American Genetic Association

Journal of Heredity, 2019, 247–256
doi:10.1093/jhered/esy068
Brief Communication
Advance Access publication December 22, 2018

OXFORD

Brief Communication

One Hundred Mitochondrial Genomes of Cicadas

Piotr Łukasik, Rebecca A. Chong, Katherine Nazario, Yu Matsuura, DeAnna C. Bublitz, Matthew A. Campbell, Mariah C. Meyer, James T. Van Leuven, Pablo Pessacq, Claudio Veloso, Chris Simon, and John P. McCutcheon

From the Division of Biological Sciences, University of Montana, Missoula, MT (Łukasik, Bublitz, Campbell, Meyer, Van Leuven, and McCutcheon); the Department of Biology, University of Hawai'i at Mānoa, Honolulu, HI (Chong); the Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT (Nazario and Simon); the Tropical Biosphere Research Center, University of the Ryukyus, Japan (Matsuura); the Centro de Investigaciones Esquel de Montaña y Estepa Patagónicas (CIEMEP), Esquel, Chubut, Argentina (Pessacq); and the Department of Ecological Sciences, Science Faculty, University of Chile, Santiago, Chile (Veloso).

Address correspondence to P. Łukasik at the address above, or e-mail: p.lukasik@gmail.com.

Received June 23, 2018; First decision July 20, 2018; Accepted December 21, 2018.

Corresponding Editor: Warren Booth

Abstract

Mitochondrial genomes can provide valuable information on the biology and evolutionary histories of their host organisms. Here, we present and characterize the complete coding regions of 107 mitochondrial genomes (mitogenomes) of cicadas (Insecta: Hemiptera: Auchenorrhyncha: Cicadoidea), representing 31 genera, 61 species, and 83 populations. We show that all cicada mitogenomes retain the organization and gene contents thought to be ancestral in insects, with some variability among cicada clades in the length of a region between the genes *nad2* and *cox1*, which encodes 3 tRNAs. Phylogenetic analyses using these mitogenomes recapitulate a recent 5-gene classification of cicadas into families and subfamilies, but also identify a species that falls outside of the established taxonomic framework. While protein-coding genes are under strong purifying selection, tests of relative evolutionary rates reveal significant variation in evolutionary rates across taxa, highlighting the dynamic nature of mitochondrial genome evolution in cicadas. These data will serve as a useful reference for future research into the systematics, ecology, and evolution of the superfamily Cicadoidea.

Keywords: Auchenorrhyncha, cicadas, endosymbiosis, intergenic spacer, mitochondria, mitogenome, organelle

While mitochondrial function is primarily controlled by genes encoded on the host nucleus, most mitochondria and mitochondria-related organelles retain their own compact genomes that are faithfully passed down in the host matriline (Rand et al. 2004; Embley and Martin 2006; Gray 2012). Analyses of the organization and contents of these genomes can provide valuable insights into the biology of the hosts, including phylogenetic inferences of relationships among groups of organisms (Simon et al. 1994, 2006; Boore and Brown 1998; Cameron 2014b; Lavrov 2014; Smith and Keeling 2015). Mitochondrial genome comparisons can also help to identify hybridization events (e.g., Good et al. 2008; Marshall et al. 2011; Toews and Brelsford 2012) infer evolutionary processes

that vary among host clades (Woolfit and Bromham 2005; Thomas et al. 2010; Li et al. 2017), and reconstruct biogeographic histories of species (e.g., Marshall et al. 2011; Ma et al. 2012). With DNA sequencing becoming cheaper and easier, the numbers of mitochondrial genomes deposited in public databases have rapidly increased in the last decade (Cameron 2014b). However, these mitochondrial genomic data are not always carefully curated and characterized, and some groups of eukaryotes lack representation in databases of mitochondrial sequences. This lack of representation is true of some groups of Hemiptera, including the superfamily Cicadoidea comprising cicadas and hairy cicadas.

