

Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont

Zakee L. Sabree^{a,b}, Srinivas Kambhampati^c, and Nancy A. Moran^{b,1}

^aCenter for Insect Science and ^bDepartment of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721-0088; and ^cDepartment of Entomology, Kansas State University, Manhattan, KS 66506-4004

Edited by Jeffrey I. Gordon, Washington University School of Medicine, St. Louis, MO, and approved September 23, 2009 (received for review July 8, 2009)

Nitrogen acquisition and assimilation is a primary concern of insects feeding on diets largely composed of plant material. Reclaiming nitrogen from waste products provides a rich reserve for this limited resource, provided that recycling mechanisms are in place. Cockroaches, unlike most terrestrial insects, excrete waste nitrogen within their fat bodies as uric acids, postulated to be a supplement when dietary nitrogen is limited. The fat bodies of most cockroaches are inhabited by *Blattabacterium*, which are vertically transmitted, Gram-negative bacteria that have been hypothesized to participate in uric acid degradation, nitrogen assimilation, and nutrient provisioning. We have sequenced completely the *Blattabacterium* genome from *Periplaneta americana*. Genomic analysis confirms that *Blattabacterium* is a member of the Flavobacteriales (Bacteroidetes), with its closest known relative being the endosymbiont *Sulcia muelleri*, which is found in many sap-feeding insects. Metabolic reconstruction indicates that it lacks recognizable uricolytic enzymes, but it can recycle nitrogen from urea and ammonia, which are uric acid degradation products, into glutamate, using urease and glutamate dehydrogenase. Subsequently, *Blattabacterium* can produce all of the essential amino acids, various vitamins, and other required compounds from a limited palette of metabolic substrates. The ancient association with *Blattabacterium* has allowed cockroaches to subsist successfully on nitrogen-poor diets and to exploit nitrogenous wastes, capabilities that are critical to the ecological range and global distribution of cockroach species.

Flavobacteria | insect endosymbiont | nitrogen conservation | uricolysis

Nitrogen acquisition for cellular metabolism is a primary concern of herbivorous animals. Cockroaches (Dictyoptera: Blattaria) typically consume decaying (nitrogen-poor) plant material but can be omnivorous, feeding opportunistically on decaying animal material and even (nitrogen-rich) conspecifics (1). Cockroaches respond to surplus nitrogen by concentrating and storing it as uric acid, which is excreted typically by terrestrial insects, within fat body cells (2). Increasing dietary nitrogen concentrations correlates with uric acid deposition in cockroach fat bodies (2, 3). These internal stores are thought to serve as metabolic reserves, providing an essential supplement when sufficient dietary nitrogen is lacking and a means for exploiting nitrogen-rich windfalls (4).

Although the ecological rationale for nitrogen conservation in phytophagous insects is clear, utilization of uric acid as a nitrogen reserve is not well understood. Many insects and other animals do not have the enzymatic capabilities required for producing ammonia from urate. Nitrogen recycling from stored uric acid in shield bugs and termites has been attributed to symbiotic microbes. The presence of *Erwinia*-like bacteria in *Parastrachia japonensis*, which are located exclusively in the gastric ceca, correlates with enzymatic uricolysis in the gastric ceca (5). Similarly, *Reticulitermes flavipes* hindgut bacteria ferment uric acid transported from the fat bodies, via the Malpighian tubules to the hindgut, into ammonia, CO₂, and acetate (6–8). Cockroaches, close relatives of termites, harbor intracellular, maternally transmitted bacteria, called *Blattabacterium* sp.,

within their fat bodies, near the uric acid deposits (9), and they have been hypothesized to participate in nitrogen recycling from stored uric acid (10, 11). *Periplaneta americana* fat bodies are comprised of two distinct cell types: adipocytes, containing fat globules and uric acid crystals, and mycetocytes, which are filled densely with *Blattabacterium* (9, 11–13). After antibiotic treatment to remove *Blattabacterium*, mycetocytes appear to be highly degraded, uric acid accumulates significantly (13–15), and production of tyrosine, phenylalanine, isoleucine, valine, arginine, and threonine is reduced (16). Consequently, production of these amino acids has been attributed to *Blattabacterium*.

Phylogenetic analysis indicates that *Blattabacterium* is a member of the Flavobacteriales, where it clusters with other insect bacteriome-associated symbionts, *Candidatus Sulcia muelleri* (hereafter referred to as *Sulcia*, hosts in numerous sap-feeding insects within Hemiptera: Auchenorrhyncha) and *Candidatus Uzinura diaspidicola* (hosts in sap-feeding scale insects, family Diaspididae), and an insect host-manipulator, the male-killing ladybird beetle endosymbiont (17–20). *Blattabacterium* appears to have infected the common ancestor of cockroaches and termites >140 Mya (21, 22). It has been maintained in all cockroaches except the cave-dwelling *Noticola* and lost in all termites except the basal species, *Mastotermes darwiniensis* (17, 22). Although the existence of *Blattabacterium* has been known for over 100 years, efforts to fully characterize it have been thwarted by its unsurprising recalcitrance to cultivation (23). Many beneficial bacterial symbionts with exclusively intracellular lifestyles (i.e., *Buchnera aphidicola*) have lost their ability to live outside their hosts but have free-living, cultivable relatives (e.g., *Escherichia coli*) (24). To elucidate further the role of *Blattabacterium* in nutrient provisioning, nitrogen conservation, and uric acid metabolism, we sequenced the genome of the *P. americana* endosymbiont *Blattabacterium* sp. BPLAN.

