

Molecular Interactions between Bacterial Symbionts and Their Hosts

Colin Dale^{1,*} and Nancy A. Moran^{2,*}

¹Department of Biology, University of Utah, Salt Lake City, UT 84112 USA

²Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721 USA

*Contact: dale@biology.utah.edu (C.D.); nmoran@u.arizona.edu (N.A.M.)

DOI 10.1016/j.cell.2006.07.014

Symbiotic bacteria are important in animal hosts, but have been largely overlooked as they have proved difficult to culture in the laboratory. Approaches such as comparative genomics and real-time PCR have provided insights into the molecular mechanisms that underpin symbiont-host interactions. Studies on the heritable symbionts of insects have yielded valuable information about how bacteria infect host cells, avoid immune responses, and manipulate host physiology. Furthermore, some symbionts use many of the same mechanisms as pathogens to infect hosts and evade immune responses. Here we discuss what is currently known about the interactions between bacterial symbionts and their hosts.

Introduction

Symbiotic bacteria are ubiquitous in animal hosts. They affect development (Braendle et al., 2003; Koropatnick et al., 2004), nutrition (Baumann, 2005; Backhed et al., 2005), reproduction and speciation (reviews in Hurst and Werren, 2001; Bandi et al., 2001; Hurst and Jiggins, 2000; Stouthamer et al., 1999), defense against natural enemies (Piel, 2002; Scarborough et al., 2005; Oliver et al., 2003), and immunity (Macdonald and Monteleone, 2005). In these associations, bacteria form persistent infections within host individuals. In many invertebrate hosts, they live an intracellular existence for much of their life cycle and undergo vertical transmission from mother to offspring via infection of eggs or embryos. "Obligate" symbionts also show highly specific interactions with host cell populations, both during colonization of progeny and during early development of host individuals (e.g., Braendle et al., 2003). Moreover, these bacteria have adapted such that they have little or no negative impact on the host's overall health; often, they are beneficial or even obligatory (Figure 1).

As more cases of chronic bacterial infection are characterized, the distinction between pathogenesis and mutualism has become increasingly blurred. Infection by a particular bacterium may be beneficial to a host under some circumstances but harmful in other hosts or environments. Furthermore, as examples in this review will illustrate, pathogenic and mutualistic bacteria often use the same core molecular mechanisms to maintain their associations with hosts.

We will use "symbiont" to encompass the variety of bacteria that form sustained infections in hosts, in which colonization and replication are controlled without compromising host vigor. We specifically focus on organisms that invade the body cavity or the cytosol of

host cells (rather than the surface or gut lumen). Given that most work has focused on maternally transmitted insect symbioses, these will represent a major focus. We note that findings of shared molecular mechanisms with pathogens have also been a theme in studies of other symbiotic systems. These include the mutualistic *Euprymna scolopes-Vibrio fischeri* light organ symbiosis, in which recent discoveries indicate that pathways involved in innate immunity are used by the squid host to recognize and facilitate colonization of the beneficial *V. fischeri* symbiont (Goodson et al., 2005; Koropatnick et al., 2004). Other relevant cases are the *Photobacterium luminescens* mutualism with entomopathogenic nematodes and the *Yersinia pestis* association with arthropod vectors and mammalian hosts. Finally, we expect that heritable symbionts of invertebrates will have many common features with bacteria that are mutualistic with vertebrates. Increasingly, studies of the microbial communities colonizing vertebrates, particularly the gut microflora, are revealing intimate coevolved associations (Backhed et al., 2004, 2005).

Categorizing Symbioses

Beginning with early descriptive studies based on microscopy (Buchner, 1965), and later molecular studies, animal symbioses have been categorized based on both apparent evolutionary age and on the extent of obligate codependence between the host and symbiont (Figure 2). At one extreme are the ancient "primary" symbionts that live only in bacteriomes (host organs specialized for housing symbiotic associates) and are required by hosts. At the other extreme, "secondary" or "facultative" symbionts usually do not reside exclusively in specialized organs and are not strictly necessary for host survival; like pathogens, they are able to engineer their own invasion of host cells and tissues. Com-

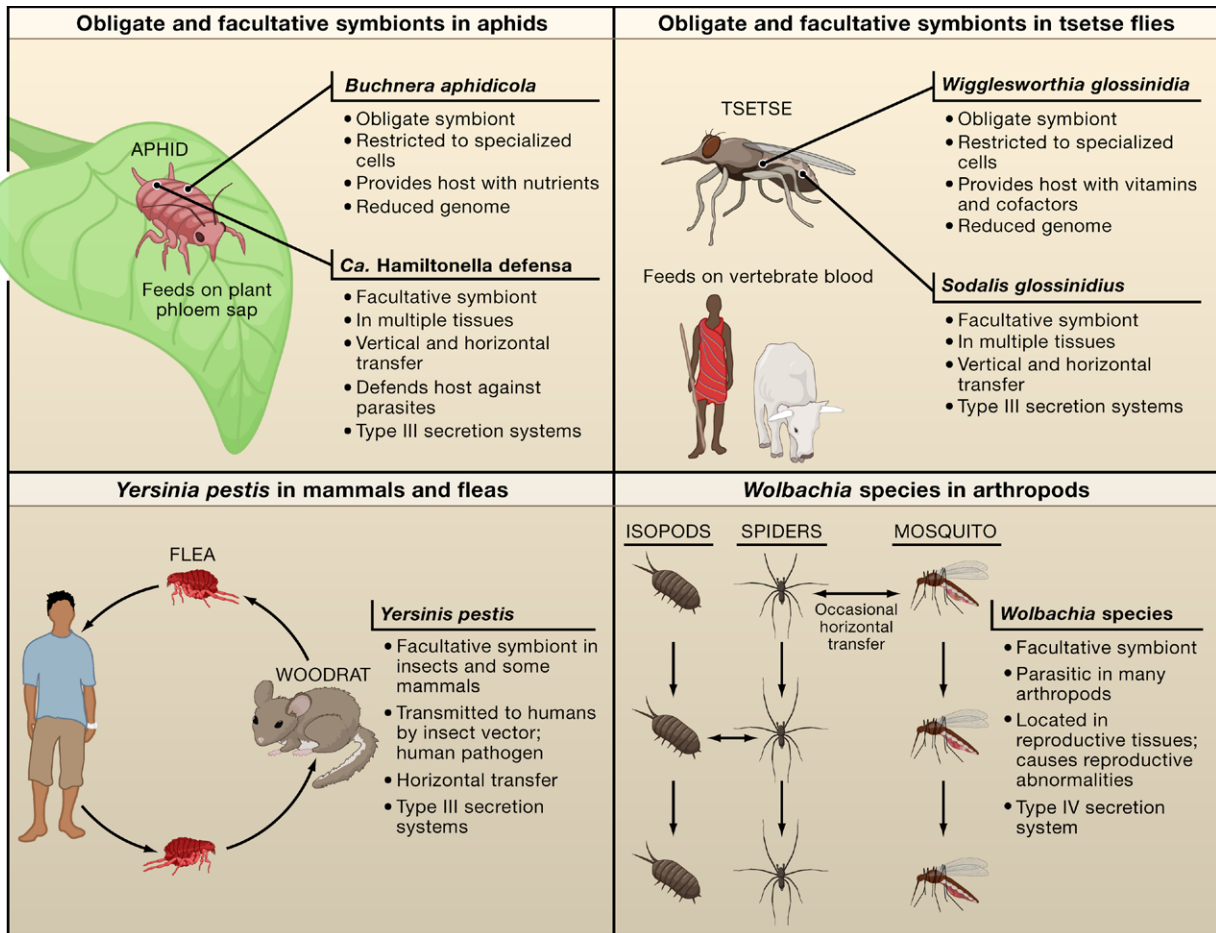


Figure 1. Important Features in the Associations between Insects and Their Symbiotic Bacteria

monly, multiple symbionts reside within the same host individual, most often in the form of one or two primary endosymbionts and one or two secondary or facultative endosymbionts.

Primary Symbionts: Ancient and Obligate

Repeatedly, primary symbionts have been shown to descend from ancient and specialized associations: their phylogenetic trees are congruent with those of their hosts over long periods of evolutionary time, encompassing the initiation of symbiosis through to the present (Figure 2; Munson et al., 1991; Baumann, 2005; Lo et al., 2003; Moran et al., 2003; Wu et al., 2006). Estimated ages of these associations range from 30 to 270 million years, with many insect symbioses predating the origin of mammals, angiosperms, and other dominant modern groups of organisms. This long period of host-dependent evolution is reflected in massive alterations in the genomes of primary symbionts (described below). In primary symbiotic associations, the hosts have adapted to maintain rather than eliminate symbionts. Therefore, these interactions bear little resemblance to those involving pathogens.

