PLOS ONE

RESEARCH ARTICLE

Genome analysis of new *Blattabacterium* spp., obligatory endosymbionts of *Periplaneta fuliginosa* and *P. japonica*

Cláudia S. L. Vicente[1,2☯], Shakhinur Islam Mondal[3,4☯], Arzuba Akter[3,5☯], Sota Ozawa[2], Tasei Kikuchi[3], Koichi Hasegawa[2]*

1 NemaLab/ICAAM—Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Departamento de Biologia, Universidade de Évora, Núcleo da Mitra, Évora, Portugal, **2** Department of Environmental Biology, College of Bioscience & Biotechnology, Chubu University, Matsumoto, Kasugai, Aichi, Japan, **3** Division of Parasitology, Faculty of Medicine, University of Miyazaki, Miyazaki, Japan, **4** Genetic Engineering and Biotechnology Department, Shahjalal University of Science and Technology, Kumargaon, Sylhet, Bangladesh, **5** Biochemistry and Molecular Biology Department, Shahjalal University of Science and Technology, Kumargaon, Sylhet, Bangladesh

☯ These authors contributed equally to this work.
* koichihasegawa@isc.chubu.ac.jp

Abstract

The successful adaptation of cockroaches is, in part, dependent of the activity of their obligatory endosymbionts, *Blattabacterium* spp., which are involved in uric acid degradation, nitrogen assimilation and nutrient provisioning. Their strategic localization, within bacteriocytes in the proximities of uric acid storage cells (urocytes), highlights their importance in the recycling of nitrogen from urea and ammonia, end-products not secreted by their host insects. In this study, we present the complete genome sequence of two new *Blattabacterium* spp. from *Periplaneta fuliginosa* (BPfu) and *P. japonica* (BPja), and detailed comparison with other *Blattabacterium* strains from different cockroach species. The genomes of BPfu and BPja show a high degree of stability as showed with for other *Blattabacterium* representatives, only presenting a 19-kb fragment inversion between BPja and BPfu. In fact, the phylogenomics showed BPja as an ancestor species of BPfu, BPLAN (*P. americana*) and BBor (*Blatta orientalis*), in congruence with their host cockroach phylogeny. Their functional profile is similar and closest to the omnivorous strain BBge (*Blattella germanica*). Interesting, BPja possesses the complete set of enzymes involved sulfate assimilatory pathway only found in BBge and BMda (*Mastotermes darwiniensis*). The newly sequenced genomes of BPja and BPfu emphasise the remarkable stability of *Blattabacterium* genomes supported by their long-term coevolution and obligatory lifestyle in their host insect.

Introduction

Considered "living-fossils", cockroaches (order Blattodea) are one of the most primitive insect group, dating back as far as the Late Carboniferous (320 Mya) [1]. They can inhabit a wide range of ecosystems, from tropical forests to deserts, coping with extremes and food scarcity

by using varied combinations of movement, habitat choice, physiological mechanisms and life-cycle strategies [2]. Cockroaches are mostly considered omnivores or generalists, and are able to select their diet according to their nutrient demands at every stage in their lifecycle [2]. Although set with sophisticated physiological mechanisms to resist and prevail under stressful conditions (more detail in 1), much of these insect's success is correlated with the close relationships established with their microbiome.

The microbiome of insects comprehends in a consortium of microbial communities colonizing the external and internal body structures, and which are involved in a wide range of important lifestyle functions such as colonization and resistance to parasites and/or pathogens, diet breakdown, nutrient recycling and production of pheromones and/or kairomones [3]. The diversity of these communities is host-depended and is mainly determined by its habitat, diet, developmental stage and phylogeny [4]. Apart from these microbial inhabitants, primary endosymbionts are particularly interesting since they live solely within specialized host cells, undergo vertical transmission to offspring and contribute extensively for host health and development [5]. These endosymbionts include the most highly constrained, stable and specialized symbioses known in the animal world [5]. Cockroaches harbour the endosymbiont *Blattabacterium* (phylum Bacteriodetes, class Flavobacteria) in their abdominal fat body, inside specialized cells (mycetocytes or bacteriocytes), which are adjacent to urocytes (cells that store uric acid) [6]. Early studies showed that aposymbiotic cockroaches (cockroaches reared with antibiotic-containing food) were less fitness that normal cockroaches, showing a smaller size, light color, reduced fecundity and increased development time [7]. These observations emphasised the importance of the endosymbionts to their roach host. *Blattabacterium* spp. are involved in nitrogen recycling from ammonia and urea, and subsequently the provision of essential amino acids and vitamins [8]. In fact, the location of the endosymbionts suggests dynamic metabolic interaction between the bacteriocytes and the urocytes [9].

The host-dependent lifestyle of *Blattabacterium* has effect on their genome composition, size and structure. Till date, several genome sequences of *Blattabacterium* have been annotated from different cockroaches and their phylogenetic closest termites: *Blatella germanica*, BBge [10]; *Periplaneta americana*, BPLAN [8]; *Cryptocercus punctulatus*, BCpu [11]; *Blaberus giganteus*, BGIGA [12]; *Mastotermes darwiniensis*, BMda [13]; *Nauphoeta cinerea*, BNCIN [14]; *Blatta orientalis*, BBor [15]; and *Panesthia angustipennis*, BPAA [16]. Here, we present the complete genomes of two *Blattabacterium* spp. isolated from the cosmopolitan *P. fuliginosa* and the endemic *P. japonica*, both cockroaches considered serious pests in Japan with contrasting environmental adaptations, and provide insights into the evolutionary stability of the endosymbionts genomic and metabolic architecture.

Materials and methods

Cockroach strains and rearing

Periplaneta fuliginosa EE and *P. japonica* Miyoshi strains are established in the Hasegawa Laboratory (Chubu University, Japan) since 2013, under the conditions described by Vicente et al. [17].

Total DNA extraction from fat body tissue and quantification of *Blattabacterium*

One adult male of each cockroach strain was surface-cleaned with 70% EtOH and dissected using sterile forceps and tweezers. The fat body tissue was carefully removed and weighed (*P. fuliginosa* EE fat body weighing = 181.9mg; *P. japonica* Miyoshi fat body weighing = 118.8mg).

Total DNA extraction of fat body tissue was performed using DNeasy Blood & Tissue kit (Qiagen, USA, California) with some modifications to the original manufacturer's instructions. After weighting, the fat body was transferred into ATL buffer, homogenized with a sterile hand-held homogenizer, and incubated for 1h at 55ºC with 20μL of Proteinase K solution (540 units/mL, Wako). The homogenate was filtered through a 0.22μM diameter filter (Millipore) and the filtrate used for DNA extraction (next steps in the kit protocol). The quality and quantity of the DNA samples was measured using NanoVue plus spectrophotometer (GE Healthcare Life Science, USA). DNA samples were stored at -20ºC.