Cicadas (Insecta: Hemiptera: Cicadidae) are a widely distributed family of large sap-feeding insects, known for their loud songs and massive, synchronized emergences in some species (Williams and Simon 1995; Moulds 2005; Marshall et al. 2018). Hairy cicadas (Tettigarctidae) are a primitive family well-represented in Mesozoic fossil records (Moulds 2018) but with only 2 extant species recognized (Moulds 2005). During a decade of work studying microbial symbioses in Cicadoidea, we have generated metagenomic datasets from more than 60 species. Assemblies of these genomic data typically contained large mitochondrial genomic contigs that encompassed the entire gene-encoding region. While these mitogenomic sequences have provided an essential phylogenetic framework for our ongoing research on cicada endosymbiont evolution (Van Leuven et al. 2014; Campbell et al. 2017; Łukasik et al. 2018; Matsuura et al. 2018), they were never the focus of our analyses per se. We reasoned that a curated mitochondrial genome collection could provide a valuable reference for researchers investigating the systematics, ecology, and evolution of various members of the superfamily Cicadoidea, but also for broader studies of mitochondrial genome evolution. Here, we present the complete sequences of the gene-encoding region of the mitochondrial genomes of 107 specimens of Cicadoidea representing 83 populations of 61 species, as well as molecular evolutionary analyses of cicada mitogenomes.

Methods

The specimens included in this study were collected between 1997 and 2017 at multiple locations in North America, South America, Asia, and Oceania (Supplementary Table S1). Cicadas were identified based on morphology and, in some cases, molecular markers. The specimens for sequencing were selected to address specific biological questions related to endosymbiosis in cicadas, resulting in sampling bias toward certain genera (*Tettigades* and *Magicicada*) and geographic areas (southwestern South America and Japan). Some cicada clades and regions are underrepresented or missing entirely from our sample set (Marshall et al. 2018). We also updated 3 cicada mitochondrial genome sequences published by our laboratory previously (Van Leuven et al. 2014), and included 7 other partial cicada mitochondrial genomes that were available from GenBank (Li et al. 2017; Song et al. 2017). These genomes are shaded in Supplementary Table S1. Finally, we included sequences from 4 divergent species from the superfamily Cercopoidea and 3 from the superfamily Membracoidea (Song et al. 2017) to serve as outgroups (Supplementary Table S1).

Whole specimens or dissected tissues were stored in ethanol, RNAlater, or acetone at −20 or −80 °C. Dissected cicada tissues (bacteriomes, fat bodies, or legs; Supplementary Table S1), were used for genomic library preparation using Illumina or NEB kits, or following previously published protocols (Meyer and Kircher 2010; Kircher et al. 2012). The libraries were sequenced using various Illumina technologies (Supplementary Table S1). Genomic reads were quality-trimmed using Trim Galore! (https://github.com/FelixKrueger/TrimGalore), merged into contigs using PEAR (Zhang et al. 2014) and used for assemblies using SPAdes v. 3.7.0. (Bankevich et al. 2012). Mitogenomic contigs were identified in the assemblies using blastn, with previously published cicada mitogenomes as references.

None of the mitogenomic contigs were completely closed into circular-mapping genomes. In all cases, the gene-encoding region was flanked by long stretches of extremely AT-rich, repetitive sequence, for which the read coverage was low. PCR reactions using outward-facing primers targeted to the gene-encoding region for several distantly related cicadas (Supplementary Table S2) and a highly processive polymerase KAPA HiFi HotStart ReadyMix (Kapa Biosystems) resulted in products that were several kilobases (kb) long (Supplementary Figure S1), and these PCR products were consistently accompanied by shorter fragments. Our repeated attempts to Sanger-sequence some of these products confirmed the assembled sequence near the ends of the gene-encoding regions, but the Sanger read quality rapidly decreased beyond these regions. We concluded that while cicada mitogenomes are almost certainly circular-mapping, obtaining the sequences of these AT-rich regions would be prohibitively complicated. Therefore, all mitogenomic contigs were trimmed immediately outside of the outermost genes (see below).

