ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Issue: *The Year in Evolutionary Biology*
REVIEW ARTICLE

Dissecting genome reduction and trait loss in insect endosymbionts

Amparo Latorre[1,2] and Alejandro Manzano-Marín[1]

[1]Institut Cavanilles de Biodiversitat I Biologia Evolutiva, Universitat de Valencia, C/Catedrático José Beltrán, Paterna, Valencia, Spain. [2]Área de Genómica y Salud de la Fundación para el fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO)-Salud Pública, València, Spain

Address for correspondence: Amparo Latorre, Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de Valencia, C/Catedrático José Beltrán 2, 46071 Paterna, Valencia, Spain. amparo.latorre@uv.es

Symbiosis has played a major role in eukaryotic evolution beyond the origin of the eukaryotic cell. Thus, organisms across the tree of life are associated with diverse microbial partners, conferring to the host new adaptive traits that enable it to explore new niches. This is the case for insects thriving on unbalanced diets, which harbor mutualistic intracellular microorganisms, mostly bacteria that supply them with the required nutrients. As a consequence of the lifestyle change, from free-living to host-associated mutualist, a bacterium undergoes many structural and metabolic changes, of which genome shrinkage is the most dramatic. The trend toward genome size reduction in endosymbiotic bacteria is associated with large-scale gene loss, reflecting the lack of an effective selection mechanism to maintain genes that are rendered superfluous by the constant and rich environment provided by the host. This genome-reduction syndrome is so strong that it has generated the smallest bacterial genomes found to date, whose gene contents are so limited that their status as cellular entities is questionable. The recent availability of data on several endosymbiotic bacteria is enabling us to form a comprehensive picture of the genome-reduction process and the phenotypic consequences for the dwindling symbiont.

Keywords: insect endosymbionts; mutualism; genome reduction; trait loss

Introduction

Beyond the origin of evolutionary leaps such as the eukaryotic cell, symbiotic associations have been documented in practically every major branch of the tree of life.[1] These observations, together with the data provided from studies of different associations established between prokaryotes and eukaryotes, reinforce the view that symbiosis is an important mechanism in the emergence of evolutionary innovations in eukaryotes. As shown throughout this review, a common feature of the maternally transmitted symbionts is genome reduction by gene loss. Thus, this process is key to understanding how the evolutionary loss of traits relates to the symbiont's genome reduction and how this can bear on the evolutionary fate of their associates.

Only recently, the genomic era has unlocked the genetic knowledge of nonculturable microbes involved in symbiosis, facilitating comparisons among different host-associated bacteria throughout the spectrum encompassing free-living organisms to obligate intracellular endosymbionts. In addition to the role played by symbiotic associations in the origin and evolution of eukaryotic cells, genomics has also revealed that the process of establishing new structures, biochemistries, or behaviors through symbiosis is an ongoing phenomenon in the evolution of life. The event that triggers a symbiotic association is more or less fortuitous. The fact that the players remain genetically, biochemically, or metabolically linked will depend on the evolutionary success of the association. It is well established that symbiotic integration is a process that profoundly changes the genome of the symbiont's free-living ancestor and, depending on the type of symbiotic relationship (e.g., mutualistic or parasitic, facultative or obligate, etc.), the age of the association (old or recent), and host

doi: 10.1111/nyas.13222

necessities (e.g., nutritional, defensive, or waste recycling), the observed changes will be more or less dramatic (reviewed in Refs. 1–6).

The first symbioses: what can we learn?

Although the nature of the first eukaryotic host cell is a matter of heated debate among cell evolutionists, mainly due to the lack of evolutionary intermediates,[7] two main scenarios have been put forward: the "mito-early" and "mito-late." The mito-early scenario proposes that mitochondrial endosymbiosis occurred in a simple proto-eukaryotic host,[8] while the mito-late suggests that a significant complexity was held by the proto-eukaryotic cell before the acquisition of the mitochondrial ancestor.[9,10] Nonetheless, the endosymbiotic origin of the organelles (plastids and mitochondria) cannot be denied.[11] While mitochondria evolved from free-living Alphaproteobacteria, which provided their host with respiration and energy-metabolism efficiency, chloroplasts evolved from what was once a free-living Cyanobacterium, which endowed its hosts with the ability to photosynthesize (chlorophyll-harboring prokaryotes became photosynthetic cells). The ability to breathe oxygen as a result of the acquisition of mitochondria led to the origin of animals, whereas the photosynthetic ability acquired subsequent to chloroplasts gave rise to plants. Therefore, both mitochondria and chloroplasts originated from free-living bacteria, whose tiny descendants are still among us.

It is well known that organellar genomes encode only a small fraction of the organelle's proteins, with the majority of these being encoded by the nuclear genome (reviewed in Ref. 12). This has resulted from the relocation of several of the organelle's genes to the host nucleus, followed by the evolution of targeting sequences in the nuclear-encoded copies and complex protein-import machineries. The latter was a crucial component for the transformation of the former endosymbionts into cellular organelles. This process has ultimately resulted in the genome reduction of the symbionts, eliminating genes that have become unnecessary in the new intracellular environment or redundant with the host genes. Additionally, genes of nuclear origin, mainly involved in replication, transcription, cell division, and signal transduction, have replaced those from the organelle, thus driving further genome reduction.[13] The similarity in gene content within contemporary plastids and mitochondria suggests that most organelle genes were transferred massively in the early evolution of both mitochondria (reviewed in Ref. 14) and plastids.[15] The subsequent tempo of gene-transfer events has been punctuated by bursts of transfer interspersed with long periods of stasis. Also, many genes show a patchy distribution across extant organelle genomes, implying recurrent transfers and convergent losses (reviewed in Ref. 16). For the symbionts, all of these processes have implied an irreversible loss of autonomy. Many nuclear genes of organellar origin were able to supply proteins to other cellular compartments and thus became "free" in terms of being able to evolve new functions.[15] In this vein, mitochondria and plastids are simply "the luckiest of a longstanding series of doomed endosymbionts who were saved by transfer of genes to the nucleus."[17]

Symbiosis in insects: an overview

In 1953, the German entomologist Paul Buchner published the first big compendium describing symbiotic associations between insects (class Insecta Linnaeus, 1758) and microorganisms[18] (translated into English in 1965[19]). He defined the term endosymbiosis as a "well-regulated and essentially undisturbed cooperative living between two differently constituted partners." In his fascinating work, he explored the endosymbionts present in many hemipteran families within the Coccoidea, Aphidoidea, Aleyrodoidea, Psylloidea, and Membracoidea. Equipped with a microscope, he was able to observe consistent infections across various individuals, as well as to detect that many of these microorganisms were hosted inside specific cells, termed mycetocytes or bacteriocytes (specialized host cells that harbor the symbiotic microorganisms), within a distinct organ-like structure, termed the bacteriome (or *mycetome* in Buchner's book). He was the first to propose "that microbes and insects not only show an amazing biodiversity in themselves, but they often come together and take evolutionary paths to persistent physical association." Since then, symbiotic associations have been broadly studied in numerous insects (including many of the ones first analyzed by Buchner), and a large number of genomic, biochemical, and physiological studies have been conducted mainly in insect endosymbionts (revised in Refs. 1, 2, 20, and 21). These

analyses have revealed that symbioses between insects and microorganisms are both diverse (in terms of both taxonomic origin and basis) and widespread. Although some fungal endosymbionts have been reported in insects, hitherto the majority of reported cases involve bacteria, and thus here we will focus on bacterial endosymbionts within insects.

In most cases, these insect–symbiont relationships have a nutrition-based foundation, with the endosymbiotic genome coding for the biosynthesis of the essential compounds lacking from the host diet, such as plant phloem (mainly deficient in essential amino acids)[22,23] or mammalian blood (mainly deficient in B vitamins).[24] For example, *Buchnera*, *Tremblaya*, and *Portiera* endosymbionts enable their hosts (aphids, mealybugs, and whiteflies, respectively) to survive on a strict phloem diet, having the capacity to synthesize essential amino acids and some vitamins.[25–27] On the other hand, *Wigglesworthia* and *Riesia* symbionts from *Glossina* flies and *Pediculus* lice, respectively, possess genomes that are capable of producing B vitamins.[28–31] In return for their contribution, the host provides a stable environment for the bacteria with a permanent supply of resources, thereby making the association a mutualistic one (a term reserved for the symbiotic relationships where each partner benefits from the activity of the other). Supporting this hypothesis, experimental studies dealing with the generation of aposymbiotic insects in the pea aphid *Acyrthosiphon pisum*,[32,33] the tsetse fly *Glossina morsitans*,[34] and different cockroaches[35] have shown that aposymbiotic females generally display reduced reproduction rates and a decrease in fertility or even complete sterility. Hence, these symbionts are required for the correct development of their hosts and thus have been termed *primary obligate endosymbionts*.

Apart from primary obligate endosymbionts, insects can establish associations with additional bacteria, termed *secondary endosymbionts*. These additional bacteria can be of facultative or obligate nature. Contrary to obligate endosymbionts, facultative ones are dispensable. However, under certain environmental conditions, some secondary facultative endosymbionts can endow the host with beneficial traits and have even been shown to somehow alter the host's biology, for example, via manipulation of reproduction (reviewed in Refs. 36–38). Finally, in some cases, a secondary endosymbiont can evolve to become an obligate partner, and if the primary is already present, a microbial co-obligate consortium can be established (reviewed in Ref. 1).

Genome reduction in endosymbionts: gene loss

The five first complete genomes from insect endosymbiotic bacteria were *Buchnera aphidicola*,[25,39,40] *Wigglesworthia glossinidia*,[28] and *Blochmania floridanus*,[41] primary obligate endosymbionts of aphids, tsetse flies, and carpenter's ants, respectively. It was immediately evident that these symbionts held highly reduced genomes (ranging from 616 to 706 kb) with a relatively similar number of genes in each functional category, based on the clusters of orthologous genes classification, pointing to similar evolutionary forces acting on these organisms. Gene losses in these endosymbiotic bacteria include many cell membrane proteins, DNA repair and recombination genes, all mobile elements, and whole biosynthetic pathways for nutrients that can be obtained through the host's diet. Comparative genomics revealed that, even though these genomes encoded between 564 and 698 genes, they shared only 313 genes, a fact leading to the proposal that this number could be close to the minimum gene set necessary to sustain endosymbiotic life.[41] In 2006, two smaller endosymbiotic bacterial genomes were published: *B. aphidicola* BCc (422 kb and 362 protein-coding genes)[42] and *Carsonella ruddii* (160 kb and 182 protein-coding genes),[43] primary endosymbionts of the aphid *Cinara cedri* and the psilid *Pachypsylla venusta*, respectively. In both cases, the gene repertoire seemed to be insufficient for these bacteria to meet their host's needs, and therefore their status as endosymbionts was questioned.[42–45] More recently, even more drastically reduced genomes have been discovered in cicadas,[46] spittlebugs,[3] mealybugs,[27] and leafhoppers,[47] sparking a renewed interest in the limits of genome reduction in endosymbiotic bacteria.[4,48,49]

Although the information derived from the extremely reduced genomes of long-term primary endosymbionts is valuable, it only provides us with information about the last steps in the process of endosymbiont integration. Comparative analysis of the bacterial genes and genomes of those primary endosymbionts revealed that, in the process toward an obligate lifestyle, bacteria experience major genetic and phenotypic changes, which

can be detected when compared against free-living relatives (reviewed in Refs. 1, 2, 4, 6, and 21). These changes mainly include a bias toward an adenine- and thymine-rich genome, an accumulation of small deleterious mutations, an accelerated sequence evolution, an increase in the number of nonsynonymous substitutions, a loss of mobile elements (mainly insertion sequences), a loss of recombination events, a high degree of synteny, and a massive reduction in genome size. Similar to organelles, the reduction in genome size is associated with the loss of a huge number of genes, mainly genes that have become unnecessary in the new nutrient-rich intracellular environment or that code for functions now carried out by the host. However, different bacteria possess a particular set of retained genes. The losses are determined by the specific host's needs (e.g., essential amino acids in phloem feeders or B vitamins in blood suckers), but can also reflect the particular processes of gene loss undergone by each symbiotic lineage, giving rise to differentially retained genes with similar or equivalent functions.[50] Regarding the retained genes, they belong to two main categories: those essential for maintenance of the bacterial cell and those that underlie the mutualistic association with its host. While the first set tend to be quite convergent (mainly genes involved in informational processes), the second set is particular to the specific nutrient requirements of the host. In some cases, the ongoing reduction process continues, and essential genes in both categories are also lost, rendering the smallest bacterial genomes ever found.[4] It is then conceivable that naturally evolved, nearly minimal gene sets may contain substantial differences.