Secondary Symbionts: Recent and Facultative

Unlike primary symbionts, secondary symbionts can colonize naive hosts through horizontal transmission among host individuals and species. These associations are typically facultative from the perspective of the host, and they can be deleterious or beneficial (Figure 1). As they are primarily maternally transmitted, these symbionts must persist and spread in host populations using mechanisms that give infected matrilineal an advantage over uninfected matrilineal. In some cases, this advantage derives from beneficial effects on host survival or reproductive rates, such as protection against parasites or stress. For example, recent experimental approaches have demonstrated a variety of these beneficial effects conferred on hosts by symbionts in the γ -Proteobacteria (Moran et al., 2005c; Oliver et al., 2003; Scarborough et al., 2005). In other cases, involving *Wolbachia*, *Rickettsia* species, *Spiroplasma* species, and others, symbionts manipulate host reproduction to enhance matrilineal transmission through parthenogenesis, the selective killing of male embryos (“son-killing”), feminization of genetic males, or reproductive incompatibility (reviewed

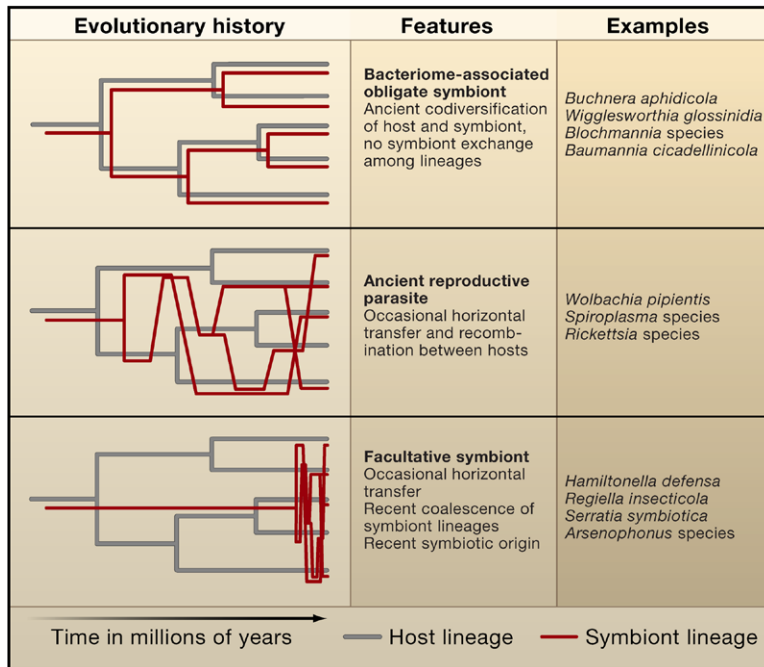


Figure 2. Distinct Evolutionary Patterns of Heritable Symbionts in Insects

in Stouthamer et al., 1999; McGraw and O'Neill, 2004). All of these influences have the effect of increasing representation of the female lineages bearing the infection, although they may lower the evolutionary fitness of hosts by compromising genetic contributions of males. Even in cases of “selfish” symbionts, i.e., reproductive manipulators causing net decreases in host fitness, effects on female survivorship and fecundity are typically negligible (e.g., Bordenstein and Werren, 2000).

Although horizontal transmission occurs, vertical inheritance often proceeds with effectively 100% efficiency under lab conditions. A recent study of *Wolbachia* sheds some light on how these organisms interact with the microtubules of developing insect oocytes in order to regulate their numbers and their subcellular position (Ferree et al., 2005).

The Genomes of Insect Symbionts

As few symbionts have been cultured *in vitro*, and many are refractory to genetic manipulation, DNA sequence data provide the primary basis for understanding symbiotic associations at the molecular level. These sequences have proved useful in the reconstruction of host-symbiont phylogenies and in the analysis of symbiont genes and genomes. Additionally, experimental strategies have been developed that exploit the foundation of gene and genome sequences in order to investigate the molecular mechanisms underlying symbiotic interactions.

Whole-genome sequencing efforts are likely to expand as the technical difficulties associated with sequencing symbiont genomes are overcome. A major

reason that several obligate mutualistic genomes were sequenced before genomes of facultative symbionts is the relative ease of isolation of DNA, as well as their very small sizes. These primary symbionts, unlike facultative symbionts, are densely packed in host tissues and often occur in discrete organs that often can be easily collected following host dissection. Several recent developments in DNA technology offer promise for improved methods of obtaining large numbers of copies of symbiont genomic DNA. For example, multiple displacement amplification (MDA) using tiny amounts of template DNA has been successfully exploited to yield genome scans of symbionts for characterization of gene sets (Moran et al., 2005a; Mavingui et al., 2005). The use of MDA may enable rapid progress in complete genome sequencing, especially with new large-scale sequencing technologies on the horizon. The development of culture techniques for facultative symbionts, either within insect cell lines (O'Neill et al., 1997; Darby et al., 2005) or in pure culture (Dale and Maudlin, 1999; Dale et al., 2006), can also facilitate genome sequencing projects, as in the case of *S. glossinidius* (Toh et al., 2006).

Primary Symbionts: “Minimal Genomes”

An important axis of variation among maternally inherited symbionts is the extent of genome degradation and reduction. At one extreme are the highly reduced bacteriome-associated primary symbionts, the category for which the largest number of published genomes are available. Examples include three genomes from *Buchnera aphidicola* strains from different aphid hosts (Shigenobu et al., 2000; Tamas et al., 2002; van Ham et al., 2003), two from *Candidatus* *Blochmannia* species from ants (Gil et al., 2004; Degnan et al., 2005), one from *Wigglesworthia glossinidia* from tsetse flies (Akman et al., 2002), and one from *Candidatus* *Baumannia cicadellinicola* from leafhoppers (Wu et al., 2006). All of these ancient symbionts are in the γ -Proteobacteria and bear close phylogenetic relationships to *Escherichia coli* and other well-studied enteric bacteria. But, whereas *E. coli* and other free-living relatives in this group have genomes of about five megabases encoding some 5000 genes, primary symbiont genomes are under one megabase in size and are known to encode as few as 500 genes, with some possibly smaller (Gomez-Valero et al., 2004). The genome sequences of primary symbionts show striking similarities beyond reduced size. In each case, few or no genes have been acquired as part of the transition to a symbiotic lifestyle; this absence of gene acquisition is a major contrast to other genomes in representa-

tives of the γ -Proteobacteria. Rather, the gene set of the ancestor has been selectively reduced so as to retain only those genes and pathways required for the symbiotic lifestyle.

A major point of interest regarding primary symbionts is the demonstration that a free-living bacterium with a large genome has the potential to undergo massive reduction, retaining only 10%–20% of the original genes. The gene sets retained in the symbiont lineages so far sequenced differ substantially. Numerous genes considered essential, based on studies in *E. coli*, are absent from one or more of the sequenced symbiont genomes; an example is *dnaA*, a gene involved in initiation of chromosome replication, which is absent from the *Wigglesworthia* and *Blochmannia* genomes (Akman et al., 2002; Gil et al., 2003), raising the possibility of replication control by the host. Likewise, *Buchnera* genomes are among the few lacking the near-universal recombinase gene, *recA*, and they also lack the genes underlying phospholipid synthesis (Shigenobu et al., 2000; van Ham et al., 2003; Zientz et al., 2004). Although transfer of genes to the host nucleus and reimportation of host-encoded gene products have not been definitively ruled out, no experimental evidence for such transfer is available. Moreover, primary symbiont genomes published to date do possess most genes that would be considered essential based on studies of model organisms, and they can therefore be considered to represent versions of “minimal” bacterial genomes (Gil et al., 2004; Zientz et al., 2004). These compact genomes lack many usual features of genomes of related free-living and pathogenic bacteria. For example, they have no phage or phage remnants, no insertion sequences, and no evidence of foreign gene uptake or homologous recombination among strains. In general, their genome sequences indicate that these obligate symbionts represent highly specialized organisms that are evolutionarily committed to dependence on their particular hosts.

Several lines of evidence indicate that during long-term coevolution, primary symbionts have relinquished control of their own genetic functions to hosts (Figure 2). First, primary symbionts show extreme depletion of ancestral regulatory genes. For example, whereas *E. coli* has more than 200 genes encoding transcriptional regulators, *Buchnera* retains fewer than 10 (Moran et al., 2005b). Regarding amino acid biosynthesis, which is central to the aphid-*Buchnera* symbiosis, the sequenced *Buchnera* genomes retain most structural genes for the 10 essential amino acid pathways, but only 0 or 1 (depending on the strain) of the 16 transcriptional regulators for these pathways (Wilcox et al., 2003).

As demonstrated in early studies, the universal heat shock chaperonin *groEL* is constitutively highly expressed in *Buchnera* and in other endosymbionts and host-restricted organisms (Kakeda and Ishikawa, 1991; Aksoy, 1995). The GroEL protein constitutes about 10% of total protein in *Buchnera* cells, even under nonstress conditions (Baumann et al., 1996), and the *groEL* tran-

script is one of the most abundant in *Buchnera* cells (Wilcox et al., 2003). Assaying transcript levels across the entire genome using microarrays revealed that constitutively elevated expression extends to most other genes that are typically part of the heat shock regulon (Wilcox et al., 2003). These observations strongly suggest that the general function of the overexpression of *groEL* and other stress-related genes involves stabilization and refolding of proteins, rather than a newly evolved function specific to GroEL. An increased investment in mechanisms for protein stabilization has been postulated to have evolved as a compensation for accumulated mutations that reduce protein stability (Moran, 1996; Wernegreen and Moran, 1999; van Ham et al., 2003; Fares et al., 2002).

