The Tiniest Tiny Genomes

Nancy A. Moran and Gordon M. Bennett

Department of Integrative Biology, University of Texas at Austin, Texas 78712; email: nancy.moran@austin.utexas.edu, gordon.bennett@utexas.edu

Annu. Rev. Microbiol. 2014. 68:195–215

First published online as a Review in Advance on June 2, 2014

The *Annual Review of Microbiology* is online at micro.annualreviews.org

This article's doi: 10.1146/annurev-micro-091213-112901

Keywords

bacteriocyte, *Buchnera*, endosymbiont, genome reduction, symbiosis

Abstract

Starting in 2006, surprisingly tiny genomes have been discovered from numerous bacterial symbionts of insect hosts. Despite their size, each retains some genes that enable provisioning of limiting nutrients or other capabilities required by hosts. Genome sequence analyses show that genome reduction is an ongoing process, resulting in a continuum of sizes, with the smallest genome currently known at 112 kilobases. Genome reduction is typical in host-restricted symbionts and pathogens, but the tiniest genomes are restricted to symbionts required by hosts and restricted to specialized host cells, resulting from long coevolution with hosts. Genes are lost in all functional categories, but core genes for central informational processes, including genes encoding ribosomal proteins, are mostly retained, whereas genes underlying production of cell envelope components are especially depleted. Thus, these entities retain cell-like properties but are heavily dependent on coadaptation of hosts, which continuously evolve to support the symbionts upon which they depend.

Contents

Minimal genome: the minimal set of genes required for growth and replication of a cell

Obligate symbionts: symbionts able to live only within hosts and required for host growth or reproduction

INTRODUCTION

The first bacterial genome sequences were published in the mid-1990s (18, 20). Soon after, the concept of the minimal genome emerged, defined as the gene set required to sustain life, imposing a lower boundary for the size of the genome (29, 39). Initially this concept appeared nicely compatible with observations on naturally occurring genomes. Until 2006, the smallest reported cellular genomes were those of the bacterium *Mycoplasma genitalium*, a host-restricted pathogen [580 kilobases (kb), 470 protein-coding genes] (20), and three strains of *Buchnera aphidicola*, obligate symbionts in insects (616–642 kb and 507–574 protein-coding genes) (**Figure 1**). A number of other host-restricted pathogens and symbionts were known to have genomes smaller than

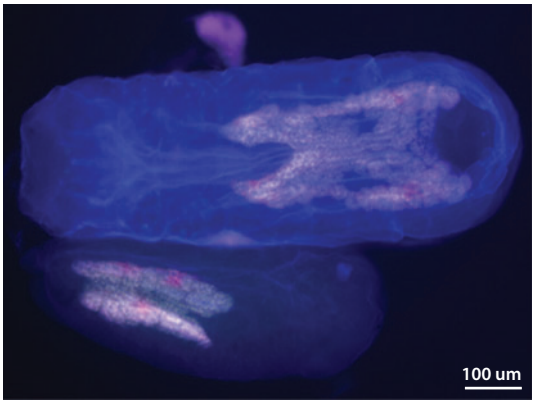


Figure 1

Bacteriocytes within the body of an embryo of the pea aphid, *Acyrthosiphon pisum*, as revealed by fluorescent in situ hybridization with rRNA of *Buchnera aphidicola*. The bacteriocytes form an organ in the abdomen of the aphid. Photo by G. Bennett, University of Texas.

1 megabase (Mb), with fewer than 1,000 protein-coding genes. Thus, for the first decade of full-genome sequencing, genomes appeared to conform to a minimal gene set of about 500 genes that consist of essential genes for cell growth and replication, plus genes linked to the specific ecology of each organism, whether mutualistic symbiont or pathogen. Furthermore, phylogenetic studies, largely based on ribosomal RNA, indicated that these small genome organisms, such as *M. genitalium*, evolved repeatedly from free-living bacteria with larger genomes (71, 81). These results provided clear evidence that genome reduction is somehow facilitated in organisms restricted to life in host cells or tissues and that it is usually a unidirectional change.

However, starting in 2006, a series of surprisingly smaller genomes have been reported, all for intracellular bacterial symbionts that are maternally transmitted between hosts. The extreme case, to date, is the genome of "*Candidatus* Nasuia deltocephalinicola," one of two obligate symbionts of the leafhopper *Macrosteles quadrilineatus*; this *Nasuia* strain possesses a mere 137 protein-coding genes and a genome of only 112 kb (2) (**Figure 2**). Several other recently discovered tiny genomes also have fewer than 200 genes, considerably less than the number once considered to be minimal for cellular life (39, 54). Extreme genome reduction is also a feature of obligate pathogens, which often possess highly reduced genomes (14, 20, 58), but pathogens have not been reported to reach the extremes found in genomes of mutualistic symbionts that are maternally transmitted and required for host development and reproduction.

The question of how these organisms can lose so many genes and still function is as yet unanswered. One key to this question is the observation that the tiniest genomes evolve exclusively in maternally inherited symbionts that are obligate mutualists and that have codiversified with hosts over millions of years. Although reduced genomes tend to retain genes underlying universal cell processes of replication, transcription, translation, and cell division (49), there are many enigmatic exceptions, especially in the smallest genomes (2, 54). The discovery of these extreme genomes challenges the premise of the minimal genome concept, that there is a lower limit of the genome size of a cellular organism.

In this review, we summarize cases of extreme genome reduction in obligate bacterial microbes as well as recent studies aimed at illuminating the processes enabling such extensive gene loss. We consider why symbionts and pathogens achieve the smallest-known genomes and what lifestyle features result in truly tiny genomes rather than simply small genomes. There has been a proliferation of newly named and sequenced tiny-genome symbionts since an earlier review (54), and we provide an updated compilation of these (2, 28, 83, 88). A major mystery regarding these organisms is how they can replicate and persist without key genes, and we compare the gene repertoires across available small genomes. The ability to live with so few genes clearly depends on coevolution by the hosts, which must provide support. We summarize results from several new studies that provide initial clues about the kinds of host contributions that enable tiny genomes to function and replicate.

Mutualists: partners of different species that provide reciprocal fitness benefits to one another

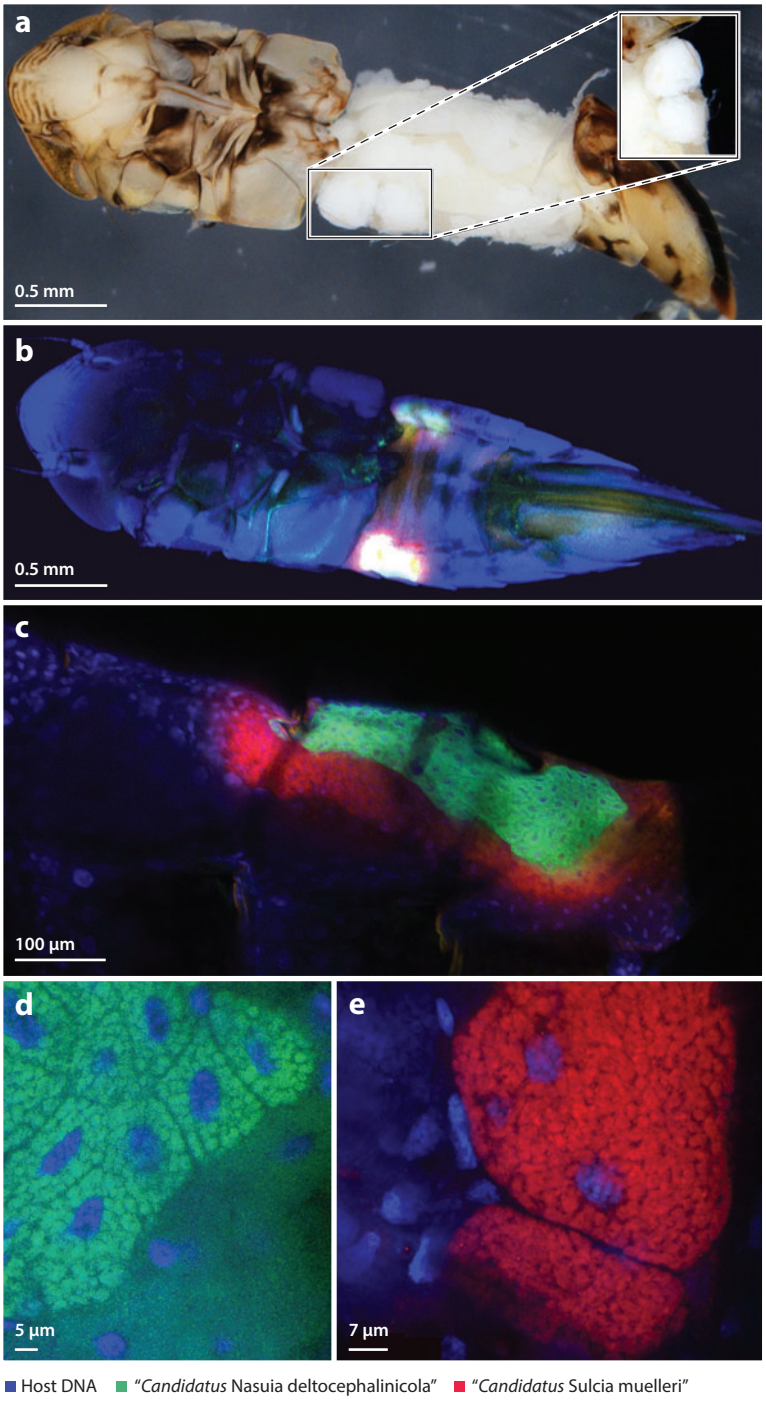
EAA: essential amino acid, one of 10 amino acids required by animals, due to their lack of the corresponding biosynthetic pathways

SMALL, TINY, AND TINIEST: WHO ARE THEY?

