Partnering With a Pest: Genomes of Hemlock Woolly Adelgid Symbionts Reveal Atypical Nutritional Provisioning Patterns in Dual-Obligate Bacteria

Kathryn M. Weglarz[1], Nathan P. Havill[2], Gaelen R. Burke[3,*], and Carol D. von Dohlen[1,*]

[1]Department of Biology, Utah State University

[2]USDA Forest Service, Northern Research Station, Hamden, Connecticut

[3]Department of Entomology, University of Georgia

*Corresponding authors: E-mails: grburke@uga.edu; carol.vondohlen@usu.edu.

Accepted: May 29, 2018

Data deposition: Genome of 'Ca. Pseudomonas adelgestsugas' is GenBank accession CP026512 and 'Ca. Annandia adelgestsuga' is CP026513).

Abstract

Nutritional bacterial symbionts enhance the diets of sap-feeding insects with amino acids and vitamins missing from their diets. In many lineages, an ancestral senior symbiont is joined by a younger junior symbiont. To date, an emergent pattern is that senior symbionts supply a majority of amino acids, and junior symbionts supply a minority. Similar to other hemipterans, adelgids harbor obligate symbionts, but have higher diversity of bacterial associates, suggesting a history of symbiont turnover. The metabolic roles of dual symbionts in adelgids and their contributions to the consortium are largely unexplored. Here, we investigate the symbionts of *Adelges tsugae*, the hemlock woolly adelgid (HWA), an invasive species introduced from Japan to the eastern United States, where it kills hemlock trees. The response of hemlocks to HWA feeding has aspects of a defensive reaction against pathogens, and some have speculated that symbionts may be involved. We sequenced the genomes of "*Ca.* Annandia adelgestsuga" and "*Ca.* Pseudomonas adelgestsugas" symbionts to detail their metabolic capabilities, infer ages of relationship, and search for effectors of plant defenses. We also tested the relationship of "*Ca.* Annandia" to symbionts of other insects. We find that both symbionts provide nutrients, but in more balanced proportions than dual symbionts of other hemipterans. The lesser contributions of the senior "*Ca.* Annandia" support our hypothesis for symbiont replacements in adelgids. Phylogenomic results were ambiguous regarding the position of "*Ca.* Annandia". We found no obvious effectors of plant defenses related to insect virulence, but hypothetical proteins in symbionts are unknown players.

Key words: *Adelges tsugae*, *Adelgidae*, "*Ca.* Annandia adelgestuga," "*Ca.* Pseudomonas adelgestsugas," hypersensitive response, nutritional endosymbiont.

Introduction

Heritable bacterial symbionts likely facilitated the exploitation of nitrogen-poor plant sap for hemipteran insects (Buchner 1965; Moran 2001). Auchenorrhynchan and sternorrhynchan hemipterans typically maintain their obligate bacterial partners within specialized organs (bacteriomes), from which they are transferred to eggs or embryos (Buchner 1965; Moran 2001). Numerous studies have detailed the roles of obligate mutualists in supplementing their hosts' diets with essential products that insects are unable to synthesize alone (Shigenobu et al. 2000; Gunduz and Douglas 2009). In certain lineages (aphids, whiteflies), a universal symbiont typically contains biosynthetic pathways for all 10 essential amino acids (EAA), as well as vitamins and cofactors (Shigenobu et al. 2000; Santos-Garcia et al. 2012). In other insect lineages (psyllids, some scales, cicadas and various hoppers), an ancestral ("senior") symbiont is typically joined by a younger ("junior") symbiont (Buchner 1965; Spaulding and von Dohlen 2001; Gruwell et al. 2010; Rosenblueth et al. 2012; Bennett and Moran 2013). Such dual symbionts cooperate in EAA synthesis, sometimes even within pathways (McCutcheon and Moran 2007, 2011; McCutcheon et al. 2009a;

McCutcheon and von Dohlen 2011; Husník et al. 2013; Koga and Moran 2014; Luan et al. 2015; Rao et al. 2015). Thus far, a universal pattern in role partitioning has emerged: The ancestral, senior symbiont performs the major role (typically, seven to nine EAA), and the junior symbiont supplies the minority remainder.

Similar to aphids and other sternorrhynchans, adelgids (Aphidoidea: Adelgidae) harbor obligate symbionts in specialized bacteriomes. Adelgids are unusual among sap-sucking hemipterans, however, in their diversity of bacterial symbionts within a comparatively small and recent crown-group lineage. Extant adelgids (*Adelges* and *Pineus* spp.) comprise ~70 species feeding solely on conifers, and share a common ancestor likely dating to the Paleogene (Havill and Foottit 2007; Havill et al. 2007) [but possibly more recently (Havill et al. 2016)]. Most adelgids possess host-alternating life cycles, in which they feed sequentially on spruce (*Picea* spp., where a gall is formed) and one of the five other conifer genera (Havill and Foottit 2007). Adelgid species diversity falls into five major lineages based on these alternate-conifer hosts (Havill et al. 2007). Although the existence of bacterial symbionts in adelgids was known for many years (Profft 1937; Buchner 1965; Steffan 1976), only recently were symbionts characterized through molecular studies. Research to date has detected eight different obligate symbionts and one facultative symbiont (Toenshoff, Gruber et al. 2012; Toenshoff, Penz et al. 2012; von Dohlen et al. 2013, 2017; Toenshoff et al. 2014). Obligate symbionts are organized into pairs of symbionts within an adelgid species; these pairs are unique to each of the five adelgid lineages (von Dohlen et al. 2017). The diversity of symbionts in Adelgidae suggests a history of repeated gains, losses, and replacements (Toenshoff, Gruber et al. 2012; Toenshoff et al. 2014; von Dohlen et al. 2017). Unlike other Hemiptera, no single, ancient symbiont is maintained universally in all adelgid taxa. However, von Dohlen et al. (2017) hypothesized that a symbiont common to the hemlock and pine lineages represents the original symbiont of Adelgidae.

Hemlock woolly adelgid (*A. tsugae* Annand) (HWA) is a complex of sexually reproducing and obligately parthenogenetic lineages dating to the Pleistocene or earlier, with ranges native to Asia and western North America (Havill et al. 2006, 2007, 2016). Populations from Japan were introduced to eastern North America in the last century, and have decimated native hemlock species there (Havill et al. 2014). In addition to retarding growth, feeding by HWA induces a systemic hypersensitive response (HR) in hemlock trees (Radville et al. 2011). The HR is a general defensive response by plants in reaction to infestation by herbivores and microbes, which induces localized cell death (Fernandes 1990; Heath 2000). Hemlocks appear to mount a systemic HR response, leading to rapid host-tree mortality (Young et al. 1995; Radville et al. 2011). What induces the systemic effect on hemlocks by HWA is unknown, but some have suggested the HR is triggered by toxic constituents in HWA saliva, possibly originating from bacterial symbionts (Radville et al. 2011; Pezet et al. 2013; von Dohlen et al. 2013).

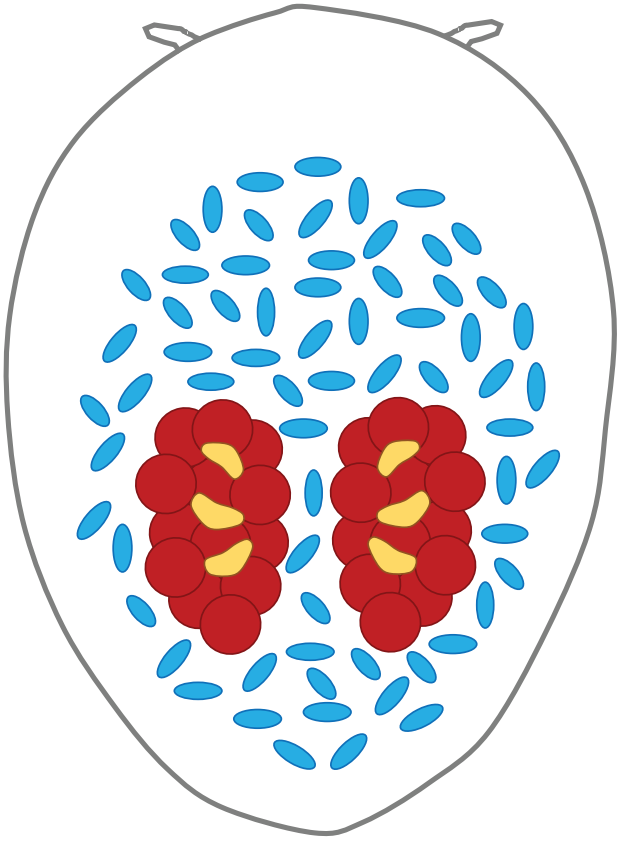

Fig. 1.—Schematic of HWA immature showing locations of symbionts. Large red circles are bacteriocytes containing cells of the obligate "*Ca*. Annandia" symbiont, grouped into paired bacteriomes. Blue ovals represent cells of the obligate "*Ca*. Pseudomonas" symbiont free-living in the hemocoel. Yellow shapes are portions of central bacteriocytes containing the facultative "*Ca*. Serratia symbiotica" bacteria. "*Ca*. Pseudomonas" cells are not drawn to scale. (See von Dohlen et al. (2013) for more information.)

As in all adelgids, HWA harbor dual-obligate symbionts. HWA are unique, however, in housing one symbiont in the bacteriome and the second symbiont in the body cavity; cells of the latter symbiont have never been detected within the bacteriome (Shields and Hirth 2005; von Dohlen et al. 2013, 2017) (fig. 1). Both symbionts are transmitted to developing eggs within the mother's body (unlike aphids, adelgids lay eggs in all generations) (von Dohlen et al. 2013). The bacteriome resident, "*Ca*. Annandia adelgestsuga," is shared by *Pineus* species, and is hypothesized to be the original, ancestral symbiont of Adelgidae (Toenshoff et al. 2014; von Dohlen et al. 2017). "*Ca*. Annandia" is most closely related to *Buchnera aphidicola* in aphids, symbionts of stinkbugs, and a symbiont of planthoppers, but its sister relationship is not well supported on the basis of ribosomal DNA sequences

(von Dohlen et al. 2013; Toenshoff et al. 2014). The hemocoel symbiont, "*Ca*. Pseudomonas adelgestsugas," belongs to a bacterial lineage that includes several plant pathogens capable of eliciting the HR (Nimchuk et al. 2003; Cui et al. 2005). This genus also includes an insect-associated species that was implicated in manipulating defensive responses in the host plant (Chung et al. 2013), as well as a toxin-producing defensive symbiont of beetles (Kellner and Dettner 1996; Kellner 2002; Piel 2002). Certain populations of HWA, namely, the introduced population in eastern North America and its source Japanese population, also harbor a facultative symbiont *Serratia symbiotica* (von Dohlen et al. 2013).