The mitogenome contigs were initially annotated with the MITOS2 on-line server, http://mitos2.bioinf.uni-leipzig.de (Bernt et al. 2013b), using the Metazoan RefSeq 81 set and default settings. The annotations were compared with each other and with the previously annotated mitogenomes of cicadas, the pea aphid *Acyrthosiphon pisum* (accession: FJ411411.1), and the fruit fly *Drosophila melanogaster* (U37541.1). Using blastn searches, we located and manually annotated a small number of tRNA genes that were missed by MITOS2. Because of the variable start and end positions of protein-coding and rRNA genes in the MITOS2 annotations, even among species within a genus, we implemented an additional annotation strategy. We aligned open reading frames and genome regions corresponding to MITOS2-annotated genes and then searched for conserved start-stop positions in the alignments. This included incomplete stop codons (T or TA) that are known to be frequent in mitogenomes of other insects (e.g., Sheffield et al. 2008; Du et al. 2017) and thought to be extended into complete stop codon (TAA) through post-transcriptional polyadenylation (Ojala et al. 1981; Stewart and Beckenbach 2009). Data on intergenic distances in pea aphid and fruit fly mitogenomes were used as a guide. We verified stop codon positions of some protein-coding genes using data on mitogenome transcript coverage for one specimen for which such data were available, *Tettigades chilensis* PL470 (P. Łukasik, unpublished). The data were visualized using Python and Processing scripts.

The annotated gene sequences were aligned using mafft v. 7.2.2.1 (Katoh and Standley 2013), in nucleotide space for tRNA and rRNA genes and in amino acid space for protein-coding genes. The alignments were verified manually using CodonCode Aligner v. 7.1.2 (CodonCode Corporation), and ends with <20% coverage trimmed. For the maximum likelihood phylogenetic analysis, we used the concatenated alignment of all genes, divided into 5 partitions: one for each of the 3 codon positions of protein-coding sequences, and one each for rRNA and tRNA genes; this partitioning scheme was based on our prior work using PartitionFinder2 (Lanfear et al. 2012; Łukasik et al. 2018). We used RAxML v. 8.2.10 (Stamatakis 2014), specifying GTR model with the gamma distribution of rates (among the best models for different partitions based on jmodeltest2—Darriba et al. 2012), and one hundred rapid bootstraps.

A Bayesian phylogenetic tree was also estimated using a concatenated dataset of all genes, divided into 3 partitions: one for the first codon position of all protein-coding sequences combined with all rRNA and tRNA genes, and one each for the second and third codon positions of protein-coding sequences (Nylander 2004; Brandley et al. 2005). Data partitions and the best-fitting nucleotide substitution model for each partition were determined using PartitionFinder2 and BIC model selection (Supplementary Table S3). We obtained posterior distributions of trees and parameters using Markov chain Monte Carlo (MCMC) sampling procedure implemented in MrBayes v. 3.2.6. (Ronquist et al. 2012) with

the selected best-fitting substitution models and default priors. We used 2 independent runs (each with 4 chains) with 10 million generations, sampling every 1000 generations, and discarded the first 25% of samples as burnin. We checked for convergence of continuous parameters using Tracer v. 1.6.1 (Rambaut et al. 2014) and considered a run to be converged when parameter had effective sample size values well over 200. Convergence of tree topology was assessed using the RWTY v. 1.0.1 package (Warren et al. 2017) implemented in R v. 3.4.4 (R Core Team 2018).

To examine relative rates of evolution across the cicada phylogeny, we used a Bayesian approach to reconstruct phylogenetic trees with BEAST2 v. 2.4.8 (Bouckaert et al. 2014). We examined relative rates of evolution for 3 different datasets representing (1) all cicada genera, (2) *Magicicada* species, and (3) *Tettigades* species, with the latter 2 datasets representing clades with denser taxonomic sampling. The cicada genera dataset included one individual for each genus while the other 2 datasets included only one representative for each species, and outgroup taxa were excluded from all datasets. Using methods described above, we determined the best-fitting substitution model for the concatenated dataset of all 13 mitochondrial coding sequences partitioned by codon position for each of the 3 datasets (Supplementary Table S3). Substitution models were unlinked between partitions but shared a clock model and tree model, and applied an uncorrelated lognormal clock model to each dataset. For the cicada genera dataset, we ran 4 independent MCMC analyses with 75 million generations each, sampling every 5000 generations, while the *Magicicada* and *Tettigades* datasets were run for 50 million generations each and sampled every 5000 generations. The first 25% of trees were discarded as burnin and convergence was assessed as described above.