Overall growth rates and densities of bacteriome-associated bacteria appear to closely track the growth and development of the host (Wolschin et al., 2004; Baumann and Baumann, 1994), implying cooperative mechanisms for coordination of host growth and symbiont replication. These control mechanisms are certainly less rigid in associations involving facultative symbionts, such as *Wolbachia* and the secondary symbionts of sap-feeding insects, because these bacteria show highly variable densities and patterns of colonization among host individuals (Koga et al., 2003; Oliver et al., 2006; Veneti et al., 2004, 2005).

Facultative Symbioses: Genomes in Transition

More recently, complete or partial genome sequences have been obtained from a number of facultative symbionts that appear to have established associations with hosts in the more recent evolutionary past. Features of these genomes may resemble those of genomes in the early stages of a transition from a free-living lifestyle to an obligate mutualism; this is the most likely interpretation of the recently published genome sequence of the tsetse fly symbiont, *S. glossinidius*. Almost half of this genome is comprised of inactivated genes in various stages of decomposition (Toh et al., 2006), resembling the case of the host-restricted pathogen *Mycobacterium leprae* (Cole et al., 2001), and illustrating the similarity in the evolutionary trajectories of some chronic pathogens and symbionts (Moran and Plague, 2004). In some cases, these organisms have also accumulated many transposable elements (Moran and Plague, 2004; Toh et al., 2006). Some facultative symbionts, such as *Wolbachia*, *Candidatus Hamiltonella defensa*, and *Candidatus Arsenophonus arthropodicus*, have genomes of intermediate size (1–3.5 megabases) containing mobile elements and phage (Moran et al., 2005a; Wu et al., 2004; Dale et al., 2006). Compared to the domesticated bacteriome-associates, these facultative symbionts may undergo high levels of recombination among strains, as is now well documented for *Wolbachia* (Figure 2; Baldo et al., 2006). Possibly, some of the repetitive sequences found in the published insect-associated *Wolbachia* genome are adaptations for promoting recombination within or among genomes (Wu et al., 2004). Additionally,

facultative symbionts undergo bacteriophage-mediated gene exchange (e.g., Moran et al., 2005a; Bordenstein and Wernegreen, 2004).

Obligate Symbioses: Domestication of Bacteria by Animal Hosts

The provision of nutrients has long been suspected of being a *raison d'être* of primary symbionts (Buchner, 1965), and this suspicion has recently been confirmed and elucidated by whole-genome sequencing and microarray analyses. Insect primary symbionts, including *Buchnera*, *Wigglesworthia*, *Blochmannia*, and *Baumannia*, have the capacity to provide their hosts with nutrients, based on the gene inventories inferred from genome sequences. An in-depth consideration of the metabolic contributions and requirements of symbionts, as inferred from complete genome sequences, is provided by Zientz et al. (2004). The sets of nutrients provided differ among the various associations. Thus, *Buchnera* retains pathways for biosynthesis of amino acids essential in animal nutrition, but has lost those for biosynthesis of nonessential amino acids (Shigenobu et al., 2000). The genes underlying production of the essential amino acids leucine and tryptophan are amplified on plasmids in the *Buchnera* strains of some aphid species (Lai et al., 1994; Baumann, 2005), probably to enable overproduction of these nutrients. The genome of *Wigglesworthia*, the obligate symbiont of tsetse flies, lacks amino acid biosynthetic genes but contains many genes underlying synthesis of a large array of essential cofactors. This observation fits with the dietary constraints of the tsetse host, which feeds exclusively on vertebrate blood, a diet that presents a full spectrum of necessary amino acids but is lacking in certain cofactors. Recently, genomic sequences for *Baumannia* and *Sulcia*, the coresident primary symbionts of xylem-feeding sharpshooters, have revealed that these symbionts have complementary sets of biosynthetic capabilities, with one (*Baumannia*) able to make many cofactors and the other (*Sulcia*) encoding pathways for synthesis of essential amino acids (Wu et al., 2006). In each of the completely sequenced primary symbionts, 10%–14% of the genome is comprised of genes devoted to host nutrition. This retention of genes underlying biosynthetic pathways provides a striking contrast to obligate pathogens with similarly small genomes, such as sequenced *Mycoplasma* or *Rickettsia* spp., which obtain nutrients from their hosts and have eliminated almost all genes underlying amino acid biosynthesis.

Some regulation of nutrient production and exchange of materials must exist within the insect-symbiont system, in order to prevent overproduction of the requisite end products. For bacteriome-associated primary symbionts, the host likely enforces such controls. A study of differentially expressed host genes in the bacteriocytes provides general support for this view, as genes underlying transport and amino acid metabolism are upregulated in these cells (Nakabachi et al., 2005).

Facultative Symbioses: Invasion and Establishment in Naive Hosts

Some symbiotic bacteria can readily invade novel host species and form stable associations, a claim that is based on inferences from phylogenies and host distributions (Werren et al., 1995; Baumann, 2005) and also from experimental transfection of naive hosts (e.g., Russell and Moran, 2005; Dobson et al., 2002). This fact implies that no host coadaptation is required. Thus, the bacteria themselves possess mechanisms for invading host cells and forming stable associations, including invasion of developing eggs or embryos. Recent data from genome sequences and experimental studies suggest that these symbionts draw on the same set of molecular tools used by relatively well-studied pathogenic bacteria. In the following sections, we describe some of the mechanisms that facultative symbionts use to establish and maintain their infections in host tissues.

Protein Secretion Systems

Many bacteria use protein secretion systems to facilitate the translocation of macromolecules from the cell cytoplasm, across the cell membrane(s), onto the surface of the bacterial cell, or into the extracellular environment. In Gram-negative bacteria that have both inner and outer membranes, five distinct secretion systems (designated type I through type V) are known to enable the capability to translocate proteins, designated effector proteins, outside of the cell. For a general review of protein secretion systems in bacteria, see Hueck (1998) and Pugsley et al. (2004). In this review, we focus on type III and type IV secretion systems (T3SSs and T4SSs) that are known to play an important role in the translocation of proteins from certain insect endosymbionts. For reviews of type III and type IV secretion pathways, respectively, see Mota et al. (2005) and Christie (2004).

Research into bacterial protein secretion systems began in earnest with the discovery of a number of so-called “pathogenicity islands” in the genomes of enteric pathogens (Groisman and Ochman, 1996). These genomic regions are composed of clusters of genes encoding virulence factors that are important or essential in the interactions between pathogenic bacteria and their eukaryotic hosts. In many cases, the proteins encoded by these gene clusters have a coordinated function such as the deployment of a multiprotein secretion apparatus on the cell surface. Although protein secretion systems, especially T3SSs, have been identified in pathogenicity islands found in the genomes of a wide range of bacteria that cause disease in both plants and animals, it is important to point out that it is not the protein secretion systems per se, but rather the effector proteins they deliver, that are potentially harmful in nature. This proves important when we consider the role of protein secretion systems in mutualistic bacterial symbionts, because these organisms are not expected to inflict harm upon their hosts.

The first evidence for involvement of a T3SS in an animal-bacterium mutualism came from a study focus-

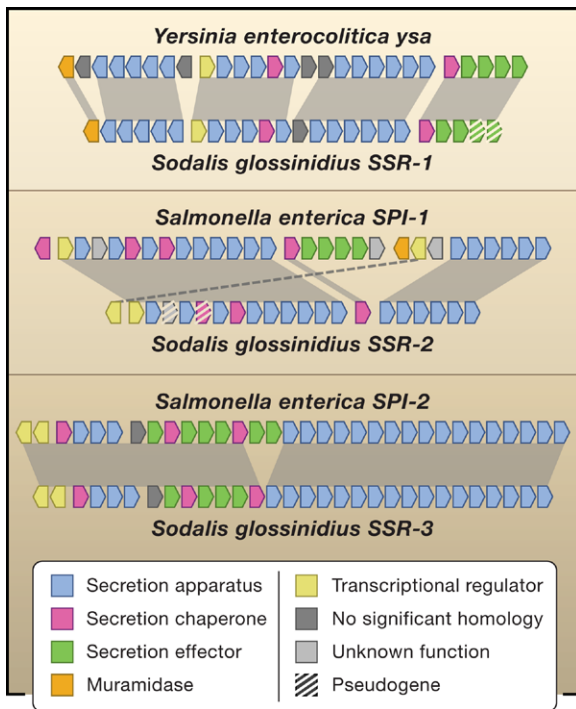


Figure 3. Organization of T3SS-Encoding Chromosomal Regions in *Sodalis glossinidius*, *Salmonella enterica*, and *Yersinia enterocolitica*