Homologous recombination between repeated elements catalyzes large inversion and deletions.[51,52] Although recombination is diminished or lost in long-term obligate endosymbionts, comparisons of genome architecture in some strains suggest a historic period of large-scale rearrangements. In this respect, the distinct genome contents, gene order, and dynamics of facultative and obligate insect mutualists may be explained by their status as recent versus ancient stages along a similar evolutionary trajectory.[39,40]

The evolutionary processes that prompt the aforementioned genomic changes are a relaxation of natural selection and the continuous bottlenecks triggered by vertical transmission. The former results from a combination of the symbiont now residing in a more stable environment (inside the host), making certain free-living functions unnecessary, and generating genetic/metabolic redundancy between the host and the symbiont (and/or another symbiont), together promoting substantial gene loss (reviewed in Refs. 1 and 2). The latter is a result of the strictly vertical transmission mode of endosymbionts, where, owing to bottlenecks, only a small bacterial subpopulation will pass to the next generation, resulting in very low effective population sizes, which favors the action of random genetic drift.[6,53] This, combined with a lack of recombination, leads to the irreversible fixation of slightly deleterious mutations, a process known as Muller's ratchet. The lack of recombination results from the loss of genes involved in DNA recombination, repair, and uptake mechanisms, all of which are common features of currently sequenced bacterial endosymbionts (reviewed in Refs. 1, 2, 4, 6, 54 and 55). The irreversible accumulation of deleterious mutations has a notable effect on many genes, as it alters the structure and function of the corresponding proteins.[40] However, radical changes in absolutely essential genes for endosymbiosis could compensate for the detrimental effect of previously fixed mutations. This is the case for the chaperone GroEL, a protein that is overexpressed in endosymbiotic bacteria and seems to participate in the correct folding of many damaged proteins, thus buffering the effect of slightly deleterious mutations.[56] In fact, a number of elegant experiments carried out in *Escherichia coli* have shown that overexpression of GroEL in bacteria evolving under continuous bottleneck, and thus with a strong genetic drift, can avoid extinction or rescue bacterial cells.[56–58]

In recent years, the genomic era has provided the opportunity to investigate a plethora of endosymbiotic associations. These have afforded snapshots of the various stages a free-living bacterium undergoes on its way to becoming an organelle-like entity. In this review, we will examine the different stages undergone by naturally reduced endosymbiotic genomes. We will discuss the characteristics defining each stage, as well as the factors and evolutionary dynamics that promote or truncate a symbiont's progression toward establishing a state of stable mutualistic intracellular symbiosis.

Major steps leading to an obligate endosymbiont: dissecting the process

As stated before, analyzing genomic features of endosymbiotic bacteria can reveal the level of integration of the symbiont. In general, early stages of genome reduction are characterized by genome sizes intermediate between those of free-living organisms and long-term endosymbionts, with a high density of insertion sequences and other mobile elements, the formation of pseudogenes, multiple genome rearrangements, and deletion of chromosome fragments. Facultative endosymbionts show most, if not all, of the characteristics of the early stages of genome reduction. However, in more anciently evolved symbionts, such as many obligate intracellular mutualistic symbioses, mobile elements and most pseudogenes have been eliminated; therefore, the genome architecture of bacteria in different host lineages tends to be highly syntenic.

The coexistence of different bacteria within the same host, one primary obligate endosymbiont and one (or more) secondary symbiont that has started its adaptation to intracellular life, raises the possibility of the primary endosymbiont being either complemented or replaced by the healthier facultative bacterium.[59] Moreover, the finding of co-obligate endosymbionts has illustrated the labile boundary between facultative and obligate endosymbiotic relationships.[42,47,60–62]

On the basis of the available genomic data from different host-associated bacteria, Toft and Andersson[5] divided the general genome reduction process of host-associated bacteria (namely intracellular) into five stages. These range from being free-living extracellular (stage 1) passing through a facultative intracellular (stage 2) and obligate stage (stages 3 and 4 (mutualist)) to becoming an organelle (stage 5). Since then, many new data have become available, showing that, although ancient events are difficult to reconstruct, different lines of evidence suggest that some primary endosymbionts may have arisen from facultative ones in insects. In this case, facultative and obligate associates may represent points along a continuous spectrum of symbiosis. Thus, we will refine the scenario analyzing the different stages (from 1 to 4) in this reduction process on the basis of up-to-date genomic and experimental analyses of different insect endosymbionts that have naturally evolved dwindling genomes (Fig. 1).

Stage 1: free-living "potential" mutualists—the newcomers

In theory, any free-living bacterium could start a symbiotic relationship that could potentially lead toward becoming an obligate mutualist. In fact, the microorganisms associated with insects are quite diverse (see Ref. 63), and we now know that they did not originate from a single infection event but rather have independent origins.[64] However, only a few phylogenetic clades contain species that have coevolved to become intracellular mutualists: Alphaproteobacteria and Gammaproteobacteria are the most widely distributed, although Betaproteobacteria and Bacteroidetes have also been found. The overrepresentation of these clades could be either an effect of sampling bias or the fact that certain bacterial groups are more prone to evolve intimate interactions with eukaryotic host cells.[5] Whatever the reason, how mutualistic bacteria have originated from a free-living ancestor has not yet been resolved.

In the first genomic studies, only an endosymbiont's free-living relatives were considered as representative of this stage. For example, *E. coli* and related enterobacterial species have been used as a model for free-living gammaproteobacterial endosymbionts, mainly as outgroups in phylogenetic reconstructions.[65,66] More recently, some free-living and/or pathogenic species from the genera *Serratia* and *Sodalis* have been key to our understanding of the changes experienced by a bacterium in its transition from being a free-living extracellular organism to becoming a co-obligate intracellular one, such as *Serratia symbiotica* in aphids[50] or *Sodalis* spp. symbionts in weevils and tsetse flies.[67]

Recent investigations into Japanese populations of the stink bug *Plautia stalii* have provided unique, and to our knowledge unprecedented, insights into the evolutionary transition from a free-living lifestyle to an obligate mutualistic one (Table 1).[68] This insect species has evolved different kinds of obligate symbiotic associations, which display a geographical pattern. Briefly, while populations from the temperate mainland harbor a fixed obligate not-yet-cultured symbiont (termed A), populations from the subtropical islands present a prevalent not-yet-cultured one (termed B), with some populations housing different culturable associates (termed C–F). As expected, symbionts A and B hold reduced

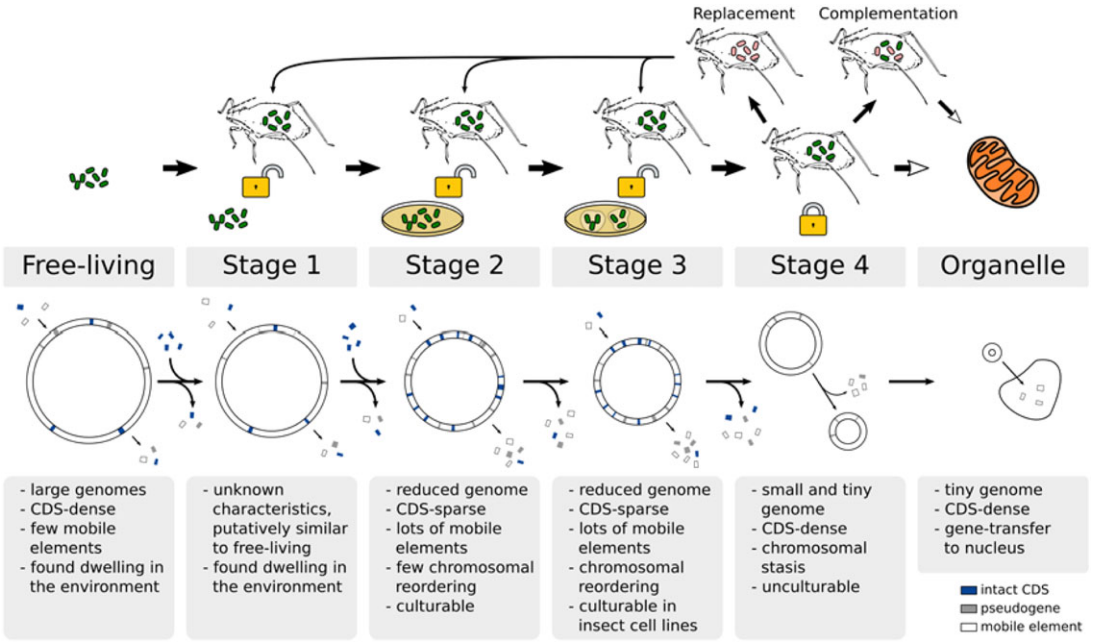

Figure 1. **Proposed stages of genome reduction and their biological traits. Graphical summary and principal characteristics of the proposed stages a free-living bacterium goes through on its way to becoming an organelle, based on current knowledge of genome reduction. Although replacement can happen at any given stage, it is represented in stage 4 to be congruent with the text and facilitate interpretation of the figure. Open locks stand for flexible associations and closed locks for deeply rooted ones. Unfilled arrowheads pointing to the organelle stage represent the lack of an identified insect endosymbiotic lineage giving rise to an organelle. CDS, coding sequences.**

genomes (2.4–3.9 Mbp) when compared with the closely related free-living *Pantoaea*, whereas symbionts C–F hold larger ones (4.7–5.5 Mbp). Among the most striking aspects of this study are the experiments performed on the environmental uptake of symbionts by surface-sterilized eggs, which prevented newborn infection. The authors found that, by rearing these symbiont-free insects on soil collected at three *P. stali* habitats, just a few individuals reached adulthood and presented a normal phenotype (7.1%), similar to natural symbiont-harboring *P. stali*. Surprisingly, while little more than half of these individuals (39 of 71) were found to have taken up symbionts C–E (revealing their free-living presence in the natural environment of *P. stali*), the remaining 32 established obligate associations with different environmental bacteria (termed X1–X6) unrelated to A–F symbionts. Interestingly, the free-living X1–X6 were not found as naturally occurring symbionts of *P. stali*, placing them as potential mutualistic symbionts, which may be a source for the evolution of new obligate symbiotic relationships.

Therefore, we consider bacteria X1–X6 (and tentatively C–E) as being in stage 1 of integration to an obligate intracellular mutualistic lifestyle. At this stage, the symbiont would retain the ability to thrive in the environment where the potential hosts prosper and, as Hosokawa and collaborators observed,[68] these would still retain large genomes (as is the case for C–E but yet unknown for X1–X6). In terms of mobile elements, no data are available from X1 to X6, but given their free-living lifestyle, it can be assumed that, similar to other free-living bacteria, they harbor a controlled amount of mobile DNA. However, symbionts C–E already show an enrichment of mobile elements. Finally, their ability to infect and establish stable associations with their newly acquired hosts would be low, as reflected in their low infection frequency.

