Title
The Complete Mitochondrial Genome Sequence of Bactericera cockerelli and Comparison with Three Other Psyloidea Species.

Permalink
https://escholarship.org/uc/item/96g2j62g

Journal
PloS one, 11(5)

ISSN
1932-6203

Authors
Wu, Fengnian
Cen, Yijing
Wallis, Christopher M
et al.

Publication Date
2016

DOI
10.1371/journal.pone.0155318

Peer reviewed
The Complete Mitochondrial Genome Sequence of Bactericera cockerelli and Comparison with Three Other Psylloidea Species

Fengnian Wu¹,², Yijing Cen¹, Christopher M. Wallis², John T. Trumble³, Sean Prager³, Ray Yokomi², Zheng Zheng¹,², Xiaoling Deng¹, Jianchi Chen²*, Guangwen Liang¹*

¹ Laboratory of Insect Ecology / Guangdong Province, Key Laboratory of Microbial Signals and Disease Control, College of Agriculture, South China Agricultural University, Guangzhou, Guangdong, China, ² San Joaquin Valley Agricultural Sciences Center, United States Department of Agriculture -Agricultural Research Service, Parlier, California, United States of America, ³ Department of Entomology, University of California Riverside, Riverside, California, United States of America

* jianchi.chen@ars.usda.gov (JC); gwliang@scau.edu.cn (GL)

Abstract

Potato psyllid (Bactericera cockerelli) is an important pest of potato, tomato and pepper. Not only could a toxin secreted by nymphs result in serious phytotoxemia in some host plants, but also over the past few years B. cockerelli was shown to transmit “Candidatus Liberibacter solanacearum”, the putative bacterial pathogen of potato zebra chip (ZC) disease, to potato and tomato. ZC has caused devastating losses to potato production in the western U.S., Mexico, and elsewhere. New knowledge of the genetic diversity of the B. cockerelli is needed to develop improved strategies to manage pest populations. Mitochondrial genome (mitogenome) sequencing provides important knowledge about insect evolution and diversity in and among populations. This report provides the first complete B. cockerelli mitogenome sequence as determined by next generation sequencing technology (Illumina MiSeq). The circular B. cockerelli mitogenome had a size of 15,220 bp with 13 protein-coding gene (PCGs), 2 ribosomal RNA genes (rRNAs), 22 transfer RNA genes (tRNAs), and a non-coding region of 975 bp. The overall gene order of the B. cockerelli mitogenome is identical to three other published Psylloidea mitogenomes: one species from the Triozidae, Paratrioza sinica; and two species from the Psyllidae, Cacopsylla coccinea and Pachypsylla venusta. This suggests all of these species share a common ancestral mitogenome. However, sequence analyses revealed differences between and among the insect families, in particular a unique region that can be folded into three stem-loop secondary structures present only within the B. cockerelli mitogenome. A phylogenetic tree based on the 13 PCGs matched an existing taxonomy scheme that was based on morphological characteristics. The available complete mitogenome sequence makes it accessible to all genes for future population diversity evaluation of B. cockerelli.
Introduction

The potato, or potato-tomato, psyllid Bactericera cockerelli (Šulc) (Hemiptera: Triozidae) is an important pest of potato, tomato and pepper production. The insect can cause damage on plants when feeding via phytotoxemia, and, in addition, also can transmit “Candidatus Liberibacter solanacearum” (a.k.a. “Ca. L. psyllaurous”), an unculturable alpha-proteobacterium associated with potato zebra chip (ZC) disease [1, 2]. B. cockerelli was first described over a hundred years ago under the name of Trioza cockerelli [3]. Thus, for over a century substantial efforts have been made to study the taxonomy, damage mechanism, and management for this psyllid [4]. However, knowledge is limited on B. cockerelli evolution and population diversity.

B. cockerelli was thought to annually migrate from warm southern regions of North America (e.g. Mexico) to the western United States (e.g. North Texas, Colorado, the Dakotas, Kansas, Nebraska) via air currents, where it then colonizes and damages solanaceous crop plants [5–7]. However, the existence of distinct regional populations of B. cockerelli (biotypes) were discovered following a 2001 outbreak throughout western North America, based on variations revealed by inter simple sequence repeat (ISSR) markers and single nucleotide polymorphism (SNP) in the mitochondrial gene of cytochrome oxidase I (coxI) [8]. That study determined B. cockerelli populations clustered into two biotypes, one from western North America and one from the central United States of America (USA) and eastern Mexico. Existence of B. cockerelli biotypes suggests adaptation to local habitats are taking place, hence, there may be less dependence on long-range dispersal and a concomitant increase in the regional pest status associated with B. cockerelli populations.

Later, Swisher et al. identified three haplotypes of B. cockerelli within USA based on their SNP analyses of a 500-bp coxI gene sequence: a Central haplotype from eastern Mexico north to Texas, Kansas, Colorado, Nebraska, Wyoming, and North Dakota; a Western haplotype, from California and New Mexico north to Washington, Oregon, and Idaho; and a Northwestern haplotype, including the Northwestern states of Washington, Oregon, and Idaho [9]. More recently, an additional haplotype (Southwestern) was identified and is found in New Mexico and southern Colorado [10]. However, some haplotypes (Central and Western) were differentiated by only one SNP. The same SNP was also described by Liu et al [8] in their biotype study. To assure a comprehensive description and haplotype/biotype stability of the B. cockerelli, we feel that there is a need to evaluate more sequences from the coxI and other genes in the psyllid genome.