To date, all genomes under 500 kb have been found in obligate symbionts of various insects, a pattern that may in part reflect sampling intensity. Tiny genome symbionts are especially concentrated in insects specialized to feed on plant saps, diets that are deficient in the essential amino acids (EAAs) required by animals. As part of their shift to such plentiful, but unbalanced, food resources, insects have evolved alliances with various bacterial lineages, mostly in *Proteobacteria* and *Bacteroidetes*. Despite their tiny genomes, these bacteria retain genes underlying pathways for biosynthesis of essential nutrients lacking in host diets, including EAAs (54, 86), vitamins (1), carotenoids (87), and defensive compounds (63). Outside of sap-feeders (Insecta:

Figure 2

Fluorescent in situ hybridization of the dual bacteriocytes of a deltocephaline leafhopper, *Nesophrosyne montium*, containing "*Candidatus* Sulcia muelleri" (*red*) and "*Candidatus* Nasuia deltocephalinicola" (*green*). Strains of these symbionts found in *Macrosteles quadrilineatus* possess two of the tiniest genomes known in bacteria. Host insect DNA is counterstained with blue. (*a*) Dissected ventral view shows the bacteriome (inset of enlarged bacteriome). (*b*) Ventral view showing fluorescence of symbiotic bacteria. (*c*) Left lateral bacteriome, showing the segregation of symbionts in their respective bacteriocyte types. (*d*,*e*) Symbionts within individual bacteriocytes. Photos by G. Bennett, University of Texas.



Hemiptera), dramatic genome reduction is typical in maternally transmitted obligate symbionts of other invertebrate hosts, including cockroaches (Insecta: Blattodea), clams (Mollusca: Bivalvia: Vesicomyidae), blood-feeding flies (Insecta: Diptera: *Glossina*), tunicates (Urochordata), and others (25, 80, 82). Symbionts that have codiversified with other host groups and that have highly reduced genomes compared to related bacterial species include *Endolissoclinum faulkneri* in tunicates, which provides polyketides linked to host defense (41, 42); *Ruthia magnifica* in clams at deep ocean vents, which enables chemoautotrophic energy production (66); and *Photodesmus katoptron* in flashlight fish (Anomalopidae), which provides luminescence believed to deter predators (25). Tiny genomes may yet be found in symbionts of other invertebrates, or of plants and fungi, and may often involve defensive or chemoautotrophic functions, rather than nutrition.

Facultative symbionts: symbionts not required for normal host development but usually able to live only within hosts

SMALL AND TINY GENOMES IN INSECT SYMBIONTS: A TO Z

Arsenophonus species and "*Candidatus* Riesia pediculicola" (*Gammaproteobacteria*) together make up a clade of facultative and obligate symbionts of arthropods (70). The *Riesia* lineage consists of maternally inherited obligate symbionts of lice (Anoplura) (26, 70) with highly reduced genomes (574 kb), but they retain capability for biosynthesis of pantothenic acid, needed to supplement the hosts' blood diet (35). Part of this biosynthetic pathway is plasmid encoded, possibly as part of a mechanism to regulate production.

"*Candidatus* Baumannia cicadellinicola" (*Gammaproteobacteria*) is the obligate symbiont of the xylem-feeding sharpshooters (Cicadellinae) (103), with which *B. cicadellinicola* and "*Candidatus* Sulcia muelleri" have both codiversified (92). It replaced *N. deltocephalinicola* found in other leafhopper groups, possibly enabling the switch to xylem from phloem feeding. Although *B. cicadellinicola* has a larger genome (686 kb) encoding pathways for many additional vitamins and cofactors, it has largely evolved to fill the same essential metabolic role as *N. deltocephalinicola*, provisioning histidine and methionine to complement products of coresident *S. muelleri*.

Blattabacterium species (*Bacteroidetes*) are the obligate symbionts of cockroaches, in which they play a role in recycling waste nitrogen from urea and ammonia into amino acids, thereby extending the hosts' nitrogen-poor diets (47, 84). *Blattabacterium* species live in specialized fat-body cells, are strictly maternally transmitted, and have codiversified with cockroaches for >130 million years (44). Genomes range from 590 to 636 kb, with size variation reflecting loss of some amino acid biosynthetic pathways in the symbionts of highly social lineages, such as the wood roach (*Cryptocercus punctulatus*) and the primitive termite *Mastotermes darwiniensis* (27, 34, 65, 82, 84, 95). In these species, gut microbes, which are dependably transmitted through behavioral interactions, assume roles in nitrogen metabolism.

"*Candidatus* Blochmannia" species (*Gammaproteobacteria*) are maternally transmitted obligate symbionts of carpenter ants (*Camponotus*). Genomes range from 705 to 791 kb and encode pathways for EAA biosynthesis and nitrogen recycling.

Buchnera aphidicola (*Gammaproteobacteria*) of the pea aphid, *Acyrthosiphon pisum*, (**Figure 1**) is the best-studied model for an insect obligate symbiont association, and these studies are facilitated by the complete genome sequence of the host insect (30). Sequenced *Buchnera* genomes range from 416 to 642 kb. The smallest, in the aphid *Cinara cedri*, has lost some pathways for EAA biosynthesis, which are complemented by pathways of an additional symbiont from the genus *Serratia* (43). *Buchnera* genomes do not show the extremes of genome reduction but nonetheless have lost some genes considered essential, such as those underlying phospholipid biosynthesis.

"*Candidatus* Carsonella ruddii" (*Gammaproteobacteria*), the obligate symbiont of psyllids (Psylloidea), was the first example of an extremely tiny genome, with 160 kb and ~182 genes for *C. ruddii* PV of the host species *Pachypsylla venusta* (64). Subsequent genome sequences for five other *C. ruddii*

species revealed that gene loss is ongoing and often involves genes underlying biosynthesis of EAAs needed by hosts (88).

Bacteriocyte: specialized host cell adapted to harbor symbionts

"*Candidatus* Hodgkinia cicadicola" (*Alphaproteobacteria*), a symbiont of cicadas (Cicadidae), is represented by a single published genome of only 143 kb, one of the tiniest reported (50). *H. cicadicola* retains genes underlying biosynthesis of methionine and histidine, and its capabilities are complemented by the coresident symbiont, *S. muelleri*, which can produce the other eight EAAs. Remarkably, *H. cicadicola* retains genes encoding steps in the pathway for biosynthesis of cobalamin, which coincides with its use of cobalamin-dependent methionine synthase (MetH), rather than cobalamin-independent methionine synthase (MetE), as encoded in other insect symbiont genomes. *H. cicadicola* lacks genes for some tRNA synthetases and tRNAs corresponding to several amino acids and has a codon reassignment of TGA from Stop to Trp (50).

"*Candidatus* Ishikawaella capsulatus" (*Gammaproteobacteria*) is a maternally transmitted gut symbiont of some stink bugs (Plataspidae) and has a reduced genome encoding pathways for EAAs and vitamins, similar to bacteriocyte associates such as *Buchnera* species (67). *I. capsulatus* shows that genome reduction does not require intracellularity.

"*Candidatus* Moranella endobia" (*Gammaproteobacteria*), a rare case of a bacterium living inside another bacterium, dwells within "*Candidatus* Tremblaya princeps" of some mealybugs (Pseudococcidae) (28, 55). The genome, at 538 kb and about 450 genes, is highly reduced but nonetheless much larger than that of *T. princeps*. More than any other known case, these partners represent a fusion of two bacterial cells into a single new cellular entity, but they are also highly dependent on contributions from the host insect.

N. deltocephalinicola, the partner of *S. muelleri* in leafhoppers of the subfamily Deltocephalinae (**Figure 2**), boasts the tiniest genome yet reported, at 112 kb (2). As for other partners of *S. muelleri*, *N. deltocephalinicola* encodes genes for methionine and histidine production. *N. deltocephalinicola* and "*Candidatus* Zinderia insecticola" form a clade (referred to as *BetaSymb*) hypothesized to have descended from a symbiont ancestor living within an ancestor of spittlebugs (Cercopoidea: Philaenini) and leafhoppers (Cicadellidae) at least 250 mya (36). Both *N. deltocephalinicola* and *Z. insecticola* share an alternative genetic code similar to that of *H. cicadicola*.