Results and Discussion

The genome of *Blattabacterium* sp. BPLAN is comprised of a 636,994-bp circular chromosome and a 3,448-bp plasmid with G + C contents of 28.2% and 28.5%, respectively. Thirty-three tRNAs corresponding to all of the amino acids are encoded on the chromosome. Additionally, a single 16S-23S-5S operon and a tmRNA gene are present. Complete transcriptional machinery, including the major sigma factor RpoD, is encoded in the genome. The origin of replication was predicted based on GC skew, and the first nucleotide in the nearest protein-coding region was designated as position 1 (Fig. S1). *Blattabacterium* encodes 581 ORFs, comprising 94% of the sequence, and 520 ORFs have an assigned function (Table 1). Four genes, encoding

Author contributions: Z.L.S., S.K., and N.A.M. designed research; Z.L.S. and S.K. performed research; Z.L.S., S.K., and N.A.M. analyzed data; and Z.L.S. and N.A.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. CP001429 and CP001430).

¹To whom correspondence should be addressed. E-mail: nmoran@email.arizona.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0907504106/DCSupplemental.

Table 1. Comparison of *Blattabacterium* sp. BPLAN genome with those of other Flavobacteriales

	<i>Blattabacterium</i> sp. BPLAN	<i>Sulcia muelleri</i> GWSS	<i>Flavobacterium johnsonii</i> UW101
Lifestyle	Intracellular symbiont	Intracellular symbiont	Free-living
Habitat	Cockroaches	Auchenorrhyncha	Soil and freshwater
Chromosome, bp	636,994	245,530	6,096,872
Plasmid [bp]	1 [3,448]	0	0
G + C content [plasmid], %	28.2 [28.5]	22	34
Total CDS	581	227	5,017
Average CDS length, bp	1,031	1,006	1,057
CDS density, %	94	87	87
CDS with assigned function	520	210	3,022
Pseudogenes	4	0	39
rRNAs	3	3	18
tRNAs	33	31	62

General features of the *Blattabacterium* sp. BPLAN genome compared with those of other completely sequenced Flavobacteriales. CDS, protein coding sequences.

a putative oxidoreductase, two ABC transporters, and a bifunctional sulfate adenylyltransferase, harbor internal stop codons and were annotated as pseudogenes.

Overall, 54% of protein-encoding genes with assigned functions are involved in amino acid metabolism, translation and biogenesis, cell wall and outer membrane biogenesis, or coenzyme metabolism and energy production (Table S1). Complete de novo purine and pyrimidine biosynthetic pathways are encoded in the *Blattabacterium* genome. Interestingly, the ribonucleotide reductase (RNR), which performs the necessary reduction of ribonucleotides to deoxyribonucleotides, is comprised of an α subunit that is encoded by a chromosome-based gene and a β subunit that is encoded by a plasmid-based gene. The *Blattabacterium* plasmid does not carry genes involved in amino acid biosynthesis, as observed for plasmids of the aphid symbiont *B. aphidicola*. Thus, the sole function of the plasmid appears to be encoding the RNR β subunit, which is required for RNR function, but why this subunit has a plasmid location is not known. On the basis of its genome, *Blattabacterium* generates essential carbon precursor metabolites by gluconeogenesis, the citric acid cycle, and the pentose phosphate pathway. Adenosine 5'-triphosphate is generated aerobically using a *cbb*₃-type cytochrome *c* oxidase-terminated electron transport chain.

Retention of genes encoding enzymes involved in vitamin and amino acid biosynthesis and loss of many genes encoding transcriptional regulatory functions and substrate-specific transporters are common features of obligate symbionts of insects (25). However, the exchange of molecules, including required metabolites, between symbiont and host cellular compartments is the foundation of the symbiotic interaction, implying that some transport mechanisms persist. In *Blattabacterium*, the NMN transporter PnuC is present and can facilitate exogenous pyridine uptake for NAD biosynthesis and recycling (26). Movement of proteins across the cell wall could be achieved by the Sec and twin-arginine targeting translocases encoded in the genome (27, 28). *Blattabacterium* also encodes multiple ATP-binding ABC transporter complexes. Although cation transport proteins are present in the *Blattabacterium* genome, few substrate-specific transporters, especially those involved in amino acid uptake or secretion, are identifiable, and this is consistent with other insect nutritional endosymbionts (18, 29).

