Co-obligate symbioses have repeatedly evolved across aphids, but partner identity and nutritional contributions vary across lineages

Alejandro Manzano-Marín[1,*], Armelle Coeur d'acier[2], Anne-Laure Clamens[2], Corinne Cruaud[3], Valérie Barbe[3], and Emmanuelle Jousselin[2]

[1]Centre for Microbiology and Environmental Systems Science, University of Vienna. Djerassiplatz 1, 1030 Vienna, Austria
[2]UMR 1062 Centre de Biologie pour la Gestion des Populations, INRA, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier. 755 avenue du Campus Agropolis, 34988 Montferrier-sur-Lez, France
[3]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay. 2 Rue Gaston Crémieux, 91057 Evry, France
[*]alejandro.manzano.marin@gmail.com

ABSTRACT

Aphids are a large family of phloem-sap feeders. They typically rely on a single bacterial endosymbiont, *Buchnera aphidicola*, to supply them with essential nutrients lacking in their diet. This association with *Buchnera* was described in model aphid species from the Aphidinae subfamily and has been assumed to be representative of most aphids. However, in two lineages, *Buchnera* has lost some essential symbiotic functions and is now complemented by additional symbionts. Though these cases break our view of aphids harbouring a single obligate endosymbiont, we know little about the extent, nature, and evolution of these associations across aphid subfamilies. Here, using metagenomics on 25 aphid species from nine subfamilies and 16S rRNA amplicon sequencing on 223 aphid samples (147 species from 12 subfamilies), we show that dual symbioses have evolved anew at least six times. We also show that these secondary co-obligate symbionts have typically evolved from facultative symbiotic taxa. Genome-based metabolic inference confirms interdependencies between *Buchnera* and its partners for the production of essential nutrients but shows contributions vary across pairs of co-obligate associates. Fluorescent *in situ* hybridisation microscopy shows a common bacteriocyte localisation of two newly acquired symbionts. Lastly, patterns of *Buchnera* genome evolution reveal that small genome reductions affecting a few key genes can be the onset of these dual systems, while, on the other hand, large gene losses can occur without any co-obligate symbiont acquisition. Hence, the *Buchnera*-aphid association, often thought of as exclusive, appears fragile, with a few metabolic losses having recurrently promoted the establishment of a new co-obligate symbiotic partner.

Keywords: symbiont replacement, metabolic complementarity, aphid, nutritional symbiosis, co-obligate symbiont, *Buchnera*.

Introduction

Nutritional symbioses between animals and microorganisms are widespread in nature, and are particularly present in hemipteran insects with a nutrient-restricted diet. Aphids (Hemiptera: Apdididae) are a family of around 5,200 species organised into 23 extant subfamilies (Favret, 2022). Their diet consists entirely of plant phloem, which is rich in sugars, but poor in essential amino acids and B vitamins (Douglas, 2006; Sandstrom and Moran, 1999; Ziegler, 1975). In order to overcome this limitation, aphids are associated to an obligate vertically-transmitted endosymbiotic bacteria, *Buchnera aphidicola*, that synthesises nutrients lacking in their diet, namely essential amino acids (EAAs) and B vitamins (Baumann, 2005; Bermingham *et al.*, 2009; Blow *et al.*, 2020; Douglas, 1998; Hansen and Moran, 2011; Shigenobu *et al.*, 2000). The long-term association of *Buchnera* and aphids is evidenced by the high degree of *Buchnera* genome synteny and general congruence of aphid and symbiont phylogenies (Baumann *et al.*, 1995; Funk *et al.*, 2000; Jousselin *et al.*, 2009; Moran *et al.*, 1993; Nováková *et al.*, 2013; van Ham *et al.*, 2003). As a result of this ancient association, *Buchnera* strains have evolved highly reduced, AT-rich, and gene-dense genomes (Chong and Moran, 2018). While genome erosion has affected many functional categories, extant *Buchnera* strains have markedly retained genes involved in the biosynthesis of essential amino acids and B vitamins. However intimate the aphid-*Buchnera* symbiosis is, long-lived associations can break down, leading to symbiont replacement or complementation.

While ancient symbiont replacement and complementation has been well documented in hemipteran taxa such as the Auchenorrhyncha (*e.g.* cicadas, treehoppers, leafhoppers, and planthoppers; Koga and Moran, 2014; Łukasik *et al.*, 2018; Matsuura *et al.*, 2018; McCutcheon and Moran, 2007; Michalik *et al.*, 2021), Pseudococcinae (mealybugs; Husnik and McCutcheon, 2016; Szabó *et al.*, 2017), Adelgidae (adelgids; Dial *et al.*, 2022; Szabó *et al.*, 2022; Toenshoff *et al.*, 2012; Weglarz *et al.*, 2018), and Psylloidea (psyllids; Nakabachi *et al.*, 2013, 2020; Sloan and Moran, 2012), these associations have not been widely reported across aphids. The mutualism between aphids and *Buchnera* has generally been seen as stable and quite exclusive. However, most of our knowledge on this symbiosis comes from one aphid lineage: the Aphidinae subfamily. This subfamily encompasses about 3,150 species (around 60% of aphid diversity) including most aphid pests, and thus, has been the one that has been most studied. Yet, microbial metagenomic data from aphids outside Aphidinae is revealing that the aphid/*Buchnera* relationship might not be as stable as widely thought. In *Geopemphigus* and within Cerataphidini aphids, *Buchnera* has been lost, and in its place, *Bacteroidota* (Chong and Moran, 2018) and yeast-like symbionts are found in these species (Fukatsu and Ishikawa, 1992; Fukatsu *et al.*, 1994), respectively. In both cases, the new symbionts have the genetic capacity to take over the nutrient provisioning role of the now defunct *Buchnera* (Chong and Moran, 2018; Vogel and Moran, 2013), which would allow the aphids to continue thriving on a nutrient-deficient diet. In addition to symbiont replacement, symbiont complementation can arise if a co-existing microbe has the metabolic capacity to rescue or take over one or more of the roles of the pre-existing associate. Such co-obligate symbioses have indeed arose in at least two aphid groups: the Lachninae (Manzano-Marín *et al.*, 2017) and the *Periphyllus* genus (Chaitophorinae: Chaitophorini;

Monnin *et al.*, 2020; Renoz *et al.*, 2022a). Very recently, an additional two reports of co-obligate symbioses have been published in two species from two different tribes: *Sipha maydis* (Chaitophorinae: Siphini; Renoz *et al.*, 2022b), which has been idenpendently analysed in the current work, and *Ceratovacuna japonica* (Hormaphidinae: Cerataphidini; Yorimoto *et al.*, 2022). These secondary co-obligate symbionts now complement evolved auxotrophies of their corresponding *Buchnera* partners, most commonly those for biotin and riboflavin and less often those for tryptophan and histidine. While Monnin *et al.* (2020) also reported alleged co-obligate symbiotic systems in two Aphidinae species, post-publication re-analysis of their work has cast important doubts regarding some of their analyses and interpretation of results (Manzano-Marín, 2020). Thus, these co-obligate associations within Aphidinae are not further discussed in the context of the current work.

In the remaining aphid subfamilies, there is no genomic evidence for the occurrence of multi-partner symbioses. However, a series of microscopic studies have revealed co-occurring bacteriocyte-associated bacteria outside of the Lachninae subfamily and *Periphyllus* genus which show co-obligate like characteristics. Early microscopic evidence showed co-existing bacteria in separate bacteriocytes to those of *Buchnera* in the aphids *Drepanosiphum* sp. (Drepanosiphinae), *Periphyllus testudinaceus* (Chaitophorinae: Chaitophorini), and *Panaphis juglandis* (Calaphidinae: Panaphidini) (Buchner, 1953). More recent work has also shown co-obligate like organisms in *Yamatocallis* spp. (Drepanosiphinae) (Fukatsu, 2001; Fukatsu and Ishikawa, 1993), *Sipha maydis* (Chaitophorinae: Siphini), *Anoecia corni* (Anoeciinae), and *Glyphina betulae* (Thelaxinae) (Kot, 2012; Michalik, 2010; Michalik *et al.*, 2014). In most above-mentioned cases, the co-obligate symbionts are not only inhabiting their own bacteriocytes, but also show a spherical cell shape, which is characteristic of many obligate symbionts of aphids and other insects with drastically reduced genomes (Dial *et al.*, 2022; Lamelas *et al.*, 2008, 2011b; Manzano-Marín *et al.*, 2016, 2017, 2020; Michalik *et al.*, 2021; Szabó *et al.*, 2022; Toenshoff *et al.*, 2012, 2014).

In this work we sought to explore the extent, nature, identity, and metabolic capabilities of nutritional endosymbiotic consortia across aphids. For this purpose, we assembled the most comprehensive and diverse set of aphid symbionts to date. This dataset included 25 newly sequenced symbiont genomes as well as 20 re-assembled and/or re-annotated previously sequenced ones. Through high-throughput 16S rRNA gene amplicon sequencing of 147 species of aphids, we were able to corroborate the identity of these symbionts, their prevalence across species from the same lineages, and explored whether new obligate symbionts necessarily emerge from frequent symbionts in the microbiota. In addition, through genome-based metabolic inference we explored the co-dependency of the co-existing symbionts and collaboration for the production of their hosts' essential nutrients. Through phylogenetic analyses we investigated the origin of co-obligate symbiont lineages. Lastly, using of Fluorescence *in situ* hybridisation (FISH), we corroborate the identity and bacteriocyte-specific localisation of two distantly-related secondary co-obligate symbionts in selected aphid species.

Results

Buchnera genome has repeatedly undergone further genome reduction

In order to reconstruct a comprehensive phylogeny of *Buchnera aphidicola* (hereafter *Buchnera*) to aid in our evolutionary interpretations, we assembled a genomic dataset of 48 strains housed by different aphid species (supplementary table S1, Supplementary Material online). This dataset represents aphid symbionts from 13 different subfamilies, including the most speciose ones, and is, to our knowledge, the largest and most diverse dataset assembled for *Buchnera*. As noted previously by Chong *et al.* (2019), *Buchnera* from Aphidinae have among the largest genome sizes as well as highest G+C content and number of coding sequences (CDSs, figure 1). Conversely, the genomes of Lachninae, known to host co-obligate symbionts (Manzano-Marín *et al.*, 2017, 2020; Meseguer *et al.*, 2017), have among the smallest genomes with under 400 CDSs and low G+C contents. With our expanded dataset, we have also observed many genomes falling along this extremes, namely those of Chaitophorinae, Thelaxinae, Neophyllaphidinae, Anoeciinae, Mindarinae, and one Eriosomatini.

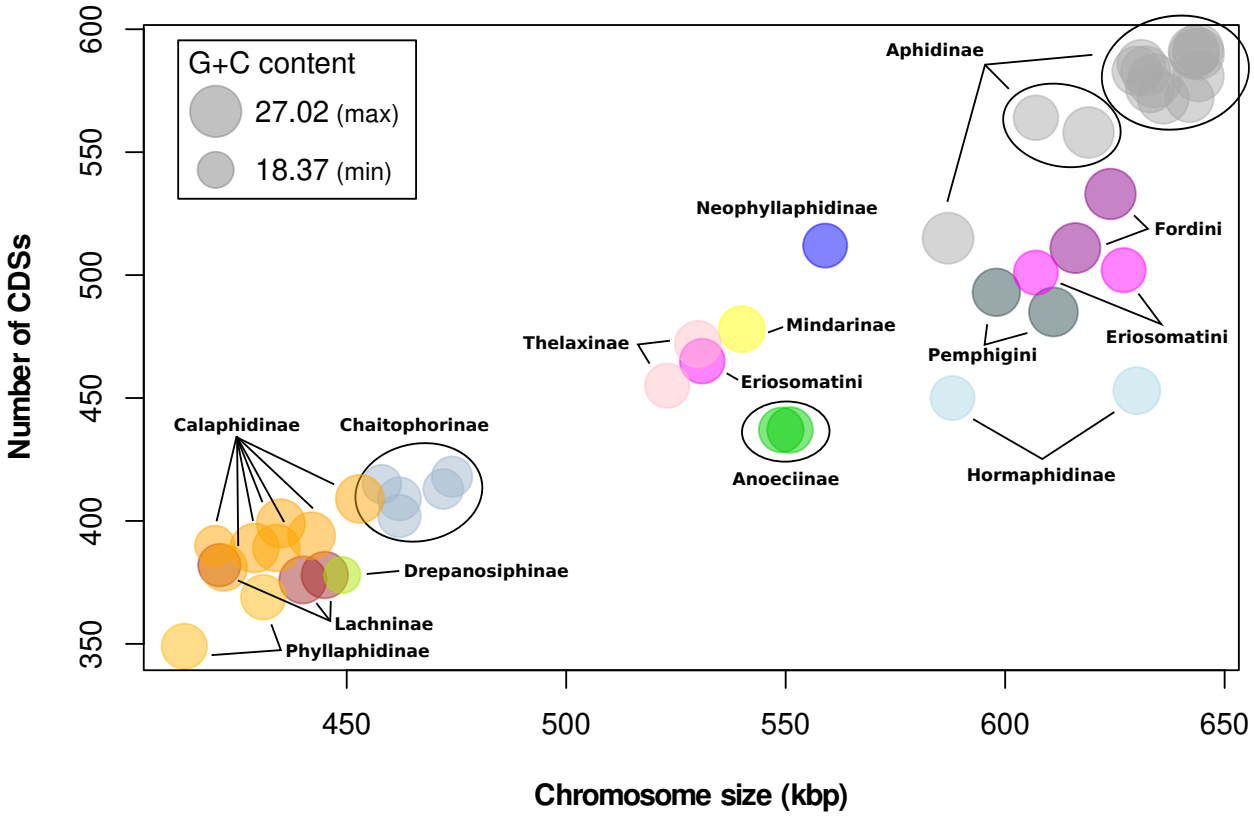

Figure 1. Relationship of genome size *vs.* CDS content in *Buchnera* genomes. Scatter plot illustrating the diversity of genome size, number of CDSs and G+C content of selected *Buchnera* genomes. Taxonomic classification of aphid host is displayed in bold black lettering. Dots are colour coded according to the host's subfamily or tribal classification. Size of circles represent the G+C content as shown in the legend box.

After thorough manual curation of the *Buchnera* genomes, we extracted the 229 single-copy core proteins and used them to reconstruct a concatenated protein phylogeny for *Buchnera* (figure 2). As previous works (Chen *et al.*, 2017; Li *et al.*, 2014; Nováková *et al.*, 2013; Ortiz-Rivas and Martínez-Torres, 2010), we found support for the non-monophyly of the Eriosomatinae subfamily. Also, the symbionts of both sequenced Phyllaphidinae species were recovered nested within those of Calaphidinae, which corroborates previous results recovering members of this subfamily as closely related to species or tribes of Calaphidinae (Nováková *et al.*, 2013; von Dohlen and Moran, 2000). It is important to note that there remains some uncertainty regarding the position of the root, given the long branch leading to the *Buchnera* lineage. However, alternative phylogenetic reconstructions with and without the more distantly related *Escherichia coli* as an outgroup, gave identical topologies with small variations in support for some internal branching (alternative trees found in https://doi.org/10.5281/zenodo.6394197).