"*Candidatus* Portiera aleyrodidarum" (*Gammaproteobacteria*), the obligate maternally transmitted symbiont of whiteflies (Aleyrodoidea), is the sister lineage to *C. ruddii* of psyllids (Psyllidae) (89), possibly because of colonization of a shared ancestor of these insects. Though *P. aleyrodidarum* has a tiny genome, 281 kb and 358 kb for the two species studied to date (89), it has genes for carotenoid biosynthesis and is the likely source for carotenoids known to occur in its host insects (87). *P. aleyrodidarum* of the pest species *Bemisia tabaci* has undergone genome rearrangement and expansion, the only such case known from long-term obligate symbionts.

"*Candidatus* Profftella armatura" (*Betaproteobacteria*), a symbiont of the Asian citrus psyllid *Diaphorina citri* is coresident with the obligate nutritional symbiont *C. ruddii*. *P. armatura* has a genome of only 465 kb, which resembles the size of genomes of nutritional symbionts such as *B. aphidicola*. However, it retains genes only for riboflavin and biotin production, whereas amino acid pathways are encoded by coresident *C. ruddii*. Instead, the major function appears to be the production of a defensive toxin, a polyketide encoded by a 70-kb locus that makes up 15% of the genome (63).

S. muelleri is an obligate symbiont in a large clade of sap-feeding insects (Hemiptera: Auchenorrhyncha). *S. muelleri* has codiversified with hosts following an ancient colonization of one of the earliest groups of insects to feed on vascular plants more than 270 mya (62, 92). In each host group, *S. muelleri* pairs with a partner symbiont, which has been lost and replaced in some host lineages (2, 36, 92, 97). *S. muelleri* and its partner together provision the 10 EAAs, and partner symbionts

show little or no overlap in genes underlying production of particular amino acids (reviewed in 2, 54). *S. muelleri* genomes are 190–285 kb and are perfectly syntenic except for ongoing gene losses.

T. princeps, the obligate symbionts of phloem-feeding mealybugs, can occur singly or paired with a companion, usually a species from the *Enterobacteriaceae* (28). *T. princeps* PCIT shows the most complex integration of two bacterial species described to date: It contains its own intracellular bacterial symbiont, *M. endobia* (98), an association that coincides with the unparalleled loss of many central informational genes, including all tRNA synthetases, resulting in the most degraded set of informational genes yet reported (28, 55). In contrast, *T. phenacola* PAVE, from *Phenacoccus avenae*, lacks a symbiont partner and retains a larger gene set, encoding a full set of tRNA synthetases and amino acid biosynthetic pathways implicated in host nutrition (28).

"*Candidatus* Uzinura diaspidicola" (*Bacteroidetes*) is the obligate symbiont of armored-scale insects and, despite a tiny genome, retains genes for nitrogen recycling and EAA biosynthesis (83).

Wigglesworthia species (*Gammaproteobacteria*) are the maternally transmitted, obligate symbionts of blood-feeding tsetse flies (*Glossinidia*), including the vectors of trypanosome agents of African sleeping sickness (91). Blood provides needed amino acids but lacks some B vitamins, and these are provisioned by *Wigglesworthia* species, as inferred from its genome sequence (1). The sequenced genomes of symbionts from two host species are highly reduced (698–720 kb) but retain more genes for cell envelope production, possibly because *Wigglesworthia* species are directly exposed to host cytoplasm with no barrier of a host-derived membrane (1, 80).

Wolbachia (*Alphaproteobacteria*) includes facultative symbionts or pathogens infecting over 40% of terrestrial arthropod species (105), in which it lowers survivorship, manipulates reproduction, or protects against viruses, depending on strain and context (e.g., 6, 101). One lineage of *Wolbachia* (8) contains obligate mutualistic symbionts of filarial parasites of animals (17) and plants (22). Genomes of this lineage are 200–300 kb smaller than genomes of the facultative symbionts in arthropods (**Figure 3**). The strain in the filarial nematode *Brugia malayi* (wBm) appears to provision hosts with vitamins and cofactors (19, 21). Gene expression studies suggest that the strain with

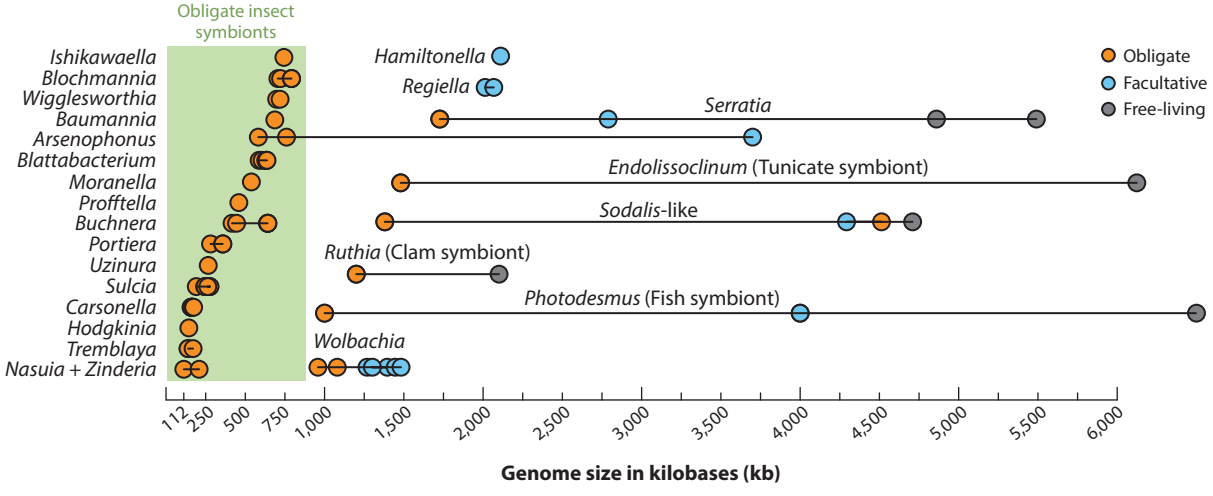


Figure 3

Genome size ranges in closely related clusters of symbiotic bacteria and relatives. Obligate symbionts are shown in orange, facultative symbionts in blue, and free-living bacteria in gray. Obligate symbionts show the tiniest genomes; those in the green box are obligate insect symbionts and include the tiniest genomes known in cellular organisms. Symbionts in younger associations with hosts are sometimes only slightly reduced in genome size.

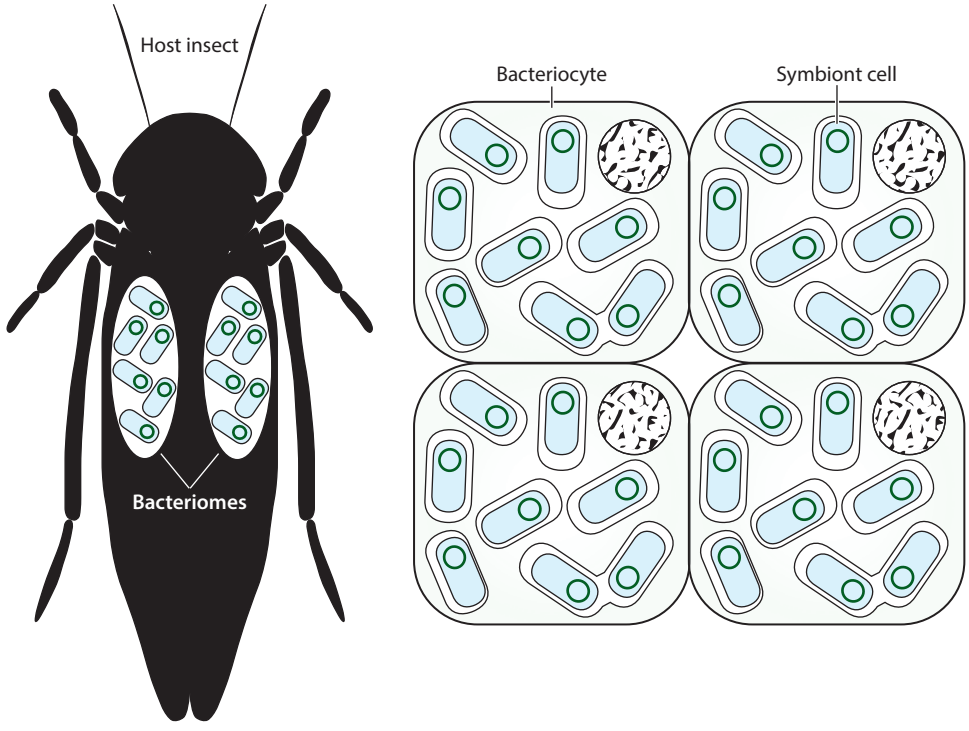


Figure 4

Diagram showing location of obligate symbionts within bacteriocytes in an insect. Individual symbiont cells are typically surrounded by a host-derived membrane within the bacteriocyte cytosol. Bacteriocytes are often clustered into a bacteriome, usually located in the insect abdomen (See **Figures 1** and **2**).