In the recent years, whole genome sequencing of mitochondria genomes (mitogenomes) has gained importance for comprehensive evolutionary and population studies of insects. This is due to the relatively small genome sizes, low levels of recombination, and variable evolution rates of genes in the mitogenomes. With few exceptions, insect mitogenomes consist of 13 protein-coding genes (PCGs), two ribosomal RNA genes (rRNA), 22 transfer RNA genes (tRNA), and a large non-coding region (also called the control region, CR) on a single circular chromosome. Arrangement of genes in mitogenomes is usually stable, retaining the ancestral pattern of gene arrangement [11]. Previous mitogenome sequencing has been heavily based on sequencing DNAs amplified using conserved PCR primers by the Sanger method [12–16]. This method is both time consuming and laborious. However, presently Next-Generation Sequencing (NGS) technologies have been developed to generate a large amount of short DNA sequences (reads) from a single insect. Mitochondrial reads can then be found based on appropriate reference sequences and used to assemble complete mitogenome sequences [17, 18].

In the superfamily Psylloidea, three complete mitogenome sequences have been published and available for public use in GenBank database: Paratrioza sinica in the family of Triozidae [19], and Cacopsylla coccinea and Pachypsylla venusta in the family of Psyllidae [12, 16].
*Sinica* is an important pest of wolfberry or Goji (*Lycium chinense*) in Northwest China, where feeding on tender shoots and buds of the host causes yellow leaves and declined growth [20]. *C. coccinae* is a pest of chocolate vine (*Akebia quinata*) in China, Japan and Korea [20]. *P. venusta* causes damage to hackberry (*Celtis occidentalis*) in the U.S. by forming woody galls on leaf petioles [21]. Morphologically, adult Triozidae are characteristic in a vein pattern in the forewing diverging from a single point, whereas adult Psyllidae are characteristic in having antennae with a second segment wider than the first and longer than the third [22]. To date, there has not been a comparative analysis among members of these two psyllid families at the mitogenome level.

While sequencing the genome of "Ca. L. solanacearum" from infected *B. cockerelli* DNA [23], a sequence contig containing an almost complete mitogenome was identified. The current report details the circularity and complete mitogenome of the *B. cockerelli*. Genes were annotated and comparative analyses were made between the mitogenome of *B. cockerelli* and those of the other three psyllids [12, 16, 19]. Efforts were made to evaluate *B. cockerelli* variations based on sequences of cox1 gene currently available in GenBank. Potential applications of the *B. cockerelli* mitogenome sequence in Psylloidea study were discussed.

**Materials and Methods**

**Sample collection and DNA preparation**

Adults of *B. cockerelli* were originally collected from the University of California South Coast Research and Extension Center in Irvine, California and maintained in a greenhouse at University of California at Riverside. DNA was extracted and purified from the individual psyllid using the DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA, USA). Briefly, individual insects were placed in microcentrifuge tubes containing 20 μL lysis buffer and homogenized by grinding with a plastic pestle (Kimble Chase, Vineland, NJ, USA). DNA isolation was carried out according to the manufacturer’s instructions. Sixty μL of DNA suspension were finally recovered from a spin column. DNA was amplified through illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Inc., Waukesha, WI, USA).

**Mitogenome sequencing and assembling**

The amplified DNA was sequenced using Illumina MiSeq format (Illumina, San Diego, CA, USA) and assembled de novo as described previously [23]. Illumina MiSeq was used because it generated a large volume of sequence data for high coverage de novo assembly. A single contig associated with mitogenome was identified by standalone BLASTn (version 2.2.30) [24] referenced to the complete mitogenome sequences of three Psylloidea species, NC_024577 (*P. sinica*), NC_027087 (*C. coccinea*) and NC_006157 (*P. venusta*) (Table 1), downloaded from GenBank database. Sequence of the contig was extracted using a Perl script. Coverage was

| Superfamily | Family | Species             | Accession number | Length of genome (bp) | Reference   |
|-------------|--------|---------------------|------------------|-----------------------|-------------|
| Psylloidea  | Triozidae | Bactericera cockerelli | KU501214         | 15,220                | This study  |
|             |         | *Paratrioza sinica* | NC_024577        | 14,863                | [19]        |
|             | Psyllidae | Cacopsylla coccinea | NC_027087        | 14,832                | [16]        |
| Aphidoidea  | Aphididae | Pachypsylla venusta | NC_006157        | 14,711                | [12]        |
| Aleyrodoidea | Aleyrodidae | Cervaphis quercus    | NC_024926        | 15,272                | [15]        |
|             |         | Aleurochiton aceris | NC_006160        | 15,388                | [12]        |