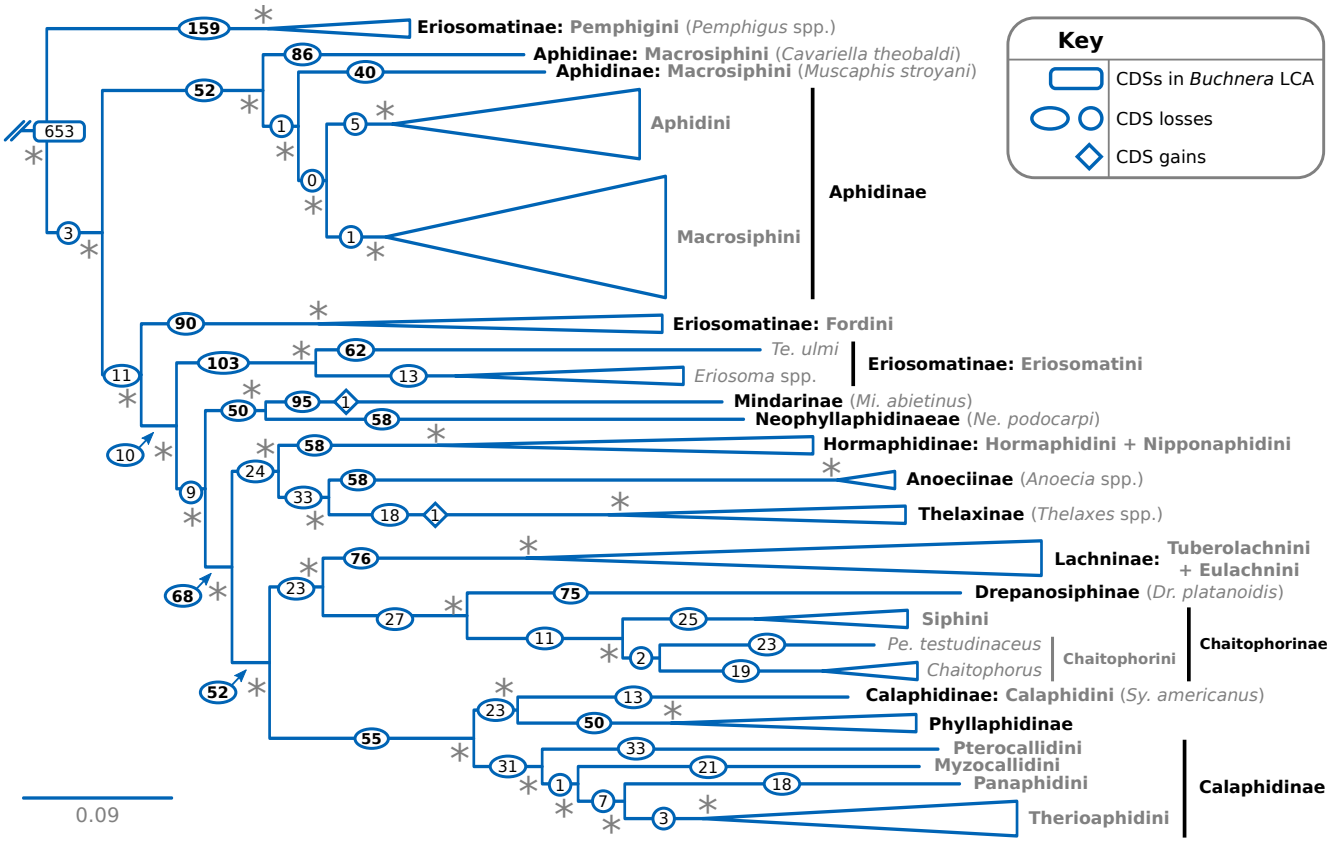

Figure 2. Phylogenetic relationships of *Buchnera*. Truncated Bayesian phylogenetic tree based on concatenated single-copy core proteins of selected *Buchnera*. Branches were collapsed on the basis of their aphid hosts' subfamily classification. At the right of leaves and clades, names of aphid subfamilies and tribes are shown in bold in black and grey colours, respectively. When a subfamily is only represented by the *Buchnera* strain of one aphid species or genus, it is indicated in between parentheses.Given that *fliM+fliN* and *fliO+fliP* often occurred as fused genes, they were each counted as one protein cluster instead of two. An asterisk at nodes stands for a posterior probability of 1. *Escherichia coli*, *Pantoea ananatis*, and *Erwinia* spp. were used as outgroups. Full phylogenetic tree can be found at (supplementary figure S1, Supplementary Material online).

The phylogenetic tree, when combined with the genomic characteristics of figure 1 shows that, in agreement with Chong *et al.* (2019), the genome of *Buchnera* has undergone multiple events of genome reduction after diversification. The largest *Buchnera* genomes (>= 580 kbp) are retained in strains from Pemphigini, Aphidinae, Eriosomatini, Fordini, and Hormaphidinae. On the other side, small genomes of under 580 kbp have evolved once within the Eriosomatini (in *Tetraneura ulmi*), and three times in the branches leading to the Mindarinae + Neophyllaphidinae, Anoeciinae + Thelaxinae, and Lachninae + Drepanosiphinae + Chaitophorinae + Calaphidinae. After manual curation of the orthologous protein clusters of *Buchnera*, we found that the last common ancestor (LCA) of this endosymbiont coded for at least 653 proteins. Similarly to genome reduction events in *Buchnera*, marked CDS losses (>=40) have independently occurred in many aphid lineages. These genomic reductions are almost always accompanied by a drastic loss of protein-coding genes (CDSs). One marked exception are both *Buchnera* strains from Hormaphidinae, which hold rather large genomes (630 and 580 kbp) with a strikingly low number of CDSs (453 and 450). Surprisingly, both of these genomes keep a small number of pseudogenes (10 and 1), revealing large portions of their genomes are ”deserted” (28.9% and 25.3%). The LCA of both *Buchnera* from Hormaphidinae is predicted to have had at least 472 CDs and a large genome, revealing small CDS losses leading to each of the Hormaphidini and Nipponaphidini branches.

Most notably, we found one gene unique to each of the *Buchnera* from Mindarinae and Thelaxinae, hinting at gene acquisition in these lineages. The first gene is Tn3-family DNA-resolvase/invertase in the pYqhA plasmid of *Mi. abietinus*. This gene is similar to *Escherichia coli*'s serine recombinase PinE (prom prophage origin), which catalyzes the inversion of an 1800-bp DNA fragment (the P region), which can exist in either orientation (Plasterk and van de Putte, 1985).The second gene is a predicted amidinotransferase (FN0238 type), with best hits against Bacteroidota, present in a *Thelaxes*-specific plasmid coding for this and a *repA1* protein, similar to that present in pLeu and pYqhA plasmids (van Ham *et al.*, 1997).

Secondary symbionts complement metabolic deficiencies in six aphid lineages

Given the well-established role of *Buchnera* as an essential amino acid- and vitamin B-provider, we analysed the genes coding for enzymes involved in the biosynthesis of these nutrients across *Buchnera* (figure 3 and supplementary figure S3, Supplementary Material online). We found that *Buchnera* strains belonging to seven aphid lineages have evolved genomes potentially unable to meet the nutritional requirements of their host, displaying losses of genes namely involved in the synthesis of EAAs and B vitamins. These seven lineages include the previously reported Lachninae, *Periphyllus* spp. (Chaitophorinae: Chaitophorini), and *Si. maydis* (Chaitophorinae: Siphini), independently sequenced in this study, as well as the newly identified members of Eriosomatinae, *Anoecia* spp. (Anoeciinae), *Dr. platanoidis* (Drepanosiphinae), *Ch. stipae* (Chaitophorinae: Siphini), and *Pa. juglandis* (Calaphidinae: Panaphidini). In all but one of the afore-mentioned aphid lineages, we were able to recover an additional symbiont genome: *Se. symbiotica* in *Periphyllus* spp., Siphini, and *Pa. juglandis*; *Fukatsuia* in *Anoecia* spp., and a

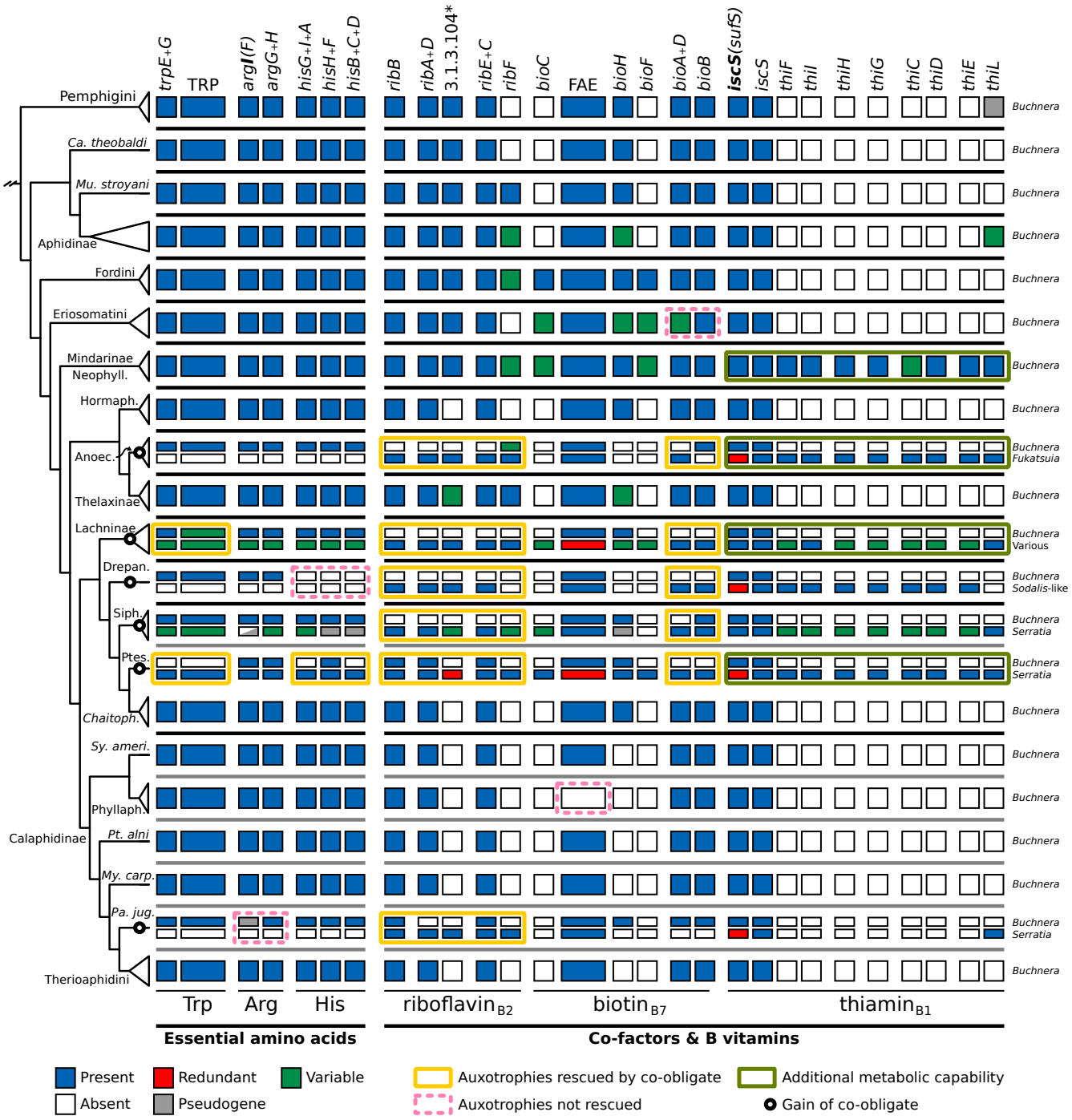

Figure 3. Metabolic complementarity of co-obligate symbiotic systems of aphids. Matrix summarising key metabolic capacities of obligate symbionts of different aphid species. On the left, dendrogram displaying phylogenetic relationships of *Buchnera* strains (see supplementary figure S1, Supplementary Material online). An empty circle at a node represent the acquisition of a co-obligate symbiont. At the leaves, abbreviation for aphid taxa. At the top of each box column, the name of the enzymes or pathway catalysing a reaction. At the bottom, the name of the compound synthesised by the enzymatic steps. At the right of each row of boxes, symbiont taxon name. Black and grey bars between boxes separate aphid subfamilies and tribes. TRP= anthranilate to tryptophan (*trpDCBA*), FAE= fatty acid elongation (*fab(B/F)GZ(I/V)*), Neophyll= Neophyllaphidinae, Hormaph.= Hormaphidinae, Anoec.= *Anoecia*, Drepan= *Dr. platanoidis*, Siph.= Siphini, Ptes.= *Pe. testudinaceus*, *Chaitoph.*= *Chaitophorus*, *Sy. ameri.*= *Sy. americanus*, Phyllaph.= Phyllaphidinae, *My. carp.*= *My. carpini*, *Pa. jug.*= *Pa. juglandis*. * EC number for reaction catalysed by promiscuous enzymes.

Sodalis-like strain in *Dr. platanoidis*. Genomes of all secondary symbionts showed several symptoms of long-term association with their hosts, exhibiting clear evidence of genome reduction in comparisson to their free living counterparts. Those symbionts associated with *Anoecia* spp., *Dr. platanoidis*, and *Pa. juglandis* show the most GC-biased and reduced genomes (supplementary table S2, Supplementary Material online). On the other hand, the *Se. symbiotica* genomes of both Siphini and *Periphyllus* were more similar in size to those of facultative strains and some co-obligate ones of Lachninae (i.e. that of *Cinara tujafilina*).

Manual curation of the genes involved in the biosynthesis of EAAs and B vitamins revealed these symbiont genomes generally bear genes that are capable of rescuing auxotrophies evolved in their *Buchnera* partners. In addition, in many cases, these bacterial genomes have also lost some of the functions that are assured by *Buchnera* (figure 3), generating a clear pattern of metabolic complementarities typical of dual symbiotic systems.

In regards to B vitamins, riboflavin- and biotin-biosynthesis is often observed to be lost in *Buchnera* and being taken over/rescued by the new co-obligate symbiont. The two exceptions are in the biotin biosynthesis in *Anoecia* spp. and *Pa. juglandis*. In the former, enzymes from both symbionts are needed: *bioA*+*bioD* from the *Fukatsuia*-related symbiont and *bioB* from *Buchnera*. In the latter, the new *Se. symbiotica* symbiont has lost all three genes needed to produce this compound from KAPA, and thus *Buchnera* has retained this role. As reported for other *Periphyllus* spp. by Monnin *et al.* (2020), we corroborate the tryptophan- and histidine-biosynthetic role of the *Se. symbiotica* endosymbiont in the newly sequenced genome from *Pe. testudinaceus*. Nonetheless, we did not observed the phenylalanine biosynthetic role being lost in *Buchnera* from *Pe. testudinaceus*. Upon closer inspection of the *Buchnera* draft genomes and annotations of *Periphyllus* spp. analysed by Monnin *et al.* (2020), we found the presence of a conserved poly(A) region across *Buchnera* from *Periphyllus* spp. in the *pheA* gene. Due to the conserved nature of this region, we considered it as likely rescued by transcriptional frameshifting, as this rescue mechanism has been experimentally demonstrated in other *Buchnera* and *Blochmannia* strains (Tamas *et al.*, 2008). In addition, similar poly(A) regions disrupting the *pheA* gene are present in all other sequenced *Buchnera* strains from Chaitophorinae. We found the thiamin-biosynthetic capacity (vitamin B_1) to be largely retained in the symbiotic consortia from Mindarinae, Neophyllaphidinae, Anoeciinae, Lachninae, Drepanosiphinae, Siphini, and *Pe. testudinaceus*. Most surprisingly, in the case of Mindarinae we found the retention of thiamin biosynthetic genes in newly sequenced *Buchnera* from Mindarinae (all) and Neophyllaphidinae (no *thiC*). The genes are present in four syntenic regions: between the *purH* and *rpoC* (*thiCEFSGH*), *dcd* and *metG* (*thiD*), *ribE* and *ribD* (*thiL*), and *dxs* and *yajR* (*thiI*). The location of the *thi* genes is syntenic with *Erwinia* spp., suggesting these genes were present in the LCA of *Buchnera* and have been repeatedly lost. Supporting this hypothesis, we observed the presence of a *thiL* gene/pseudogene in *Buchnera* from Pemphiginae and many Aphidinae, with top BLASTp hits *vs.* NCBI's nr to *Erwnia* and *Pantoea* bacteria, which are the closest free-living relatives of *Buchnera* (Husník *et al.*, 2011). Finally, *Buchnera* from Phyllaphidinae have lost the genes involved in fatty acid elongation, which would make

Buchnera dependant on the host for the production of its own membrane. However, as is the case for other strains, it would be able to synthesise biotin from KAPA.

Regarding essential amino acids, we corroborated the collaboration of both symbiotic partners in the production of tryptophan in some Lachninae (Lamelas *et al.*, 2011b; Manzano-Marín *et al.*, 2016) as well as the takeover of this role and the histidine biosynthesis by the co-obligate *Se. symbiotica* in *Periphyllus* spp. (Monnin *et al.*, 2020). In addition, we uncovered the complete loss of histidine-biosynthetic genes in *Buchnera* from Drepanosiphinae and a loss of the *argI* gene, rendering *Buchnera* and potentially the aphid host dependant on histidine and/or L-citruline for the production of these two essential amino acids.