Horizontal transmission: acquisition of symbionts from individuals other than parents

Maternal transmission: transmission of symbionts to progeny by mothers

Pseudogene: a former gene inactivated by mutation

Mobile elements: Genetic elements that insert themselves into genomes, often resulting in repetitive DNA regions

the smaller genome (wOo) provides ATP to hosts and modulates the mammalian response to nematode infection (9). The mutualistic lifestyle of wOo may be yielding a more stable genome, as it has been stripped of most insertion sequences (IS) found in other *Wolbachia* genomes.

Z. insecticola (*Betaproteobacteria*) is an obligate maternally transmitted symbiont of spittlebugs and provisions histidine, methionine, and tryptophan to hosts, with the other seven EAAs provisioned by its partner *Sulcia*. *Z. insecticola* has one of the smallest known genomes (208 kb) (53) and is the sister group to *N. deltocephalinicola*, the *Sulcia* partner in some leafhoppers (2, 36).

THE ROAD TO A TINY GENOME

The tiniest genomes are all found in symbionts that are strictly maternally transmitted, restricted to living in bacteriocytes of hosts (**Figures 1**, **2**, **4**), and obligate for normal host development. Symbionts that undergo occasional horizontal transmission such as *Arsenophonus* species, "*Candidatus* Hamiltonella defensa," *Serratia symbiotica* (all *Enterobacteriaceae*), and *Wolbachia* species, are usually facultative for their hosts and typically have reduced genomes compared to nonsymbiotic relatives (**Figure 3**). Obligate symbiosis and strict maternal transmission lead to even more extreme reduction, with intermediate stages often showing proliferation of pseudogenes and mobile elements (54). For example, the nonsymbiotic species *Serratia proteamaculans* and *Serratia marcescens* have genomes of ~5.5 Mb, whereas *S. symbiotica* genomes are ~1.8 and ~2.8 Mb (5, 43). Furthermore, the smaller *S. symbiotica* genome is found in a lineage that has become essential

for hosts, whereas the larger one is facultative for hosts (43). Similarly, genomes of *Wolbachia*, which are facultative symbionts or pathogens of many arthropods, are generally reduced, at 1.3–1.5 Mb, but a lineage that has become essential for development of filarial nematode hosts (8) has smaller genomes (1.0–1.1 Mb) (9, 19). The cluster corresponding to the symbiont genera *Arsenophonus*, *Riesia*, and *Aschnera* includes facultative or obligate symbionts in various arthropods (70). In wasps of the genus *Nasonia*, *Arsenophonus nasoniae* is facultative or pathogenic and has a genome >3.7 Mb (10), whereas two lineages in the same cluster, *Riesia pediculicola* and *Aschnera chinzeii*, have become obligate symbionts of body lice and bat flies (Diptera: Nycteribiidae), respectively, and have much smaller genomes (0.6 Mb and ~0.8 Mb, respectively) (35).

Purifying selection: selection acting against mutations that lower function or inactivate a gene or functional genomic region (same as negative selection)

The genus *Sodalis* is another example of a widespread clade represented by sequenced genomes of closely related strains that differ in symbiotic status and genome size (summary in 38, 91). A nonsymbiotic strain isolated from a human wound infection (*Sodalis* HS) has a genome typical for free-living *Enterobacteriaceae* species, at 5.2 Mb, with few pseudogenes or repetitive elements (7). *Sodalis glossinidius*, a cultivable facultative symbiont of blood-feeding tsetse flies, exemplifies a genome undergoing gene loss and pseudogene formation following a recent switch to a symbiotic lifestyle (94). An obligate symbiont in this group that also appears to have originated relatively recently, SOPE of grain weevils (genus *Sitophilus*), has a somewhat reduced genome (4.5 Mb) and more dramatic gene loss due to abundant pseudogenes and proliferation of IS elements. Finally, a *Sodalis* strain that is an obligate symbiont in one group of spittlebugs has a drastically smaller genome, at ~1.4 Mb, and fewer intact genes (38).

Host restriction, in pathogens or mutualists, results in other distinctive genomic changes, including elevated rates of sequence evolution, high rates of amino acid replacement in protein sequences, and biased nucleotide base composition. In some extreme cases, the genetic code of these symbionts has been altered (2, 51, 53). These changes, along with genome reduction, are primarily driven by lack of sufficient purifying selection to maintain genes, resulting in mutations that lower gene functionality, followed by deletion of the corresponding DNA (59, 61, 99). As a result, lineages that have become host restricted in the recent evolutionary past possess many pseudogenes and inert DNA, whereas long-established symbionts display the tight gene packing characteristic of most bacteria.

Thus, we can posit two phases of genome reduction, corresponding to two general causes. First, organisms that become host restricted experience narrow habitat ranges, causing some genes to be superfluous, along with small populations and clonality, resulting in ineffective selection for the maintenance of useful but nonessential genes in many functional categories (61, 99). Some genes also may be actively removed by selection (40). The net result is gene erosion and deletion of DNA corresponding to nonessential genes. Because these genes make up the bulk of the genome of a typical bacterium, this can result in massive gene loss. These processes apply to both symbionts and pathogens, explaining why they display so many genomic similarities. But strikingly, the tiniest genomes are limited to symbionts required by hosts and inherited strictly vertically for millions of years, implying that this lifestyle facilitates even greater reduction. The second phase applies only if hosts benefit from symbiont presence and functionality and are thus selected to prop up the symbiont even as the latter is losing capabilities. The evolution of host support enables further slippage in symbiont capabilities due to gene erosion and loss, leading to ever-increasing host control and metabolic contributions and ever-decreasing autonomy of the symbiont. This process explains why bacteria beneficial to hosts may attain extreme genome reduction beyond that found in host-restricted pathogens or facultative symbionts. Mitochondria and plastids represent extreme cases in which gene loss is so extreme that they are not usually considered to be cellular entities, although clearly they once were bacteria. In the case of tiny genomes of insect symbionts, a distinct cellular

nature is still evident, as they retain most genes required for transcription, translation, and replication (49, 54), with the exception of some *Tremblaya* species (see below, What Genes Are Lost?).

EVER TINIER: ONGOING GENE LOSS IN REDUCED GENOMES

The initial reported cases of obligate symbiosis were represented by a single sequenced genome for each symbiont (usually corresponding to a particular clade of insect hosts), so it was difficult to assess whether tiny genomes attain a final, static state or whether reduction is ongoing. Thanks to next-generation sequencing technologies, there are now numerous instances in which full genomes have been sequenced and compared for multiple representatives of the same obligate symbiont group (**Figure 3**). Currently, these include two or more representatives for *S. muelleri* (2, 54), *C. ruddii* (88), *T. princeps* (28, 46), *N. deltocephalinicola* + *Z. insecticola* (2, 53), *B. aphidicola* (12, 32, 74), *Wigglesworthia* species (80), *Blattabacterium* species (65, 73, 82, 96), and *P. aleyrodidarum* (33, 89). In each case, representatives differ in the extent of genome reduction, and reconstructions of genome evolution using a phylogenetic framework reveal ongoing gene loss in individual lineages. For example, sequenced *S. muelleri* genomes range from 190 to 285 kb and include 190 to 246 protein-coding genes, and *C. ruddii* genomes range from 157 to 174 kb and include 182 to 207 protein-coding genes, with differences due to differential loss of particular genes. In each case, except for *P. aleyrodidarum*, gene order is identical or nearly identical across genomes, except for the deletions (2, 28, 73, 80, 82, 88, 89). Thus, tiny genomes show ongoing gene loss, no gene uptake, and a complete or almost complete absence of gene rearrangements. Although genome architecture is static, sequence divergence of orthologous genes is high, reflecting both the age of these clades (over 270 million years for *S. muelleri* and the *Betaproteobacteria* clade represented by *N. deltocephalinicola* + *Z. insecticola*) and rapid sequence evolution (e.g., 2, 88). Multiple genomes are also available for the less extreme genomes of *B. aphidicola*, *Blattabacterium* species, *Blochmannia* species, and *Wigglesworthia* species, and these consistently are characterized by lack of mobile elements, gene order stability, lack of gene uptake, and ongoing loss of individual genes and pathways (reviewed in 54). *P. aleyrodidarum* BT has low gene density, at 358 kb and 256 coding genes, and displays many genomic rearrangements and repetitive elements and thus is an exception to the genomic stability of other obligate maternally transmitted symbionts (89).

In some cases, the ongoing gene loss in individual genomes involves elimination of a major function. For example, the smallest reported *S. muelleri* genome, that of *S. muelleri* ALF from the leafhopper *M. quadrilineatus*, has lost nearly all genes involved in oxidative phosphorylation, including those encoding the complete cytochrome C oxidase complex, NADH dehydrogenase, menaquinone and ubiquinone biosynthesis proteins, and the molybdopterin oxioreductase complex. As a result, it is not clear how ATP synthesis can occur (2). These genes are present in *S. muelleri* of other host species, suggesting a major metabolic shift in this particular lineage.

