Serial horizontal transfer of vitamin-biosynthetic genes enables the establishment of new nutritional symbionts in aphids' di-symbiotic systems

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

[1]UMR 1062 Centre de Biologie pour la Gestion des Populations, INRA, CIRAD, IRD, Montpellier SupAgro, Univ. Montpellier, Montpellier, France
[2]Institut de Biologie François-Jacob, CEA, Genoscope, Évry Cedex, France
[*]alejandro.manzano.marin@gmail.com

ABSTRACT

Many insects with a nutrient-restricted diet depend on obligate mutualistic bacteria for the provisioning of essential nutrients lacking from their food source, namely essential amino acids and B vitamins. Most aphids (Hemiptera: Aphididae), whose diet consists of phloem, rely on the bacterial endosymbiont *Buchnera* for the supply of the aforementioned compounds. However, in some aphid lineages *Buchnera* have lost the capability of producing these nutrients and thus the symbiotic consortium has accommodated an extra bacterial partner to supplement *Buchnera*'s deficiencies. In this work, we explore the di-symbiotic nutritional endosymbiosis of a group of *Cinara* aphids which has been found to harbour both *Buchnera* and an *Erwinia*-related symbiont. Using fluorescence *in situ* hybridisation, we have located this symbiont to the bacteriome where it inhabits its own bacteriocytes. Through whole-genome sequencing of the endosymbionts of 9 species of *Erwinia*-associated *Cinara* aphids, we have found that *Ewrinia* genomes are highly syntenic and all show significant genome reduction. Additionally, *Erwinia* symbionts display phylogenetic congruency with *Buchnera*, suggesting long-term co-divergence. Most significantly, we found that not only is *Erwinia* capable of complementing *Buchnera*'s auxotrophies, but that the genes involved in the biosynthesis of two B vitamins have actually been horizontally acquired from a *Sodalis*-related bacterium. Finally, this B-vitamin biosynthetic genes have been further transferred to a new *Hamiltonella* co-obligate symbiont in a specific *Cinara* lineage, thus displaying a tri-symbiotic system. These results highlight the important role horizontal gene transfer plays in the establishment of new obligate nutritional symbionts.

Keywords: symbiosis, mutualism, horizontal gene transfer, vitamin biosynthesis, *Erwinia haradaeae*.

Introduction

Obligate symbiotic bacteria are pervasive among different groups of animals whose diets lack essential nutrients. These include blood- (Manzano-Marín *et al.*, 2015; Smith *et al.*, 2015; Říhová *et al.*, 2017) and sap-feeders (xylem and phloem) (Husnik and McCutcheon, 2016; Matsuura *et al.*, 2018; Santos-Garcia *et al.*, 2014). Aphids (Hemiptera: Aphididae) are plant-sucking insects whose diet consists of plant phloem, which is poor in essential amino acids and B vitamins (Akman Gündüz and Douglas, 2009; Sandstrom and Moran, 1999; Ziegler, 1975). Accordingly, they depend on obligate symbiotic bacteria from the *Buchnera* genus to supply these nutrients. *Buchnera* localises to the bacteriome (Buchner, 1953), a specialised organ made up mainly of cells called bacteriocytes which house the *Buchnera* bacteria intracellularly and (Griffiths and Beck, 1975; Munson *et al.*, 1991). Similarly to other "long-term" vertically-inherited obligate endosymbionts, *Buchnera* strains have highly reduced genomes, with sizes ranging between 0.656 and 0.422 Mega base pairs (**Mbp**). Additionally, they all show a high degree

of genome synteny (Tamas *et al.*, 2002; van Ham *et al.*, 2003) and display strong phylogenetic congruency with their hosts (Baumann *et al.*, 1995; Funk *et al.*, 2000; Jousselin *et al.*, 2009), evidencing an early drastic genome reduction preceding the diversification of aphids and followed by co-divergence.

Aphids can also harbour additional facultative endosymbionts, which are neither required for normal development nor are generally fixed across populations of the species (Oliver *et al.*, 2010). These secondary bacterial associates have beneficial effects for their hosts, such as protection against parasitoids (Hansen *et al.*, 2012; Oliver *et al.*, 2003) or survival after heat stress (Chen *et al.*, 2000; Montllor *et al.*, 2002). Some aphids have established new obligate nutritional-based associations with additional bacterial lineages. The best known cases occur in species from the Lachninae subfamily of aphids, where *Buchnera* symbionts have ancestrally lost the capability to synthesise two essential B vitamins: biotin (B_7) and riboflavin (B_{12}) (Manzano-Marín *et al.*, 2016; Meseguer *et al.*, 2017). As a consequence, both the aphids and their *Buchnera* now rely on fixed secondary endosymbionts to complement the nutritional deficiencies. These younger co-obligate symbionts belong to diverse gammaproteobacterial taxa, including *Sodalis*, *Fukatsuia*, and *Serratia symbiotica* (Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017). Contrary to *Buchnera*, these symbionts present more diverse cell shapes and tissue tropism. They can be round or rod-shaped confined to the cytoplasm of bacteriocytes, or they can display long filamentous cells that are found surrounding bacteriocytes, co-infecting *Buchnera*'s, or inhabiting their own (Buchner, 1953; Fukatsu and Ishikawa, 1993; Manzano-Marín *et al.*, 2017; Pyka-Fościak and Szklarzewicz, 2008).

In *Cinara* aphids (the most species-rich genus within the Lachninae), several cases of secondary co-obligate symbiont replacement have been proposed (Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017). Based on an ancestral reconstruction of the bacterial symbiotic associations within this aphid clade, Meseguer *et al.* (2017) suggested that the common ancestor of these aphids harboured a *Se. symbiotica* endosymbiont which has been subsequently replaced several times by other bacterial taxa. These include the aforementioned *Sodalis* and *Fukatsuia*, as well as the gammaproteobacteria *Erwinia*. While the formers have closely related relatives found living as facultative or early co-obligate symbionts from aphid species (Guay *et al.*, 2009; Heyworth and Ferrari, 2015; Manzano-Marín *et al.*, 2017; Meseguer *et al.*, 2017), the latter has not, to our knowledge, been identified as more than part of the gut flora of *Acyrthosiphon pisum*.

Erwinia bacteria are primarily known as plant-pathogenic and plant-associated bacteria (Kado, 2006). Their genomes sizes range between 3.83 and 5.07 Mbps (based on currently available fully closed genomes). Regarding *Erwinia* and aphids, early studies isolated and axenically cultured an aerobic *Erwinia* bacterium from the gut of the pea aphid *Ac. pisum*, named *Erwinia aphidicola* (Harada *et al.*, 1997). This bacterial species was found to be the largest in number among the seven bacterial groups which were found to constitute the major flora of the aphid gut (Harada and Ishikawa, 1993; Harada *et al.*, 1996). In *Cinara* aphids, *Erwinia* bacteria have been found to be fixed secondary symbionts in a particular monophyletic group of *Cinara* aphids (Meseguer *et al.*, 2017). Moreover, the study showed that at least one strain of these *Erwinia* symbionts has a reduced genome with the capability to rescue *Buchnera*'s biotin and riboflavin auxotrophy. Given the above, this association is of particular interest due the possible recent transition of *Erwinia* from being a free-living gut associated bacteria to becoming a co-obligate vertically-transmitted nutritional endosymbiont. This sparks several questions: Are these symbionts derived from a single event of settlement of a free-living *Erwinia* lineage? Have they since cospeciated with their aphid hosts (as shown for the primary symbiont Buchnera)? or have they been acquired and lost several times, as is probably the case for obligate *Se. symbiotica* (Manzano-Marín and Latorre, 2016)? What changes has *Erwinia* undergone through this transition (e.g. genome reduction, metabolic pathway losses)? What is the localisation of these symbionts within the aphid? and finally, what factors are pivotal to the *Buchnera-Erwinia* association?

To explore these questions, we first used fluorescence *in situ* hybridisation (**FISH**) microscopy to explore the localisation of *Erwinia* within the aphid. Then, we sequenced, assembled, and annotated the genomes of the co-existing endosymbionts of ten *Erwinia*-associated *Cinara* species. Additionally, we explored the genome-based metabolic complementation displayed by *Buchnera-Erwinia* symbiont pairs. These approaches revealed yet a third co-obligate *Hamiltonella* endosymbiont present in two species. Through a combination of BLAST searches and phylogenetic analyses, we determined that key genes involved

in the symbionts' nutritional roles have been horizontally acquired by *Erwinia* and subsequently passed on to *Hamiltonella*. Our results show that *Erwinia* symbionts derive from a single event of lifestyle shift, have then co-diverged with their hosts and associated *Buchnera* for millions of years, and that serial horizontal gene transfer (**HGT**) has played a crucial role in the development and maintenance of both the secondary co-obligate *Erwinia* and the tertiary co-obligate *Hamiltonella* symbiont.

Materials and Methods

Aphid collection, DNA extraction, and sequencing

Following Meseguer *et al.* (2017), we collected nine different species of *Erwinia*-associated *Cinara* aphid species across the north-western and central USA and the south-east of France (supplementary table S1, Supplementary Material online) and then kept in 70% ethanol at 6°C. For whole-genome sequencing, we prepared DNA samples enriched with bacteria following a slightly modified version of the protocol by Charles and Ishikawa (1999) as described in Jousselin *et al.* (2016). For this filtration protocol 15 aphids for one colony were pooled together. Extracted DNA was used to prepare 2 custom paired-end libraries in France Génomique. 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 then purified by Ampure XP (Beckman Coulter, USA) and DNA fragments (>200 bp) 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.