**Table 1. General information of mitogenomes used in this study.**

doi:10.1371/journal.pone.0155318.t001
calculated by mapping to the contig using paired reads of the MiSeq data by CLC Genomics Workbench 7.5 (CLC Bio, Denmark), with the following parameters: mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.8, and similarity fraction = 0.9. The mitogenome circularity was verified by conventional PCR using the primers BC-mito-F (5′-GGTATC TAA TCC TGG TTT AGC GC-3′) and BC-mito-R (5′-TTG TCT TCT AAC ATT GGA GTG GGG-3′) designed by Primer3 [25] based on sequence from both end regions of the mitogenome contig. For PCR, reaction mixture (25 μL) contained: 20 ng of DNA template, 0.2 μM of each primer, 5 mM dNTP mixture, 2.5 μL 10X buffer and 1 U of TaKaRa Taq HS enzyme. PCR amplification was: initial denaturation for 3 min at 95°C, followed by 35 cycles of denaturation for 45 s at 95°C, annealing for 30 s at 55°C, elongation for 2 min 30 s at 72°C, and a final extension step of 72°C for 10 min. Amplicons were electrophoresed on 1% agarose gel, collected and purified using NucleoSpin Gel and PCR Clean-up kit (QIAGEN, Valencia, USA). The purified DNA was sequenced using ABI 3130 DNA sequencer (ABI, Foster, CA, USA). The mitogenome of B. cockerelli was enclosed manually.

Annotation and Sequence Analyses
PCGs of B. cockerelli were identified by ORF Finder software available at the website of NCBI (National Center for Biotechnology Information) with the invertebrate mitochondrial genetic codons. Gene boundaries were compared and confirmed with the annotated sequences of the three published psyllid mitogenomes (Table 1) using ClustalW as implemented in MEGA 6 [26]. tRNAs were predicted by their cloverleaf secondary structure using tRNAscan-SE 1.21 [27], ARWEN v1.2 [28] and MITOS [29]. Sequence tandem repeats were analyzed by Tandem Repeats Finder [30] (http://tandem.bu.edu/trf/trf.html). Sequence secondary structure was predicted by Mfold [31]. Nucleotide composition and codon usage were analyzed with MEGA 6 [26]. AT and GC-skew were calculated according to the formulae: AT skew = (fa – ft) / (fa + ft) and GC skew = (fg – fc) / (fg + fc), where fa, ft, fc and fg are proportions of each nucleotide [32]. Sliding window analyses were performed using DnaSP v5 [33]. Nucleotide diversities (Pi’s) among PCGs and rRNA genes were estimated by sliding window analyses (a 250 bp window in 25 bp overlapping steps) across the alignment of mitogenome sequences among the four Psylloidea members (Table 1). For cox1 locus evaluation and phylogenetic tree construction, available B. cockerelli sequences were downloaded from GenBank.

Phylogenetic analysis
Phylogenetic relationships among the four Psylloidea members were analyzed based on mitogenome sequences along with two mitogenome sequences (NC_024926 for Cervaphis quercus and NC_006160 for Aleurochiton aceris) downloaded from GenBank as out groups (Table 1). The nucleotide sequences of 13 PCGs in the mitogenomes were collected and translated into amino acid sequences. The amino acid sequences were aligned by Clustal X [34] following the method of Jeyaprakash and Hoy [35]. Briefly, gap opening penalty = 35 and extension = 0.75 were set for pairwise sequence alignment; Gap opening penalty = 15 and extension = 0.3 were set for multiple sequence alignment; and the Gonnet protein weight matrix was used. The Gonnet matrix is an extension of the commonly used PAM matrix, but more suitable for large data set with more taxa or characters. With the help of a Perl script and referring to the amino acid sequences, the third nucleotide in each codon was removed in each gene sequence to minimize the effect of synonymous substitution leading to evaluation bias or skew [11]. For a cross-checking purpose, the procedure of Jeyaprakash and Hoy [35] was followed. Two software programs, PHYML 3.0 [36] for the maximum likelihood (ML) method and MrBayes (version: 3.2.5) [37] for the Bayesian inference (BI) method, were utilized to construct phylogenetic analysis.
trees. For ML analyses, the optimal substitution model obtained from jModelTest [38] was used. Nodal support among branches was evaluated by bootstrap analysis with 100 replicates [39]. For BI analyses, two sets of four chains were allowed to run simultaneously for 1,000,000 generations, with sampling every 100 generations. After discarding the first 25% samples as burn-in, Bayesian posterior probability values were calculated in a consensus tree [40].

Results

Mitogenome organization of B. cockerelli

From de novo assembly of MiSeq data, a single contig of 15,263 bp with a 43-bp duplication at both ends was identified by the mitogenome sequences of P. sinica (78%), C. coccinea (77%), and P. venusta (76%). Primer set BC-mito-F/BC-mito-R amplified a fragment of 1,285 bp, proving the circularity of the psyllid mitogenome (Figs 1 and 2). The mitogenome of B. cockerelli was determined to be 15,220 bp. Based on annotation, the B. cockerelli mitogenome included the entire set of 37 genes as those in the ancestral mitogenomes of insects [11, 41]. The average nucleotide coverage of the B. cockerelli mitogenome was 10,604 X, with the lowest (2,074 X) in control region and the highest (16,468 X) in the nad6 gene. Twenty-three genes were on the majority strand (J-strand), and the other 14 genes on the minority strand (N-strand) (Fig 2). Sizes of all intergenic regions ranged from 3 to 31 bp with the exceptions of CR as discussed later. Gene overlaps were also observed: four between PCG and PCG (1–7 bp), seven between tRNA and tRNA (1–12 bp), and three between PCG and tRNA gene (2–3 bp) (Fig 2).