WHAT GENES ARE LOST?

Tiny genomes have lost genes in all functional categories but tend to retain core genes underlying the central informational processes of DNA replication, transcription, and translation, indicating some self-reliance (49, 54). Despite this trend, losses of seemingly essential genes have occurred, raising the question of how replication and growth are possible. It has been suggested that symbionts with tiny genomes be relegated to a novel category and not be considered cells (79, 93). However, there is no clear gap between reduced-genome symbionts with gene repertoires similar to those of pathogens and tiny-genome symbionts: A continuum of genome sizes is now apparent (**Figure 3**). Also, this categorization does not solve the mystery of how tiny-genome symbionts

replicate and persist. Certainly, coadaptation by the host is key, but the nature of host support is not yet evident.

One caveat in a discussion of missing genes is that sometimes a gene homolog may be present but not recognized, because of extreme sequence divergence and base compositional bias. However, such cases are likely very few, as there are few orphan genes in these tiny genomes that could be candidates for the missing genes.

Homolog: a gene that is descended from the same ancestral gene and often retains the same function

Orphan gene: a gene with no known homologs and (usually) unknown function

Peptidoglycan: polymer of sugars and amino acids that makes up a component of bacterial cell walls

Central Informational Processes

Whereas *dnaE*, which underlies DNA polymerization, is retained by all of the tiny genomes, genes for other subunits of the DNA polymerase holoenzyme (e.g., *holA*, *dnaQ*, *dnaN*, *dnaX*) are often missing. DNA repair genes are one of the most depleted functional categories, with no repair genes in the tiniest genomes (**Figure 5**). Genes encoding subunits of RNA polymerase are also retained. Thus, even the smallest genomes maintain the central catalytic machinery for replication and transcription. Patterns regarding genes underlying translation are more mixed. Symbionts with small and tiny genomes usually encode a single copy of rRNA genes and retain a minimal set of tRNAs for using the full genetic code (11, 24). Additionally, they retain genes for most ribosomal proteins, although sometimes a gene for a particular ribosomal protein, usually a smaller one located on the exterior of the ribosome, such as RpmC, cannot be detected (**Figure 6**). These observations indicate that symbionts are responsible for producing their own ribosomes, which could be a criterion for considering them cellular entities. But, whereas most reduced-genome symbionts retain a set of tRNA synthetases corresponding to the full set of protein amino acids, the tiniest genomes appear to lack tRNA synthetases for particular amino acids (52, 64) (**Figure 5**). Possibly, existing tRNA synthetases are able to catabolize multiple aminoacylation reactions, as has been demonstrated in other systems (e.g., 104).

Biosynthesis of the Cellular Envelope

One of the most significant losses in obligate symbionts involves genes responsible for producing components of the bacterial cell envelope (54). Loss of ability to synthesize phospholipids or lipopolysaccharides was noted for the first sequenced symbiont genome, the moderately reduced genome of *B. aphidicola* of the pea aphid (86). In the smallest genomes, those of *N. deltocephalinicola*, *Z. insecticola*, *T. princeps*, *H. cicadicola*, and *C. ruddii*, genes for making peptidoglycan, phospholipids, lipopolysaccharide, and other components are almost completely lost. Transporters are also absent or few, with most substrate-specific transporter genes eliminated.

Biosynthetic Abilities

Capabilities to synthesize small molecules are drastically curtailed in all highly reduced genomes. For example, many obligate symbionts cannot synthesize purines and pyrimidines, many amino acids, and B vitamins, and similar losses are observed in host-restricted pathogens such as the mycoplasmas and phytoplasmas (72), reflecting dependence on the host in both cases. But obligate symbionts differ strikingly from pathogens in retaining genes that enable biosynthesis of nutrients needed by hosts. Repeatedly, obligate-symbiont genomes have been found to retain capabilities for nutrient production that are complementary with those of the host, and sometimes with those of coresident symbionts (54) (see individual symbiont descriptions above for details).

Energy Metabolism

Some of the most striking gene losses in reduced symbiont and pathogen genomes are those underlying pathways for oxidative phosphorylation and ATP production (**Figure 7**). In part, these

Figure 5

Presence and absence in obligate insect endosymbionts of genes underlying DNA replication and repair, transcription and translation, and tRNA synthesis. Genomes are ordered by size, from larger to smaller (see **Figure 4**); presence is indicated by a filled block and absence by an empty block. Half-filled blocks indicate cases in which sequenced genomes within the group differ in gene presence.

losses may relate to massive changes in membranes, in view of the missing membrane components in many small genomes. In some cases, symbionts may be ATP parasites. At the least, these losses indicate a lack of flexibility in energy acquisition, due to the constant host environment, which must directly provide some necessary metabolites to symbionts.

Figure 6

Presence and absence in obligate insect endosymbionts of genes underlying protein folding and stability, and ribosomal proteins. Genomes are ordered by size, from larger to smaller (see **Figure 4**); presence is indicated by a filled block and absence by an empty block. Half-filled blocks indicate cases in which sequenced genomes within the group differ in gene presence.

SUSTAINING SYMBIONTS WITH DEGENERATE GENOMES

How do these small genomes function despite losing genes that would be essential in most bacteria? Although this question is far from answered, studies are beginning to give clues and suggest the importance of coadaptations of both symbiont and host genes. Bacterial genes

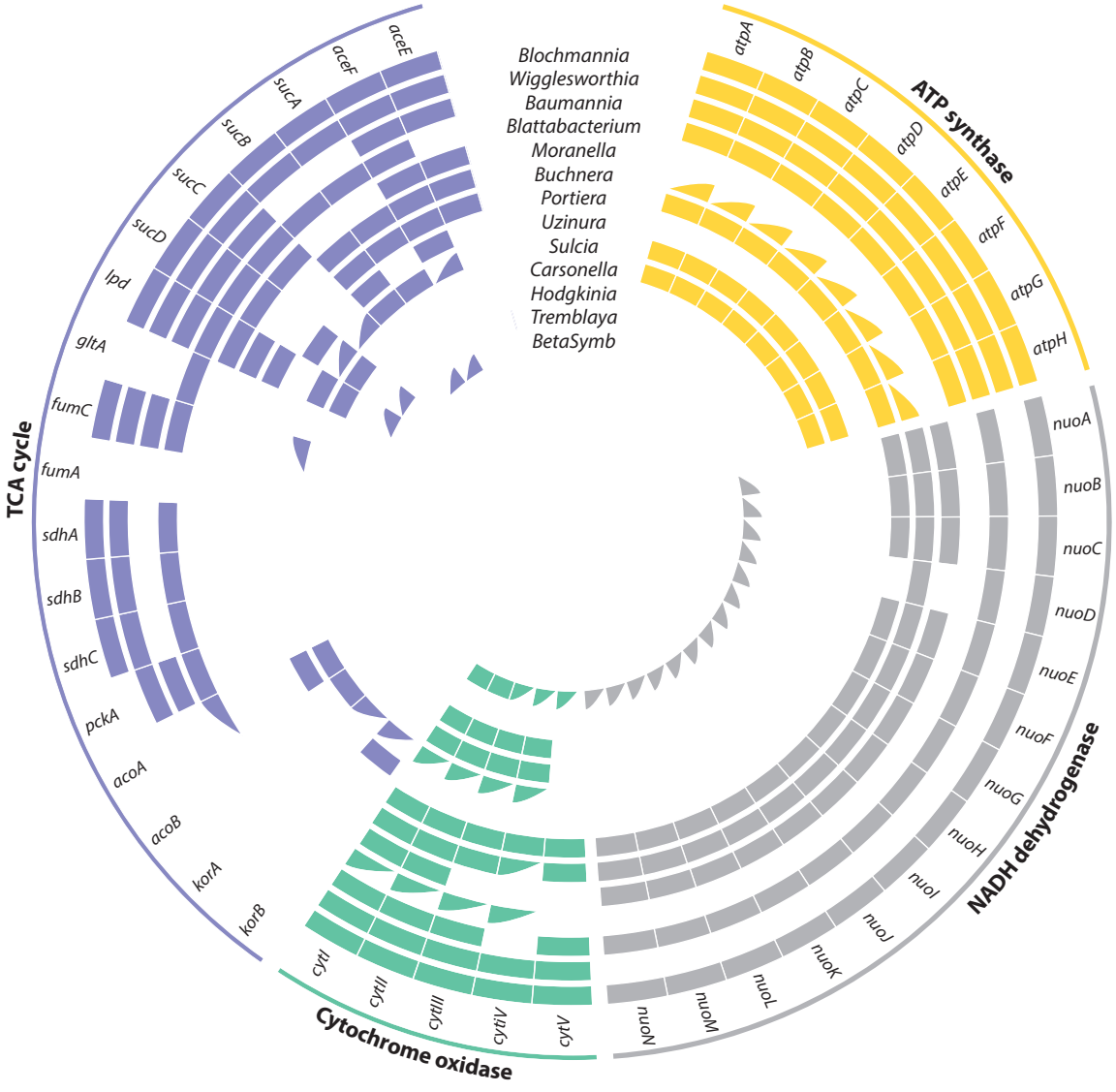


Figure 7

Presence and absence in obligate insect endosymbionts of genes underlying energy metabolism, including the TCA cycle, cytochrome oxidase, NADH dehydrogenase, and ATP synthase. Genomes are ordered by size, from larger to smaller (see **Figure 4**); presence is indicated by a filled block and absence by an empty block. Half-filled blocks indicate cases in which sequenced genomes within the group differ in gene presence.

might take on novel roles and multitask, or novel protein interactions might stabilize complexes, including the ribosome, and enable loss of other genes. Regulation of gene expression might be shifted from transcriptional regulation to posttranscriptional regulation via interactions of transcripts with small RNAs, as suggested by conserved motifs in *Buchnera* genomes (12). Genes underlying components of the cellular envelope are especially depleted, and host contributions

seem most likely in this arena. Here we briefly summarize recent experiments suggesting mechanisms for compensating losses.