Although the role of symbionts in adelgids is presumed to encompass nutritional functions, their metabolic contributions to the consortium are essentially unexplored, except for an 85 kb fragment from a single junior symbiont (Toenshoff, Penz, et al. 2012). In this study we investigate the functional roles and evolution of the obligate symbionts of adelgids, focusing on HWA. The primary goal of this study was to describe the metabolic contributions of "*Ca*. A. adelgestsuga" and "*Ca*. P. adelgestsugas" to the consortium, as inferred from their genomes. We determined whether the genomic composition of symbionts exhibit characteristics of long-term, obligate associations, and whether symbionts provide nutritional functions similar to other sap-feeding insects—either alone, or sharing pathways with each other and/or the host. We further investigated whether "*Ca*. P. adelgestsugas" might perform protective or defensive functions for its host. We initiated this inquiry by sequencing and characterizing the complete genomes of both symbionts and reconstructing the putative metabolic capabilities of symbionts and putative contributions from HWA. Two additional goals of the study were to determine whether features of the "*Ca*. Annandia adelgestsuga" genome support the idea that it is the original, ancestral symbiont of Adelgidae, and whether "*Ca*. Annandia" might share a most-recent common ancestor with *Buchnera* of Aphididae.

Materials and Methods

Insect Samples, DNA Preparation, and Sequencing

Adelgid samples (adults and egg masses) were collected in New Haven County, Connecticut, USA from eastern hemlock (*Tsuga canadensis*) on 22 April 2013 by Nathan Havill (voucher ID #13-050). Genomic DNA was extracted from eggs (masses from two females), to maximize the amount of symbiont DNA in the extraction using the High Pure PCR template kit (Roche Diagnostics, Indianapolis), and treated with DNase–free RNAse (Roche Diagnostics). DNA concentration was quantified with a Qubit fluorometer. Paired-end Illumina library construction and sequencing was performed at the Yale Center for Genome Analysis (New Haven, CT) a single lane of an Illumina HiSeq 2500 with 150 nt paired-end reads.

Genome Assembly and Annotation

Quality assessment of raw Illumina reads was performed with the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html; last accessed March 15, 2018). Reads were filtered such that a read was eliminated if it had fewer than 90% of bases with a Phred quality score of 30 or more. The resulting quality-filtered data set consisted of 120,202,263 reads totaling more than 18 billion nts of sequence. The data were assembled *de novo* using SPAdes v.3.7.0 with k-mers 21, 33, 55, 77, 99, 127 and the "–meta" flag (Bankevich et al. 2012). We identified symbiont-derived contigs using coverage and GC content statistics in addition to BLAST against the coding portion of genomes for *Pseudomonas aeruginosa* (NC_002516.2) and *B. aphidicola* (NC_002528.1). Closure of the single "*Ca*. Pseudomonas adelgestsugas" (CP026512) scaffold into a circular bacterial chromosome was achieved via alignment of terminal overlapping regions of ~200 bp. The order and orientation of six "*Ca*. Annandia adelgestsuga" (CP026513) scaffolds was determined using PCR with primers specific for each scaffold. Due to the high AT content and presence of low complexity repetitive sequences in the "*Ca*. Annandia adelgestsuga" genome, sequencing across these gaps was not feasible and the genome was closed by inserting "Ns" to indicate gaps of unknown size.

Hemlock woolly adelgid symbiont genome size, GC content, and coding capacity was compared with other symbiont and free-living bacterial genomes. Genome statistics were downloaded for all bacterial "reference genomes" from NCBI in June 2017 and additional bacteria for the comparisons were added to the data set. Symbiont designation was determined by consulting the literature. Plots were created in R using ggplot2 (Wickham 2009).

The origin of replication for both genomes was determined using GC skew with GenSkew v.1.0. The "*Ca*. Annandia adelgestsuga" genome had weak GC skew overall, thus the origin was designated at a noncoding site with the strongest signal. Genomes were initially annotated with PROKKA v1.11. The annotation was manually checked for genes interrupted by ambiguous bases. Any genes annotated as hypothetical proteins were searched against the nr database with BLAST, and if a function could be assigned, the annotation was adjusted manually. We designated all fragmented coding sequences (with length less than 80% of full-length homologs belonging to other species in the S-PROT database) as pseudogenes using tblastn as previously described (Lerat and Ochman 2004; Burke and Moran 2011). Both genomes were checked for insertion sequence elements by uploading the fasta nucleotide files to the ISSaga2 web-based interface (Varani et al. 2011). Amino acid and vitamin biosynthesis pathways were reconstructed using the BioCyc, EcoCyc,

and KAAS databases (Moriya et al. 2007; Keseler et al. 2013; Caspi et al. 2016). Lists of genes for vital functions were compiled, and presence and absence were compared with other obligate symbionts. Clusters of orthologous groups were determined using the online eggNOG-mapper tool (DIAMOND mapping mode and default choices for other settings) to provide functional annotation of protein sequences and categorize sequences into 17 functional categories (Tatusov et al. 2000; Huerta-Cepas et al. 2016).

Synteny Analyses

Synteny was examined between each symbiont and one or two related taxa. "*Ca*. Annandia adelgestsuga" was compared with "*Ca.* Ishikawaella capsulata" (NZ_AP010872.1) and *B. aphidicola* APS (NC_002528.1), and "*Ca.* Pseudomonas adelgestsugas" was compared with *P. aeruginosa* PAO1 (NC_002516.2). An all-against-all BLAST of amino acid sequences (*e*-value cut-off = 1*e*−10) served as input for MCScanX to identify collinear blocks between genomes involving greater than five genes (parameters: gap_penalty = 5) (Wang et al. 2012). Synteny plots were generated using VGSC 2.0 (Xu et al. 2016).

Phylogenomic Analyses

Reconstructing the evolutionary relationships between free-living bacteria and bacterial symbionts is notoriously difficult due to their high rates of substitution and low G + C content, resulting in long-branch attraction (Husník et al. 2011). Sophisticated phylogenetic methods must be used to overcome these issues, even when large data sets are available (Husník et al. 2011). We use the two best-performing approaches established by Husník et al. (2011), using Phylobayes on an amino acid data set, and nhPhyML on a nucleotide data set (Boussau and Gouy 2006; Lartillot et al. 2013). Forty-five single-copy orthologous genes derived from 50 γ-proteobacteria taxa (including 14 symbiont taxa) from Husník et al. were combined with data from four endosymbiont genomes, "*Ca*. Annandia adelgestsuga" (CP026513), "*Ca.* Moranella endobia" (NC_015735.1), "*Ca.* Buchnera aphidicola" str Sc (NZ_CP011299.1), and "*Ca.* Trabutinella endobia" (NZ_LT594522.1). Amino acid sequences from each gene were aligned in SeaView version 4 and aligned by the MAFFT version 7L-INS-i algorithm, followed by trimming with GBlocks version 0.91b with the following parameters: Minimum number of sequences for a conserved position: 28; minimum number of sequences for a flanking position: 46; maximum number of contiguous nonconserved positions: 8; minimum length of a block: 5; allowed gap positions: with half. These trimmed alignments were concatenated in SeaView, and represented an alignment of 13,063 amino acid positions. The amino acid data set was recoded using the dayhoff6 scheme in Phylobayes MPI v1.3b (Lartillot et al. 2013). Phylobayes was run on this data set with a CAT + GTR model and two chains for over 4000 generations until convergence. Markov Chain Monte Carlo convergence was evaluated with Tracer v1.6 and Phylobayes bpcomp (Rambaut et al. 2014). For phylogenetic reconstruction with nhPhyML, the concatenated amino acid data set was back-translated to nucleotide data and the third codons were removed using SeaView v4 (Gouy et al. 2010). GBlocks v. 0.91b was used to refine the alignment with the default parameters in SeaView v4 to a total of 27,600 nucleotide positions. nhPhyML was used to apply a nonhomogeneous nonstationary model of sequence evolution to the nucleotide data set using a best starting tree with taxa placed according to the results from the Phylobayes analysis.

Results

Genomes of "*Ca*. Annandia adelgestsuga" and "*Ca*. Pseudomonas adelgestsugas"

The "*Ca*. Annandia adelgestsuga" (referred to hereinafter as *Annandia*) genome is 334,746 bp in length with G + C content of 17.8%, and the genome size of "*Candidatus P. adelgestsugas*" (referred herein as *Pseudomonas* or *P. adelgestsugas*) is 1,835, 598 bp, with G + C content of 39.5% (fig. 2*A*). Bacterial species typically feature a tight correlation between genome size and the number of protein-coding genes (fig. 2*B*). Although *Annandia* and *P. adelgestsugas* have genome sizes that fall within the size range representative of other obligate bacterial symbionts of insects, the *Pseudomonas* genome has a larger genome than expected given the number of protein-coding genes in the genome (coding density of 45.5%, compared with 83.9% for *Annandia*).

The *Annandia* genome contains 343 genes, of which 285 are protein-coding sequences (CDS) that could be assigned to known genes based on homology to other bacteria, and eight encode hypothetical proteins of unknown function. The genome contains 29 tRNAs, 25 predicted pseudogenes, and a single ribosomal RNA operon; however, the 16S and 23S genes are unusually long, at 2,149 and 2,974 bp, respectively (von Dohlen et al. 2013). The *Pseudomonas* genome contains 1,107 genes comprising 944 CDS of predicted known identity, 198 hypothetical proteins, a single ribosomal RNA operon, 38 tRNAs, and 31 pseudogenes. No insertion sequence elements were identified in either genome, however, one false positive insertion sequence ORF was identified in the *Annandia* genome.