Nucleotide compositions

The A+T content, GC-skew and AT-skew of the B. cockerelli and three other Psylloidea mitogenomes are listed in Table 2. The mitogenome of B. cockerelli was highly A+T biased with the control region being the highest (82.7%). Whole genome-wide, PCGs in N-strand PCGs were more AT-skewed (-0.248) than the J-strand (-0.084), i.e. PCGs in J-strand had nearly equal A and T. The tRNA genes in the N-strand was more GC-skewed (0.429) than those in the J-strand (0.022). No significant differences were found in both PCGs and tRNA genes between B. cockerelli and the other three psyllids. For rRNA genes, the A+T content of B. cockerelli and P. sinica (75.8% and 75.1%) were lower than those of C. coccinea and P. venusta (77.1% and 77.8).

Codon usage of protein coding genes (PCGs)

A total of 3,595 amino acids were coded for in the 13 PCGs of the B. cockerelli mitogenome. In terms of usage, the most frequently occurring amino acids were isoleucine (10.79%),

![Fig 1. A schematic representation of de novo assembling and coverage estimate of the 15,263 bp mitogenome of Bactericera cockerelli using MiSeq data.](https://example.com/fig1.png)

In the top Contig line, the blue and red arrows represent the forward and reversed primers (BC-mito-F/BC-mito-R) used to verify the circularity of B. cockerelli mitogenome by PCR. Numbers are nucleotides in bp. In the Coverage section, the pink area represents nucleotide coverage with the highest of 16,468 X in nad6 and the lowest of 2,074 X in CR. In the Reads section, ten top read assemblings from MiSeq data were representatively shown with blue as pair reads, red as forward reads only and green as reversed reads only.

doi:10.1371/journal.pone.0155318.g001
phenylalanine (9.60%), and leucine (the UUR codon) (8.98%) (Fig 3). It should be noted that leucine and serine are each coded by two separate codon formats (CUN or UUR for leucine, and AGN or UCN for serine). The two codons formats for each were listed separately because they corresponded to different tRNAs [41]. Regardless, isoleucine, phenylalanine and leucine (UUR codon) also were in the top three amino acids in the mitogenome of the other three Psyloidea mitogenomes. On the other hand, arginine and cysteine were the least used amino acids in all four Psylloidea mitogenomes (1.31% and 1.20% for *B. cockerelli*, respectively) (Fig 3).

Relative synonymous codon usage (RSCU, defined as the ratio of the observed frequency of codons to the expected frequency given that all the synonymous codons for the same amino acids are used equally) in *B. cockerelli* were shown in Fig 4. Codons ending in U and A were the most frequent, which was associated with the high A+T content mitogenome. Similar RSCUs were also observed in the other three Psyloidea mitogenomes (S1 Table).

All 13 PCGs of *B. cockerelli* start with the typical "ATN" codon, similar to those of *P. sinica*, *C. coccinea* and *P. venusta*, with a lone exception for nad5 in *P. sinica*, where "TTG" was used.
Four genes (cox1, cox2, nad5 and nad4) of*B. cockerelli* also had incomplete stop codons with only a “T” present. Incomplete “T” stop codons occurred in all three other Psylloidea mitogenomes, albeit the number of genes with the incomplete stop codon varied between species (Table 3). With the exception of two genes (cox2 and nad1), length of other 11 PCGs genes varied among the four Psylloidae mitogenomes (Table 3). The most variable gene was nad4 with *C. coccinea* being 1,287 bp and *P. sinica, P. venusta* being 1,240 bp, i.e. a nine amino acid difference.

### Ribosomal RNA genes

The length of the two ribosomal RNA gene *rrnL* (16S) and *rrnS* (12S) in *B. cockerelli* are 1,181 bp and 800 bp, respectively. The *rrnL* of *B. cockerelli* is over 20 bp longer than other three Psylloidea species (Table 3). BLASTn searches using *rrnL* of *B. cockerelli* as query against complete

![Fig 3. Percentages of amino acid usage in mitochondrial proteins of four Psylloidea species. Each amino acid is represented by its three letter abbreviation. Note that leucine and serine are each coded by two different genetic codons, and listed separately.](https://doi.org/10.1371/journal.pone.0155318.g003)
mitogenome database in GenBank identified three top hits: *P. sinica* (Query cover: 92%; identity: 84%), *C. coccinea* (Query cover: 89%; identity: 79%), and *P. venusta* (Query cover: 55%; identity: 82%). BLASTn searches of the *rrnS* shows that there is a fairly homologous core with *P. sinica* (Query cover: 97%; identity: 85%), *C. coccinea* (Query cover: 97%; identity: 78%) and *P. venusta* (Query cover: 96%; identity: 78%).