Host DNA was extracted from the muscle tissue of each cockroach strain using DNeasy Blood & Tissue kit (Qiagen, USA, California) following the manufacturer's instructions. The quality and quantity was assessed as above mentioned. The quantification of *Blattabacterium sp.* in the fat body of each cockroach strain (BPful, *Blattabacterium sp.* from *P. fuliginosa*; and BPjap, *Blattabacterium sp.* from *P. japonica*) was inferred by the ratio between the DNA host sample and the respective DNA endosymbiont sample. For both, a standard curve (log DNA concentration plotted against Cq-values) was determined by quantitative real-time PCR (qPCR) using single copy genes, urease *ureA* (for *Blattabacterium sp.*) and wingless *wg* (for *Periplaneta sp.*) [18]. The DNA concentrations tested ranged between 0.63-10ng/μL for the endosymbiont, and 2-50ng /μL for the host. *ureA* primers were designed based on the multiple sequence alignment (MSA) of *ureA* gene of all *Blattabacterium spp.* with known genomes (Table 1): universal *ureA*_for (5'-ATG CAY TTA AMT TYT TAT GAA-3') and universal *ureA*_rev (5'-TCA TAT WKY YRT ATT YTC TYT TWC C-3'). MSA was conducted with default parameters in BioEdit version 7.2.5 using ClustalW. *wg* primers designed for *Periplaneta genus* were: *wg*_for (5'-TGG TCT ACT TGG AGC CTT CC-3') and *wg*_rev (5'-ATC CAC GCC TAT CGA CGT AT-3'). qPCR was performed using CFX96TM Real-Time (Bio-Rad), and SYBR Premix Ex TaqTM II (Tli RNAse H Plus) kit (Takara Bio Inc., Japan) following the manufacturer's indications. The thermal cycling conditions were: initial denaturation at 95ºC for 30 secs; 39 cycles of denaturation at 95ºC for sec, annealing and extension at 60ºC for 30 secs; followed by the melting curve. For accuracy, the standard curve was three

Table 1. Characterization of BPfu and BPja genomes in comparison with other published *Blattabacterium* strains.

Strain	BPfu	BPju	BNCIN	BGIGA	BBge	BPLAN	BBor	BPAA	BCpu	Bmda
Genome size (bp)	645 082	636 644	626 627	632 588	640 335	640 442	638 184	632 490	609 561	590 554
Plasmids	1	1	1	1	1	1	1	*	1	1
Plasmid size (bp)	4127	3781	3675	3423	3485	3448	3735	*	3816	3306
Chromosome size (bp)	640 955	632 863	622 952	629 165	636 850	636 994	634 449	*	605 745	587 248
G + C content (%)	28.1/30.8	27.6/29.7	26.2/20.6	25.7/30.9	27.1/29.8	28.2/28.5	28.2/30.6	26.4/-*	23.8/30.5	27.5/31.9
Total number of genes	640	623	627	616	631	634	627	624	589	597
Total CDS	590 + 6	572 + 6	581 + 5	573 + 4	586 + 4	587 + 4	572 + 7	575*	545 + 3	544 + 4
rRNAs	3	3	3	3	3	3	3	3	3	3
tRNAs	34	34	32	34	34	33	33	34	32	34
Other ncRNAs	1	0	1	1	3	1	3	3	3	3
Pseudogenes	6	8	5	1	1	6	9	9	3	9

Host species abbreviations and accession numbers: BPfu, *Periplaneta fuliginosa*; BPja, *Periplaneta japonica*; BNCIN, *Nauphoeta cinereal* (NC_022550.1); BGIGA, *Blaberus giganteus* (NC_017924.1); BBge, *Blattella germanica* (NC_013454.1); BPLAN, *Periplaneta americana* (NC_013418.2); BBor, *Blatta orientalis* (NC_020195.1); BPAA, *Panesthia angustipennis* (NC_020510.1); BCpu, *Cryptocercus punctulatus* (NC_016621.1); BMda, *Mastotermes darwiniensis* (NC_016146.1). The following features result from the analysis of both bacterial chromosome and plasmids (if present): genome size, number of base pairs, total number of genes, G+C content and total CDS. The asterisk (*) indicates no information given by the authors of the original study [16].

https://doi.org/10.1371/journal.pone.0200512.t001

times repeated and the efficiency calculated using the equation (1) PCR efficiency = $(10^{(-1/S)}-1)$ × 100, where S is the slope of the standard curve.

Sequencing, assembly and annotation

Paired-end sequencing libraries were prepared from 100 ng of endosymbiont enriched DNA using the Nextera DNA Library Prep kit (Illumina) according to the manufacturer's instructions. Libraries were sequenced on the Illumina MiSeq sequencer with the v3 kit (301 cycles x 2). The raw sequence data were analysed using the RTA 1.12.4.2 analysis pipeline and were used for genome assembly after removal of adapters, low quality, and duplicate reads to produce a total of 2.5 and 4.2 Gb sequence data for *Blattabacterium* from *P. fuliginosa*; and *P. japonica*, respectively. Illumina sequence reads were assembled using the Platanus assembler version 1.2.4 [19] with options (assembly; -k 91, scaffolding and gapclosing; default parameters). The Platanus assemblies were further improved using MITObim and manually curated using gap5 [20]. The assembled genomes were annotated using the Prokaryotic Genome Annotation System (PROKKA) [21] and the Rapid Annotations using Subsystems Technology (RAST) server [22]. The annotations were, then, manually curated using *in silico* Molecular Cloning software (*In-silico* biology, https://www.insilicobiology.jp). Raw sequences, genome assembly and gene annotation were deposited in DDBJ (DNA Data Bank of Japan) under Bio-Project PRJDB6862 (*Blattabacterium* sp. *P. japonica*, SAMD00114544 and DRX119986; and *Blattabacterium* sp. *P. fuliginosa*, SAMD00114545 and DRX119985).

Functional categorization and pathway reconstruction

The protein-coding genes in the genome were subjected to COG analysis using WebMGA [23] and IMG [24]. The COG profile was displayed using the 'Heatmapper' web tool [25]. KEGG website, BlastKOALA and KEGG Mapper, was used for metabolic pathway reconstruction [26]. Re-examination and verification of the pathways were performed according to the pathway descriptions in the EcoCyc [27] and MetaCyc [28] databases.

Results and discussion

To ensure a complete genome coverage of both BPful and BPjap, an estimation of ratio host: endosymbiont DNA from the total fat body DNA was determined using independent standard curves for host (*wg* gene) and endosymbiont (*ureA* gene) (S1 Fig). The ratio of *P. japonica*: BPjap was 1:3 and *P. fuliginosa*: BPful was 1:1.8. These ratios indicated that both endosymbionts genomes were 3x BPjap and 1.18x BPful represented in total host DNA.