Phylogenetic relationships of *A. adelgestsuga*

Phylogenomic analysis using 45 genes from representative Enterobacteriaceae taxa placed *Annandia* in a well-supported clade with several *Buchnera* lineages and *Ishikawaella capsulata* (supplementary fig. S1, Supplementary Material online). However, the branching

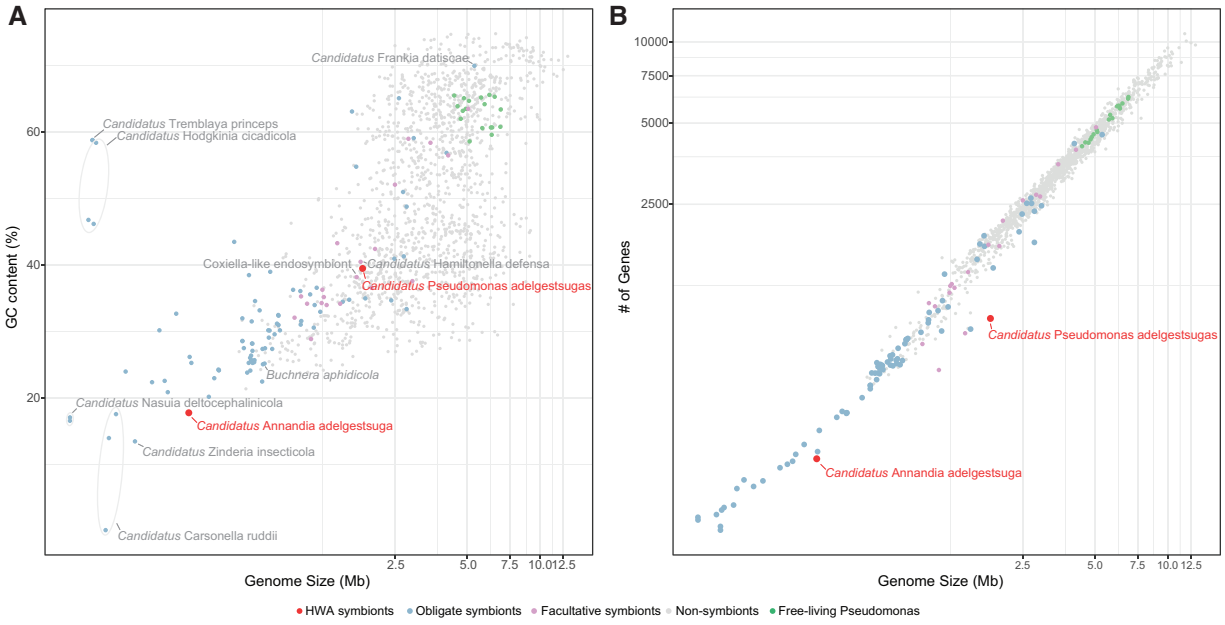

Fig. 2.—Relationships between (*A*) genome size and GC content and (*B*) genome size and total gene number in HWA symbionts and other representative bacterial genomes.

pattern between these symbionts was poorly supported, and the *Annandia* branch is notably long. In addition, there was very little conservation of synteny between *Annandia* and relatives *Ishikawaella capsulata* and *Buchnera* str. APS (fig. 3). *Annandia* shares 13 syntenous blocks with *Buchnera* and 14 blocks with *Ishikawaella*, but these blocks are small, containing 5–29 genes each. Of these conserved blocks, only eight are shared across all three genomes. In contrast, a similar comparison between *P. adelgestsuga* and *P. aeruginosa* PAO1, which are distantly related within the *Pseudomonas* group (von Dohlen et al. 2013), indicates some conservation of synteny, particularly at the 5′ end of the symbiont genome, with a total of 34 conserved blocks of five genes or more.

Nutritional Capabilities

Both symbiont genomes possess genes for the production of EAAs and intermediates, as well as some nonessential amino acids (NAAs) (fig. 4; supplementary fig. S2, Supplementary Material online). *Annandia* retains 57 intact genes involved in EAA biosynthesis, whereas the *Pseudomonas* genome has 66 genes. With these genes, both *Annandia* and *Pseudomonas* are each able to produce lysine and threonine and the intermediate chorismate. As in some other consortia (Wilson et al. 2010; Hansen and Moran 2011; McCutcheon and von Dohlen 2011), symbionts of *A. tsugae* presumably rely on the insect host to provide certain enzymes to complete several EAA pathways. These include branched-chain amino acid aminotransferase, cystathionine gamma-lyase (CGL), ornithine aminotransferase, and possibly aspartate aminotransferase (AAT) (fig. 4); transcripts from these genes have been identified in an unpublished *A. tsugae* transcriptome (NCBI BioProject PRJNA242203). In addition to lysine and threonine, *Annandia* could produce arginine, valine, leucine, phenylalanine, and isoleucine with input from host genes. In addition to lysine and threonine, *Pseudomonas* encodes pathways for synthesizing histidine, phenylalanine, and methionine, the latter with host contribution of CGL. The tryptophan biosynthetic pathway has been divided between the symbionts, such that *Annandia* performs the first two rate-limiting steps with *trpEG* and *Pseudomonas* performs the remainder with *trpDFCAB*. *Annandia* contains a redundant *trpC* gene, which encodes the bifunctional fused indole-3-glycerol phosphate synthase/phosphoribosylanthranilate isomerase. This performs both reactions EC: 5.3.1.24 and EC: 4.1.1.48. In *Pseudomonas* bacteria, EC: 5.3.1.24 and EC: 4.1.1.48 are catalyzed separately by the monofunctional enzymes phosphoribosylanthranilate isomerase (TrpF) and indole-3-glycerol phosphate synthase (TrpC), respectively. Both symbionts show partial redundancy in other EAA biosynthetic pathways. *Pseudomonas* retains an incomplete set of genes for the biosynthesis of arginine (produced in full by *Annandia* with a host gene), and *Annandia* lacks one gene in the phenylalanine biosynthesis pathway (produced in full by *Pseudomonas*), although a host gene AAT may be able to compensate for this loss. Whereas *Pseudomonas* retains vestiges of the valine, leucine, and isoleucine pathways, many of the genes required for these pathways are pseudogenes. Finally, both symbionts retain a partial set of genes to produce the NAA. Both *Annandia* and *Pseudomonas* can make alanine and glycine, *Pseudomonas* can produce aspartic acid,

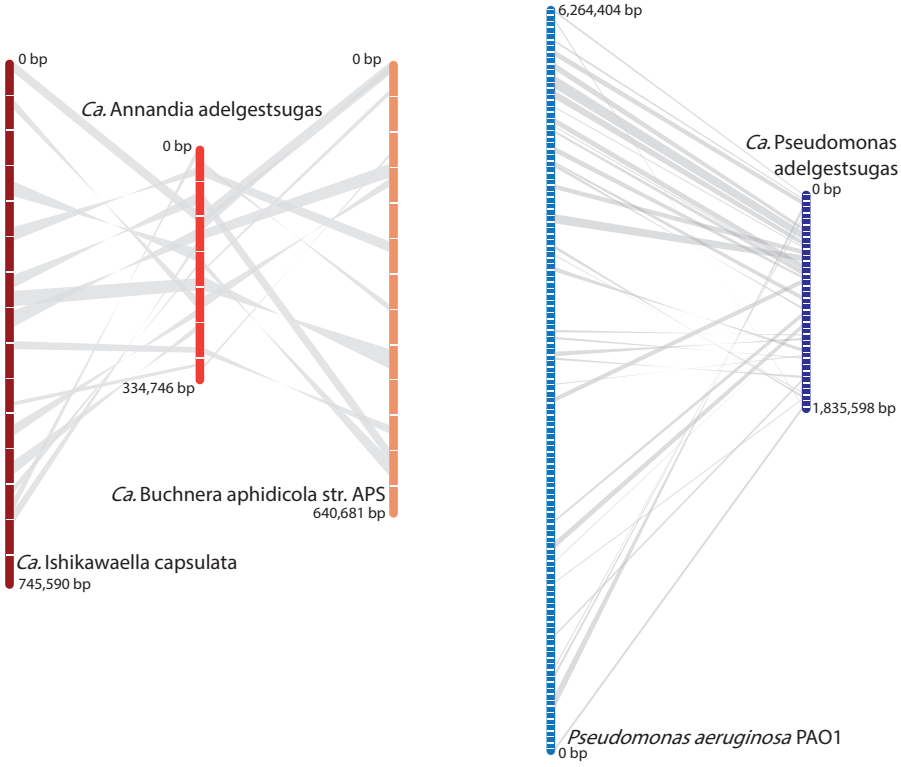

Fig. 3.—Conserved syntenic blocks and rearrangements between HWA symbiont genomes and relatives. Chromosomes are marked every 50 kb and inverted when appropriate for clarity. Ribbons between chromosomes indicate position and length of collinear syntenic blocks that are composed of a minimum of five genes. (*A*) Comparison of "*Ca*. Annandia adelgestsuga" to "*Ca*. Buchnera aphidicola" str. APS and "*Ca*. Ishikawaella *capsulata*". (*B*) Comparison of "*Ca*. Pseudomonas adelgestsugas" and *Pseudomonas aeruginosa* PAO1.

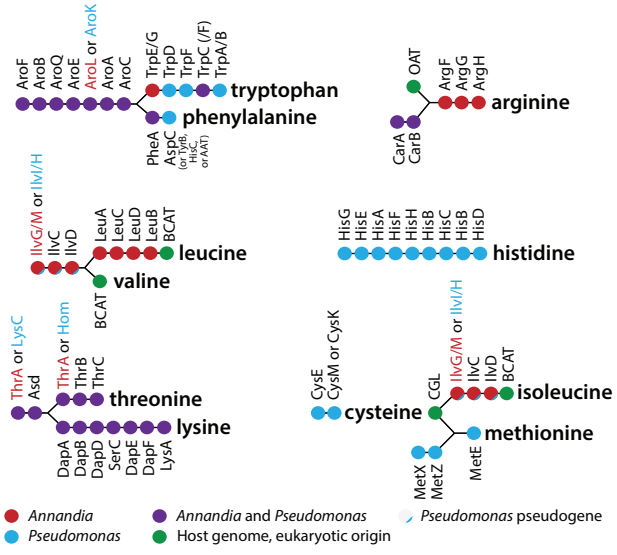

Fig. 4.—Reconstruction of essential amino acid pathways based on genes present in "*Ca*. Annandia adelgestsuga," "*Ca*. Pseudomonas adelgestsugas," and an unpublished transcriptome of HWA.

cysteine, and glutamine, and tyrosine could be synthesized cooperatively (supplementary table S1, Supplementary Material online). In further support of amino acid production, *Pseudomonas* retains all genes for a pathway of assimilatory sulfate reduction (sulfur assimilation).