Aphids are associated with a large and diverse pool of secondary symbionts

Most known co-obligate symbiotic lineages have facultative counterparts (Manzano-Marín *et al.*, 2018, 2020; Meseguer *et al.*, 2017), and thus, exploring the symbiotic microbiota from a large diverse aphid pool can provide clues into the available potential source for new co-obligate symbionts. Our 16S rRNA amplicon survey of 223 aphid samples yielded ∼5.4 million sequencing reads with an average of ∼11,300 reads per aphid sample kept. Those reads were distributed into 287 clusters, with most of these assigned to *Buchnera aphidicola* and nine known facultative or obligate symbionts of aphids (figure 4 and supplementary figure

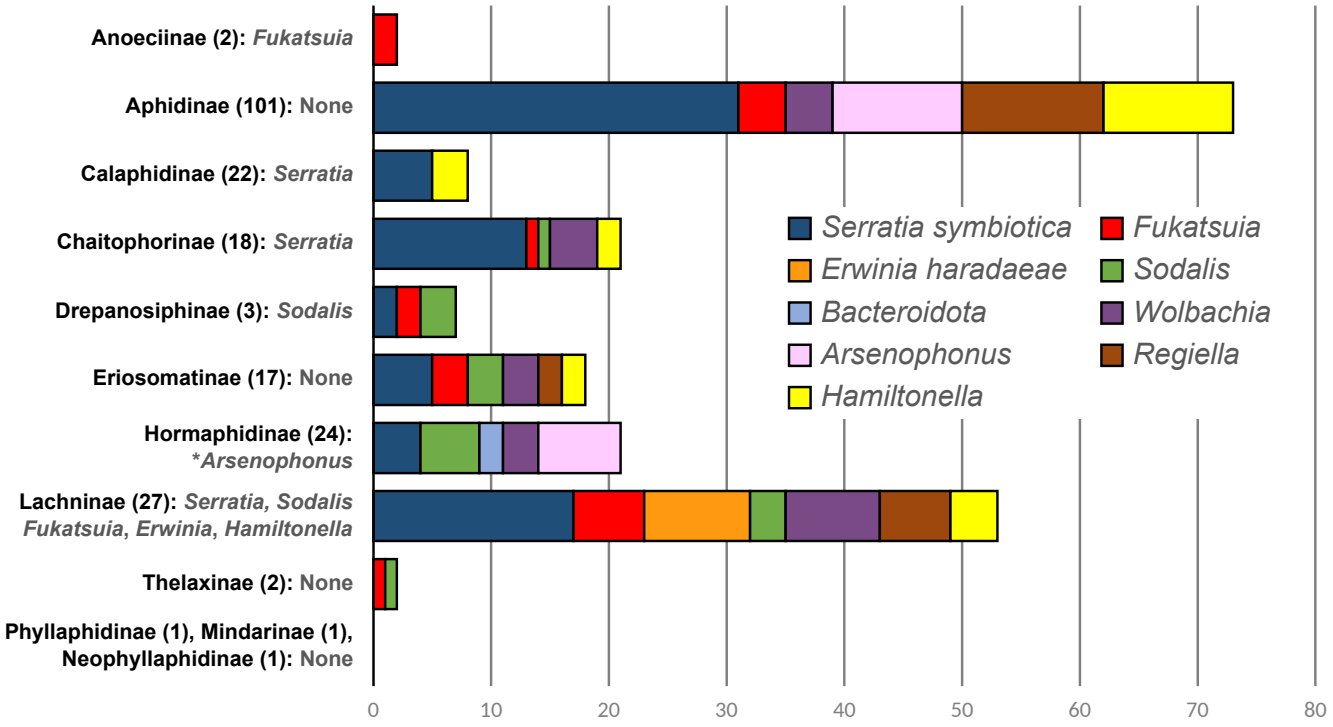

Figure 4. Secondary symbiont diversity across aphid subfamilies. Stacked barplot displaying absolute count of symbiont taxa found across samples of specific subfamilies. To the left of the bars, the subfamily name is indicated followed by the number of samples included in this survey in parentheses. Family names and counts are followed by the putative and known co-obligate symbiont taxa as inferred from the present study or previous genomic analysis studies. * Very recently published by Yorimoto *et al.* (2022)

S2, Supplementary Material online), including the recently described *Erwinia haradaeae* (Manzano-Marín *et al.*, 2020), *Fukatsuia* in some *Cinara* spp. (Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017), and *Bacteroidota* which has been found replacing *Buchnera* in *Geopemphigus* spp. (Chong and Moran, 2018). In addition, some clusters were designated as *Acetobacteraceae*- and *Gillamella*-related bacteria, known as gut-associated symbionts of diverse insects (Brown and Wernegreen, 2019; Holley *et al.*, 2022; Kwong and Moran, 2013; Pais *et al.*, 2018; Smith and Newton, 2020). The remaining reads were assigned to ubiquitous bacteria that could be of an environmental source and not representative of the aphid microbiota. In fact, these bacterial taxa were not always found across PCR replicates of the same sample and were sometimes found in the controls (supplementary figure S2, Supplementary Material online). From the bacterial taxa, *Se. symbiotica* was the most common secondary endosymbiont detected in our sampling, being present in 77 out of 223 samples and across seven subfamilies. In addition to *Se. symbiotica*; *Hamiltonella*, *Wolbachia*, *Fukatsuia*, and *Sodalis*-like bacteria were found in 23 (across five subfamilies), 22 (across five subfamilies), 18 (across seven subfamilies), and 15 samples (across six subfamilies), respectively.

This 16S amplicon survey is in agreement with our analysis of genomic data: when a co-obligate association has been inferred in an aphid species by genome-based metabolic inference, the symbiont taxon was found in all samples of that species and often in closely related ones. For example, we detected the presence of *Se. symbiotica* in all species of *Periphyllus* spp. (Chaitophorinae: Chaitophorini), where it has previously been reported as a co-obligate symbiont (Monnin *et al.*, 2020). This symbiont was also detected in Siphini from our sampling (Chaitophorinae: Siphini), as well as in various Lachninae species, where it is known as an obligate partner of many species (Lamelas *et al.*, 2011b; Manzano-Marín and Latorre, 2014; Manzano-Marín *et al.*, 2016, 2017; Meseguer *et al.*, 2017). The description of the dual symbiotic system of *Si. maydis* has been very recently independently reported by Renoz *et al.* (2022b), which resulted from parallel sequencing efforts and showing similar findings to ours for this system. *Erwinia* was also found in all samples from a specific clade of *Cinara* (Lachninae: Eulachnini), validating its co-obligate status (Manzano-Marín *et al.*, 2020). In addition, we found *Se. symbiotica*, *Fukatsuia* and a *Sodalis*-like bacterium in all samples of *Pa. juglanidis*, Anoeciinae and Drepanosiphinae aphids, respectively.

Our analysis also corroborated the absence of *Buchnera* in the one sample of *Cerataphis brasiliensis* (Hormaphidinae: Cerataphidini), a species in which the ancient obligate symbiont has likely been replaced by a ”Yeast-like symbiont” (Vogel and Moran, 2013). Further, although a *Bacteroidota* symbiont has been found replacing *Buchnera* in a member of the Fordini (*Geopemphigus* spp.; Chong and Moran, 2018), we did not recover any secondary bacterial 16S rRNA sequence belonging to this bacterial taxon in the sampled Fordini. In fact, *Bacteroidota* 16S rRNA sequences were only found in two samples of Hormaphidinae aphids in our survey (*Astegopteryx bambusae* ACOE3753 and *Reticulaphis mirabilis* ACOE3769). *Arsenophonus* was found in many samples of Hormaphidinae, where it has been previously found to be common within this subfamily and, very recently, even found to be establishing a co-obligate

association in *Ceratovacuna japonica* (Hormaphidinae: Cerataphidini; Yorimoto *et al.* 2022). Aphidinae hosted many secondary symbionts, however, none appear to be fixed at least at the tribe or even genus level.

Newly identified co-obligate symbionts have repeatedly evolved from within well-known symbiotic lineages

Phylogenetic placement of the newly sequenced *Serratia*-related co-obligate symbionts confirmed their taxonomic assignment as *Se. symbiotica* (figure 5A). The recovered phylogenetic tree suggests at least two independent origins of the co-obligate *Se. symbiotica* endosymbionts in *Periphyllus* spp., with support for a single origin of this symbiont lineage in the LCA of *Pe. aceris* and *Pe. acericola*. Additionally, the resulting phylogeny does not support a common origin for the co-obligate symbionts of Siphini aphids, suggesting repeated replacements within this tribe. Regarding the co-obligate symbionts of *Anoecia* spp., they are recovered as a sister group to *Fu. symbiotica*. The 16S identity of the *Fukatsuia*-like

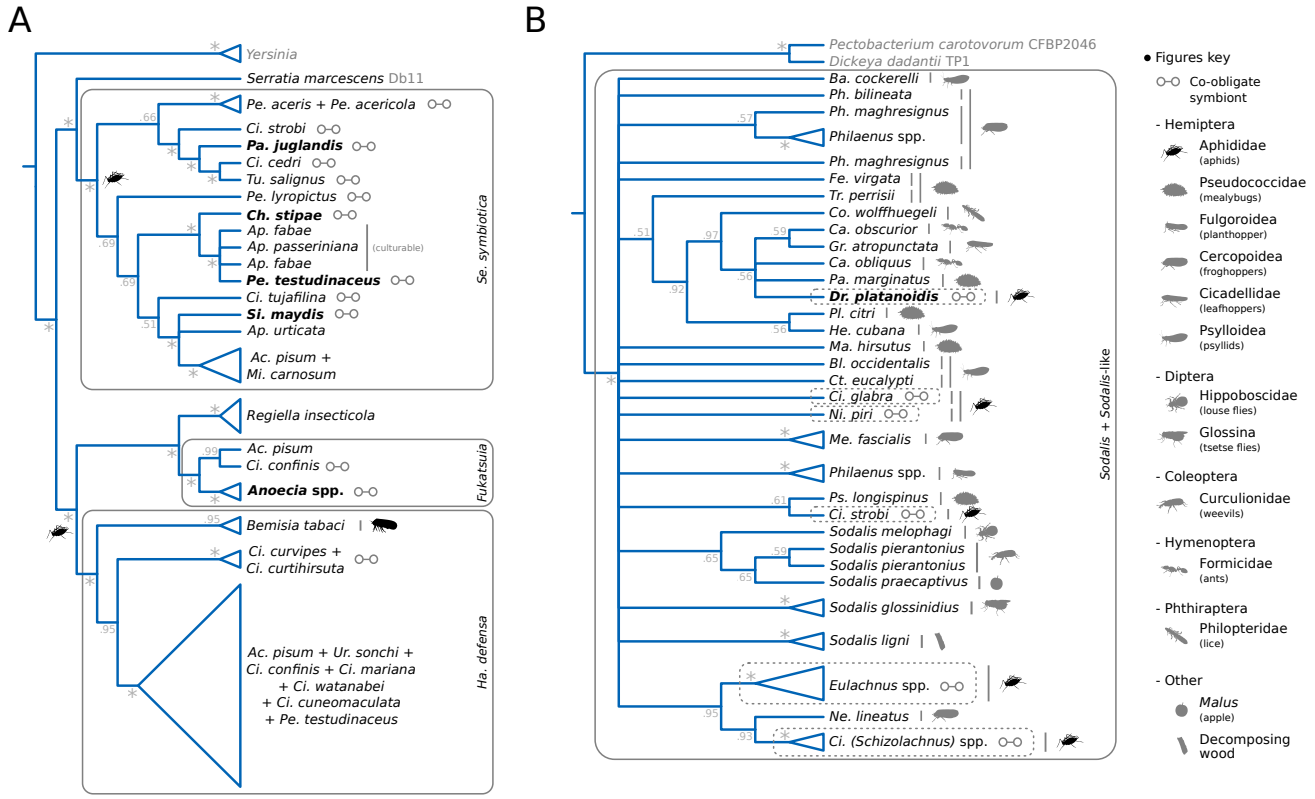

Figure 5. Phylogenetic relationships of aphid co-obligate endosymbionts. **(A)** Phylogenetic placement of *Yersinia*- and *Serratia*-related co-obligate endosymbionts. Silhouettes at nodes or at leaves indicate host relationship of clade. Grey boxes with binomial names on the side indicate aphid symbiotic taxa of interest. **(B)** Phylogenetic placement of *Sodalis*-related co-obligate endosymbionts. Grey box delimits the *Sodalis* + *Sodalis*-related clade. Silhouettes at branch ends indicate identity of host relationship of leaf/clade. Dotted boxes highlight *Sodalis*-like symbionts of aphids. For both A and B, abbreviated nomenclature is used when referring to the symbiont's host name. Names in bold, highlight newly sequenced endosymbiont lineages. Names in grey mark taxa used as outgroups.

co-obligate symbionts of Anoeciinae aphids falls below the recommended species threshold of 98.7% (95.79% and 96.04%; Chun *et al.*, 2018), but above the recommended minimum threshold for being classified as the same genus (Yarza *et al.*, 2014). Nonetheless, both *Anoecia* spp. symbionts show a 16S sequence identity of 98.77%, suggesting they belong to the same molecular species. Co-obligate *Sodalis*-like symbionts have been identified in several aphid species (Manzano-Marín *et al.*, 2017, 2018). The newly identified *Sodalis*-like co-obligate symbiont of *Dr. platanoidis* does not cluster with any of the previously identified aphid *Sodalis*-like symbionts (figure 5B). In fact, the tree suggests most co-obligate *Sodalis*-like endosymbionts from aphids have evolved independently, with those of *Eulachnus* and *Cinara (Schizolachnus)* spp. having likely evolved from closely related ancestor strains.

Co-obligate symbionts in *An. corni* and *Si. maydis* reside inside bacteriocyte cells

In order to investigate the tissue tropism of the newly identified co-obligate endosymbionts, we analysed available preserved specimens of *An. corni* and *Si. maydis* embryos through FISH miscroscopy using specific probes for each symbiont (figure 6). We corroborated and taxonomically identified the secondary symbionts previously observed by transmission electron microscopy by Michalik (2010); Michalik *et al.* (2014). The *Fukatsuia* symbiont of *An. corni* resides in separate bacteriocytes to those of *Buchnera* and shows a spherical pleomorphic shape. Similarly, we observed the *Se. symbiotica* symbiont of *Si. maydis*

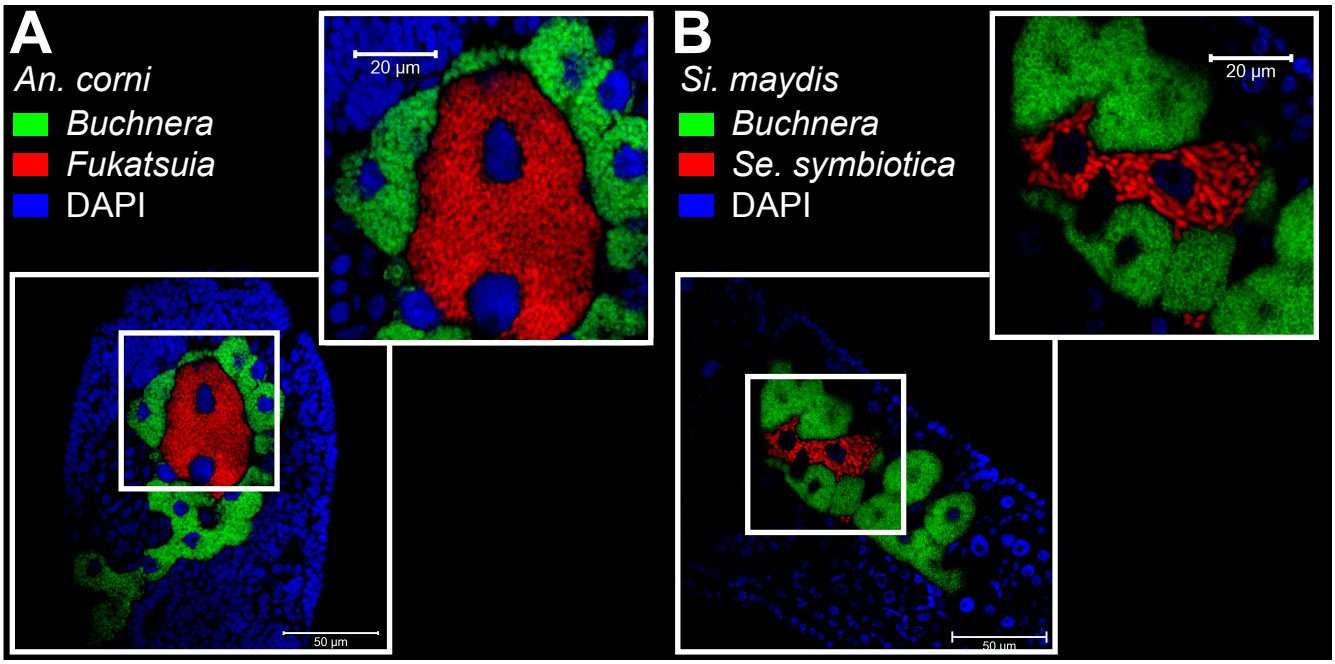

Figure 6. Location and morphology of co-obligate symbionts in selected aphid species. Merged FISH microscopic images of aphid embryos from **(A)** *An. corni* (lateral) **(B)** *Si. maydis* (tilted lateral). Co-obligate symbionts' signal is shown in red, *Buchnera*'s in green, and DAPI's (staining DNA, highlighting host nuclei) in blue. Thick white boxes indicate the magnified region depicted in the top-right of each panel. The scientific name for each species along with the false colour code for each fluorescent probe and its target taxon are shown at the top-left of each panel. Unmerged images can be found in https://doi.org/10.5281/zenodo.6394197.

inhabiting different bacteriocytes to those of *Buchnera*. However, the co-obligate endosymbiont shows an elongated rod-shape, more typical of facultative and early co-obligate symbionts (Manzano-Marín and Latorre, 2014; Moran *et al.*, 2005).