Fluoresence *in situ* hybridisation microscopy

From the aforementioned nine species of *Erwinia*-associated *Cinara* species, we were able to store a variable number of individuals from *Cinara cuneomaculata*, *Cinara kochiana kochiana* (hereafter referred to as *C. kochiana*), and *Cinara curvipes* 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 2 weeks (changing the solution every 3 days). Then, embryos were washed twice with absolute ethanol. Hybridization was performed overnight at 28°C in standard hybridization buffer (20 mM Tris–HCl [pH 8.0], 0.9 M NaCl, 0.01% SDS, and 30% formamide) and then washed (20 mM Tris–HCl [pH 8.0], 5 mM EDTA, 0.1 M NaCl, and 0.01% SDS) before slide preparation. When enough individuals were available, the embryos of up to 10 individuals were viewed under an ZEISS LSM700 confocal microscope. We used two competitive specific probes for *Buchnera* (**BLach-FITC** [5'-FITC-CCCGTTYGCCGCTCGCCGTCA-FITC-3'] from Manzano-Marín *et al.* 2017) and *Erwinia* symbionts in *Cinara* (**ErCinara-Cy3** [5'-ErCinara-Cy3-CCCGTTCGCCACTCGTCGBCA-3']). TIFF-formatted images were exported using the **ZEN** v2.3 SP1 software from ZEISS automatically setting the intensity minimum to Black and the intensity maximum to White value (Auto Min/Max option). Exported images were imported into **Inkscape** v0.92.4 for building the published images.

Genome Assembly and Annotation

Illumina reads were right-tail clipped (minimum quality threshold of 20) using **FASTX-Toolkit** v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/, last accessed February 7 2019). Reads shorted than 75 after the aforementioned clipping were dropped. Additionally, **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. The resulting reads were assembled using **SPAdes** v3.10.1 (Bankevich *et al.*, 2012) with the options --only-assembler option and k-mer sizes of 33,

55, 77, 99, and 127. From the resulting contigs, those that were shorter than 200 bps were dropped. The remaining contigs were binned using results from a **BLASTX** (Altschul, 1997) search (best hit per contig) against a database consisting of the Pea aphid's proteome and a selection of aphid's symbiotic bacteria, and free-living relatives, proteomes (supplementary table S2, Supplementary Material online). When no genome was available for a certain lineage, closely related bacteria were used. The assigned contigs were manually screened using the **BLASTX** web server (searching against the nr database) to insure correct assignment. This binning process confirmed the presence of the previously reported putative *Erwinia* and *Hamiltonella* co-obligate symbionts (Jousselin *et al.*, 2016; Meseguer *et al.*, 2017) as well as other additional symbionts, including *Hamiltonella* in *C. curvipes*. The resulting contigs were then used as reference for read mapping and individual genome assembly using **SPAdes**, as described above, with read error correction.

The resulting genomes were annotated using **Prokka** v1.12 (Seemann, 2014). In order to validate start codons, ribosomal binding sites were predicted using **RBSfinder** (Suzek *et al.*, 2001). This was followed by non-coding RNA prediction using **infernal** v1.1.2 (Nawrocki and Eddy, 2013) (against the **Rfam** v12.3 database (Nawrocki *et al.*, 2015)), **tRNAscan-SE** v1.3.1 (Lowe and Eddy, 1997), and **ARAGORN** v1.2.36 (Laslett, 2004). This annotation was followed by manual curation of the genes on **UGENE** v1.28.1 (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. Priority for the BLAST searches was as follows: **(1)** against *Escherichia coli* K-12 substrain MG1655, **(2)** against **Erwinia amylovora** CFBP1430 (for *Erwinia* symbionts), and **(3)** against the whole nr database. The resulting coding sequences (CDSs) were considered to be putatively functional if all essential domains for the function were found, if a literature search supported the truncated version of the protein as functional in a related organism, or if the CDS displayed truncations but retained identifiable domains (details of the literature captured in the annotation file). For *Erwinia* and *Buchnera* symbionts, pseudogenes were also searched based on synteny against closely related available genomes. This prediction was performed using a combination of sequence alignment with **m-coffee** (Wallace, 2006) and **BLASTX** searches against the NCBI's nr database (restricted to *Erwinia* or *Buchnera* taxon ID). This allowed the identification of pseudogenes missed by the previous searches. The annotated genomes have been submitted to the European Nucleotide Archive with project numbers PRJEB15506, PRJEB31183, PRJEB31187, PRJEB31188, PRJEB31190, PRJEB31191, PRJEB31194, PRJEB31195, and PRJEB31197 and are on queue to be accessioned. They are temporarily available in supplementary file S3 (Supplementary Material online).

Phylogenetic Analyses

For performing both phylogenetic inferences and analysing the genetic differences in *Buchnera* and *Erwinia* from the different aphids, we first ran an orthologous protein clustering analysis using **OrthoMCL** v2.0.9 (Chen *et al.*, 2007; Li, 2003) on two sets: i) *Erwinia* and *Pantoea* strains for *Erwinia*, and ii) Different *Buchnera* strains (supplementary table S3, Supplementary Material online). We then extracted the single copy-core proteins of the selected genomes for phylogenetic reconstruction (351 for *Buchnera* and 320 protein groups for *Erwinia*). We aligned the single-copy core protein sets, gene by gene, using **MAFFT** v7.271 (Katoh and Standley, 2013) (L-INS-i algorithm). We then removed divergent and ambiguously aligned blocks using **Gblocks** v0.91b (Talavera and Castresana, 2007) and concatenated the resulting alignments into a single one for following phylogenetic inference. Maximum-likelihood phylogenetic inference were performed in **IQ-TREE** v1.6.8 using the LG+PMSF+R4 amino acid substitution model (**AASM**) and ultrafast bootstrap approximation with 1000 replicates (Hoang *et al.*, 2018; Nguyen *et al.*, 2015; Wang *et al.*, 2018).

For phylogenetic inference of putative horizontally-transferred genes, we collected orthologous genes across a selection of enterobacterial species (supplementary table S4, Supplementary Material online). Bayesian inference was performed in **MrBayes** v3.2.5 (Ronquist *et al.*, 2012) using the GTR+I+G AASM running two independent analyses with four chains each for 3,000,000 generations and checked for convergence. For phylogenetic inference of the *Hamiltonella* symbionts, we followed Chevignon *et al.* (2018) and used the genes *accD*, *dnaA*, *gyrB*, *hrpA*, *murE*, *ptsI*, and *recJ*. The original authors of the aforementioned study kindly provided the alignment files, we then followed to add the newly acquired sequences and re-aligned

using **MUSCLE** v3.8.31 (Edgar, 2004) and then removed divergent and ambiguously aligned blocks using **Gblocks** v0.91b. Bayesian inference was then performed as previously described. Finally, to infer the origin of the Tn3 mobile elements found in *Erwinia* symbionts' plasmids, we performed on-line **BLASTP** *vs.* NCBI's nr database and collected the top 50 non-redundant hits. We aligned the sequences using **MAFFT** v7.271 (L-INS-i algorithm) and then removed divergent and ambiguously aligned blocks using **Gblocks** v0.91b. Bayesian inference was done in **MrBayes** v3.2.5 using the LG+G AASM running two independent analyses with four chains each for 3,000,000 generations and checked for convergence.

All resulting trees were visualized and exported with **FigTree** v1.4.1 (http://tree.bio.ed.ac.uk/software/figtree/, last accessed February 7 2019) and edited in **Inkscape** v0.92.4. All alignments, NEXUS files, and NEWICK-/NEXUS-formatted trees can be found in https://doi.org/10.5281/zenodo.2566355 (last accessed February 7 2019).

Results

Localisation of *Cinara*-associated *Erwinia* symbionts

To investigate the localisation of *Erwinia* symbionts inside the aphid body, we performed FISH microscopy using competitive specific probes targeting the 16S rRNA from *Buchnera* and *Erwinia* symbionts of 3 species of aphids: *Cinara cuneomaculata* (fig. 1*A*), *Cinara kochiana* (fig. 1*B*), and *Cinara curvipes* (fig. 1*C*). Similarly to *Buchnera*, the *Erwinia* symbionts were found exclusively distributed inside bacteriocytes distinct from those of *Buchnera*. They present a small round cell shape contrasting the larger round cells observed for *Buchnera*. This localisation resembles that of the *Se. symbiotica* symbionts from *C. cedri* and *Tuberolachnus salignus* (both holding a highly reduced genome) and contrasts that of the *Se. symbiotica* symbiont of *Cinara tujafilina* (holding a mildly reduced genome) (Manzano-Marín *et al.*, 2017).

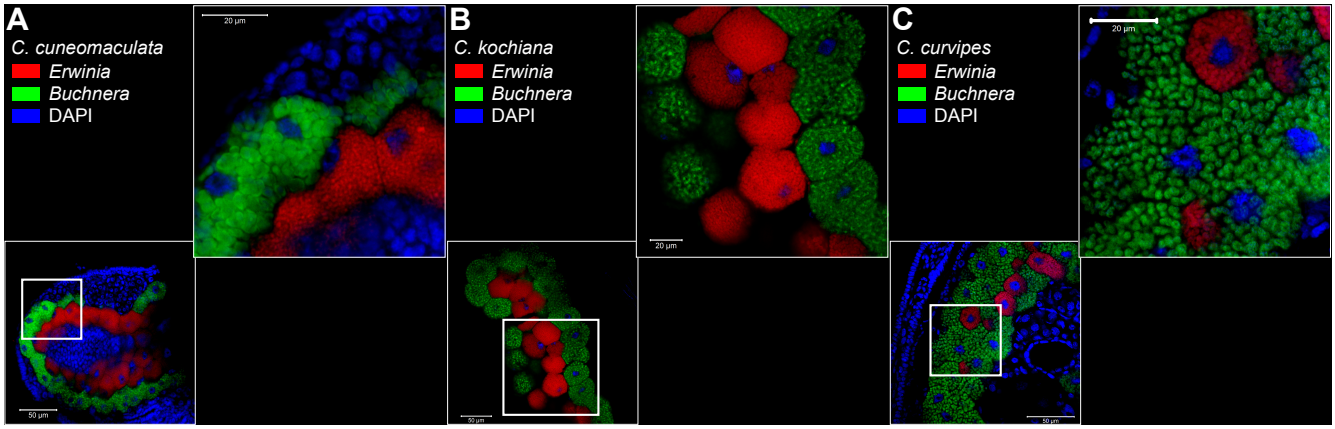

Figure 1. Location and morphology of *Erwinia* and *Buchnera* symbionts in selected *Cinara* aphids. Merged FISH microscopic images of aphid embryos from selected *Cinara* aphids. *Erwinia* signal is shown in red, *Buchnera*'s in green, and DAPI's in blue. Thick white boxes indicate the magnified region 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 group are shown at the top-left of each panel. Unmerged images can be found in supplementary fig. S1 (Supplementary Material online).