The three *S. glossinidius* islands (SSR-1, SSR-2, and SSR-3) are most closely related to the *Y. enterocolitica* *ysa* island and the *S. enterica* SPI-1 and SPI-2 islands, respectively. Individual genes are depicted in color boxes with colors corresponding to predicted functions shown in the key. Regions of synteny are highlighted by the connecting gray bars, and the translocation of *hilA* in *S. glossinidius* SSR-2 and *S. enterica* SPI-1 is highlighted by a dotted line. Note that each of the *S. glossinidius* T3SS-encoding islands is reduced in gene content relative to the pathogens. Notably, all three *S. glossinidius* islands lack one or more genes encoding putative effector proteins.

ing on the tsetse fly endosymbiont, *Sodalis glossinidius* (Dale et al., 2001). *S. glossinidius* is one of only a handful of insect endosymbionts that have been cultured in vitro, both in an insect cell line (Welburn et al., 1987) and on solid media (Dale and Maudlin, 1999). In *S. glossinidius*, transposon (Tn5) mutagenesis revealed the presence of T3SS genes, such as *invC*, that were required for *S. glossinidius* to enter insect cells in vitro. A *S. glossinidius* *invC* mutant was also incapable of establishing infection in vivo in its natural host, the tsetse fly, and analysis of the genic regions flanking the *invC* homolog in *S. glossinidius* revealed additional T3SS genes.

In a subsequent study, the complete gene inventories of two *Sodalis* T3SS-encoding regions were obtained from BAC clones (Dale et al., 2005). The two regions were found to carry genes encoding phylogenetically distinct T3SS systems, designated SSR-1 and SSR-2. The SSR-1 region carries the genes identified in the Tn5 screen, which are most closely related to homologs in the *ysa* island in *Yersinia enterocolitica* (Foultier et al., 2002). The SSR-2 island carries a paralogous set of

T3SS genes most closely related to the SPI-1 genes found in *Salmonella enterica* and *Chromobacterium violaceum*. Although both SSR-1 and SSR-2 have a complete set of genes encoding a functional T3SS export apparatus, these regions lack many genes encoding effector proteins found in the corresponding *ysa* island in *Y. enterocolitica* and the SPI-1 island in *S. enterica*. Indeed, the only intact genes encoding putative effector proteins in *S. glossinidius* are located in SSR-1, and they likely encode effector proteins that specifically facilitate the host cell cytoskeletal rearrangements necessary for bacterial entry. This suggests that the functions of the *S. glossinidius* T3SSs have been reduced in the transition from parasitism to mutualism. Transcript analyses indicate that SSR-1 and SSR-2 have distinct temporal patterns of gene expression (Dale et al., 2005). Whereas SSR-1 transcripts are detected when bacteria are entering insect cells, SSR-2 transcripts are detected after an intracellular infection is established. Mutant *S. glossinidius*, lacking an essential component of SSR-2, can invade cells, but fail to replicate once they become intracellular. The recent completion of the *S. glossinidius* whole genome sequence (Toh et al., 2006) will provide the opportunity to search for new T3SS effectors and has already revealed the presence of a third T3SS encoding region in the *S. glossinidius* chromosome (designated SSR-3) whose function is currently unknown. The gene organization of the three T3SS encoding regions in the *S. glossinidius* genome is shown in Figure 3.

A primary question that remains relates to the origin of T3SS genes in the *Sodalis* genome. Given that the T3SSs have important functions in the mutualism between *S. glossinidius* and its host, did the acquisition of the T3SS genes predate the origin of the symbiotic association? Moreover, are T3SSs found in mutualistic insect endosymbionts that are not close relatives of *Sodalis*? To date, T3SS components have been identified in the genomes of three facultative endosymbionts, including *Sodalis*, and in two closely related primary symbionts (discussed below). In general, T3SS gene clusters are found only in the genomes of insect endosymbionts that have established relatively recent associations with hosts. The more ancient primary endosymbionts, including *Buchnera*, *Blochmannia*, and *Wigglesworthia*, have no T3SS components in their genomes, although it is conceivable that they possessed T3SSs during earlier stages of their evolution.

The two primary insect endosymbionts known to have T3SSs are found in the coleopteran grain weevils, *Sitophilus oryzae* and *S. zeamais*. These bacteria have no formal nomenclature and are known as SOPE and SZPE, respectively. Their association with hosts is estimated to be recent in origin (about 25 million years; Lefevre et al., 2004). PCR screens have revealed the presence of SSR-1 and SSR-2 homologs in the weevil symbiont genomes, and comparative analyses of pairwise sequence divergence between the T3SS genes and chromosomal informational genes indicate that both SSR-1 and SSR-2 were

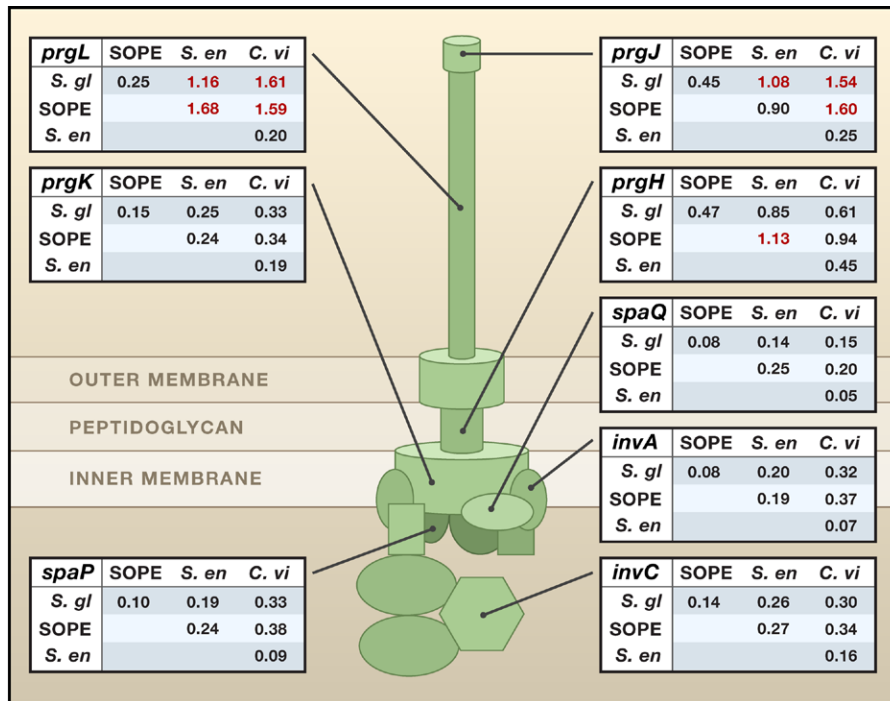


Figure 4. Relationship between Cellular Location and Evolutionary Rate in T3SS-Encoding Proteins

The model SPI-1/SSR-2 secretome reveals the approximate cellular locations of T3SS proteins relative to the bacterial cell wall. The tables contain estimates of protein evolutionary rates relative to neutral sequence changes (dN/dS) computed for selected T3SS proteins in pairwise comparisons among the mutualistic endosymbionts *Sodalis glossinidius* (*S. gl*) and SOPE, and the pathogens *Salmonella enterica* (*S. en*) and *Chromobacterium violaceum* (*C. vi*). The dN/dS values are elevated in pairwise comparisons between mutualistic endosymbionts and pathogens, and the highest dN/dS ratios (highlighted in red) are observed in T3SS proteins that have an extracellular location.

present in a common ancestor of the lineage leading to *S. glossinidius* and the weevil endosymbionts (Dale et al., 2002). The complete gene inventories of SSR-1 and SSR-2 have recently been obtained from both SOPE and SZPE, and the analysis of these sequences provides further support for common ancestry of both T3SS gene clusters in all three symbionts (C.D., unpublished data). The functions of the SSR-1 and SSR-2 islands in SOPE and SZPE are not well defined because the bacteria have not been cultured. However, real-time PCR assays of transcripts in vivo suggests that the T3SS genes play an important role during a specific stage of weevil development (Dale et al., 2002).

The availability of the complete sequences of SSR-1 and SSR-2 from SOPE and SZPE also provides some insight into the mode of evolution of the T3SS genes in the insect endosymbionts. This is of particular relevance to genes encoding protein components of the needle substructure in SSR-2 that were initially predicted to be nonfunctional due to extremely high levels of amino acid substitutions (Dale et al., 2005). Molecular evolutionary analyses indicate that rates of protein sequence evolution are elevated in protein components of the T3SS exposed to the extracellular environment (Figure 4). Host-pathogen conflicts were predicted to be a driving force for high rates of polypeptide evolution in the extracellular components of the *Salmonella* SPI-1 T3SS (Li et al., 1995; Boyd et al., 1997). Surprisingly, the relationship between cellular location and protein evolutionary rate holds true for the SSR-2 T3SSs in mutualistic endosymbionts, despite the fact that these organisms are not expected to engage in conflict with their hosts.

