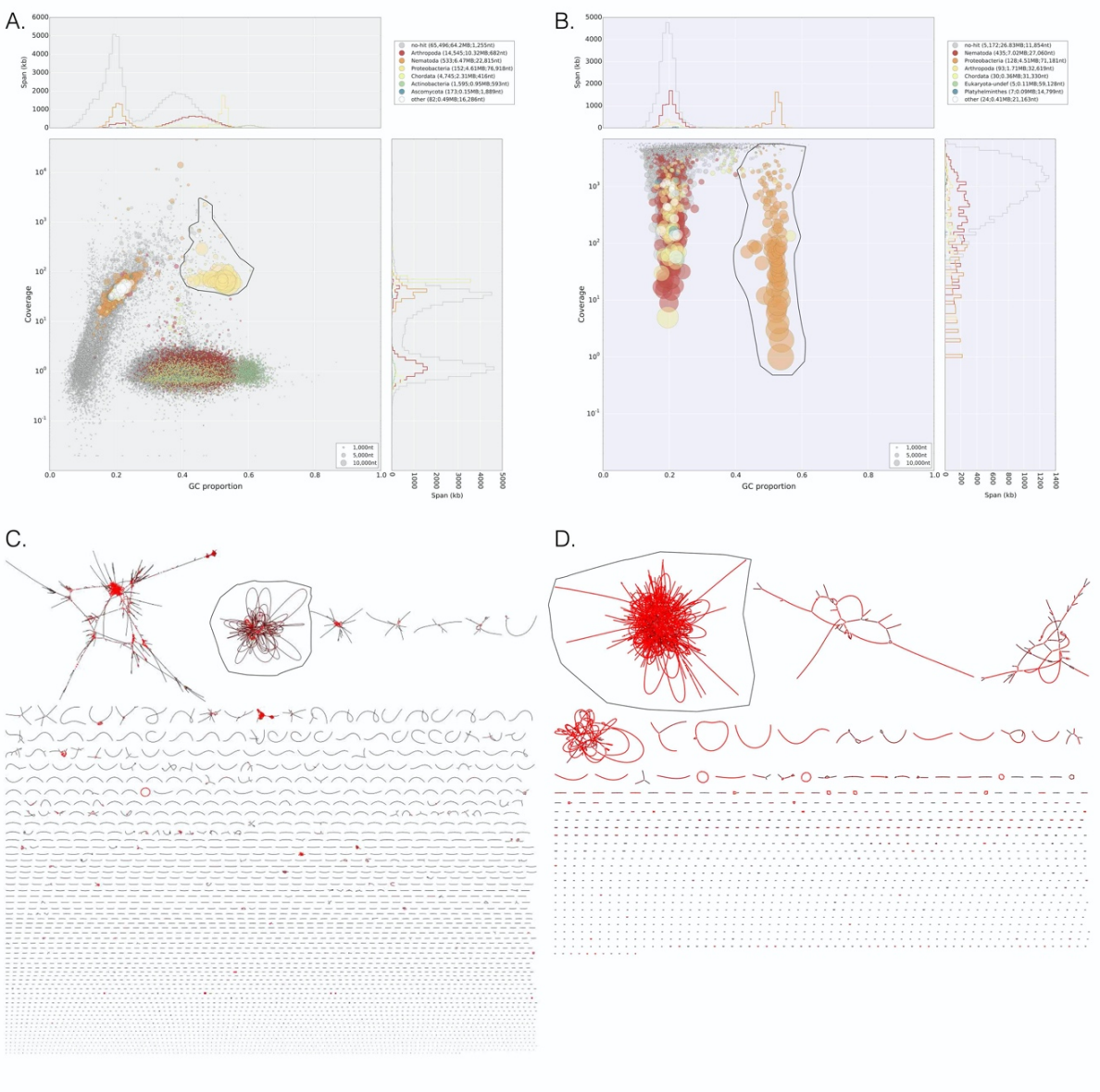

Figure S7. Identifying the symbiont genome contigs within complex assemblies. A) Blobtools image of *Symbiopectobacterium* contigs among the *Howardula aoronymphium* A) metaSPAdes genome assembly and B) Unicycler. Visualization of the De Bruijn graphs for C) *Howardula* nematode symbiont (SyHa) and D) *Chilacis* bulrush bug symbiont (SyCt) constructed in Unicycler assembly showing links between contigs. Color indicates relative coverage of each contig (red = highest coverage).

