Contrasting genomic trajectories of *Bartonellaceae* symbionts of planthoppers

Mingjie Ma[1,2 *], Anna Michalik[3], Junchen Deng[2,4], Yi Hu[1], Piotr Łukasik[2 *]

1. Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, College of Life Sciences, Beijing Normal University, Beijing 100875, China
2. Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland
3. Department of Developmental Biology and Morphology of Invertebrates, Institute of Zoology and Biomedical Research, Faculty of Biology, Jagiellonian University, Kraków, Poland
4. Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland

* to whom correspondence should be addressed. E-mail: mamingjie@mail.bnu.edu.cn; p.lukasik@uj.edu.pl

Abstract

Symbioses with microorganisms have shaped the nutritional biology and evolution of many insects. For example, several ant clades have adapted to nutrient-poor diets through symbiosis with a specific clade of bacteria in the family Bartonellaceae (Hyphomicrobiales), notorious for also including virulent vertebrate pathogens. Here we show that Bartonellaceae phylogenetically placed within the clade that has only encompassed ant symbionts to date – *Candidatus* genus Tokpelaia – have established as symbionts in four different clades of planthoppers (Insecta: Hemiptera: Fulgoromorpha). Genome size and contents indicate different levels of integration of these strains into the planthopper host biology, and their diverse roles. Symbionts infecting one of the clades have some of the largest genomes among Bartonellaceae, at ca. 2 Mb, two others are under 700 kb, and the fourth is reduced to barely 158 kb. The planthopper-associated *Tokpelaia* strains with larger genomes, similarly to ant symbionts, encode multiple amino acid and vitamin biosynthesis genes, complementing the degraded nutritional capabilities of their planthoppers' ancient heritable endosymbionts. Strikingly, the smallest *Tokpelaia* genome lacks any genes linked to essential amino acid biosynthesis, in contrast to all other known insect-associated bacteria with genomes of comparable size. We identified a single vitamin biosynthesis gene and Fe-S cluster assembly genes as its only putative contributions to the host biology.

Our results broaden the host spectrum of non-pathogenic Bartonellaceae, indicating that they have contributed to nutrition and symbiotic consortium function in diverse diet-restricted host clades. They also highlight an unexpectedly broad range of evolutionary outcomes for this important bacterial group.

Key words: nutritional endosymbiont, amino acid biosynthesis, vitamins, Auchenorrhyncha, Rhizobiales, *Bartonella*

1. Introduction

Microbial symbionts can play dramatically important and highly diverse roles in the biology of their animal hosts, shaping their life history traits, ecological niches, biological interactions, and evolutionary processes (Moran 2007; Douglas 2014; Lukasik and Kolasa 2024). In particular, host long-term evolutionary trajectories have repeatedly been shaped by the microbial contributions to the host nutrition (Moran 2007; Sudakaran, et al. 2017; Brentassi and de la Fuente 2024). Microbial symbionts frequently help address challenges related to the digestion of recalcitrant foods, the neutralization of toxic compounds, and the acquisition of limiting nutrients that animals require but cannot make on their own. For example, plant-based foods typically have low nitrogen contents and specifically lack essential amino acids as well as vitamins, and many herbivores rely on symbiotic bacteria for nitrogen recycling, upgrading, and sometimes fixation, and for the biosynthesis of lacking nutrients (Hansen and Moran 2014). Insects provide many independent examples of associations between heritable symbiont acquisitions and shifts to limiting diets (Cornwallis, et al. 2023).

Ants (Hymenoptera: Formicidae) are often quoted as an example of a clade that has undergone extensive dietary diversification, including multiple independent transitions from more balanced omnivorous diets to herbivory (Russell, et al. 2009; Smith, et al. 2023). These nutritional changes have typically been enabled by bacterial symbionts, most often in the family Bartonellaceae (Alphaproteobacteria: Hyphomicrobiales) (Russell, et al. 2009; Bisch, et al. 2018). Genomics and experiments have confirmed Bartonellaceae roles in herbivorous genera *Cephalotes* (Hu, et al. 2018), *Dolichoderus* (Bisch, et al. 2018), and *Tetraponera* (Ma, et al. 2025), where they help recycle nitrogenous waste and supplement the host with nitrogen-rich compounds, especially amino acids. On the other hand, Bartonellaceae were also found in carnivores, and in the predatory ant *Harpegnathos saltator* encode genes related to protein degradation, amino acid transport, and riboflavin biosynthesis, apparently assisting the hosts with the digestion of protein-rich diets and using amino acids as energy source, in addition to provisioning vitamins (Neuvonen, et al. 2016; Bisch, et al. 2018). Phylogenomic analyses have shown that the functionally diverse ant symbionts form a monophyletic clade – *Candidatus* genus *Tokpelaia* – that is sister to the genus *Bartonella* that contains notorious human and mammalian pathogens (Segers, et al. 2017; Wagner and Dehio 2019), but also symbionts of bees and other arthropods (Bisch, et al. 2018; Ma, et al. 2025).

Planthoppers (Hemiptera: Fulgoromorpha) showcase a very different nutritional history and symbiotic make-up. They have ancestrally fed on phloem sap, and to counteract its imbalanced nutritional contents, have repeatedly formed symbiotic associations with microorganisms that provide the lacking nutrients (Bennett and Moran 2013; Bell-Roberts, et al. 2019; Brentassi and de la Fuente 2024). Specifically, this 263-million-year-old infraorder (Deng, et al. 2024) has ancestrally associated with a Bacteroidetes symbiont *Candidatus* Sulcia muelleri (further referred to as *Sulcia*) and a Betaproteobacteria symbiont *Ca.* Vidania fulgoroidea (thereafter: *Vidania*) (Urban and Cryan 2012; Deng, et al. 2023) . These two symbionts, transmitted strictly maternally and co-diversifying with their planthopper hosts (Moran, et al. 2005), jointly synthesize the 10 essential amino acids. In most planthoppers, *Sulcia* produces 3 [leucine, valine and isoleucine],

and *Vidania* 7 [methionine, histidine, tryptophan, threonine, lysine, arginine and phenylalanine] (Michalik, et al. 2021; Michalik, et al. 2025). However, in several planthopper clades, these microbes have been replaced by Hypocreales fungi (Siehl, et al. 2024; Michalik, et al. 2025), or else, joined by additional microbes that established residence within host tissues alongside the ancient endosymbionts (Michalik, et al. 2023; Brentassi and de la Fuente 2024). These additional symbionts, representing the genera *Sodalis*, *Asenophonus*, *Serratia*, and others, can complement the amino acid biosynthesis pathways of the ancestral symbionts and frequently encode genes involved in the biosynthesis of B vitamins (Bennett and Mao 2018; Michalik, et al. 2021; Michalik, et al. 2024). However, we do not have a good understanding of how the nutritional functions vary among symbiont or planthopper clades, how they have evolved, and how they complement the functions encoded by other members of the symbiotic partnership. As the genomics-based characterization of these more recently acquired symbionts in Auchenorrhyncha has been largely limited to a few species carrying *Sodalis* (Koga, et al. 2013; McCutcheon, et al. 2019; Michalik, et al. 2021), we know little about other microbes forming relationships with these sap feeders.

In our recent metagenomics study spanning 149 species of taxonomically diverse planthoppers from different geographic areas (Deng, et al. 2024; Michalik, et al. 2025), in some of the surveyed species we identified, for the first time, Bartonellaceae symbionts. This included a strain infecting *Saccharosydne* sp., a Delphacidae planthopper, which had lost *Sulcia* and whose *Vidania* had the smallest genome of any known bacteria, of a mere 50-kilobases genome, reduced to provisioning a single amino acid, phenylanine (Michalik, et al. 2025). We thus postulated that in this and other systems, the nutritional functions lost from the ancient symbionts' genomes may have been taken over by the Bartonellaceae. This conclusion has prompted a closer investigation into the nature and significance of the associations between Bartonellaceae and their planthopper hosts.

Specifically, we asked about Bartonellaceae distribution across planthopper clades and the phylogenetic relationships among the newly identified strains and other members of the family, including mammalian pathogens and ant symbionts. We also asked about the genomic characteristics and contents of these newly identified strains, particularly in relation to nutritional provisioning that had been identified as a key function of previously studied *Tokpelaia* symbionts of ants. Finally, we asked how these microbes fit into multi-partner symbioses that provide hosts with essential nutrients. We addressed these questions by exploring the extensive recent metagenomics data for diverse global planthoppers, and through comparative genomics analyses of the newly identified planthopper-associated and previously published Bartonellaceae strains.