'*Candidatus* Fukatsuia anoeciicola' sp. nov.

'*Candidatus* Fukatsuia anoeciicola' (an.oe.ci.i'co.la. N.L. fem. n. *Anoecia*, an aphid genus; L. masc./fem. n. suff. -cola, inhabitant, dweller; N.L. fem. n. *anoeciicola*, inhabiting *Anoecia* aphids).

We propose the specific name '*Candidatus* Fukatsuia anoeciicola' for the monophyletic lineage of enterobacterial endosymbionts from the '*Candidatus* Fukatsuia' genus hitherto exclusively found affiliated as co-obligate nutritional symbionts in *Anoecia* species (Hemiptera: Aphididae: Anoeciinae). In embryos of *Anoecia corni*, '*Candidatus* Fukatsuia anoeciicola' is found co-inhabiting the bacteriome intracellularly in different bacteriocytes to those of *Buchnera aphidicola* (figure 6; Michalik *et al.*, 2014). In oviparous *An. corni*, '*Candidatus* Fukatsuia anoeciicola' has a spherical shape (mean cell diameter of 2.08μm), it is located intracellularly in bacteriocytes surrounded by those of *Buchnera*, and can also be found in oocytes (Michalik *et al.*, 2014).

Discussion

Long-term associations with strict vertically-transmitted bacteria often results in drastic genome reduction of the bacterial symbiont, which can sometimes lead to impairment of the symbiotic functions (*e.g.* host's nutrition) (McCutcheon *et al.*, 2019). In this context, replacement or complementation of the primary symbiont with a new one can maintain the stability of the symbiotic system. There are now many descriptions of symbiont replacement and complementation (Sudakaran *et al.*, 2017), in both sap feeding insects and, recently, in blood-feeding ticks (Buysse *et al.*, 2021). The nutritional association of *Buchnera aphidicola* with aphids has persisted for millions of years and has widely been regarded, until very recently, as mostly existing in one-to-one obligate associations with its host, in comparison to nutritional symbioses observed in other hemipterans. While both early and recent microscopic investigations into diverse aphids has revealed several cases of likely dual symbiotic systems (Buchner, 1953; Fukatsu, 2001; Fukatsu and Ishikawa, 1993; Michalik, 2010; Michalik *et al.*, 2014), little is known about the diversity and nature of these associations. Here we add to the recent evidence that the aphid-*Buchnera* evolutionary history is marked by repeated losses of essential metabolic capacities and complementation by a new co-obligate partner. We show that 12 aphid species (8 sequenced by us) from six subfamilies in which *Buchnera* has lost one or several essential symbiotic functions. This study, confirms recent observations in Chaitophorinae on two different genera, *Periphyllus* and *Sipha* (Monnin *et al.*, 2020; Renoz *et al.*, 2022b), the latter being published just months before the preparation of this manuscript. Our study shows that *Se. symbiotica* also occurs as a co-obligate nutritional symbiont in the related *Chatosiphella* genus. However, we show that *Buchnera* from the closely related *Chaitophorus* genus preserves an intact

repertoire to synthesise all EAAs and B vitamins.Along the same line, very recently published evidence strongly suggests that members of the *Ceratovacuna* genus (Hormaphidinae: cerataphidini) relay on two nutritional symbionts, *Buchnera* and *Arsenophonus* (Yorimoto *et al.*, 2022). According to our analyses, representative from the other two tribes of Hormaphidinae, do not depend on additional symbionts from nutritional complementation. Thus, in both Hormaphidinae and Chaitophorinae, losses of symbiotic functions happened after the diversification of the subfamily, and in the latter, after diversification of the tribe Chaitophorini. In most cases, *Buchnera*'s auxotrophies are rescued by a new bacterial partner, with the marked exceptions of two Eriosomatini (biotin), Drepanosiphinae (histidine), and Calaphidinae (Arginine), for which horizontal gene-transfer event to the host might play a role (see below for expanded discussion).

We found *Se. symbiotica* associated as the new co-obligate partner in now four subfamilies with at least four independent acquisitions (figure 7). However, as for Lachninae aphids (Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017), repeated replacement of co-obligate endosymbionts has likely occurred in *Periphyllus* and Siphini aphids: the *Se. symbiotica* strains from these aphids do not form a well-supported monophyletic clade (figure 5). In addition, repeated replacement has likely also occurred within Drepanosiphinae: while we found *Drepanosiphum* and *Drepanosiphoniella* hosted *Sodalis*-like symbionts, *Yamatocallis* spp. have been shown to harbour unrelated gammaproteobacterial obligate symbionts (Fukatsu, 2001). These patterns of repeated symbiont replacement would thus be similar to those observed in mealybugs (Husnik and McCutcheon, 2016), psyllids (Sloan and Moran, 2012), or Auchenorrhyncha (Bennett and Moran, 2013), where the primary obligate endosymbiont is retained, but the secondary co-obligate one shows less stability over evolutionary time.

Se. symbiotica was the most common bacterial lineage detected in our 16s rRNA survey, confirming its predominance as a natural facultative symbiont of aphids (Pons *et al.*, 2022; Zytynska and Weisser, 2016). In fact, *Se. symbiotica* has not only been recorded in aphids, but has also been identified as a possible facultative endosymbiont in Japanese populations of the hemlock woolly adelgid *Adelges tsugae* (von Dohlen *et al.*, 2013). Strains of this symbiont have been recorded as protecting against heat-stress (Chen *et al.*, 2000; Montllor *et al.*, 2002; Russell and Moran, 2006) and providing resistance against parasitoid wasps (Oliver *et al.*, 2003). In fact, strain IS has and even been shown to compensate, under laboratory conditions, for the elimination of the primary obligate endosymbiont *Buchnera* in the pea aphid *Acyrthosiphon pisum*, albeit for only 24 generations (Koga *et al.*, 2003) (suggesting some functional redundancy with *Buchnera*). Given the multiple benefits conferred by this endosymbiont, the potential for vertical transmission through the host plant of the aphid host (Pons *et al.*, 2019), and the capacity that some strains might have to take-over (at least) part of the symbiotic contributions of *Buchnera*, it is not unexpected to find it widespread across aphid taxa and populations, where it might easily become more common when the right environmental and biological conditions arise. The pervasiveness of *Se. symbiotica* strains in aphid populations has probably facilitated their adaptation as new co-obligate partners upon *Buchnera*'s gene losses. However, *Se. symbiotica* is not the only bacterial lineage that

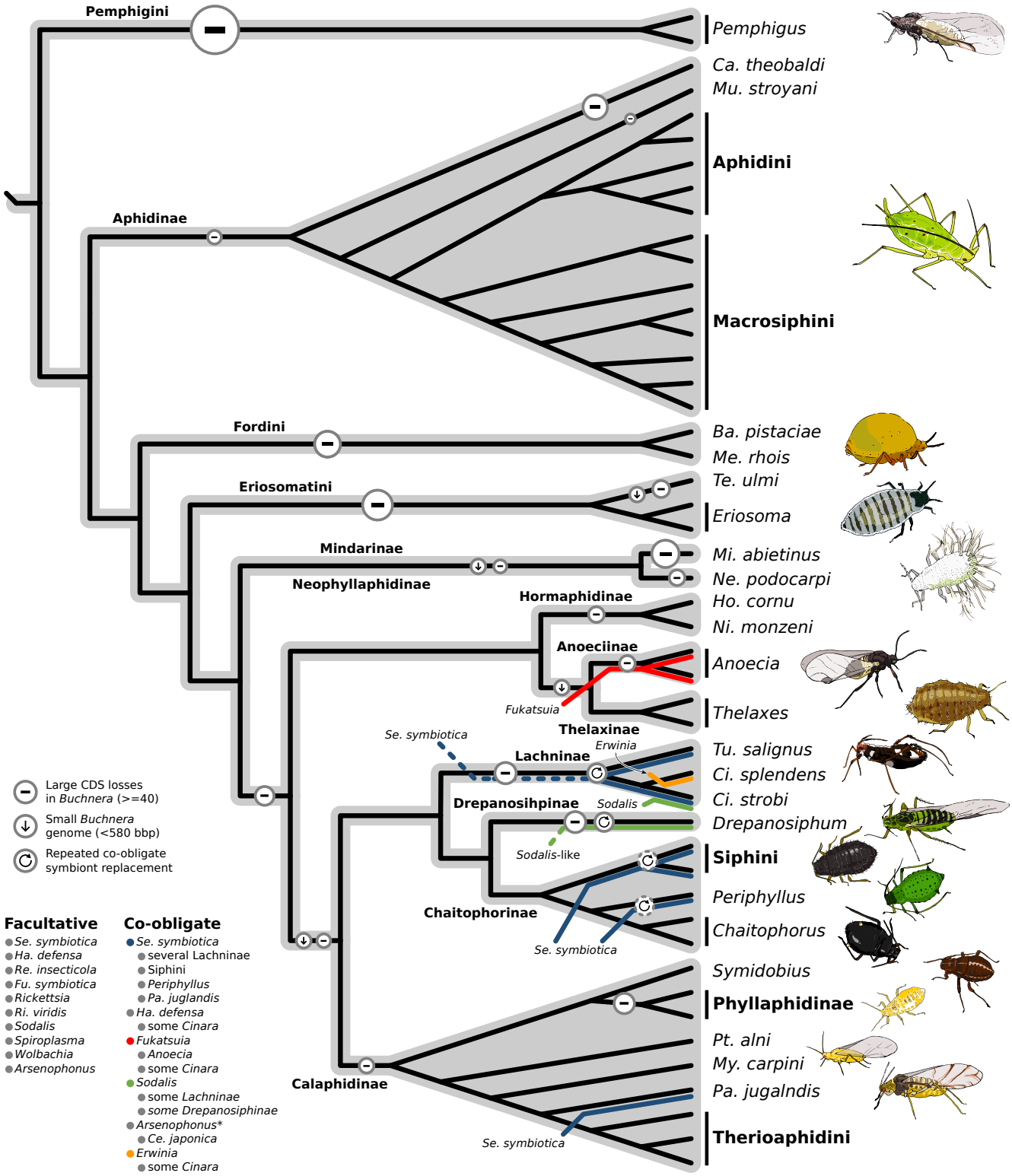

Figure 7. Evolutionary history of co-obligate symbioses in Aphididae. Diagram summarising results obtained from the genomic analysis presented in this work. Cladogram was made using the reconstructed phylogeny for *Buchnera* symbionts. Black lines in the cladogram represent *Buchnera*. Coloured lines are used to represent the acquisition of co-obligate symbionts. Dotted lines represent uncertainty in the acquisition of a specific taxa. Dotted circles represent uncertainty in symbiont replacement dynamics within the aphid lineage. On the right, aphid cartoons depicting variety of species across the Aphididae.

has become essential to its aphid hosts. A *Fukatsuia*-related symbiont (*Fukatsuia anoeciicola*) was revealed as a co-obligate partner of Anoeciine aphids. Another species of this symbiotic lineage, *Fukatsuia symbiotica*, has already been reported as a co-obligate endosymbiont of Lachninae (Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017). Our 16S rRNA amplicon survey revealed that *Fukatsuia*-related symbionts are found in populations of aphid species from different subfamilies, but much frequently than *Se. symbiotica*. *Fukatsuia* symbionts have been associated with several benefits to the pea aphid (i.e resistance against parasitoid wasps, fungal pathogens, and increased reproduction after heat-stress), however only when co-infecting with other symbiont species (Donald *et al.*, 2016; Doremus and Oliver, 2017; Heyworth and Ferrari, 2015). Additionally a *Sodalis*-like symbiont was identified as a co-obligate partner of Drepanosiphinae. Despite *Sodalis*-related strains being widely found across arthropods, they have been rarely reported in aphids. Some of the few strains found in aphids have been described as being in an obligate association with some Lachninae, mainly *Cinara* spp. (Manzano-Marín *et al.*, 2017, 2018; Meseguer *et al.*, 2017). We have nonetheless identified this symbiont occasionally in six subfamilies. With limited knowledge on *Sodalis* and *Fukatsuia* potential benefits/costs as facultative symbionts, it is hard to speculate on the environmental conditions favouring their expansion in populations, and thus, their adoption as a co-obligate partners. In any case, our data on *Sodalis*, along with the uncovering of *Erwinia haradaeae* as secondary co-obligate symbiont in some *Cinara* spp. and the *Bacteroidota* symbiont replacing *Buchnera* in *Geompemphigus* spp., shows that, even relatively rare symbionts (in extant populations) and environmental bacteria not known as symbionts of the host, can eventually become essential to it. This mirrors what has been observed in other sap- and blood-feeding lineages hosting dual symbiotic systems. For example, in whiteflies, a widespread but rare arthropod symbiont, *Arsenophonus*, has evolved to complement its degenerated *Portiera* symbiont. Conversely, phylogenetic placement of tick symbionts suggests they have evolved from common pathogens found in blood meals (Duron *et al.*, 2018). Hence, our study adds to the current evidence that obligate symbionts can derive not only from facultative beneficial symbionts but also sometimes from less common bacterial symbionts as well as strains that occur in the environment (Hosokawa *et al.*, 2016).

Among the very speciose Aphidinae subfamily, which shelters a large diversity of facultative symbionts (figure 4), and where 56% of the samples was positive for at least one facultative symbiotic taxon, we found no evidence of putative co-obligate endosymbionts, including in the banana aphid *Pentalonia nigronervosa*, corroborating previous genomic re-analyses (Manzano-Marín, 2020). As previously mentioned, nutritonal co-obligate associations have been suggested for two Aphidinae species (Monnin *et al.*, 2020). However, extensive review of this work casts important doubts regarding this conclusion (Manzano-Marín, 2020).

A common pattern in the dual nutritional symbioses revealed here is the riboflavin biosynthetic-role takeover by the new co-obligate symbiont. The biosynthetic genes of this pathway have been found to be particularly upregulated in aphid embryos (Bermingham *et al.*, 2009), which suggests the importance of this nutrient in the early life stages of the aphid but not in later ones. This could relax selection for the retention of these genes in *Buchnera* populations at later life stages of their hosts (i.e. during the

reproductive phase and transmission of *Buchnera* cells to the next generation), and thus open a window for facilitating the takeover of the nutritional role by a co-existing symbiont. Biotin was the second most common nutrient predicted to be supplied by the secondary co-obligate endosymbionts: it was found in all cases except in the co-obligate association observed in *Pa. juglandis*. A similar vitamin-biosynthetic role for the new co-obligate endosymbionts can be found in some psyllids (Sloan and Moran, 2012) as well as some adelgids (Dial *et al.*, 2022; Szabó *et al.*, 2022). In contrast, in other adelgids, retention of the B vitamin biosynthetic genes can also be observed in the primary symbionts, and not in the newly evolved symbiotic partners. In mealybugs, genes from these pathways can even be found missing in all symbiont genomes, while being found either partially (riboflavin) or entirely (biotin) in the hosts' genomes, originating from horizontal gene-transfer events (Husnik and McCutcheon, 2016). Therefore, this pattern seems to be system specific, and likely dependent on the background of the previously existing primary obligate endosymbiont. To our knowledge, no information is available on the specific life stage on which the aphid might depend on this nutrient; which precludes elaborating scenarios on how this pathway can be lost in *Buchnera*. In addition, thiamin was found to be retained in different lineages of co-obligate symbionts with a reduced genome, while lost in others as well as in all *Buchnera* genomes but the ones associated with Mindarinae and Neophyllaphidinae (figure 3). Consequently, complete and almost-complete pathways (lacking the thiamine-monophosphate kinase gene *thiL*) are found in obligate symbiotic systems of five subfamilies. The retention of this pathway in highly reduced genomes of endosymbionts suggest an important role for this nutritional role. This has been previously proposed in the endosymbiotic systems of a monophyletic group of *Cinara* spp., where the thiamin-biosynthetic genes have been serially horizontally transferred from a *Sodalis*-like bacteria to the now co-obligate *Er. haradaeae*, and further to a tertiary co-obligate *Ha. defensa* (Manzano-Marín *et al.*, 2020). Interestingly, the *Buchnera* strains that retain this pathway are associated exclusively with aphids feeding on conifers, as *Cinara* spp. Sap feeders are generally assumed to have very similar nutritional requirements, relying exclusively on a sugar rich diet that is deficient in EAAs and vitamins. However, based on the plant taxon they feed on, phloem sap might vary in their vitamin and amino acid abundance. Unfortunately, and to our knowledge, no comparative phloem sap analysis of conifers *vs.* angiosperms is available which could aid in inferring whether this symbiotic function can be related to host plant association. Further, other di-symbiotic systems maintaining thiamin genes analysed here are associated with more diverse botanical families (i.e. some herbaceus plants for *Si. maydis*). However, the presence of this pathway in the *Se. symbiotica* of *Si. maydis* (Siphini) can be related to a recent acquisition of this co-obligate strain: its genome size remains similar to that of facultative strains associated, for instance, with *Ac. pisum* (Burke and Moran, 2011; Nikoh *et al.*, 2019). Therefore, the retention of this pathway might not necessarily be related to any essential nutrient supplementation role. More studies elucidating the importance of thiamin during aphid development and its abundance in the phloem-sap of diverse host plant families are necessary to fully understand the evolutionary processes leading to its losses and gains throughout the diversification of aphids.