Endosymbiont genomes

We successfully assembled the genomes of 9 *Buchnera-Erwinia* symbiont pairs from *Cinara* aphids. The genomes of *Buchnera* are highly conserved in terms of number of genes, and have a size of between 442.57 and 458.25 kilo base pairs with an average G+C content of 23.01% (supplementary table S5, Supplementary Material online). They code a mean of 377 proteins, which are largely a subset of those coded by the previously sequenced *Buchnera* strains harboured by *C. cedri*, *C. tujafilina*, and

Cinara strobi.

Regarding the *Erwinia* endosymbionts, we first performed a phylogenetic reconstruction using a set of single-copy core concatenated protein sequences shared across selected *Erwinia* representing the currently-available diversity of the genus and *Pantoea* species (fig. 2*A*). All *Cinara*-associated *Erwinia* endosymbionts form a well-supported monophyletic group sister to that of a group of free-living *Erwinia* species that have been isolated as both pathogenic and non-pathogenic Rosaceae plant symbionts. When comparing the *Erwinia* endosymbionts' tree topology *vs.* that of *Buchnera*, we observed a congruent evolutionary history between the two endosymbionts.

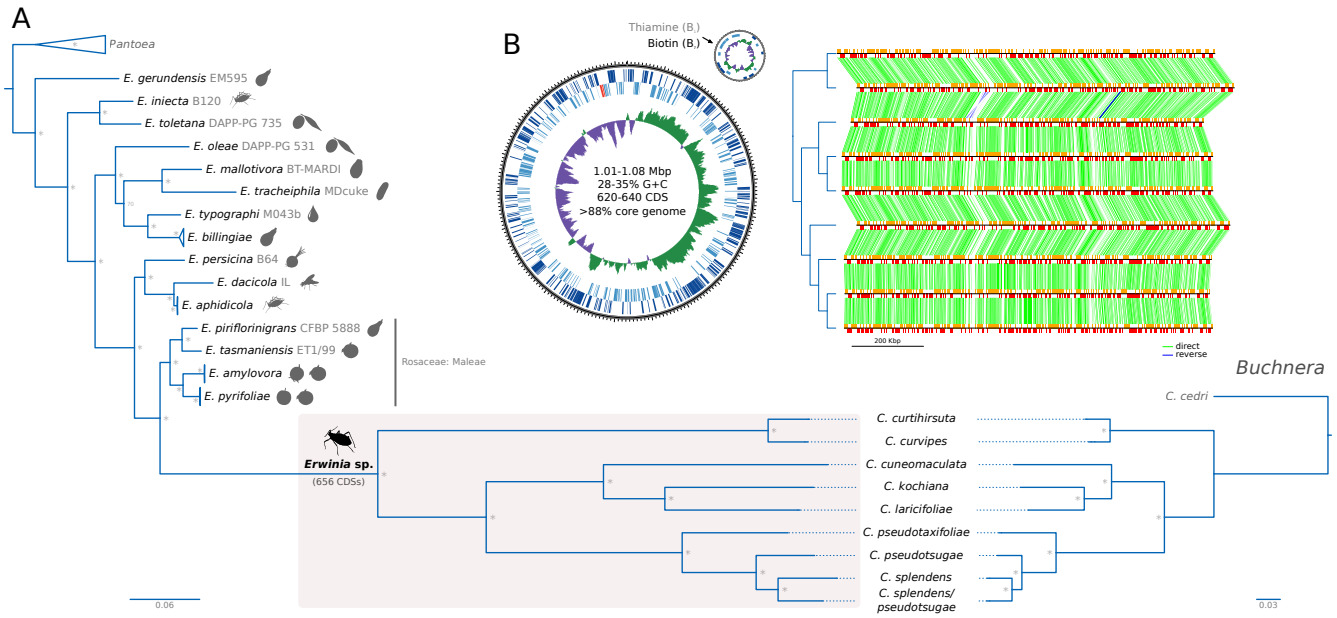

Figure 2. Phylogenetic reconstruction of *Erwinia* spp. and genome properties of *Cinara*-associated *Erwinia* endosymbionts. (A) On the left, phylogenetic reconstruction for *Erwinia* spp. using *Pantoea* spp. as an outgroup. *Erwinia* endosymbionts of *Cinara* species form a well supported monophyletic group. Asterisks at nodes stand for a posterior probability equal to 1. Silhouettes next to the leafs represents the isolation source for the strains. For the *Erwinia* symbionts, the name of the *Cinara* host is used. On the bottom right, the phylogeny of the corresponding *Buchnera* strains using the *Buchnera* associated with *C. cedri* as an outgroup is shown. **(B)** On the left, the genome plot for the endosymbiont of *C. pseudotaxifoliae* along with statistics for sequenced *Erwinia* genomes. From outermost to innermost ring, the features on the direct strand, the reverse strand, and GC-skew plot are shown. On the right, pairwise synteny plots of *Erwinia* endosymbionts with a dendogram on the left displaying the phylogenetic relationships among *Erwinia* symbionts.

Concerning their genomes, we found the *Erwinia* symbionts all show a very similar genome size of around 1 Mbp, an increased A+T content of around 31.50% (when compared to free-living *Erwinia*), an average of 627 CDSs (fig. 2*B*; supplementary table S6, Supplementary Material online), and a lack of mobile elements. Regarding ncRNAs, they all code for a set of 36 tRNAs with charging potential for all of the 20 standard amino acids. A second copy of the *trnX* gene (tRNA-formylmethionine) is found in the symbionts of *C. kochiana* and *C. laricifoliae*, likely arising from a recent duplication event in the ancestor of these two species. All have one tmRNA gene, one 4.5S sRNA component of the Signal Recognition Particle, and the RNase P M1 RNA component. Each genome has an average of 14 pseudogenes, which are largely found as intact CDSs in other *Erwinia* endosymbiont strains. This, in combination with an analysis of shared genes among the *Erwinia* endosymbionts, revealed that the common ancestor of these strains coded for at least 656 distinct CDSs. We also identified a further event of gene duplication in the chromosome involving the *hflC* and *dapB* genes, occurring in the symbiont of *Cinara pseudotaxifoliae*. However, one copy of each gene has been pseudogenised, thus removing redundancy. Regarding genome organisation, the genomes of all sequenced *Erwinia* endosymbionts are highly syntenic (fig. 2*B*), with one major identifiable

event involving the inversion of the region delimited by the *dapD* and *fldA* genes.

All *Erwinia* strains were found to have a plasmid which mainly codes for proteins involved in the biosynthesis of biotin (*bioA* and *bioB*) and thiamine (*thiC*, *thiF*, *thiS*, *thiG*, *thiH*, *thiD*, and *thiE*) as well as a FAD:protein FMN transferase (*apbE*), a 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (*gpmA*), a nucleoside permease (*nupC*), a putative heat shock protein (similar to IbpA/IbpB from *Escherichia coli*), and a PLD-like domain-containing protein. The thiamine-biosynthetic genes are notably missing in the plasmids of the *Eriwnia* symbionts of *Cinara curtihirsuta* and *C. curvipes*.

Regarding their metabolic potential, *Erwinia* symbionts have a highly conserved metabolism. They preserve complete pathways for the synthesis of purine nucleotides from PRPP, L-glycine, and L-aspartate; while they preserve salvage pathways for pyrimidines from thymidine and uridine. They encode for an intact FoF1-ATP synthase, NADH:quinone oxidoreductase 1, and a cytochrome *bo* oxidase. They preserve a fructose PTS permase and can perform glycolysis from glucose-1-P. In terms of the of essential amino acids (**EAAs**), *Erwinia* symbionts do not retain the capability to *de novo* synthesise for any of them. The most intact route is that of L-lysine, where all strains lack the *lysA* gene and show differential retention of the *dapF* gene. Regarding vitamins and cofactors, *Erwinia* symbionts can synthesise riboflavin (from GTP and ribulose-5P, biotin from KAPA, and thiamine pyrophosphate (the biologically active form of thiamine) from PRPP, L-cysteine, and L-tyrosine. This last vitamin biosynthetic pathway is notably missing from the endosymbionts of *C. curtihirsuta* and *C. curvipes*, which form a monophyletic clade.