Diversity of protein coding genes (PCGs) and rRNA genes within Psylloidea

Nucleotide diversity of the 13 PCGs among the four mitogenomes of Psyllidae species is shown in Fig 5. On average, genes *nad2* (Pi = 0.38), *nad4L* (Pi = 0.37) and *nad6* (Pi = 0.38) displayed...
the highest variability. However, the most variable region (window) was in \textit{nad5} (Pi = 0.46) despite of its mean value of 0.35. On the other hand, \textit{cox1} (Pi = 0.21), \textit{cox2} (Pi = 0.24), and \textit{cox3} (Pi = 0.26) were the most conserved PCGs. Both \textit{rrnL} (Pi = 0.25) and \textit{rrnS} (Pi = 0.22) were also highly conserved.

Fig 6A shows the sequence alignment of a 500 bp regions in the \textit{cox1} among the four Psyllioidea species. This was the region used for establishment of \textit{B. cockerelli} haplotypes [9, 10] and the California psyllid (RSTM) in this study fit in the Western haplotype. Currently, there are 14 \textit{cox1} sequences of \textit{B. cockerelli} deposit in GenBank (release 211.0). By focusing the sequences corresponding to the 500-bp region, all \textit{cox1} sequences fit into one of the four haplotypes with the exception two sequences represented by AY971885, designated as M, that were identical to the Northwestern haplotype with a SNP (Fig 6A). Regardless of the numbers of SNPs (from 1 to 19), when the corresponding amino acid sequences were compared, all \textit{cox1} sequences from different psyllids shared a 100% identity with the only exception being that of the Northwestern haplotype and sequence M, which had isoleucine (I) in the place of threonine (T) at position 134 (Fig 6B). Therefore, this amino acid switch could have an impact on the structure of the \textit{cox1} protein.

As expected, significantly more variations were found among the four Psyllioidea species at the nucleotide level (Fig 6A). Total similarity (BLASTn identity) between \textit{B. cockerelli} and \textit{P. sinica} (Triozidae) was 85%, higher than that between \textit{B. cockerelli} and the two Psyllidae members [\textit{C. coccinea} (81%) and \textit{P. venusta} (82%)]. Accordingly, there were 8 to 23 substitutions that occurred at the amino acid level (Fig 6B).

Transfer RNA genes

Twenty-two tRNA genes ranging in size from 54 to 75 bp were identified in the mitogenome of \textit{B. cockerelli}. They all had standard cloverleaf structure, which was a typical feature of metazoan mitogenomes [42] with the exception of \textit{trnSer}^{AGN} with a missing dihydrouridine (DHU) arm (Fig 7). In conducting a comparison among the four Psyllioidea species, two differences were noted (Fig 7): 1) the incomplete DHU arm in \textit{trnCys} (\textit{P. venusta}); and 2) the loss of the incomplete variable loop in \textit{trnAsp} and \textit{trnHis} (\textit{P. venusta}), \textit{trnThr} (\textit{C. coccinea}), and \textit{trnPro} (\textit{B. cockerelli}).
Fig 6. Alignment of a 500-bp nucleotide sequence in cox1 (A) and the corresponding amino acid sequences (B) among four Psyllidea mitogenomes including Bactericera cockerelli haplotypes and other sequences. A: BC, B. cockerelli; RSBC, the B. cockerelli used in the study. W: Western haplotype (JQ708095 and AY971885). C: Central haplotype (JQ708094, EF372597 and AY971888). SW: Southwestern haplotype (KJ305359). NW: Northwestern haplotype (JQ708093, KR534769, KR534767, KR534766 and KR534765). M, KR534768. PS, Paratrioza sinica. CC, Cacopsylla coccinea. PV, Pachypsylla venusta. Nucleotide variations at non-third codon position and the corresponding amino acids are in red.

doi:10.1371/journal.pone.0155318.g006
Fig 7. Secondary structures of the three types of tRNA genes identified in the mitogenome of four Psylloidea species. Bar “-”, Watson-Crick base pairing. Others are canonical base pairings in tRNA: Plus sign “+”, a paring between G and U; Dot “••” a paring between U and U; And Hollow dot “◦”, a paring between A and G. Bases highlighted in red indicated the different structure among the four Psylloidea members.

doi:10.1371/journal.pone.0155318.g007
The length of the *B. cockerelli* control region is 975 bp, longer than that of *C. coccinea* (671 bp), *P. venusta* (597 bp) and *P. sinica* (700 bp) [12, 16, 19]. To further confirm the length consistency of *B. cockerelli*, ten adults were selected for conventional PCR using primer set BC-mito-F/BC-mito-R, resulting in the same amplified DNA fragment. There were no tandem repeats with a repeat unit size > 2 bp detected in *B. cockerelli*, *C. coccinea*, and *P. venusta*. For *P. sinica*, 3.2 tandem repeats with the unit size of 14 and 85% unit similarity were identified. Alignment of the control regions from four Psylloidea mitogenomes revealed an insertion sequence of 260 bp (from position 14,373 to position 14,632) in the *B. cockerelli* mitogenome (Fig 8). This region could be folded into a three stem-loop secondary structure (Fig 8).

**Phylogenetic relationships**

As shown in Fig 9, from both ML and BI methods, *B. cockerelli* was grouped with *P. sinica*, another member of the family Triozidae, and separated to *C. coccinea* and *P. venusta*, both in...
the Psyllidae. Both the bootstrap value (100%) in ML analyses and Bayesian posterior probabilities (1.00) in BI analyses indicated the high level of reliability of the analyses. The phylogenetic analyses of the Psylloidea indicated a close relationship with *Cervaphis quercus*, which belongs to Aphidoidea, even though that latter was comprised of partial mitogenome, albeit one with high reliability [12, 43].