Protein Stability

Genes involved in protein folding and stability are among those most consistently retained by reduced genomes: For example, *groL* and *dnaK* are among the few genes retained by all small genomes (**Figure 6**). Moreover, these genes are very highly expressed based on transcriptomic and proteomic studies in several symbiont systems (9, 15, 51, 54, 100). The elevated production of heat-shock proteins, which refold and recycle degraded cellular proteins, can compensate for protein instability due to mutation accumulation (60, 99). Nonetheless, obligate insect symbionts are heat sensitive and can limit the heat tolerances of hosts (4, 13, 16).

Heat-shock proteins: stress proteins involved in the folding and recycling of cellular proteins

Bacteriome: an organ formed by a collection of bacteriocytes, may contain one or several symbiont types

Symbiosome: organelle consisting of symbiont cells and surrounding host-derived membrane

Adaptations of Host Bacteriocytes

Key elements in host support for small-genome symbionts are adaptations of the bacteriocytes, which are often arranged into a larger organ called the bacteriome. Within bacteriocytes, symbionts are usually enclosed in a host-derived membrane, forming the symbiosome (**Figure 4**). Studies in the aphid-*Buchnera* system show that bacteriocytes display distinctive gene expression during early embryonic development (3) and participate in specific mechanisms for transmission to progeny (37). Bacteriocytes have evolved independently in each obligate symbiosis, from different cell types, and acquisition of a novel symbiont likely requires the evolution of a corresponding new bacteriocyte type (36).

Exploration of the bacteriocyte-symbiont interface is in the early stages. Studies of the gene expression patterns and proteomics of bacteriocytes in the aphid-*Buchnera* system have revealed that genes involved in amino acid metabolism and transport that complement the enzymatic machinery of *B. aphidicola* are highly upregulated in bacteriocytes as compared to other host tissues (23, 48, 75, 76). Similar complementarity of symbiont amino acid biosynthetic pathways has been found in studies of bacteriocyte transcriptomes in both psyllids and mealybugs (28, 90). One problem posed by nutritional symbioses is how host-symbiont exchange of products and substrates is regulated; this issue is especially perplexing, as symbionts have lost most regulatory genes for the biosynthetic pathways used to produce nutrients for hosts. In *B. aphidicola*, this regulation may be focused at the symbiosomal membrane, as suggested by the finding that the transporter for glutamine, the major substrate for amino acid biosynthesis, is negatively regulated by arginine, one of the EAAs donated by *B. aphidicola* (77).

Another issue is how hosts regulate symbiont numbers and confine them to bacteriocytes. In the aphid-*Buchnera* system, a class of novel aphid proteins that contain signal peptides is specifically expressed in bacteriocytes, potentially playing a role in symbiont control (85). In the symbiotic *Wolbachia* species of parasitic nematodes, symbiont surface proteins have specific interactions with host cytoskeleton and glycolytic enzymes, suggesting a mechanism for controlling the growth and distribution within hosts (57). And in the obligate symbionts (SOPE, a member of the *Sodalis* clade) living in bacteriocytes of grain weevils, an antimicrobial peptide, coleoptericin-A, inhibits cell division of symbionts and is required to prevent symbiont invasion of other tissues (45). In carpenter ants, pattern recognition receptors, overexpressed in bacteriocytes in the midgut wall, are hypothesized to play a role in regulating *Blochmannia* species (78). The broad picture from these studies is that symbiont-dependent hosts have acquired a variety of mechanisms to communicate with and control their symbionts and that lysosomal systems are commonly involved.

Horizontal Transfer to Host Genomes, and Control of Obligate Symbionts

One remarkable finding from studies aimed at detecting host genes critical to bacteriocyte function is the discovery that some of these genes are themselves horizontally transferred into the host genome from bacterial donors. The sequencing of the pea aphid genome enabled a search for genes originating from bacteria, initially motivated by the hypothesis that some former *Buchnera* genes would be transferred to the aphid nuclear genome. In fact, the aphid genome acquired a number of microbial genes, and some are highly expressed specifically in the bacteriocytes (68, 69). However, the donor lineages are close to *Wolbachia* and *Rickettsia* lineages, and not *Buchnera* lineages. Other studies revealed very similar situations in psyllids and mealybugs, in which insect genes originating from bacteria are highly expressed in bacteriocytes (28, 55, 90). Some transferred genes encode lysozymes and might play a role in regulation of symbiont populations (68); others are related to biosynthesis of cell envelope components, including peptidoglycan, and likely play a role in establishing the envelope surrounding symbionts (28). Donor bacteria include a number of groups but mostly correspond to symbionts that are not currently present in the host species, indicating that ancestral symbiotic associations have served as sources for genes that play a role in governing the current obligate symbionts.

Finally, gene transfer from obligate symbiotic *Wolbachia* species to filarial nematode hosts has also occurred, with several instances documented in which transferred genes appear active (31, 56, 102). In one case, a *Wolbachia* species donated a gene encoding the final step in the heme pathway, now an essential gene in the nematode *Brugia malayi*, agent of river blindness; this gene has been suggested as a target for treatment of filariasis (102).

CONCLUSIONS AND PERSPECTIVES

The tiniest genomes are found in ancient symbioses in which symbionts benefit hosts and hosts have in turn evolved to support and control symbionts. The evolution of tiny genomes is a continuous process, and all stages are represented among sequenced genomes. Recent experiments suggest that host adaptations are critical in enabling extreme gene loss. Surprisingly, some of these host adaptations appear to involve acquisition of novel host genes via horizontal gene transfer largely from other bacteria, including ancestral symbiont associations. Adaptations within symbiont genomes themselves are also important. Better understanding of these systems will reveal how genomes from divergent lineages, such as insect and bacteria, can intermingle and fuse to yield a functional metaorganism with a mosaic, compartmentalized genome. These instances can also potentially give some insight into processes hypothesized to underlie the origin of the eukaryotic cell.

SUMMARY POINTS

1. Very tiny genomes, less than 500 kb, have been discovered in symbionts of numerous animals and occur in diverse lineages of *Proteobacteria* and *Bacteroidetes*.
2. Tiny genomes tend to retain genes for central cellular processes, but some seemingly essential genes are missing, presenting a mystery as to how they replicate.
3. Reduced genomes occur in many host-restricted symbionts and pathogens, but tiny genomes occur only in obligate symbionts that produce molecules needed by hosts, live in specialized host cells, and diversify jointly with hosts for millions of years.

4. Obligate symbionts present a continuum of genome sizes, and even the tiniest genomes appear to continue to eliminate genes in some lineages.
5. Whereas reduced genomes can result from relaxed selection and genetic drift, tiny genomes appear to require coadaptation by hosts to replace or support symbiont functions.
6. Multiple symbionts in the same host can evolve support systems for one another, as evidenced in the case of *T. princeps* and *M. endobia*.
7. A surprising mechanism for host adaptations to support symbionts appears to involve host acquisition of genes from bacterial sources.