Table S1. Primers and conditions for PCR and qPCR used in this project (42-48).

Table S1. Primers and conditions for PCR and qPCR

Primer reference	Target group	Target gene	Target length	Primer Name	Primer sequence (5'- to -3')	Cycling conditions
Folmer *et al.* 1994	*Drosophila* & sphaerocerids	*COI*	~700 bp	LCO1490 HCO2198	GGTCAACAAATCATAAAGATATTGG TGATTTTTTGGTCACCCTGAAGTTTA	
Jaenike & Brekke 2011	*Howardula*	*COI*	~300 bp	aorF aorR	TTCGTTTGGAGCTTTCCAAACCTGG AACACTYGAACCCACATGACCCAA	
Martinson *et al.* 2011 Sandström *et al.* 2001	Bacteria	*16S rRNA*	~1500 bp	27f-short 1507r	GAGTTTGATCCTGGCTCA TACCTTGTTACGACTTCACCCCAG	PCR •94°C for 2 min •35 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 90 s •72°C for 10 min
Current study	*Ca.* Symbiopectobacterium	*purK* spacer	~1350 bp	n8_purK-p1498_F n8_purK-p1498_R	GTGACGCGCAGAACCATTTT GCTGTCGAGCTCTTCGTCTT	
Current study	*Ca.* Symbiopectobacterium	*gyrA*	~1150 bp	SymbiopectogyrA_2f SymbiopectogyrA_2r	GTGACGCGCAGAACCATTTT GCTGTCGAGCTCTTCGTCTT	
Floyd 2005	*Howardula*	*18S rRNA*	~900 bp	Nem_18S_F Nem_18S_R	CGCGAATRGCTCATTACAACAGC GGGCGGTATCTGATCGCC	
Current study	*Ca.* Symbiopectobacterium	*groEL*	158 bp	Symbiopecto_groEL_F Symbiopecto_groEL_R	CGGTGTGAACGTTCTTGCTA ATGTCGATCATGTTGCCGTA	PCR •94°C for 3 min •35 cycles of 94°C for 10 s, 55°C for 15 s, 72°C for 30 s •72°C for 5 min
Current study	*Ca.* Symbiopectobacterium	*16S rRNA*	1207 bp	Symbiopecto_16S_F Symbiopecto_16S_R	GAGCTTGCTCTCTGGGTGAC TATGAGGTCCGCTTGCTCTT	PCR •94°C for 3 min •35 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 90 s •72°C for 7 min
Marchesi *et al.* 1998 Sandström *et al.* 2001	Bacteria	*16S rRNA*	1444 bp	63F 1507r	CAGGCCTAACACATGCAAGTC TACCTTGTTACGACTTCACCCCAG	
Current study	*Ca.* Symbiopectobacterium	*rplQ*	158 bp	Symbiopecto_rplQ_F Symbiopecto_rplQ_R	GTTGTTGAACCGCTGATTACTCT TTCAGAATGCGAGTATAACCACC	qPCR Primer efficiency: 106.4%; $r^2 = 0.993$
Ballinger *et al.* 2019	*Drosophila*	*rpl32*	161 bp	DrosRpL32-qF DrosRpL32-qR	CGCACCAAGCACTTCATCCG AGCGACGACGTACTCTGTTGT	qPCR

Table S2. Sequence information and collection source for members of *Symbiopectobacterium* (29-41, 49).