2. Results

2.1. The distribution and phylogenetic relationships of planthopper-associated Bartonellaceae

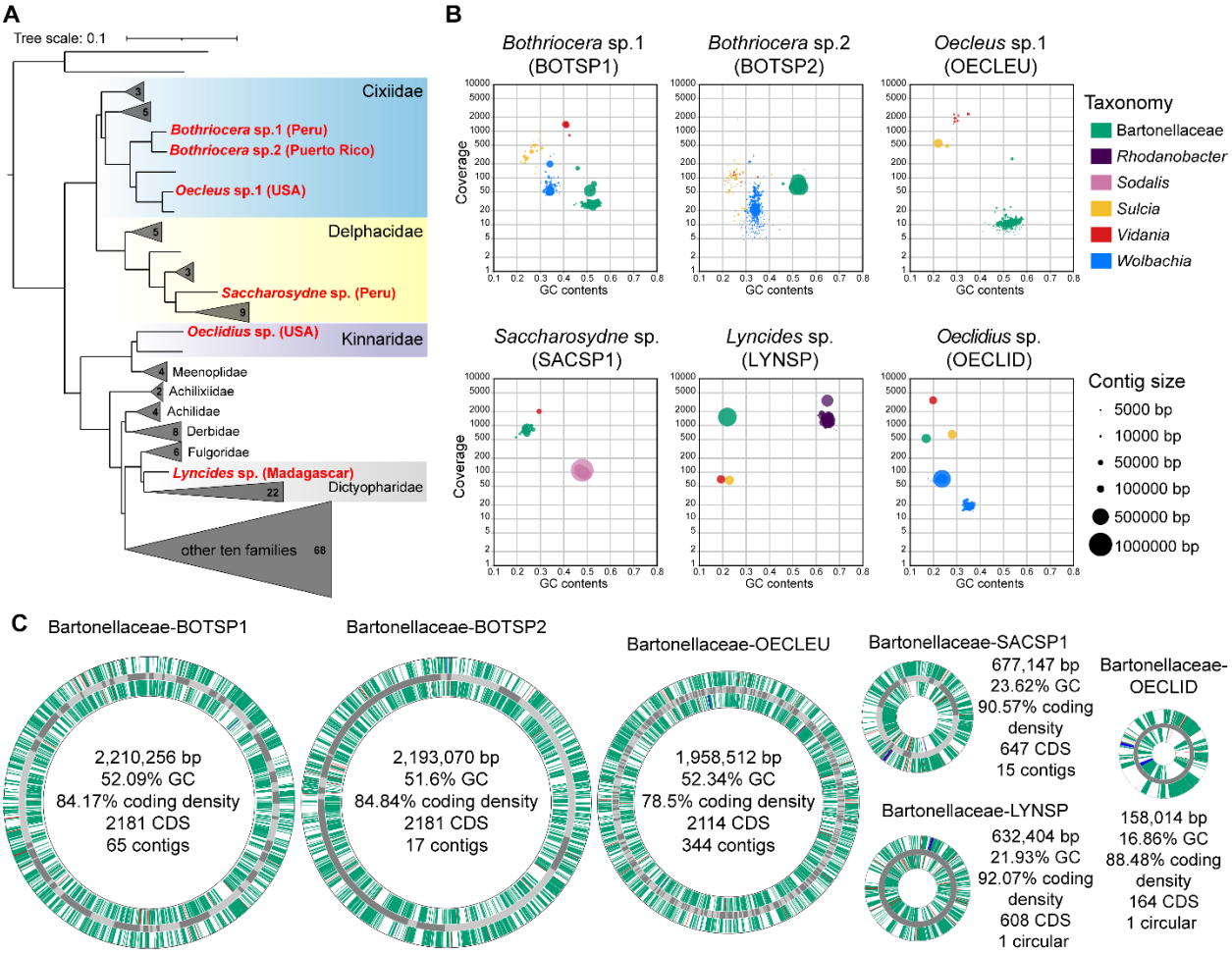

Figure 1. Bartonellaceae symbiont distribution across planthoppers, their representation within metagenomic assemblies, and genomic characteristics. (A) The genome-based phylogeny of 149 planthopper species representing 19 families, modified from Deng, et al. (2024), with species harboring Bartonellaceae indicated with red font. Clades where these symbionts were not detected were collapsed, with the number of species in these clades shown. **(B)** Taxon-annotated GC content-coverage plots for metagenomic assemblies of six planthopper species where Bartonellaceae were identified. Colored blobs represent contigs corresponding to the identified genomic fragments of different symbiotic bacteria. The size of each blob represents contig size. **(C)** Circular diagrams of *Tokpelaia* genomes from five species, showing gene positions on forward and reverse strands and basic genome characteristics. Note that only two genomes were circularized - most comprise multiple contigs, connected for the ease of illustration. The inner and outer circles represent coding sequences (CDS), and the functional categories are shown in green. tRNA, rRNA, and tmRNA genes are shown in red, blue, and purple, respectively. The dark gray and light gray in the middle circle represent different contigs. The size of the first five circles correlates with genome size, the sixth (OECLID) is magnified three times.

The systematic survey of 149 phylogenetically diverse planthopper species (Deng, et al. 2024; Michalik, et al. 2025) using 16S rRNA sequences reconstructed from metagenomes using phyloFlash (Gruber-Vodicka, et al. 2020), revealed Bartonellaceae in six species from four planthopper families, all classified to the genus level only. They included three members of the family Cixiidae: two representatives of the genus *Bothriocera* (BOTSP1, BOTSP2) and *Oecleus* sp. (hereafter OECLEU); a Dictyopharidae species *Lyncides* sp. (LYNSP); a Delphacidae planthopper *Saccharosydne* sp. (SACSP1), and a Kinnaridae species *Oeclidius* sp. (OECLID) (Fig. 1A, Tab. S1). The phylogenetic distance among Bartonellaceae-hosting planthoppers indicated at least four independent host-microbe associations – one in each family. In turn, the Cixiidae clade that included three Bartonellaceae-hosting species also contained two species with no traces of this bacterium, and based on currently available data, we cannot distinguish among scenarios such as independent infections, an ancestral gain followed by subsequent losses in some clades, or a facultative association.

The metagenome assemblies for these five planthopper species contained a set of contigs, between 1 and 344, with significant sequence similarity to the genome of *Candidatus* Tokpelaia hoelldobblerii (Bartonellaceae) from a predatory ant *Harpegnathos saltator* (Neuvonen, et al. 2016), and forming a distinct cluster in GC contents and coverage space (Fig. 1B, Tab. S2, S3). The Bartonellaceae symbionts from Cixiidae planthoppers (BOTSP1, BOTSP2, and OECLEU) all possessed relatively large genomes, ranging from 1.96 to 2.21 Mb, predicted to encode 2,114–2,181 proteins, and with the coding density from 78.5% to 84.84% (Fig. 1C). Their genomic GC content values were relatively high, between 51.6% and 52.34%. The Bartonellaceae symbionts of SACSP1 and LYNSP had genomes approximately three-fold smaller, 0.68 Mb (SACSP1) and 0.63 Mb (LYNSP), with 647 and 608 predicted protein-coding genes, respectively. These reduced genomes also exhibited markedly higher coding densities—90.57% for SACSP1 and 93.2% for LYNSP—as well as much lower GC contents, 23.62% and 21.93%, respectively (Fig. 1C, Tab. S2). In turn, for the Kinnaridae strain OECLID, we have identified a single circularly mapping contig of 158 kb, with a GC content of 16.86% and a coding density of 88.48%, encoding predicted 164 genes that included the large majority of those conserved in other extremely reduced bacterial genomes (McCutcheon, et al. 2024). Despite the low sequence similarity to any previously known Bartonellaceae (84.13 - 86.39% at 16S rRNA gene, AAI across value: 47.03 - 51.47%, Tab. S4, S5) amino acid sequences of protein-coding genes primarily resembled Bartonellaceae among all NCBI records. No further contigs within the assembly could be classified to the same genome.

Figure 2. Phylogenetic placement of Bartonellaceae symbionts of planthoppers. (A) Maximum likelihood analysis of Bartonellaceae strains with sequenced genomes, based on 68 homologous single-copy genes. The purple bars represent the genome size, and the blue bars - the GC content. **(B)** Maximum likelihood phylogeny of full-length 16S rRNA gene sequences of Bartonellaceae associated with planthoppers, ants, and selected other hosts, derived from metagenomes and the NCBI database. In both panels, the *Tokpelaia* clade is indicated with green boxes; bold black labels indicate ant-associated, and bold red – planthopper-associated strains. Bootstrap support values above 50% are shown by the respective nodes. The very long branch leading to the OECLID strain was truncated in both figures.

Phylogenomic analysis of the planthopper-associated Bartonellaceae reliably placed them among previously characterized ant symbionts forming the genus *Tokpelaia* (Fig. 2A, S1). Specifically, the planthopper-associated Bartonellaceae form two well-supported subclades within the genus *Tokpelaia*. One of them, on a relatively short branch, comprises symbionts of Cixiidae planthoppers (BOTSP1, BOTSP2, and OECLEU) and forms a sister clade with strains of *Tetraponera* ants. The second clade, monophyletic but with very long branches among strains, comprises symbionts of LYNSP, SACSP1, and OECLID. These small-genome strains show a closer relationship to smallest-genome symbionts of *Dolichoderus* ants, with strong support but long branches. This indicates an accelerated rate of sequence evolution, which combined with a strong AT bias, could have affected the accuracy of phylogenomic reconstructions through systematic problems such as long branch attraction, especially for the extremely reduced OECLID

(Susko and Roger 2021). 16S rRNA gene-based phylogeny, including a much broader sampling of *Tokpelaia* sequences, produced a topology broadly consistent with the genome-based analysis, again placing the planthopper-associated symbionts among ant-associated *Tokpelaia* strains and forming two separate clades (Fig. 2B, S2). However, due to the limited phylogenetic resolution of the short 16S rRNA gene, finer relationships among these lineages could not be confidently resolved.

The genome synteny comparisons did not support phylogenetics-based conclusions (Fig. S3). The highly fragmented assemblies of two of Cixiidae symbiont genomes prevented comparisons within that clade. There was little similarity among strains comprising the second clade. Surprisingly, large syntenic regions were detected between BOTSP2 and SACSP1, despite them representing different clades. Also, symbionts from BOTSP1 and BOTSP2 shared limited but distinct regions of synteny with *Tokpelaia* from *Tetraponera* ants.