FUTURE ISSUES

1. Replacement of functions lost from symbiont genomes likely involves coadaptation at both host and symbiont loci, and future experiments may reveal which kinds of functions are more likely to be replaced in which way.
2. Focus on what molecules are exchanged between host and symbiont compartments would illuminate how hosts support symbionts: The import of host-derived molecules into symbiont cells, which is required for many potential mechanisms, is so far undocumented.
3. Most protein-coding genes underlying regulatory mechanisms are lost, and more attention to other means of regulation, especially small RNAs, might reveal how these tiny genomes function.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

1. Akman L, Yamashita A, Watanabe H, Oshima K, Shiba T, et al. 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat. Genet.* 32:402–7
2. Bennett GM, Moran NA. 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biol. Evol.* 5:1675–88
3. Braendle C, Miura T, Bickel R, Shingleton AW, Kambhampati S, Stern DL. 2003. Developmental origin and evolution of bacteriocytes in the aphid–*Buchnera* symbiosis. *PLoS Biol.* 1:E21
4. Burke G, Fiehn O, Moran N. 2010. Effects of facultative symbionts and heat stress on the metabolome of pea aphids. *ISME J.* 4:242–52
5. Burke GR, Moran NA. 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol. Evol.* 3:195–208
6. Chrostek E, Marialva MSP, Esteves SS, Weinert LA, Martinez J, et al. 2013. *Wolbachia* variants induce differential protection to viruses in *Drosophila melanogaster*: a phenotypic and phylogenomic analysis. *PLoS Genet.* 9:e1003896
7. Clayton AL, Oakeson KF, Gutin M, Pontes A, Dunn DM, et al. 2012. A novel human-infection-derived bacterium provides insights into the evolutionary origins of mutualistic insect-bacterial symbioses. *PLoS Genet.* 8:e1002990

8. Comandatore F, Sassera D, Montagna M, Kumar S, Koutsovoulos G, et al. 2013. Phylogenomics and analysis of shared genes suggest a single transition to mutualism in *Wolbachia* of nematodes. *Genome Biol. Evol.* 5:1668–74
9. Darby AC, Armstrong SD, Bah GS, Kaur G, Hughes MA, et al. 2012. Analysis of gene expression from the *Wolbachia* genome of a filarial nematode supports both metabolic and defensive roles within the symbiosis. *Genome Res.* 22:2467–77
10. Darby AC, Choi JH, Wilkes T, Hughes MA, Hurst GD, Colbourne JK. 2010. Characteristics of the genome of *Arsenophonus nasoniae*, son-killer bacterium of the wasp *Nasonia*. *Insect Mol. Biol.* 19:75–89
11. de Crécy-Lagard V, Marck C, Grosjean H. 2012. Decoding in Candidatus *Riesia pediculicola*, close to a minimal tRNA modification set? *Trends Cell Mol. Biol.* 7:11–34
12. Degnan PH, Ochman H, Moran NA. 2011. Sequence conservation and functional constraint on intergenic spacers in reduced genomes of the obligate symbiont *Buchnera*. *PLoS Genet.* 7:e1002252
13. Dunbar HE, Wilson ACC, Ferguson NR, Moran NA. 2007. Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol.* 5:e96
14. Dybvig K, Lao P, Jordan DS, Simmons WL. 2010. Fewer essential genes in mycoplasmas than previous studies suggest. *FEMS Microbiol. Lett.* 311:51–55
15. Fan Y, Thompson JW, Dubois LG, Moseley MA, Wernegreen JJ. 2013. Proteomic analysis of an unculturable bacterial endosymbiont (*Blochmannia*) reveals high abundance of chaperonins and biosynthetic enzymes. *J. Proteome Res.* 12:704–18
16. Fan Y, Wernegreen JJ. 2013. Can't take the heat: High temperature depletes bacterial endosymbionts of ants. *Microb. Ecol.* 66:727–33
17. Ferri E, Bain O, Barbuto M, Martin C, Lo N, et al. 2011. New insights into the evolution of *Wolbachia* infections in filarial nematodes inferred from a large range of screened species. *PLoS ONE* 6:e20843
18. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
19. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, et al. 2005. The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* 3:e121
20. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, et al. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397–403
21. Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, et al. 2007. Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* 317:1756–60
22. Haegeman A, Vanholme B, Jacob J, Vandekerckhove TTM, Claeys M, et al. 2009. An endosymbiotic bacterium in a plant-parasitic nematode: member of a new *Wolbachia* supergroup. *Int. J. Parasitol.* 39:1045–54
23. Hansen AK, Moran NA. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. USA* 108:2849–54
24. Hansen AK, Moran NA. 2012. Altered tRNA characteristics and 3′ maturation in bacterial symbionts with reduced genomes. *Nucleic Acids Res.* 40:7870–84
25. Hendry TA, de Wet JR, Dunlap PV. 2014. Genomic signatures of obligate host dependence in the luminous bacterial symbiont of a vertebrate. *Environ. Microbiol.* 16:2611–22
26. Hosokawa T, Nikoh N, Koga R, Satô M, Tanahashi M, et al. 2012. Reductive genome evolution, host-symbiont co-speciation and uterine transmission of endosymbiotic bacteria in bat flies. *ISME J.* 6:577–87
27. Huang CY, Sabree ZL, Moran NA. 2012. Genome sequence of *Blattabacterium* sp. strain BGIGA, endosymbiont of the *Blaberus giganteus* cockroach. *J. Bacteriol.* 194:4450–51
28. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, et al. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153:1567–78
29. Hutchison CA, Peterson SN, Gill SR, Cline RT, White O, et al. 1999. Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* 286:2165–69
30. Int. Aphid Genomics Consort. 2010. Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* 8:e1000313
31. Ioannidis P, Johnston KL, Riley DR, Kumar N. 2013. Extensively duplicated and transcriptionally active recent lateral gene transfer from a bacterial *Wolbachia* endosymbiont to its host filarial nematode *Brugia malayi*. *BMC Genomics* 4:639

32. Jiang Z, Jones DH, Khuri S, Tsinoremas NF, Wyss T, et al. 2013. Comparative analysis of genome sequences from four strains of the *Buchnera aphidicola* Mp endosymbiont of the green peach aphid, *Myzus persicae*. *BMC Genomics* 14:917
33. Jiang Z-F, Xia F, Johnson KW, Brown CD, Bartom E, et al. 2013. Comparison of the genome sequences of "Candidatus *Portiera aleyrodidarum*" primary endosymbionts of the whitefly *Bemisia tabaci* B and Q biotypes. *Appl. Environ. Microbiol.* 79:1757–59
34. Kambhampati S, Alleman A, Park Y. 2013. Complete genome sequence of the endosymbiont *Blattabacterium* from the cockroach *Nauphoeta cinerea* (Blattodea: Blaberidae). *Genomics* 102:479–83
35. Kirkness EF, Haas BJ, Sun W, Braig HR, Perotti MA, 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. USA* 107:12168–73
36. Koga R, Bennett GM, Cryan JR, Moran NA. 2013. Evolutionary replacement of obligate symbionts in an ancient and diverse insect lineage. *Environ. Microbiol.* 15:2073–81
37. Koga R, Meng X-Y, Tsuchida T, Fukatsu T. 2012. Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte-embryo interface. *Proc. Natl. Acad. Sci. USA* 109:E1230–37
38. Koga R, Moran NA. 2014. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. *ISME J.* 8:1237–46
39. Koonin EV. 2000. How many genes can make a cell: the minimal-gene-set concept. *Annu. Rev. Genomics Hum. Genet.* 1:99–116
40. Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012. Selection-driven gene loss in bacteria. *PLoS Genet.* 8:e1002787
41. Kwan JC, Donia MS, Han AW, Hirose E, Haygood MG, Schmidt EW. 2012. Genome streamlining and chemical defense in a coral reef symbiosis. *Proc. Natl. Acad. Sci. USA* 109:20655–60
42. Kwan JC, Schmidt EW. 2013. Bacterial endosymbiosis in a chordate host: long-term co-evolution and conservation of secondary metabolism. *PLoS ONE* 8:e80822
43. Lamelas A, Gosalbes MJ, Manzano-Marín A, Peretó J, Moya A, Latorre A. 2011. *Serratia symbiotica* from the aphid *Cinara cedri*: a missing link from facultative to obligate insect endosymbiont. *PLoS Genet.* 7:e1002357
44. Lo N. 2003. Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. *Mol. Biol. Evol.* 20:907–13
45. Login FH, Balmand S, Vallier A, Vincent-Monégat C, Vigneron A, et al. 2011. Antimicrobial peptides keep insect endosymbionts under control. *Science* 334:362–65
46. López-Madrigal S, Latorre A, Porcar M, Moya A, Gil R. 2013. Mealybugs nested endosymbiosis: going into the "matryoshka" system in *Planococcus citri* in depth. *BMC Microbiol.* 13:74
47. López-Sánchez MJ, Neef A, Peretó J, Patiño-Navarrete R, Pignatelli M, et al. 2009. Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*. *PLoS Genet.* 5:e1000721
48. Macdonald SJ, Lin GG, Russell CW, Thomas GH, Douglas AE. 2012. The central role of the host cell in symbiotic nitrogen metabolism. *Proc. Biol. Sci.* 279:2965–73
49. McCutcheon JP. 2010. The bacterial essence of tiny symbiont genomes. *Curr. Opin. Microbiol.* 13:73–78
50. McCutcheon JP, McDonald BR, Moran NA. 2009. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc. Natl. Acad. Sci. USA* 106:15394–99
51. McCutcheon JP, McDonald BR, Moran NA. 2009. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.* 5:e1000565
52. McCutcheon JP, Moran NA. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc. Natl. Acad. Sci. USA* 104:19392–97
53. McCutcheon JP, Moran NA. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. *Genome Biol. Evol.* 2:708–18
54. McCutcheon JP, Moran NA. 2012. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10:13–26
55. McCutcheon JP, von Dohlen CD. 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. *Curr. Biol.* 21:1366–72