Amino Acid Biosynthesis and Nitrogen Recycling. The central role hypothesized for *Blattabacterium* is amino acid biosynthesis and waste nitrogen recycling, and this role is supported by past experimental studies (13, 16, 30, 31). Metabolic pathway reconstruction from the genome sequence reveals complete biosynthetic pathways for all amino acids except asparagine, glutamine,

methionine, and serine, implying that *Blattabacterium* can make most amino acids from a few substrates, such as glutamate, urea, and ammonia (Fig. 1). Genes involved in asparagine and glutamine biosynthesis are entirely absent. Previous experimental evidence suggests that *Blattabacterium* generates methionine by cysteine transsulfuration (30, 32). Genomic analysis reveals that all of the enzymes for methionine biosynthesis by this route are encoded by the genome except for the acetylhomoserine-yielding homoserine O-acetyltransferase. Alternatively, cystathionine γ -synthase produces homocysteine using cysteine and O-phosphohomoserine, a threonine biosynthesis metabolite, potentially circumventing the need for acetylhomoserine.

Although genomic evidence for serine biosynthetic capability exists, *Blattabacterium* lacks the enzymes that catalyze the final serine-yielding reactions. Alternative mechanisms for endogenous methionine and serine biosynthesis may be used, or the metabolic intermediates or the actual amino acids are imported from the host cytoplasm.

Proline biosynthesis in some insects can proceed by arginine conversion in fat bodies, but enzymes for generating arginine from citrulline are thought to be absent in insects, implying an incomplete ornithine-urea cycle (33, 34). *Blattabacterium* can make arginine from glutamate, and it encodes the ornithine-urea cycle and enzymes to produce pyrroline-5-carboxylate, a proline biosynthetic intermediate that can be reduced further by a host-provided pyrroline-5-carboxylate reductase to yield proline. *Blattabacterium* must acquire proline from the cockroach, because it lacks the enzyme for the final biosynthetic step. Proline biosynthesis by arginine degradation also produces urea and ammonia, both of which can be used by the *Blattabacterium*-encoded urease and glutamate dehydrogenase to generate glutamate, effectively conserving nitrogen. Thus, our genomic data support the hypothesis that *Blattabacterium* recycles nitrogenous wastes and provides amino acids.

Opportunities for nitrogen scavenging and recycling are abundant due to various metabolic processes in *Blattabacterium*. Cystathionine γ -synthase and threonine dehydratase both produce ammonia while generating amino acid biosynthesis intermediates. Vitamin biosynthesis and modification processes also yield ammonia during biosynthesis. Enzymes involved in degradation and salvage pathways of ribonucleosides, ribonucleotides, pyridoxal 5'-phosphate, and nicotinamide produce ammonia as a by-product. Ammonia is toxic, but glutamate dehydrogenase (GdhA) can detoxify the intracellular environment and recycle ammonia into glutamate. Bacteria typically employ GdhA for nitrogen assimilation when ammonia is abundant and ATP is limited, because nitrogen assimilation by this route does not consume ATP and thus is physiologically advantageous (35–37).

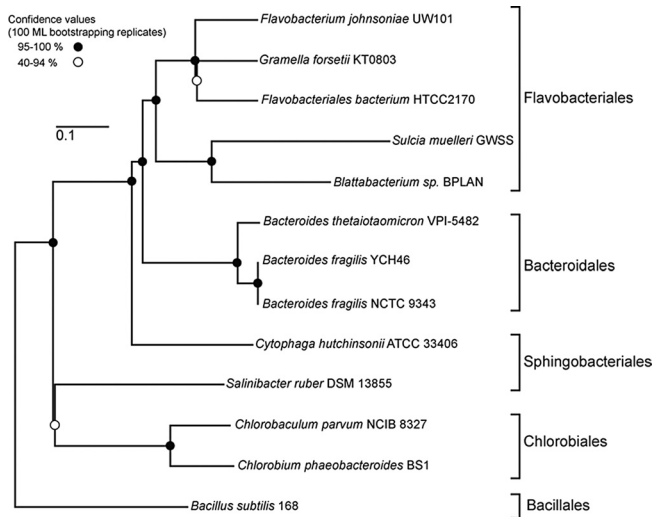


Fig. 2. Phylogenetic relatedness of *Blattabacterium* sp. BPLAN to *Cytophaga*–*Flavobacterium*–*Bacteroidetes* bacteria with completely sequenced genomes. (Scale bar, 0.1 substitutions per site.)

The hypothesis that uricolysis is carried out by *Blattabacterium* using proteins annotated as having an alternative or no known function cannot be ruled out entirely. A second possibility is that the insect host supplies uricolytic enzymes. A BLAST analysis of expressed sequences from two cockroach species (*P. americana* and *Blattella germanica*) reveals partial transcripts encoding homologues of proteins participating in uricolysis (GenBank accession nos. FG125694, FG128011, and FG132115). Xanthine dehydrogenase is variably present among animals but is found in most insects, based on genomes sequenced to date. An Xdh homologue is encoded in *B. germanica* (GenBank accession no. FG127839), based on blastx analysis of expressed sequences from *B. germanica*. Weighing against the hypothesis that the cockroach supplies the uricolytic enzymes are results showing that uricase activity is absent in termite tissues (43) and that antibiotics dramatically reduce uricolytic activity (8).