Beyond amino acid production and support, both symbionts appear to produce vitamins and cofactors, although *Annandia*'s capacity here is limited. Its genome contains genes only for lipoate (an essential cofactor) and 5-phospho-α-D-ribosyl 1-pyrophosphate (PRPP) biosynthesis (supplementary table S2, Supplementary Material online). *Pseudomonas* retains 53 genes related to vitamin and coenzyme biosynthesis; with complete pathways for pyridoxal 5′ phosphate (vitamin B6), heme compounds, PRPP, glutathione, and lipoate, and partial pathways for biotin, pyridine, folate, riboflavin, thiamine, ubiquinone (coenzyme Q), and coenzyme A biosynthesis (supplementary table S2, Supplementary Material online). *Pseudomonas* encodes *gdhB*, and so is putatively able to recycle nitrogen from ammonia.

Annandia and *Pseudomonas* differ substantially in their inventory of small-molecule and macromolecule transporters. They share only one transporter in common, bnrQ, a secondary carrier for branched-chain amino acids. Beyond this, *Annandia* possesses few others; annotations included galP (transport of monosaccharides), yhgN (putative antibiotic transporter), and yadH (integral membrane subunit of a putative polyketide drug exporter). *Pseudomonas* possesses over

30 CDS annotated as transporters. Most are categorized as ATP-binding cassette transporters. These include several amino acid transporters: aotPQ of the aot operon for transport of arginine and ornithine; sdaC, a secondary carrier for serine and threonine; and hisJ, part of the histidine permease transporter. Others include transporters for lipopolysaccharides (LPS), metal ions, antibiotics, toxins, and putrescine. *Pseudomonas* may contain a functional Sec protein export system (Beckwith 2013), as most of the sec genes (secA, secYEG, secDF) are retained except for secB, which could be functionally replaced by another chaperone (Zientz et al. 2004).

Core Cellular Processes

Functional assignments of genes, as determined by COG (clusters of orthologous groups) categorization (Tatusov et al. 2001), show *Annandia* is similar to primary/senior symbionts such as *Sulcia muelleri* in Auchenorrhyncha, *Buchnera* from the aphid *Cinara cedri* (BCc), and *Portiera aleyrodidarum* in whiteflies (supplementary fig. S2, Supplementary Material online). *Annandia* devotes the majority of protein-coding genes to translation-related functions (category J, 37%), amino acid biosynthesis functions (category E, 14%), and energy-related functions (category C, 12.5%) (supplementary fig. S2, Supplementary Material online). Similar to BCc, it appears incapable of nucleotide metabolism and transport (category F). The suite of core genes involved in central cellular processes, as compiled in Moran and Bennett (2014) and Bennett *et al.* (2014), is also characteristic of long-term, obligate symbionts (figs. 5 and 6). *Annandia* has lost homologs of several genes important for DNA replication and repair, including the DNA polymerase holoenzyme gene, *holA* (which has been pseudogenized) (fig. 5*A*). It has lost most genes for DNA replication initiation (retaining only *dnaB*, a replicative DNA helicase), and all core genes involved in cell division (fig. 6*A*). With respect to protein synthesis, *Annandia* retains most genes corresponding to tRNA synthetases and approximately 80% of core genes involved in transcription and translation (retaining *rpoABD* of the core RNA polymerase but losing *rpoC*) (fig. 5*A*). It retains 51/54 ribosomal proteins (fig. 5*B*). In the category of protein folding and stability, it retains *groL*, *groS*, *dnaJ*, and *dnaK*, but has lost *grpE* (fig. 5*B*). In energy production and respiration, *Annandia* contains a complete set of ATP synthase genes and all but one of NADH dehydrogenase and cytochrome oxidase subunits; however, it has lost most core genes in the TCA cycle (fig. 5*C*). *Annandia* has little apparent capacity to synthesize or maintain a cell wall/envelope, or to produce the associated membrane proteins (fig. 6). *Annandia* retains no genes for peptidoglycan synthesis, and has lost *mrcB*, which is essential for cell growth in free-living bacteria (fig. 5*B*). Of the core genes involved in phospholipid and fatty acid synthesis (fig. 6*B*), lipid A synthesis, lipopolysaccharide (LPS) core synthesis, and LPS antigen synthesis (fig. 6*C*), outer membrane protein assembly/transport and cell envelope shape and integrity (fig. 6*D*), it retains only three genes, a functional *secY* (involved in protein translocation across the cell membrane), and *secE* and *secG* as pseudogenes.

Composition of the *Pseudomonas* genome with respect to COGs resembles insect symbionts with larger genome sizes. *Pseudomonas* devotes the majority of protein-coding genes to amino acid and coenzyme biosynthesis functions (categories E and H, 18%), translation-related functions (category J, 16.8%), cell wall/membrane/envelope biogenesis (category M, 11.2%), and energy-related functions (category C, 8.6%) (supplementary fig. S2, Supplementary Material online). It retains a number of key genes for nucleotide metabolism (category F), and may therefore be capable of most of this function. In many COG categories, counts of *Pseudomonas* genes are closest to *Baumannia cicadellinicola* str. Hc (the obligate junior symbiont of Cicadellinae leafhoppers), *I. capsulata* str. Mpkobe (an extracellular, nutritional symbiont of Plataspidae stinkbugs), *Hamiltonella defensa* str. 5AT (a facultative symbiont of aphids and whiteflies), *Regiella insecticola* (a facultative symbiont of pea aphids), and *Wigglesworthia glossinidia* (the obligate symbiont of tsetse flies) (supplementary fig. S2, Supplementary Material online). *Pseudomonas* possesses a more complete core gene repertoire than *Annandia*, yet with substantial losses in certain categories (figs. 5 and 6). *Pseudomonas* retains most genes for DNA replication, replication initiation, and repair; however, it has lost over two-thirds of core genes controlling cell division (figs. 5*A* and 6*A*). *Pseudomonas* has functional genes for all but two tRNA synthetases, and contains all core genes controlling transcription and translation (fig. 5*A*). It retains all ribosomal protein genes and core genes involved in protein folding and stability (fig. 5*B*). In energy production and respiration, it has an inventory identical to *Annandia*; however, unlike *Annandia*, it retains a majority of genes in the TCA cycle (fig. 5*C*). Gene losses are more extensive in categories related to the cell envelope (fig. 6). Oddly, although *Pseudomonas* is clearly rod shaped (von Dohlen et al. 2013), it has lost all but four functional core genes controlling cell envelope shape and integrity (fig. 6*D*). It has lost one-third of genes involved in peptidoglycan synthesis, almost half of genes in phospholipid and fatty acid synthesis, 30% of genes contributing to lipid A synthesis, all but one gene involved in LPS antigen synthesis, and 25% of genes for outer membrane protein assembly and transport (fig. 6). Notably, it has also lost the Tol-Pal system, which confers outer membrane stability and is involved in transport of certain macromolecule across the cell envelopes. Although this system is typically lost in symbionts enclosed in a host-derived membrane, it is present in other cytosolic residents such as "*Ca*. Blochmannia" and *Wigglesworthia* (Zientz et al. 2004).

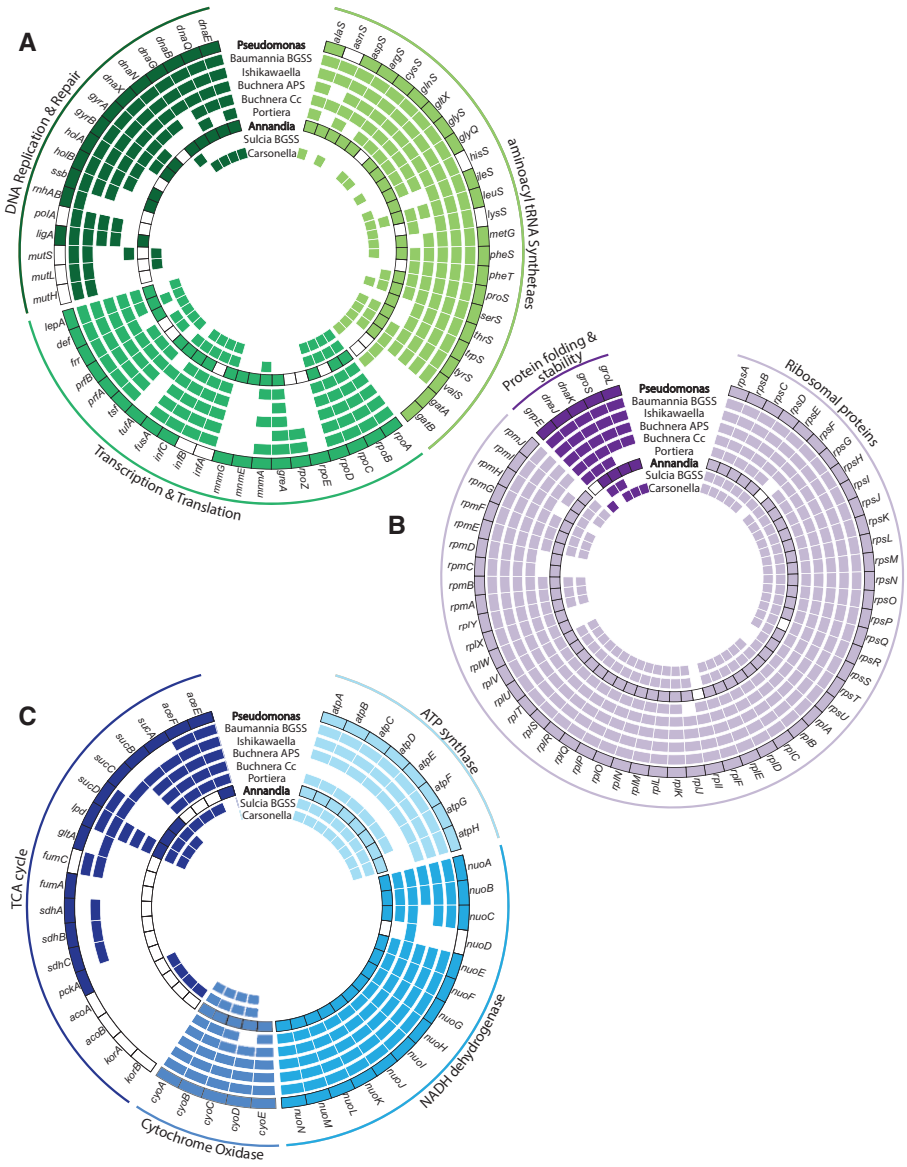

Fig. 5.—Presence and absence of genes involved in central cellular processes related to protein synthesis and energy production for the symbionts of HWA, in comparison to representative obligate hemipteran symbionts. Colored boxes indicate gene presence. HWA symbionts are bolded and their gene boxes are outlined for clarity. Baumannia BGSS is "*Ca.* Baumannia cicadellinicola" str. BGSS, Ishikawaella is "*Ca.* Ishikawaella capsulata," Buchnera APS and Cc are "*Ca.* Buchnera aphidicola" strains APS and Cc, respectively, Portiera is "*Ca.* Portiera aleyrodidarum," Sulcia BGSS is "*Ca.* Sulcia muelleri" str. BGSS, and Carsonella is "*Ca.* Carsonella ruddii" from *Heteropsylla texana*. Genomes are arranged by size from largest (outer ring) to smallest (inner ring).