Furthermore, the estimated rates of polypeptide evolution in extracellular components of the T3SS are highest in pairwise comparisons between mutualistic endosymbionts and pathogens (Figure 4). Based on these observations, we propose that the extracellular protein components of the SPI-1 and SSR-2 T3SSs undergo bursts of rapid polypeptide evolution simply to facilitate the mechanical adaptations necessary for the T3SS to interface with different eukaryotic cell types.

The genes encoding T3SSs have also recently been identified in two facultative endosymbionts that are not close relatives of *S. glossinidius*. The first of these is *Ca. Hamiltonella defensa*, a facultative symbiont of numerous insects that, in the pea aphid, plays a role in the defense of hosts against parasitoid attack (Oliver et al., 2003, 2005; Moran et al., 2005c). In a preliminary sequence analysis of the *Ca. Hamiltonella defensa* genome, 22 putative T3SS genes were recovered from a total of 473 predicted coding sequences belonging to two T3SS subfamilies corresponding to SPI-1 and SPI-2 in *S. enterica* (Moran et al., 2005a). T3SS genes have also been identified in the genome of *Ca. Arsenophonus arthropodicus*, a facultative endosymbiont that is found in a wide range of arthropod hosts and has recently been isolated in pure culture (Dale et al., 2006; C.D., unpublished data).

Type IV secretion systems are known to play an important role in a number of plant and animal pathogens that secrete proteins and nucleoproteins into host cells including *Agrobacterium tumefaciens*, *Helicobacter pylori*, *Bordetella pertussis*, *Brucella* spp., *Bartonella henselae*, *Rickettsia prowazekii*, and *Campylo-*

bacter jejuni (Sexton and Vogel, 2002). A complete set of T4SS genes were found in the sequenced genomes of the mutualistic *Wolbachia* (wBm) from the nematode *Brugia malayi* and the parasitic *Wolbachia* (wMel) from the insect *Drosophila melanogaster* as well as in a number of *Wolbachia* strains isolated from other insects (Wu et al., 2004; Foster et al., 2005). Although no definitive functions have been identified for the *Wolbachia* T4SS to date, it does seem likely that the T4SS has an important function based on the fact that it is maintained in so many *Wolbachia* strains. Indeed, one can easily envisage a role for the T4SS in the manifestation of the reproductive distortions that accompany *Wolbachia* infections in many arthropod hosts, because these distortions are likely to be mediated by proteins secreted by *Wolbachia*.

Dynamic Processes in Genomes of Facultative Symbionts

The numerical distribution of mobile DNA in host-associated bacteria, like that of secretion systems, is skewed in favor of those (mostly facultative) symbionts that have established recent associations with hosts. This is particularly notable in the case of transposon and bacteriophage sequences, which are largely absent from the genomes of ancient primary endosymbionts.

Plasmids have been identified in some primary endosymbionts, including *Buchnera* and *Wigglesworthia* (e.g., Lai et al., 1994). However, the presence of genes on plasmids likely reflects a need to modulate gene copy number rather than to facilitate horizontal transfer of DNA, as noted earlier in the case of *Buchnera* plasmids bearing genes for leucine and tryptophan biosynthesis (Wernegreen and Moran, 2001; Latorre et al., 2005). Although this may seem like a trivial point, it does illustrate that many mobile elements and plasmids have impacts on bacterial genomes that are unrelated to their ability to vector genetic material. These auxiliary impacts include the modification of genome organization and the modulation of gene copy number and expression.

Mobile Genetic Elements and Repetitive DNA in Symbiont Genomes

Insertion sequence (IS) elements represent the simplest forms of mobile genetic elements found in the genomes of bacteria. They use a simple transposase to effect "cut and paste" excision and insertion of transposon DNA in the bacterial genome. Although these transposons have no ability to direct their own replication, they are known to utilize host replication machinery to effect "cut, copy, and paste" transposition (Mahillon and Chandler, 1998). In their simplest form, IS elements are considered selfish because they carry only those genes necessary to effect their own transposition. However, even in this form, they can impact the host genome by silencing genes, mediating rearrangements, and modulating gene expression.

Several authors have noted that intracellular bacteria have a tendency to accumulate IS element sequences in their genomes during the early stages of host association (Moran and Plague, 2004; Bordenstein and Reznikoff,

2005). As host restriction is known to reduce the acquisition of mobile DNA, it seems likely that the increase in IS element numbers in these genomes reflects an increase in the replicative transposition of elements that are resident at the onset of symbiosis. To explain this phenomenon, Moran and Plague (2004) postulated a neutral hypothesis for IS element proliferation in the early stages of symbiosis. This takes into account changes predicted to occur in the population structure of bacteria switching from a free-living stage with a large effective population size to a symbiotic stage with a small effective population size. Thus, as effective population size decreases, the efficiency of purifying selection acting on genic sequences is reduced and more IS elements become fixed in the population. Genes whose functions are redundant or only mildly beneficial become effectively neutral and act as space into which IS elements can transpose with no deleterious effect. As IS element copy number increases, we would expect a concomitant increase in the frequency of IS element transposition. However, as transposition frequency increases, so does the opportunity for transpositions to have deleterious effects, and for this reason IS element expansions are expected to be inherently self-limiting.

The genome sequences of the mutualistic and parasitic strains of *Wolbachia*, wBm and wMel, are notable because they have markedly different quantities of repetitive DNA, composed mostly of inactivated IS elements. Whereas the genome of wBm has only 5.4% repetitive DNA (Foster et al., 2005), the genome of wMel contains 14.2% repetitive DNA (Wu et al., 2004). The increased level of repetitive DNA in the wMel genome is correlated with a history of genome instability, characterized by an aberrant pattern of G+C skew across the entire wMel chromosome (Wu et al., 2004). This has led some authors to speculate that IS element proliferation might confer an adaptive advantage by increasing the frequency of genome rearrangements (Foster et al., 2005; Brownlie and O'Neill, 2005). This hypothesis is only plausible under circumstances in which genome rearrangements promote phenotypic diversity, which may be advantageous for *Wolbachia* during host switches. However, the more evident and direct impact of IS transposition is deleterious (because most events will interrupt open reading frames or transcription units), and the more parsimonious explanation of increased IS frequency in recently host-restricted bacteria is a reduced effectiveness of natural selection in limiting their spread.

Bacteriophage and Toxins in Symbiotic Systems

Bacteriophages represent the most abundant organisms on Earth (Wommack and Colwell, 2000), and perhaps the most dynamic of the mobile genetic elements found in prokaryotes. They have a unique ability to transfer genetic information between isolated populations and therefore make a substantial contribution to the bacterial gene pool within ecosystems.

In the bacterial genomes sequenced to date, prophage sequences are particularly abundant in the bac-

terial pathogens (Brussow et al., 2004). In contrast, even prophage remnants are entirely absent from the genomes of ancient host associated endosymbionts (Bordenstein and Reznikoff, 2005). Again, this likely reflects the genome streamlining associated with long-term degenerative evolution and does not exclude the possibility that these bacteria harbored phages in the earlier stages of their symbiotic associations. However, the genome sequences of more recently established host associated bacteria, including *Wolbachia* spp., *S. glossinidius*, and *Ca. Hamiltonella defensa*, are known to contain complete prophage genomes and substantial numbers of prophage remnants (Wu et al., 2004; Foster et al., 2005; Moran et al., 2005a; Toh et al., 2006). Therefore, it is clear that bacteriophages are capable of transducing genetic material into the genomes of insect endosymbionts during the early stages of symbiotic life, but what are the consequences of these phage transduction events?

To illustrate the role of phage in the dissemination of genes encoding toxins in insect endosymbionts, we focus on two examples in which bacteria harbor active prophage genomes carrying genes expected to encode protein toxins capable of affecting eukaryotic cells. The first example centers on the parasitic strains of *Wolbachia*, many of which are known to harbor a bacteriophage known as phage WO (Masui et al., 2000). Phage WO is widespread among strains of parasitic *Wolbachia*, and there is evidence of recent lateral transfer of WO prophage DNA between these strains (Bordenstein and Wernegreen, 2004). Phage WO particles have been visualized in insect tissues (Masui et al., 2001) and purified in the laboratory (Fujii et al., 2004), further supporting the notion that phage WO undertakes infectious transmission.