2.2. Genomic features and metabolic functions in *Tokpelaia* symbionts from planthoppers and ants

The genomic contents comparisons among planthopper and ant symbionts revealed large differences among bacterial strains and host clades (Fig. 3). Typically, strains devote a large portion of their genome contents to the metabolism of carbohydrates, amino acids, cofactors and vitamins, nucleotides, and energy metabolism, as well as to membrane transport. Interestingly, Cixiidae symbiont strains consistently encode much greater numbers of carbohydrate metabolism genes than others. In turn, the numbers of amino acid and cofactor/vitamin biosynthesis genes were relatively consistent except for the genome-reduced planthopper symbiont clade, where they were substantially reduced, and amino acid gene set in a single ant strain (JR021). Interestingly, secondary metabollite metabolism genes are consistently much less abundant in planthopper symbionts than in ant symbionts.

As the number of translation-related genes is comparable across all but the tiniest genome, their relative proportion negatively correlates with the genome size. As a result, genetic information processing genes make up 61.8% of the tiny OECLID genome, as despite it losing many genes in that category, much greater losses occurred in most other categories. This included the elimination of all genes involved in essential amino acid biosynthesis, with only two non-essential amino acid metabolism genes (cysE and cysK) and two vitamin biosynthesis genes (cobA and iscS) remaining.

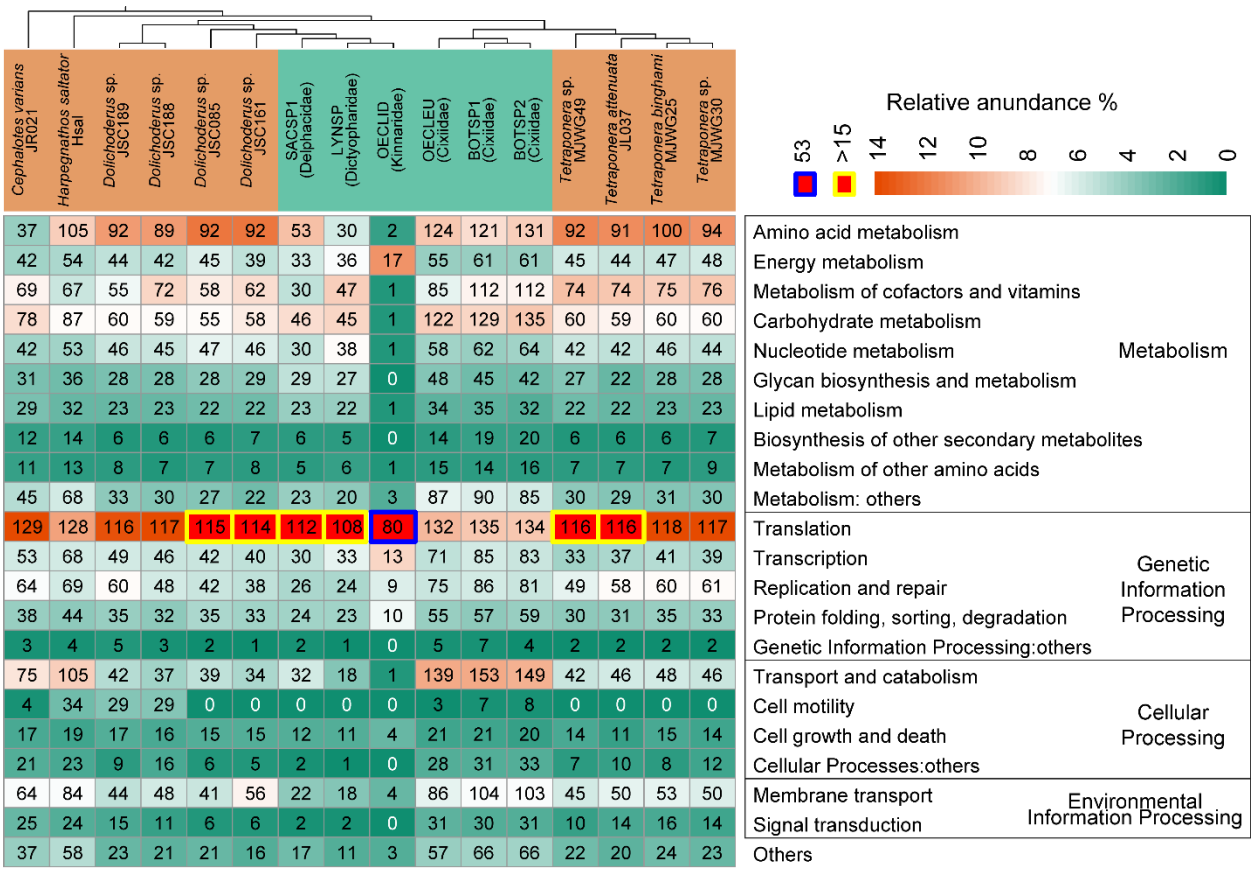

Figure 3. Relative abundance of KEGG pathways in *Tokpelaia* genomes. For each genome and each of selected categories, the numbers in the squares represent the number of genes in that category, and the background colors represent the proportion of the genes in that category relative to all genes included in the classification.

Comparative analysis of nutrient biosynthesis pathways across the sequenced *Tokpelaia* genomes indicated a strong similarity among symbionts of Cixiidae planthoppers (BOTSP1, BOTSP2, and OECLEU) and those of herbivorous *Dolichoderus* and *Tetraponera* ants (Fig. 3), whereas strains from LYNSP and SACSP1 had substantially reduced biosynthetic repertoire. Most strikingly, the highly reduced OECLID retained very little nutrient biosynthetic capacity.

Specifically, amino acid biosynthesis pathway comparisons revealed that the Cixiidae planthopper and *Dolichoderus* and *Tetraponera* ant symbionts retain very similar gene sets,spanning nearly complete pathways for the biosynthesis of non-essential and essential amino-acids with the exception of alanine-synthesizing transaminase (alaA) and asparagine synthase (asnB). *Tokpelaia* of BOTSP2 is the only missing prephenate dehydratase (pheA2), which catalyzes the decarboxylation of prephenate to phenylpyruvate in the phenylalanine pathway. The *Tokpelaia* of OECSP lacks three additional genes: aspB and glyA for aspartate and glycine synthesis, and argF involved in arginine biosynthesis. Notably, however, many of the genes in *Tokpelaia* genomes of BOTSP1, BOTSP2, and OECLEU, scattered across pathways, appeared pseudogenized (Fig. 4). In contrast, *Tokpelaia* of LYNSP and SACSP1 have both lost most amino acid biosynthesis genes. The LYNSP strain lacks the complete pathways for any

essential amino acid, retaining six of nine genes for lysine biosynthesis and complete pathways for four non-essential amino acids: glutamate, aspartate, glycine, and serine. The SACSP1 strain retains complete pathways for three essential amino acids (lysine, histidine and threonine) and three non-essential amino acids (glutamate, aspartate, and glycine), along with partial pathways for leucine, isoleucine, and valine. This reduced biosynthetic capacity resembles strain JR021, a culturable member of conserved multi-partite microbiota of a herbivorous *Cephalotes* ant (Hu, et al. 2018). The highly reduced OECLID strain encoded only two genes involved in the biosynthesis of one non-essential amino acid, cysteine.