Discussion

Annandia adelgestsuga is the Putative Ancestral Adelgid Symbiont

The genome of *Annandia adelgestsuga* shares the characteristics of extreme AT bias, reduced size, and high coding density observed in many ancient, obligate endosymbionts of Hemiptera, such as *B. aphidicola*, "*Ca.* Portiera aleyrodidarum," "*Ca.* Moranella endobia," and "*Ca.* Sulcia muelleri". In addition, the core genomic repertoire of *Annandia* resembles these long-term, obligate symbionts. It has lost most genes for central processes of cell division, production of a cell envelope, protein assembly, and transport across the outer membrane. Across the set of COG categories, *Annandia* is more similar to the strain of *Buchnera* found in *Cinara cedri* (Cc) than the APS strain. Unlike the strains in most aphids, *Buchnera* str. Cc coexists with a junior symbiont, *S. symbiotica*, to accomplish essential functions. *Annandia's* higher similarity to this particular *Buchnera* suggests that partnership plays a role in the nature of overall genome degradation. The extent of shrinkage and other genomic distortions in *Annandia*, together with its presence in the sister lineage to

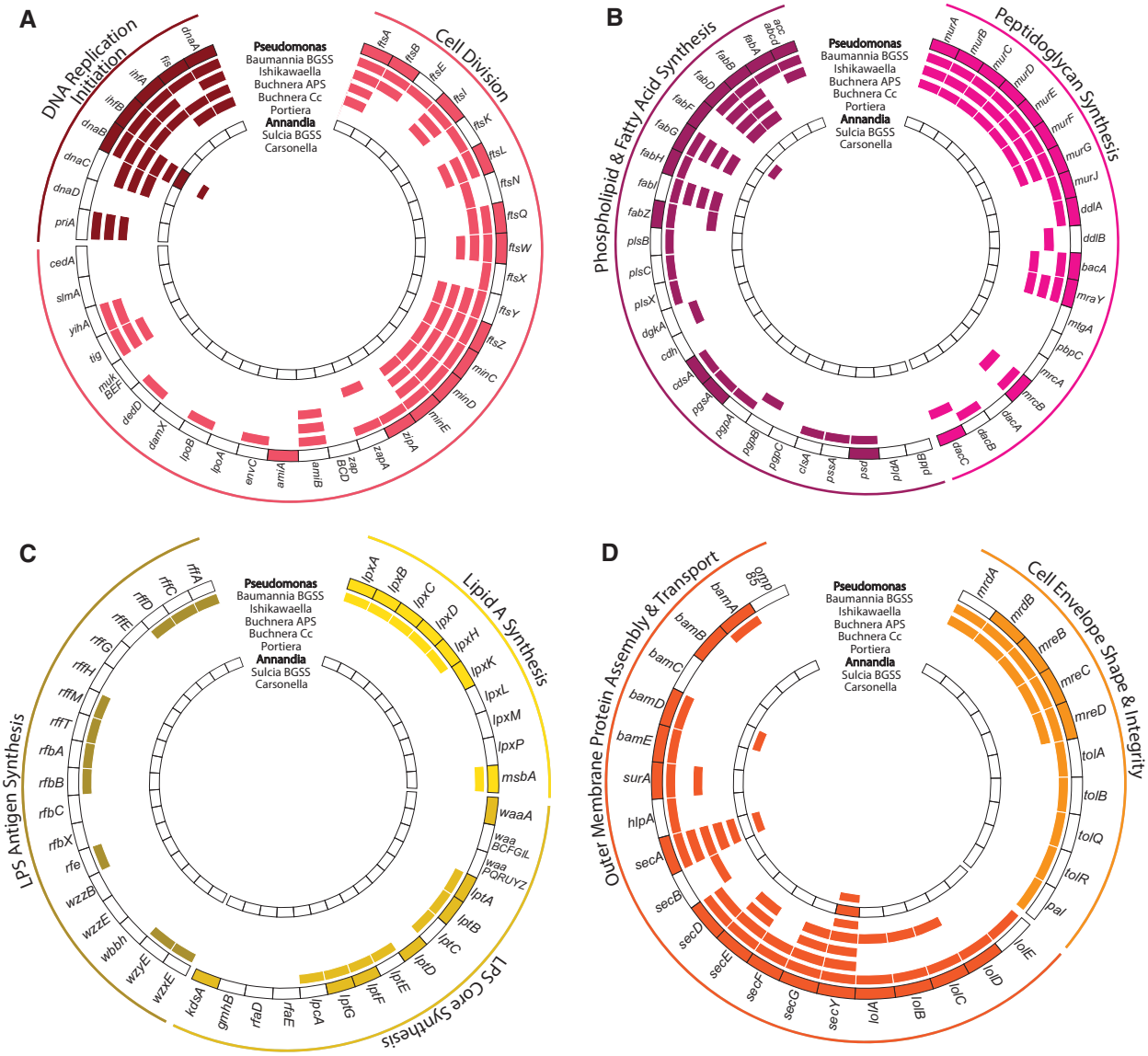

Fig. 6.—Presence and absence of genes involved in central cellular processes related to bacterial replication, including DNA replication initiation, cell division, and synthesis of cell membrane components. For figure details, refer to legend for figure 5.

Adelges, comprising *Pineus* species (Toenshoff et al. 2014; Weglarz et al., in preparation), supports the assignation of this bacterium as the ancestral symbiont of Adelgidae, possibly acquired in the late Cretaceous (Toenshoff et al. 2014; von Dohlen et al. 2017).

Whether a continuous association of *Annandia* ancestors extends deeper into Aphidoidea phylogeny has been in question. Aphidoidea comprises three lineages: Aphididae, Adelgidae, and Phylloxeridae. *Buchnera aphidicola* is the near-universal and ancestral obligate symbiont of Aphididae (Munson et al. 1991; Moran et al. 1993); phylloxerans contain no bacteriome nor obligate symbionts (Vorwerk et al. 2007). The relationship between *Annandia* and other obligate symbionts, particularly *B. aphidicola*, has been difficult to resolve (von Dohlen et al. 2013; Toenshoff et al. 2014). Previous phylogenetic studies based on ribosomal RNA genes disagreed regarding the placement of *Annandia* within a clade of related symbionts. von Dohlen *et al.* (2013) found a weakly supported sister relationship between *Annandia* and *Buchnera* within a set of symbionts including "*Ca.* Purcelliella pentastirinorum" and "*Ca.* Ishikawaella capsulata". Were this relationship true, then the age of the symbiotic association in aphidoids (represented by an ancestral *Annandia*/*Buchnera* lineage) would date at least to the common ancestor of Aphididae and Adelgidae, approximately 120–150 Ma (Heie 1987; Heie and Pike 1996). Toenshoff et al. (2014), however, found a weakly supported sister relationship between *Annandia* and

"*Ca.* Purcelliella". We sought to clarify the position of *Annandia* using genomic data from the set of symbionts with complete genomes. Of the set of species placed closest to *Annandia* in the previous studies, however, genomes were available only for a single strain of "*Ca.* Ishikawaella capsulata" and five *Buchnera* strains. Although we employed strategies recommended to compensate for the rapidly evolving genomes of ancient symbionts (Husník et al. 2011), our phylogenomic analyses failed to support the placement of *Annandia* as sister to *Buchnera* with confidence. Synteny analysis between *Annandia, Ishikawaella*, and *Buchnera* also indicated that these genomes have been highly rearranged with respect to a common ancestor they may have shared. These analyses suggest that *Buchnera* and *Annandia* do not share a most recent common ancestor. Instead of a single, ancient symbiont introduction in the ancestor of Aphidoidea, with a subsequent loss in Pylloxeridae, both *Annandia* and *Buchnera* were more likely acquired independently in their respective host lineages. It remains unclear whether phylloxerans also had a symbiont that was subsequently lost, or simply never acquired one.

Pseudomonas adelgestsugas is an Obligate Symbiont Housed in an Unusual Location

The location of an obligate symbiont in the hemocoel is highly atypical for Sternorrhyncha and other sap-feeding insects (Buchner 1965). *Pseudomonas adelgestsugas* has been found solely in the hemocoel of HWA nymphs and adults, and never in the bacteriome, leading to speculation about its functions and the age of its association (von Dohlen et al. 2013, 2017). We found that the genome of *P. adelgestsugas* exhibits many characteristics of obligate, vertically transmitted symbionts. Compared with free-living *Pseudomonas* relatives, its genome size is reduced by ~75%, its gene count by ~20%, and coding density by ~50%. Although *P. adelgestsugas* has a larger genome and higher GC content than many other obligate sternorrhynchan symbionts, this may be explained by the genomic characteristics of free-living *Pseudomonas* ancestors from which it was descended. Environmental *Pseudomonas* have genomes of approximately 5–7 Mb and 60–65% GC (Stover et al. 2000).