Stage 2: facultative intracellular (early stage)—still culturable

Little research has focused on the recently derived endosymbionts, and examples are scarce. At this stage, a free-living bacterium has effectively become

Table 1. Characteristics of free-living and stage 1 (potential mutualist) insect symbionts

Species (INSDC accession)	Strain or isolate	Genome size (Mb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
Free-living							
Escherichia coli (U00096)	K-12 MG1655	4.64	4165	Few	Few	Rod[69]	Derived from original K-12 strain isolated from *Homo sapiens* (Primates: Hominidae)[70]
Pantoea ananatis (CP003085-6)	PA13	4.87	4372	Few	Few	*Rod*[71,a]	Plant pathogen of unspecified *Oryza* sp. (Poales: Poaceae)[72]
Serratia marcescens (HG326223)	Db11	5.11	4709	Few	Few	Rod[73,a]	Streptomycin-resistant mutant of Db10 strain isolated from *Drosophila melanogaster* (Diptera: Drosophilidae)[74]
Sodalis praecaptivus (CP006569-70)	HS1	5.16	4282	Few	Few	Rod[75]	Isolated from a human wound caused by impalement with a branch of dead crab apple tree (*Malus* sp.) (Rosales: Rosaceae)[67]
Stage 1 potential mutualists							
Plautia stali symbionts X1–X6	–	–	–	–	–	–	Found in soil from *P. stali*'s (Hemiptera: Pentatomidae) natural environment[68]
P. stali symbiont C (BBOB00000000)	Ps-ISGKf53	5.14	–	Many	–	–	Both isolated as a naturally occurring mutualist of *P. stali* and found in soil from the insect's natural environment[68]
P. stali symbiont D (BBOC00000000)	Ps-ISGKm56	5.54	–	Many	–	–	Both isolated as a naturally occurring mutualist of *P. stali* and found in soil from the insect's natural environment[68]
P. stali symbiont E (BBOD00000000)	Ps-ISGKf70	5.41	–	Many	–	–	Both isolated as a naturally occurring mutualist of *P. stali* and found in soil from the insect's natural environment[68]

MEs, mobile elements; ψ , pseudogenes.
[a]The given characteristic is derived from common characteristics of related strains belonging to the same species.

a facultative intracellular one, but it is still able to grow on artificial medium if cultured by standard techniques.

Examples have been documented of endosymbiotic bacteria that still retain the ability to grow axenically on complex culture media (Table 2). These include *Sodalis glossinidius* (facultative symbiont of the tsetse fly and the first endosymbiont to be successfully cultured),[76] *S. symbiotica* strain CWBI-2.3^T (facultative endosymbiont of the aphid *Aphis fabae*),[78] *Arsenophonus arthropodicus* (secondary symbiont of the pigeon louse fly),[80] and

Table 2. Characteristics of stage 2 facultative intracellular (early-stage) insect symbionts

Species (INSDC accession)	Strain or isolate	Genome size (Mb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
Sodalis glossinidius (AP008232-5)	Morsitans	4.26	2516	Many	Many	Filamentous[76,77,a]	Facultative endosymbiont found both intracellularly in midgut epithelium and free in the midgut lumen of *Glossina morsitans morsitans* (Diptera: Glossinidae)[77]
Serratia symbiotica (CCES00000000)	CWBI-2.3^T	3.58	3398	Many	Many	Filamentous[78]	Facultative endosymbiont found intracellularly in sheath cells of the bacteriome of *Aphis fabae* (Hemiptera: Aphididae)[78]
Frankliniella occidentalis symbiont BFo1 (JMSO00000000)	BFo1	5.13	4829	–	Many	–	Symbiont associated with *F. occidentalis* (Thysanoptera: Thripidae)[79]
F. occidentalis symbiont BFo2 (JMSP00000000)	BFo2	3.10	3068	–	Many	–	Symbiont associated with *F. occidentalis* (Thysanoptera: Thripidae)[79]
Arsenophonus arthropodicus	–	–	–	–	–	–	Facultative endosymbiont found intracellularly in hemocytes of *Pseudolynchia canariensis* (Diptera: Hippoboscidae)[80]

MEs, mobile elements; ψ, pseudogenes.
[a]The given characteristic is derived from common characteristics of related strains belonging to the same species harbored by the same host species.

two endosymbionts from the western flower thrips, termed BFo1 and BFo2.[80] However, full genomes are available only for the first two endosymbionts (BFo1 and BFo2 have only highly fragmented, low-coverage draft genomes available). While the first two show genomes that are close in size to those of their free-living relatives (*S. glossinidius* (4.17 Mbp)[81] vs. *Sodalis praecaptivus* (4.7 Mb),[72] and *S. symbiotica* strain CWBI-2.3^T (3.6 Mb)[82] vs. *Serratia marcescens* strain Db11 (5.1 Mb)),[83] they show a great enrichment in mobile elements and pseudogenes. While the large number of mobile elements observed in these genomes could reflect an ineffectual purifying selection for insertion events, as well as occasional horizontal transfer, the massive number of pseudogenes might reflect a faster gene inactivation relative to DNA deletion.[92] These genomes also show a moderate number of rearrangements, relative to their free-living counterparts, a feature that is putatively a result of the increase in

mobile elements that promote genomic shuffling. These elements could also be promoting gene inactivation, as it has been described for *S. glossinidius*,[84] thus further contributing to the genomic erosion. Even though both *S. glossinidius* and *S. symbiotica* strain CWBI-2.3^{T} are of a facultative nature, they could endow their hosts with an adaptive advantage. For example, the specific elimination of *S. glossinidius* from the tsetse fly by the antibiotic streptozotocin seems to cause a reduction in longevity of F1 flies from treated females.[85] These conditional benefits to the host could lie behind the fixation and triggering of long-term vertical transfer of these symbionts within a population, thus promoting host–symbiont coevolution. In the case of *S. symbiotica* CWBI-2.3^{T}, to our knowledge, no experiments have been conducted to assess the contribution of this endosymbiont to its aphid host. Additionally, these symbionts are not confined to bacteriocytes but rather are found in close proximity to these (e.g., in sheath cells, syncytial cells located at the periphery of primary bacteriocytes)[78] and other tissues.[86] These histological properties reflect a more recent coevolutionary history with their hosts, contrasting the exclusive intrabacteriocyte localization of many long-term insect endosymbionts.

Therefore, stage 2 would include bacteria such as *S. glossinidius* and *S. symbiotica* strain CWBI-2.3^{T}. Unlike bacteria in stage 1, these would not be found thriving in the hosts' natural environment but would be confined to their hosts as endosymbionts. However, the endosymbiont would still retain the ability to grow axenically on complex media and establish stable cultures. These bacteria would have diverged little from their free-living counterparts and thus show few rearrangements and incidents of gene loss. However, they would already show a massive enrichment in mobile elements and pseudogenes, evidencing that they have started to become accommodated to their symbiotic system: host (tsetse fly and aphids) and primary endosymbiont (*Wigglesworthia* and *Buchnera*), respectively. Also, these bacteria tend to show a broad tissue tropism, not being necessarily confined to bacteriocytes. In summary, although generally facultative in nature, these bacteria could, under certain environmental conditions, provide advantages to the host, which could eventually lead to the evolution of obligate associations.

Stage 3: obligate intracellular mutualist (advanced stage)—no way back

The early sequencing of diverse insect facultative endosymbionts revealed interesting insights into these (Table 3). Deciphering the genome of *Wolbachia pipientis* strain wMel, an obligate intracellular reproductive manipulator (nonmutualist) from the fruit fly *Drosophila melanogaster*, evidenced the advanced stage of genome reduction of endosymbionts.[87] This genome is a mere 1.3 Mb and shows a massive enrichment in mobile elements, mainly insertion sequences. Since then, many more *Wolbachia* strains have been sequenced, revealing similar properties (see Ref. 88). The genome sequencing of facultative endosymbionts from the aphid *A. pisum* evidenced the generality of these characteristics in recently derived endosymbionts (*Hamiltonella defensa* strain 5AT,[89] *Regiella insecticola* strains LSR1 and 5.15,[90,91] and *S. symbiotica* strain Tucson[92]). These symbionts have already become dependent on the obligate primary endosymbiont *Buchnera* for some essential functions, having developed auxotrophies for various essential amino acids. However, although facultative for the host and primary endosymbiont, different strains of these symbionts have been shown to confer on their host a variety of convenient traits, such as acting as defensive symbionts against parasitoid wasps (*S. symbiotica*, *H. defensa*, and *R. insecticola*)[91,93,94] and fungal parasites (*R. insecticola*, *Rickettsiella* spp., and *Spiroplasma* spp.),[95,96] conferring survival after heat stress (*Rickettsia* spp. and *S. symbiotica*)[97,98] and even influencing their insect host preference for host plant.[99–101] These traits could lead to the eventual fixation of a facultative symbiont in a particular population if the environmental conditions providing that advantage are maintained during sufficient evolutionary time.[102,103]

Given the right triggers, facultative symbionts at this stage can become obligate endosymbionts. This is so for the early obligate associates *Sodalis pierantonius* strain SOPE (primary endosymbiont of the rice weevil[104–107]) and *S. symbiotica* strain SCt-VLC (secondary co-obligate endosymbiont of the cypress pine aphid).[62] Comparative analysis of the latter versus the facultative *S. symbiotica* from *A. pisum* revealed that, while their genomes possess very similar gene content and metabolic

Table 3. Characteristics of stage 3 obligate intracellular mutualist (advanced-stage) insect symbionts

Species (INSDC accession)	Strain or isolate	Genome size (Mb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
Serratia symbiotica (AENX00000000)	Tucson	2.57[a]	2098[a]	Many	Many	Filamentous[103]	Facultative endosymbiont found intracellularly in both sheath cells and bacteriocytes of *A. pisum* (Hemiptera: Aphididae)[103]
S. symbiotica (FR904230-48, HG934887-9)	SCt-VLC	2.49	1601	Many	Many	Filamentous[62,66]	Co-obligate mutualistic endosymbiont found intracellularly in both sheath cells and bacteriocytes of *C. tujafilina* (Hemiptera: Aphididae)[62,66]
Hamiltonella defensa (CP001277-8)	5AT	2.17	2158	Many	Many	Filamentous[103,b]	Facultative defensive endosymbiont[89] found intracellularly in both sheath cells and bacteriocytes of *A. pisum* (Hemiptera: Aphididae)[103]
Regiella insecticola (ACYF00000000)	LSR1	2.07[a]	1769[a]	Many	Many	Filamentous[103,b]	Facultative endosymbiont found intracellularly in both sheath cells and bacteriocytes of *A. pisum* (Hemiptera: Aphididae)[103]
R. insecticola (AGCA00000000)	5.15	2.01[a]	2313[a]	Many	Many	Filamentous[103,b]	Facultative defensive endosymbiont[87] found intracellularly in both sheath cells and bacteriocytes of *A. pisum* (Hemiptera: Aphididae)[103,b]
Sodalis pierantonius (CP006568)	SOPE	4.51	4080	Many	Many	Filamentous[104]	Obligate endosymbiont found intracellularly in bacteriocytes from *Sitophilus oryzae*[104]

MEs, mobile elements; ψ, pseudogenes.

[a]Numbers are based on highly fragmented genome assemblies and thus might be overestimated.

[b]The given characteristic is derived from common characteristics of related strains belonging to the same species harbored by the same host species.