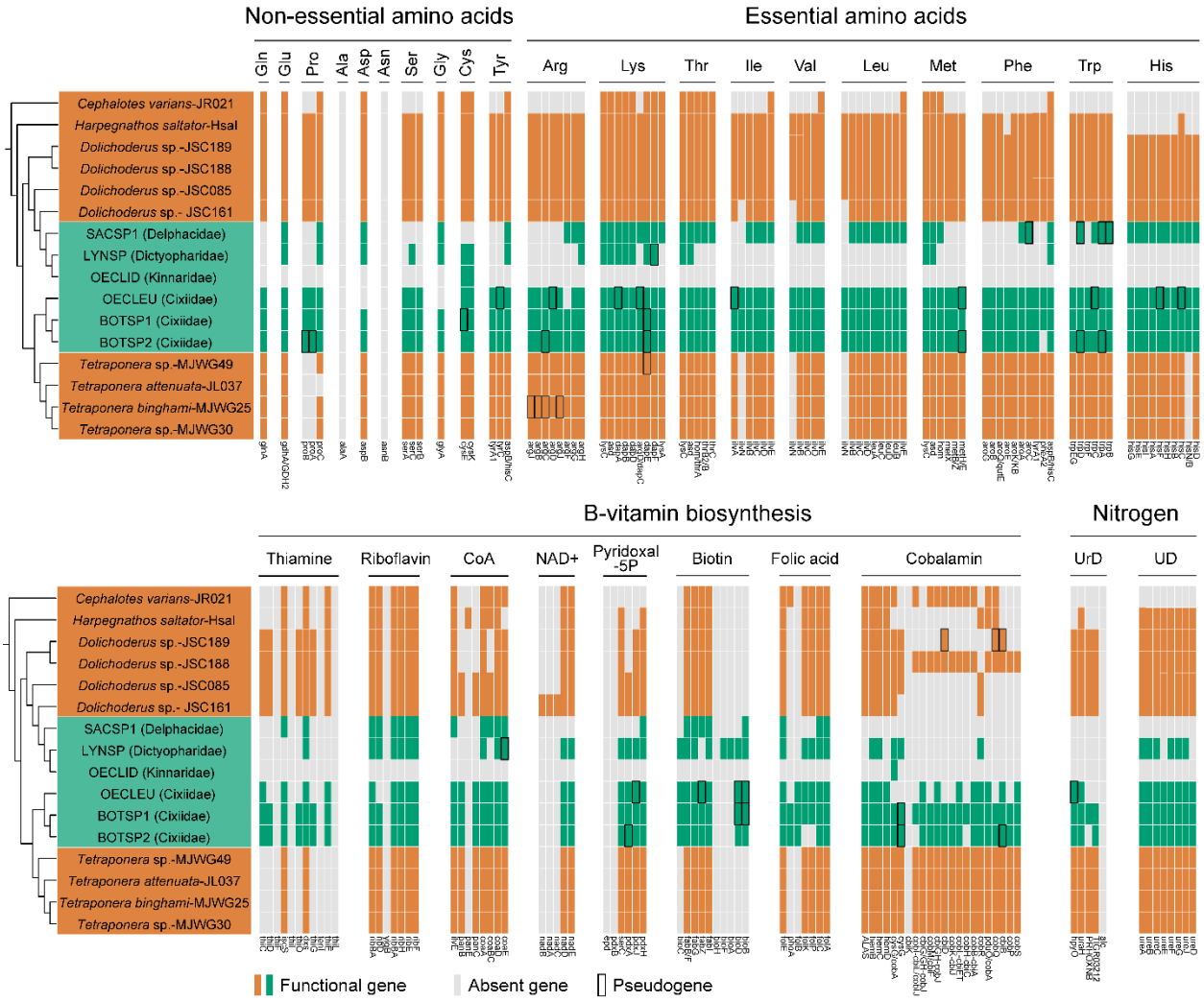

Figure 4. Summary of *Tokpelaia* biosynthetic functions. Genes associated with nitrogen cycling and cofactor biosynthesis in the genomes of planthopper (green) and ant (orange)-associated *Tokpelaia*. Each bar represents a single gene with an abbreviation at the bottom. Genes are classified as present (green or orange) and absent (light gray); putative pseudogenes, with disrupted open reading frames, are placed in a black frame. A cladogram based on the phylogenomic relationships of *Tokpelaia* is shown on the left, with labels in the same order as in Fig. 2A (where full host names are provided). UrD: Uric acid degradation, UD: Urea degradation.

266 Besides amino acids, the *Tokpelaia* symbionts display a broad vitamin biosynthetic capability,
267 with substantial variation among strains, and again a substantial similarity among Cixiidae,
268 *Dolichoderus,* and *Tetraponera* symbionts. For thiamine biosynthesis, BOTSP1 & BOTSP2
269 symbionts encode the same set of seven genes as *Dolichoderus* strains, whereas all other strains
270 encode a more reduced set. All studied *Tokpelaia* strains retain genes for riboflavin (vitamin B2)
271 biosynthesis, only missing yigB gene, with OECLEU additionally lacking ribD. For CoA and
272 pyridoxal-5P, *Tokpelaia* symbionts of Cixiidae planthoppers encode similar gene sets as those of
273 *Tetraponera* ants, with *Dolichoderus* and LYNSP-SACSP1 symbionts encoding fewer. In turn, the
274 set of four biotin (vitamin B7) biosynthesis genes is also highly conserved among *Tokpelaia*
275 strains (with a single gene loss in LYNSP), but all planthopper symbionts encode a set of
276 additional genes involved in the production of this vitamin. Interestingly, BOTSP1's *Tokpelaia* is
277 the only one retaining the full pathway from purines to folic acid, with all other strains having
278 experienced gene losses. Finally, cobalamin, a key coenzyme involved in the metabolism of
279 branched-chain amino acids and methionine, has the longest synthesis pathway of 22 steps, all
280 but one of which (cbiK) are encoded by *Tetraponera* and BOTSP1 symbionts. Here, BOTSP2 and
281 one of the *Dolichoderus* strains lost more genes, and all others have a more substantially reduced
282 set. The highly reduced OECLID symbiont only encodes one vitamin biosynthesis gene, cobA,
283 in the cobalamin biosynthesis pathway.
284
285 Similarly, *Tokpelaia* from Cixiidae planthoppers and most ants (*Dolichoderus*, *Tetraponera*,
286 and *Harpegnathos saltator*) harbor a complete urease gene cluster (ureABCDEFGJ) for urea
287 degradation. However, the LYNSP strain lacks two accessory components (ureE and ureJ), and
288 SACSP1 has lost the entire gene cluster, same as the *Cephalotes* ant symbiont JR021 (Fig. 3).
289 On the other hand, while the pathway for uric acid degradation is retained in *Tokpelaia* from
290 *Dolichoderus* and *Tetraponera* ants, BOTSP1 is the only planthopper whose symbiont retains the
291 entire pathway, with partial or complete losses in other species (Fig. 4). Again, none of the
292 nitrogen recycling genes were found in *Tokpelaia*-OECLID.
293
294 The close inspection of the OECLID symbiont gene set revealed, outside of genetic
295 information processing genes, two substantial functions that were retained (Fig. S4, Tab. S7). The
296 first is energy metabolism, including components of the electron transport chain (17 genes related
297 to cytochrome c oxidase, ATP synthase, and electron transfer), the second is iron–sulfur cluster
298 assembly (Fig. S5).

2.3. The degeneration of ancient symbiont degeneration is often associated with the *Tokpelaia* acquisition

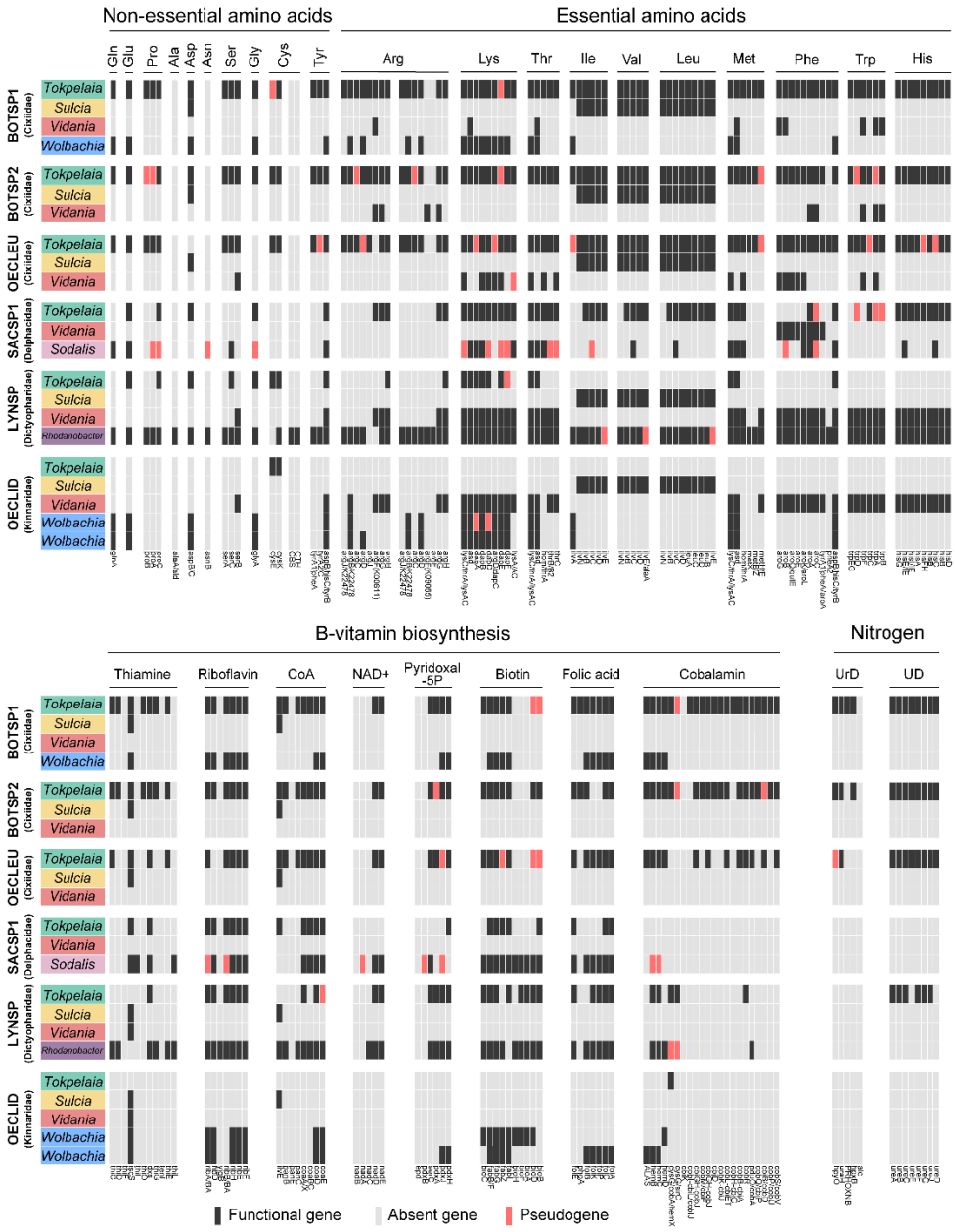

Figure 5. Genomic complementarity in nutritional provisioning among planthopper symbionts. Amino acid and B vitamin biosynthesis gene distribution among genomes of different symbionts that were identified in metagenomes of **six** *Tokpelaia*-hosting planthopper species. Black rectangles represent genes present in the genome, red square recognizable pseudogenes, grey represent genes that were not detected.

Within metagenomic assemblies of the six *Tokpelaia*-hosting planthopper species we also identified other bacteria (Fig. 1B, Tab. S1) (Michalik, et al. 2025). Specifically, within all studied insects, we identified rRNA sequences of *Vidania*, and in five of them (all except SACSP1), of *Sulcia* – two ancestral symbionts of planthoppers. BOTSP1, BOTSP2 and OECLID also contained the facultative symbiont *Wolbachia*. LYNSP also carried *Rhodanobacter*, whereas SACSP1 hosted *Sodalis* as an additional symbiont (Fig. 1B). For most of these symbionts, we recovered genomic bins with acceptable completeness and contamination statistics (Tab. S2). Intriguingly, in Cixiidae species, *Vidania* contigs were fragmented and incomplete, unlike in the three other *Tokpelaia*-hosting species and the vast majority of previously studied planthoppers (Michalik, et al. 2025). Further, in BOTSP2, *Wolbachia was* not successfully reconstructed through binning; we identified matching contigs via blastN, but because of high contamination scores for this setas computed by CheckM (72.81%) we decided to exclude it from analyses.