**Discussion**

Nucleotide sequences from mitochondria have been successfully used to evaluate population variations of *B. cockerelli* albeit only 500 bp region of the 1,531 bp *cox1* gene were considered [8–10] (Table 3). As shown in Fig 6B and also discussed by Swisher et al [10], all currently identified SNPs are in the third position of the codons. In the context of this study (Fig 6A and 6B), the nucleotide substitutions at the third position nucleotide would not affect the gene function, suggesting the substitutions or at least some of the substitutions may not be stable and undermining its use for haplotyping. For example, sequence M (KR534768) is identical to the Northwestern haplotype but a SNP (Fig 6A). Likewise, the sequences for Central and Western haplotypes differ only by a SNP (Fig 6A). A natural question that arises is whether a SNP in a single gene should be cause for considering defining a novel haplotype. In this consideration, it deems to be necessary to further evaluate the population of *B. cockerelli* based on more sequence variations. The mitogenome sequence published here will potentially be a guide that can be used for selecting other genes for establishing haplotypes. If there are additional suitable regions for haplotype definition, additional data can be collected to verify existing schemes to separate psyllid populations. There likely is not a requirement to sequence the mitogenomes of all defined haplotypes of *B. cockerelli*, as the careful selection of both conservative and highly variable regions together should yield robust data to separate populations. That said, NGS techniques are becoming more and more accessible, and in the near future, NGS-based mitogenome sequencing according to the methods employed herein could be used to facilitate sequencing all or the majority of proposed psyllid haplotypes.

Thao et al. published a 3,077 bp sequence of the *B. cockerelli* mitogenome (AY601890) for a genome organization study of three Sternorrhyncha members (whiteflies, aphids and psyllids) [12]. The sequence covered two complete PCGs, *nad1* and *rrnL*. An analysis on *rrnL* sequence revealed SNPs among the four *B. cockerelli* haplotypes [44]. Interestingly, there was also only one SNP between the Western and Central haplotypes. Furthermore, AY601890 was not identical to any of the four haplotypes [44]. Powell et al [45] also explored the use of AY601890 but they detected no variations among the psyllid samples collected from Colorado, Washington and Texas. The complete mitogenome sequence presented in this study provides feasible access to the sequence of every mitochondrial gene for future *B. cockerelli* -SNP analyses. For example, 31 SNPs were identified by comparison of this mitogenome sequence with AY601890 (S1 Fig) outside the *cox1* gene. All *B. cockerelli* haplotypes/biotypes in future will be screened for these newly discovered SNPs.

In view with gene diversity among the four Psyllidea members, the most conserved gene is *cox1* with the mean Pi value of 0.21 (Fig 5). The large number of SNPs among the four Psyllidea members suggests that sequence in this region could be effective for differentiation of the four species. Referenced to the *B. cockerelli* mitogenome sequence, the 500 bp region used for existing biotyping / haplotyping studies [8–10] was located in the most conserved part of the *cox1* gene (Fig 5). As such, populations could likely be described with establishment of SNPs as observed in that region first. However, other SNPs in more variable portions could reveal
additional haplotypes, as it is easy to speculate that if the more variable portion of the \textit{cox1} gene were used, then, more SNPs could be found.

In addition to \textit{cox1}, other genes also have potential to be used for population analyses. For example, \textit{nad1} had been successfully used to analyze the phylogeny of aphid parasitoids (Hymenoptera: Braconidae: Aphidiinae) [46]. Interestingly, 14 SNPs were revealed by the comparison between the \textit{nad1} sequence from this study to that of AY601890 [12] (S1 Fig) Inferred by the data at the family/genus level, \textit{nad4} sequence has a higher Pi value of 0.32 than that of \textit{cox1} (Fig 5). Interestingly, \textit{nad4} shows a sequence that varies up to 47 bp among the four Psyllloidea members (Table 3). Therefore, \textit{nad4} could be candidate for future evaluation for diversity study among \textit{B. cockerelli} members. Another sequence to consider is that of \textit{rrnL}, which was tested with limited number of samples [44], and suggested by the inter-family/genus analysis (Fig 5) that \textit{rrnL} had a 25-bp region with the lowest Pi value of 0.14. Sequence of this gene had been used to explore phylogenetics of Dictyoptera insects [47].

Due to existing technical difficulty, only a limited number of insect mitogenomes have been sequenced The fact that only four mitogenome sequences were used for phylogenetic evaluation of Psyllloidea could be considered to lack robustness across the wide variety of insects for whom also could have been included in this study. However, the clustering of the two Triozidae members and the two Psyllidae members suggests that mitogenome-based grouping is in line with current morphology-based taxonomy of two sub-families in Psyllidea (Fig 9), and it is likely adding more members would further confirm this. For \textit{B. cockerelli}, the complete mitogenome from this study could provide a template for primer design to study gene or even complete mitogenome sequence variations through the traditional PCR approach [12–16] although the advancement of NGS technology and the drop of sequence cost are expected to be the main technique for future whole mitogenome sequencing.