A third hypothesis is that uric acid is transported to the gut, where it is catabolized by uricolytic microorganisms. In *R. flavipes*, uric acid is transported actively from hemolymph, via the Malpighian tubules, to the hindgut, where anaerobic *Bacteroides* and *Streptococcus* degrade uric acid to ammonia (8). Cockroaches and termites are close relatives, with similar dietary preferences and excretory physiology (43). Uricolytic *Clostridium* have been isolated consistently from *P. americana* (47). Thus, members of the cockroach gut microflora can degrade uric acid (2). Similarly, antibiotic treatment of shield bugs (Insecta: Hemiptera) eliminates uricolytic bacteria from the gastric cecum and reduces uricolytic activity (5). If microbial degradation of uric acid occurs in the gut lumen of *P. americana*, then the role of *Blattabacterium* may be to use the products of uricolysis (urea and ammonia) for amino acid production.

Blattabacterium Phylogenetics and Comparative Genomics. According to maximum likelihood trees of concatenated protein alignments from completely sequenced members of Bacteroidetes/Chlorobi, *Blattabacterium* sp. BPLAN belongs to the Flavobacteriales and is related most closely to *Sulcia*, another obligate insect symbiont of (Fig. 2). These results are consistent with 16S rRNA gene analyses that included more taxa (19, 21, 22). Many genes, including some expected to enable metabolic plasticity, are eliminated from obligate endosymbiont genomes (Table S1), leaving genes essential to maintaining the host–symbiont lifestyle (i.e., amino acid biosynthesis) (48). *Blattabacterium* is typical of insect

nutritional symbionts in possessing a highly reduced genome at 637 kb; the even tinier *Sulcia* genome is one of the smallest yet discovered at 245 kb (18). Thus, *Blattabacterium*'s genome provides clues to stages of genome reduction, potentially resembling an ancestral state for the tiny *Sulcia* genome. Consistent with this view, the *Sulcia* gene set is largely a subset of that of *Blattabacterium*. Only eight *Sulcia* ORFs lack homologues in *Blattabacterium*. Five are annotated as conserved hypothetical proteins of unknown function. The other three *Sulcia*-specific ORFs encode the α and β subunits of the ferredoxin oxidoreductase complex, which is involved in carbon dioxide assimilation, and the glutamate metabolic enzyme pyrroline-5-carboxylate dehydrogenase.

Sulcia and *Blattabacterium* have similar roles within insect hosts. Both are capable of synthesizing essential amino acids, with *Blattabacterium* able to make all 10. *Sulcia* from the glassy-winged sharpshooter can make eight, excluding histidine and methionine, for which pathways are encoded by the co-resident symbiont, *Baumannia cicadellincola* (18).

Comparisons of numbers of genes assigned to categories based on functional roles are also consistent with the view that *Blattabacterium* represents an intermediate stage between larger genomes of free-living bacteria, represented within the Bacteroidetes by *Flavobacterium johnsonii*, and the extremely tiny genomes of some obligate symbionts, represented by *Sulcia*. *Blattabacterium* and *F. johnsonii* show highly significant differences in distributions of genes among functional categories ($\chi^2 = 330$, $df = 13$, $P < 0.00001$). In *Blattabacterium*, genes involved in transcriptional regulation, signal transduction, secondary metabolites, and defensive activities are disproportionately lost, whereas those involved in translation, amino acid metabolism, and coenzyme metabolism are retained more often, a pattern resembling that noted previously for symbiotic Gammaproteobacteria (25, 48). *Blattabacterium* and *Sulcia* also differ in representation of different functional categories ($\chi^2 = 49$, $df = 13$, $P < 0.00001$), and this comparison illuminates the patterns of gene loss during the transition from small (*Blattabacterium*) to tiny (*Sulcia*) genomes. *Sulcia* has retained differentially genes in the categories corresponding to translation and amino acid biosynthesis and has differentially eliminated genes in the categories of unknown function, biosynthesis of lipids, and cell wall components.

Loss of genes determining cell size and shape are likely to contribute to morphological changes associated with the most extreme stages of genome reduction. *Sulcia*'s small genome does not encode the rod-shape-determining proteins MreB and MreC or the pathways for producing a cell wall, and these losses correlate with irregular cell morphology (Fig. 3A). This cell morphology and lack of genes for synthesis of cell wall components also are observed in *Carsonella ruddii* (Gammaproteobacteria) and *Hodgkinia cicadicola* (Alphaproteobacteria), two unrelated symbionts with even tinier genomes (145 and 160 kb) (49, 50). Microscopy confirms that *Blattabacterium* alone inhabits cockroach fat bodies, where it populates the cytoplasm (Fig. 3C). Additionally, it has a well-defined, rod-shaped cellular morphology and an intact cell wall (Fig. 3B) (13, 51), consistent with the presence of MreB and MreC and a complete peptidoglycan biosynthetic pathway.