The ancient nutritional endosymbionts' genome organization and functions varied among the studied species (Fig. 5). *Sulcia* was conserved in the five planthoppers that retained it, always encoding full pathways for valine, leucine, and isoleucine biosynthesis. In contrast, *Vidania* genome retained ancestral organization and the ability to produce seven essential amino acids in only two species, LYNSP and OECLID. In the third, SACSP1, *Vidania* genome was reduced to provisioning phenylalanine (Michalik, et al. 2025). In turn, in all three Cixiidae species, the unusually fragmented *Vidania* strains only retained a small and variable subset of genes scattered across biosynthetic pathways. Then, in Cixiidae and SACSP1, genes encoded by other bacteria fill in the gaps in conserved essential amino acid biosynthesis pathways (Fig. 5). Specifically, in Cixiidae, the largely complete biosynthetic gene sets encoded in *Tokpelaia* genomes seem to fully compensate for *Vidania* gene losses. With few exceptions (lysine and phenylalanine), this is also the case in SACSP1; in this species, 17 genes overlapping with those of *Tokpelaia* are also encoded by *Sodalis*. In LYNSP, the conserved gene sets of *Sulcia* and *Vidania* are almost entireely duplicated by *Rhodanobacter*, with *Tokpelaia* providing further scattered functions. Finally, in OECLID, two *Wolbachia* strains encode a set of biosynthetic genes that largely overlap with those of ancient symbionts.

Species differed substantially in terms of B vitamin and cofactor biosynthesis gene sets and their distribution across symbiont genomes (Fig. 5). In Cixiidae species, *Tokpelaia* encodes large numbers of genes across different pathways, with Wolbachia duplicating these functions in BOTSP1. In SACSP1 and LYNSP, *Sodalis* and *Rhodanobacter*, respectively, encode the greatest number of genes, with *Tokpelaia* duplicating many of them. Finally, in OECLID, its two *Wolbachia* genomes encode multiple functions, and *Tokpelaia* contributes two genes. Across all these species, few biosynthetic pathways appear complete, but we note that the assembly fragmentation may have prevented the detection of some genes, and they may also be encoded within the host genomes.

3. Discussion

When famously writing about 'a veritable fairyland of insect symbiosis', Buchner (1965) referred to the huge variety of bacterial symbionts of Auchenorrhyncha that accompany, and

351 sometimes have replaced, their ancient, degraded nutritional endosymbionts. Of these, the
352 enterobacterial genus *Sodalis* – apparently, the most broadly distributed across the taxonomic
353 diversity of insects – may be the best understood in terms of genomic evolution and nutritional
354 contributions to the host biology (McCutcheon, et al. 2019; Renoz, et al. 2023). Through our
355 research, the family Bartonellaceae joins the ranks of Auchenorrhynchan co-symbionts, showing
356 patterns of the distribution, evolution, and function not unlike *Sodalis*, but casting a new
357 perspective on microbial symbioses of hemipterans, ants, and other insects.

358 3.1. Bartonellaceae host range broadens

359 Bartonellaceae are primarily known as pathogens of mammals, often vectored by blood-
360 sucking insects (Engel, et al. 2011). As an example, *Bartonella bacilliformis* causes the
361 devastating Carrion disease in humans (Pons, et al. 2016), while *Bartonella henselae* causes
362 persistent and often asymptomatic infections in cats, vectored by cat fleas (Chomel, et al. 1996).
363 The discovery of the bee gut symbiont *Bartonella apis* (Segers, et al. 2017) and the identification
364 of more divergent strains colonizing diverse ants, classified to the *Candidatus* genus *Tokpelaia*
365 (Russell, et al. 2009; Bisch, et al. 2018; Ma, et al. 2025), have changed the understanding of the
366 biology of this bacterial group. Our discovery of planthopper-associated Bartonellaceae, fitting
367 right within the *Tokpelaia* clade and with at least four independent original host association events,
368 further broadens the list of confirmed hosts. The plant-based diet of many Bartonellaceae-hosting
369 ants and planthoppers, but also the mutualistic interactions among many ants and hemipterans,
370 including planthoppers (Ivens, et al. 2017; Bourgoin, et al. 2023), suggest potential avenues for
371 parallel acquisition of related microbes by these groups, or their exchange among groups.
372 However, we already know that Bartonellaceae host list is longer: for example, a divergent clade
373 commonly colonizes stored-product mites (Xiong, et al. 2024), and Bartonellaceae 16S rRNA
374 sequences were identified in a scorpion and in heteropteran bugs (Matsuura, et al. 2012; Martinez,
375 et al. 2019). We have no doubts that a systematic survey across diverse arthropods will reveal
376 these bacteria in a substantially broader range of hosts. At the same time, as shown recently,
377 systematic screens across related species and within species, microscopy, transcriptomics,
378 and/or experiments, can clarify the nature of their associations with diverse hosts (Xiong, et al.
379 2024; Ma, et al. 2025).
380
381 Overall, our findings support the view of Bartonellaceae as relatively broadly distributed
382 across animals, and perhaps a wider range of environments. We suspect that some strains are
383 relatively versatile and capable of synthesizing a wide range of nutrients, which enable them to
384 move across and colonize new host lineages and then potentially evolve into more specialized
385 associations. Some of these, including the mammal-pathogenic *Bartonella* clade, retain the ability
386 to switch hosts, while others may become strictly heritable, promoting genomic reduction and the
387 loss of functions not essential to the specific host. This proposed evolutionary trajectory agrees
388 with the emerging consensus on the evolution of other key insect-symbiotic bacteria, exemplified
389 by the enterobacterial genus *Sodalis* (McCutcheon, et al. 2019). It includes strains such as
390 *Sodalis ligni* that live in dead wood (Tlaskal, et al. 2021)*; Sodalis praecaptivus* isolated from a
391 human wounds following impalement on a branch (Clayton, et al. 2012); and diverse strains that
392 have independently established heritable nutritional symbiotic relationships with a wide variety of

insect lineages (Renoz, et al. 2023). Such heritable associations typically lead to rampant genomic changes, increased evolutionary rates, and rapid gene losses. Some of the degenerating symbionts may reach a stable state where gene loss slows to nearly a halt – as evidenced by high similarity among genomes of *Buchnera*, *Blochmannia*, or *Baumannia*, all likely derived from *Sodalis* tens or hundreds of millions years ago (McCutcheon, et al. 2019).

3.2. Planthopper symbionts represent contrasting genomic evolutionary outcomes

Variation in planthopper-associated *Tokpelaia* genome size and contents indicates different lifestyles and varying effects on the host biology, paralleling the range of associations of Sodalis-allied symbionts.

On the one hand, the three *Tokpelaia* strains infecting Cixiidae planthoppers have the largest genomes (~2Mb) among the *Tokpelaia* strains studied so far (0.96-1.9Mb), retaining a broad range of amino acid and vitamin biosynthesis genes. This suggests a substantial contribution to the host nutrition, comparable to that of ant-associated *Tokpelaia* (Bisch, et al. 2018; Ma, et al. 2025). Further, these strains' positions on short branches in the *Tokpelaia* phylogeny suggests proximity to the putative ancestral state of other associations with diverse ants and planthoppers. This resembles the versatile *Sodalis praecaptivus*, whose function-rich genome is much larger than most insect-symbiotic *Sodalis* strains, and which consistently falls near the base of *Sodalis* phylogenies, among variable-length branches representing insect-symbiotic strains (McCutcheon, et al. 2019).

On the other hand, the *Tokpelaia* genomes of LYNSP and SACSP1 planthoppers are substantially reduced (632-677 Kb), to the level comparable with *Sodalis*-derived long-term obligate nutritional symbionts of aphids (B*uchnera*, 412-671 kb), carpenter ants (B*lochmannia*, 705-792 kb), sharpshooter leafhoppers (B*aumannia*, 633-759 kb), or tse-tse flies (Wigglesworthia, ~700 kb) (Bennett, et al. 2015; Bing, et al. 2017; Chong, et al. 2019). Despite the massive variation in their age of association, tissue localization, transmission mechanisms, and biosynthetic capacity, all these gamma-proteobacterial symbionts provide deficient nutrients that make them virtually indispensable to their hosts (Feldhaar, et al. 2007; McCutcheon, et al. 2009; Bing, et al. 2017). We suspect that LYNSP and SACSP1 *Tokpelaia* strain plays a comparable roles in planthoppers. As both strains have lost multiple nutritional genes that are retained by other symbionts, their specific situation may particularly resemble that of *Baumannia*, or *Buchnera* from *Cinara* aphids (Manzano-Marin, et al. 2016), or endobacterial symbiont of mealybugs (Husnik and McCutcheon 2016): all of them are complementary in their amino acid biosynthetic capacity to other symbionts residing within the same host insect.