Genome characteristics

The genomes of primary endosymbionts reveal a high degree of stability as a result of convergent patterns of evolution such as genome reduction and AT-richness [5]. The genomes of the newly sequenced *Blattabacterium* genomes from *P. fuliginosa* (BPfu) and *P. japonica* (BPja) also show the same trend in comparison with other *Blattabacterium* (Table 1). The BPfu genome is 645,082 bp in length and is composed of a 640,955 bp chromosome and a 4,127 bp plasmid, with a GC content of 28.1% and 30.8%, respectively. The BPfu chromosome encodes 634 protein-coding genes, including 6 pseudogenes, 3 rRNA genes, 1 ncRNA gene, and 34 tRNA genes assigning all 20 standard amino acids (S2 Table). The pseudogenes were *metG* (methionine-tRNA ligase), *lolD* (outer membrane-specific lipoprotein transporter subunit), *mvaK* (mevalonate kinase), *ygaD* (ABC transporter ATP-binding protein) and two genes coding for hypothetical proteins, and are involved in protein synthesis, lipoprotein translocation,

isoprenoid/sterol synthesis and multidrug export respectively. The BPja genome is 636,644 bp in length consisting of a 632,863 bp chromosome and a 3,781 bp plasmid. The GC content of chromosome and plasmid is 27.6% and 29.7%, respectively. The BPja chromosome encodes 618 protein-coding genes, distributed as follows: 8 pseudogenes, 3 rRNA genes, and 34 tRNA genes assigning all 20 standard amino acids (S3 Table). The BPja pseudogenes were *gltX* (gluta-myl-tRNA synthetase), *atpG* (ATP synthase, gamma subunit), *lolD*, *dnaN* (DNA polymerase III), *resP* (resolvase) and three genes coding for hypothetical proteins, which are related, respectively, with protein synthesis, ATP synthesis, lipoprotein translocation, DNA synthesis and protein degradation.

Both BPfu and BPja plasmids, named pBPFU1 and pBPJA1, respectively, encode 6 protein-coding genes: *nrdB* (ribonucleotide-diphosphate reductase subunit beta), *dut* (deoxyuridine triphosphatase), and four genes coding for hypothetical proteins, which belong to metabolic pathways for DNA replication and nucleotide biosynthesis. The same genes were also identi-fied in BNCIN [14]. In general, the number of genes in the *Blattabacterium* plasmids varied between 3 to 7 and also share a high identity between *Blattabacterium* from different host spe-cies [14].

Phylogenomics and gene synteny

The phylogenetic analysis between BPfu and BPja and other endosymbiont representatives of phylum Bacteroidetes is presented in Fig 1. Within Flavobacteriales order, all *Blattabacterium* strains are grouped in the same clade and organized in different three sub-clades. One sub-clade groups BBge and BNCIN, BGIGA and BPAA. Another sub-clade is composed by the *Blattabacterium* strains from wood-feeding cockroach *C. punctulatus* (BCpu) and primitive termite *M. darwiniensis* (BMda). The third and monophyletic sub-clade groups BPja, BPfu, BPLAN and BBor, all *Blattabacterium* strains from omnivorous and worldwide pest cock-roaches. The phylogenies of endosymbionts and their hosts are generally congruent evidenced by their long-term evolution [10, 29–30]. In this study, the topology of this third sub-clade shows BPja as an ancestor species of BPfu, BPLAN and BBor, which was also seen in the topol-ogy of their host cockroach [31].

The intragenomic changes are quite rare in most primary endosymbionts [5, 8, 15]. The synteny between all sequenced *Blattabacterium* genomes (Table 1) is presented in Fig 2. The newly sequenced BPja and BPfu genomes reinforce the remarkable stability, in terms of genome structure and content, of the other *Blattabacterium* genomes. Only the *Blattabacter-ium* BCpu and BMda show more divergence compare to other sequenced *Blattabacterium* genomes [11, 13, 15]. Within the sub-clade II, one inversion occurred, a 19-kb fragment between BPja and BPfu (Fig 2). In clade III, another inversion occurred between BBor and BCpu (20-kb fragment), already denoted by other authors [11, 13, 15] (Fig 2). Other recombi-nation events, such as the presence of large and repetitive intergenic regions observed in the obligate endosymbiont *Portiera* from whitefly host (*Trialeurodes vaporarium*) [32], were absented suggesting the general model for genome evolution of gene inactivation and subse-quent loss [33–34].