The amino acid-biosynthetic genes are almost always retained in *Buchnera*, while lost in the secondary

co-obligate symbionts (figure 3). Marked exceptions are those of the split pathway between the co-obligate symbionts of the Lachninae aphids *Cinara cedri* and *Tu. salignus* (Lamelas *et al.*, 2008; Manzano-Marín *et al.*, 2016) as well as the takeover of the biosynthesis of this compound plus histidine in *Periphyllus* spp. (Monnin *et al.*, 2020; Renoz *et al.*, 2022b). Most surprising was the identification of evolved auxotrophies in *Buchnera* that were not rescued by the co-obligate symbionts retrieved in our dataset. This was the case for arginine in *Pa. juglandis*, histidine in *Dr. platanoidis*, fatty acid elongation in Phyllaphidinae, and biotin in *Er. grossulariae* and *Te. ulmi*. We hypothesise that either the aphid is able to supplement precursors through horizontally transferred genes or the specific diet of the aphid is able to rescue these auxotrophies. In fact, horizontally transferred genes from diverse bacteria have been predicted and/or shown to support nutrient biosynthesis in Pseudococcinae (Husnik and McCutcheon, 2016; Husnik *et al.*, 2013; Szabó *et al.*, 2017), psyllids (Sloan *et al.*, 2014), and the whitefly *Bemisia tabaci* (Bao *et al.*, 2021; Luan *et al.*, 2015; Ren *et al.*, 2020; Xie *et al.*, 2018). In particular, certain genes whose product are involve in the biosynthesis of arginine and biotin have been shown to be acquired by diverse hemipterans through horizontal gene-transfer (HGT) and hypothesised to be involved in complementing their nutritional symbiont's biosynthetic pathways (Husnik and McCutcheon, 2016; Luan *et al.*, 2015). Another scenario would be the presence of an additional symbiont not detected by our genomic sequencing. In this vein, it is important to note that the presence of a *Fukatsuia* strain was detected in both *Dr. platanoidis* samples, by 16S rRNA amplicon sequencing (supplementary table S2, Supplementary Material online). However, in the case of both *Dr. platanoidis* and *Pa. juglandis*, microscopic studies have revealed the presence of only two different bacteria (Buchner, 1953). Additionally, a *Sodalis*-like strain was also detected in the related *Drepanosiphoniella fugans*, and while no deeper investigation of this genus/species was performed, it hints to a possible ancestral association with a *Sodalis*-like lineage. It is noteworthy to mention that both currently available sequenced strains of *Fukatsuia* symbionts do not preserve any genes in the histidine biosynthetic pathway (Patel *et al.*, 2019). Similarly, we observed a lack of these genes in the assembled genome for the *Fukatsuia* strain associated with *Dr. platanoidis*. Lastly, both *Sodalis*-like and *Se. symbiotica* symbionts of the aforementioned species show typical characteristics of well established obligate vertically transmitted symbionts: a lower G+C content and smaller genome size than their facultative/free-living symbiotic counterparts as well as clear metabolic complementarities with their corresponding symbiotic partner. Therefore, currently available evidence points towards these endosymbiotic systems being made up of *Buchnera* and one additional secondary co-obligate symbiont.

Our study also revealed patterns of *Buchnera* degradation: this primary symbiont has undergone multiple events of CDS losses, which accordingly is most commonly accompanied by a reduction in genome size. While the acquisition of co-obligate endosymbionts is often associated to a drastic gene loss (figure 7), there are also some exceptions to this pattern. In Calaphidinae and Phyllaphidinae, *Buchnera* are generally below 450 Mb (as small as in Lachninae) but this reduction was not concomitant to dependency on a new partner, as gene losses did not affect symbiotic functions. On the other hand, very few genes are actually lost in the branch leading to *Pa. juglandis* (18, out of which 8 are related to essential amino acid- and B vitamin-biosynthesis). A similar case is observed in the breanch leading to *Periphyllus* spp.

This shows that it is the inactivation of very few functional genes that actually triggers dependency on an additional symbiont and not necessarily large-scale genome decay. We also observed large genomes with a low coding density in Hormaphidinae aphids. Such large amounts of ”deserted” DNA are more commonly observed in some transitional symbiont genomes (Manzano-Marín and Latorre, 2016; Manzano-Marín *et al.*, 2018). To our knowledge, such low coding densities in such a small symbiont genome has only been reported for the primary obligate endosymbiont of mealybugs, *Tremblaya princeps* (McCutcheon and von Dohlen, 2011), as well as for several of their co-obligate *Sodalis*-related endosymbionts (Husnik and McCutcheon, 2016). This suggests that, although rarely observed, large gene losses can precede genome reduction, even in a small compact symbiont genome. More surprisingly, we found the recent acquisition of genes to *Buchnera* plasmids in *Mi. abietinus* and *Thelaxes* spp. The role these two newly gained genes might play are quite different. The DNA-resolvase/invertase found in the pYqhA plasmid of the *Buchnera* from *Mi. abietinus* might play a role in promoting inversions within this plasmid. On the other hand, the predicted amidinotransferase carried by *Buchnera* from *Thelaxes* spp. might play a yet unknown metabolic role. Regarding the DNA-resolvase/invertase, it keeps similarity to the PinE protein of *E. coli* which is responsible for a naturally occurring inversion of the so-called P-region (Plasterk and van de Putte, 1985). In fact, historical inversions are observable among pLeu plasmids of *Buchnera* from different aphid subfamilies (Gil *et al.*, 2006; Van Ham *et al.*, 2000). In addition, similar proteins have been found in the pBioThi plasmids of the co-obligate symbiont *Er. haradaeae*, which show a duplicated inverted region causing a duplication of biotin- and thiamin-biosynthetic genes (Manzano-Marín *et al.*, 2020). Therefore, these DNA-resolvase/invertases could play a role in rearrangement of plasmidic genes in different endosymbiotic lineages. Both of these novel genes were likely acquired through horizontal gene transfer (HGT) events, from a hitherto unknown origin. HGT events have been observed in the aphid endosymbiont *Er. haradaeae*, where biotin- and thiamin-biosynthetic genes have been likely horizontally transferred from a once co-existing symbiont related to *Sodalis* bacteria (Manzano-Marín *et al.*, 2020). Similar cases of HGT have been observed in other small vertically transmitted symbiont genomes. These include the aforementioned example of *Er. haradaeae*, the transfer of an antifungal lagriamide biosynthetic gene cluster in the *Burkholderia* symbiont of the beetle *Lagria villosa* (Waterworth *et al.*, 2020), and the acquisition of a biotin-biosynthetic gene cluster in *Wolbachia* (Driscoll *et al.*, 2020; Gerth and Bleidorn, 2017; Nikoh *et al.*, 2014), *Legionella polyplacis*, (Říhová *et al.*, 2017), and the *Midichloria* symbiont of the tick *Hyalomma marginatum* (Buysse *et al.*, 2021), among others. HGT events in *Buchnera* and other endosymbionts evidence that, despite a degraded machinery needed for homologous recombination and their highly compact and degraded genomes, these symbionts are still able to horizontally acquire genes from unrelated bacteria. While the route(s) facilitating this HGT events remain, to our knowledge, unknown, the co-infection of bacteriocyte cells by facultative bacteria (Fukatsu *et al.*, 2000) could represent an opportunity for such an event.

In conclusion, we found that co-obligate symbiotic systems are more widespread in aphids than previously thought, potentially existing in at least 11.4% of aphid species, given known species diversity in aphid lineages that are now inferred as hosting a dual nutritional symbiosis. One recent study corroborates

our finding of a co-obligate *Se. symbiotica* endosymbiont in *Si. maydis* (Renoz *et al.*, 2022b), and another one shows the evolution of a co-obligate symbiotic system in *Ceratovacuna japonica* (Hormaphidinae: Cerataphidini; Yorimoto *et al.*, 2022). These results, in light of the large-scale approach taken in this work, show that these multi-partner symbioses have evolved several times independently, even within subfamilies. A few major aphid lineages are still missing from this reconstruction, most markedly members of two well-diversified subfamily Greenideinae and Hormaphidinae where more dual symbiotic systems might be revealed. Our analyses also show that a few key gene losses (i.e. those in the riboflavin biosynthetic pathway), rather than large-scale ones, are at the onset of co-obligate symbioses and that genome reduction can be decoupled from new nutritional symbiont acquisitions. It has sometimes been suggested that it is the pervasiveness of facultative symbionts the one that relax selective pressure on genes from *Buchnera* and make it plunge further down the rabbit hole of endosymbiosis leading to co-dependency on a third partner (Lamelas *et al.*, 2011a). Our data on *Buchnera* genome degradation, co-obligate symbiont acquisition, and symbiont prevalence across subfamilies, gives support to a scenario where it is not the presence of a facultative symbiont that triggers nor accelerates *Buchnera* decay by relaxing selective constraints of this obligate symbiont. Given the ongoing *Buchnera* genome degradation, even in mono-symbiotic systems (reported here and in Chong *et al.*, 2019), it is likely that the decay of *Buchnera* genomes is inherent to its lifestyle, and when it affects gene loss and enzymatic activity, it opens a niche for the reliance on a new symbiont: likely any bacterial lineage that is present and potentially capable of filling that niche.

Finally, the role of additional or lost metabolic capabilities (such as the reacquisition of the thiamin biosynthetic pathway) in aphid symbiotic systems remains to be explored. We expect that comparative analyses into ingested phloem by specific aphid species as well as assessment of the specific dependence on nutrients by aphids at different life stages will shed light on whether these gain/losses of metabolic potential can facilitate the establishment of new obligate symbionts.

Materials and Methods

Aphid collection and sequencing

Twenty-five different species of aphids were sourced from the *CBGP - Continental Arthropod Collection* (Centre de Biologie pour la Gestion des Populations, 2018; supplementary table S3, Supplementary Material online), where the specimens were preserved in ethanol 70% at 6°C until extraction. In this collection, a specimen (with its unique voucher) corresponds to individuals from a single aphid colony. Bacteria-enriched DNA extractions were performed following Jousselin *et al.* (2016). When possible, 10-15 individuals were used for extraction.

For genomic sequencing, extracted DNA was used to prepare two custom paired-end libraries in Genoscope. Briefly, 5ng of genomic DNA were sonicated using the E220 Covaris instrument (Covaris, USA). Fragments were end-repaired, 3'-adenylated, and NEXTflex PCR free barcodes adapters (Bioo

Scientific, USA) were added by using NEBNext Ultra II DNA library prep kit for Illumina (New England Biolabs, USA). Ligation products were purified by Ampure XP (Beckman Coulter, USA) and DNA fragments (>200bp) were PCR- amplified (2 PCR reactions, 12 cycles) using Illumina adapter-specific primers and NEBNext Ultra II Q5 Master Mix (NEB). After library profile analysis by Agilent 2100 Bioanalyser (Agilent Technologies, USA) and qPCR quantification using the KAPA Library Quantification Kit for Illumina Libraries (Kapa Biosystems, USA), the libraries were sequenced using 251 bp paired-end reads chemistry on a HiSeq2500 Illumina sequencer. In addition, Nanopore sequencing was done for the aphid species *Anoecia corni*. For long-read sequencing, the Rapid Low Input by PCR Sequencing Kit (SQK_RLI001) was used. The library was sequenced on an R9.4 flow cell. In addition to the newly produced sequencing reads, data for an additional 19 species was downloaded from the NCBI's Short Read Archive (supplementary table S4, Supplementary Material online), with a selection based on two criteria: sequencing data quality and phylogenetic coverage of aphid subfamilies. We therefore did not include all of the previously reported symbiont genomes of Lachninae, as the co-obligate associations have been extensively investigated in previous studies, nor Aphidinae, which so far have only been shown to host *Buchnera* as an obligate nutritional symbiont.

For amplicon sequencing, we investigated symbiont diversity in 223 aphid samples, comprised of 147 species of 75 genera belonging to 12 subfamilies sourced from the aforementioned Aphididae collection. These species included those used for symbiont genome data as well as species closely related to them. DNA was extracted from single individuals and a 251bp portion of the V4 region of the 16S rRNA gene was chose for amplification following Mizrahi-Man *et al.* (2013). The 16S fragment was amplified using primers 16S-V4F (5'-GTGCCAGCMGCCGCGGTAA-3') and 16S-V4R (5'-GGACTACHVGGGTWTCTAAT-3') following the dual-index sequencing strategy developed by Kozich *et al.* (2013) and the protocol described in Jousselin *et al.* (2016). Each DNA extract was amplified twice along with negative controls (DNA extraction and PCR controls) using distinct 96-well microplates for PCR replicates. We obtained a total of 485 PCR products, which were pooled together and subjected to gel electrophoresis. The bands corresponding to the PCR products were excised from the gel, purified with a PCR clean-up and gel extraction kit (Macherey-Nagel) and quantified with the Kapa Library Quantification Kit (Kapa Biosystems). The DNA pool was then paired-end sequenced on an Illumina MiSeq flowcell with a 500-cycle Reagent Kit v2 (Illumina).

16S amplicon sequencing and analysis

16S amplicon sequences were first filtered through Illumina's quality control procedure. We then used a pre-processing script from Sow *et al.* (2019) which uses FLASH v1.2.11 (Magoč and Salzberg, 2011) and CUTADAPT v1.9.1 (Martin, 2011) to merge paired sequences into contigs and trim primers, respectively. We then used the FROGS pipeline (Escudié *et al.*, 2018) to generate an abundance table of symbiont lineages across samples. Briefly, to generate the tables, we first filtered out sequences >261 and <241 bp. We then clustered variants with Swarm (Mahé *et al.*, 2014) using a maximum aggregation distance of 1.

Lastly, we identified and removed chimeric variants with VSEARCH (Rognes *et al.*, 2016).

Taxonomic assignments of clusters was carried out using RDPtools v2.0.3 (https://github.com/rdpstaff-/RDPTools, last accessed July 18, 2022; Cole *et al.*, 2014) and BLASTn+ (Camacho *et al.*, 2009) against the Silva database release 138 (Quast *et al.*, 2013) as implemented in FROGS. Following taxonomic affiliation, we aggregated clusters when they shared the same taxonomy with at least 98% of identity (FROGS' affiliation_postprocess step). From the abundance table of clusters across samples, we transformed read numbers per aphid samples into percentages and sequences accounting for <0.5% of all the reads for a given sample were excluded using an R script following Jousselin *et al.* (2016). Clusters were kept only if present in both PCR replicates of the same sample. For final description of endosymbiont diversity we only kept the PCR replicate that yielded the highest number of reads. We refined ambiguous taxonomic assignations using BLASTn+ against 16S rRNA gene sequences extracted from whole endosymbiont genomes from this study and retrieved from GenBank. We also used the webtool leBIBI IV (https://umr5558-proka.univ-lyon1.fr/lebibi/lebibi.cgi, last accessed July 18, 2022) that provides automatic phylogenetic placement of bacterial 16S sequences (Flandrois *et al.*, 2015). The resulting relative abundance table (supplementary table S5, Supplementary Material online) was used to produced a heatmap in R of presence/absence of endosymbiotic taxa across aphid subfamilies.