In our view, the most striking discovery is the tiny, 158-kb genome of the *Tokpelaia* symbiont of the Kinnaridae planthopper *Oeclidius* sp. (OECLID), comparable in size to the most reduced known insect endosymbionts: *Sulcia*, *Vidania*, *Nasuia*, *Hodgkinia*, *Tremblaya*, or *Carsonella* (McCutcheon, et al. 2024). All these bacteria colonize sap-sucking hemipterans, where they have well-established roles as provisioners of essential amino acids and sometimes vitamins. This

does not appear to be the case with *Tokpelaia*-OECLID – even if, as shown in a mealybug, one nutrition-related biosynthetic gene (cobA in the current case) might be enough to justify long-term symbiosis (Garber, et al. 2024). But in fact, OECLID's symbiont genetic repertoire resembles much more closely the endosymbionts of certain protists. Specifically, the 158-kb genome of a Verrucomicrobium, *Ca.* Organicella extenuata, symbiont of an *Euplotes* protist inhabiting a hypersaline Antarctic lake, also retains the capacity for genetic information processing but not for the biosynthesis of amino acids or vitamins (Williams, et al. 2021), with fatty acid synthesis and iron-sulfur (Fe-S) cluster assembly identified as its likely contribution to the host biology. Fe-S cluster assembly has not been identified as critical to sap-sucking insects' biology, and thus the nature of the association, significance, and evolutionary history of this unusual bacterium that conflicts with the current understanding of insect heritable endosymbioses remains a mystery that cannot be resolved based on a single mid-coverage metagenome alone. To further complicate conclusions, from OECLID metagenome, we also reconstructed a low-coverage rRNA sequence representing an Entomophthorales fungus with 96.6% identity to entomopathogens *Pandora delphacis* and *Furia pieris*, and we cannot exclude the possibility that *Tokpelaia*-OECLID actually associates with the planthopper's pathogen (Matsuura, et al. 2018; Siehl, et al. 2024).

3.3. Bartonellaceae significance in the biology of planthoppers and their other symbionts

We know that heritable nutritional endosymbionts in diverse insect systems produce essential nutrients deficient in their hosts' diets, thus being indispensable to the hosts. At the same time, they drive the evolution of hosts' dependence on symbiont nutritional contributions, even as these symbionts degrade and loose efficiency – a process aptly described as 'the evolutionary rabbit hole' (Bennett and Moran 2015). Further, they drive the complementary loss of functions by their co-symbionts (McCutcheon, et al. 2009; Husnik and McCutcheon 2016), forcing the hosts to manage an increasingly specialized symbiotic consortium. We think that these general statements accurately describe the biology and significance of most planthopper-associated *Tokpelaia* strains that we have identified, which in all cases, belong to multi-partite symbiotic consortia with complementary biosynthetic capabilities (Michalik, et al. 2021). In a delphacid SACSP1, the Bartonellaceae symbiont largely fills up the nutritional gap resulting from *Sulcia* loss and *Vidania* genome reduction to the smallest-known state, while in Cixiidae, they seem to compensate for the apparent degeneration of *Vidania.* While such complementarity cannot be seen in LYNSP, and in OECLID it is limited to a single gene, it is tempting to think of *Tokpelaia* as the common driver of the functional loss by ancient symbioses. At the same time, other symbionts are likely to be driving the gene loss in *Tokpelaia*,

We are not aware of any major ecological transitions in planthopper lineages represented by *Tokpelaia*-infected species. Such changes, exemplified by Typhlocybinae leafhoppers that switched to feeding on parenchymal cell cytoplasm (Buchner 1965; Kobiałka, et al. 2025) and in Derbidae and Achilidae planthoppers that switched to feeding on fungi during nymphal stages (Michalik, et al. 2025) – seem to have enabled major changes and functional losses in their microbiota. Here, we think that through alternative evolutionary routes, symbiotic complexes involving *Tokpelaia* have reached, or are on a path towards, complementarity in the biosynthesis

473 of largely the same set of nutrients as most other planthoppers. Depending on the order of arrival
474 and the properties of bacteria that were in planthoppers before, the independently acquired
475 Bartonellaceae have assumed different positions and roles in the symbiotic consortium.

476 However, in all presented systems, achieving the full understanding of host-microbe
477 associations would require a much broader set of data that what we had available - a combination
478 within-species microbiome surveys, genomic comparisons among closely related species,
479 microscopy, and, ideally, experiments. Unfortunately, these approaches are not easy to
480 implement on unidentified non-model species sampled on different continents years ago, which
481 was the case here. Nevertheless, this study demonstrates that a broad genomics survey alone
482 can help discover new biological states and processes that change our understanding of biology
483 and evolution of significant clades of organisms. In the age of genomics, such broad surveys
484 should become an important tool for the discovery and characterization of biodiversity and
485 biological interactions.

486 4. Methods

487 4.1. Study material and data

488 For the study, we used a set of metagenomic datasets for 149 taxonomically diverse
489 planthoppers, presented in our previous work that focused on phylogenomic relationships of these
490 planthoppers and comparative genomics of their *Sulcia* and *Vidania* symbionts (Deng, et al. 2024;
491 Michalik, et al. 2025). Briefly, insects were wild-caught from their natural habitats around the world
492 between 1996 and 2022. Following identifications by taxonomic experts, the DNA extracted from
493 dissected abdominal tissue of single individuals, or in a few cases, dissected bacteriome tissue,
494 was used for metagenomic library preparation and sequencing on Illumina platforms. Collection
495 metadata, reagents, and sequencing platforms for the entire metagenomic collection are provided
496 in the abovementioned publications. For the six specimens within which we detected *Tokpelaia*,
497 listed in Tab. S1, abdominal tissue was used for DNA extraction using the Bio-Trace DNA
498 Purification Kit (Eurx, Poland), and sequencing libraries were prepared and sequenced by
499 Novogene Ltd. on NovaSeq 6000 platform using a 300-cycle kit.

500 4.2. Bioinformatics analysis

501 To detect Bartonellaceae within the sequenced metagenomes, we systematically examined
502 16S rRNA sequences assembled using phyloFlash (Gruber-Vodicka, et al. 2020), validating all
503 those without close matches to recognized bacteria through phylogenetics and blastn
504 comparisons against NCBI databases. In all cases where these data suggested
505 Hyphomicrobiales infections, we followed with the analysis of metagenomic assemblies obtained
506 using Megahit v1.2.9 (maximum k-mer size = 255, min contig size= 1000) (Li, et al. 2015) To
507 obtain the draft genomes of Bartonellaceae and other bacteria present in these samples, we
508 binned the metagenomic assembly data using the metaWRAP pipeline (Uritskiy, et al. 2018),
509 incorporating three binning tools: CONCOCT, MaxBin, and metaBAT. The resulting bins were

510 refined with metaWRAP's "Bin_refinement" module, applying a minimum completeness threshold
511 of 70% and a maximum contamination limit of 10%. To avoid the ommission of genomic contigs,
512 assemblies were further examined using BLASTn and BLASTx searches against the NCBI
513 database, and separately against published and unpublished genomes. Only contigs consistent
514 with the bin classification, and meeting criteria for read coverage and GC contents were retained
515 (Tab. S2). Bin completeness and contamination were evaluated using CheckM v21 (Parks, et al.
516 2015), and taxonomic classification was performed using the Genome Taxonomy Database via
517 GTDB-Tk v2.1.122 (Chaumeil, et al. 2019). Average amino acid identity (AAI) was calculated
518 using CompareM v0.1.2 (https://github.com/dparks1134/CompareM) to assess the genomic
519 relatedness among *Tokpelaia* symbionts.
520
521 Contigs of bacteria other than *Sulcia* and *Vidania* were annotated with Prokka v1.14.6
522 (Seemann 2014), and gene functions were assigned to KEGG Orthology (KO) pathways using
523 KofamScan (Aramaki, et al. 2020) and the online tool eggNOG-mapper (Tab. S3 4)
524 (Cantalapiedra, et al. 2021). Based on these annotations, metabolic pathways of the symbiotic
525 bacteria were manually reconstructed using the KEGG database. Genomes were visualized using
526 Proksee (Grant, et al. 2023). Genome synteny analysis was performed using Promer tool for
527 Mummer v3.07, with default settings (Kurtz, et al. 2004). Pseudofinder v1.1.0 (Syberg-Olsen, et
528 al. 2022) software was used to predict pseudogenes in the genomes of symbiotic bacteriaIn turn,
529 all confirmed *Sulcia* and *Vidania* contigs and genomes were annotated using a custom pipeline
530 (Łukasik, et al. 2018) adopted for planthopper symbionts recently (Michalik, et al. 2025). The
531 pipeline first extracted all Open Reading Frames (ORFs) and translated them into amino acid
532 sequences. These ORFs were then searched using HMMER v3.3.1 (Eddy 2011) (hmmer.org)
533 against a custom database containing manually curated protein-coding and rRNA genes from
534 previously characterized *Sulcia* and *Vidania* genomes. Ribosomal RNA genes were detected with
535 nhmmer implemented in HMMER v3.3.1, and tRNA genes were identified with tRNAscan-SE
536 v2.0.7 (Chan, et al. 2021). Based on the relative length to the reference genes, protein-coding
537 genes were classified as functional (>90%), putative pseudogenes (>60%), or pseudogenes
538 (<60%). Any ORFs longer than 300 bp with no significant similarity to any reference genes were
539 further examined by querying UniProt (The UniProt 2021) and NCBI databases and compared
540 carefully to the top hits. Genes lacking any annotations were marked as "hypothetical".