Table S2. Sequence information and collection source for members of *Symbiopectobacterium*

GenBank Accession	Citation	Genome	Symbiosis Type	Host	Common name	Phylum	Order	Suborder	Superfamily	Notes
JQ013482	Unpublished	–	Facultative	*Kolla paulula*	Leafhopper	Arthropoda	Hemiptera	Auchenorrhyncha	Membracoidea	Described as *Xylella*.
Z14096	Campbell *et al*. 1993	Degnan et al. 2011	Facultative	*Euscelidius variegatus*	Leafhopper	Arthropoda	Hemiptera	Auchenorrhyncha	Membracoidea	
EU727119	Duron *et al*. 2008	–	Facultative	*Cicadella viridis*	Green leafhopper	Arthropoda	Hemiptera	Auchenorrhyncha	Membracoidea	
MK943676	Iasur-Kruh *et al*. 2017	–	Facultative	*Hyalesthes obsoletus*	Planthopper	Arthropoda	Hemiptera	Auchenorrhyncha	Fulgoroidea	
AB772265	Koga *et al*. 2014	–	Facultative	*Aeneolamia contigua*	Froghopper	Arthropoda	Hemiptera	Auchenorrhyncha	Cercopoidea	Not mentioned in the publication.
AB772259	Koga *et al*. 2014	–	Facultative	*Aphrophora quadrinotata*	Spittlebug	Arthropoda	Hemiptera	Auchenorrhyncha	Cercopoidea	Described as *Sodalis*.
KF742539	Husnik & McCutcheon 2016	Husnik & McCutcheon 2016	Obligate	*Pseudococcus longispinus*	Long-tailed mealybug	Arthropoda	Hemiptera	Sternorrhyncha	Coccoidea	
DQ418491	Unpublished	–	Facultative	*Ericsoma lanigerum*	Wooly apple aphid	Arthropoda	Hemiptera	Sternorrhyncha	Aphidoidea	
JQ726826-31	Matsuura *et al*. 2012	–	Facultative	*Kleidocerys resedae*	Birch catkin bug	Arthropoda	Hemiptera	Heteroptera	Lygaeoidea	
FR729479	Kuechler *et al*. 2012	This study	Obligate	*Chilacis typhae*	Bulrush bug	Arthropoda	Hemiptera	Heteroptera	Lygaeoidea	
HE586117	Kuechler *et al*. 2012	This study	Obligate	*Belonochilus numenius*	Sycamore seed bug	Arthropoda	Hemiptera	Heteroptera	Lygaeoidea	16S sequence has deletion
AB475137-140	Hosokawa *et al*. 2010	Benoit et al. 2016	Faclutative	*Cimex lectularius*	Bed bug	Arthropoda	Hemiptera	Heteroptera	Cimicoidea	
JQ410807 sequence group	da Mota *et al*. 2012	–	Facultative	*Rhodnius* sp., *Dipetalogaster* sp.	Kissing bug	Arthropoda	Hemiptera	Heteroptera	Reduvioidea	
KX011884	Díaz *et al*. 2016	–	Facultative	*Rhodnius* sp., *Panstrongylus* sp.	Kissing bug	Arthropoda	Hemiptera	Heteroptera	Reduvioidea	
EU727118	Duron *et al*. 2008	–	Facultative	*Pieris brassicae*	Cabbage butterfly	Arthropoda	Lepidoptera	Neolepidoptera	Papilionoidea	
	This study	Tvedte et al. 2019	Facultative	*Diachasma alloeum*	Rhagoletis parasitoid wasp	Arthropoda	Hymenoptera	Apocrita	Ichneumonoidea	
	This study	This study	Obligate	*Howardula aoronymphium, H. neocosmis, H. sp.*	*Drosophilla*-parasitic nematode	Nematoda	Rhabditida	Tylenchina	Sphaerularioidea	
KY124187	Salmonova *et al*. 2018	–	Facultative	*Pectinatella magnifica*	Freshwater bryozoan	Bryozoa	Plumatellida		Pectinatellidae	
JQ046878	Yashiro *et al*. 2012	–	?	Plant, *Malus domestica*	Apple	Angiospermophyta	Rosales	Rosaceae	Amygdaloideae	
EF530551, EF178669, EF178670	Pitman *et al*. 2008	–	?	Plant, *Solanum tuberosum*	Potato	Angiospermophyta	Solanales	Solanaceae	Solanoideae	Described as *Dickeya*.
AB729551 sequence group	Someya *et al*. 2013	–	?	Plant, Solanum tuberosum	Potato	Angiospermophyta	Solanales	Solanaceae	Solanoideae	