Functional profile and metabolic pathways

Selective pressure across insect hosts and historical contingencies are responsible by specific changes in gene content and pathways among closest endosymbiont strains [35]. The func-tional profile of BPja and BPful and the other *Blattabacterium* strains, categorized in COG, are presented in Fig 3 and S4 Table. Notwithstanding, the clustering diagram with the *Capnocyto-phaga canimorsu* as outgroup, shows the functional profile of BPfu and BPja are closer to the
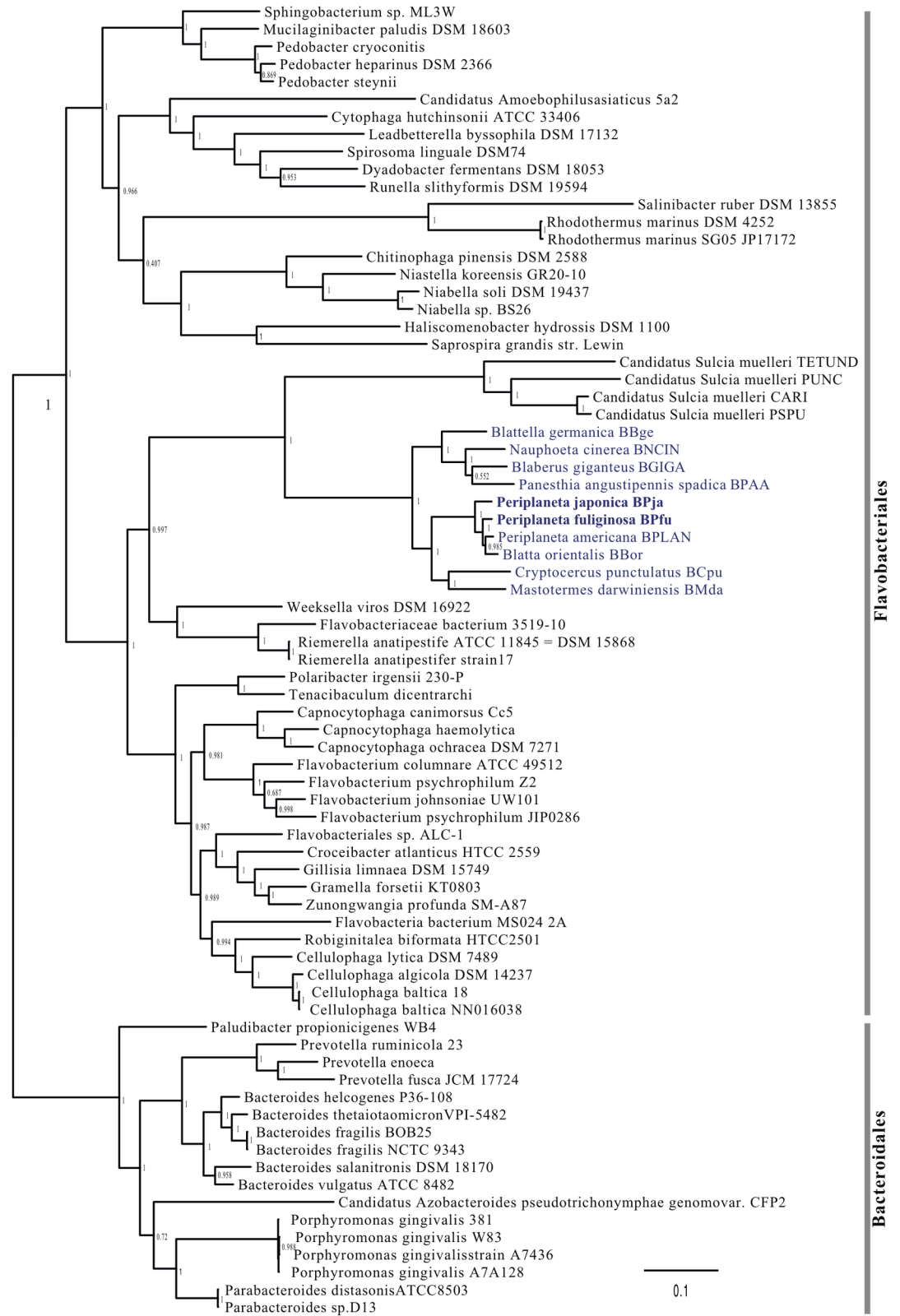
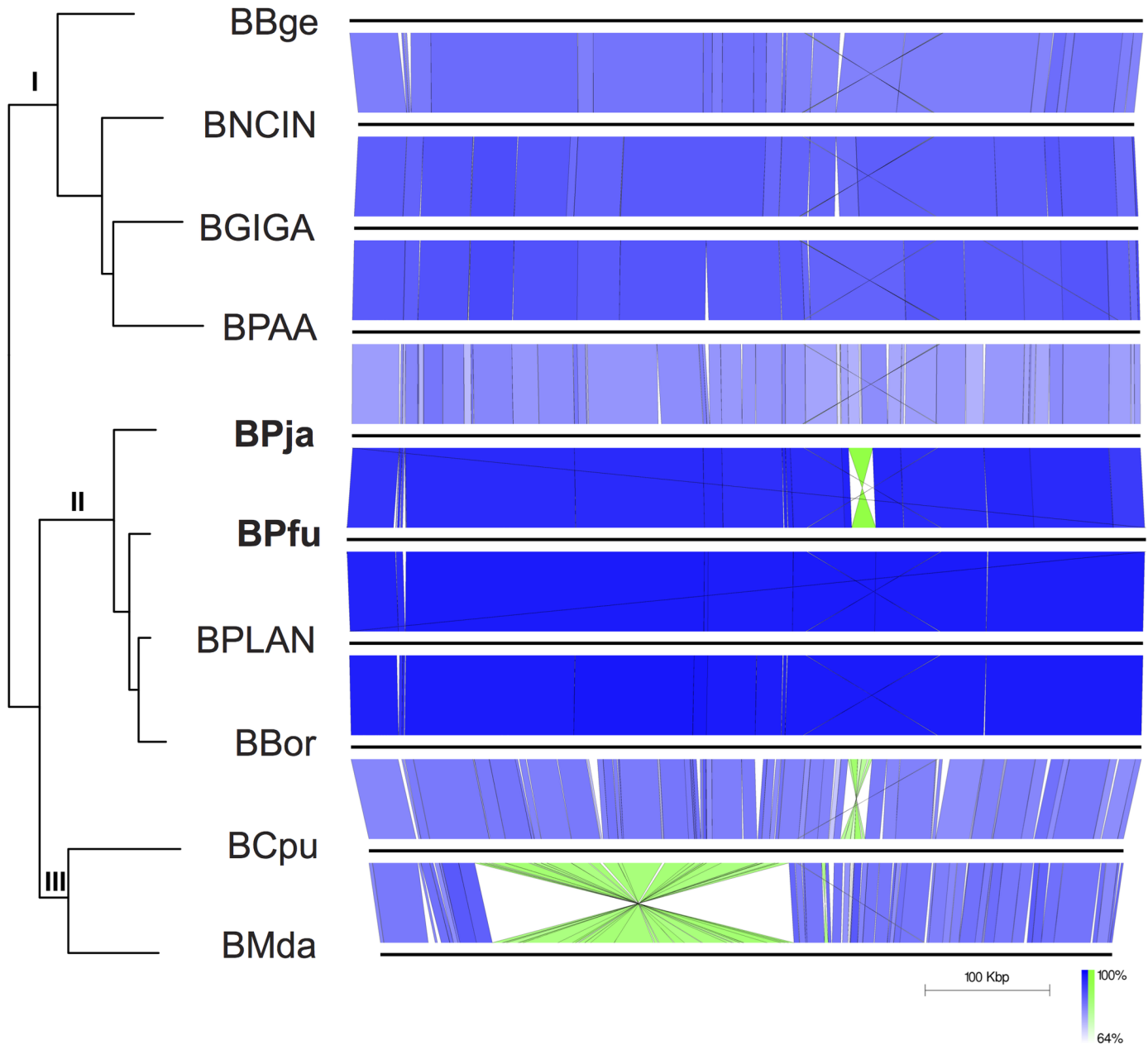**Fig 1. Phylogenomics analysis of 75 Bacteroidetes including *Blattabacterium* strains.**

Fig 2. Gene order comparison between all *Blattabacterium* strains. Lines between genomes connect orthologous genes in blue if genes are in the same orientation, or in green if they are inverted. The first gene in all strains is *yidC* (membrane protein insertase C).