One of the interesting features of the WO prophage genome is the presence of ANK genes encoding proteins with ankyrin repeat domains (Masui et al., 2000; Wu et al., 2004). The genomes of *Wolbachia* strains that cause reproductive distortions in their hosts have unusually high numbers of ANK genes relative to the mutualistic *Wolbachia* wBm strain and their other close relatives in the α -Proteobacteria (Iturbe-Ormaetxe et al., 2005). Proteins containing ankyrin repeat domains act as scaffolds in a variety of protein-protein interactions in eukaryotic cells (Mosavi et al., 2004). They participate in a wide range of cellular functions, including cell signaling, cytoskeleton integrity, cell-cycle regulation, gene regulation, and development. In *Wolbachia*, ANK genes have been linked to the cytoplasmic incompatibility (CI) phenotype through a number of potential mechanisms, including the ANK protein-mediated inhibition of host cyclin-dependant kinase activity (Iturbe-Ormaetxe et al., 2005; Tram et al., 2003). A notable recent study focused on a comparative sequence analysis of ANK genes in two closely related *Wolbachia* strains infecting geographically distinct populations of *Culex* mosquitoes that are incompatible as a result of *Wolbachia*-induced

CI (Sinkins et al., 2005). Surprisingly, the *Wolbachia* strains infecting these incompatible mosquitoes are so closely related that they share identical sequences in their surface protein-encoding *wsp* gene, which is known to contain hypervariable regions and therefore evolves rapidly (Baldo et al., 2005). The close genetic relationship between the two strains is further exemplified by the comparative sequence analysis of their ANK genes; 16 out of a total of 18 different ANK genes are predicted to encode identical polypeptides in the two incompatible *Wolbachia* strains. The remaining two ANK genes found in each genome are notable because they show divergence between the two incompatible strains (8.4% and 3.4% amino acid sequence divergence). These two divergent ANK genes happen to be located within a WO prophage region and are coexpressed with a WO phage capsid gene that is actively transcribed in adult mosquitoes. Curiously, expression of one of these ANK genes was demonstrated to occur specifically in female mosquitoes and represents the only *Wolbachia* gene known so far to display sex-dependent expression in the host insect. The observation of sex-dependent phage gene expression was corroborated in a more recent study focusing on *Culex* mosquitoes (Sanogo and Dobson, 2006). Taken together, these results provide strong support for the direct involvement of the ANK genes in the induction of the CI phenotype in *Wolbachia*. They also implicate bacteriophage WO in the dissemination and transcriptional regulation of genes involved in the CI phenotype.

The second example of phage and toxin gene transduction in an insect endosymbiont involves *Ca. Hamiltonella defensa*, a facultative bacterial symbiont residing in aphids and other insects. Bacteriophage particles, named APSE-1, were initially discovered in the tissues of the aphid *Acyrtosiphon pisum* and subjected to complete genome sequence analysis (van der Wilk et al., 1999) several years before the identity of the bacterial host, *Ca. Hamiltonella defensa*, was established (Moran et al., 2005c). As a close relative of the *S. enterica* phage P22, APSE-1 is classified as a temperate phage and the APSE-1 and P22 genomes both maintain a full complement of genes anticipated to be necessary for both lysogeny and lysis. The APSE-1 genome is notable because it also carries a gene homologous to the Shiga toxin (*Stx*) encoding genes found in prophages residing in important human and animal pathogens (Herold et al., 2004). The potential significance of the *stx* gene in the APSE-1 genome was not highlighted until Moran et al. (2005a) discovered additional haplotypes of the APSE-1 phage (named APSE-2a and APSE-2b) carrying homologs of *cdtB* instead of *stx*; *cdtB* is expected to encode the active subunit of a holotoxin known as cytolethal distending toxin (Thelestam and Frisan, 2004). The complete genome sequences of the two haplotypes of APSE-2 were derived from an assembly of a partial *Ca. Hamiltonella defensa* genome sequence obtained through whole genome amplification of symbiont DNA

from an iso-female line of aphids. It therefore seems likely that APSE-2a and APSE-2b are lysogenized in the same bacterial genome, and Southern hybridization confirmed that at least one copy was lysogenized. The copy number of APSE-2 genes (including *cdtB*) was estimated to be 40-fold higher than single-copy chromosomal genes based on quantitative PCR analyses and on the bias observed in representation of phage genes in the sequence assembly. Given that the APSE phages cannot replicate autonomously, the elevated copy number of the APSE-2 sequences must result from the continuous production of phage particles in the symbiont cells. But what advantages might be conferred by this process given that the opportunities for phage transmission are likely to be restricted in the insect body? Perhaps ongoing lysis and particle production provide an opportunity for the enhanced production and release of certain phage-associated proteins, such as Stx and CdtB. Indeed, real-time PCR experiments have demonstrated that the *cdtB* gene is expressed at a high level *in vivo* in *Ca. Hamiltonella defensa*, some 5-fold higher than the highly expressed chromosomal gene *dnaK* and 15-fold higher than the APSE-2 structural gene P2. It is also known that the sequestration and delivery of Shiga toxin by bacterial pathogens is mediated by the ongoing induction of Shiga toxin-encoding prophages (Wagner et al., 2002). Despite the presence of an ongoing lytic cycle, the symbiotic infection is stable within individual hosts and within lineages of aphids, indicating that a portion of the symbiont population in each host escapes lysis and colonizes developing progeny or eggs. Although there is no direct evidence of a functional role for toxin production in the aphid, it has been suggested that toxins might play a role in the defense of the host insect against eukaryotic parasites. Indeed, *Ca. Hamiltonella defensa* isolates have been shown to protect their aphid hosts against attack by hymenopteran parasites and to effect different levels of protection in aphid hosts (Oliver et al., 2003, 2005).

Concluding Remarks and Future Directions

Recent attempts to characterize symbiotic bacteria in just a few insect groups have revealed a diverse range of intimate associations that appear to have evolved repeatedly from a number of different ancestral bacterial lineages. Stable symbiotic associations, such as those considered in this review, are likely to be least as common in nature as those associations that revolve around parasitism, or pathogenesis. Over the course of the last few decades, the focus of mainstream microbiology has been directed toward vertebrate pathogens in an effort to reduce the burden of human and animal disease. However, it is naive to assume that the molecular mechanisms of pathogenesis evolved specifically to facilitate disease in vertebrates. Rather, it seems more likely that the mechanisms of pathogenesis, including protein secretion systems, evolved in bacterial symbionts that undertook interactions with more ancient

eukaryotes, including arthropods and other invertebrate phyla (McFall-Ngai, 2002). This leads to two important points: First, there is much to learn about the evolution of virulence and pathogenesis from the study of interactions between bacteria and invertebrate hosts. Second, we may discover that many more vertebrate pathogens have the ability to infect and reside within nonvertebrate hosts. Evidence of the latter was recently obtained from a study focusing on *Y. enterocolitica*, which was found to produce insecticidal proteins at low temperature, despite the fact that no insect host is known to harbor this bacterium (Bresolin et al., 2006). Finally, vertebrates possess many beneficial associations with bacteria, which have been shown to provide both nutritional and defensive advantages, and invertebrate models can be used to elucidate the role of these bacteria in human and animal health.

ACKNOWLEDGMENTS

We thank Becky Nankivell and Kelly Oakeson for helping to edit the manuscript. The authors are supported by NSF awards 0315815 and 0313737 (to N.M.) and EF-58501127 (to C.D.).

REFERENCES

- Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., and Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat. Genet.* *32*, 402–407.
- Aksoy, S. (1995). Molecular analysis of the endosymbionts of tsetse flies: 16S rDNA locus and over-expression of a chaperonin. *Insect Mol. Biol.* *4*, 23–29.
- Backhed, F., Ding, H., Wang, T., Hooper, L.V., Koh, G.Y., Nagy, A., Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. USA* *101*, 15718–15723.
- Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. *Science* *307*, 1915–1920.
- Baldo, L., Lo, N., and Werren, J.H. (2005). Mosaic nature of the *Wolbachia* surface protein. *J. Bacteriol.* *187*, 5406–5418.
- Baldo, L., Bordenstein, S., Wernegreen, J.J., and Werren, J.H. (2006). Widespread recombination throughout *Wolbachia* genomes. *Mol. Biol. Evol.* *23*, 437–449.
- Bandi, C., Dunn, A.M., Hurst, G.D., and Rigaud, T. (2001). Inherited microorganisms, sex-specific virulence and reproductive parasitism. *Trends Parasitol.* *17*, 88–94.
- Baumann, P. (2005). Biology of bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu. Rev. Microbiol.* *59*, 155–189.
- Baumann, L., and Baumann, P. (1994). Growth kinetics of the endosymbiont *Buchnera aphidicola* in the aphid *Schizaphis graminum*. *Appl. Environ. Microbiol.* *60*, 3440–3443.
- Baumann, P., Baumann, L., and Clark, M.A. (1996). Levels of *Buchnera aphidicola* chaperonin GroEL during growth of the aphid *Schizaphis graminum*. *Curr. Microbiol.* *32*, 279–285.
- Bordenstein, S.R., and Reznikoff, W.S. (2005). Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* *3*, 688–699.
- Bordenstein, S.R., and Wernegreen, J.J. (2004). Bacteriophage flux in endosymbionts (*Wolbachia*): Infection frequency, lateral transfer, and