56. McNulty SN, Foster JM, Mitreva M, Dunning Hotopp JC, Martin J, et al. 2010. Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *PLoS ONE* 5:e11029
57. Melnikow E, Xu S, Liu J, Bell AJ, Ghedin E, et al. 2013. A potential role for the interaction of *Wolbachia* surface proteins with the *Brugia malayi* glycolytic enzymes and cytoskeleton in maintenance of endosymbiosis. *PLoS Negl. Trop. Dis.* 7:e2151
58. Merhej V, Raoult D. 2011. Rickettsial evolution in the light of comparative genomics. *Biol. Rev. Camb. Philos. Soc.* 86:379–405
59. Mira A, Ochman H, Moran NA. 2001. Deletional bias and the evolution of bacterial genomes. *Trends Genet.* 17:589–96
60. Moran NA. 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA* 93:2873–78
61. Moran NA. 2002. Microbial minimalism: genome reduction in bacterial pathogens. *Cell* 108:583–86
62. Moran NA, Tran P, Gerardo NM. 2005. Symbiosis and insect diversification: an ancient symbiont of sap-feeding insects from the bacterial phylum *Bacteroidetes*. *Appl. Environ. Microbiol.* 71:8802–10
63. Nakabachi A, Ueoka R, Oshima K, Teta R, Mangoni A, et al. 2013. Defensive bacteriome symbiont with a drastically reduced genome. *Curr. Biol.* 23:1478–84
64. Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, et al. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267
65. Neef A, Latorre A, Peretó J, Silva FJ, Pignatelli M, Moya A. 2011. Genome economization in the endosymbiont of the wood roach *Cryptocercus punctulatus* due to drastic loss of amino acid synthesis capabilities. *Genome Biol. Evol.* 3:1437–48
66. Newton ILG, Woyke T, Auchtung TA, Dilly GF, Dutton RJ, et al. 2007. The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* 315:998–1000
67. Nikoh N, Hosokawa T, Oshima K, Hattori M, Fukatsu T. 2011. Reductive evolution of bacterial genome in insect gut environment. *Genome Biol. Evol.* 3:702–14
68. Nikoh N, McCutcheon JP, Kudo T, Miyagishima S-Y, Moran NA, Nakabachi A. 2010. Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. *PLoS Genet.* 6:e1000827
69. Nikoh N, Nakabachi A. 2009. Aphids acquired symbiotic genes via lateral gene transfer. *BMC Biol.* 7:12
70. Nováková E, Hypsa V, Moran NA. 2009. *Arsenophonus*, an emerging clade of intracellular symbionts with a broad host distribution. *BMC Microbiol.* 9:143
71. Ochman H. 2005. Genomes on the shrink. *Proc. Natl. Acad. Sci. USA* 102:11959–60
72. Oshima K, Maejima K, Namba S. 2013. Genomic and evolutionary aspects of phytoplasmas. *Front. Microbiol.* 4:230
73. Patiño-Navarrete R, Moya A, Latorre A, Peretó J. 2013. Comparative genomics of *Blattabacterium cuenoti*: the frozen legacy of an ancient endosymbiont genome. *Genome Biol. Evol.* 5:351–61
74. Pérez-Brocal V, Gil R, Ramos S, Lamelas A, Postigo M, et al. 2006. A small microbial genome: the end of a long symbiotic relationship? *Science* 314:312–13
75. Poliakov A, Russell CW, Ponnala L, Hoops HJ, Sun Q, et al. 2011. Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. *Mol. Cell Proteomics* 10:M110.007039
76. Price DRG, Duncan RP, Shigenobu S, Wilson ACC. 2011. Genome expansion and differential expression of amino acid transporters at the aphid/*Buchnera* symbiotic interface. *Mol. Biol. Evol.* 28:3113–26
77. Price DRG, Feng H, Baker JD, Bavan S, Luetje CW, Wilson ACC. 2014. Aphid amino acid transporter regulates glutamine supply to intracellular bacterial symbionts. *Proc. Natl. Acad. Sci. USA* 111:320–25
78. Ratzka C, Gross R, Feldhaar H. 2013. Gene expression analysis of the endosymbiont-bearing midgut tissue during ontogeny of the carpenter ant *Camponotus floridanus*. *J. Insect Physiol.* 59:611–23
79. Reyes-Prieto M, Latorre A, Moya A. 2014. Scanty microbes, the 'symbionelle' concept. *Environ. Microbiol.* 16:335–38
80. Rio RVM, Symula RE, Wang J, Lohs C, Wu Y-N, et al. 2012. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: Glossinidae) obligate symbiont *Wigglesworthia*. *MBio* 3:e00240–11
81. Rogers MJ, Simmons J, Walker RT, Weisburg WG, Woese CR, et al. 1985. Construction of the mycoplasma evolutionary tree from 5S rRNA sequence data. *Proc. Natl. Acad. Sci. USA* 82:1160–64

82. Sabree ZL, Huang CY, Arakawa G, Tokuda G, Lo N, et al. 2012. Genome shrinkage and loss of nutrient-providing potential in the obligate symbiont of the primitive termite *Mastotermes darwiniensis*. *Appl. Environ. Microbiol.* 78:204–10
83. Sabree ZL, Huang CY, Okusu A, Moran NA, Normark BB. 2012. The nutrient supplying capabilities of *Uzinura*, an endosymbiont of armoured scale insects. *Environ. Microbiol.* 15:1988–99
84. Sabree ZL, Kambhampati S, Moran NA. 2009. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc. Natl. Acad. Sci. USA* 106:19521–26
85. Shigenobu S, Stern DL. 2013. Aphids evolved novel secreted proteins for symbiosis with bacterial endosymbiont. *Proc. Biol. Sci.* 280:20121952
86. Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids, *Buchnera* sp. APS. *Nature* 407:81–86
87. Sloan DB, Moran NA. 2012. Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biol. Lett.* 8:986–89
88. Sloan DB, Moran NA. 2012. Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. *Mol. Biol. Evol.* 29:3781–92
89. Sloan DB, Moran NA. 2013. The evolution of genomic instability in the obligate endosymbionts of whiteflies. *Genome Biol. Evol.* 5:783–93
90. Sloan DB, Nakabachi A, Richards S, Qu J, Murali SC, et al. 2014. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol. Biol. Evol.* 31:857–71
91. Snyder AK, Rio RVM. 2013. Interwoven biology of the tsetse holobiont. *J. Bacteriol.* 195:4322–30
92. Takiya DM, Tran PL, Dietrich CH, Moran NA. 2006. Co-cladogenesis spanning three phyla: leafhoppers (Insecta: Hemiptera: Cicadellidae) and their dual bacterial symbionts. *Mol. Ecol.* 15:4175–91
93. Tamames J, Gil R, Latorre A, Peretó J, Silva FJ, Moya A. 2007. The frontier between cell and organelle: genome analysis of Candidatus *Carsonella ruddii*. *BMC Evol. Biol.* 7:181
94. Toh H, Weiss BL, Perkin SAH, Yamashita A, Oshima K, et al. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res.* 16:149–56
95. Tokuda G, Elbourne LDH, Kinjo Y, Saitoh S, Sabree Z, et al. 2013. Maintenance of essential amino acid synthesis pathways in the *Blattabacterium cuenoti* symbiont of a wood-feeding cockroach. *Biol. Lett.* 9:20121153
96. Tokuda G, Lo N, Watanabe H, Slaytor M, Matsumoto T, Noda H. 1999. Metazoan cellulase genes from termites: intron/exon structures and sites of expression. *Biochim. Biophys. Acta* 1447:146–59
97. Urban JM, Cryan JR. 2012. Two ancient bacterial endosymbionts have coevolved with the planthoppers (Insecta: Hemiptera: Fulgoroidea). *BMC Evol. Biol.* 12:87
98. von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature* 412:433–36
99. Wernegreen JJ. 2002. Genome evolution in bacterial endosymbionts of insects. *Nat. Rev. Genet.* 3:850–61
100. Wilcox JL, Dunbar HE, Wolfinger RD, Moran NA. 2003. Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol. Microbiol.* 48:1491–500
101. Woolfit M, Iturbe-Ormaetxe I, Brownlie JC, Walker T, Riegler M, et al. 2013. Genomic evolution of the pathogenic *Wolbachia* strain, wMelPop. *Genome Biol. Evol.* 5:2189–204
102. Wu B, Novelli J, Jiang D, Dailey HA, Landmann F, et al. 2013. Interdomain lateral gene transfer of an essential ferrochelatase gene in human parasitic nematodes. *Proc. Natl. Acad. Sci. USA* 110:7748–53
103. Wu D, Daugherty SC, Van Aken SE, Pai GH, Watkins KL, et al. 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol.* 4:e188
104. Zhang C-M, Hou Y-M. 2004. Synthesis of cysteinyl-tRNACys by a prolyl-tRNA synthetase. *RNA Biol.* 1:35–41
105. Zug R, Hammerstein P. 2012. Still a host of hosts for *Wolbachia*: Analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS ONE* 7:e38544