Conclusions

The present study presents the first complete mitogenome of \textit{B. cockerelli}. The mitogenome contains 37 genes (13 PCGs, 2 rRNA and 22 tRNA) and a control region (CR). The gene order of the mitogenome matches with that of the common ancestral insect mitogenome. However, sequence variations existed when compared with other mitogenomes of three other Psyllloidea members, mostly in the form of SNPs and small nucleotide insertions / deletions, particularly the CR. Knowledge gained from this study will allow future phylogenetic studies to proceed, and improved understanding of the population genetics of psyllids should help development and testing of new management options.

Supporting Information

S1 Table. A: Codon usage of protein coding genes of \textit{Bactericera cockerelli}, B: \textit{Paratrioza sinica}, C: \textit{Cacopsylla coccinea}, and D: \textit{Pachypsylla venusta}. (DOCX)

S1 Fig. Alignment between AY601890 (3,077 bp, Thao et al., 2004) and the corresponding region in the mitogenome sequence of \textit{Bactericera cockerelli} (KU501214). \textit{cob} = cytochrome oxidase b, \textit{nad} = NADH dehydrogenase subunits, \textit{rrnL} = large ribosomal RNA subunit \textit{rrnS} = small ribosomal RNA subunit. Sequence codes: yellow = protein coding genes, green = rRNA genes, and underlined = tRNA genes. Identical nucleotides are represented by dot “.”, and single nucleotide polymorphisms are indicated by letters in red. (TIF)
Acknowledgments

We thank Ayyamperumal Jeyaprakash and Sonia Vargas for technical support. This research was supported by Chinese Modern Agricultural Technology Systems (CARS-27) and California Citrus Research Board.

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

Author Contributions

Conceived and designed the experiments: JC FW. Performed the experiments: FW SP CW. Analyzed the data: FW CW ZZ XD GL JC. Contributed reagents/materials/analysis tools: YC JT SP RY. Wrote the paper: FW JC CW RY.

References

1. Hansen AK, Trumble JT, Stouthamer R, Paine TD. A new huanglongbing species, “Candidatus Liberibacter psyllaurous,” found to infect tomato and potato, is vectored by the psyllid Bactericera cockerelli (Sulc). Appl Environ Microbiol. 2008; 74(18): 5862–5865. doi:10.1128/AEM.01268-08 PMID: 18676707

2. Liefting LW, Weir BS, Pennycook SR, Clover GR. ’Candidatus Liberibacter solanacearum’, associated with plants in the family Solanaceae. Int J Syst Evol Micr. 2009; 59(9): 2274–2276.

3. Šulc K. Trioza cockerelli n. sp., a novelty from North America, being also of economic importance. Acta Societatis Entomologicae Bohemiae. 1909; 6: 102–108.

4. Butler CD, Trumble JT. The potato psyllid, Bactericera cockerelli (Sulc)(Hemiptera: Triozidae): life history, relationship to plant diseases, and management strategies. Terr Arthropod Rev. 2012; 5(2): 87–111.

5. Knowlton GF, Janes MJ. Studies on the biology of Paratrioza cockerelli (Sulc). Ann Entomol Soc Am. 1931; 24(2): 283–291.

6. Glick PA. The distribution of insects, spiders, and mites in the air. USDA Technical Bulletin. 1939; 673: 150.

7. Papp RP, Johnson J B. Origins of psyllid fallout in the Central Sierra Nevada of California (Homoptera). Pan-Pac Entomol. 1979; 55: 95–98.

8. Liu D, Trumble JT, Stouthamer R. Genetic differentiation between eastern populations and recent introductions of potato psyllid (Bactericera cockerelli) into western North America. Entomol Exp Appl. 2006; 118(3): 177–183.

9. Swisher KD, Munyaneza JE, Crosslin JM. High resolution melting analysis of the cytochrome oxidase I gene identifies three haplotypes of the potato psyllid in the United States. Environ Entomol. 2012; 41: 1019–1028.

10. Swisher KD, Henne DC, Crosslin JM. Identification of a fourth haplotype of Bactericera cockerelli (Hemiptera: Triozidae) in the United States. J Insect Sci. 2014; 14(1): 161.

11. Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. Ann Rev Entomol. 2014; 59: 95–117.

12. Thao ML, Baumann L, Baumann P. Organization of the mitochondrial genomes of whiteflies, aphids, and psyllids (Hemiptera, Sternorrhyncha). BMC Evol Biol. 2004; 4(1): 25.

13. Wang HL, Xiao N, Yang J, Wang XW, Colvin J, Liu SS. The complete mitochondrial genome of Bemisia afer (Hemiptera: Aleyrodidae). Mitochondrial DNA. 2014; (0: ): 1–2.

14. Zhang S, Luo J, Wang C, Lv L, Li C, Jiang W, et al. Complete mitochondrial genome of Aphis gossypii Glover (Hemiptera: Aphididae). Mitochondrial DNA. 2014; (0: ): 1–2.