pan-genome and the omnivorous strains, BBge, whereas the strains from the two wood-feeding species, BCpu and BMda, cluster with the core genome. Several expected features arose from this analysis, such as the predominance of genes in COG category J (translation) within the core and it is the most represented functional category in the endosymbiont genomes. This was expected as most of the conserved universal COGs are contained within this category [36]. The second most represented functional category, especially in the omnivorous species, contains genes involved in amino acid transport and metabolism (E) and shows the main gap between the wood-feeding species and the others, because it harbours most of the gene losses
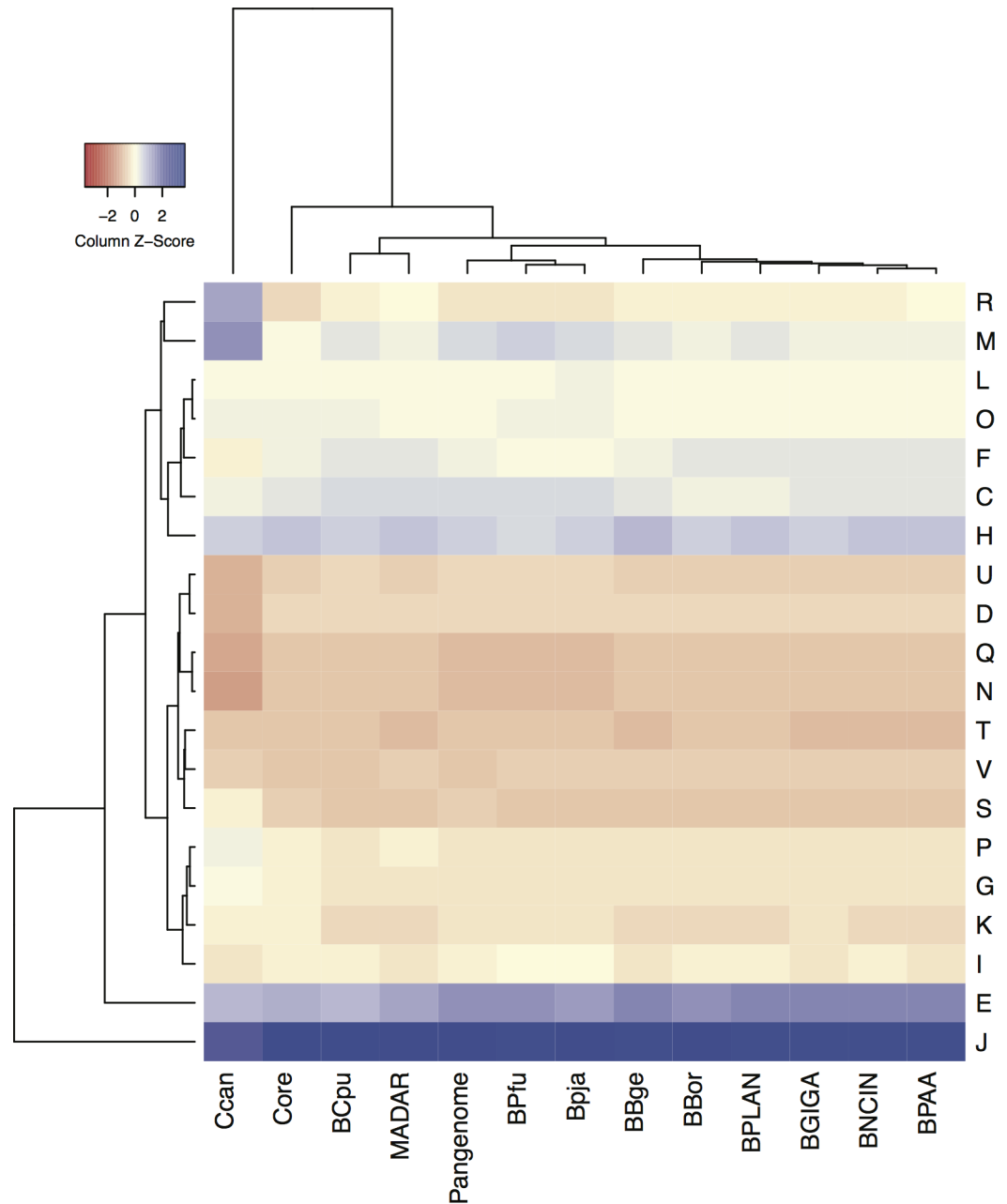
Fig 3. COG frequency heat map of different *Blattabacterium* strains with their pan- and core-genome, and the free-living *Capnocytophaga canimorsu*. By alphabetic order: C, energy production and conversion; D, cell cycle control; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; J, translation; K, transcription; L, replication, recombination, and repair; M, cell/wall membrane biogenesis; N, cell motility; O, post-translational modification, protein turnover, chaperones; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport, and catabolism; R, general function prediction only; S, function unknown; T, signal transduction mechanism; U, intracellular trafficking and secretion; and V, defense mechanism.