[Corrections added on November 23, 2016, after first online publication: Reference citations were corrected in Table 3: row 2, column 7, "Filamentous[62,75]" was changed to "Filamentous[62,66]"; row 2, column 8, "[. . .](Hemiptera: Aphididae)[62,75]" was changed to "[. . .](Hemiptera: Aphididae)[62,66]".]

capabilities, their obligate and facultative status, respectively, are determined by gene losses in the primary endosymbiont *Buchnera*, which are compensated by *S. symbiotica*. This example reveals a possible route to becoming a fixed symbiont. These two early obligate symbionts contribute to

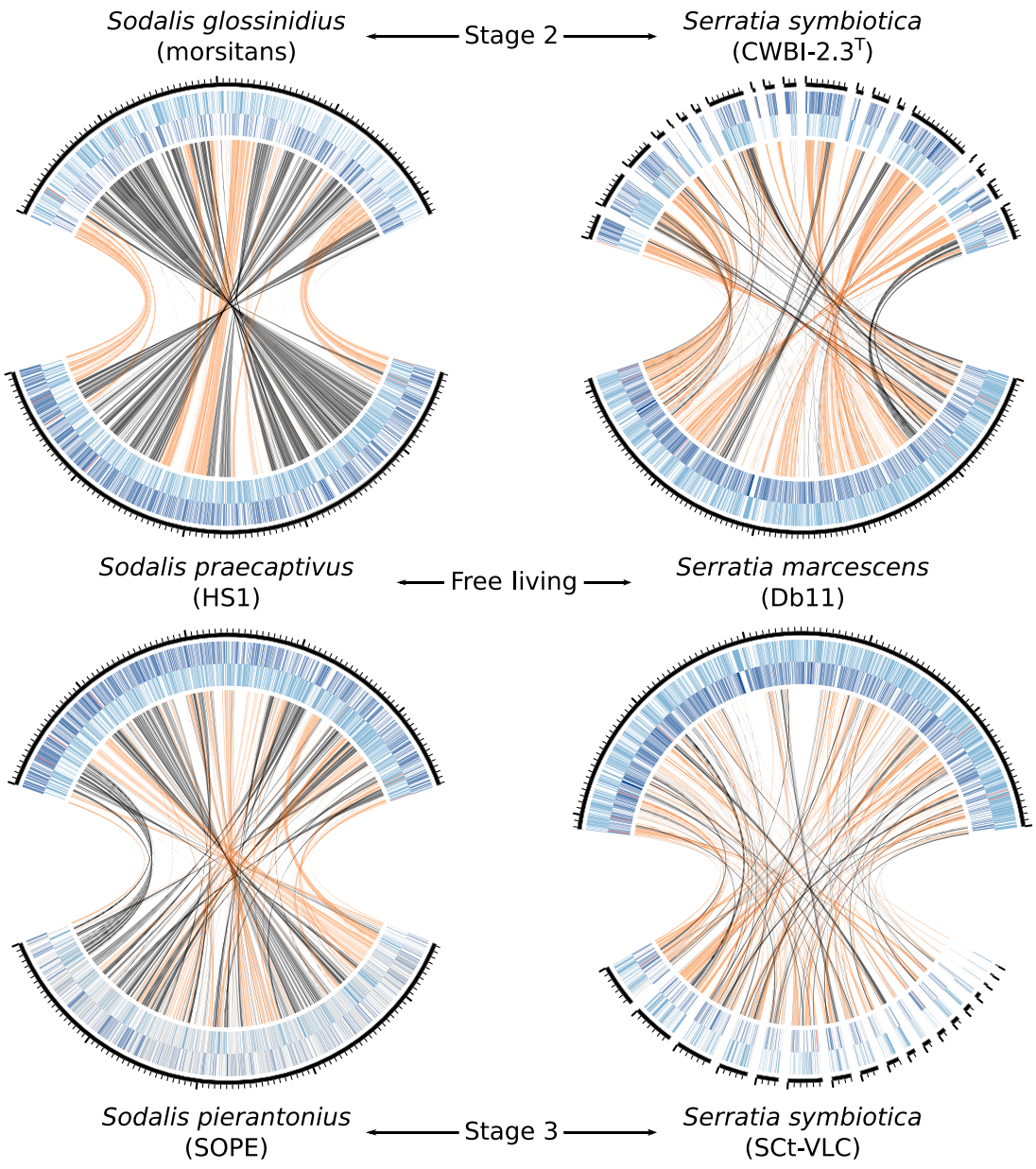

Figure 2. Genome rearrangement in stage 2 and stage 3 endosymbionts. Circular plots representing two examples, *Sodalis* (left) and *Serratia* (right), of free-living versus insect endosymbiont genome-wide synteny. Rings, from outer to innermost, represent features on the forward and reverse strands of the genomic contigs, respectively. Coding sequence features are shown in blue, pseudogenes in gray, rRNAs in red, and tRNAs in black. Orange and black lines connecting different contigs represent conserved single-copy genes in direct or inverse orientation, respectively.

the nutrition of their hosts and, similarly to facultative endosymbionts, hold large genomes (4.5[104] and 2.5 Mb,[62] respectively) with a great enrichment in mobile elements. Also, contrary to early culturable facultatives, their genomes have already undergone a massive number of rearrangements when compared with their free-living relatives (Fig. 2). This greater genome reduction compared with stage 2 symbionts could be the result of a longer and tighter evolutionary history with their hosts

in which they undergo longer periods of vertical transmission. Finally, while there are no reports of growing *S. pierantonius* strain SOPE or *S. symbiotica* strain SCt-VLC symbionts in pure culture, some facultative endosymbionts are able to establish stable cultures inside insect cell lines. For example, strains of *Wolbachia*,[108] *H. defensa*, and *R. insecticola*[109] have been successfully cultured in mosquito cell lines C6/36. [Correction added on November 23, 2016, after first online publication: In the preceding sentence, "*H. defense*, and *R. insecticola*[109] have been successfully cultured" was changed to "*H. defensa*, and *R. insecticola*[109] have been successfully cultured".] Additionally, the latter two also showed infection in *D. melanogaster* S2 and *Spodoptera frugiperda* (Lepidoptera: Noctuidae) SF9 cell lines. This apparent dependence on insect cells for the sustained growth of the aforementioned symbionts could be a result of further genetic erosion, promoting the development of auxotrophies that would impair its free-living capacity.

Thus, at this stage, the intracellular symbiont has lost its ability to grow axenically on artificial medium. However, cultivation in insect cell lines is possible, putatively reflecting their now obligate intracellular lifestyle. Their genomes are characterized by an extensive pseudogenization and gene loss, as well as a high level of genome rearrangement, when compared with their free-living relatives. These symbionts tend to reside in either sheath cells (as stage 2 endosymbionts) or inside bacteriocytes, whereas others have even developed dependence on the host's primary obligate endosymbiont. Finally, while some can be of facultative nature, others can already be of obligate (e.g., *S. pierantonius* strain SOPE) or even of co-obligate nature (e.g., *S. symbiotica* strain SCt-VLC).

Stage 4: obligate intracellular mutualist (extreme stage)—fragile symbiosis

Once the bacterium is locked into an intracellular habitat and the interaction with the host becomes stronger, genetic and metabolic losses or adjustments can enable the transition to a drastically reduced obligate intracellular mutualist, the fourth stage. At this stage, the ontogeny of a host can undergo reprogramming to produce specialized cells (bacteriocytes) to house these symbionts. Such relationships impose dramatic changes in both the host and the bacterium. It is important to note that there are exceptions to this intracellularity. Namely, stinkbugs from the Plataspidae family harbor a drastically reduced extracellular obligate mutualistic bacteria (*Ishikawaella capsulata*[110]) in crypts in the posterior of the midgut, which can be regarded as a "pseudo-bacteriome."[111,112] However, this bacterium shares the strict vertical transmission of endosymbionts, leading to continuous strong bottlenecks like those experienced by obligate intracellular symbionts, thereby leading to an attenuated purifying selection.

To date, many examples have been found of drastically reduced primary obligate intracellular mutualistic insect endosymbionts (Table 4). These symbionts commonly share traits, such as a reduced and gene-rich genome, fewer pseudogenes, and an inability to establish stable cultures, either in insect cell lines or in pure culture (reviewed in Ref. 63). The determination of the first whole *Buchnera* genome (also the first of an endosymbiont) revealed a high specialization of its genetic repertoire toward the production of essential amino acids,[25] which are lacking from its host's diet. This symbiotic-function specialization is not unique to *Buchnera*, since this has also been observed for other symbionts from phloem-feeder insects[26,110,113,114] and even blood suckers.[28,29,31,115] Beyond these two groups of organisms, other insects that feed on more complex diets (omnivorous), such as cockroaches and ants from the Campotini tribe, have also been found to house the obligate drastically reduced mutualistic intracellular symbionts from the genera *Blattabacterium*[116,117] and *Blochmannia*,[118,119] respectively. Through the study of these endosymbionts, it was found that their genomes are not only generally specialized for producing amino acids for their hosts, but also hold genes for a functional urease[41,120,121] and are thus involved in nitrogen recycling and thereby provide a nutritional upgrade to their hosts.[122–124]

Further sequencing of some of the host genomes and/or transcriptomes of these obligate endosymbionts has revealed intricate patterns of metabolic complementation between the host genome and the endosymbionts' genes, sometimes even involving ancient putative horizontal transfers from diverse bacteria.[125–127] The sequencing of various strains belonging to the same endosymbiotic species has revealed features about their genome architecture: *Buchnera*, *Blochmannia*, and *Blattabacterium*

Table 4. Characteristics of stage 4 obligate intracellular mutualist (extreme-stage) insect symbionts

Species (INSDC accession)	Strain or isolate	Genome size (kb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
Small genomes							
Buchnera aphidicola (BA000003, AP001070-1)	APS	655.73	574	None	Few	Spherical[103,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *A. pisum* (Hemiptera: Aphididae)[103,a]
B. aphidicola (AE016826, AAF492591)	BP	618.38	507	None	Few	Spherical[103,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *B. pistaciae* (Hemiptera: Aphididae)[103,a]
B. aphidicola (CP000263, AY438025, EU660486)	BCc	425.23	364	None	Few	Spherical[42]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Cinara cedri* (Hemiptera: Aphididae)[42]
Blattabacterium cuenoti (CP001487, CP002849)	Bge	640.94	591	None	Few	Rod[133]	Obligate mutualistic endosymbiont found intracellularly inside bacteriocytes within the fat body of *Blatella germanica* (Blattodea: Blattidae)[121]
Blochmannia pennsylvanicus (CP000016)	BPEN	791.65	610	None	Few	Filamentous[6,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Camponotus pennsylvanicus* (Hymenoptera: Formicidae)[6,a]
B. floridanus (BX248583)	–	405.56	583	None	Few	Filamentous[119,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Camponotus floridanus* (Hymenoptera: Formicidae)[119,a]
Baumannia cicadellinicola (CP008985)	BGSS	759.43	696	None	Few	Rod[134]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Graphocephala atropunctata* (Hemiptera: Cicadellidae)[134]