An analysis of the gene order of homologous proteins indicates little conservation of gene order between *Sulcia* and *Blattabacterium* (Fig. S2). Thus, one or both of these lineages continued to undergo extensive chromosomal rearrangements after their divergence.

Conclusions

The *Blattabacterium* genome confirms and elucidates its central role in nitrogen metabolism within the cockroach system, including nitrogen recycling from ammonia and urea and extensive capabilities for amino acid provisioning. These metabolic capabilities are likely to have been important in allowing cockroaches

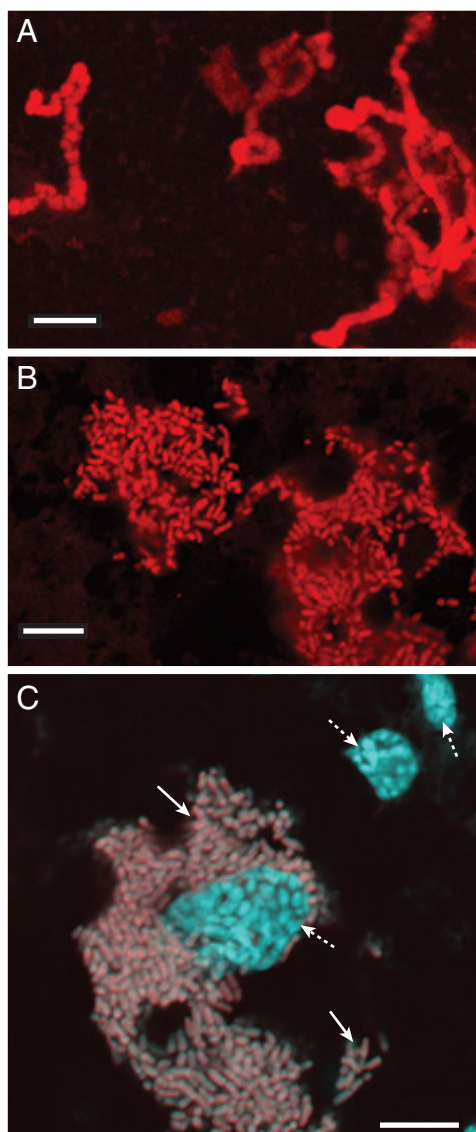


Fig. 3. Comparison of the cell morphology of *Blattabacterium* sp. BPLAN and *Sulcia muelleri*. Confocal micrographs of Cy3-labeled *S. muelleri* (A) and *Blattabacterium* sp. BPLAN (B). (C) A FISH micrograph of *Blattabacterium* sp. BPLAN within *Periplaneta americana*. A composite image of *P. americana* abdominal tissue hybridized with a Cy3-labeled *Blattabacterium*-specific probe and counterstained with DAPI. Dashed arrows, insect nuclei; solid arrows, *Blattabacterium* sp. BPLAN cells. (Scale bar, 10 μ m.)

to expand their ecological niche by exploiting food sources in which nitrogen sources are variable and often limited. The first step in the utilization of uric acid as a nitrogen storage compound appears to be provided not by *Blattabacterium*, which lacks a uricase homologue, but by either the insect host or, as known to occur in termites, bacteria in the gut lumen. Previous hypotheses suggesting that *Blattabacterium* is capable of uricolysis are not supported by the genomic evidence, raising the question of why *Blattabacterium* is located in such close proximity to uric acid deposits in fat bodies. Possibly, a *Blattabacterium* ancestor was uricolytic but, in the course of genome reduction, lost its uricase, transferring this function to uric acid degraders among the gut microflora. *Blattabacterium* resembles other insect nutritional symbionts in having lost almost all specific transporters and regulatory genes but is unique for insect symbionts in retaining the transcriptional regulator RpoN.

Materials and Methods

DNA Preparation. Whole fat bodies were dissected from five adult *P. americana* (Carolina Biological Supply) in 1 \times PBS and lightly homogenized by hand in 200 μ L of 1 \times PBS. Homogenate was passed through 20- and 11- μ m filters and centrifuged at 8,000 \times g at 4 $^{\circ}$ C for 10 min. The pellet was resuspended in 200 μ L of Buffer AL from the Qiagen DNEasy kit, and DNA was prepared according to the manufacturer's protocols. Extracted DNA was suspended in 1 \times PBS to a concentration of 104 ng/ μ L. Quantitative PCR using the bacterial 16S rRNA and host EF1- α genes was performed to determine the ratio of *Blattabacterium*-to-cockroach total DNA. The final ratio was \approx 610:1. We confirmed that *Blattabacterium* was the only bacterium represented in the sample by sequencing amplicons obtained from a PCR using universal 16S rDNA bacterial primers. Five micrograms of total DNA was submitted to the University of South Carolina and sequenced using a Roche/454 GS-FLX 100.