Several aspects of the *P. adelgestsugas* genome suggest that, although it has evolved an obligate association, this condition is more recent than that of *Annandia*. *P. adelgestsugas* was likely acquired in the stem *A. tsugae* lineage or the common ancestor of extant *A. tsugae* species complex (von Dohlen et al. 2017). A fossil-calibrated phylogeny of representative Adelgidae estimated this common ancestor to be approximately 20–35 Ma (Havill et al. 2007); however, a more in-depth study of the *A. tsugae* species complex estimated the ancestor to be considerably younger, at less than 1 Ma (Havill et al. 2016). Although we cannot be certain when *P. adelgestsugas* was acquired, its low coding density suggests that it is actively in the process of transition to a smaller genome, which still contains large amounts of intergenic sequence relative to most obligate symbionts for which genomes are available. Intergenic sequences may be the degenerated remains of inactivated genes that have not yet been removed from the genome via sequence deletions. This feature is shared with the facultative symbionts *Sodalis glossinidius* from tsetse flies, *S. symbiotica* from the pea aphid *Acyrthosiphon pisum*, and other recently acquired *Sodalis*-like obligate symbionts in insects (Plague et al. 2008; Burke and Moran 2011; Clayton et al. 2012; Koga and Moran 2014). In contrast to *S. symbiotica* and *S. glossinidius*, though, *Pseudomonas* is much smaller and only contains a single rRNA operon, suggesting an intermediate state of genome reduction between older obligate and facultative symbionts. A comparison between *P. adelgestsugas* and *P. aeruginosa* PAO1 revealed conservation of order in certain regions, even in light of the reduced size and gene deletions in the symbiont genome. The core gene repertoire retained by *P. adelgestsugas* is also larger and more comprehensive than that of its partner, *Annandia*. This may be attributed both to its presumed younger age and that it lives unbound in the hemolymph, that is, it is not enclosed within host cells or a host-derived membrane (symbiosome) (von Dohlen et al. 2013, 2017). In particular, *P. adelgestsugas* might be expected to maintain greater capabilities related to cell envelope biogenesis, as reflected in the retention of certain genes involved in peptidoglycan biosynthesis, outer-membrane transport, and cell division. Nevertheless, losses of genes or functionality in these categories suggests that the host must be involved in controlling the *P. adelgestsugas* population.

Atypical Pattern of Nutritional Provisioning in HWA Symbionts

Characteristic of other dual-symbiont partnerships in sap-feeding insects (Moran and Bennett 2014), the major role of both obligate symbionts in HWA appears to be that of nutritional mutualist. Together, *Annandia* and *Pseudomonas* have the capacity to provision their host with all ten EAAs. *Annandia* could potentially produce seven EAAs on its own or with input from host genes, and *Pseudomonas* could synthesize five EAAs on its own or with host genes. Of the seven EAAs that *Annandia* could make, three are redundantly produced by *Pseudomonas*. Two EAAs are produced by *Pseudomonas* alone. Also similar to other insects with nutritional symbiont partners, metabolic interdependence for EAA synthesis has evolved, in which each symbiont possesses a different subset of genes necessary to complete the tryptophan biosynthetic pathway (Gosalbes et al. 2008; McCutcheon and von Dohlen 2011; Sloan and Moran 2012). As in other hemipteran-symbiont consortia, the mechanism of metabolite transport over symbiont membranes is

uncertain, but likely involves the host to some degree (Wilson and Duncan 2015). Given the paucity of obvious transporters in the *Annandia* genome, most traffic must be controlled by the host. However, *Pseudomonas* retains greater capability in this regard, and may have control over certain metabolites, including a few EAAs.

Hemlock woolly adelgid symbionts have the capability to produce several NAAs, vitamins, and cofactors, most of which is contributed by *Pseudomonas*. This is consistent with other dual-nutritional symbioses, in which the junior partner typically contributes the majority of non-EAA nutrition (Moran and Bennett 2014). One unusual feature of NAA production is the apparently shared pathway (and gene redundancy) for tyrosine, in which *Annandia* encodes TyrA, both symbionts encode PheA, and *Pseudomonas* encodes the aromatic aminotransferases, AspC and TyrB. Tyrosine is the principal precursor important for sclerotization (hardening) of the insect cuticle after molting. Complete or near-complete pathways for tyrosine are found in *Blochmannia* symbionts of carpenter ants and *Nardonella* symbionts of weevils, in which expression of tyrosine genes is elevated in preadult stages (Zientz et al. 2006; Anbutsu 2017). A newly characterized *Sodalis* symbiont of a lygaeoid bug retains tyrosine as one of two complete AA pathways in its reduced genome, and the TyrA enzyme has possibly undergone alteration to produce higher titres of this amino acid (Santos-Garcia et al. 2017). In contrast, the genomes of obligate symbionts surveyed in supplementary figure S2, Supplementary Material online show little capacity to produce tyrosine, either alone or in cooperation by senior and junior symbionts. Only *I. capsulata* of plataspid stinkbugs possesses *tyrA*, and none possess *tyrB*. Hemlock woolly adelgid individuals are protected by waxy secretions, but underneath this wax their dark-brown cuticle appears heavily sclerotized (for an aphidoid). It is conceivable that symbionts contribute to the production of tyrosine for this additional protection.

The most striking feature regarding nutrient provisioning in the HWA consortium is the more equitable balance of contributions from the two symbionts. In all other dual nutritional symbionts examined so far, the senior symbiont synthesizes the great majority of EAAs (typically eight–nine), or contributes a high proportion of genes to integrated pathways, whereas the junior symbiont retains pathways only for one or two EAAs. This is the case in auchenorrhynchans (Wu et al. 2006; McCutcheon and Moran 2007, 2010; McCutcheon et al. 2009a; Bennett and Moran 2013; Husník et al. 2013; Bennett et al. 2014; Koga and Moran 2014; Mao et al. 2017; Łukasik et al. 2018), as well as sternorrhynchans (Nakabachi et al. 2006; Lamelas et al. 2011; Sloan and Moran 2012; Rosas-Pérez et al. 2014; De Clerck et al. 2015). Even in the most metabolically and physically integrated consortium of certain mealybugs, the senior symbiont possesses twice as many genes in EAA pathways than the junior symbiont (McCutcheon and von Dohlen 2011). Another important way in which the HWA consortium differs from those in other insects is the redundancy present in EAA pathways between *Annandia* and *Pseudomonas*. It is possible that this redundancy will be lost over time to create complementarity of metabolic roles, a theme that has evolved independently in other dual-partner obligate symbioses in insects (McCutcheon and Moran 2011; Bennett and Moran 2013).

Acquisition of a junior symbiont may be an effective compensatory mechanism for loss of function in the original senior symbiont (Bennett and Moran 2015). Highly reduced and otherwise modified genomes are hallmarks of long-term nutritional symbionts (Wernegreen 2002), and gene deletions may occur in essential nutritional pathways that benefit hosts (Moran et al. 2008). Such genomic degradation is thought to be a consequence of several factors, including elevated genetic drift from population bottlenecks, mutational bias toward deletions, lack of recombinational opportunities, and relaxed selection in the host environment (Moran 1996; Rispe and Moran 2000; Moran et al. 2009). Strong host-level selection may counter symbiont genome degradation in various ways, through elevated expression of chaperonins, host–insect expression of genes functionally equivalent to those lost, and horizontal transfer of genes from transient facultative symbionts (Sato and Ishikawa 1997; Wilson et al. 2010; Husník et al. 2013; Sloan et al. 2014). A more comprehensive solution, however, is to acquire a partner symbiont: While the former compensations must evolve as individual selective events, the acquisition of a junior symbiont that possesses a genome with complete nutritional pathways potentially solves many deficiencies simultaneously.

The examples above illustrate that host-level selection acts as a strong counterbalance to symbiont-level forces of genome decay, because ancient senior symbionts have maintained much of their nutritional functionality over hundreds of millions of years. This seems not to be the case in HWA, in which the ancestral senior symbiont, *Annandia*, has lost greater amino acid capabilities over a much shorter time period. What could be different about the adelgid case, where the high rate of symbiont replacement suggests a repeated cycle of accelerated senior symbiont decline? We do not expect, *a priori*, that the action of drift, mutational bias, or recombination should be substantially different in adelgids. What may be different is the nature of host-level selection. One way that loss of nutritional function in symbionts might be accelerated is if the host's diet provides those nutrients. Fluctuations in nutrient levels within the host environment could periodically relax selection on symbiont pathways, such that formerly essential genes could be inactivated (Bennett and Moran 2015). We have previously proposed a hypothesis incorporating this idea (von Dohlen et al. 2017). Historical changes in dietary dependence on phloem versus parenchyma during the evolution of host-alternating life cycles in adelgids may have imposed fluctuating selection on symbiont functions and accelerated gene inactivation in

nutrient pathways. The substantial losses we see in nutritional capabilities of *Annandia* support our hypothesis, which posits that this metabolism was lost before a junior symbiont was gained. We suggest that this functionality may have degraded during an evolutionary period after gaining a gall phase (high-nutrient parenchyma) but before evolving host alternation and low-nutrient phloem-dependent generations (von Dohlen et al. 2017). Hemlock woolly adelgid is atypical among adelgids in that it feeds on xylem ray parenchyma cells on hemlock, its alternate host (Young et al. 1995). Many populations in the HWA species complex are exclusively (or nearly so) parthenogenetic on hemlock, but in some populations a complete cycle to galls on spruce may occur (Havill et al. 2016). Certain generations in these complete cycles may still depend on phloem feeding, which could explain why the dual symbionts of HWA maintain full EAA synthesis pathways. Our ongoing genomic studies are investigating whether this partitioning of nutrient production holds for dual symbionts of other adelgid lineages.

No Obvious Role for Obligate Symbionts in HWA Virulence

Feeding by HWA elicits both a localized and systemic HR in host hemlock trees (Radville et al. 2011). As a plant defense against herbivores, the HR consists of elevated levels of reactive oxygen species and can lead to tissue death at the feeding site (Heath 2000; Liu et al. 2010). We suggested previously that toxins or other molecules from bacterial symbionts in HWA might induce the HR in hemlock (von Dohlen et al. 2013). A recent study found that HWA feeding in hemlock increases emission of volatile compounds typically elicited by plant pathogens (Schaeffer et al. 2018). Because products from insect hemolymph can be transported into saliva (Miles 1967), it is conceivable that molecules derived from "*Ca*. P. adelgestsugas," in particular, might make their way into host–plant tissue and trigger the plant defense. In several other systems, herbivore-associated microbes may manipulate plant physiology through effector molecules such as flagellin, porin-like proteins, or cytokinins (Kaiser et al. 2010; Body et al. 2013; Chung et al. 2013; Guo et al. 2013; Acevedo et al. 2015). Plant-pathogenic *Pseudomonas* and other bacteria elicit an HR through the type III secretion system (TTSS), whereby molecules such as flagellin, LPS, peptidoglycan, and elongation factor TU are injected into plant tissues (Chisholm et al. 2006; Cunnac et al. 2009; Deslandes and Rivas 2012).