Continued

Table 4. *Continued*

Species (INSDC accession)	Strain or isolate	Genome size (kb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
B. cicadellinicola (CP000238)	Hc	686.20	595	None	Few	Spherical[134]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Homalodisca vitripennis* (Hemiptera: Cicadellidae)[134]
Riesia pediculicola (CP001085-6)	USDA	582.13	556	None	Few	Rod[135]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Bemisia tabaci* (Phthiraptera: Pediculidae)[135]
Portiera aleyrodidarum (CP003835)	BT-QVLC	357.47	247	None	Few	Pleomorphic blob[136,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Bemisia tabaci* (Hemiptera: Aleyrodidae)[136,a]
Serratia symbiotica (CP002295)	SCc	1762.77	677	None	Many	Spherical[60]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Cinara cedri* (Hemiptera: Aphididae)[60]
S. symbiotica (LN890288)	STs-Pazieg	650.32	495	None	Few	Spherical[50]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Tuberolachnus salignus* (Hemiptera: Aphididae)[50]
Wigglesworthia glossinidia (CP003315)[b]	Morsitans	724.73	620	None	Few	Rod[30]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Glossina morsitans morsitans* (Diptera: Glossinidae)[30]
W. glossinidia (BA000021)	Brevipalpis	703.00	616	None	*Few*	Rod[137,a]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Glossina brevipalpis* (Diptera: Glossinidae)[137,a]
Tiny genomes							
Sulcia muelleri (CP000770)	GWSS	245.53	227	None	Few	Pleomorphic tubular[138,a]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Homalodisca vitripennis* (Hemiptera: Cicadellidae)[138,a]

Continued

Table 4. *Continued*

Species (INSDC accession)	Strain or isolate	Genome size (kb)	Number of CDSs	MEs	ψ	Cell shape	Isolation source or niche
Tremblaya princeps (CP002244)	PCIT	138.93	125	None	Few	Pleomorphic[139,a]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Planococcus citri* (Hemiptera: Pseudococcidae)[139,a]
Zinderia insecticola (CP002161)	CARI	208.56	202	None	Few	Pleomorphic[3]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Clastoptera arizonana* (Hemiptera: Clastopteridae)[3]
Nasuia deltocephalinicola (CP006059)	ALF	112.09	137	None	Few	Pleomorphic[140,a]	Co-obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Macrosteles quadrilineatus* (Hemiptera: Cicadellidae)[140,a]
Carsonella ruddii (AP009180)	PV	159.66	182	None	Few	Pleomorphic tubular[43]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Pachypsylla venusta* (Hemiptera: Psyllidae)[43]
C. ruddii (CP003467)	DC	174.01	207	None	Few	Pleomorphic blob[44]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Diaphorina citri* (Hemiptera: Psyllidae)[44]
Hodgkinia cicadicola (CP001226)	Dsem	143.80	169	None	Few	Pleomorphic tubular[46]	Obligate mutualistic endosymbiont found intracellularly in bacteriocytes of *Diceroprocta semicincta* (Hemiptera: Cicadidae)[46]

MEs, mobile elements; ψ , pseudogenes.
[a]The given characteristic is derived from common characteristics of related strains belonging to the same or a related species.
[b]Missing reported plasmid accession.

genomes each display general genome-wide synteny, with the exception of a few inversions.[40,124,128] This characteristic results from their rapid genome reduction following their fixation as obligate endosymbionts in their respective hosts and confirms the monophyletic origin of the infection in the respective host's ancestors.

As mentioned earlier, the aforesaid symbionts hold small and gene-dense genomes with highly specialized roles. Given this metabolic specialization,

their genomes have become fragile, in the sense that gene losses can be particularly detrimental to their symbiotic role. Such losses have been found in specific lineages, for example, the loss in the ammonia assimilation capability in *Blochmannia vafer*[129] and of various essential amino acid–production genes in some *Blattabacterium* strains.[130,131] While these losses might be overcome by the complex diet of their hosts, other hosts with more restricted diets (such as strict phloem, xylem, or blood feeders) can suffer a great fitness reduction. Thus, these gene losses can trigger important changes in the symbiotic consortium, leading to either the complementation of the symbiotic function by a present facultative endosymbiont or the loss or the primary endosymbiont and subsequent replacement of it by a healthier endosymbiont (see Refs. 1, 60, and 132).

Complementation

The early sequencing of the small genomes of *Buchnera* from the aphid *C. cedri*[42] and *C. ruddii* from the psyllid *P. venusta*[43] revealed that their highly reduced genomes have lost some capabilities assumed to be essential for their hosts (e.g., biosynthesis of tryptophan) and their own cellular maintenance (e.g., cell envelope biogenesis). In the case of the former, it was later discovered that this aphid species harbored a secondary putatively co-obligate endosymbiont, *S. symbiotica*.[141] Later, the determination of the full genome sequences from both endosymbionts corroborated that they had indeed established a metabolic complementation for the biosynthesis of tryptophan and other nutrients.[60,142] Further analyses of other endosymbiotic systems from the Lachninae aphids revealed that this secondary co-obligate endosymbiosis was putatively triggered by the ancient loss of the riboflavin biosynthetic capability in the *Buchnera* from the Lachninae last common ancestor, rendering the secondary endosymbiont (capable of synthesizing riboflavin) of co-obligate nature for the aphid–*Buchnera* consortium within this subfamily.[50,62] In the case of the *Carsonella*-harboring psyllids, in-depth sequencing of five species belonging to three different genera (*Ctenarytania*, *Heteropsylla*, and *Pachypsylla*) revealed that at least two of them harbored distinctive genomically reduced secondary endosymbionts[61] capable of metabolically complementing the lost enzymatic steps in different metabolic pathways. However, the sequencing data for the rest of the psyllid species did not yield any obvious signs of the presence of a secondary endosymbiont. The authors speculate that this apparent absence of endosymbionts could be due to putative differences in the amino acid profiles of the phloem sap the insects feed on or an increased role of the host genome (horizontal gene transfer from either *C. ruddii* or another bacterial source). However, the presence of another unidentified symbiont cannot be completely ruled out.

A more extreme case of metabolic complementation is found in the endosymbionts of mealybugs from the Pseudococcinae subfamily. They harbor the betaproteobacterial *Tremblaya princeps* inside bacteriocytes, which, unlike their Phenacoccinae relative (*Phenacoccus avenae*), in turn harbor their own intracellular endosymbionts.[139,143] Recent sequencing of some of these bacterial consortia, along with their hosts, demonstrated a high level of metabolic complementarity between the endosymbionts and even a dependency of *T. princeps* on the import of translational machinery from its bacterial resident.[27,144–146] A common feature of these co-obligatory endosymbioses is that, once the secondary endosymbiont is fixed as obligate, there is further gene loss and an evolutionary acceleration of its co-obligate bacterial partner (Fig. 3). This can be explained by the new selective pressures imposed by a new player on the already-present symbiont, namely gene redundancy.

Within the hemipteran suborder Auchenorrhyncha, there is a very ancient dual endosymbiotic system (presumably established over 270 Mya) in most extant families: *Sulcia muelleri* plus a Betaproteobacteria, named the *BetaSymb* lineage (see Ref. 47). Although various genus names have been proposed to designate the different betaproteobacterial secondary associates (*Nasuia*, *Vidania*, and *Zinderia*) on the basis of phylogenetic analyses, it has been suggested that all three putatively originated from a single "beta symbiont" infection.[147] Additionally, *Sodalis*-like tertiary associates have also been found within the Cercopoidea superfamily coexisting with both *Sulcia* and *Zinderia*. It is also important to note that the smallest endosymbiont genomes have been found within those housed by insects within these suborders, such as the 144-kb genome of *Hodgkinia cicadicola*[46] and the tiny 112-kb genome of *Nasuia deltocephalinicola*,[47] co-obligate endosymbionts of the cicada *Diceroprocta semicincta* and the leafhopper *Macrosteles*

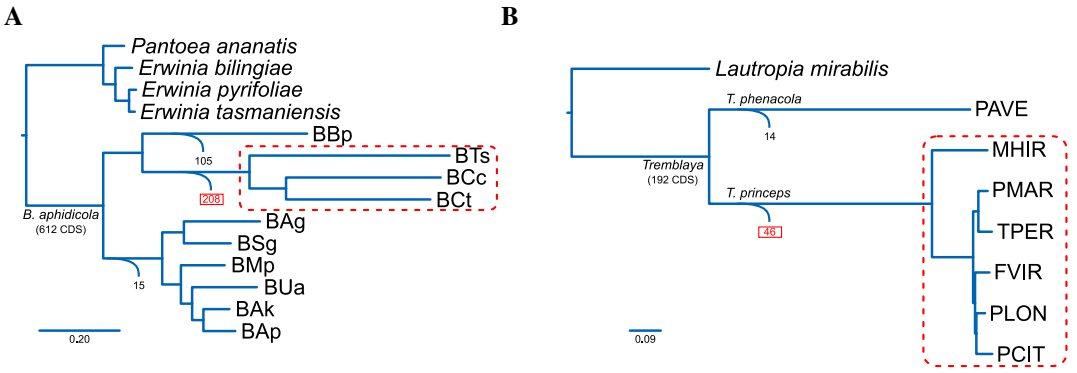

Figure 3. Gene loss associated with the complementation of an obligate endosymbiont. Maximum likelihood phylogenetic reconstruction of *Buchnera* (A) and *Tremblaya* (B) genomes on the basis of single-copy conserved genes, as calculated by OrthoMCL.[169] The genus or species name for each endosymbiotic lineage is displayed under the branch leading to it. In parentheses is shown the number of CDSs calculated to be present in the ancestor of each endosymbiotic lineage, according to Manzano-Marin et al.[50] (A) and Husnik and McCutcheon[146] (B). Arrows emanating from branches in the tree represent CDS losses in each lineage, with those harboring obligate secondary endosymbionts marked with a red box and lettering. Endosymbiotic clades that have an obligate secondary associate are grouped with a red-dotted box. BMp, Bua, and BAk refer to *Buchnera* from *Mizus persicae*, *Uroleucon ambrosiae*, and *Acyrthosiphon kondoi*, respectively. Strain names were used for *Tremblaya*.

quadrilineatus, respectively. These tiny genomes demonstrate the drastic consequences of co-obligate associations, driving genome reduction to the extreme.

Replacement

We can consider Buchner as the first author to propose the replacement of an endosymbiont.[18] His studies of the aforementioned Auchenorrhyncha (based on his student H. J. Müller's work) led him to propose that one symbiotic association was ancestral to the suborder, with different host lineages acquiring and losing additional symbionts during the radiation of this clade.[148,149] This resulted in a mosaic of different symbiont combinations across modern subfamilies and tribes. Recent molecular studies have refined the hypothesis:[47,150–152] while the almost universal presence of *Sulcia* and the widespread presence of *BetaSymb* indicate that these two were coresident in the common ancestor of Auchenorrhyncha, there is evidence for multiple replacements by *Hodgkinia*, *Sodalis*-like, and *Baumannia* bacteria in the Cicadoidea and within the Cercopoidea and the Cicadellinae, respectively.[47] These symbionts have evolved a metabolically similar repertoire[3,150] and, in the case of *Baumannia*[134,151] and *Hodgkinia*,[46,152] a drastically reduced genome.

Symbiont replacement of the primary and only endosymbiont from an insect has been observed in both Curculionidae (weevils)[153,154] and Aphididae (aphids).[156] In the case of weevils, phylogenetic and microscopic studies have shown that there was an original infection by a *Nardonella* symbiont in their common ancestor, followed by symbiont replacement by, for example, a *Sodalis* bacterium in the genus *Sitophilus*[153,154] and *Curculioniphilus buchneri* in the genus *Curculio*.[155] In *Sitophilus* weevils, the sequencing of its primary obligate endosymbiont (*S. pierantonius*) has revealed characteristics congruent with the proposed recent acquisition of this symbiont, including a large genome, an abundance of insertion sequences, and a high level of pseudogenization.[104] Regarding symbiont replacement in aphids, the substitution of *Buchnera* by a yeast-like symbiont was first described in the Cerataphidini aphid *Tuberaphis styraci*[156] and later molecularly characterized in other monophyletic aphids within this tribe.[157] Surprisingly, the loss of *Buchnera* has been accompanied by a loss of the bacteriome, and thus the symbiont resides in the hemocoel and fat body of its host. The sequencing of the yeast-like symbiont from the aphid *Cerataphis brasiliensis* revealed, as expected, that the symbiont was able to synthesize the essential nutrients that are generally supplied by *Buchnera* in other aphids,[158] corroborating the functional replacement of the original primary endosymbiont.