Metabolic complementation and the tertiary co-obligate *Hamiltonella* symbiont

In previously analysed co-obligate *Buchnera-Se. symbiotica* nutritional endosymbiotic systems in aphids, metabolic complementation is observed for some essential nutrients, including pathway splits regarding the biosynthesis of tryptophan and biotin (Manzano-Marín *et al.*, 2016). In these systems, similarly to the di-symbiotic systems of Lachninae (Meseguer *et al.*, 2017), *Buchnera* lacks key biotin- and riboflavin-biosynthetic genes. As would be expected, the co-obligate associate preserves these genes, complementing the gene losses from *Buchnera*. To infer the metabolic complementation between the *Buchnera* and *Erwinia* partners, we searched for the genes involved in the biosynthesis of EAAs, vitamins, and co-factors (fig. 3; supplementary figs. S2 and S3, Supplementary Material online). Similarly to *Se. symbiotica* co-obligate symbionts of Lachninae, the *Erwinia* symbionts preserve an intact pathway for the biosynthesis of riboflavin as well as the *bioA*, *bioD*, and *bioB* genes; thus being able to complement *Buchnera*'s auxotrophy for these compounds. While most preserve intact pathways for ubiquinol-8 and thiamine biosynthesis; the symbionts of *C. curtihirsuta* and *C. curvipes* have lost the *ubiI* gene as well as most thiamine-biosynthetic genes. Additionally, we have found a third bacterial partner from the *Hamiltonella* genus in both aforementioned *Cinara* species, which was previously determined to be a fixed symbiont in at least *C. curtihirsuta* and possibly *C. curvipes* (Meseguer *et al.*, 2017). These *Hamiltonella* strains show highly syntenic reduced genomes of around 1.4 Mbp, contrasting the 2.2 Mbp genomes of the *Hamiltonella defensa* facultative symbionts of *Ac. pisum* (Chevignon *et al.*, 2018). Based on 16S, their closes relatives are *H. defensa* strains ZA17, AS3, and A2C (99% identity, 2737 bit-score). Additionally, a multi-gene phylogenetic reconstruction following Chevignon *et al.* (2018) shows that these *Hamiltonella* symbionts are nested within *H. defensa* symbionts (supplementary fig. S4, Supplementary Material online), confidently assigning these newly sequenced symbionts within the taxon. Based on a preliminary genome annotation of the assembled scaffolds, both *Hamiltonella* symbionts have around 940 CDSs, 50 possible pseudogenes, 38 tRNAs, 1 tmRNA, two copies of the rRNA operon, and 6 other ncRNAs (supplementary table S7, Supplementary Material online). The *Hamiltonella* symbionts associated with *Cinara* and *Erwinia* show a drastically reduced genetic repertoire, namely of CDSs and ncRNAs. Regarding EAAs, vitamins, and co-factors, they retain the ability to code for L-threonine, pyridoxal 5'-P, riboflavin, and lipoic acid. Compared to the facultative *H. defensa* symbionts, it has lost the ability to synthesise biotin, having lost the *bioA* gene and undergoing pseudogenisation of the *bioC*, *bioH*, *bioF*, *bioD*, and *bioB* genes. Also, they lack the typical APSE phage found in facultative *H. defensa* strains.

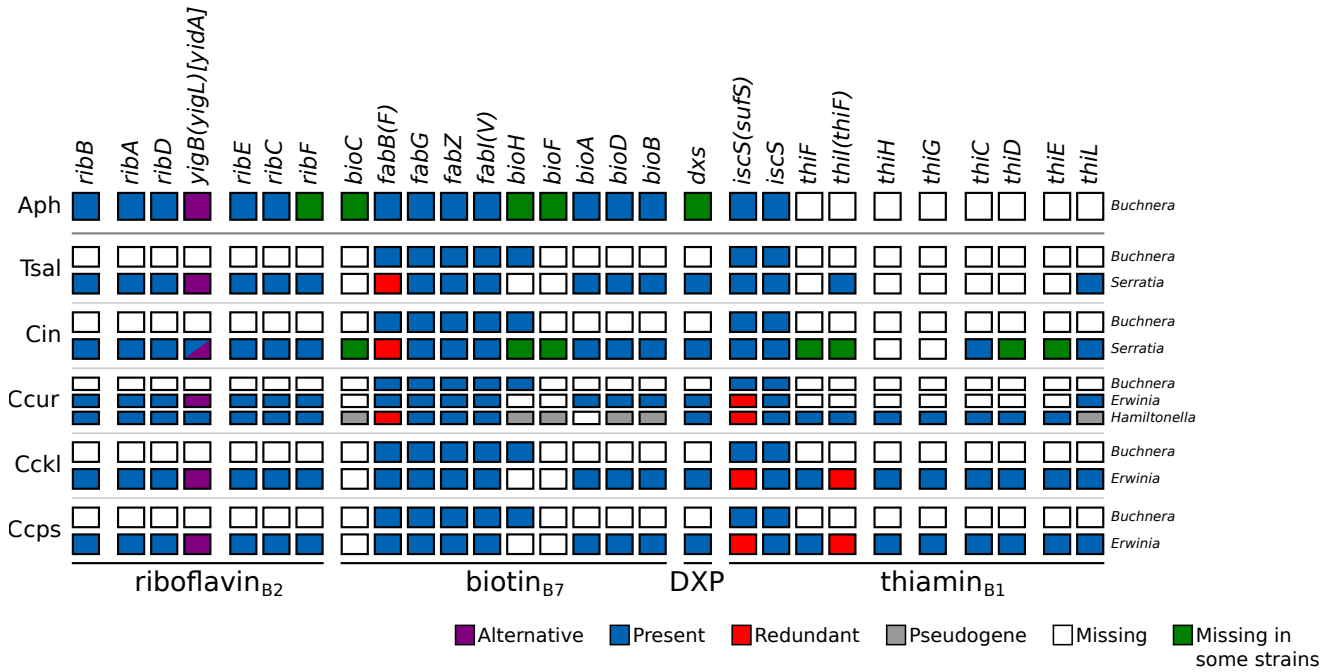

Figure 3. B-vitamin and cofactor biosynthetic metabolic complementation of obligate symbiotic consortia of different aphid species. Diagram summarising the metabolic complementarity of the fixed endosymbiotic consortia of co-obligate symbiotic systems of analysed *Erwinia*-associated *Cinara* aphids. For comparison, a collapsed representation of Aphididae *Buchnera*-only and Lachninae *Buchnera-Se. symbiotica* systems are used as outgroups. The names of genes coding for enzymes involved in the biosynthetic pathway are used as column names. Each row's boxes represent the genes coded by a symbiont's genome. At the right of each row, the genus for the corresponding symbiont. Abbreviations for the collapsed group of aphids harbouring the symbionts is shown at the left of each group of rows and goes as follows. Aph= Aphididae, Tsal= *T. salignus*, Cct= *C. cedri* + *C. tujafilina*, Ccur= *C. curtihirsuta* + *C. curvipes*, Cckl= *C. cuneomaculata* + *C. kochiana* + *C. laricifoliae*, Cpps= *C. pseudotaxifoliae* + *C. pseudotsugae* + *C. splendens* + *C.* cf. *splendens/pseudotsugae*. On the bottom, lines underlining the genes involved in the pathway leading to the compound specified by the name underneath the line.

Serial horizontal gene transfer underlies multi-parter co-obligate mutualistic association

During the manual curation of the *Erwinia* symbiont plasmids, we noted that both the bion- and thiamine-biosynthetic genes (*bioA*, *bioB*, *thiC*, *thiE*, *thiF*, *thiS*, *thiG*, *thiH*, and *thiD*) consistently showed top **BLASTP** hits against *Sodalis* and *Sodalis*-like bacteria, hinting at HGT. Surprisingly, the thiamine-biosynthetic genes present in the *Hamiltonella*'s symbiont genome also showed best BLAST hits against *Sodalis*-related bacteria. To test for HGT of these vitamin-biosynthetic genes, we ran **BLASTP** similarity searches of the proteins of *Erwinia* symbionts *vs.* a database built from the proteomes of *Erwinia* and *Sodalis* species. The search and manual analysis revealed 11 proteins being from putative HGT origin from *Sodalis*-related bacteria (*bioA*, *bioD*, *bioB*, *thiC*, *thiE*, *thiF*, *thiS*, *thiG*, *thiH*, *thiI*, and *thiD*). To test this, we collected orthologous genes across different enterobacterial species and reconstructed Bayesian phylogenies (fig. 4). All 11 genes showed strong support for a single event of HGT for both *Erwinia* and *Hamiltonella* symbionts, consistently being recovered as a monophyletic group nested within or sister to *Sodalis* spp. This is in clear contrasts with what is observed for putative non-HGT *gpmA* and *nupC* genes, where they are confidently recovered nested within *Erwinia* spp. (supplementary fig. S5, Supplementary Material online). Additionally, the majority of the genes' subtrees are congruent with the topology of the *Erwinia* symbionts' subtree (fig. 4: Bottom right; supplementary fig. S6, Supplementary Material online).