We searched the genomes of both symbionts of HWA for putative effectors of the HR, but found no obvious candidates. "*Ca*. Annandia adelgestsuga" seems devoid of any potential effectors. The genome of "*Ca*. P. adelgestsugas" contains no apparent TTSS, nor any recognizable genes for flagellin (*fli* genes) or cytokinin (e.g., *ipt*). Our annotations revealed no genes for other known pathogen effectors (e.g., *avr, hop, pop, xop, pep13, epl*) (Chisholm et al. 2006; Cunnac et al. 2009; Deslandes and Rivas 2012). "*Ca*. P. adelgestsugas" is also missing several core genes for peptidoglycan and LPS core synthesis and lacks all core genes for LPS antigens. Both symbionts of HWA presumably rely on host input to synthesize a cell envelope, as in other symbionts lacking these genes (Nakabachi et al. 2006; Perez-Brocal et al. 2006; Wu et al. 2006; McCutcheon et al. 2009a, 2009b; McCutcheon and Moran 2010; McCutcheon and von Dohlen 2011; Sloan and Moran 2012; Bennett and Moran 2013; Bennett et al. 2014; Moran and Bennett 2014). Whether such hybrid membrane products would be recognized by plants as signals of bacterial invasion remains in question. "*Ca*. P. adelgestsugas" does contain genes for two outer-membrane porin proteins, *nicP* (*ybfM*) and *oprD* (*oprQ*), although neither of these has been identified specifically as an effector.

Effectors of the HR in hemlock might derive from symbiont sources as yet unidentified. Genomes of both HWA symbionts contain hypothetical genes of unknown function. Although "*Ca*. Annandia adelgestsuga" has fewer than 10 such genes, "*Ca*. P. adelgestsugas" contains almost 200. Another possible source of effectors could be the third, facultative symbiont in certain HWA populations, identified as *S. symbiotica* (von Dohlen et al. 2013). This symbiont has been found only in the introduced HWA populations in eastern North America (ENA), and the Japanese source population (Havill et al. 2006; von Dohlen et al. 2013). Based on 16S rDNA sequences, *S. symbiotica* in HWA is sister to the facultative strain found in aphids but is not identical. Although *S. symbiotica* is not universally present in all individuals of ENA populations, it has been detected in at least some individuals from every tree sampled (Mech et al. 2017). We are currently working to assemble and annotate the *S. symbiotica* genome to characterize its metabolism.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank John McCutcheon for bioinformatics assistance, Gordon Bennett for template figure files of core-gene comparisons, and Thomas Crowley and Cody Bills for laboratory assistance. We would like to acknowledge the Arnold Arboretum of Harvard University for providing financial support through a Deland Award to K.M.W. K.M.W. was also supported by a grant from the Lewis and Clark Fund for Exploration and Field Research from the American Philosophical Society, and a Doctoral Dissertation Improvement Grant from the National Science Foundation to C.D.v.D. and K.M.W. (DEB-1601802). Research was further supported by National Science Foundation grants to

C.D.v.D. (DEB-1655182) and G.R.B. (DEB-1655177), the Utah Agricultural Experiment Station (approved as journal paper 9074), the USDA Forest Service Forest Health Protection (#15-DG-11083150-068), and the University of Georgia's Research Foundation, Office for the Vice President for Research, and Agricultural Experiment Station. The support and resources from the Center for High Performance Computing at the University of Utah are also gratefully acknowledged.

Literature Cited

Acevedo FE, Rivera-Vega LJ, Chung SH, Ray S, Felton GW. 2015. Cues from chewing insects—the intersection of DAMPs, HAMPs, MAMPs and effectors. Curr Opin Plant Biol. 26:80–86.

Anbutsu H. 2017. Small genome symbiont underlies cuticle hardness in beetles. Proc Natl Acad Sci U S A. 114(40):E8382–E8391.

Bankevich A, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Beckwith J. 2013. The Sec-dependent pathway. Res Microbiol. 164(6):497–504.

Bennett GM, McCutcheon JP, MacDonald BR, Romanovicz D, Moran NA. 2014. Differential genome evolution between companion symbionts in an insect-bacterial symbiosis. mBio 5(5):e01697–e01614.

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(9):1675–1688.

Bennett GM, Moran NA. 2015. Heritable symbiosis: the advantages and perils of an evolutionary rabbit hole. Proc Natl Acad Sci U S A. 112(33):10169–10176.

Body M, Kaiser W, Dubreuil G, Casas J, Giron D. 2013. Leaf-miners co-opt microorganisms to enhance their nutritional environment. J Chem Ecol. 39(7):969–977.

Boussau B, Gouy M. 2006. Efficient likelihood computations with non-reversible models of evolution. Syst Biol. 55(5):756–768.

Buchner P. 1965. Endosymbiosis of Animals with Plant Microorganisms. New York: John Wiley and Sons.

Burke GR, Moran NA. 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. Genome Biol Evol. 3:195–208.

Caspi R, et al. 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 44(D1):D471–D480.

Chisholm ST, Coaker G, Day B, Staskawicz BJ. 2006. Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124(4):803–814.

Chung SH, et al. 2013. Herbivore exploits orally secreted bacteria to suppress plant defenses. Proc Natl Acad Sci U S A. 110(39):15728–15733.

Clayton AL, et al. 2012. A novel human-infection-derived bacterium provides insights into the evolutionary origins of mutualistic insect–bacterial symbioses. PLoS Genet. 8(11):e1002990.

Cui J, et al. 2005. *Pseudomonas syringae* manipulates systemic plant defenses against pathogens and herbivores. Proc Natl Acad Sci U S A. 102(5):1791–1796.

Cunnac S, Lindeberg M, Collmer A. 2009. *Pseudomonas syringae* type III secretion system effectors: repertoires in search of functions. Curr Opin Microbiol. 12(1):53–60.

De Clerck C, et al. 2015. A metagenomic approach from aphid's hemolymph sheds light on the potential roles of co-existing endosymbionts. Microbiome 3(1):1–11.

Deslandes L, Rivas S. 2012. Catch me if you can: bacterial effectors and plant targets. Trends Plant Sci. 17(11):644–655.

Fernandes GW. 1990. Hypersensitivity: a neglected plant resistance mechanism against insect herbivores. Environ Entomol. 19(5):1173–1182.

Gosalbes MJ, Lamelas A, Moya A, Latorre A. 2008. The striking case of tryptophan provision in the cedar aphid *Cinara cedri*. J Bacteriol. 190(17):6026–6029.

Gouy M, Guindon S, Gascuel O. 2010. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol. 27(2):221–224.

Gruwell ME, Hardy NB, Gullan PJ, Dittmar K. 2010. Evolutionary relationships among primary endosymbionts of the mealybug subfamily phenacoccinae (Hemiptera: coccoidea: pseudococcidae). Appl Environ Microbiol. 76(22):7521–7525.

Gunduz EA, Douglas AE. 2009. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. Proc R Soc B Biol Sci. 276(1658):987–991.

Guo H, et al. 2013. A porin-like protein from oral secretions of *Spodoptera littoralis* larvae induces defense-related early events in plant leaves. Insect Biochem Mol Biol. 43(9):849–858.

Hansen AK, Moran NA. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proc Natl Acad Sci U S A. 108(7):2849–2854.

Havill NP, Foottit RG. 2007. Biology and evolution of Adelgidae. Annu Rev Entomol. 52:325–349.

Havill NP, Foottit RG, von Dohlen CD. 2007. Evolution of host specialization in the Adelgidae (Insecta: hemiptera) inferred from molecular phylogenetics. Mol Phylogenet Evol. 44(1):357–370.

Havill NP, Montgomery ME, Yu GY, Shiyake S, Caccone A. 2006. Mitochondrial DNA from hemlock woolly adelgid (Hemiptera: adelgidae) suggests cryptic speciation and pinpoints the source of the introduction to eastern North America. Ann Entomol Soc Am. 99:195–203.

Havill NP, et al. 2016. Ancient and modern colonization of North America by hemlock woolly adelgid, *Adelges tsugae* (Hemiptera: adelgidae), an invasive insect from East Asia. Mol Ecol. 25(9):2065–2080.

Havill NP, Vieira LC, Salom SM. 2014. Biology and Control of Hemlock Woolly Adelgid, Vol. FHTET-2014-05, USDA Forest Service, Forest Health Technology Enterprise Team.

Heath MC. 2000. Hypersensitive response-related death. Plant Mol Biol. 44(3):321–334.

Heie OE. 1987. Paleontology and phylogeny. In: Minks, AK, Harrewijn, P, editors. Aphids. Their Biology, Natural Enemies and Control, World Crop Pests Vol. 2A. Amsterdam: Elsevier. p. 367–391.

Heie OE, Pike EM. 1996. Reassessment of the taxonomic position of the fossil aphid family Canadaphididae based on two additional specimens of *Canadaphis carpenteri* (Hemiptera: aphidinea). Eur J Entomol. 93:617–622.

Huerta-Cepas J, et al. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 44(D1):D286–D293.

Husník F, Chrudimský T, Hypsa V. 2011. Multiple origins of endosymbiosis within the Enterobacteriaceae (gamma-Proteobacteria): convergence of complex phylogenetic approaches. BMC Biol. 9:87.

Husník F, et al. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell 153(7):1567–1578.

Kaiser W, Huguet E, Casas J, Commin C, Giron D. 2010. Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. Proc R Soc B Biol Sci. 277(1692):2311–2319.

Kellner RLL. 2002. Molecular identification of an endosymbiotic bacterium associated with pederin biosynthesis in *Paederus sabaeus* (Coleoptera: staphylinidae). Insect Biochem Mol Biol. 32(4):389–395.

Kellner RLL, Dettner K. 1996. Differential efficacy of toxic pederin in deterring potential arthropod predators of *Paederus* (Coleoptera: staphylinidae) offspring. Oecologia 107(3):293–300.