In summary, endosymbionts in this very advanced stage of genome reduction hold drastically reduced, gene-dense genomes, are generally housed within specialized cells (bacteriocytes), and have developed an intricate metabolic complementation with their host. Their genomes are fragile, and further gene losses can have a strong impact on their symbiotic role, leading to symbiont replacement or complementation. Host shifts to richer diets can also trigger further genome reduction, as proposed for *Carsonella* in psyllids[61] or *Westeberhardia* in ants.[159] Also, horizontal gene transfers from unrelated bacteria are readily detectable in the hosts, facilitating further genome reduction[27,146,160] and even enabling host–symbiont interactions.[161] If the symbiont successfully survives to this stage, functional transfer of informational genes from its genome to the host nucleus can lead to the bacterial associate turning into an organelle, similar to what triggered the establishment of an Alphaproteobacterium and a Cyanobacterium as the mitochondrion and chloroplast organelles, respectively.

Perspectives

With the continuous sequencing and analyses of endosymbiotic systems from diverse taxa, we are discovering exciting and unique relationships. In particular, the discovery of symbionts belonging to a unique genus that are present in contrastingly different stages of genome reduction is shedding light on genome reduction and its link to phenotypic traits. Such examples include *Arsenophonus*,[115,162] *Coxiella*,[163–165] *Serratia*,[50,60,62,82,92,166] and *Sodalis*[67,81,104] symbionts. Particularly, *Coxiella* and *Serratia* symbionts from ticks and aphids, respectively, offer the unique opportunity to dissect the genome-reduction process within a single symbiotic lineage and closely related hosts (ticks or aphids).

However, several key questions remain unanswered. These are related to the origin and evolution of the symbiosis and to the ultimate fate of the endosymbionts. Regarding origin, as yet we do not know which specific changes occur in a free-living bacterium to trigger its transition to an endosymbiont. Another unanswered question is why some microorganisms become intracellular whereas others remain extracellular (ectosymbionts or exosymbionts), as well as the role played by the hosts in this outcome. In general, eukaryotes, including some insects, harbor a rich intestinal microbiota participating in many host functions,[167] indicating that the extracellular route is more abundant on the phylogenetic scale than the intracellular one. However, they are not necessarily alternative routes, and, for example, in cockroaches, both intracellular and extracellular systems coexist.[124,167,168] Regarding the evolution of the symbiotic association, the determining factors behind the genotypic and phenotypic changes undergone by the endosymbionts in the different stages of genome reduction have yet to be determined. The genome-reduction stages proposed in this review are by no means definite, and we expect that further research into symbiotic systems in diverse animals will expand and even blur the lines separating each stage.

Regarding the ultimate fate of the endosymbionts, we have presented extreme cases of genome reduction and shown that they are dependent both on their hosts and on the presence of additional endosymbionts. This raises some important questions: What do these organelle-like entities represent? What roles do they play in the evolution of eukaryotes? What changes does the host's immune system undergo to allow bacterial accommodation? In this review, we have extensively described the profound transformations experienced by a bacterium on becoming an intracellular endosymbiont. However, more in-depth research should now focus on the host, exploring the host-derived factors that could promote the endosymbiont's genomic reduction and lifestyle changes. Finally, the integration of all these branches of knowledge will have a great impact on synthetic biology, enabling us to learn how these microorganisms can be manipulated and transformed, mainly those that are still culturable. In doing so, we would be able to reprogram their genomes for specific studies.

In the near future, research should tackle all the above-mentioned and ambitious research properly. It will be necessary to (1) undertake more systematic study into the genomics and other omics of both eukaryotes and their intra- and extracellular symbionts, (2) study the molecular biology of the host's tolerance mechanisms for endosymbionts, and (3) set up experimental studies to ascertain how a stage 1 or potential mutualist symbiont is altered by a forced continuous association with a particular host.

Acknowledgments

We thank Andrés Moya for valuable comments on the manuscript. This work was supported by Grant BFU2015-64322-C2-2-R (cofinanced by FEDER funds and the Ministerio de Economía y Competitividad (Spain) to A.L.), the European Commission (Marie Curie FP7 PITN-GA-2010-264774-SYMBIOMICS to A.M.M.), and the Consejo Nacional de Ciencia y Tecnología (Mexico) doctoral scholarship (CONACYT 327211/381508 to A.M.M.)

Conflict of interest

The authors declare no conflicts of interest.

References

1. Moya, A., J. Peretó, R. Gil & A. Latorre. 2008. Learning how to live together: genomic insights into prokaryote–animal symbioses. *Nat. Rev. Genet.* **9:** 218–229.
2. Moran, N.A., J.P. McCutcheon & A. Nakabachi. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* **42:** 165–190.
3. McCutcheon, J.P. & N.A. Moran. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 million years of evolution. *Genome Biol. Evol.* **2:** 708–718.
4. Moran, N.A. & G.M. Bennett. 2014. The tiniest tiny genomes. *Annu. Rev. Microbiol.* **68:** 195–215.
5. Toft, C. & S.G.E. Andersson. 2010. Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* **11:** 465–475.
6. Wernegreen, J.J. 2015. Endosymbiont evolution: predictions from theory and surprises from genomes. *Ann. N.Y. Acad. Sci.* **1360:** 16–35.
7. Koonin, E.V. 2010. The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol.* **11:** 209.
8. Embley, T.M. & W. Martin. 2006. Eukaryotic evolution, changes and challenges. *Nature* **440:** 623–630.
9. Spang, A. *et al.* 2015. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521:** 173–179.
10. Pittis, A.A. & T. Gabaldón. 2016. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531:** 101–104.
11. Margulis, L. 1993. *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons.* New York: W.H. Freeman & Co. Ltd.
12. Timmis, J.N., M.A. Ayliffe, C.Y. Huang & W. Martin. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5:** 123–135.
13. Gabaldón, T. & M.A. Huynen. 2007. From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PLoS Comput. Biol.* **3:** e219.
14. Kurland, C.G. & S.G.E. Andersson. 2000. Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* **64:** 786–820.
15. Martin, W. *et al.* 2002. Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **99:** 12246–12251.
16. Keeling, P.J. & J.D. Palmer. 2008. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9:** 605–618.
17. Keeling, P.J., J.P. McCutcheon & W.F. Doolittle. 2015. Symbiosis becoming permanent: survival of the luckiest. *Proc. Natl. Acad. Sci. U.S.A.* **112:** 10101–10103.
18. Buchner, P. 1953. *Endosymbiose der Tiere mit Pflanzlichen Mikroorganismen.* Basel: Birkhäuser.
19. Buchner, P. 1965. *Endosymbiosis of Animals with Plant Microorganisms.* New York: Interscience Publishers.
20. Baumann, P. 2005. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* **59:** 155–189.
21. McCutcheon, J.P. & N.A. Moran. 2012. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10:** 13–26.
22. Ziegler, H. 1975. Transport in plants I. In *Encyclopedia of Plant Physiology.* Vol. 1. M.H. Zimmermann & J.A. Milburn, Eds.: 59–100. Berlin, Heidelberg: Springer.
23. Sandstrom, J. & N. Moran 1999. How nutritionally imbalanced is phloem sap for aphids? *Entomol. Exp. Appl.* **91:** 203–210.
24. Lehane, M.J. 2005. *The Biology of Blood-Sucking in Insects.* Cambridge: Cambridge University Press.
25. Shigenobu, S. *et al.* 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407:** 81–86.
26. Sloan, D.B. & N.A. Moran. 2012. Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biol. Lett.* **8:** 986–989.
27. Husnik, F. *et al.* 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153:** 1567–1578.
28. Akman, L. *et al.* 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia. Nat. Genet.* **32:** 402–407.
29. Kirkness, E.F. *et al.* 2010. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* **107:** 12168–12173.
30. Rio, R.V.M. *et al.* 2012. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: Glossinidae) obligate symbiont *Wigglesworthia. MBio* **3:** e00240-11.
31. Boyd, B.M., J.M. Allen, V. de Crecy-Lagard & D.L. Reed. 2014. Genome sequence of *Candidatus* Riesia pediculischaeffi, endosymbiont of chimpanzee lice, and genomic comparison of recently acquired endosymbionts from human and chimpanzee lice. *G3 (Bethesda)* **4:** 2189–2195.
32. Koga, R., T. Tsuchida & T. Fukatsu. 2003. Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid. *Proc. Biol. Sci.* **270:** 2543–2550.

33. Ohtaka, C. & H. Ishikawa. 1991. Effects of heat treatment on the symbiotic system of an aphid mycetocyte. *Symbiosis* **11:** 19–30.
34. Nogge, G. 1976. Sterility in tsetse flies (*Glossina morsitans* Westwood) caused by loss of symbionts. *Experientia* **32:** 995–996.
35. Sacchi, L., A. Grigolo, G. Biscaldi & U. Laudani. 1993. Effects of heat treatment on the symbiotic system of Blattoidea: morphofunctional alterations of bacteriocytes. *Boll. di Zool.* **60:** 271–279.
36. Engelstädter, J. & G.D.D. Hurst. 2009. The ecology and evolution of microbes that manipulate host reproduction. *Annu. Rev. Ecol. Evol. Syst.* **40:** 127–149.
37. Oliver, K.M., P.H. Degnan, G.R. Burke & N.A. Moran. 2010. Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annu. Rev. Entomol.* **55:** 247–266.
38. Oliver, K.M., A.H. Smith & J.A. Russell. 2014. Defensive symbiosis in the real world—advancing ecological studies of heritable, protective bacteria in aphids and beyond. *Funct. Ecol.* **28:** 341–355.
39. Tamas, I. *et al.* 2002. 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **296:** 2376–2379.
40. van Ham, R.C. *et al.* 2003. Reductive genome evolution in *Buchnera aphidicola*. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 581–586.
41. Gil, R. *et al.* 2003. The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 9388–9393.
42. Pérez-Brocal, V. *et al.* 2006. A small microbial genome: the end of a long symbiotic relationship? *Science* **314:** 312–313.
43. Nakabachi, A. *et al.* 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* **314:** 267–267.
44. Nakabachi, A. *et al.* 2013. Defensive bacteriome symbiont with a drastically reduced genome. *Curr. Biol.* **23:** 1478–1484.
45. Tamames, J. *et al.* 2007. The frontier between cell and organelle: genome analysis of *Candidatus* Carsonella ruddii. *BMC Evol. Biol.* **7:** 181.
46. McCutcheon, J.P., B.R. McDonald & N.A. Moran. 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* **5:** e1000565.
47. Bennett, G.M. & N.A. Moran. 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biol. Evol.* **5:** 1675–1688.
48. McCutcheon, J.P. 2010. The bacterial essence of tiny symbiont genomes. *Curr. Opin. Microbiol.* **13:** 73–78.
49. Reyes-Prieto, M., A. Latorre & A. Moya. 2014. Scanty microbes, the 'symbionelle' concept. *Environ. Microbiol.* **16:** 335–338.
50. Manzano-Marín, A., J.C. Simon & A. Latorre. 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**: 1440–1458.
51. Aras, R.A. *et al.* 2003. Extensive repetitive DNA facilitates prokaryotic genome plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 13579–13584.
52. Shen, P. & H.V Huang. 1986. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* **112:** 441–457.
53. Moran, N.A. 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 2873–2878.
54. Moran, N.A. & J.J. Wernegreen. 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol. Evol.* **15:** 321–326.
55. Silva, F.J., A. Latorre & A. Moya. 2003. Why are the genomes of endosymbiotic bacteria so stable? *Trends Genet.* **19:** 176–180.
56. Fares, M.A., A. Moya & E. Barrio. 2004. GroEL and the maintenance of bacterial endosymbiosis. *Trends Genet.* **20:** 413–416.
57. Fares, M.A. *et al.* 2002. Endosymbiotic bacteria: *GroEL* buffers against deleterious mutations. *Nature* **417:** 398.
58. Sabater-Muñoz, B. *et al.* 2015. Fitness trade-offs determine the role of the molecular chaperonin GroEL in buffering mutations. *Mol. Biol. Evol.* **32:** 2681–2693.
59. Moya, A., R. Gil & A. Latorre. 2009. The evolutionary history of symbiotic associations among bacteria and their animal hosts: a model. *Clin. Microbiol. Infect.* **15:** 11–13.
60. Lamelas, A. *et al.* 2011. *Serratia symbiotica* from the aphid *Cinara cedri*: a missing link from facultative to obligate insect endosymbiont. *PLoS Genet.* **7:** e1002357.
61. Sloan, D.B. & N.A. Moran. 2012. Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. *Mol. Biol. Evol.* **29:** 3781–3792.
62. Manzano-Marín, A. & A. Latorre. 2014. Settling down: the genome of *Serratia symbiotica* from the aphid *Cinara tujafilina* zooms in on the process of accommodation to a cooperative intracellular life. *Genome Biol. Evol.* **6:** 1683–1698.
63. Kikuchi, Y. 2009. Endosymbiotic bacteria in insects: their diversity and culturability. *Microbes Environ.* **24:** 195–204.
64. Husník, F., T. Chrudimský & V. Hypša. 2011. Multiple origins of endosymbiosis within the Enterobacteriaceae (γ-Proteobacteria): convergence of complex phylogenetic approaches. *BMC Biol.* **9:** 87.
65. Gil, R. *et al.* 2002. Extreme genome reduction in *Buchnera* spp.: toward the minimal genome needed for symbiotic life. *Proc. Natl. Acad. Sci. U.S.A.* **99:** 4454–4458.
66. Lamelas, A. *et al.* 2008. Evolution of the secondary symbiont "*Candidatus* Serratia symbiotica" in aphid species of the subfamily Lachninae. *Appl. Environ. Microbiol.* **74:** 4236–4240.
67. Clayton, A.L. *et al.* 2012. A novel human-infection-derived bacterium provides insights into the evolutionary origins of mutualistic insect–bacterial symbioses. *PLoS Genet.* **8:** e1002990.
68. Hosokawa, T. *et al.* 2016. Obligate bacterial mutualists evolving from environmental bacteria in natural insect populations. *Nat. Microbiol.* **1:** 15011.