- recombination rates. *Mol. Biol. Evol.* **21**, 1981–1991.
- Bordenstein, S.R., and Werren, J.H. (2000). Do *Wolbachia* influence fecundity in *Nasonia vitripennis*? *Heredity* **84**, 54–62.
- Boyd, E.F., Li, J., Ochman, H., and Selander, R.K. (1997). Comparative genetics of the inv-spa invasion gene complex of *Salmonella enterica*. *J. Bacteriol.* **179**, 1985–1991.
- Braendle, C., Miura, T., Bickel, R., Shingleton, A.W., Kambhampati, S., and Stern, D.L. (2003). Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. *PLoS Biol.* **1**, E21.
- Bresolin, G., Morgan, J.A., Ilgen, D., Scherer, S., and Fuchs, T.M. (2006). Low temperature-induced insecticidal activity of *Yersinia enterocolitica*. *Mol. Microbiol.* **59**, 503–512.
- Brownlie, J.C., and O'Neill, S.L. (2005). *Wolbachia* genomes: Insights into an intracellular lifestyle. *Curr. Biol.* **15**, R507–R509.
- Brussow, H., Canchaya, C., and Hardt, W.D. (2004). Phages and the evolution of bacterial pathogens: From genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **68**, 560–602.
- Buchner, P. (1965). *Endosymbiosis of Animals with Plant Microorganisms*. (New York: John Wiley)
- Christie, P.J. (2004). Type IV secretion: The Agrobacterium VirB/D4 and related conjugation systems. *Biochim. Acta* **1694**, 219–234.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Garnier, T., Churcher, C., Harris, D., et al. (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007–1011.
- Dale, C., and Maudlin, I. (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.
- Dale, C., Young, S.A., Haydon, D.T., and Welburn, S.C. (2001). The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. *Proc. Natl. Acad. Sci. USA* **98**, 1883–1888.
- Dale, C., Plague, G.R., Wang, B., Ochman, H., and Moran, N.A. (2002). Type III secretion systems and the evolution of mutualistic endosymbiosis. *Proc. Natl. Acad. Sci. USA* **99**, 12397–12402.
- Dale, C., Jones, T., and Pontes, M. (2005). Degenerative evolution and functional diversification of type-III secretion systems in the insect endosymbiont *Sodalis glossinidius*. *Mol. Biol. Evol.* **22**, 758–766.
- Dale, C., Beeton, M., Harbison, C., Jones, T., and Pontes, M. (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.
- Darby, A.C., Chandler, S.M., Welburn, S.C., and Douglas, A.E. (2005). Aphid-symbiotic bacteria cultured in insect cell lines. *Appl. Environ. Microbiol.* **71**, 4833–4839.
- Degnan, P.H., Lazarus, A.B., and Wernegreen, J.J. (2005). Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects. *Genome Res.* **15**, 1023–1033.
- Dobson, S.L., Marsland, E.J., Veneti, Z., Bourtzis, K., and O'Neill, S.L. (2002). Characterization of *Wolbachia* host cell range via the in vitro establishment of infections. *Appl. Environ. Microbiol.* **68**, 656–660.
- Fares, M.A., Ruiz-Gonzalez, M.X., Moya, A., Elena, S.F., and Barrio, E. (2002). Endosymbiotic bacteria: *groEL* buffers against deleterious mutations. *Nature* **417**, 398.
- Ferree, P.M., Frydman, H.M., Li, J.M., Cao, J., Wieschaus, E., and Sulivan, W. (2005). *Wolbachia* utilizes host microtubules and Dynein for anterior localization in the *Drosophila* oocyte. *PLoS Pathog.* **1**, e14.
- Foster, J., Ganatra, M., Kamal, I., Ware, J., Makarova, K., Ivanova, N., Bhattacharyya, A., Kapatral, V., Kumar, S., Posfai, J., et al. (2005). The *Wolbachia* genome of *Brugia malayi*: Endosymbiont evolution within a human pathogenic nematode. *PLoS Biol.* **4**, e121.
- Foutier, B., Troisfontaines, P., Muller, S., Opperdoes, F.R., and Cornelis, G.R. (2002). Characterization of the *ysa* pathogenicity locus in the chromosome of *Yersinia enterocolitica* and phylogeny analysis of type III secretion systems. *J. Mol. Evol.* **55**, 37–51.
- Fujii, Y., Kubo, T., Ishikawa, H., and Sasaki, T. (2004). Isolation and characterization of the bacteriophage WO from *Wolbachia*, an arthropod endosymbiont. *Biochem. Biophys. Res. Commun.* **317**, 1183–1188.
- Gil, R., Silva, F.J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., Rausell, C., Kamerbeek, J., Gadau, J., Holldobler, B., et al. (2003). The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. USA* **100**, 9388–9393.
- Gil, R., Silva, F.J., Pereto, J., and Moya, A. (2004). Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.* **68**, 518–537.
- Gomez-Valero, L., Soriano-Navarro, M., Perez-Brocal, V., Heddi, A., Moya, A., Garcia-Verdugo, J.M., and Latorre, A. (2004). Coexistence of *Wolbachia* with *Buchnera aphidicola* and a secondary symbiont in the aphid *Cinara cedri*. *J. Bacteriol.* **186**, 6626–6633.
- Goodson, M.S., Kojadinovic, M., Troll, J.V., Scheetz, T.E., Casavant, T.L., Soares, M.B., and McFall-Ngai, M.J. (2005). Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. *Appl. Environ. Microbiol.* **71**, 6934–6946.
- Groisman, E.A., and Ochman, H. (1996). Pathogenicity islands: Bacterial evolution in quantum leaps. *Cell* **87**, 791–794.
- Herold, S., Karch, H., and Schmidt, H. (2004). Shiga toxin-encoding bacteriophages—genomes in motion. *Int. J. Med. Microbiol.* **294**, 115–121.
- Hueck, C.J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.* **62**, 379–433.
- Hurst, G.D., and Jiggins, F.M. (2000). Male-killing bacteria in insects: Mechanisms, incidence, and implications. *Emerg. Infect. Dis.* **6**, 329–336.
- Hurst, G.D., and Werren, J.H. (2001). The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.* **2**, 597–606.
- Iturbe-Ormaetxe, I., Burke, G.R., Riegler, M., and O'Neill, S.L. (2005). Distribution, expression, and motif variability of ankyrin domain genes in *Wolbachia pipientis*. *J. Bacteriol.* **187**, 5136–5145.
- Kakeda, K., and Ishikawa, H. (1991). Molecular chaperon produced by an intracellular symbiont. *J. Biochem. (Tokyo)* **110**, 583–587.
- Koga, R., Tsuchida, T., and Fukatsu, T. (2003). Changing partners in an obligate symbiosis: A facultative endosymbiont can compensate for loss of the essential *Buchnera* in an aphid. *Proc. R. Soc. Lond. B. Biol. Sci.* **276**, 2543–2550.
- Koropatnick, T.A., Engle, J.T., Apicella, M.A., Stabb, E.V., Goldman, W.E., and McFall-Ngai, M.J. (2004). Microbial factor-mediated development in a host-bacterial mutualism. *Science* **306**, 1186–1188.
- Lai, C.Y., Baumann, L., and Baumann, P. (1994). Amplification of *trpEG*: Adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proc. Natl. Acad. Sci. USA* **91**, 3819–3823.
- Latorre, A., Gil, R., Silva, F.J., and Moya, A. (2005). Chromosomal stasis versus plasmid plasticity in aphid endosymbiont *Buchnera aphidicola*. *Heredity* **95**, 339–347.
- Lefevre, C., Charles, H., Vallier, A., Delobel, B., Farrell, B., and Heddi,