Keseler IM, et al. 2013. EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res. 41(D1):D605–D612.
Koga R, Moran NA. 2014. Swapping symbionts in spittlebugs: evolutionary replacement of a reduced genome symbiont. ISME J. 8:1237–1246.
Lamelas A, et al. 2011. *Serratia symbiotica* from the aphid *Cinara cedri*: a missing link from facultative to obligate insect endosymbiont. PLoS Genet. 7(11):e1002357.
Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol. 62(4):611–615.
Lerat E, Ochman H. 2004. Ψ-Φ: exploring the outer limits of bacterial pseudogenes. Genome Res. 14(11):2273–2278.
Liu X, et al. 2010. Reactive oxygen species are involved in plant defense against a gall midge. Plant Physiol. 152(2):985–999.
Luan JB, et al. 2015. Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. Genome Biol Evol. 7(9):2635–2647.
Łukasik P, et al. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of cicadas. Proc Natl Acad Sci U S A. 115(2):E226–E235.
Mao M, Yang X, Poff K, Bennett G. 2017. Comparative genomics of the dual-obligate symbionts from the treehopper, *Entylia carinata* (Hemiptera: membracidae), provide insight into the origins and evolution of an ancient symbiosis. Genome Biol Evol. 9(6):1803–1815.
McCutcheon JP, McDonald BR, Moran NA. 2009a. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. Proc Natl Acad Sci U S A. 106(36):15394–15399.
McCutcheon JP, McDonald BR, Moran NA. 2009b. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. PLoS Genet. 5(7):e1000565.
McCutcheon JP, Moran NA. 2007. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. Proc Natl Acad Sci U S A. 104(49):19392–19397.
McCutcheon JP, Moran NA. 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 my of evolution. Genome Biol Evol. 2:708–718.
McCutcheon JP, Moran NA. 2011. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 10(1):13–26.
McCutcheon JP, von Dohlen CD. 2011. An interdependent metabolic patchwork in the nested symbiosis of mealybugs. Curr Biol. 21(16):1366–1372.
Mech AM, Harper SJ, Havill NP, von Dohlen CD, Burke GR. 2017. Ecological factors influencing the beneficial endosymbionts of the hemlock woolly adelgid (Hemiptera: Adelgidae). Insect Sci. p. 1–11, https://doi.org/10.1111/1744-7917.12514.
Miles PW. 1967. Studies on the salivary physiology of plant-bugs: transport from haemolymph to saliva. J Insect Physiol. 13(12):1787–1801.
Moran NA. 1996. Accelerated evolution and Müller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A. 93(7):2873–2878.
Moran NA. 2001. The coevolution of bacterial endosymbionts and phloem-feeding insects. Ann Mol Bot Gard. 88(1):35–44.
Moran NA, Bennett GM. 2014. The tiniest tiny genomes. Annu Rev Microbiol. 68:195–215.
Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet. 42:165–190.
Moran NA, McLaughlin HJ, Sorek R. 2009. The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. Science 323(5912):379–382.
Moran NA, Munson MA, Baumann P, Ishikawa H. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proc R Soc B Biol Sci. 253(1337):167–171.
Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. 35(Web Server issue): W182–W185.
Munson MA, et al. 1991. Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. J Bacteriol. 173(20):6321–6324.
Nakabachi A, et al. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314(5797):267.
Nimchuk Z, Eulgem T, Holt BF III, Dangl JL. 2003. Recognition and response in the plant immune system. Annu Rev Genet. 37(1):579–609.
Perez-Brocal V, et al. 2006. A small microbial genome: the end of a long symbiotic relationship? Science 314(5797):312–313.
Pezet J, et al. 2013. Hemlock woolly adelgid and elongate hemlock scale induce changes in foliar and twig volatiles of eastern hemlock. J Chem Ecol. 39(8):1090–1100.
Piel J. 2002. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. Proc Natl Acad Sci U S A. 99(22):14002–14007.
Plague GR, Dunbar HE, Tran PL, Moran NA. 2008. Extensive proliferation of transposable elements in heritable bacterial symbionts. J Bacteriol. 190(2):777–779.
Profft J. 1937. Beitrage zur symbiose der aphiden und psylliden. Z Morphol Ökol Tiere 32(2):289–326.
Radville L, Chaves A, Preisser EL. 2011. Variation in plant defense against invasive herbivores: evidence for a hypersensitive response in eastern hemlocks (*Tsuga canadensis*). J Chem Ecol. 37(6):592–597.
Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6. Available from: http://tree.bio.ed.ac.uk/software/tracer/
Rao Q, et al. 2015. Genome reduction and potential metabolic complementation of the dual endosymbionts in the whitefly *Bemisia tabaci*. BMC Genomics 16(1):226.
Rispe C, Moran NA. 2000. Accumulation of deleterious mutations in endosymbionts: Muller's ratchet with two levels of selection. Am Nat. 156(4):425–441.
Rosas-Pérez T, Rosenblueth M, Rincón-Rosales R, Mora J, Martínez-Romero E. 2014. Genome sequence of '*Candidatus* Walczuchella monophlebidarum' the flavobacterial endosymbiont of *Llaveia axin axin* (Hemiptera: coccoidea: monophlebidae). Genome Biol Evol. 6(3):714–726.
Rosenblueth M, Sayavedra L, Sámano-Sánchez H, Roth A, Martínez-Romero E. 2012. Evolutionary relationships of flavobacterial and enterobacterial endosymbionts with their scale insect hosts (Hemiptera: coccoidea). J Evol Biol. 25(11):2357–2368.
Santos-Garcia D, et al. 2012. Complete genome sequence of '*Candidatus* Portiera aleyrodidarum' BT-QVLC, an obligate symbiont that supplies amino acids and carotenoids to *Bemisia tabaci*. J Bacteriol. 194(23):6654–6655.
Santos-Garcia D, et al. 2017. The all-rounder *Sodalis*: a new bacteriome-associated endosymbiont of the lygaeoid bug *Henestaris halophilus* (Heteroptera: henestarinae) and a critical examination of its evolution. Genome Biol Evol. 9(10):2893–2910.
Sato S, Ishikawa H. 1997. Expression and control of an operon from an intracellular symbiont which is homologous to the groE operon. J Bacteriol. 179(7):2300–2304.
Schaeffer RN, Wang Z, Thornber CS, Preisser EL, Orians CM. 2018. Two invasive herbivores on a shared host: patterns and consequences of phytohormone induction. Oecologia 186(4):973–982.
Shields KS, Hirth RT. 2005. Bacterial endosymbionts of *Adelges tsugae* Annand: potential targets for biocontrol? In: Onken, B, Reardon, R, editors. Third Symposium on Hemlock Woolly Adelgid in the Eastern United States. USDA Forest Service, FHTET-2005-01, p. 357–359.
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(6800):81–86.

Sloan DB, et al. 2014. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. Mol Biol Evol. 31(4):857–871.
Sloan DB, Moran NA. 2012. Genome reduction and co-evolution between the primary and secondary bacterial symbionts of psyllids. Mol Biol Evol. 29(12):3781–3792.
Spaulding AW, von Dohlen CD. 2001. Psyllid endosymbionts exhibit patterns of co-speciation with hosts and destabilizing substitutions in ribosomal RNA. Insect Mol Biol. 10(1):57–67.
Steffan AW. 1976. Evolution of morphological characters and of endosymbionts in the aphid family Adelgidae (Homoptera: aphidina). Verh Dtsch Zool Ges. 69:232.
Stover CK, et al. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. Nature 406(6799):959–964.
Tatusov RL, et al. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 29(1):22–28.
Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28(1):33–36.
Toenshoff ER, Gruber D, Horn M. 2012. Co-evolution and symbiont replacement shaped the symbiosis between adelgids (Hemiptera: adelgidae) and their bacterial symbionts. Environ Microbiol. 14(5):1284–1295.
Toenshoff ER, et al. 2012. Bacteriocyte-associated gammaproteobacterial symbionts of the *Adelges nordmannianae/picea*e complex (Hemiptera: adelgidae). ISME J. 6(2):384–396.
Toenshoff ER, Szabo G, Gruber D, Horn M. 2014. The pine bark adelgid, *Pineus strobi*, contains two novel bacteriocyte-associated gammaproteobacterial symbionts. Appl Environ Microbiol. 80(3):878–885.
Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. 2011. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol. 12(3):R30.
von Dohlen CD, et al. 2013. Diversity of proteobacterial endosymbionts in hemlock woolly adelgid (*Adelges tsugae*) (Hemiptera: adelgidae) from its native and introduced range. Environ Microbiol. 15(7): 2043–2062.
von Dohlen CD, et al. 2017. Dynamic acquisition and loss of dual-obligate symbionts in the plant-sap-feeding Adelgidae (Hemiptera: sternorrhyncha: aphidoidea). Front Microbiol. 8:1–15.
Vorwerk S, Martinez-Torres D, Forneck A. 2007. *Pantoea agglomerans*-associated bacteria in grape phylloxera (*Daktulosphaira vitifoliae*, Fitch). Agric For Entomol. 9(1):57–64.
Wang Y, et al. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40(7):e49.
Wernegreen JJ. 2002. Genome evolution in bacterial endosymbionts of insects. Nat Rev Genet. 3(11):850–861.
Wickham H. 2009. ggplot2: elegant Graphics for Data Analysis. New York: Springer-Verlag.
Wilson ACC, et al. 2010. Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol Biol. 19:249–258.
Wilson ACC, Duncan RP. 2015. Signatures of host/symbiont genome coevolution in insect nutritional endosymbioses. Proc Natl Acad Sci U S A. 112:10255–10261.
Wu D, et al. 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. PLoS Biol. 4(6):e188.
Xu Y, et al. 2016. VGSC: A Web-Based Vector Graph Toolkit of Genome Synteny and Collinearity. BioMed Res. Int. 2016:7.
Young RF, Shields KS, Berlyn GP. 1995. Hemlock woolly adelgid (Homoptera: adelgidae): Stylet bundle insertion and feeding sites. Ann Entomol Soc Am. 88(6):827–835.
Zientz E, Beyaert I, Gross R, Feldhaar H. 2006. Relevance of the endosymbiosis of *Blochmannia floridanus* and carpenter ants at different stages of the life cycle of the host. Appl Environ Microbiol. 72(9):6027–6033.
Zientz E, Dandekar T, Gross R. 2004. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol Mol Biol Rev. 68(4):745–770.

Associate editor: Esther Angert