Genome assembly

For all Illumina datasets, reads were right-tail clipped (using a minimum quality threshold of 20) using FASTX-Toolkit v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/, last accessed July 18, 2022) and those shorter than 75 after were dropped. PRINSEQ v0.20.4 (Schmieder and Edwards, 2011) was used to remove reads containing undefined nucleotides as well as those left without a pair after the filtering and clipping process. Clean reads were assembled using SPAdes v3.11.1 (Bankevich *et al.*, 2012) with the --only-assembler option and k-mer sizes of 33, 55, 77, 99, and 127 (depending on the read length). Assembled contigs shorter than 200 bps were dropped. The surviving contigs were binned using results from a BLASTx (Altschul *et al.*, 1997) search (best hit per contig) against a database consisting of the proteomes of the Pea aphid and a selection of aphid's symbiotic and expected free-living bacteria (supplementary table S6, Supplementary Material online). When no genome was available for a certain lineage, closely related bacteria were used.

The binned contigs were manually screened using the BLASTx web server (*vs.* the nr database) to insure correct taxonomic assignment. The resulting contigs were then used as reference for read-mapping and individual genome re-assembly using SPAdes, as described above. Binned and re-assembled contigs were checked for circularisation and completion through assembly graphs. For those *Buchnera* strains that have evolved a low-complexity replication start, we closed the ends with "N" stretches after checking for the conserved proteins around the end and truncation of low-complexity sequencing artefacts. Details of any sequence modification post-assembly is captured in the GenBank-formatted annotation files.

Genome annotation

The resulting *Buchnera* genomes were annotated as follows. First, open reading frame (ORF) prediction was done using Prokka v1.14.6 (Seemann, 2014). This ORF prediction was followed by non-coding RNA prediction with infernal v1.1.2 (Nawrocki and Eddy, 2013) (against the Rfam v14.1 database; Kalvari *et al.*, 2021), tRNAscan-SE v2.0.9 (Chan *et al.*, 2021), and ARAGORN v1.2.36 (Laslett, 2004). This annotation was followed by careful manual curation of the genes on UGENE v37.1 (Okonechnikov *et al.*, 2012) through on-line BLASTx searches of the intergenic regions as well as through BLASTp and DELTA-BLAST (Boratyn *et al.*, 2012) searches of the predicted ORFs against NCBI's nr database. The resulting coding sequences (CDSs) were considered to be putatively functional if all essential domains for the function were found or if a literature search supported the truncated version of the protein as functional in a related organism, or if the the CDS displayed truncations but retained identifiable domains (details of the literature captured in the annotation file). Pseudogenes were further searched for based on synteny with closely related *Buchnera* strains on which previous manual curation was done. This prediction was performed based on nucleotide sequence using a combination of sequence alignment (with m-coffee; Wallace et al. 2006) and BLASTx searches against the NCBI's nr database. This last check allowed for the identification of pseudogenes missed by the previous searches. Once a *Buchnera* was manually curated, it was used as a Prokka reference protein set for the next *Buchnera* in order to maintain naming congruent across genomes. Following experimental evidence presented in Tamas *et al.* (2008), genes interrupted by a frameshift in a low complexity A- or T-homopolymeric region, were annotated as coding for a functional protein, with special attention to conservation of frameshifted region and amino acid sequence within closely related *Buchnera* strains.

For the genomes of co-obligate symbionts, draft Prokka annotations were performed and genes of interest to metabolic complementarity for nutrient provisioning underwent a manual curation as described above. For genome statistics, an estimated count for pseudogene features was done as follows. Contiguous proteins detected as truncated and with same assignment as well as proteins under 100 amino acids, were considered pseudogenes. tRNAs, ncRNAs, and rRNAs were predicted as described above.

Phylogenetics

In order to reconstruct a phylogenetic hypothesis for *Buchnera*, we used OrthoMCL v2.0.9 (Chen *et al.*, 2007; Li, 2003) to build clusters of orthologous proteins. To the newly acquired, reassembled, and re-annotated symbiont genomes; we added three already annotated ones from Lachninae (supplementary table S4, Supplementary Material online). These orthologous groups were then manually curated using the annotation to keep same genes together in clusters. Count v10.04 (Csuos, 2010)was used to calculate the most parsimonious scenario for gene losses in *Buchnera*. *Escherichia coli* K-12 MG 1655 and a selection of *Pantoea* and *Eriwnia* strains were used as outgroups (supplementary table S7, Supplementary Material online). We then retrieved the single copy-core proteins of the selected genomes for phylogenetic reconstruction. We aligned the single-copy core protein sets, gene by gene, using MAFFT v7.453 (Katoh

and Standley, 2013) (L-INS-i algorithm). Divergent and ambiguously aligned blocks were removed using Gblocks v0.91b (Talavera and Castresana, 2007). The resulting alignments were concatenated for phylogenetic inference. Bayesian inference was performed in MrBayes v3.2.7a (Ronquist *et al.*, 2012) running two independent analyses with four chains each for up to 300,000 generations and checked for convergence with a burn-in of 25%. JModelTest v2.1.10 (Darriba *et al.*, 2012) was used to select the best model for phylogenetic reconstruction based on the Akaike's Information Criterion (cpREV+I+G4).

To analyse the phylogenetic relations of newly sequenced aphid co-obligate endosymbionts, we built two phylogenies. We first reconstructed a phylogeny for *Yersinia*- and *Serratia*-related endosymbionts. For this, we used the dataset built by (Rouïl *et al.*, 2020) and added data for *Serratia symbiotica* symbionts and the newly sequenced symbiotic strains. Given that many *Se. symbiotica* genomes lacked an intact *hrpA*, this gene was excluded from our dataset. Each gene set was individually aligned using MUSCLE v3.8.31 (Edgar, 2004). Then, we removed divergent and ambiguously aligned blocks with Gblocks. Bayesian inference was performed in MrBayes, using the GTR+I+G4 substitution model running two independent analyses with four chains each for 1,000,000 generations and checked for convergence. The second phylogeny was reconstructed for *Sodalis*-like endosymbionts. Given that many of these endosymbionts remain without a sequenced genome, we used 16S rRNA gene sequences following (Manzano-Marín *et al.*, 2017). We used SSU-ALIGN v0.1 (Nawrocki, 2009) to align the rRNA gene sequences. GBlocks v0.91b was then used to eliminate poorly aligned positions and divergent regions with the option '-b5=h' to allow half of the positions with a gap. MrBayes v3.2.7 was used for phylogenetic inference under the GTR+I+G4 model running two independent analyses with four chains each for up to 10,000,000 generations and checked for convergence, discarding the first 25% as burn-in. Visualization and tree-editing for all analyses was done in FigTree v1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/, last accessed July 18, 2022).

Fluoresence *in situ* hybridisation microscopy

Live aphid specimens from *Sipha maydis* and *Anoecia corni* were fixed in modified Carnoy's fixative (6 chloroform: 3 absolute ethanol: 1 glacial acetic acid) and left overnight, following the protocol of (Koga *et al.*, 2009). Individuals were then dissected in absolute ethanol to extract embryos and transferred into a 6% solution of H_2O_2 diluted in absolute ethanol and were then left in this solution for two weeks (changing the solution every three days). Embryos were then washed twice with absolute ethanol. Hybridization was performed overnight at 28°C in standard hybridisation buffer (20 mM Tris-HCl [pH 8.0], 0.9M NaCl, 0.01% SDS, and 30% formamide) and then washed (20 mM Tris-HCl [pH 8.0], 5 mM EDTA, 0.1M NaCl, and 0.01% SDS) before slide preparation. Competitive probes were adapted for this specific symbionts from (Manzano-Marín *et al.*, 2017) (supplementary table S8, Supplementary Material online).

All files relating to orthologous protein grouping as well as phylogenetic reconstructions can be found at https://doi.org/10.5281/zenodo.6394197.

Supplementary Material & Data Availability

Supplementary figures S1-S3 and tables S1-8 and have been included in this submission. All auxiliary files for phylogenetics and other analyses, unmerged FISH microscopy images, as well as the genomes of *Buchnera* and co-obligate endosymbionts are available online at https://doi.org/10.5281/zenodo.6394197. Newly sequenced and annotated genomes are in the process of being submitted for accessioning to the European Nucleotide Archive (ENA).

Acknowledgements

We would like to acknowledge the talented artist/scientist Jorge Mariano Collantes Alegre for the aphid cartoons in figure 7. We would also like to thank Prof. Aharon Oren for his help and advice in bacterial species nomenclature. This work was supported by the Marie-Curie AgreenSkills+ fellowship programme co-funded by the EU's Seventh Framework Programme (FP7-609398) to A.M.M. The Marie-Skłodowska-Curie H2020 Programme (H2020-MSCA-IF-2016) to E.J., the Agropolis foundation/Labex Agro ("Cinara's microbiome") to E.J, the France Génomique National Infrastructure, funded as part of the *Investissement d'Avenir* program managed by the *Agence Nationale pour la Recherche* (ANR-10-INBS-0009) to E.J., C.C., and V.B. This publication has been written with the support of the AgreenSkills + fellowship programme which has received funding from the EU's Seventh Framework Programme under grant agreement No. FP7-609398 (AgreenSkills + contract) and the Horizon 2020 Programme under grant agreement 746189 (H2020-MSCA-IF-2016-MicroPhan). We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) for providing help and/or computing and/or storage resources. The authors are grateful to the CBGP-HPC computational platform and the staff maintaining the *Life Science Compute Cluster* (LiSC; http://cube.univie.ac.at/lisc) at the University of Vienna which were used for computational analysis. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Altschul S. F, Madden T. L, Schäffer A. A, Zhang J, Zhang Z, Miller W, and Lipman D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17): 3389–3402. **URL** https://dx.doi.org/10.1093/nar/25.17.3389.

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

Bao X.-Y, Yan J.-Y, Yao Y.-L, Wang Y.-B, Visendi P, Seal S, and Luan J.-B. 2021. Lysine provisioning by horizontally acquired genes promotes mutual dependence between whitefly and two intracellular symbionts. *PLOS Pathogens*, 17(11): e1010120. **URL** https://dx.doi.org/10.1371/journal.ppat.1010120.

Baumann P. 2005. Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annual Review of Microbiology*, 59(1): 155–189. **URL** https://dx.doi.org/10.1146/annurev.micro.59.030804.121041.

Baumann P, Baumann L, Lai C.-Y, Rouhbakhsh D, Moran N. A, and Clark M. A. 1995. Genetics, physiology, and evolutionary relationships of the genus *Buchnera*: Intracellular symbionts of aphids. *Annual Review of Microbiology*, 49(1): 55–94. **URL** https://dx.doi.org/10.1146/annurev.mi.49.100195.000415.

Bennett G. M and Moran N. A. 2013. Small, smaller, smallest: The origins and evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biology and Evolution*, 5(9): 1675–1688. **URL** https://dx.doi.org/10.1093/gbe/evt118.

Bermingham J, Rabatel A, Calevro F, Viñuelas J, Febvay G, Charles H, Douglas A, and Wilkinson T. 2009. Impact of host developmental age on the transcriptome of the symbiotic bacterium *Buchnera aphidicola* in the Pea aphid (*Acyrthosiphon pisum*). *Applied and Environmental Microbiology*, 75(22): 7294–7297. **URL** https://dx.doi.org/10.1128/AEM.01472-09.

Blow F, Bueno E, Clark N, Zhu D. T, Chung S. H, Güllert S, Schmitz R. A, and Douglas A. E. 2020. B-vitamin nutrition in the pea aphid-*Buchnera* symbiosis. *Journal of Insect Physiology*, 126: 104092. **URL** https://dx.doi.org/10.1016/j.jinsphys.2020.104092.

Boratyn G. M, Schäffer A. A, Agarwala R, Altschul S. F, Lipman D. J, and Madden T. L. 2012. Domain enhanced lookup time accelerated BLAST. *Biology Direct*, 7(1): 12. **URL** https://dx.doi.org/10.1186/1745-6150-7-12.

Brown B. P and Wernegreen J. J. 2019. Genomic erosion and extensive horizontal gene transfer in gut-associated *Acetobacteraceae*. *BMC Genomics*, 20(1): 472. **URL** https://dx.doi.org/10.1186/s12864-019-5844-5.

Buchner P. 1953. *Endosymbiose der tiere mit pflanzlichen mikroorganismen*. Birkhäuser Basel, Basel, 1 edition. **URL** https://dx.doi.org/10.1007/978-3-0348-6958-4.

Burke G. R and Moran N. A. 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biology and Evolution*, 3(0): 195–208. **URL** https://dx.doi.org/10.1093/gbe/evr002.

Buysse M, Floriano A. M, Gottlieb Y, Nardi T, Comandatore F, Olivieri E, Giannetto A, Palomar A. M, Makepeace B. L, Bazzocchi C, Cafiso A, Sassera D, and Duron O. 2021. A dual endosymbiosis supports nutritional adaptation to hematophagy in the invasive tick *Hyalomma marginatum*. *eLife*, 10. **URL** https://dx.doi.org/10.7554/eLife.72747.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and Madden T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, 10(1): 421. **URL** https://dx.doi.org/10.1186/1471-2105-10-421.

Centre de Biologie pour la Gestion des Populations. 2018. CBGP - Continental arthropod collection. **URL** https://dx.doi.org/10.15454/D6XAKL.

Chan P. P, Lin B. Y, Mak A. J, and Lowe T. M. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Research*, 49(16): 9077–9096. **URL** https://dx.doi.org/10.1093/nar/gkab688.

Chen D.-Q, Montllor C. B, and Purcell A. H. 2000. Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, *Acyrthosiphon pisum*, and the blue alfalfa aphid, *A. kondo*i. *Entomologia Experimentalis et Applicata*, 95(3): 315–323. **URL** https://dx.doi.org/10.1046/j.1570-7458.2000.00670.x.

Chen F, Mackey A. J, Vermunt J. K, and Roos D. S. 2007. Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS ONE*, 2(4): e383. **URL** https://dx.doi.org/10.1371/journal.pone.0000383.

Chen J, Wang Y, Jiang L, and Qiao G. 2017. Mitochondrial genome sequences effectively reveal deep branching events in aphids (Insecta: Hemiptera: Aphididae). *Zoologica Scripta*, 46(6): 706–717. **URL** https://dx.doi.org/10.1111/zsc.12241.

Chong R. A and Moran N. A. 2018. Evolutionary loss and replacement of *Buchnera*, the obligate endosymbiont of aphids. *The ISME*

Journal, 12(3): 898–908. **URL** https://dx.doi.org/10.1038/s41396-017-0024-6.

Chong R. A, Park H, and Moran N. A. 2019. Genome evolution of the obligate endosymbiont *Buchnera aphidicola*. *Molecular Biology and Evolution*, 36(7): 1481–1489. **URL** https://dx.doi.org/10.1093/molbev/msz082.

Chun J, Oren A, Ventosa A, Christensen H, Arahal D. R, da Costa M. S, Rooney A. P, Yi H, Xu X.-W, De Meyer S, and Trujillo M. E. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 68(1): 461–466. **URL** https://dx.doi.org/10.1099/ijsem.0.002516.

Cole J. R, Wang Q, Fish J. A, Chai B, McGarrell D. M, Sun Y, Brown C. T, Porras-Alfaro A, Kuske C. R, and Tiedje J. M. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic acids research*, 42(Database issue): D633–42. **URL** https://dx.doi.org/10.1093/nar/gkt1244.

Csuos M. 2010. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics*, 26(15): 1910–1912. **URL** https://dx.doi.org/10.1093/bioinformatics/btq315.

Darriba D, Taboada G. L, Doallo R, and Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9(8): 772–772. **URL** https://dx.doi.org/10.1038/nmeth.2109.

Dial D. T, Weglarz K. M, Aremu A. O, Havill N. P, Pearson T. A, Burke G. R, and von Dohlen C. D. 2022. Transitional genomes and nutritional role reversals identified for dual symbionts of adelgids (Aphidoidea: Adelgidae). *The ISME Journal*, 16(3): 642–654. **URL** https://dx.doi.org/10.1038/s41396-021-01102-w.

Donald K. J, Clarke H. V, Mitchell C, Cornwell R. M, Hubbard S. F, and Karley A. J. 2016. Protection of pea aphids associated with coinfecting bacterial symbionts persists during superparasitism by a braconid wasp. *Microbial Ecology*, 71(1): 1–4. **URL** https://dx.doi.org/10.1007/s00248-015-0690-8.

Doremus M. R and Oliver K. M. 2017. Aphid heritable symbiont exploits defensive mutualism. *Applied and Environmental Microbiology*, 83(8): e03276–16. **URL** https://dx.doi.org/10.1128/AEM.03276-16.