15. Wang Y, Huang XL, Qiao GX. The complete mitochondrial genome of Cervaphis quercus (Insecta: Hemiptera: Aphiidae: Greendineinae). Insect science. 2014; 21(3): 278–290. doi: 10.1111/1744-7917.12112 PMID: 24482299

16. Que S, Yu L, Xin T, Zou Z, Hu L, Xia B. Complete mitochondrial genome of Cacopsylla coccinae (Hemiptera: Psyllidae). Mitochondrial DNA. 2015; (0: ): 1–2.
17. Cantacessi C, Jex AR, Hall RS, Young ND, Campbell BE, Joachim A, et al. A practical, bioinformatic workflow system for large data sets generated by next-generation sequencing. Nucleic Acids Res. 2010; 38(17): e171. doi: 10.1093/nar/gkq667 PMID: 20682560

18. Williams ST, Foster PG, Littlewood DTJ. The complete mitochondrial genome of a turbidid vetigastropod from MiSeq Illumina sequencing of genomic DNA and steps towards a resolved gastropod phylogeny. Gene. 2014; 533(1): 38–47. doi: 10.1016/j.gene.2013.10.005 PMID: 24126225

19. Zhang QL, Guo ZL, Yuan ML. The complete mitochondrial genome of Poratrioza sinica (Insecta: Hemiptera: Psyllidae). Mitochondrial DNA. 2014; (0): 1–2.

20. Li FS. Psyllidomorpha in China (Insecta: Hemiptera). Volume I. Beijing: Science Press. 2011. (In Chinese)

21. Hall DW, Butler JF, Cromroy HL. Hackberry Petiole Gall Psyllid, Pachypsylla venusta (Osten-Saken). University of Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, EDIS. 1998.

22. Dolling WR. Hemiptera. Oxford University Press. 1991.

23. Wu F, Deng X, Liang G, Cen Y, Wallis C, Trumble JT, et al. Draft genome sequence of “Candidatus Liberibacter solanacearum” strain RSTM from a potato psyllid in California. Genome Announc. 2015; 3(6): e01500–15.

24. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST: architecture and applications. BMC Bioinformatics. 2009; 10: 421. doi: 10.1186/1471-2105-10-421 PMID: 20003500

25. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. Nucleic Acids Research. 2012; 40(15): e115. PMID: 22730293

26. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30: 2725–2729. doi: 10.1093/molbev/mst197 PMID: 24132122

27. Lowe TM, Eddy SR. tRNAscan-SE: a program to detect tRNA genes in genomic sequences. Nucleic Acids Res. 1997; 25: 955–64. PMID: 9023104

28. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, et al. MITOS: Improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013; 69(2): 313–319. doi: 10.1016/j.ympev.2012.08.023 PMID: 22982435

29. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999; 27: 573. PMID: 9862982

30. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic acids research. 2003; 31(13): 3406–3415. PMID: 12824337

31. Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol. 1995; 41(3): 353–358. PMID: 7563121

32. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25: 404–405. PMID: 19346325

33. Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ. Multiple sequence alignment with Clustal X. Trends Biochem Sci. 1998; 23(10): 403–405. PMID: 9810230

34. Jeyaprakash A, Hoy MA. First divergence time estimate of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. Exp Appl Acarol. 2009; 47(1): 1–18. doi: 10.1007/s10493-008-9203-5 PMID: 18931924

35. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003; 52: 696–704. PMID: 14530136

36. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19(12): 1572–1574. PMID: 12912839

37. Posada D. jModelTest: phylogenetic model averaging. Mol Biol Evol. 2008; 25(7): 1253–1256. doi: 10.1093/molbev/msn083 PMID: 18397919

38. Wang Y, Huang XL, Gao GX. Comparative analysis of mitochondrial genomes of five aphid species (Hemiptera: Aphididae) and phylogenetic implications. PLoS ONE. 2013; 8(10): e77511. doi: 10.1371/journal.pone.0077511 PMID: 24147014

39. Ye F, Liu T, King SD, You P. Mitochondrial genomes of two phlebotomine sand flies, Phlebotomus chimensis and Phlebotomus papatasi (Diptera: Nematocera), the first representatives from the family Psychodidae. Parasite Vector. 2015; 8(1): 1–13.

40. Boore JL. Animal mitochondrial genomes. Nucleic Acids Res. 1999; 27(8): 1767–1780. PMID: 10101183
42. Wolstenholme DR. Animal mitochondrial DNA: Structure and evolution. Int Rev Cytol. 1992; 141: 173–216. PMID:1452431
43. Forero D. The systematics of the Hemiptera. Rev Col Entomol. 2008; 34(1): 1–21.
44. Swisher KD, Crosslin JM. Restriction digestion method for haplotyping the potato psyllid, Bactericera cockerelli. Southwest Entomol, 2014; 39(1): 49–56.
45. Powell CM, King J, Bextine BR. Cytochrome B Sequences of Potato Psyllids, Bactericera cockerelli (Sulc), from North and Central America. Southwest Entomol, 2012; 37(4): 521–524.
46. Smith PT, Kambhampati S, Völk W, Mackauer M. A phylogeny of aphid parasitoids (Hymenoptera: Braconidae: Aphidiinae) inferred from mitochondrial NADH 1 dehydrogenase gene sequence. Mol Phylogenet Evol. 1999; 11(2): 236–245. PMID:10191068
47. Kambhampati S. A phylogeny of cockroaches and related insects based on DNA sequence of mitochondrial ribosomal RNA genes. PNAS. 1995; 92(6): 2017–2020. PMID:7534409