541 4.3. Phylogenetic analyses

542 To investigate the evolutionary origin of Bartonellaceae bacteria associated with
543 planthoppers, we compared Phyloflash-reconstructed 16S rRNA sequences against top hits of
544 insect and environmental origin, especially published ant-associated *Tokpelaia* sequences. All
545 sequences were aligned using MAFFT v7.526 (Rozewicki, et al. 2019), with divergent regions
546 removed using Gblocks v0.91b (Talavera and Castresana 2007). Phylogenetic analysis was
547 performed using RAxML v8.2.12 (Stamatakis 2014) under the GTRGAMMA model, with *Brucella*
548 *pseudintermedia* (accession number: CP099968) designated as the outgroup. The analysis
549 included 1000 bootstrap replicates and was visualized using iTOL (Letunic and Bork 2021).
550 To clarify the evolutionary relationships among the newly identified strains and other
551 Bartonellaceae, we conducted a genome-wide phylogenetic analysis that included all currently

552 published *Tokpelaia* or *Bartonella* genomes representing the major recognized clades. The
553 proteins were clustered into families using OrthoFinder v2.5.5 (Emms and Kelly 2019), resulting
554 in a data set of 68 single-copy orthologs. As for rRNA data, proteins from each family were aligned
555 using MAFFT v7.526 (Rozewicki, et al. 2019), with divergent and ambiguously aligned blocks
556 trimmed using Gblocks v0.91b (Talavera and Castresana 2007) . We concatenated the resulting
557 alignments. Prottest 3.4.2 (Guindon and Gascuel 2003; Darriba, et al. 2011) was used to select
558 the optimal amino acid substitution model, JTT+I+G+F, according to AIC and BIC criterions.
559 Phylogenetic analysis was performed with RAxML v8.2.12 (Stamatakis 2014) using the best
560 model and 1000 bootstrap replicates, and results were visualized and annotated using iTOL
561 (Letunic and Bork 2021). All images were adjusted using Adobe Illustrator CC 2019.

562 5. Data availability

563 All raw data and genomic assemblies used in comparisons have been uploaded to NCBI
564 databases. BioProject and BioSample accession numbers are listed in Tab.s S1-S2.

565 6. Acknowledgements

566 This study was supported by the Polish National Science Centre grant
567 2018/30/E/NZ8/00880 to P.Ł. and the China Scholarship Council grant no. 202406040164 to
568 M.M.

569 7. References

570 Aramaki T, et al. 2020. KofamKOALA: KEGG ortholog assignment based on profile HMM and
571 adaptive score threshold. Bioinformatics 36: 2251-2252. doi: 10.1093/bioinformatics/btz859

572 Bell-Roberts L, Douglas AE, Werner GDA 2019. Match and mismatch between dietary switches
573 and microbial partners in plant sap-feeding insects. Proc. Biol. Sci. 286: 20190065. doi:
574 10.1098/rspb.2019.0065

575 Bennett GM, Mao M 2018. Comparative genomics of a quadripartite symbiosis in a planthopper
576 host reveals the origins and rearranged nutritional responsibilities of anciently diverged bacterial
577 lineages. Environ. Microbiol. 20: 4461-4472. doi: 10.1111/1462-2920.14367

578 Bennett GM, McCutcheon JP, McDonald BR, Moran NA 2015. Lineage-specific patterns of
579 genome deterioration in obligate symbionts of sharpshooter leafhoppers. Genome Biol. Evol. 8:
580 296-301. doi: 10.1093/gbe/evv159

581 Bennett GM, Moran NA 2015. Heritable symbiosis: The advantages and perils of an evolutionary
582 rabbit hole. Proc. Natl. Acad. Sci. USA 112: 10169-10176. doi: 10.1073/pnas.1421388112

583 Bennett GM, Moran NA 2013. Small, smaller, smallest: the origins and evolution of ancient dual
584 symbioses in a Phloem-feeding insect. Genome Biol. Evol. 5: 1675-1688. doi:
585 10.1093/gbe/evt118

586 Bing X, et al. 2017. Unravelling the relationship between the tsetse fly and its obligate symbiont
587 *Wigglesworthia*: transcriptomic and metabolomic landscapes reveal highly integrated
588 physiological networks. Proc. R. Soc. B 284. doi: 10.1098/rspb.2017.0360

589 Bisch G, et al. 2018. Genome evolution of Bartonellaceae symbionts of ants at the opposite ends
590 of the trophic scale. Genome Biol. Evol. 10: 1687–1704. doi: 10.1093/gbe/evy126

591 Bourgoin T, et al. 2023. When cockroaches replace ants in trophobiosis: a new major life-trait
592 pattern of Hemiptera planthoppers behaviour disclosed when synthesizing photographic data.
593 Diversity 15: 356. doi: 10.3390/d15030356

594 Brentassi ME, de la Fuente D 2024. Obligate mutualistic heritable symbiosis in sap-feeding
595 insects: an intricate relationship in nature. Symbiosis 92: 159-185. doi: 10.1007/s13199-023-
596 00965-1

597 Buchner P. 1965. Endosymbiosis of animals with plant microorganisms. Interscience Publishers,
598 Inc, New York.

599 Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J 2021. eggNOG-mapper
600 v2: functional annotation, orthology assignments, and domain prediction at the metagenomic
601 scale. Mol. Biol. Evol. 38: 5825-5829. doi: 10.1093/molbev/msab293

602 Chan Patricia P, Lin Brian Y, Mak Allysia J, Lowe Todd M 2021. tRNAscan-SE 2.0: improved
603 detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49: 9077-9096.
604 doi: 10.1093/nar/gkab688

605 Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH 2019. GTDB-Tk: a toolkit to classify genomes
606 with the Genome Taxonomy Database. Bioinformatics 36: 1925–1927. doi:
607 10.1093/bioinformatics/btz848

608 Chomel BB, et al. 1996. Experimental transmission of *Bartonella henselae* by the cat flea. J. Clin.
609 Microbiol. 34: 1952-1956. doi: 10.1128/jcm.34.8.1952-1956.1996

610 Chong RA, Park H, Moran NA 2019. Genome evolution of the obligate endosymbiont *Buchnera*
611 *aphidicola*. Mol. Biol. Evol. 36: 1481-1489. doi: 10.1093/molbev/msz082

612 Clayton AL, et al. 2012. A novel human-infection-derived bacterium provides insights into the
613 evolutionary origins of mutualistic insect-bacterial symbioses. PLoS Genet. 8: e1002990. doi:
614 10.1371/journal.pgen.1002990

615 Cornwallis CK, et al. 2023. Symbioses shape feeding niches and diversification across insects.
616 Nat. Ecol. Evol. doi: 10.1038/s41559-023-02058-0

617 Darriba D, Taboada GL, Doallo R, Posada D 2011. ProtTest 3: fast selection of best-fit models of
618 protein evolution. Bioinformatics 27: 1164–1165. doi: 10.1093/bioinformatics/btr088

619 Deng J, et al. 2023. Genome comparison reveals inversions and alternative evolutionary history
620 of nutritional endosymbionts in planthoppers (Hemiptera: Fulgoromorpha). Genome Biol. Evol. 15:
621 evad120. doi: 10.1093/gbe/evad120

622 Deng J, et al. 2024. Phylogenomic insights into the relationship and the evolutionary history of
623 planthoppers (Insecta: Hemiptera: Fulgoromorpha). Syst. Entomol. 50: 495–518. doi:
624 10.1111/syen.12666

625 Douglas AE 2014. Symbiosis as a general principle in eukaryotic evolution. Cold Spring Harbor
626 Perspect. Biol. 6. doi: 10.1101/cshperspect.a016113

627 Eddy SR 2011. Accelerated profile HMM searches. PLoS Comp. Biol. 7: e1002195. doi:
628 10.1371/journal.pcbi.1002195

629 Emms DM, Kelly S 2019. OrthoFinder: phylogenetic orthology inference for comparative
630 genomics. Genome Biol. 20: 238. doi: 10.1186/s13059-019-1832-y

631 Engel P, et al. 2011. Parallel evolution of a type IV secretion system in radiating lineages of the
632 host-restricted bacterial pathogen *Bartonella*. PLoS Genet. 7: e1001296. doi:
633 10.1371/journal.pgen.1001296

634 Feldhaar H, et al. 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont
635 *Blochmannia*. BMC Biol. 5: 48–59. doi: 10.1186/1741-7007-5-48

636 Garber AI, et al. 2024. Retention of an endosymbiont for the production of a single molecule.
637 Genome Biol. Evol. 16. doi: 10.1093/gbe/evae075

638 Grant JR, et al. 2023. Proksee: in-depth characterization and visualization of bacterial genomes.
639 Nucleic Acids Res. 51: W484-w492. doi: 10.1093/nar/gkad326

640 Gruber-Vodicka HR, Seah Brandon KB, Pruesse E 2020. phyloFlash: rapid small-subunit rRNA
641 profiling and targeted assembly from metagenomes. mSystems 5: 10.1128/msystems.00920-
642 00920. doi: 10.1128/msystems.00920-20

643 Guindon S, Gascuel O 2003. A simple, fast, and accurate algorithm to estimate large phylogenies
644 by maximum likelihood. Syst. Biol. 52: 696-704. doi: 10.1080/10635150390235520

645 Hansen AK, Moran NA 2014. The impact of microbial symbionts on host plant utilization by
646 herbivorous insects. Mol. Ecol. 23: 1473-1496. doi: 10.1111/mec.12421

647 Hu Y, et al. 2018. Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-
648 recycling gut microbiome. Nat. Commun. 9: 964–978. doi: 10.1038/s41467-018-03357-y

649 Husnik F, McCutcheon JP 2016. Repeated replacement of an intrabacterial symbiont in the
650 tripartite nested mealybug symbiosis. Proc. Natl. Acad. Sci. USA 113. doi:
651 10.1073/pnas.1603910113

652 Ivens ABF, Gadau A, Kiers ET, Kronauer DJC 2017. Can social partnerships influence the
653 microbiome? Insights from ant farmers and their trophobiont mutualists. Mol. Ecol. 27: 1898-1914.
654 doi: 10.1111/mec.14506

655 Kobiałka M, Świerczewski D, Walczak M, Urbańczyk W 2025. Extremely distinct microbial
656 communities in closely related leafhopper subfamilies: Typhlocybinae and Eurymelinae
657 (Cicadellidae, Hemiptera). mSystems 10: e00603-00625. doi: 10.1128/msystems.00603-25