https://doi.org/10.1371/journal.pone.0200512.g003

described in these two strains [11, 13]. The exception is in other wood feeding species BPAA which is close to BNCIN and BGIGA as described by Tokuda et al. [16].
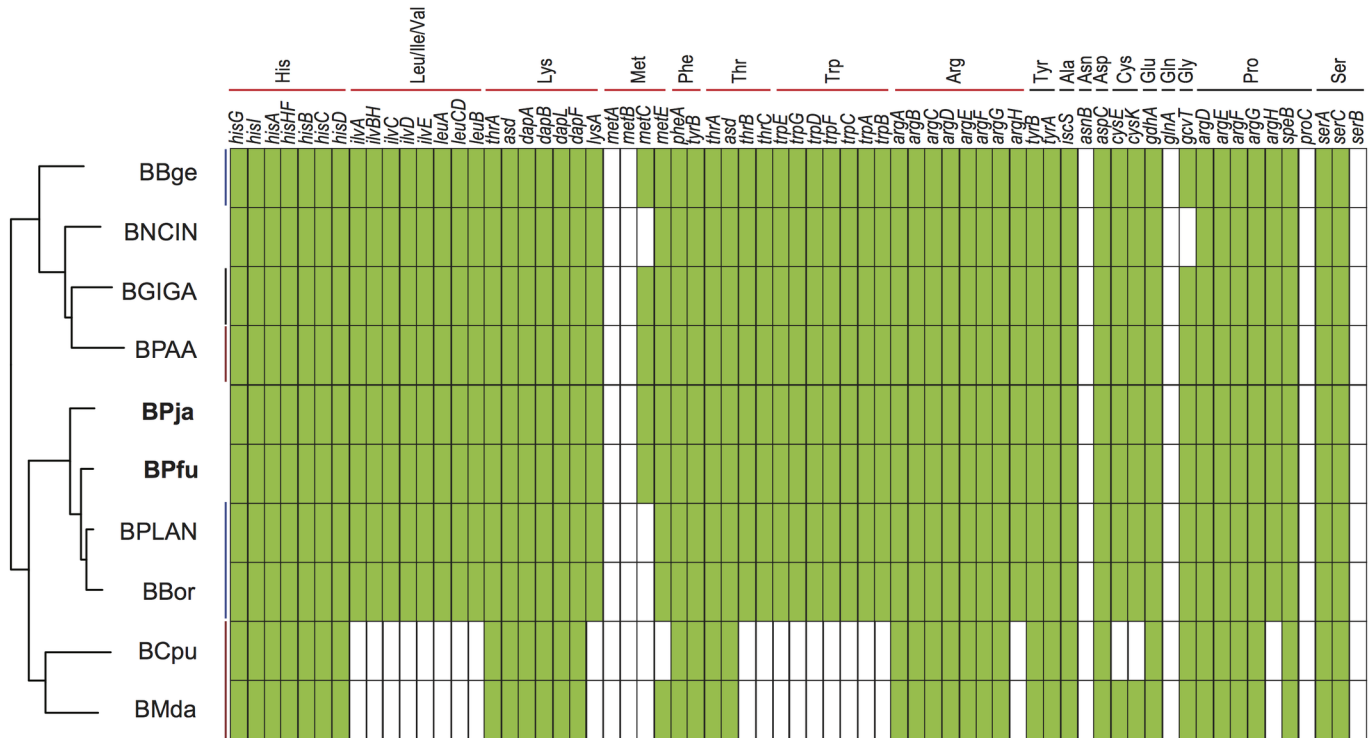
Fig 4. List of genes associated with amino acid synthesis in all sequenced *Blattabacterium* strains. The absent genes are represented by white boxes. Host species abbreviations: BPfu, *Periplaneta fuliginosa*; BPja, *Periplaneta japonica*; BNCIN, *Nauphoeta cinerea*; BGIGA, *Blaberus giganteus*; BBge, *Blattella germanica*; BPLAN, *Periplaneta americana*; BBor, *Blatta orientalis*; BPAA, *Panesthia angustipennis*; BCpu, *Cryptocercus punctulatus*; BMda, *Mastotermes darwiniensis*. Vertical bars indicate omnivorous (blue), wood-feeding (red), and litter-feeding (black) hosts, respectively. EAA are indicated by orange horizontal bars, and non-EAA indicated by black horizontal bars.

Cockroaches are incapable of *de novo* essential amino acid (EAA) synthesis, relying mostly in their diet and in their *Blattabacterium* endosymbionts [2, 8]. The putative metabolic pathway reconstruction from the genome sequence revealed that both BPja and BPfu have the complete biosynthetic pathways for nine EAA (His, Leu, Ile, Val, Lys, Phe, Thr, Trp and Arg) and six non-EAA (Tyr, Ala, Asp, Cys, Glu and Gly) amino acids ([Fig 4]). The gene set for Asn and Gln biosynthesis is absent in both strains. Whereas only the terminal step enzyme is missing, in case of Pro and Ser biosynthetic pathways. The comparison of amino acid biosynthetic pathways of BPja and BPfu with the other eight-sequenced *Blattabacterium* genomes [8, 10, 12, 14, 16] revealed that none of these strains are able to synthesize all the amino acids. However, the complete biosynthetic pathways of His, Phe, Tyr, Asp, and Glu are present in all *Blattabacterium* strains. All the strains are capable to synthesize all the EAA, except BCpu and BMda that are unable to synthesize Leu, Ile, Val, Thr and Trp.

Most insect species are unable to use reduce oxidized sulfur compounds and incorporate them into biomolecules, depending thus on their diet or the activity of their endosymbionts [37]. The endosymbiont *Buchnera aphidicola*, the primary endosymbiont of aphids, can assimilate sulfate into sulfur-containing amino acids as Cys or Met [36]. In *Blattabacterium*, only BBge [38, 39] and BMda [15] are able to assimilate sulfate as sulfur donor, possessing a cluster of 7 genes coding for all enzymes involved (*cysNDHIJEK*), in exception of *cysC* (5'-phosphosulfate kinase). Likewise, we found that BPfu also possess all genes for sulfate assimilation like BBge and BMda ([Fig 5]). In contrast, BPja have lost most of the genes of this pathway (*cysDCHI*).
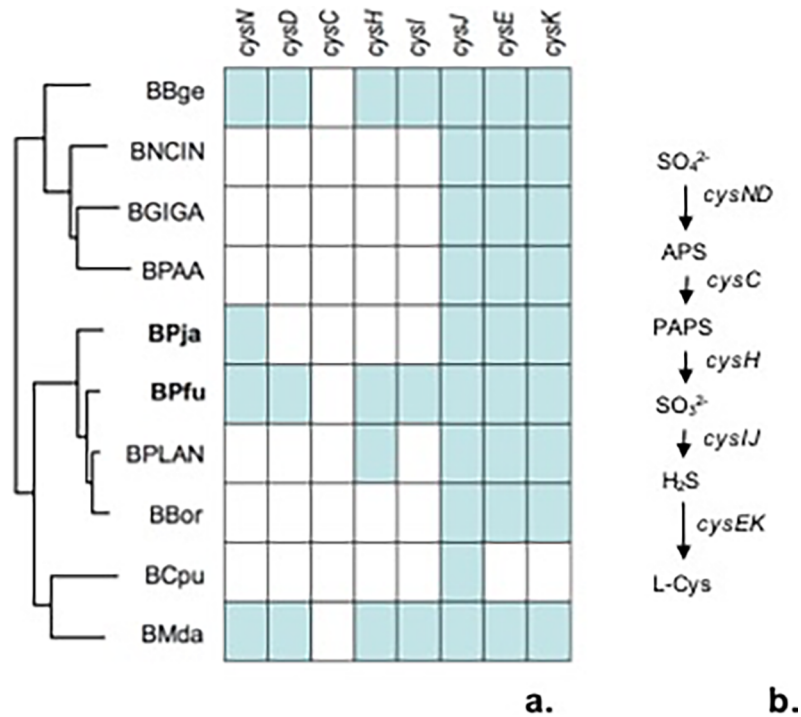
Fig 5. The pathway of sulfate assimilation from different *Blattabacterium* strains. (a). Genes required for sulfur assimilation (b) include *cysN* and *cysD* coding for two subunits of sulfate adenyltransferase; the adenosine 5'-phosphosulfate (APS) reductase *cysH* and the sulfite reductase *cysIJ*. There is a missing step for the conversion of adenosine-5'-phosphosulfate (APS) into 3'-phospho adenosine-5'-phosphosulfate (PAPS). The generated sulfite is reduced to hydrogen sulfide further on assimilated into sulfur-containing amino acids L-cysteine and L-methionine.