Genome Assembly. Two sequencing runs yielded 225,047 reads with an average read length of 255 nt. Sequence reads were assembled into 11,737 contigs by the Newbler assembler, version 1.1.03, and visualized with Consed (52) and SeqMan (DNASTAR). Fourteen contigs were distinct in having greater length and greater read average depth of coverage. Thirteen contigs, ranging from 1,019 to 146,043 nt and together comprising 110,232 reads (49% of all reads), were ordered, using "-to" and "-from" information appended to the trace filenames by the Newbler assembler, into a putative circular genome of 636,578 nt with an average 28.4% G + C content and average read depth of 44.2 \times . Sanger sequencing across gaps between the 13 contigs enabled assembly of a single circular contig of 636,994 nt with a 28.2% G + C content. The remaining 3,249 nt contig displayed no significant match to any other contig over its entire length and had an average read depth of coverage of 78.1 \times , which was almost 2 \times that of the genome. ORF Finder was used to search for ORFs, followed by blastp search of the putative ORFs against the GenBank nr database, yielding regions encoding four proteins with high similarity to bacterial proteins. Given the read depth, its lack of assembly with the rest of the genome, and the presence of bacterial genes, we hypothesized that this fragment corresponded to a plasmid. By designing outward-oriented PCR primers based on the sequences at the ends of this contig, we obtained a single product that linked the ends of this fragment and confirmed a plasmid.

Genome Annotation. Prodigal was used to predict protein-coding regions. Functional annotation of the coding regions proceeded by combining results of blastp searches against the GenBank nr (downloaded November 20, 2008) and COG myva (53) databases and hmmpfam (HMMER, version 2.3.2) searches against TIGRFAM, version 8.0 (54), and Pfam, version 23.0 (55), databases. The 16S rRNA coding operon was identified by analyzing intergenic regions by blastn searches against GenBank. The blastx searches of the intergenic regions against GenBank nr did not yield additional protein-coding regions. The tRNAs were identified by tRNAscan-SE, version 1.23 (56), using the bacterial model. The tmRNAs were identified using BRUCE (57) and confirmed by blastn search against GenBank. The origin of replication was predicted using originx (58). Metabolic pathway prediction was performed using Pathologic within Pathway Tools, version 13.0 (59). Circular chromosome representation was constructed using DNAPlotter, version 1.1 (60). Promoter prediction was performed using the promscan Perl script (61) to score 16-mers using a weight matrix generated from 186 RpoN-dependent promoters (41).

Comparative Genomic Analyses. The *Blattabacterium* genome was compared with that of *F. johnsonii*, a related free-living bacterium, and to *S. muelleri*, a related obligately symbiotic bacterium with a genome of 225 kb. Genes were assigned to COG functional categories (53) for each genome (Table S1), and significance of differences was evaluated using χ^2 goodness-of-fit tests for the pairwise comparisons of *Flavobacterium-Blattabacterium* and *Blattabacterium-Sulcia*. To reduce the number of cells with low expected values, we combined categories corresponding to poorly characterized genes (general predicted function and unknown function) and defense and secondary metabolites, and we eliminated remaining categories in which the expected number of genes was <100 for the largest (*Flavobacterium johnsonii*) genome. This resulted in 14 functional categories, corresponding to COGs C, E, G-M, O, P, T, R + S, and Q + V. Two-tailed tests were performed. Gene synteny between the *Sulcia* and *Blattabacterium* genomes was evaluated by plotting the genomic locations of *Sulcia* genes against the locations of their *Blattabacterium* homologues.

Phylogenetic Analysis. We used MAFFT, version 6.624b, with the L-INS-i algorithm (62), to align individual protein datasets. Customized Perl scripts were used to remove columns within the alignments containing gap characters and

concatenate all of the aligned protein sequence regions before phylogenetic analysis. A maximum likelihood tree, based on 100 bootstrap replications, was constructed using PhyML, version 2.4.4, with the JTT model of amino acid substitution (63). The proteins used for the alignments and the genomes from which they were obtained can be found in *SI Text*.

Fluorescence in Situ Hybridization. A third-instar *P. americana* specimen was euthanized and then beheaded to facilitate reagent infiltration. The specimen was fixed in 4% paraformaldehyde and embedded in paraffin before sectioning (64). In situ hybridization was performed on 4- μ m sagittal tissue sections affixed to microscope slides with a cyanine 3 fluorophore (Cy3)-labeled *Blattabacterium*-specific oligonucleotide probe, BLATTA303 (CCAGT-

GTGGGGGATCAC). 4',6-Diamidino-2-phenylindole was used as a counterstain. *Sulcia muelleri* cells were obtained from *Clastoptera arizonana* bacteriomes and hybridized with a Cy3-labeled *S. muelleri*-specific oligonucleotide probe, SULCIA303 (CCAATGTGGGGWACGC). Protocols for tissue preparation and FISH hybridization were adapted from ref. 65. Images were obtained using a Zeiss LSM 501 META laser confocal microscope.