658 Koga R, Bennett GM, Cryan JR, Moran NA 2013. Evolutionary replacement of obligate symbionts
659 in an ancient and diverse insect lineage. Environ. Microbiol. 15: 2073-2081. doi: 10.1111/1462-
660 2920.12121

661 Kurtz S, et al. 2004. Versatile and open software for comparing large genomes. Genome Biol. 5:
662 R12. doi: 10.1186/gb-2004-5-2-r12

663 Letunic I, Bork P 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
664 display and annotation. Nucleic Acids Res. 49: W293–W296. doi: 10.1093/nar/gkab301

665 Li D, Liu C-M, Luo R, Sadakane K, Lam T-W 2015. MEGAHIT: an ultra-fast single-node solution
666 for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:
667 1674-1676. doi: 10.1093/bioinformatics/btv033

668 Lukasik P, Kolasa MR 2024. With a little help from my friends: the roles of microbial symbionts in
669 insect populations and communities. Philos. Trans. R. Soc. Lond., Ser. B: Biol. Sci. 379:
670 20230122. doi: 10.1098/rstb.2023.0122

671 Łukasik P, et al. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of
672 cicadas. Proc. Natl. Acad. Sci. USA 115: E226-E235. doi: 10.1073/pnas.1712321115

673 Ma M, Zhu B, Zhang D, Łukasik P, Hu Y 2025. Symbiotic solutions for colony nutrition: conserved
674 nitrogen recycling within the bacterial pouch of *Tetraponera* ants. bioRxiv. doi:
675 10.1101/2025.03.13.643113

676 Manzano-Marin A, Simon JC, Latorre A 2016. Reinventing the wheel and making it round again:
677 evolutionary convergence in *Buchnera-Serratia* symbiotic consortia between the distantly related
678 Lachninae aphids *Tuberolachnus salignus* and *Cinara cedri*. Genome Biol. Evol. 8: 1440-1458.
679 doi: 10.1093/gbe/evw085

680 Martinez AJ, et al. 2019. Angiosperm to Gymnosperm host-plant switch entails shifts in microbiota
681 of the *Welwitschia* bug, *Probergrothius angolensis* (Distant, 1902). Mol. Ecol. 28: 5172-5187. doi:
682 10.1111/mec.15281

683 Matsuura Y, Kikuchi Y, Meng XY, Koga R, Fukatsu T 2012. Novel clade of alphaproteobacterial
684 endosymbionts associated with stinkbugs and other arthropods. Appl. Environ. Microbiol. 78:
685 4149-4156. doi: 10.1128/AEM.00673-12

686 Matsuura Y, et al. 2018. Recurrent symbiont recruitment from fungal parasites in cicadas. Proc.
687 Natl. Acad. Sci. USA 115: E5970-E5979. doi: 10.1073/pnas.1803245115

688 McCutcheon JP, Boyd BM, Dale C 2019. The life of an insect endosymbiont from the cradle to
689 the grave. Curr. Biol. 29: R485-R495. doi: 10.1016/j.cub.2019.03.032

690 McCutcheon JP, Garber AI, Spencer N, Warren JM 2024. How do bacterial endosymbionts work
691 with so few genes? PLoS Biol. 22: e3002577. doi: 10.1371/journal.pbio.3002577

692 McCutcheon JP, McDonald BR, Moran NA 2009. Convergent evolution of metabolic roles in
693 bacterial co-symbionts of insects. Proc. Natl. Acad. Sci. USA 106: 15394-15399.

694 Michalik A, et al. 2021. Alternative transmission patterns in independently acquired nutritional
695 cosymbionts of Dictyopharidae planthoppers. mBio 12: e0122821. doi: 10.1128/mBio.01228-21

696 Michalik A, Diego CF, Szklarzewicz T, Stroiński A, Łukasik P 2024. Facultatively intrabacterial
697 localization of a planthopper endosymbiont as an adaptation to its vertical transmission.
698 mSystems 9: e00634-00624. doi: 10.1128/msystems.00634-24

699 Michalik A, et al. 2025. The tiniest genomes shrink much further: extreme reductive evolution in
700 planthopper symbionts. bioRxiv. doi: 10.1101/2025.05.16.654412

701 Michalik A, et al. 2023. Variable organization of symbiont-containing tissue across planthoppers
702 hosting different heritable endosymbionts. Front Physiol. 14: 1135346. doi:
703 10.3389/fphys.2023.1135346

704 Moran NA 2007. Symbiosis as an adaptive process and source of phenotypic complexity. Proc.
705 Natl. Acad. Sci. USA 104: 8627-8633.

706 Moran NA, Tran P, Gerardo NM 2005. Symbiosis and insect diversification: an ancient symbiont
707 of sap-feeding insects from the bacterial phylum Bacteroidetes. Appl. Environ. Microbiol. 71:
708 8802-8810. doi: 10.1128/AEM.71.12.8802-8810.2005

709 Neuvonen MM, et al. 2016. The genome of Rhizobiales bacteria in predatory ants reveals urease
710 gene functions but no genes for nitrogen fixation. Sci. Rep. 6: 39197. doi: 10.1038/srep39197

711 Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW 2015. CheckM: assessing the
712 quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome
713 Res. 25: 1043–1055. doi: 10.1101/gr.186072.114

714 Pons MJ, Gomes C, Del Valle-Mendoza J, Ruiz J 2016. Carrion's disease: more than a sand fly-
715 vectored illness. PLoS Pathog. 12: e1005863. doi: 10.1371/journal.ppat.1005863

716 Renoz F, Arai H, Pons I 2023. The genus *Sodalis* as a resource for understanding the multifaceted
717 evolution of bacterial symbiosis in insects. Symbiosis 92: 187-208. doi: 10.1007/s13199-023-
718 00966-0

719 Rozewicki J, Li S, Amada KM, Standley DM, Katoh K 2019. MAFFT-DASH: integrated protein
720 sequence and structural alignment. Nucleic Acids Res. 47: W5–W10. doi: 10.1093/nar/gkz342

721 Russell JA, Moreau CS, Goldman-Huertas B, Fujiwara M, Lohman DJ 2009. Bacterial gut
722 symbionts are tightly linked with the evolution of herbivory in ants. Proc. Natl. Acad. Sci. USA 106:
723 21236–21241. doi: 10.1073/pnas.0907926106

724 Seemann T 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30: 2068-2069.
725 doi: 10.1093/bioinformatics/btu153

726 Segers FH, Kesnerova L, Kosoy M, Engel P 2017. Genomic changes associated with the
727 evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J. 11: 1232-
728 1244. doi: 10.1038/ismej.2016.201

729 Siehl R, Vyhnal K, Goffredi SK 2024. Friendly fungi: tropical insect families form partnerships with
730 intracellular fungi related to pathogens. iScience 27: 110674. doi: 10.1016/j.isci.2024.110674

731 Smith EJ, Vizueta J, Younger MA, Mullen SP, Traniello JFA 2023. Dietary diversity, sociality, and
732 the evolution of ant gustation. Front. Ecol. Evol. 11. doi: 10.3389/fevo.2023.1175719

733 Stamatakis A 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
734 phylogenies. Bioinformatics 30: 1312–1313. doi: 10.1093/bioinformatics/btu033

735 Sudakaran S, Kost C, Kaltenpoth M 2017. Symbiont acquisition and replacement as a source of
736 ecological innovation. Trends Microbiol. 25: 375-390. doi: 10.1016/j.tim.2017.02.014

737 Susko E, Roger AJ 2021. Long branch attraction biases in phylogenetics. Syst. Biol. 70: 838-843.
738 doi: 10.1093/sysbio/syab001

739 Syberg-Olsen MJ, Garber AI, Keeling PJ, McCutcheon JP, Husnik F 2022. Pseudofinder:
740 detection of pseudogenes in prokaryotic genomes. Mol. Biol. Evol. 39. doi:
741 10.1093/molbev/msac153

742 Talavera G, Castresana J 2007. Improvement of phylogenies after removing divergent and
743 ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56: 564–577. doi:
744 10.1080/10635150701472164

745 The UniProt C 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res.
746 49: D480-D489. doi: 10.1093/nar/gkaa1100

747 Tlaskal V, Pylro VS, Zifcakova L, Baldrian P 2021. Ecological divergence within the
748 enterobacterial genus *Sodalis*: from insect symbionts to inhabitants of decomposing deadwood.
749 Front. Microbiol. 12: 668644. doi: 10.3389/fmicb.2021.668644

750 Urban JM, Cryan JR 2012. Two ancient bacterial endosymbionts have coevolved with the
751 planthoppers (Insecta: Hemiptera: Fulgoroidea). BMC Evol. Biol. 12: 87. doi: 10.1186/1471-2148-
752 12-87

753 Uritskiy GV, DiRuggiero J, Taylor J 2018. MetaWRAP—a flexible pipeline for genome-resolved
754 metagenomic data analysis. Microbiome 6: 158–171. doi: 10.1186/s40168-018-0541-1

755 Wagner A, Dehio C 2019. Role of distinct type-IV-secretion systems and secreted effector sets in
756 host adaptation by pathogenic *Bartonella* species. Cell. Microbiol. 21: e13004. doi:
757 10.1111/cmi.13004

758 Williams TJ, et al. 2021. Genome analysis of a Verrucomicrobial endosymbiont with a tiny genome
759 discovered in an Antarctic Lake. Front. Microbiol. 12: 674758. doi: 10.3389/fmicb.2021.674758

760 Xiong Q, Sopko B, Klimov PB, Hubert J 2024. A novel *Bartonella*-like bacterium forms an
761 interdependent mutualistic symbiosis with its host, the stored-product mite *Tyrophagus*
762 *putrescentiae*. mSystems 9: e0082923. doi: 10.1128/msystems.00829-23
763