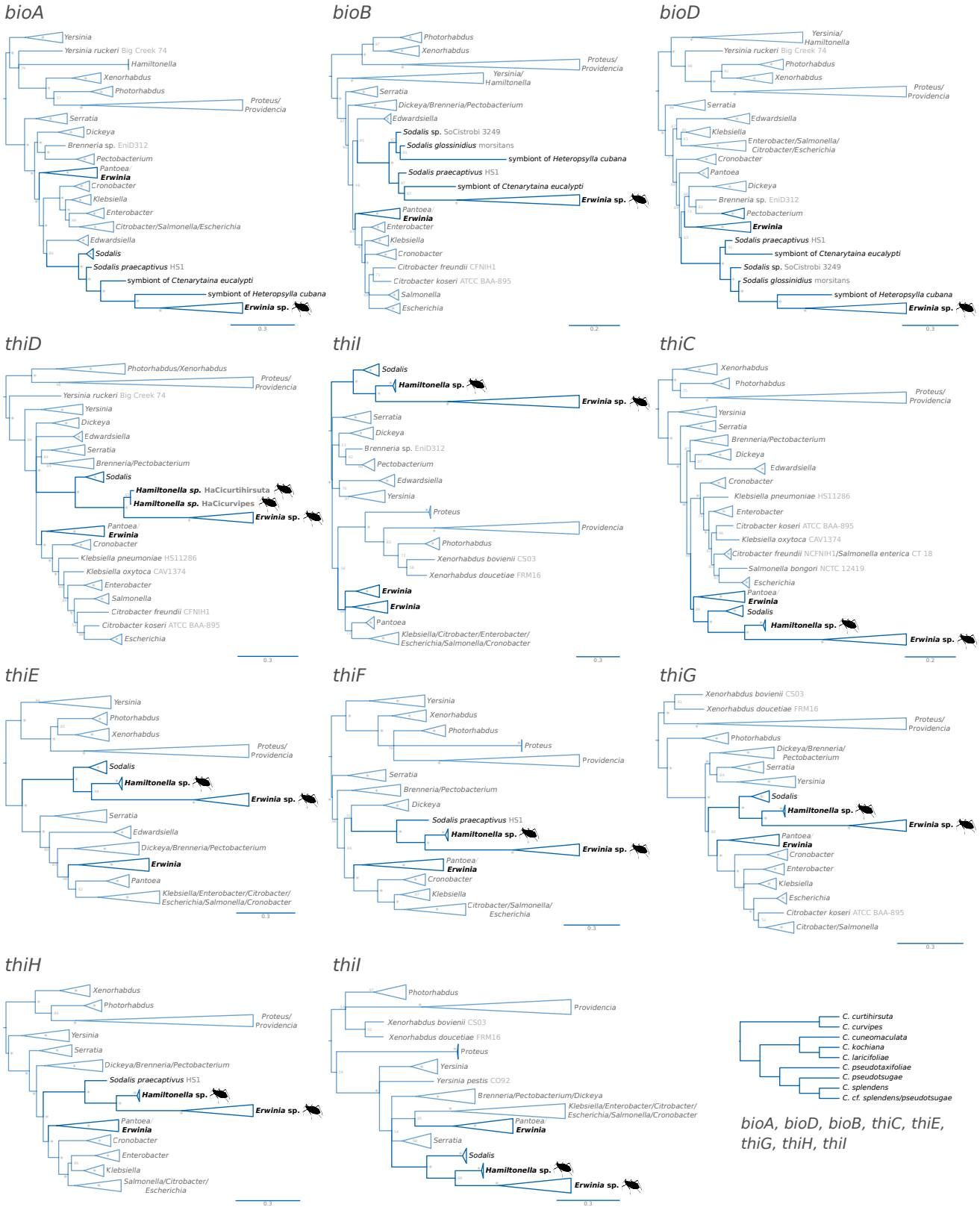

Figure 4. Phylograms of putatively horizontally transferred genes in *Erwinia* and *Hamiltonella* symbionts. Bayesian phylograms of the identified putative horizontally transferred genes present in the symbionts of the newly sequenced *Cinara* species. Taxon labels show the strain name in grey. Asterisks at nodes stand for a posterior probability equal to 1. On the bottom right, cladogram showing the relationships of genes congruent with the *Erwinia* phylogenetic tree.

To further explore the evolutionary history of the horizontally transferred genes, we analysed the location and gene order of them in the plasmids of *Erwinia* symbionts and the genomes of *Hamiltonella* symbionts and the free-living *Sodalis praecaptivus* (fig. 5). We chose this *Sodalis* strain since it is the only currently sequenced member of its genus that preserves all thiamine-biosynthetic genes. Regarding the *Erwinia* plasmids, while they are very well conserved in terms of gene content, they display important changes in genome architecture. We observed that the symbionts of *C. cuneomaculata* and those of *C. pseudotaxifoliae*, *Cinara pseudotsugae*, *Cinara splendens*, and *Cinara* cf. *splendens/pseudotsugae* show the same gene order. However, within the monophyletic clade made up of the symbionts of *C. kochiana* and *Cinara laricifoliae*, we observed an ancestral duplication of the full gene repertoire. This has lead to a plasmid with one intact and one pseudogenised copy of each gene in *C. kochiana* (except for the heat shock protein) and a new gene order in *C. laricifoliae*. All *Erwinia* symbionts belonging to the larger clade comprised of the aforementioned *Cinara* species show both *thiI* and *bioD* genes inserted into the same location in the chromosome. Within the monophyletic cluster made of *C. curtihirsuta* and *C. curvipes*, the *Erwinia* symbionts show a much different plasmid architecture than those of their sister clade. Their plasmid shows an inverted duplication which includes the *bioA* and *bioB* genes as well as the genes coding for the PLD-like domain-containing protein, a putative heat shock protein, a replication-associated protein (*repA2*), and the Tn3 family resolvase/invertase (a mobile element). The presence of this last is puzzling, given that the other sequenced *Erwinia* symbionts completely lack mobile elements. A Bayesian phylogenetic analyses revealed this Tn3 family resolvase/invertase is most closely related to Tn3 resolvase/invertases encoded by other *H. defensa* symbionts (supplementary fig. S7, Supplementary Material online), hinting at their origin. As mentioned before, the *Erwinia* symbionts of *C. curtihirsuta* and *C. curvipes* do not host the "typical" thiamine-biosynthetic genes in their plasmid. In turn, these genes are now hosted in the *Hamiltonella* genome, flanked by insertion sequence (hereafter **IS**) elements and a glycerol dehydrogenase pseudogenes. Weather or not these genes reside in a plasmid or not remains unclear. However, the scaffolds in which these genes reside in each *Cinara* species has roughly double the coverage as the rest of the large scaffolds. When comparing the genomic context of the horizontally transferred genes against that of the same genes in *So. praecaptivus*, several rearrangements have taken place. Namely, the *thiD* gene is now hosted by the aphid symbionts in close proximity of the rest of the "thi" genes and the *bioD* gene has been distanced from the other biotin-biosynthetic genes.

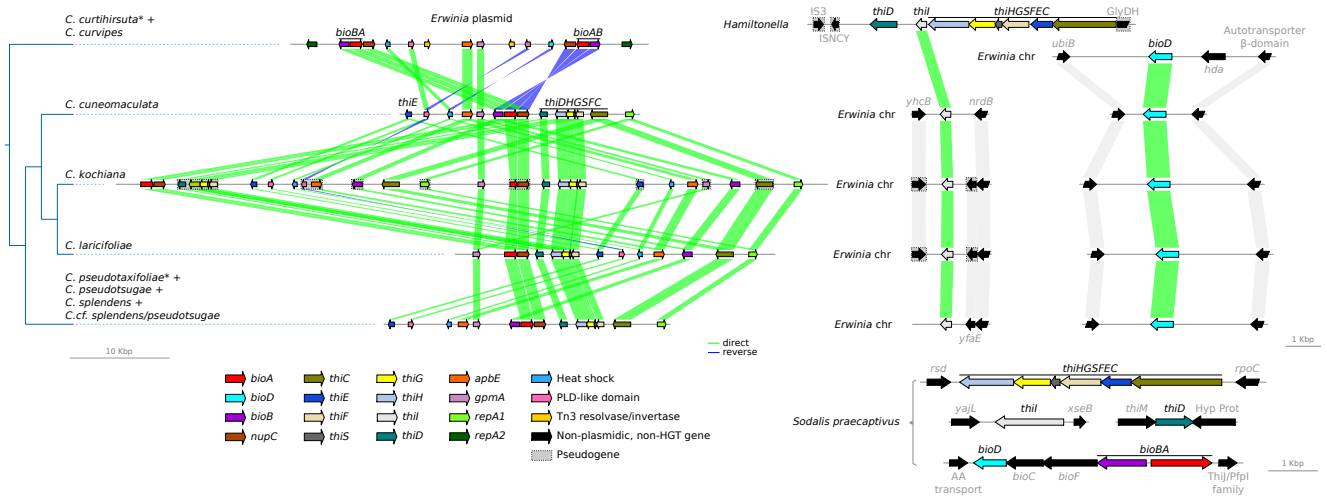

Figure 5. Evolution of the horizontally-transferred biotin- and thiamine-biosynthetic genes. Linear plots of the genomic context for the genes originating from a HGT from *Sodalis*-related bacteria. On the left, dendogram showing the phylogenetic relations of the different *Erwinia*-associated *Cinara* groups of species. On the bottom-right, genomic context for the aforementioned genes in *So. praecaptivus* for comparison.

'*Candidatus* Erwinia haradaeae' sp. nov.

We propose the specific name '*Candidatus* Erwinia haradaeae' for the monophyletic lineage of enterobacterial endosymbionts exclusively found affiliated as co-obligate symbionts in the monophyletic group of *Cinara* aphid species analysed in this paper, although its presence in other closely related species cannot be discarded. The closest relative of the symbiont of *C. pseudotaxifoliae* by 16S rRNA gene sequence identity is the *Erwinia pyrifoliae* strain EpK1/15 (INSDC accession number CP023567.1) with which it shares 94 % sequence identity (2350 bit-score). The specific epithet 'haradaeae' alludes to Hosami Harada, who performed research suggesting the origin of the aphid's obligate *Buchnera* symbiont as derived from a habitant of the plant on which the host fed (Harada and Ishikawa, 1993; Harada *et al.*, 1996). These studies put forward the idea of the aphids' symbionts possibly evolving from bacteria originally inhabiting the plant, then transitioning to gut associates, and finally obligate symbionts.

In *C. cuneomaculata*, *C. kochiana*, and *C. curvipes*, '*Ca.* Erwinia haradaeae' is found inhabiting the bacteriome inside bacteriocytes distinct from those of *Buchnera*. Previous work referring to these symbionts as '*Erwinia*' or '*Erwinia*-related' include Jousselin *et al.* (2016) and Meseguer *et al.* (2017). Currently available sequences that correspond to '*Ca.* Erwinia haradaeae' are deposited under INSDC accession numbers LT670851.1 and LT670852.1. All currently known '*Ca.* Erwinia haradaeae' species harbour biotin- and thiamine-biosynthetic horizontally transferred genes from a *Sodalis* or *Sodalis*-related bacteria both in their chromosome and plasmid. Based on genome-based metabolic inference and their parallel evolutionary history with *Buchnera*, they all are co-obligate endosymbionts along with *Buchnera*, and *Ca.* Hamiltonella defensa in *C. curtihirsuta* and *C. curvipes*, in the *Cinara* species sequenced and presented in this study.