ACKNOWLEDGMENTS. We thank Joe Jones at the University of South Carolina for 454 sequencing, Sheena Brown for cockroach specimens, and John McCutcheon for technical assistance. This work was funded by the National Science Foundation Grant 0626716 (to N.A.M.), and the University of Arizona Center for Insect Science through National Institutes of Health Training Grant 2 K12 GM000708 (to Z.L.S.).

- Bell WJ, Roth LM, Nalepa CA (2007) *Cockroaches: Ecology, Behavior, and Natural History*. (Johns Hopkins Univ Press, Baltimore).
- Mullins DE, Cochran DG (1975) Nitrogen metabolism in American cockroach. I. An examination of positive nitrogen-balance with respect to uric-acid stores. *Comp Biochem Physiol* 50:489–500.
- Mullins DE, Cochran DG (1974) Nitrogen metabolism in the American cockroach: An examination of whole body and fat body regulation of cations in response to nitrogen balance. *J Exp Biol* 61:557–570.
- Cochran DG (1985) Nitrogen excretion in cockroaches. *Annu Rev Entomol* 30:29–49.
- Kashima T, Nakamura T, Tojo S (2006) Uric acid recycling in the shield bug, *Parastrachia japonensis* (Hemiptera: Parastrachiidae) during diapause. *J Insect Physiol* 52:816–825.
- Potrikus CJ, Breznak JA (1980) Anaerobic degradation of uric-acid by gut bacteria of termites. *Appl Environ Microbiol* 40:125–132.
- Potrikus CJ, Breznak JA (1980) Uric acid-degrading bacteria in guts of termites [*Reticulitermes flavipes* (Kollar)]. *Appl Environ Microbiol* 40:117–124.
- Potrikus CJ, Breznak JA (1981) Gut bacteria recycle uric-acid nitrogen in termites—a strategy for nutrient conservation. *Proc Natl Acad Sci USA* 78:4601–4605.
- Gier HT (1936) The morphology and behavior of the intracellular bacteroids of roaches. *Biol Bull* 71:433–452.
- Donnella JF, Kilby BA (1967) Uric acid metabolism by symbiotic bacteria from the fat body of *Periplaneta americana*. *Comp Biochem Physiol* 22:235–252.
- Lanham UN (1968) The Blochmann bodies: Hereditary intracellular symbionts of insects. *Biol Rev Camb Philos Soc* 43:269–289.
- Blochmann F (1887) Occurrence of bacteria-like particles in tissues and eggs of different insects. *Biol Zentralbl* 7:606–608 (German).
- Brooks MA (1970) Comments on classification of intracellular symbionts of cockroaches and description of the species. *J Invertebr Pathol* 16:249–258.
- Brooks MA, Richards AG (1955) Intracellular symbiosis in cockroaches. I. Production of aposymbiotic cockroaches. *Biol Bull* 109:22–39.
- Malke H (1964) Production of aposymbiotic cockroaches by means of lysozyme. *Nature* 204:1223–1224.
- Henry SM (1962) The significance of microorganisms in nutrition of insects. *Trans NY Acad Sci* 24:676–683.
- Bandi C, et al. (1995) The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proc R Soc London Ser B* 259:293–299.
- McCutcheon JP, Moran NA (2007) Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc Natl Acad Sci USA* 104:19392–19397.
- Gruwell ME, Morse GE, Normark BB (2007) Phylogenetic congruence of armored scale insects (Hemiptera: Diaspididae) and their primary endosymbionts from the phylum Bacteroidetes. *Mol Phylogenet Evol* 44:267–280.
- Majerus TMO, et al. (1998) Extreme variation in the prevalence of inherited male-killing microorganisms between three populations of *Harmonia axyridis* (Coleoptera: Coccinellidae). *Heredity* 81:683–691.
- Clark JW, Hossain S, Burnside CA, Kambhampati S (2001) Coevolution between a cockroach and its bacterial endosymbiont: A biogeographical perspective. *Proc R Soc London Ser B* 268:393–398.
- Lo N, Bandi C, Watanabe H, Nalepa C, Beninati T (2003) Evidence for cladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. *Mol Biol Evol* 20:907–913.
- Brooks MA, Richards K (1966) On the *in vitro* culture of intracellular symbionts of cockroaches. *J Invertebr Pathol* 8:150–157.
- Ochman H, Moran NA (2001) Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science* 292:1096–1099.
- Wernegreen JJ (2005) For better or worse: Genomic consequences of intracellular mutualism and parasitism. *Curr Opin Genet Dev* 15:572–583.
- Grose JH, et al. (2005) Assimilation of nicotinamide mononucleotide requires periplasmic AphA phosphatase in *Salmonella enterica*. *J Bacteriol* 187:4521–4530.
- Lee PA, Tullman-Ercek D, Georgiou G (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* 60:373–395.
- Kuhn A (2009) From the Sec complex to the membrane insertase YidC. *Biol Chem* 390:701–706.
- Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* 407:81–86.