BPja and BPfu are able synthesize some vitamins and cofactors including riboflavin (vitamin B2), pyridoxine (vitamin B6), folate (vitamin B9) and the cofactors FAD, NADP⁺, Coenzyme A, lipoate, FE-S cluster, molybdopterin (Fig 6). Like other endosymbionts, the missing
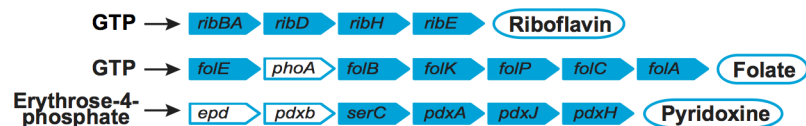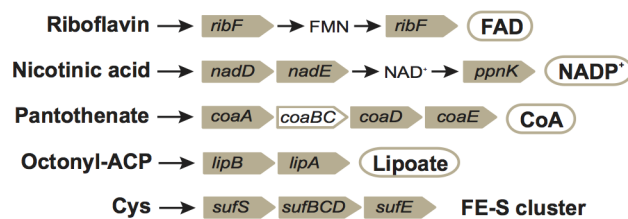

Fig 6. Reconstruction of pathways for biosynthesis of vitamins (a) and cofactors (b) in *Periplaneta fuliginosa* and *P. japonica*. Gene names are indicated in coloured rectangles. White rectangles indicate missing genes and circles indicate products.

enzymes of vitamin and cofactor biosynthesis might compensate from either the host or from co-occurring endosymbionts [40–41].

Conclusions

Aside from their ability to live indoors, *P. japonica* and *P. fuliginosa* are adapted to contrasting environmental conditions being the first native from cooler northern climates and the latter from warmer climates and with low or absent cold-resistance [42–43]. These habitats preferences also impose differences in their diet source, which as in other insect species lead to slight differences in the genome content of their endosymbionts. Although the newly sequenced genomes of BPja and BPful emphasise the remarkable stability of *Blattabacterium* genomes, we were able to detect changes in terms of genome synteny and content. The most surprising result was observed in BPja in the sulfur assimilatory pathway, which is similar to BBge and BMda, endosymbionts of contrasting hosts respectively the cockroach *B. germanica* and termite *M. darwiniensis*.

Supporting information

S1 Fig. Standard curves for qPCR quantification of *Blattabacterium (ureA)* from *Periplaneta japonica* (a) and *P. fuliginosa* (b); and host (*wg*) *P. japonica* (c) and *P. fuliginosa* (d). (TIF)

S1 Table. Predicted genes in the *Blattabacterium sp*. from *Periplaneta fuliginosa* (BPfu) genome. (XLSX)

S2 Table. Predicted genes in the *Blattabacterium sp*. from *Periplaneta japonica* (BPja) genome. (XLSX)

S3 Table. Accession number of all genomes used for phylogenomics analysis. (XLSX)

S4 Table. Protein-coding genes have been assigned COG numbers for all *Blattabacterium* strains. (XLSX)

Author Contributions

Conceptualization: Cláudia S. L. Vicente, Sota Ozawa, Koichi Hasegawa.

Data curation: Arzuba Akter.

Formal analysis: Cláudia S. L. Vicente, Shakhinur Islam Mondal, Arzuba Akter.

Funding acquisition: Tasei Kikuchi, Koichi Hasegawa.

Investigation: Cláudia S. L. Vicente, Tasei Kikuchi, Koichi Hasegawa.

Methodology: Cláudia S. L. Vicente, Sota Ozawa, Tasei Kikuchi, Koichi Hasegawa.

Project administration: Tasei Kikuchi, Koichi Hasegawa.

Resources: Tasei Kikuchi, Koichi Hasegawa.

Software: Shakhinur Islam Mondal, Arzuba Akter.

Supervision: Koichi Hasegawa.

Writing – original draft: Cláudia S. L. Vicente, Koichi Hasegawa.

Writing – review & editing: Cláudia S. L. Vicente, Shakhinur Islam Mondal, Arzuba Akter, Tasei Kikuchi, Koichi Hasegawa.

References

1. Mullins DE. Physiology of environmental adaptations and resource acquisition in cockroaches. Ann Rev Entomol. 2015; 60:473–92.

2. Bell WJ, Roth LM, Nalepa CA, editors. Cockroaches: Ecology, Behavior, and Natural History. Johns Hopkins University Press. USA; 2007.

3. Engel P, Moran NA. 2013. The gut microbiota of insects—diversity in structure and function. FEMS Microbiol Rev. 2013; 37:699–735 https://doi.org/10.1111/1574-6976.12025 PMID: 23692388

4. Yun JH, Roh SW, Whon TW, Jung MJ, Kim MS, Park DS, et al. (2014) Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. App Environ Microbiol. 2014; 80:5254–5264

5. Wernegreen JJ. Endosymbiont evolution: predictions from theory and surprises from genomes. Ann N Y Acad Sci. 2015; 1360: 16–35. https://doi.org/10.1111/nyas.12740 PMID: 25866055

6. Bandi C, Sironi M, Damiani G, Magrassi L, Nalepa CA, Laudani U, et al. The establishment of intra-cellular symbiosis in an ancestor of cockroaches and termites. Proc R Soc Lond [Biol]. 1995; 259: 293–299.

7. Brooks MA, Richards AG, Intracellular symbiosis in cockroaches. 1. Production of aposymbiotic cock-roaches. Biolog Bull. 1994; 109: 22–39.

8. Sabree ZL, Kambhampati S, Moran NA. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. PNAS. 2009; 106:19521–6. https://doi.org/10.1073/pnas.0907504106 PMID: 19880743

9. Sacchi L, Nalepa CA, Bigliardi E, Corona S, Grigolo A, et al. Ultrastructural studies of the fat body and bacterial endosymbionts of *Cryptocercus punctulatus* Scudder (Blattaria: Cryptocercidae). Symbiosis. 1998; 25:251–69

10. López-Sánchez MJ, Neef A, Pereto J, Patiño-Navarrete R, Pignatelli M, Latorre A, et al. Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cock-roach *Blattella germanica*. PLoS Genetics. 2009; 5:e1000721. https://doi.org/10.1371/journal.pgen.1000721 PMID: 19911043

11. Neef A, Latorre A, Peretó J, Silva FJ, Pignatelli M, Moya A. Genome economization in the endosymbi-ont of the wood roach *Cryptocercus punctulatus* due to drastic loss of amino acid synthesis capabilities. Genome Biol Evol. 2011; 3:1437–1448. https://doi.org/10.1093/gbe/evr118 PMID: 22094859

12. Huang CY, Sabree ZL, Moran NA. Genome sequence of *Blattabacterium* sp. strain BGIGA, endosymbi-ont of the *Blaberus giganteus* cockroach. J Bacteriol. 2012; 194:4450. https://doi.org/10.1128/JB.00789-12 PMID: 22843586