Discussion

Obligate nutritional microbial endosymbioses are among the most integrated forms of beneficial associations between eukaryotes and bacteria. Within these, the relationships that arise between animals with nutrient-restricted diets and their nutritional symbionts provide valuable data for studying these emergent associations. Of particular interest to understand the evolutionary steps leading to these associations (where the partners life-habits, fitness and life cycles are completely intertwine), are those partnerships that have evolved recently and for which closely related free-living or facultative endosymbiotic lineages are available.

In this work we have analysed the recently evolved co-obligate association between *Buchnera* and *Ca.* Erwinia haradaeae (hereafter *E. haradaeae*) by sequencing nine symbiont pairs hosted by a monophyletic group of *Cinara* aphids (Meseguer *et al.*, 2017). Through phylogenetic and genome rearrangement analyses of *Erwinia* genomes, we have determined that these symbionts most likely arose, similarly to *Buchnera*, from a single event of host fixation and have been co-diverging with their *Buchnera* partners ever since. Given the age of the aphid clade hosting these symbionts, as estimated by Meseguer *et al.* (2015), these new *Erwinia* obligate mutualistic parterns have co-diverged with their aphid hosts and associated *Buchnera* for about 20-25 million years. Like their co-obligate *Buchnera* partner, *E. haradaeae* strains generally lack mobile elements and all share a very similar genetic repertoire, with around 90% of their gene content being shared across all strains. Additionally, *E. haradaeae* symbionts share a similar localisation and cell shape across different *Cinara* species. This provides strong support for an evolutionary history involving an early drastic genome reduction and loss of mobile elements before the diversification of their *Cinara* hosts' group, as suggested for other obligate mutualistic association (Pérez-Brocal *et al.*, 2006; Williams and Wernegreen, 2015). Regarding their evolutionary origin, we found that *E. haradaeae* is most closely related to a group of *Erwinia* species that have been isolated as both as pathogenic and non-pathogenic plant symbionts. *Erwinia* are generally described as as plant-associated bacteria (Kado, 2006). Nonetheless, *Erwinia* species have been isolated from insects, namely the olive fruit fly (*Bactrocera oleae*) (Capuzzo, 2005), the western flower thrips (*Frankliniella occidentalis*) (de Vries *et al.*, 2001) and aphids (Fischer-Le Saux *et al.*, 2015; Harada and Ishikawa, 1993; Harada *et al.*, 1996, 1997). Concerning the latter, *Erwinia* was isolated both from artificial diets exposed to probing of *Diuraphis noxia* biotype 2 sylets through a Parafilm

membrane (*Erwinia iniecta*) and from the gut of *Ac. pisum* (*Erwinia aphidicola*). The existence of this bacteria within the aphid digestive tract is supported by an early experiment demonstrating that the aphid *Aphis pomi* was capable of ingesting a pathogenic *Erwinia amylovora* after a short feeding period (5 minutes) and this would persist in the bodies of the insects for at least 72 hours (duration of the longest experiment) (Plurad *et al.*, 1965). Thus, aphid feeding is a likely source for the presence and persistence of *Erwinia* bacteria inside the aphid digestive tract. It is important to note that the aforementioned *Erwinia* symbionts described for *F. occidentalis* (Facey *et al.*, 2015) (INSDC accessions JMSO00000000.1, LAGQ00000000.1, and LAGP00000000.1) belong to the *E. aphidicola* species (AB681773.1; type strain X 001=CIP 106296=IAM 14479=JCM 21238=LMG 24877=NBRC 102417), based on 16S sequence identity (between 99.72-100.00%). This indicates that strains from these bacterial species can actually infect very distantly related insect species such as aphids and thrips.

As suggested in Meseguer *et al.* (2017), *E. haradaeae* symbionts have evolved a very compact genome that has lost the capacity to synthesise any essential amino acid, but have retained routes for complementing *Buchnera* in regards to the biosynthesis of namely two B vitamins: biotin and riboflavin. This pattern is consistent with what is observed in another two independently-evolved *Buchnera-Se. symbiotica* co-obligate systems found Lachninae aphids (Manzano-Marín *et al.*, 2016). Unlike these previously analysed nutritional symbiotc systems, we found that the capacity to synthesise thiamine, another essential B vitamin, was conserved in all newly sequenced symbiotic consortia. Interestingly, Ehrhardt (1968) found that in *Neomyzus circumflexus* aphids fed on an artificial diet missing thiamine, mortality in the first generation was normal at around 5%. The adults were found to be much smaller and gave 54% less larvae than the control animals. Additionally, the development of the 2nd generation was stagnating, being the larvae stuck in the 2nd stage after 20 days. These results support the essential nature of this vitamin and suggest that the diet from the *E. haradaeae*-associated *Cinara* might have a lower concentration of thiamine than that of the previously sequenced aphid symbiotic systems. To our knowledge, no studies have looked at sap vitamin contents in the conifer genera on which this monophyletic group of aphids feed (i.e. mainly *Pseudotsuga*, *Larix*, *Abies*, and a few *Picea* species; see Jousselin *et al.* 2013 and Meseguer *et al.* 2015 for a study of aphid host plant association in this clade).

Most strikingly, we found strong evidence supporting the horizontal transfer of not only three biotin-biosynthetic genes (*bioA*, *bioD*, and *bioB*), but also most thiamine-biosynthetic ones (except for *thiL*). Through phylogenetic analyses, we were able to confidently determine that the likely origin of the HGT genes in *Erwinia* and *Hamiltonella* was a *Sodalis* or *Sodalis*-related bacterium. *Sodalis*-allied bacteria are found across different insect taxa as facultative or obligate endosymbionts (Koga and Moran, 2014; Oakeson *et al.*, 2014; Santos-Garcia *et al.*, 2017; Sloan and Moran, 2012) as well as, likely, plant tissues (Chari *et al.*, 2015). Furthermore, bacteria from these taxa have been previously identified as a bacterium associated to aphids (Manzano-Marín *et al.*, 2017, 2018; Meseguer *et al.*, 2017). Therefore, it is likely that the co-inhabitation of *Erwinia* and a *Sodalis*-allied bacterium facilitated the Horizontal transfer of these genes to the *Erwinia* plasmid. HGT involving genes in vitamin biosynthetic pathways has been previously documented in *Candidatus* Legionella polyplacis, the endosymbiont of the blood-feeding louse *Polyplax serrata* (Říhová *et al.*, 2017). This *Legionella* symbiont has acquired a complete biotin operon (*bioADCHFB*) likely from a rickettsial bacterium. A similar example is observed in the *Cardinium* endosymbionts of the parasitoid wasp *Encarsia pergandiella* (Penz *et al.*, 2012) and the whitefly *Bemisia tabaci* (Říhová *et al.*, 2017).

While most *E. haradaeae* retain the thiamin-biosynthetic genes, the monophyletic group of *Erwinia* symbionts from *C. curtihirsuta* and *C. curvipes* have lost all horizontally-transferred thiamine-biosynthetic genes. Our results suggest that these genes have been further horizontally-transferred to a tertiary co-obligate *H. defensa* symbiont, thus locking a new tri-symbiotic co-obligate nutritional symbiosis. The handing on of the thiamine-biosynthetic genes from *Erwinia* to yet a newcomer, *Hamiltonella*, that has now become obligate, further supports the hypothesis that these genes have a functional significance for the mutualistic association between aphids and their consortium of bacteria. Although Meseguer *et al.* (2017) suggested that the replacement of co-obligate symbionts in *Cinara* might not be adaptive for the aphid hosts (they do not correlate with broad ecological niche shifts of aphids), the fact these new symbionts carry and retain genes with new metabolic functions might suggest otherwise. The shift in ecological niche enabled by the replacement of symbionts alongside *Buchnera*, might be

detectable at a finer scale than the one explored by Meseguer *et al.* (2017).

Our analyses allow us to propose an evolutionary scenario for the *Buchnera-Erwinia* co-obligate symbiotic consortium in *Cinara* aphids (fig. 6). First, the *Erwinia* bacterium associated with aphids, likely originating from a plant-affiliated symbiont, replaced the previously existing secondary co-obligate symbiont (*Se. symbiotica*, according to Meseguer *et al.*, 2017) and underwent a drastic genomic reduction before the diversification of the group. The ability to replace the former co-obligate symbiont was facilitated by the presence of horizontally transferred biotin- and thiamine-biosynthetic genes in its plasmid from a co-existing *Sodalis* symbiont. Subsequently, the gene *bioD* was transferred into the chromosome of *E. haradaeae*. Afterwards, in the lineage leading to *C. curtihirsuta* and *C. curvipes*, *E. haradaeae* lost the capacity to synthesise thiamine due to the loss of the *thiCEFSGHID* cluster of genes. Nonetheless, this loss was accompanied by the acquisition by HGT of the aforementioned genes from *E. haradaeae* to the co-existing *H. defensa*. This second HGT event locked the tri-symbiotic system together. In the meantime, in the lineage leading to the sister clade of *Hamiltonella*-associated *Cinara* aphids, *E. haradaeae*'s genome underwent a further event of plasmid to chromosome transfer of the gene of HGT origin *thiI*.