- Block RJ, Henry SM (1961) Metabolism of sulphur amino-acids and of sulphate in *Blattella germanica*. *Nature* 191:392–393.
- Douglas AE (1989) Mycetocyte symbiosis in insects. *Biol Rev Camb Philos Soc* 64:409–434.
- Henry SM, Block RJ (1961) Sulfur metabolism of insects. 6. Metabolism of sulfur amino acids and related compounds in German cockroach, *Blattella germanica* (L.). *Contrib Boyce Thompson Inst* 21:129–145.
- Inokuchi T, Horie Y, Ito T (1969) Urea cycle in silkworm *Bombyx mori*. *Biochem Biophys Res Commun* 35:783–787.
- Reddy SR, Campbell JW (1977) Enzymic basis for the nutritional requirement of arginine in insects. *Experientia* 33:160–161.
- Helling RB (1994) Why does *Escherichia coli* have two primary pathways for synthesis of glutamate? *J Bacteriol* 176:4664–4668.
- Helling RB (1998) Pathway choice in glutamate synthesis in *Escherichia coli*. *J Bacteriol* 180:4571–4575.
- Reitzer L (2003) Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu Rev Microbiol* 57:155–176.
- Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD (2000) The bacterial enhancer-dependent σ^{54} (σ^N) transcription factor. *J Bacteriol* 182:4129–4136.
- Szabo A, et al. (1994) The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system DnaK, DnaJ, and GrpE. *Proc Natl Acad Sci USA* 91:10345–10349.
- Horwich AL, Farr GW, Fenton WA (2006) GroEL–GroES-mediated protein folding. *Chem Rev* 106:1917–1930.
- Dombrecht B, Marchal K, Vanderleyden J, Michiels J (2002) Prediction and overview of the RpoN-regulon in closely related species of the *Rhizobiales*. *Genome Biol* 3:RESEARCH0076.
- Skibinski DAG, et al. (2002) Regulation of the hydrogenase-4 operon of *Escherichia coli* by the σ^{54} -dependent transcriptional activators FhIA and HyfR. *J Bacteriol* 184:6642–6653.
- Potrikus CJ, Breznak JA (1980) Uric acid in wood-eating termites. *Insect Biochem* 10:19–27.
- Wu XW, Lee CC, Muzny DM, Caskey CT (1989) Urate oxidase: Primary structure and evolutionary implications. *Proc Natl Acad Sci USA* 86:9412–9416.
- Ramazzina I, Folli C, Secchi A, Berni R, Percudani R (2006) Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat Chem Biol* 2:144–148.
- Lee Y, et al. Transthyretin-related proteins function to facilitate the hydrolysis of 5-hydroxyisourate, the end product of the uricase reaction. *FEBS Lett* 579:4769–4774.
- Cruden DL, Markovetz AJ (1987) Microbial ecology of the cockroach gut. *Annu Rev Microbiol* 41:617–643.
- Moran NA (2002) Microbial minimalism: Genome reduction in bacterial pathogens. *Cell* 108:583–586.
- Nakabachi A, et al. (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science* 314:267.
- McCutcheon JP, McDonald BR, Moran NA (2009) Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc Natl Acad Sci USA* 106:15394–15399.
- Radek R (2000) Light and electron microscopic study of a *Rickettsiella* species from the cockroach *Blattella orientalis*. *J Invertebr Pathol* 76:249–256.
- Gordon D, Abajian C, Green P (1998) *Consed*: A graphical tool for sequence finishing. *Genome Res* 8:195–202.
- Tatusov RL, et al. (2003) The COG database: An updated version includes eukaryotes. *BMC Bioinformatics* 4:41.
- Haft DH, et al. (2001) TIGRFAMs: A protein family resource for the functional identification of proteins. *Nucleic Acids Res* 29:41–43.
- Finn RD, et al. (2008) The Pfam protein families database. *Nucleic Acids Res* 36:D281–D288.
- Lowe TM, Eddy SR (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964.
- Laslett D, Canback B, Andersson S (2002) BRUCE: A program for the detection of transfer-messenger RNA genes in nucleotide sequences. *Nucleic Acids Res* 30:3449–3453.
- Worning P, Jensen LJ, Hallin PF, Staerfeldt HH, Ussery DW (2006) Origin of replication in circular prokaryotic chromosomes. *Environ Microbiol* 8:353–361.
- Keseler IM, et al. (2009) EcoCyc: A comprehensive view of *Escherichia coli* biology. *Nucleic Acids Res* 37:D464–D470.
- Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: Circular and linear interactive genome visualization. *Bioinformatics* 25:119–120.
- Studholme DJ, Dixon R (2003) Domain architectures of σ^{54} -dependent transcriptional activators. *J Bacteriol* 185:1757–1767.
- Katoh K, Asimeno G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* 537:39–64.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704.
- Barbosa P (1974) *Manual of Basic Techniques in Insect Histology* (Autumn Publishers, Amherst, MA).
- Daims H, Stoecker K, Wagner M (2005) In *Molecular Microbial Ecology*, eds Osborn AM, Smith C (BIOS Scientific Publications, London), pp 213–240.