69. Welch, R.A. 2006. Proteobacteria: gamma subclass. In *The Prokaryotes*. Vol. 6. S. Falkow, E. Rosenberg, K.H. Schleifer & E. Stackebrandt, Eds.: 60–71. New York: Springer.
70. Bachmann, B.J. 1996. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. F.C. Neidhardt & R. Curtiss, Eds.: 2460–2488. Richmond, CA: ASM Press.
71. Walterson, A.M. & J. Stavrinides. 2015. Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiol. Rev.* **39:** 968–984.
72. Choi, O. *et al.* 2012. Complete genome sequence of the rice pathogen *Pantoea ananatis* strain PA13. *J. Bacteriol.* **194:** 531.
73. Grimont, F. & P.A.D. Grimont. Proteobacteria: gamma subclass. In *The Prokaryotes*. Vol. 6. S. Falkow, E. Rosenberg, K.H. Schleifer & E. Stackebrandt, Eds.: 219–244. New York: Springer.
74. Flyg, C., K. Kenne & H.G. Boman. 1980. Insect pathogenic properties of *Serratia marcescens*: phage-resistant mutants with a decreased resistance to cecropia immunity and a decreased virulence to *Drosophila*. *Microbiology* **120:** 173–181.
75. Chari, A. *et al.* 2015. Phenotypic characterization of *Sodalis praecaptivus* sp. nov., a close non-insect-associated member of the *Sodalis*-allied lineage of insect endosymbionts. *Int. J. Syst. Evol. Microbiol.* **65:** 1400–1405.
76. Dale, C. & I. Maudlin. 1999. *Sodalis* gen. nov. and *Sodalis glossinidius* sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. *Int. J. Syst. Bacteriol.* **49:** 267–275.
77. Pinnock, D.E. & R.T. Hess. 1974. The occurrence of intracellular *Rickettsia*-like organisms in the tsetse flies, *Glossina morsitans*, *G. fuscipes*, *G. brevipalpis* and *G. pallidipes*. *Acta Trop.* **31:** 70–79.
78. Sabri, A. *et al.* 2011. Isolation, pure culture and characterization of *Serratia symbiotica* sp. nov., the R-type of secondary endosymbiont of the black bean aphid *Aphis fabae*. *Int. J. Syst. Evol. Microbiol.* **61:** 2081–2088.
79. Facey, P.D. *et al.* 2015. Draft genomes, phylogenetic reconstruction, and comparative genomics of two novel cohabiting bacterial symbionts isolated from *Frankliniella occidentalis*. *Genome Biol. Evol.* **7:** 2188–2202.
80. Dale, C., *et al.* 2006. Isolation, pure culture, and characterization of "*Candidatus* Arsenophonus arthropodicus," an intracellular secondary endosymbiont from the hippoboscid louse fly *Pseudolynchia canariensis*. *Appl. Environ. Microbiol.* **72:** 2997–3004.
81. Toh, H. *et al.* 2005. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res.* **16:** 149–156.
82. Foray, V. *et al.* 2014. Whole-genome sequence of *Serratia symbiotica* strain CWBI-2.3T, a free-living symbiont of the black bean aphid *Aphis fabae*. *Genome Announc.* **2**. pii: e00767-14.
83. Iguchi, A. *et al.* 2014. Genome evolution and plasticity of *Serratia marcescens*, an important multidrug-resistant nosocomial pathogen. *Genome Biol. Evol.* **6:** 2096–2110.
84. Belda, E., A. Moya, S. Bentley & F.J. Silva. 2010. Mobile genetic element proliferation and gene inactivation impact over the genome structure and metabolic capabilities of *Sodalis glossinidius*, the secondary endosymbiont of tsetse flies. *BMC Genomics* **11:** 449.
85. Dale, C. & A. Welburn. 2001. The endosymbionts of tsetse flies: manipulating host–parasite interactions. *Int. J. Parasitol.* **31:** 628–631.
86. Cheng, Q. & A. Aksoy. 1999. Tissue tropism, transmission and expression of foreign genes *in vivo* in midgut symbionts of tsetse flies. *Insect Mol. Biol.* **8:** 125–132.
87. Wu, M. *et al.* 2004. Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2:** e69.
88. Werren, J.H., L. Baldo & M.E. Clark. 2008. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6:** 741–751.
89. Degnan, P.H. *et al.* 2009. *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proc. Natl. Acad. Sci. U.S.A.* **106:** 9063–9068.
90. Degnan, P.H. *et al.* 2009. Dynamics of genome evolution in facultative symbionts of aphids. *Environ. Microbiol.* **12:** 2060–2069.
91. Hansen, A.K., C. Vorburger & N.A. Moran. 2012. Genomic basis of endosymbiont-conferred protection against an insect parasitoid. *Genome Res.* **22:** 106–114.
92. Burke, G.R. & N.A. Moran. 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol. Evol.* **3:** 195–208.
93. Oliver, K.M., J.A. Russell, N.A. Moran & M.S. Hunter. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 1803–1807.
94. Oliver, K.M., N.A. Moran & M.S. Hunter. 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 12795–12800.
95. Scarborough, C.L. 2005. Aphid protected from pathogen by endosymbiont. *Science* **310:** 1781.
96. Łukasik, P.M. *et al.* 2013. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecol. Lett.* **16:** 214–218.
97. Chen, D.Q., C.B. Montllor & A.H. Purcell. 2000. Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, *Acyrthosiphon pisum*, and the blue alfalfa aphid, *A. kondoi*. *Entomol. Exp. Appl.* **95:** 315–323.
98. Montllor, C.B., A. Maxmen & A.H. Purcell. 2002. Facultative bacterial endosymbionts benefit pea aphids *Acyrthosiphon pisum* under heat stress. *Ecol. Entomol.* **27:** 189–195.
99. Ferrari, J., C.L. Scarborough & H.C.J. Godfray. 2007. Genetic variation in the effect of a facultative symbiont on host-plant use by pea aphids. *Oecologia* **153:** 323–329.
100. Leonardo, T.E. & G.T. Muiru. 2003. Facultative symbionts are associated with host plant specialization in pea aphid populations. *Proc. Biol. Sci.* **270**(Suppl. 2): S209–S212.