- A. (2004). Endosymbiont phylogenesis in the dryophthoridae weevils: Evidence for bacterial replacement. *Mol. Biol. Evol.* *21*, 965–973.
- Li, J., Ochman, H., Groisman, E.A., Boyd, E.F., Solomon, F., Nelson, K., and Selander, R.K. (1995). Relationship between evolutionary rate and cellular location among the Inv/Spa invasion proteins of *Salmonella enterica*. *Proc. Natl. Acad. Sci. USA* *92*, 7252–7256.
- Lo, N., Bandi, C., Watanabe, H., Nalepa, C., and Beninati, T. (2003). Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. *Mol. Biol. Evol.* *20*, 907–913.
- McGraw, E.A., and O'Neill, S.L. (2004). *Wolbachia pipientis*: Intracellular infection and pathogenesis in *Drosophila*. *Curr. Opin. Microbiol.* *7*, 7–10.
- Mahillon, J., and Chandler, M. (1998). Insertion sequences. *J. Bacteriol.* *180*, 5306–5312.
- Masui, S., Kamoda, S., Sasaki, T., and Ishikawa, H. (2000). Distribution and evolution of bacteriophage WO in *Wolbachia*, the endosymbiont causing sexual alterations in arthropods. *J. Mol. Evol.* *51*, 491–497.
- Masui, S., Kuroiwa, H., Sasaki, T., Inui, M., Kuroiwa, T., and Ishikawa, H. (2001). Bacteriophage WO and virus-like particles in *Wolbachia*, an endosymbiont of arthropods. *Biochem. Biophys. Res. Commun.* *283*, 1099–1104.
- Mavingui, P., Tran Van, V., Labeyrie, E., Rances, E., Vavre, F., and Simonet, P. (2005). Efficient procedure for purification of obligate intracellular *Wolbachia pipientis* and representative amplification of its genome by multiple-displacement amplification. *Appl. Environ. Microbiol.* *71*, 6910–6917.
- Macdonald, T.T., and Monteleone, G. (2005). Immunity, inflammation, and allergy in the gut. *Science* *307*, 1920–1925.
- McFall-Ngai, M. (2002). Unseen forces: The influence of bacteria on animal development. *Dev. Biol.* *242*, 1–14.
- Moran, N.A. (1996). Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA* *93*, 2873–2878.
- Moran, N.A., and Plague, G.R. (2004). Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* *14*, 627–633.
- Moran, N.A., Plague, G., Sandström, J., and Wilcox, J. (2003). A genomic perspective on nutrient-provisioning by bacterial symbionts of insects. *Proc. Natl. Acad. Sci. USA* *100*, 14543–14548.
- Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E., and Ochman, H. (2005a). The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. *Proc. Natl. Acad. Sci. USA* *102*, 16919–16926.
- Moran, N.A., Dunbar, H.E., and Wilcox, J.L. (2005b). Regulation of transcription in a reduced bacterial genome: Nutrient-provisioning genes of the obligate symbiont, *Buchnera aphidicola*. *J. Bacteriol.* *187*, 4229–4237.
- Moran, N.A., Russell, J.A., Koga, R., and Fukatsu, T. (2005c). Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Appl. Environ. Microbiol.* *71*, 3302–3310.
- Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., and Peng, Z.Y. (2004). The ankyrin repeat as molecular architecture for protein recognition. *Protein Sci.* *13*, 1435–1448.
- Mota, L.J., Sorg, I., and Cornelis, G.R. (2005). Type III secretion: The bacteria-eukaryotic cell express. *FEMS Microbiol. Lett.* *252*, 1–10.
- Munson, M.A., Baumann, P., Clark, M.A., Baumann, L., Moran, N.A., Voegtlin, D.J., and Campbell, B.C. (1991). Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *J. Bacteriol.* *173*, 6321–6324.
- Nakabachi, A., Shigenobu, S., Sakazume, N., Shiraki, T., Hayashizaki, Y., Carninci, P., Ishikawa, H., Kudo, T., and Fukatsu, T. (2005). Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. *Proc. Natl. Acad. Sci. USA* *102*, 5477–5482.
- Oliver, K.M., Russell, J.A., Moran, N.A., and Hunter, M.S. (2003). Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl. Acad. Sci. USA* *100*, 1803–1807.
- Oliver, K.M., Moran, N.A., and Hunter, M.S. (2005). Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proc. Natl. Acad. Sci. USA* *102*, 12795–12800.
- Oliver, K.M., Moran, N.A., and Hunter, M. (2006). Costs and benefits of a superinfection of facultative symbionts in aphids. *Proc. R. Soc. Lond. B. Biol. Sci.* *273*, 1273–1280.
- O'Neill, S.L., Pettigrew, M.M., Sinkins, S.P., Braig, H.R., Andreadis, T.G., and Tesh, R.B. (1997). *In vitro* cultivation of *Wolbachia pipientis* in an *Aedes albopictus* cell line. *Insect Mol. Biol.* *6*, 33–39.
- Piel, J. (2002). A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc. Natl. Acad. Sci. USA* *99*, 14002–14007.
- Pugsley, A.P., Francetic, O., Driessen, A.J., and de Lorenzo, V. (2004). Getting out: Protein traffic in prokaryotes. *Mol. Microbiol.* *52*, 3–11.
- Russell, J.A., and Moran, N.A. (2005). Horizontal transfer of bacterial symbionts: Heritability and fitness effects in a novel aphid host. *Appl. Environ. Microbiol.* *71*, 7987–7994.
- Sanogo, Y.O., and Dobson, S.L. (2006). WO bacteriophage transcription in *Wolbachia*-infected *Culex pipiens*. *Insect Biochem. Mol. Biol.* *36*, 80–85.
- Scarborough, C.L., Ferrari, J., and Godfray, H.C. (2005). Aphid protected from pathogen by endosymbiont. *Science* *310*, 1781.
- Sexton, J.A., and Vogel, J.P. (2002). Type IVB secretion by intracellular pathogens. *Traffic* *3*, 178–185.
- Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000). Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. *APS. Nature* *407*, 81–86.
- Sinkins, S.P., Walker, T., Lynd, A.R., Steven, A.R., Makepeace, B.L., Godfray, H.C., and Parkhill, J. (2005). *Wolbachia* variability and host effects on crossing type in *Culex* mosquitoes. *Nature* *436*, 257–260.
- Stouthamer, R., Breeuwer, J.A., and Hurst, G.D. (1999). *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* *53*, 71–102.
- Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A., and Andersson, S.G. (2002). 50 million years of genomic stasis in endosymbiotic bacteria. *Science* *296*, 2376–2379.
- Thelestam, M., and Frisan, T. (2004). Cytolethal distending toxins. *Rev. Physiol. Biochem. Pharmacol.* *152*, 111–133.
- Toh, H., Weiss, B.L., Perkin, S.A., Yamashita, A., Oshima, K., Hattori, M., and Aksoy, S. (2006). Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res.* *16*, 149–156.
- Tram, U., Ferree, P.M., and Sullivan, W. (2003). Identification of *Wolbachia*–host interacting factors through cytological analysis. *Microbes Infect.* *5*, 999–1011.
- van der Wilk, F., Dulleman, A.M., Verbeek, M., and van den Heuvel, J.F. (1999). Isolation and characterization of APSE-1, a bacteriophage infecting the secondary endosymbiont of *Acyrtosiphon pisum*. *Virology* *262*, 104–113.
- van Ham, R.C., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernandez, J.M., Jimenez, L., Postigo, M., and Silva, F.J. (2003). Reductive genome evolution in *Buchnera aphidicola*. *Proc. Natl. Acad. Sci. USA* *100*, 581–586.

- Veneti, Z., Clark, M.E., Karr, T.L., Savakis, C., and Bourtzis, K. (2004). Heads or tails: Host-parasite interactions in the *Drosophila-Wolbachia* system. *Appl. Environ. Microbiol.* **70**, 5366–5372.
- Veneti, Z., Bentley, J.K., Koana, T., Braig, H.R., and Hurst, G.D. (2005). A functional dosage compensation complex required for male killing in *Drosophila*. *Science* **307**, 1461–1463.
- Wagner, P.L., Livny, J., Neely, M.N., Acheson, D.W., Friedman, D.I., and Waldor, M.K. (2002). Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Mol. Microbiol.* **44**, 957–970.
- Welburn, S.C., Maudlin, I., and Ellis, D.S. (1987). In vitro cultivation of rickettsia-like organisms from *Glossina* spp. *Ann. Trop. Med. Parasitol.* **81**, 331–335.
- Wernegreen, J.J., and Moran, N.A. (1999). Evidence for genetic drift in endosymbionts: Analyses of protein coding genes. *Mol. Biol. Evol.* **16**, 83–97.
- Wernegreen, J.J., and Moran, N.A. (2001). Vertical transmission of biosynthetic plasmids in aphid endosymbionts (*Buchnera*). *J. Bacteriol.* **183**, 785–790.
- Werren, J.H., Zhang, W., and Guo, L.R. (1995). Evolution and phylogeny of *Wolbachia*: Reproductive parasites of arthropods. *Proc. R. Soc. Lond. B. Biol. Sci.* **261**, 55–63.
- Wilcox, J.L., Dunbar, H.E., Wolfinger, R.D., and Moran, N.A. (2003). Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol. Microbiol.* **48**, 1491–1500.
- Wolschin, F., Holldobler, B., Gross, R., and Zientz, E. (2004). Replication of the endosymbiotic bacterium *Blochmannia floridanus* is correlated with the developmental and reproductive stages of its ant host. *Appl. Environ. Microbiol.* **70**, 4096–4102.
- Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: Viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* **64**, 69–114.
- Wu, M., Sun, L.V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J.C., McGraw, E.A., Martin, W., Esser, C., Ahmadinejad, N., et al. (2004). Phylogenomics of the reproductive parasite *Wolbachia pipiensis* wMel: A streamlined genome overrun by mobile genetic elements. *PLoS Biol.* **2**, E69.
- Wu, D., Daugherty, S.C., Van Aken, S.E., Pai, G.H., Watkins, K.L., Khouri, H., Tallon, L.J., Zaborsky, J.M., Dunbar, H.E., Tran, P.L., et al. (2006). Metabolic complementarity and genomics of the dual symbiosis of sharpshooters. *PLoS Biol.* **4**, E188.
- Zientz, E., Dandekar, T., and Gross, R. (2004). Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiol. Mol. Biol. Rev.* **68**, 745–777.