13. Sabree ZL, Huang CY, Arakawa G, Tokuda G, Lo N, Watanabe H, Moran NA. Genome shrinkage and loss of nutrient-providing potential in the obligate symbiont of the primitive termite *Mastotermes darwi-niensis*. App Environ Microbiol. 2012; 78:204–210

14. Kambhampati S, Alleman A, Park Y. Complete genome sequence of the endosymbiont *Blattabacterium* from the cockroach *Nauphoeta cinerea* (Blattodea: Blaberidae). Genomics. 2013; 102:479–83. https://doi.org/10.1016/j.ygeno.2013.09.003 PMID: 24071059

15. Patiño-Navarrete R, Moya A, Latorre A, Pereto J. Comparative genomics of *Blattabacterium cuenoti*: the frozen legacy of an ancient endosymbiont genome. Genome Biol Evol. 2013; 5:351–61. https://doi.org/10.1093/gbe/evt011 PMID: 23355305

16. Tokuda G, Elbourne LD, Kinjo Y, Saitoh S, Sabree Z, Hojo M, et al. Maintenance of essential amino acid synthesis pathways in the *Blattabacterium cuenoti* symbiont of a wood-feeding cockroach. Biol Let-ters. 2013; 9:20121153.

17. Vicente CSL, Ozawa S, Hasegawa K. Composition of the Cockroach Gut Microbiome in the Presence of Parasitic Nematodes. Microbes Environ. 2016; 31:314–320. https://doi.org/10.1264/jsme2.ME16088 PMID: 27524304

18. Elias-Neto M, Belles X. Tergal and pleural structures contribute to the formation of ectopic prothoracic wings in cockroaches. R Soc open sci. 2016; 3: 160347 https://doi.org/10.1098/rsos.160347 PMID: 27853616

19. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res. 2014; 24:1384–95. https://doi.org/10.1101/gr.170720.113 PMID: 24755901

20. Hahn C, Bachmann L, Chevreux B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nuclei Acids Res. 2013; 41: e129

21. Seemann T. PROKKA: rapid prokaryotic genome annotation. Bioinformatics. 2014; 30:2068–9. https://doi.org/10.1093/bioinformatics/btu153 PMID: 24642063

22. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genomics. 2008; 9:75. https://doi.org/10.1186/1471-2164-9-75 PMID: 18261238

23. Wu S, Zhu Z, Fu L, Niu B, Li W. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics. 2011; 12:444. https://doi.org/10.1186/1471-2164-12-444 PMID: 21899761

24. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, et al. IMG: the Integrated Microbial Genomes database and comparative analysis system. Nuclei Acids Res. 2012; 40:D115–22.

25. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic acids research. 2016:W147–53. https://doi.org/10.1093/nar/gkw419 PMID: 27190236

26. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016; 428:726–31. https://doi.org/10.1016/j.jmb.2015.11.006 PMID: 26585406

27. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, et al. EcoCyc: fusing model organism databases with systems biology. Nucleic acids research. 2013; 41: D605–12. https://doi.org/10.1093/nar/gks1027 PMID: 23143106

28. Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic acids research. 2016; 44: D471–80. https://doi.org/10.1093/nar/gkv1164 PMID: 26527732

29. Lo N, Bandi C, Watanabe H, Nalepa C, Beninati T. Evidence for cocladogenesis between diverse dictyopteran lineages and their intracellular endosymbionts. Mol Biol Evol. 2013; 20: 907–913.

30. Conord C, Despres L, Vallier A, Balmand S, Miquel C, Zundel S, et al. Long-term evolutionary stability of bacterial endosymbiosis in Curculionoidea: Additional evidence of symbiont replacement in the Dryophthoridae family. Molecular Biol Evol. 2008; 25: 859–868.

31. Chaumot A, Da Lage JL, Maestro O, Martin D, Iwema T, Brunet F, et al. Molecular adaptation and resilience of the insect's nuclear receptor USP. BMC Evol Biol. 2002; 12: 199.

32. Sloan DB, Moran NA. The evolution of genomic instability in the obligate endosymbionts of whiteflies. Genome Biol Evol. 2013; 5:783–793. https://doi.org/10.1093/gbe/evt044 PMID: 23542079

33. Moran NA, McCutcheon JP, Nakabachi A. Genomics and evolution of heritable bacterial symbionts. Ann Rev Genet. 2008; 42:165–90. https://doi.org/10.1146/annurev.genet.41.110306.130119 PMID: 18983256

34. Moya A, Pereto J, Gil R, Latorre A. Learning how to live together: genomic insights into prokaryote-animal symbioses. Nature Rev Genet. 2008; 9:218–29. https://doi.org/10.1038/nrg2319 PMID: 18268509

35. Pal C, Papp B, Lercher MJ, Csermely P, Oliver SG, Hurst LD. Chance and necessity in the evolution of minimal metabolic networks. Nature. 2006; 440: 667–670. https://doi.org/10.1038/nature04568 PMID: 16572170

36. Ciccarelli FD, Doerks T, Mering C, Creevey C, Snel B, Bork P. Toward automatic reconstruction of a highly resolved tree of life. Science. 2006; 311:1283–1287 https://doi.org/10.1126/science.1123061 PMID: 16513982

37. Zientz E, Dandekar T, Gross R. Metabolic interdependence of obligate intracellular bacteria and their insect host. Microbiol Mol Biol Rev. 2014; 68: 745–770

38. González-Domenech C, Belda E, Patiño-Navarette R, Moya A, Peretó J, Latorre A. Metabolic basis in an ancient symbiosis: genome-scale metabolic networks from two *Blattabacterium cuenoti* strains, primary endosymbionts of cockroaches. BMC Microbiol. 2012; 12:55. https://doi.org/10.1186/1471-2180-12-55

39. Ponce-de-León M, Montero F, Peretó J. Solving gap metabolites and blocked reactions in genome-scale models: application to the metabolic network of Blattabacterium cuenoti. BMC Sys Biol. 2013; 7: 114.

40. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, et al. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell. 2013; 153:1567–78. https://doi.org/10.1016/j.cell.2013.05.040 PMID: 23791183

41. Sloan DB, Nakabachi A, Richards S, Qu J, Murali SC, Gibbs RA, et al. Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. Mol Biol Evol. 2014; 31:857–71. https://doi.org/10.1093/molbev/msu004 PMID: 24398322

42. Tanaka K, Tanaka S. Winter survival and freeze tolerance in a Northern cockroach, *Periplaneta japonica* (Blattidae: Dictyoptera). Zoo Science. 1997; 14:849–853.

43. Appel AG, Smith LM. Biology and management of the smokybrown cockroach. Ann Rev Entomol. 2002; 47:33–55.