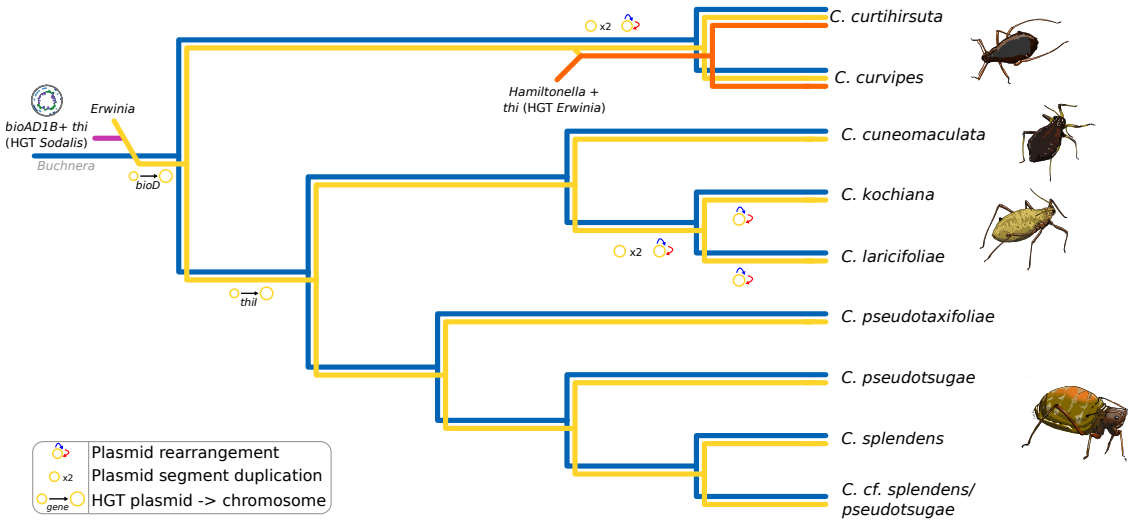

Figure 6. Proposed evolutionary scenario for the establishment and evolution of co-obligate symbionts of *E. haradaeae*-associated *Cinara*. Cladogram displaying the relationships of sequenced *E. haradaeae*-associated *Cinara* lineages. Coloured lines in the dendogram are used to represent the persistence of a symbiont in the aphid lineage. Incoming lines on branches symbolise the acquisition of co-obligate secondary symbionts. Lines joining two symbiotic lineages symbolise HGT events. At the leaves, cartoons of selected aphids form the different groups of genera are shown.

Conclusion

Symbiont replacement is a phenomenon observed in several obligate symbiotic systems (Husnik and McCutcheon, 2016; Koga and Moran, 2014; Toenshoff *et al.*, 2012). The presented results reveal the crucial role that horizontal transfer of pivotal metabolic genes across co-existing symbionts can have in the emergence of new obligate associations between insects and bacterial species. First explored by Harada *et al.* (1996), the current work provides support for the bacteria present in the aphid's diet as being a likely source of obligate symbiotic bacteria. Finally, this work also raises new questions about the role of these new symbionts in aphids' ecological niches: our genomic enquiry suggests that thiamine biosynthesis genes might be pivotal in the establishment of both *Erwinia* and then *Hamiltonella*. We expect that further exploration of the intricacies of the co-obligate di-symbiotic *Buchnera-Erwinia* and tri-symbiotic *Buchnera-Erwinia-Hamiltonella* systems will continue to provide important clues into the emergence and maintenance of these complex endosymbiotic consortia as well as on the process of genome re-organisation and evolution in endosymbiotic bacteria.

Acknowledgements

We would like to acknowledge the talented artist/scientist Jorge Mariano Collantes Alegre for the aphid cartoons in fig. 6. 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 *Agropolis foundation/Labex Agro* ("*Cinara*'s microbiome") to E.J, the the *France Génomique* National Infrastructure, funded as part of the *Investissemnt d'Avenir* program managed by the *Agence Nationale pour la Recherche* (ANR-10-INBS-09) to C.O, 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). 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Akman Gündüz E and Douglas A. E. 2009. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proc R Soc Lond B Biol Sci*, 276(1658): 987–91. **URL** http://rspb.royalsocietypublishing.org/content/276/1658/987.long.

Altschul S. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25(17): 3389–3402. **URL** http://nar.oxfordjournals.org/lookup/doi/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. *J Comput Biol*, 19(5): 455–77. **URL** http://online.liebertpub.com/doi/abs/10.1089/cmb.2012.0021.

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** http://www.annualreviews.org/doi/abs/10.1146/annurev.mi.49.100195.000415.

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. *Biol Direct*, 7(1): 12. **URL** http://biologydirect.biomedcentral.com/articles/10.1186/1745-6150-7-12.

Buchner P. 1953. *Endosymbiose der Tiere mit Pflanzlichen Mikroorganismen*. Birkhäuser Basel, Basel. **URL** http://link.springer.com/10.1007/978-3-0348-6958-4.

Capuzzo C. 2005. '*Candidatus* Erwinia dacicola', a coevolved symbiotic bacterium of the olive fly *Bactrocera oleae* (Gmelin). *Ineternational Journal of Systematic and Evolutionary Microbiology*, 55(4): 1641–1647. **URL** http://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.63653-0.

Chari A, Oakeson K. F, Enomoto S, Jackson D. G, Fisher M. A, and Dale C. 2015. Phenotypic characterization of *Sodalis praecaptivus* sp. nov., a close non-insect-associated member of the *Sodalis*-allied lineage of insect endosymbionts. *International journal of systematic and evolutionary microbiology*, 65(5): 1400–1405. **URL** http://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.000091.

Charles H and Ishikawa H. 1999. Physical and Genetic Map of the Genome of *Buchnera*, the Primary Endosymbiont of the Pea Aphid *Acyrthosiphon pisum*. *J Mol Evol*, 48(2): 142–150. **URL** http://link.springer.com/10.1007/PL00006452.

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. kondoi*. *Entomologia Experimentalis et Applicata*, 95(3): 315–323. **URL** http://doi.wiley.

com/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** http://dx.plos.org/10.1371/journal.pone.0000383.

Chevignon G, Boyd B. M, Brandt J. W, Oliver K. M, and Strand M. R. 2018. Culture-Facilitated Comparative Genomics of the Facultative Symbiont *Hamiltonella defensa*. *Genome Biology and Evolution*, 10(3): 786–802. **URL** https://academic.oup.com/gbe/advance-article/doi/10.1093/gbe/evy036/4857210.

de Vries E. J, Breeuwer J. A, Jacobs G, and Mollema C. 2001. The Association of Western Flower Thrips, *Frankliniella occidentalis*, with a near *Erwinia* Species Gut Bacteria: Transient or Permanent? *Journal of Invertebrate Pathology*, 77(2): 120–128. **URL** http://linkinghub.elsevier.com/retrieve/pii/S0022201101950098.

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

Ehrhardt P. 1968. Der Vitaminbedarf einer siebröhrensaugenden Aphide, *Neomyzus circumflexus* Buckt. (Homoptera, Insecta). *Zeitschrift für Vergleichende Physiologie*, 60(4): 416–426. **URL** http://link.springer.com/10.1007/BF00297936.

Facey P. D, Méric G, Hitchings M. D, Pachebat J. A, Hegarty M. J, Chen X, Morgan L. V, Hoeppner J. E, Whitten M. M, Kirk W. D, Dyson P. J, Sheppard S. K, and Sol R. D. 2015. Draft Genomes, Phylogenetic Reconstruction, and Comparative Genomics of Two Novel Cohabiting Bacterial Symbionts Isolated from *Frankliniella occidentalis*. *Genome Biology and Evolution*, 7(8): 2188–2202. **URL** https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evv136.

Fischer-Le Saux M, Luna E, Lapitan N, Portier P, Tisserat N. A, Campillo T, and Leach J. E. 2015. *Erwinia iniecta* sp. nov., isolated from Russian wheat aphid (*Diuraphis noxia*). *International Journal of Systematic and Evolutionary Microbiology*, 65(10): 3625–3633. **URL** http://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.000466.

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** http://link.springer.com/10.1007/BF00556361.

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** http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2000.1314.

Griffiths G. W and Beck S. D. 1975. Ultrastructure of pea aphid mycetocytes: Evidence for symbiote secretion. *Cell Tissue Res*, 159(3): 351–367. **URL** http://link.springer.com/10.1007/BF00221782.

Guay J.-F, Boudreault S, Michaud D, and Cloutier C. 2009. Impact of environmental stress on aphid clonal resistance to parasitoids: Role of *Hamiltonella defensa* bacterial symbiosis in association with a new facultative symbiont of the pea aphid. *Journal of Insect Physiology*, 55(10): 919–926. **URL** http://linkinghub.elsevier.com/retrieve/pii/S0022191009002364.

Hansen A. K, Vorburger C, and Moran N. A. 2012. Genomic basis of endosymbiont-conferred protection against an insect parasitoid. *Genome Research*, 22(1): 106–114. **URL** http://genome.cshlp.org/cgi/doi/10.1101/gr.125351.111.

Harada H and Ishikawa H. 1993. Gut microbe of aphid closely related to its intracellular symbiont. *Biosystems*, 31(2-3): 185–191. **URL** http://linkinghub.elsevier.com/retrieve/pii/030326479390048H.

Harada H, Oyaizu H, and Ishikawa H. 1996. A consideration about the origin of aphid intracellular symbiont in connection with gut bacterial flora. *The Journal of General and Applied Microbiology*, 42(1): 17–26. **URL** http://joi.jlc.jst.go.jp/JST.Journalarchive/jgam1955/42.17?from=CrossRef.

Harada H, Oyaizu H, Kosako Y, and Ishikawa H. 1997. *Erwinia aphidicola*, a new species isolated from pea aphid, Acyrthosiphon pisum. *The*

Journal of General and Applied Microbiology, 43(6): 349–354. **URL** http://joi.jlc.jst.go.jp/JST.Journalarchive/jgam1955/43.349?from=CrossRef.

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** http://doi.wiley.com/10.1111/jeb.12705.

Hoang D. T, Chernomor O, von Haeseler A, Minh B. Q, and Vinh L. S. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2): 518–522. **URL** https://academic.oup.com/mbe/article/35/2/518/4565479.

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** http://www.pnas.org/lookup/doi/10.1073/pnas.1603910113.

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** http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2008.0679.

Jousselin E, Cruaud A, Genson G, Chevenet F, Foottit R. G, and Cœur d'acier A. 2013. Is ecological speciation a major trend in aphids? Insights from a molecular phylogeny of the conifer-feeding genus *Cinara*. *Frontiers in Zoology*, 10(1): 56. **URL** http://frontiersinzoology.biomedcentral.com/articles/10.1186/1742-9994-10-56.