101. Tsuchida, T., R. Koga & T. Fukatsu. 2004. Host plant specialization governed by facultative symbiont. *Science* **303:** 1989.
102. Henry, L.M. *et al.* 2013. Horizontally transmitted symbionts and host colonization of ecological niches. *Curr. Biol.* **23:** 1713–1717.
103. Moran, N.A., J.A. Russell, R. Koga & T. Fukatsu. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Appl. Environ. Microbiol.* **71:** 3302–3310.
104. Oakeson, K.F. *et al.* 2014. Genome degeneration and adaptation in a nascent stage of symbiosis. *Genome Biol. Evol.* **6:** 76–93.
105. Wicker, C. 1983. Differential vitamin and choline requirements of symbiotic and aposymbiotic *S. oryzae* (coleoptera: Curculionidae). *Comp. Biochem. Physiol. Part A Physiol.* **76:** 177–182.
106. Gasnier-Fauchet, F. & P. Nardon. 1987. Comparison of sarcosine and methionine sulfoxide levels in symbiotic and aposymbiotic larvae of two sibling species, *Sitophilus oryzae* L. and *S. zeamais* mots (coleoptera: Curculionidae). *Insect Biochem.* **17:** 17–20.
107. Grenier, A.M., C. Nardon & P. Nardon. 1994. The role of symbiotes in flight activity of *Sitophilus* weevils. *Entomol. Exp. Appl.* **70:** 201–208.
108. O'Neill, S.L. *et al.* 1997. *In vitro* cultivation of *Wolbachia pipientis* in an *Aedes albopictus* cell line. *Insect Mol. Biol.* **6:** 33–39.
109. Darby, A.C., S.M. Chandler, S.C. Welburn & A.E. Douglas. 2005. Aphid-symbiotic bacteria cultured in insect cell lines. *Appl. Environ. Microbiol.* **71:** 4833–4839.
110. Nikoh, N. *et al.* 2011. Reductive evolution of bacterial genome in insect gut environment. *Genome Biol. Evol.* **3:** 702–714.
111. Hosokawa, T., Y. Kikuchi, X.Y. Meng & T. Fukatsu. 2005. The making of symbiont capsule in the plataspid stinkbug *Megacopta punctatissima*. *FEMS Microbiol. Ecol.* **54:** 471–477.
112. Hosokawa, T. *et al.* 2006. Strict host–symbiont cospeciation and reductive genome evolution in insect gut bacteria. *PLoS Biol.* **4:** e337.
113. Santos-Garcia, D. *et al.* 2012. Complete genome sequence of "*Candidatus* Portiera aleyrodidarum" BT-QVLC, an obligate symbiont that supplies amino acids and carotenoids to *Bemisia tabaci*. *J. Bacteriol.* **194:** 6654–6655.
114. Santos-Garcia, D. *et al.* 2014. Small but powerful, the primary endosymbiont of moss bugs, "*Candidatus* Evansia muelleri", holds a reduced genome with large biosynthetic capabilities. *Genome Biol. Evol.* **6:** 1875–1893.
115. Nováková, E., F. Husník, E. Šochová & V. Hypša. 2015. *Arsenophonus* and *Sodalis* symbionts in louse flies: an analogy to the *Wigglesworthia* and *Sodalis* system in tsetse flies. *Appl. Environ. Microbiol.* **81:** 6189–6199.
116. Mercier, L.M. 1906. Les corps bacteroides de la blatte (*Periplanata orientalis*: *Bacillus cuenoti* (n. sp. L. Mercier). *C. R. Seances Soc. Biol. Fil.* **58:** 682–684.
117. Hollande, A.C. & R. Favre. 1931. La structure cytologique de *Blattabacterium cuenoti* (Mercier) N.G., symbiote du tissu adipeux des Blattides. *C. R. Seances Soc. Biol. Fil.* **107:** 752–754.
118. Schröder, D. *et al.* 1996. Intracellular endosymbiotic bacteria of *Camponotus* species (carpenter ants): systematics, evolution and ultrastructural characterization. *Mol. Microbiol.* **21:** 479–489.
119. Sauer, C., D. Dudaczek, B. Holldobler & R. Gross. 2002. Tissue localization of the endosymbiotic bacterium "*Candidatus* Blochmannia floridanus" in adults and larvae of the carpenter ant *Camponotus floridanus*. *Appl. Environ. Microbiol.* **68:** 4187–4193.
120. Sabree, Z.L., S. Kambhampati, & N.A. Moran. 2009. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc. Natl. Acad. Sci. U.S.A.* **106:** 19521–19526.
121. López-Sánchez, M.J. *et al.* 2009. Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*. *PLoS Genet.* **5:** e1000721.
122. Feldhaar, H. *et al.* 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. *BMC Biol.* **5:** 48.
123. González-Domenech, C.M. *et al.* 2012. Metabolic stasis in an ancient symbiosis: genome-scale metabolic networks from two *Blattabacterium cuenoti* strains, primary endosymbionts of cockroaches. *BMC Microbiol.* **12** (Suppl. 1): S5.
124. Patiño-Navarrete, R., A. Moya, A. Latorre & J. Peretó. 2013. Comparative genomics of *Blattabacterium cuenoti*: the frozen legacy of an ancient endosymbiont genome. *Genome Biol. Evol.* **5:** 351–361.
125. International Aphid Genomics Consortium. 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* **8:** e1000313.
126. Hansen, A.K. & N.A. Moran. 2011. Aphid genome expression reveals host–symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **108:** 2849–2854.
127. Luan, J.B. *et al.* 2015. Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. *Genome Biol. Evol.* **7:** 2635–2647.
128. Williams, L.E. & J.J. Wernegreen. 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.
129. Williams, L.E. & J.J. Wernegreen. 2010. Unprecedented loss of ammonia assimilation capability in a urease-encoding bacterial mutualist. *BMC Genomics* **11:** 687.
130. Tokuda, G. *et al.* 2013. Maintenance of essential amino acid synthesis pathways in the *Blattabacterium cuenoti* symbiont of a wood-feeding cockroach. *Biol. Lett.* **9:** 20121153.
131. Neef, A. *et al.* 2011. Genome economization in the endosymbiont of the wood roach *Cryptocercus punctulatus* due to drastic loss of amino acid synthesis capabilities. *Genome Biol. Evol.* **3:** 1437–1448.
132. Bennett, G.M. & N.A. Moran. 2015. Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. *Proc. Natl. Acad. Sci. U.S.A.* **112:** 10169–10176.
133. López-Sánchez, M.J. *et al.* 2008. Blattabacteria, the endosymbionts of cockroaches, have small genome sizes and

high genome copy numbers. *Environ. Microbiol.* **10:** 3417–3422.
134. Bennett, G.M. *et al.* 2014. Differential genome evolution between companion symbionts in an insect–bacterial symbiosis. *MBio* **5:** e01697-14.
135. Sasaki-Fukatsu, K. *et al.* 2006. Symbiotic bacteria associated with stomach disc of human lice. *Appl. Environ. Microbiol.* **72:** 7349–7352.
136. Santos-García, D. *et al.* 2014. No exception to the rule: *Candidatus* Portiera alyrodidarum cell wall revisited. *FEMS Microbiol. Lett.* **360:** 132–136.
137. Reinhardt, C., R. Steiger & H. Hecker. 1972. Ultrastructural study of the midgut mycetome-bacteroids of the tsetse flies *Glossina morsitans*, *G. fuscipes*, and *G. brevipalpis* (Diptera, Brachycera). *Acta Trop.* **29:** 280–288.
138. Moran, N.A., P. Tran & N.M. Gerardo. 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum Bacteroidetes. *Appl. Environ. Microbiol.* **71:** 8802–8810.
139. von Dohlen, C.D., S. Kohler, S.T. Alsop & W.R. McManus. 2001. Mealybug beta-proteobacterial endosymbionts contain gamma-proteobacterial symbionts. *Nature* **412:** 433–436.
140. Kobiałka, M. *et al.* 2016. *Sulcia* symbiont of the leafhopper *Macrosteles laevis* (Ribaut, 1927) (Insecta, Hemiptera, Cicadellidae: Deltocephalinae) harbors *Arsenophonus* bacteria. *Protoplasma* **253:** 903–912.
141. Gómez-Valero, L. *et al.* 2004. Coexistence of *Wolbachia* with *Buchnera aphidicola* and a secondary symbiont in the aphid *Cinara cedri*. *J. Bacteriol.* **186:** 6626–6633.
142. Gosalbes, M.J., A. Lamelas, A. Moya & A. Latorre. 2008. The striking case of tryptophan provision in the cedar aphid *Cinara cedri*. *J. Bacteriol.* **190:** 6026–6029.
143. Thao, M.L., P.J. Gullan & P. Baumann. 2002. Secondary (gamma-Proteobacteria) endosymbionts infect the primary (beta-Proteobacteria) endosymbionts of mealybugs multiple times and coevolve with their hosts. *Appl. Environ. Microbiol.* **68:** 3190–3197.
144. McCutcheon, J.P. & C.D. von Dohlen. 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr. Biol.* **21:** 1366–1372.
145. Lopez-Madrigal, S. *et al.* 2011. Complete genome sequence of "*Candidatus* Tremblaya princeps" strain PCVAL, an intriguing translational machine below the living-cell status. *J. Bacteriol.* **193:** 5587–5588.
146. Husnik, F. & J.P. McCutcheon. 2016. Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *bioRxiv*. doi: 10.1101/042267.
147. Koga, R. *et al.* 2013. Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. *Environ. Microbiol.* **15:** 2073–2081.
148. Müller, H.J. 1940. *Die Symbiose der Fulgoroiden (Homoptera-Cicadina)*. Stuttgart: Schweizerbart Science Publishers.
149. Müller, H.J. 1962. Neuere vorstellungen über verbreitung und phylogenie der endosymbiosen der zikaden. *Z. Morphol. Oekol. Tiere* **51:** 190–210.
150. Koga, R. & N.A. Moran. 2014. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. *ISME J.* **8:** 1237–1246.
151. Wu, D. *et al.* 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol.* **4:** e188.
152. Campbell, M.A. *et al.* 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. *Proc. Natl. Acad. Sci. U.S.A.* **112:** 10192–10199.
153. Lefèvre, C. *et al.* 2004. Endosymbiont phylogenesis in the Dryophthoridae weevils: evidence for bacterial replacement. *Mol. Biol. Evol.* **21:** 965–973.
154. Conord, C. *et al.* 2008. Long-term evolutionary stability of bacterial endosymbiosis in Curculionoidea: additional evidence of symbiont replacement in the Dryophthoridae family. *Mol. Biol. Evol.* **25:** 859–868.
155. Toju, H. *et al.* 2010. "*Candidatus* Curculioniphilus buchneri", a novel clade of bacterial endocellular symbionts from weevils of the genus *Curculio*. *Appl. Environ. Microbiol.* **76:** 275–282.
156. Fukatsu, T. & H. Ishikawa. 1992. A novel eukaryotic extracellular symbiont in an aphid, *Astegopteryx styraci* (Homoptera, Aphididae, Hormaphidinae). *J. Insect Physiol.* **38:** 765–773.
157. Fukatsu, T., S. Aoki, U. Kurosu & H. Ishikawa. 1994. Phylogeny of Cerataphidini aphids revealed by their symbiotic microorganisms and basic structure of their galls—implications for host–symbiont coevolution and evolution of sterile soldier castes. *Zoo. Sci.* **11:** 613–623.
158. Vogel, K.J. & N.A. Moran. 2013. Functional and evolutionary analysis of the genome of an obligate fungal symbiont. *Genome Biol. Evol.* **5:** 891–904.
159. Klein, A. *et al.* 2016. A novel intracellular mutualistic bacterium in the invasive ant *Cardiocondyla obscurior*. *ISME J.* **10:** 376–388.
160. Sloan, D.B. *et al.* 2014. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol. Biol. Evol.* **31:** 857–871.
161. Nakabachi, A. *et al.* 2014. Aphid gene of bacterial origin encodes a protein transported to an obligate endosymbiont. *Curr. Biol.* **24:** R640–R641.
162. Xue, J. *et al.* 2014. Genomes of the rice pest brown planthopper and its endosymbionts reveal complex complementary contributions for host adaptation. *Genome Biol.* **15:** 521.
163. Seshadri, R. *et al.* 2003. Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 5455–5460.
164. Smith, T.A., T. Driscoll, J.J. Gillespie & R. Raghavan. 2015. A *Coxiella*-like endosymbiont is a potential vitamin source for the lone star tick. *Genome Biol. Evol.* **7:** 831–838.
165. Gottlieb, Y., I. Lalzar & L. Klasson. 2015. Distinctive genome reduction rates revealed by genomic analyses of two *Coxiella*-like endosymbionts in ticks. *Genome Biol. Evol.* **7:** 1779–1796.

166. Abebe-Akele, F. *et al.* 2015. Genome sequence and comparative analysis of a putative entomopathogenic *Serratia* isolated from *Caenorhabditis briggsae*. *BMC Genomics* **16:** 531.

167. Sabree, Z.L. & N.A. Moran. 2014. Host-specific assemblages typify gut microbial communities of related insect species. *Springerplus* **3:** 138.

168. Pérez-Cobas, A. *et al.* 2015. Diet shapes the gut microbiota of the omnivorous cockroach *Blatella germanica*. *FEMS Microbiol. Ecol.* **91**. pii: fiv022.

169. Li, L., C.J. Stoeckert & D.S. Roos. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13:** 2178–2189.

[Corrections added on November 23, 2016, after first online publication: Multiple reference citations were corrected throughout the article. On page 7, "[. . .] symbiont of the pigeon louse fly),[79]" was changed to "[. . .] symbiont of the pigeon louse fly),[80]"; on page 8, "[. . .] relative to DNA deletion.[62]" was changed to "[. . .] relative to DNA deletion.[92]"; on page 17, "[. . .] refined hypothesis:[47,147,150]" was changed to "[. . .] refined hypothesis:[47,150,152]"; "[. . .] and Aphididae (aphids).[155]" was changed to "[. . .] and Aphididae (aphids).[156]"; "[. . .] proposed for *Carsonella* in psyllids[60]" was changed to "[. . .] proposed for *Carsonella* in psyllids[61]"; on page 18, "[. . .]*Serratia*,[49,59,61,82,91,166]" was changed to "[. . .]*Serratia*,[50,60,62,82,92,166]"; and "*Sodalis*[71,80,103]" was changed to "*Sodalis*[67,81,104]".]