Douglas A. E. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria Buchnera. *Annu Rev Entomol*, 43(1): 17–37. **URL** https://dx.doi.org/10.1146/annurev.ento.43.1.17.

Douglas A. E. 2006. Phloem-sap feeding by animals: problems and solutions. *Journal of Experimental Botany*, 57(4): 747–754. **URL** https://dx.doi.org/10.1093/jxb/erj067.

Driscoll T. P, Verhoeve V. I, Brockway C, Shrewsberry D. L, Plumer M, Sevdalis S. E, Beckmann J. F, Krueger L. M, Macaluso K. R, Azad A. F, and Gillespie J. J. 2020. Evolution of *Wolbachia* mutualism and reproductive parasitism: Insight from two novel strains that co-infect cat fleas. *PeerJ*, 8: e10646. **URL** https://dx.doi.org/10.7717/PEERJ.10646.

Duron O, Morel O, Noël V, Buysse M, Binetruy F, Lancelot R, Loire E, Ménard C, Bouchez O, Vavre F, and Vial L. 2018. Tick-bacteria mutualism depends on B vitamin synthesis pathways. *Current Biology*, 28(12): 1896–1902.e5. **URL** https://dx.doi.org/10.1016/j.cub.2018.04.038.

Edgar R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5): 1792–1797. **URL** https://dx.doi.org/10.1093/nar/gkh340.

Escudié F, Auer L, Bernard M, Mariadassou M, Cauquil L, Vidal K, Maman S, Hernandez-Raquet G, Combes S, and Pascal G. 2018. FROGS: Find, rapidly, OTUs with Galaxy solution. *Bioinformatics*, 34(8): 1287–1294. **URL** https://dx.doi.org/10.1093/bioinformatics/btx791.

Favret C. 2022. Aphid Species File. *Version 5.0/5.0*. **URL** http://aphid.speciesfile.org.

Flandrois J.-P, Perrière G, and Gouy M. 2015. leBIBIQBPP: a set of databases and a webtool for automatic phylogenetic analysis of

prokaryotic sequences. *BMC Bioinformatics*, 16(1): 251. **URL** https://dx.doi.org/10.1186/s12859-015-0692-z.

Fukatsu T. 2001. Secondary intracellular symbiotic bacteria in aphids of the genus *Yamatocallis* (Homoptera: Aphididae: Drepanosiphinae). *Applied and Environmental Microbiology*, 67(11): 5315–5320. **URL** https://dx.doi.org/10.1128/AEM.67.11.5315-5320.2001.

Fukatsu T and Ishikawa H. 1992. A novel eukaryotic extracellular symbiont in an aphid, *Astegopteryx styraci* (Homoptera, Aphididae, Hormaphidinae). *Journal of Insect Physiology*, 38(10): 765–773. **URL** https://dx.doi.org/10.1016/0022-1910(92)90029-D.

Fukatsu T and Ishikawa H. 1993. Occurrence of Chaperonin 60 and Chaperonin 10 in primary and secondary bacterial symbionts of aphids: Implications for the evolution of an endosymbiotic system in aphids. *Journal of Molecular Evolution*, 36(6): 568–577. **URL** https://dx.doi.org/10.1007/BF00556361.

Fukatsu T, Aoki S, Kurosu U, and Ishikawa H. 1994. Phylogeny of Cerataphidini aphids revealed by their symbiotic microorganisms and basic structure of their galls - Implications for host-symbiont coevolution and evolution of sterile soldier castes. *Zoological Science*, 11(4): 613–623. **URL** http://ci.nii.ac.jp/naid/110003323229.

Fukatsu T, Nikoh N, Kawai R, and Koga R. 2000. The secondary endosymbiotic bacterium of the Pea aphid *Acyrthosiphon pisum* (Insecta: Homoptera). *Applied and Environmental Microbiology*, 66(7): 2748–2758. **URL** 10.1128/AEM.66.7.2748-2758.2000.

Funk D. J, Helbling L, Wernegreen J. J, and Moran N. A. 2000. Intraspecific phylogenetic congruence among multiple symbiont genomes. *Proceedings of the Royal Society B: Biological Sciences*, 267(1461): 2517–2521. **URL** https://dx.doi.org/10.1098/rspb.2000.1314.

Gerth M and Bleidorn C. 2017. Comparative genomics provides a timeframe for *Wolbachia* evolution and exposes a recent biotin synthesis operon transfer. *Nature Microbiology*, 2(3): 16241. **URL** https://dx.doi.org/10.1038/nmicrobiol.2016.241/.

Gil R, Sabater-Muñoz B, Perez-Brocal V, Silva F. J, and Latorre A. 2006. Plasmids in the aphid endosymbiont *Buchnera aphidicola* with the smallest genomes. A puzzling evolutionary story. *Gene*, 370(1): 17–25. **URL** https://dx.doi.org/10.1016/j.gene.2005.10.043.

Hansen A. K and Moran N. A. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proceedings of the National Academy of Sciences*, 108(7): 2849–2854. **URL** https://dx.doi.org/10.1073/pnas.1013465108.

Heyworth E. R and Ferrari J. 2015. A facultative endosymbiont in aphids can provide diverse ecological benefits. *Journal of Evolutionary Biology*, 28(10): 1753–1760. **URL** https://dx.doi.org/10.1111/jeb.12705.

Holley J.-a. C, Jackson M. N, Pham A. T, Hatcher S. C, and Moran N. A. 2022. Carpenter bees (*Xylocopa*) harbor a distinctive gut microbiome related to that of oney bees and bumble bees. *Applied and Environmental Microbiology*, 88(13). **URL** https://dx.doi.org/10.1128/aem.00203-22.

Hosokawa T, Ishii Y, Nikoh N, Fujie M, Satoh N, and Fukatsu T. 2016. Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. *Nature Microbiology*, 1(1): 15011. **URL** https://dx.doi.org/10.1038/nmicrobiol.2015.11.

Husnik F and McCutcheon J. P. 2016. Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proceedings of the National Academy of Sciences*, 113(37): E5416–E5424. **URL** https://dx.doi.org/10.1073/pnas.1603910113.

Husník F, Chrudimský T, and Hypša V. 2011. Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. *BMC Biology*, 9(1): 87. **URL** https://dx.doi.org/10.1186/1741-7007-9-87.

Husnik F, Nikoh N, Koga R, Ross L, Duncan R. P, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson A. C, von Dohlen C. D, Fukatsu T,

and McCutcheon J. P. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*, 153(7): 1567–1578. **URL** https://dx.doi.org/10.1016/j.cell.2013.05.040.

Jousselin E, Desdevises Y, and Coeur d'acier A. 2009. Fine-scale cospeciation between *Brachycaudus* and *Buchnera aphidicola*: bacterial genome helps define species and evolutionary relationships in aphids. *Proceedings of the Royal Society B: Biological Sciences*, 276(1654): 187–196. **URL** https://dx.doi.org/10.1098/rspb.2008.0679.

Jousselin E, Clamens A.-L, Galan M, Bernard M, Maman S, Gschloessl B, Duport G, Meseguer A. S, Calevro F, and Cœur d'Acier A. 2016. Assessment of a 16S rRNA amplicon Illumina sequencing procedure for studying the microbiome of a symbiont-rich aphid genus. *Molecular Ecology Resources*, 16(3): 628–640. **URL** https://dx.doi.org/10.1111/1755-0998.12478.

Kalvari I, Nawrocki E. P, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, Rivas E, Eddy S. R, Finn R. D, Bateman A, and Petrov A. I. 2021. Rfam 14: Expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research*, 49(D1): D192–D200. **URL** https://dx.doi.org/10.1093/nar/gkaa1047.

Katoh K and Standley D. M. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4): 772–780. **URL** https://dx.doi.org/10.1093/molbev/mst010.

Koga R and Moran N. A. 2014. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. *The ISME journal*, 8(6): 1237–46. **URL** https://dx.doi.org/10.1038/ismej.2013.235.

Koga R, Tsuchida T, and Fukatsu T. 2003. Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid. *Proceedings of the Royal Society B: Biological Sciences*, 270(1533): 2543–2550. **URL** https://dx.doi.org/10.1098/rspb.2003.2537.

Koga R, Tsuchida T, and Fukatsu T. 2009. Quenching autofluorescence of insect tissues for in situ detection of endosymbionts. *Applied Entomology and Zoology*, 44(2): 281–291. **URL** https://dx.doi.org/10.1303/aez.2009.281.

Kot M. 2012. Ovariole structure in viviparous and oviparous generations of *Glyphina betulae* (Linnaeus 1758) (Insecta, Hemiptera, Aphidinea: Thelaxidae). *Aphids and Other Hemipterous Insects*, (18): 13–20. **URL** https://www.kul.pl/files/658/aphids18/02_Kot.pdf.

Kozich J. J, Westcott S. L, Baxter N. T, Highlander S. K, and Schloss P. D. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17): 5112–5120. **URL** https://dx.doi.org/10.1128/AEM.01043-13.

Kwong W. K and Moran N. A. 2013. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the *Betaproteobacteria*, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., *Orbales* ord. nov., a sister taxon to the order '*Enterobacteriales*' of the *Gammaproteobacteria*. *International journal of systematic and evolutionary microbiology*, 63(Pt 6): 2008–2018. **URL** https://dx.doi.org/10.1099/ijs.0.044875-0.

Lamelas A, Perez-Brocal V, Gomez-Valero L, Gosalbes M. J, Moya A, and Latorre A. 2008. Evolution of the secondary symbiont "*Candidatus* Serratia symbiotica" in aphid species of the subfamily Lachninae. *Applied and Environmental Microbiology*, 74(13): 4236–4240. **URL** https://dx.doi.org/10.1128/AEM.00022-08.

Lamelas A, Gosalbes M. J, Moya A, and Latorre A. 2011a. New clues about the evolutionary history of metabolic losses in bacterial endosymbionts, provided by the genome of *Buchnera aphidicola* from the aphid *Cinara tujafilina*. *Applied and Environmental Microbiology*, 77(13): 4446–4454. **URL** https://dx.doi.org/10.1128/AEM.00141-11.

Lamelas A, Gosalbes M. J, Manzano-Marín A, Peretó J, Moya A, and Latorre A. 2011b. *Serratia symbiotica* from the aphid *Cinara cedri*: A missing link from facultative to obligate insect endosymbiont. *PLoS Genetics*, 7(11): e1002357. **URL** https://dx.doi.org/10.1371/journal.pgen.1002357.

Laslett D. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research*, 32(1): 11–16. **URL** https://dx.doi.org/10.1093/nar/gkh152.

Li L. 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, 13(9): 2178–2189. **URL** https://dx.doi.org/10.1101/gr.1224503.

Li X.-Y, Jiang L.-Y, and Qiao G.-X. 2014. Is the subfamily Eriosomatinae (Hemiptera: Aphididae) monophyletic? *Turkish Journal of Zoology*, 38(3): 285–297. **URL** http://dergipark.ulakbim.gov.tr/tbtkzoology/article/view/5000022742.

Luan J.-B, Chen W, Hasegawa D. K, Simmons A. M, Wintermantel W. M, Ling K.-S, Fei Z, Liu S.-S, and Douglas A. E. 2015. Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. *Genome Biology and Evolution*, 7(9): 2635–2647. **URL** https://dx.doi.org/10.1093/gbe/evv170.

Łukasik P, Nazario K, Van Leuven J. T, Campbell M. A, Meyer M, Michalik A, Pessacq P, Simon C, Veloso C, and McCutcheon J. P. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of cicadas. *Proceedings of the National Academy of Sciences*, 115(2): E226–E235. **URL** https://dx.doi.org/10.1073/pnas.1712321115.

Magoč T and Salzberg S. L. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, 27(21): 2957–2963. **URL** https://dx.doi.org/10.1093/bioinformatics/btr507.

Mahé F, Rognes T, Quince C, de Vargas C, and Dunthorn M. 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ*, 2: e593. **URL** https://dx.doi.org/10.7717/peerj.593.

Manzano-Marín A. 2020. No evidence for *Wolbachia* as a nutritional co-obligate endosymbiont in the aphid *Pentalonia nigronervosa*. *Microbiome*, 8(1): 72. **URL** https://dx.doi.org/10.1186/s40168-020-00865-2.

Manzano-Marín A and Latorre A. 2014. Settling down: The genome of *Serratia symbiotica* from the aphid *Cinara tujafilina* zooms in on the process of accommodation to a cooperative intracellular life. *Genome Biology and Evolution*, 6(7): 1683–1698. **URL** https://dx.doi.org/10.1093/gbe/evu133.

Manzano-Marín A and Latorre A. 2016. Snapshots of a shrinking partner: Genome reduction in *Serratia symbiotica*. *Scientific Reports*, 6: 32590. **URL** https://dx.doi.org/10.1038/srep32590.

Manzano-Marín A, Simon J.-C, and Latorre A. 2016. Reinventing the wheel and making it round again: Evolutionary convergence in *Buchnera - Serratia* symbiotic consortia between the distantly related Lachninae aphids *Tuberolachnus salignus* and *Cinara cedri*. *Genome Biology and Evolution*, 8(5): 1440–1458. **URL** https://dx.doi.org/10.1093/gbe/evw085.

Manzano-Marín A, Szabó G, Simon J.-C, Horn M, and Latorre A. 2017. Happens in the best of subfamilies: establishment and repeated replacements of co-obligate secondary endosymbionts within Lachninae aphids. *Environmental Microbiology*, 19(1): 393–408. **URL** https://dx.doi.org/10.1111/1462-2920.13633.

Manzano-Marín A, Coeur d'acier A, Clamens A.-L, Orvain C, Cruaud C, Barbe V, and Jousselin E. 2018. A freeloader? The highly eroded yet large genome of the *Serratia symbiotica* symbiont of *Cinara strobi*. *Genome Biology and Evolution*, 10(9): 2178–2189. **URL** https://dx.doi.org/10.1093/gbe/evy173.

Manzano-Marín A, Coeur d'acier A, Clamens A.-L, Orvain C, Cruaud C, Barbe V, and Jousselin E. 2020. Serial horizontal transfer of vitamin-biosynthetic genes enables the establishment of new nutritional symbionts in aphids' di-symbiotic systems. *The ISME Journal*, 14(1): 259–273. **URL** https://dx.doi.org/10.1038/s41396-019-0533-6.

Manzano-Marín A. 2020. Post-publication review of Monnin *et. al.* (2020), "Parallel evolution in the integration of a co-obligate aphid symbiosis". *PubPeer*. **URL** https://pubpeer.com/publications/51CF51BF8B00F33721655DB14425F4.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1): 10. **URL** http://journal.embnet.org/index.php/embnetjournal/article/view/200.

Matsuura Y, Moriyama M, Łukasik P, Vanderpool D, Tanahashi M, Meng X.-Y, McCutcheon J. P, and Fukatsu T. 2018. Recurrent symbiont recruitment from fungal parasites in cicadas. *Proceedings of the National Academy of Sciences*, 115(26): E5970–E5979. **URL** https://dx.doi.org/10.1073/pnas.1803245115.

McCutcheon J. P and Moran N. A. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proceedings of the National Academy of Sciences*, 104(49): 19392–19397.

McCutcheon J. P and von Dohlen C. D. 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Current Biology*, 21(16): 1366–1372. **URL** https://dx.doi.org/10.1016/j.cub.2011.06.051.

McCutcheon J. P, Boyd B. M, and Dale C. 2019. The life of an insect endosymbiont from the cradle to the grave.

Meseguer A. S, Manzano-Marín A, Coeur d'Acier A, Clamens A.-L, Godefroid M, and Jousselin E. 2017. *Buchnera* has changed flatmate but the repeated replacement of co-obligate symbionts is not associated with the ecological expansions of their aphid hosts. *Molecular Ecology*, 26(8): 2363–2378. **URL** https://dx.doi.org/10.1111/mec.13910.

Michalik A. 2010. Ovary structure and transovarial transmission of endosymbiotic microorganisms in *Clethrobius comes*, *Myzocallis walshii* and *Sipha maydis* /Hemiptera, Aphididae: Drepanosiphinae/. *Aphids and Other Hemipterous Insects*, 16: 5–12. **URL** http://www.kul.pl/files/658/aphids16/01_Michalik.pdf.

Michalik A, Szklarzewicz T, Janakowska W, and Wieczorek K. 2014. Endosymbiotic microorganisms of aphids (Hemiptera: Sternorrhyncha: Aphidoidea): Ultrastructure, distribution and transovarial transmission. *European Journal of Entomology*, 111(1): 91–104. **URL** http://www.eje.cz/doi/10.14411/eje.2014.011.html.