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. *Mol Ecol Resour*, 16(3): 628–640. **URL** http://doi.wiley.com/10.1111/1755-0998.12478.

Kado C. I. 2006. *Erwinia* and Related Genera. In E. Dworkin, Martin and Falkow, Stanley and Rosenberg, Eugene and Schleifer, Karl-Heinz and Stackebrandt, editor, *The Prokaryotes. Volume 6: Proteobacteria: Gamma Subclass*, chapter 3, pages 443–450. Springer New York, 3rd edition. **URL** https://link.springer.com/referenceworkentry/10.1007/0-387-30746-X{_}15.

Katoh K and Standley D. M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol*, 30(4): 772–780. **URL** http://mbe.oxfordjournals.org/cgi/doi/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** http://dx.doi.org/10.1038/ismej.2013.235.

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** http://joi.jlc.jst.go.jp/JST.JSTAGE/aez/2009.281?from=CrossRef.

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

Li L. 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Res*, 13(9): 2178–2189. **URL** http://www.genome.org/cgi/doi/10.1101/gr.1224503.

Lowe T. M and Eddy S. R. 1997. tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Research*, 25(5): 955–964. **URL** http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/25.5.0955.

Manzano-Marín A and Latorre A. 2016. Snapshots of a shrinking partner: Genome reduction in *Serratia symbiotica*. *Sci Rep*, 6: 32590. **URL** http://www.nature.com/articles/srep32590.

Manzano-Marín A, Oceguera-Figueroa A, Latorre A, Jiménez-García L. F, and Moya A. 2015. Solving a Bloody Mess: B-Vitamin Independent Metabolic Convergence among Gammaproteobacterial Obligate Endosymbionts from Blood-Feeding Arthropods and the

Leech *Haementeria officinalis*. *Genome Biol Evol*, 7(10): 2871–2884. **URL** http://gbe.oxfordjournals.org/lookup/doi/10.1093/gbe/evv188.

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 Biol Evol*, 8(5): 1440–1458. **URL** http://gbe.oxfordjournals.org/lookup/doi/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. *Environ Microbiol*, 19(1): 393–408. **URL** http://doi.wiley.com/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** http://gbe.oxfordjournals.org/lookup/doi/10.1093/gbe/evy173.

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** http://www.pnas.org/lookup/doi/10.1073/pnas.1803245115.

Meseguer A. S, Coeur d'acier A, Genson G, and Jousselin E. 2015. Unravelling the historical biogeography and diversification dynamics of a highly diverse conifer-feeding aphid genus. *Journal of Biogeography*, 88(8): 1482–1492. **URL** http://doi.wiley.com/10.1111/jbi.12531.

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. *Mol Ecol*, 26(8): 2363–2378. **URL** http://doi.wiley.com/10.1111/mec.13910.

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** http://dx.doi.org/10.1046/j.1365-2311.2002.00393.x.

Munson M. A, Baumann P, and Kinsey M. G. 1991. *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a Taxon Consisting of the Mycetocyte-Associated, Primary Endosymbionts of Aphids. *Int J Syst Bacteriol*, 41(4): 566–568. **URL** http://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-41-4-566.

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

Nawrocki E. P, Burge S. W, Bateman A, Daub J, Eberhardt R. Y, Eddy S. R, Floden E. W, Gardner P. P, Jones T. A, Tate J, and Finn R. D. 2015. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res*, 43(D1): D130–D137. **URL** http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gku1063.

Nguyen L.-T, Schmidt H. A, von Haeseler A, and Minh B. Q. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, 32(1): 268–274. **URL** https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msu300.

Oakeson K. F, Gil R, Clayton A. L, Dunn D. M, Von Niederhausern A. C, Hamil C, Aoyagi A, Duval B, Baca A, Silva F. J, Vallier A, Jackson D. G, Latorre A, Weiss R. B, Heddi A, Moya A, and Dale C. 2014. Genome degeneration and adaptation in a nascent stage of symbiosis. *Genome Biology and Evolution*, 6(1): 76–93. **URL** http://gbe.oxfordjournals.org/cgi/doi/10.1093/gbe/evt210.

Okonechnikov K, Golosova O, and Fursov M. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, 28(8): 1166–1167. **URL** http://bioinformatics.oxfordjournals.org/cgi/doi/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** http://www.pnas.org/cgi/doi/10.1073/pnas.0335320100.

Oliver K. M, Degnan P. H, Burke G. R, and Moran N. A. 2010. Facultative Symbionts in Aphids and the Horizontal Transfer of Ecologically Important Traits. *Annual Review of Entomology*, 55(1): 247–266. **URL** http://www.annualreviews.org/doi/abs/10.1146/annurev-ento-112408-085305.

Penz T, Schmitz-Esser S, Kelly S. E, Cass B. N, Müller A, Woyke T, Malfatti S. A, Hunter M. S, and Horn M. 2012. Comparative Genomics Suggests an Independent Origin of Cytoplasmic Incompatibility in *Cardinium hertigii*. *PLoS Genetics*, 8(10): e1003012. **URL** http://dx.plos.org/10.1371/journal.pgen.1003012.

Pérez-Brocal V, Gil R, Ramos S, Lamelas A, Postigo M, Michelena J. M, Silva F. J, Moya A, and Latorre A. 2006. A small microbial genome: The end of a long symbiotic relationship? *Science*, 314(5797): 312–313. **URL** http://www.sciencemag.org/cgi/doi/10.1126/science.1130441.

Plurad S. B, Goodman R. N, and Enns W. R. 1965. Persistence of *Erwinia amylovora* in the Apple Aphid (*Aphis pomi* DeGeer), a Probable Vector. *Nature*, 205(4967): 206–206. **URL** http://www.nature.com/articles/205206a0.

Pyka-Fościak G and Szklarzewicz T. 2008. Germ cell cluster formation and ovariole structure in viviparous and oviparous generations of the aphid *Stomaphis quercus*. *The International Journal of Developmental Biology*, 52(2-3): 259–265. **URL** http://www.intjdevbiol.com/paper.php?doi=072338gp.

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. *Syst Biol*, 61(3): 539–542. **URL** http://sysbio.oxfordjournals.org/cgi/doi/10.1093/sysbio/sys029.

Sandstrom J and Moran N. 1999. How nutritionally imbalanced is phloem sap for aphids? *Entomol Exp Appl*, 91(1): 203–210. **URL** http://doi.wiley.com/10.1046/j.1570-7458.1999.00485.x.

Santos-Garcia D, Latorre A, Moya A, Gibbs G, Hartung V, Dettner K, Kuechler S. M, and Silva F. J. 2014. Small but powerful, the primary endosymbiont of moss bugs, *Candidatus* Evansia muelleri, holds a reduced genome with large biosynthetic capabilities. *Genome biology and evolution*, 6(7): 1875–1893. **URL** https://academic.oup.com/gbe/article-lookup/doi/10.1093/gbe/evu149.

Santos-Garcia D, Silva F. J, Morin S, Dettner K, and Kuechler S. M. 2017. The All-Rounder Sodalis: A New Bacteriome-Associated Endosymbiont of the Lygaeoid Bug Henestaris halophilus (Heteroptera: Henestarinae) and a Critical Examination of Its Evolution. *Genome Biology and Evolution*, 9(10): 2893–2910. **URL** http://academic.oup.com/gbe/article/doi/10.1093/gbe/evx202/4237403.

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

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30(14): 2068–2069. **URL** https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu153.

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** http://mbe.oxfordjournals.org/cgi/doi/10.1093/molbev/mss180.

Smith T. A, Driscoll T, Gillespie J. J, and Raghavan R. 2015. A *Coxiella*-Like Endosymbiont Is a Potential Vitamin Source for the Lone Star Tick. *Genome Biology and Evolution*, 7(3): 831–838. **URL** http://gbe.oxfordjournals.org/cgi/doi/10.1093/gbe/evv016.

Suzek B. E, Ermolaeva M. D, Schreiber M, and Salzberg S. L. 2001. A probabilistic method for identifying start codons in bacterial genomes. *Bioinformatics*, 17(12): 1123–1130. **URL** http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/17.12.1123.

Talavera G and Castresana J. 2007. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst Biol*, 56(4): 564–577. **URL** http://sysbio.oxfordjournals.org/cgi/doi/10.1080/10635150701472164.

Tamas I, Klasson L, Canbäck B, Näslund K. A, Eriksson A.-S, Wernegreen J. J, Sandström J. P, Moran N. A, and Andersson S. G. E. 2002. 50 Million Years of Genomic Stasis in Endosymbiotic Bacteria. *Science*, 296(5577): 2376–2379. **URL** http://www.sciencemag.org/cgi/doi/10.1126/science.1071278.

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** http://doi.wiley.com/10.1111/j.1462-2920.2012.02712.x.

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*. *Proc Natl Acad Sci U S A*, 100(2): 581–586. **URL** http://www.pnas.org/cgi/doi/10.1073/pnas.0235981100.

Wallace I. M. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res*, 34(6): 1692–1699. **URL** https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkl091.

Wang H.-C, Minh B. Q, Susko E, and Roger A. J. 2018. Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation. *Systematic Biology*, 67(2): 216–235. **URL** https://academic.oup.com/sysbio/article/67/2/216/4076233.

Williams L. E and Wernegreen J. J. 2015. Genome evolution in an ancient bacteria-ant symbiosis: parallel gene loss among *Blochmannia* spanning the origin of the ant tribe Camponotini. *PeerJ*, 3: e881. **URL** https://peerj.com/articles/881.

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** http://link.springer.com/10.1007/978-3-642-66161-7.

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** http://academic.oup.com/gbe/article/doi/10.1093/gbe/evx217/4562435.