Michalik A, Castillo Franco D, Kobiałka M, Szklarzewicz T, Stroiński A, and Łukasik P. 2021. Alternative transmission patterns in independently acquired nutritional cosymbionts of Dictyopharidae planthoppers. *mBio*, 12(4): 2021.04.07.438848. **URL** https://dx.doi.org/10.1128/mBio.01228-21.

Mizrahi-Man O, Davenport E. R, and Gilad Y. 2013. Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: Evaluation of effective study designs. *PLoS ONE*, 8(1): e53608. **URL** https://dx.doi.org/10.1371/journal.pone.0053608.

Monnin D, Jackson R, Kiers E. T, Bunker M, Ellers J, and Henry L. M. 2020. Parallel evolution in the integration of a co-obligate aphid symbiosis. *Current Biology*, 30(10): 1949–1957.e6. **URL** https://dx.doi.org/10.1016/j.cub.2020.03.011.

Montllor C. B, Maxmen A, and Purcell A. H. 2002. Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. *Ecological Entomology*, 27(2): 189–195. **URL** https://dx.doi.org/10.1046/j.1365-2311.2002.00393.x.

Moran N. A, Munson M. A, Baumann P, and Ishikawa H. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proceedings of the Royal Society B: Biological Sciences*, 253(1337): 167–171. **URL** https://dx.doi.org/10.1098/rspb.1993.0098.

Moran N. A, Russell J. A, Koga R, and Fukatsu T. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Applied and Environmental Microbiology*, 71(6): 3302–3310. **URL** https://dx.doi.org/10.1128/AEM.71.6.3302-3310.2005.

Nakabachi A, Ueoka R, Oshima K, Teta R, Mangoni A, Gurgui M, Oldham N. J, van Echten-Deckert G, Okamura K, Yamamoto K, Inoue H, Ohkuma M, Hongoh Y, Miyagishima S.-Y, Hattori M, Piel J, and Fukatsu T. 2013. Defensive bacteriome symbiont with a drastically reduced genome. *Current Biology*, 23(15): 1478–1484. **URL** https://dx.doi.org/10.1016/j.cub.2013.06.027.

Nakabachi A, Piel J, Malenovský I, and Hirose Y. 2020. Comparative genomics underlines multiple roles of *Profftella*, an obligate symbiont of psyllids: Providing toxins, vitamins, and carotenoids. *Genome Biology and Evolution*, 12(11): 1975–1987. **URL** https://dx.doi.org/10.1093/gbe/evaa175.

Nawrocki E. P. 2009. *Structural RNA homology search and alignment using covariance models*. Ph.d. thesis, Washington University in Saint Louis. **URL** http://eddylab.org/publications/Nawrocki09b/Nawrocki09b-phdthesis.pdf.

Nawrocki E. P and Eddy S. R. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*, 29(22): 2933–2935. **URL** https://dx.doi.org/10.1093/bioinformatics/btt509.

Nikoh N, Hosokawa T, Moriyama M, Oshima K, Hattori M, and Fukatsu T. 2014. Evolutionary origin of insect-Wolbachia nutritional mutualism. *Proceedings of the National Academy of Sciences*, 111(28): 10257–10262. **URL** https://dx.doi.org/10.1073/pnas.1409284111.

Nikoh N, Koga R, Oshima K, Hattori M, and Fukatsu T. 2019. Genome sequence of "*Candidatus* Serratia symbiotica" strain IS, a facultative bacterial symbiont of the pea aphid *Acyrthosiphon pisum*. *Microbiology Resource Announcements*, 8(19): e00272–19.

Nováková E, Hypša V, Klein J, Foottit R. G, von Dohlen C. D, and Moran N. A. 2013. Reconstructing the phylogeny of aphids (Hemiptera: Aphididae) using DNA of the obligate symbiont *Buchnera aphidicola*. *Molecular phylogenetics and evolution*, 68(1): 42–54. **URL** https://dx.doi.org/10.1016/j.ympev.2013.03.016.

Okonechnikov K, Golosova O, and Fursov M. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8): 1166–1167. **URL** https://dx.doi.org/10.1093/bioinformatics/bts091.

Oliver K. M, Russell J. A, Moran N. A, and Hunter M. S. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences*, 100(4): 1803–1807. **URL** https://dx.doi.org/10.1073/pnas.0335320100.

Ortiz-Rivas B and Martínez-Torres D. 2010. Combination of molecular data support the existence of three main lineages in the phylogeny of aphids (Hemiptera: Aphididae) and the basal position of the subfamily Lachninae. *Molecular Phylogenetics and Evolution*, 55(1): 305–317. **URL** https://dx.doi.org/10.1016/j.ympev.2009.12.005.

Pais I. S, Valente R. S, Sporniak M, and Teixeira L. 2018. *Drosophila melanogaster* establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. *PLOS Biology*, 16(7): e2005710. **URL** https://dx.doi.org/10.1371/journal.pbio.2005710.

Patel V, Chevignon G, Manzano-Marín A, Brandt J. W, Strand M. R, Russell J. A, and Oliver K. M. 2019. Cultivation-assisted genome of *Candidatus* Fukatsuia symbiotica; the enigmatic 'X-type' symbiont of aphids. *Genome Biology and Evolution*, 11(12): 3510–3522. **URL** https://doi.org/10.1093/gbe/evz252.

Plasterk R and van de Putte P. 1985. The invertible P-DNA segment in the chromosome of *Escherichia coli*. *The EMBO Journal*, 4(1): 237–242. **URL** https://dx.doi.org/10.1002/j.1460-2075.1985.tb02341.x.

Pons I, Renoz F, Noël C, and Hance T. 2019. Circulation of the cultivable symbiont *Serratia symbiotica* in aphids is mediated by plants. *Frontiers in Microbiology*, 10. **URL** 10.3389/fmicb.2019.00764.

Pons I, Scieur N, Dhondt L, Renard M.-E, Renoz F, and Hance T. 2022. Pervasiveness of the symbiont *Serratia symbiotica* in the aphid natural environment: distribution, diversity and evolution at a multitrophic level. *FEMS Microbiology Ecology*, 98(1). **URL** https://dx.doi.org/10.1093/femsec/fiac012/6526308.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, and Glockner F. O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1): D590–D596. **URL** https://dx.doi.org/10.1093/nar/gks1219.

Ren F.-R, Sun X, Wang T.-Y, Yao Y.-L, Huang Y.-Z, Zhang X, and Luan J.-B. 2020. Biotin provisioning by horizontally transferred genes from bacteria confers animal fitness benefits. *The ISME Journal*, 14(10): 2542–2553. **URL** https://dx.doi.org/10.1038/s41396-020-0704-5.

Renoz F, Lopes M. R, Gaget K, Duport G, Eloy M.-C, Geelhand de Merxem B, Hance T, and Calevro F. 2022a. Compartmentalized into bacteriocytes but highly invasive: the puzzling case of the co-obligate symbiont *Serratia symbiotica* in the aphid *Periphyllus lyropictus*. *Microbiology Spectrum*, 10(3). **URL** https://dx.doi.org/10.1128/spectrum.00457-22.

Renoz F, Ambroise J, Bearzatto B, Fakhour S, Parisot N, Lopes M. R, Gala J.-L, Calevro F, and Hance T. 2022b. The di-symbiotic systems in the aphids *Sipha maydis* and *Periphyllus lyropictus* provide a contrasting picture of recent co-obligate nutritional endosymbiosis in aphids. *Microorganisms 2022, Vol. 10, Page 1360*, 10(7): 1360. **URL** https://dx.doi.org/10.3390/MICROORGANISMS10071360.

Rognes T, Flouri T, Nichols B, Quince C, and Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4: e2584. **URL** https://dx.doi.org/10.7717/peerj.2584.

Ronquist F, Teslenko M, van der Mark P, Ayres D. L, Darling A, Hohna S, Larget B, Liu L, Suchard M. A, and Huelsenbeck J. P. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3): 539–542. **URL** https://dx.doi.org/10.1093/sysbio/sys029.

Rouïl J, Jousselin E, Coeur d'acier A, Cruaud C, and Manzano-Marín A. 2020. The protector within: Comparative genomics of APSE phages across aphids reveals rampant recombination and diverse toxin arsenals. *Genome Biology and Evolution*, 12(6): 878–889. **URL** https://dx.doi.org/10.1101/2020.03.17.996009.

Russell J. A and Moran N. A. 2006. Costs and benefits of symbiont infection in aphids: variation among symbionts and across temperatures. *Proceedings of the Royal Society B: Biological Sciences*, 273(1586): 603–10. **URL** https://dx.doi.org/10.1098/rspb.2005.3348.

Sandstrom J and Moran N. 1999. How nutritionally imbalanced is phloem sap for aphids? *Entomologia Experimentalis et Applicata*, 91(1): 203–210. **URL** https://dx.doi.org/10.1046/j.1570-7458.1999.00485.x.

Schmieder R and Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27(6): 863–864. **URL** https://dx.doi.org/10.1093/bioinformatics/btr026.

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

Shigenobu S, Watanabe H, Hattori M, Sakaki Y, and Ishikawa H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature*, 407(6800): 81–6. **URL** https://dx.doi.org/10.1038/35024074.

Sloan D. B and Moran N. A. 2012. Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. *Molecular Biology and Evolution*, 29(12): 3781–3792. **URL** https://dx.doi.org/10.1093/molbev/mss180.

Sloan D. B, Nakabachi A, Richards S, Qu J, Murali S. C, Gibbs R. A, and Moran N. A. 2014. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Molecular Biology and Evolution*, 31(4): 857–871. **URL** https://dx.doi.org/10.1093/molbev/msu004.

Smith E. A and Newton I. L. G. 2020. Genomic signatures of Honey bee association in an acetic acid symbiont. *Genome Biology and Evolution*, 12(10): 1882–1894. **URL** https://dx.doi.org/10.1093/gbe/evaa183.

Sow A, Brévault T, Benoit L, Chapuis M.-P, Galan M, Coeur d'acier A, Delvare G, Sembène M, and Haran J. 2019. Deciphering host-parasitoid interactions and parasitism rates of crop pests using DNA metabarcoding. *Scientific Reports*, 9(1): 3646. **URL** https://dx.doi.org/10.1038/s41598-019-40243-z.

Sudakaran S, Kost C, and Kaltenpoth M. 2017. Symbiont acquisition and replacement as a source of ecological innovation. *Trends in Microbiology*, 25(5): 375–390. **URL** https://dx.doi.org/10.1016/j.tim.2017.02.014.

Szabó G, Schulz F, Toenshoff E. R, Volland J.-M, Finkel O. M, Belkin S, and Horn M. 2017. Convergent patterns in the evolution of mealybug symbioses involving different intrabacterial symbionts. *The ISME Journal*, 11(3): 715–726. **URL** https://dx.doi.

org/10.1038/ismej.2016.148.

Szabó G, Schulz F, Manzano-Marín A, Toenshoff E. R, and Horn M. 2022. Evolutionarily recent dual obligatory symbiosis among adelgids indicates a transition between fungus- and insect-associated lifestyles. *The ISME Journal*, 16(1): 247–256. **URL** https://dx.doi.org/10.1038/s41396-021-01056-z.

Talavera G and Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, 56(4): 564–577. **URL** https://dx.doi.org/10.1080/10635150701472164.

Tamas I, Wernegreen J. J, Nystedt B, Kauppinen S. N, Darby A. C, Gomez-Valero L, Lundin D, Poole A. M, and Andersson S. G. E. 2008. Endosymbiont gene functions impaired and rescued by polymerase infidelity at poly(A) tracts. *Proceedings of the National Academy of Sciences of the United States of America*, 105(39): 14934–9. **URL** https://dx.doi.org/10.1073/pnas.0806554105.

Toenshoff E. R, Gruber D, and Horn M. 2012. Co-evolution and symbiont replacement shaped the symbiosis between adelgids (Hemiptera: Adelgidae) and their bacterial symbionts. *Environmental Microbiology*, 14(5): 1284–1295. **URL** https://dx.doi.org/10.1111/j.1462-2920.2012.02712.x.

Toenshoff E. R, Szabo G, Gruber D, and Horn M. 2014. The pine bark adelgid, *Pineus strobi*, contains two novel bacteriocyte-associated gammaproteobacterial symbionts. *Applied and Environmental Microbiology*, 80(3): 878–885. **URL** https://dx.doi.org/10.1128/AEM.03310-13.

van Ham R. C. H. J, Moya A, and Latorre A. 1997. Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *Journal of bacteriology*, 179(15): 4768–4777. **URL** https://dx.doi.org/10.1128/jb.179.15.4768-4777.1997.

Van Ham R. C. H. J, González-Candelas F, Silva F. J, Sabater-Muñoz B, Moya A, and Latorre A. 2000. Postsymbiotic plasmid acquisition and evolution of the *repA1*-replicon in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20): 10855–60. **URL** https://dx.doi.org/10.1073/pnas.180310197.

van Ham R. C. H. J, Kamerbeek J, Palacios C, Rausell C, Abascal F, Bastolla U, Fernandez J. M, Jimenez L, Postigo M, Silva F. J, Tamames J, Viguera E, Latorre A, Valencia A, Moran F, and Moya A. 2003. Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Sciences*, 100(2): 581–586. **URL** https://dx.doi.org/10.1073/pnas.0235981100.

Vogel K. J and Moran N. A. 2013. Functional and evolutionary analysis of the genome of an obligate fungal symbiont. *Genome Biology and Evolution*, 5(5): 891–904. **URL** https://dx.doi.org/10.1093/gbe/evt054.

von Dohlen C. D and Moran N. A. 2000. Molecular data support a rapid radiation of aphids in the Cretaceous and multiple origins of host alternation. *Biological Journal of the Linnean Society*, 71(4): 689–717. **URL** https://dx.doi.org/10.1111/j.1095-8312.2000.tb01286.x.

von Dohlen C. D, Spaulding U, Shields K, Havill N. P, Rosa C, and Hoover K. 2013. Diversity of proteobacterial endosymbionts in hemlock woolly adelgid (*Adelges tsugae*) (Hemiptera: Adelgidae) from its native and introduced range. *Environmental Microbiology*, 15(7): 2043–2062. **URL** https://dx.doi.org/10.1111/1462-2920.12102.

Waterworth S. C, Flórez L. V, Rees E. R, Hertweck C, Kaltenpoth M, and Kwan J. C. 2020. Horizontal gene transfer to a defensive symbiont with a reduced genome in a multipartite beetle microbiome. *mBio*, 11(1). **URL** https://dx.doi.org/10.1128/mBio.02430-19.

Weglarz K. M, Havill N. P, Burke G. R, and von Dohlen C. D. 2018. Partnering with a pest: Genomes of Hemlock woolly adelgid symbionts reveal atypical nutritional provisioning patterns in dual-obligate bacteria. *Genome Biology and Evolution*, 10(6): 1607–1621. **URL** https://dx.doi.org/10.1093/gbe/evy114/5026599.

Xie W, Yang X, Chen C, Yang Z, Guo L, Wang D, Huang J, Zhang H, Wen Y, Zhao J, Wu Q, Wang S, Coates B. S, Zhou X, and

Zhang Y. 2018. The invasive MED/Q *Bemisia tabaci* genome: A tale of gene loss and gene gain. *BMC Genomics*, 19(1): 68. **URL** https://dx.doi.org/10.1186/s12864-018-4448-9.

Yarza P, Yilmaz P, Pruesse E, Glöckner F. O, Ludwig W, Schleifer K.-H, Whitman W. B, Euzéby J, Amann R, and Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, 12(9): 635–645. **URL** https://dx.doi.org/10.1038/nrmicro3330.

Yorimoto S, Hattori M, Kondo M, and Shigenobu S. 2022. Complex host/symbiont integration of a multi-partner symbiotic system in the eusocial aphid *Ceratovacuna japonica*. *iScience*, page 105478. **URL** 10.1016/j.isci.2022.105478.

Ziegler H. 1975. Nature of transported substances. In M. H. Zimmermann and J. A. Milburn, editors, *Encyclopedia of plant physiology. Volume 1. Transport in Plants I*, chapter 3, pages 59–100. Springer Berlin Heidelberg, Berlin, Heidelberg. **URL** https://dx.doi.org/10.1007/978-3-642-66161-7.

Zytynska S. E and Weisser W. W. 2016. The natural occurrence of secondary bacterial symbionts in aphids. *Ecological Entomology*, 41(1): 13–26. **URL** https://dx.doi.org/10.1111/een.12281.

Říhová J, Nováková E, Husník F, and Hypša V. 2017. *Legionella* becoming a mutualist: adaptive processes shaping the genome of symbiont in the louse *Polyplax serrata*. *Genome Biology and Evolution*, 9(11): 2946–2957. **URL** https://dx.doi.org/10.1093/gbe/evx217/4562435.