The strength and direction of selection (ω, the ratio of nonsynonymous to synonymous substitutions, or *d*N/*d*S) for mitochondrial genes was computed using codeml implemented in PAML v. 4.8 (Yang 2007). Analyses were performed for the same 3 datasets used in the BEAST2 analyses (i.e., all cicada genera, *Magicicada* species, and *Tettigades* species). We calculated ω for individual mitochondrial protein-coding genes as well as the concatenated data of all 13 coding sequences for each of the 3 datasets. Specifically, we generated maximum likelihood phylogenies for each gene and the concatenated data partitioned by codon position for all 3 datasets using RAxML v. 8.2.10, and estimated nonsynonymous and synonymous substitution rates along each phylogeny.

Results and Discussion

Cicada Mitogenomes Retain the Canonical Arthropod Organization, with Some Intergenic Length Variation

We find that all 117 mitochondrial genomes of cicadas analyzed here retain the conserved set of metazoan mitochondrial genes, comprising 13 protein-coding genes, 2 rRNA genes, and 22 tRNA genes (Boore 1999). The gene order and orientation within the gene-encoding region is identical to the reconstructed ancestral insect mitogenome (Cameron 2014b) (Figure 1A). The length of the gene-encoding region ranges from 14 350 basepairs (bp) in *Vagitanus terminalis* to 14 740 bp in *Magicicada tredecassini*. Most of the length variation is contained within the region between genes *nad2* and *cox1*, which encodes 3 tRNA genes, for tryptophan (W), cysteine (C), and tyrosine (Y) (Figure 1B). This region ranges from 183 bp (*Tettigarcta crinita*) to 552 bp (*M. tredecassini*) (Figure 1B). This variation seems to result from intergenic spacer expansions that have happened independently in different cicada clades, as indicated by differences in relative length of spacers and varying, sometimes extreme, nucleotide composition (Figure 1B). In species or species groups where intergenic spacer expansion has occurred (*Magicicada cassini*/*M. tredecassini*, *Hyalessa maculaticollis*), individuals can differ substantially, suggesting that the onset of spacer expansion may be associated with broader genomic instability (Burger et al. 2003). The length of other spacer regions was less variable and larger expansions only occurred in isolated cicada clades (Figure 1C). Intergenic spacers in mitochondrial genomes are variable among and within animal species. In vertebrates, the W-A-N-C-Y tRNA region (corresponding to the cicada variable-length W-C-Y region) is the site of the origin of light-strand replication but in insects both strands have their origin of replication in the A+T-rich control region (Saito et al. 2005), so the expansion of the W-C-Y region must have some other explanation. Some turtles and snakes have a lengthy noncoding sequence between *trnN* and *trnC* genes whereas crocodilians, sphenodon, and birds lack this sequence (Seutin et al. 1994). Blackspot seabream (fish) have high levels of length heteroplasmy in the W-A-N-C-Y region within and among individuals (Ponce *et al.* 2008). Differences in other intergenic region lengths have also been reported. For example, the mitogenome of honeybee is characterized by a spacer (250–650 bp) between the COI and COII genes, the length of which varies within and between subspecies (Cornuet et al. 1991). However, changes in the gene order and genome organization have received more attention. Departures from the ancestral insect gene order have been reported in at least 13 insect orders, with some orders, such as Hymenoptera, clearly more variable than others (Dowton et al. 2009; Simon and Hadrys 2013; Cameron 2014b). The instability of the W-C-Y region is evident in cicadellid planthoppers, where *Japananus hyalinus* exhibits the only gene rearrangement so far known in Auchenorrhyncha; W-C-Y becomes Y-W-C (Du et al. 2017). A different transposition in this region can be seen in some whiteflies (Hemiptera: Sternorrhyncha: Aleyrodidae) where W-C-Y changes to W-Y-C (Thao et al. 2004) and in Neuroptera where it changes to C-W-Y (Negrisolo et al. 2011). In many cases, changes to the mitogenome organization were much more dramatic (Smith and Keeling 2015). Among the best-known examples is the fragmentation of the mitochondrial genome into 18 mini-circles encoding 1–3 genes each in the human body louse (Shao et al. 2009). Thus, in comparison to mitogenomes of many other animals, those of cicadas can be regarded as stable.

Cicada Mitogenomes Have Large Control Regions

As explained in the Methods section, we were not able to reconstruct the sequence of the AT-rich control region for any of the mitogenomes, but for several cicadas we obtained PCR products that appeared to span that region. The products typically consisted of multiple bands, the longest of which was often the most intensely stained in ethidium bromide agarose gels (Supplementary Figure S1). This suggests that cicada mitochondrial genomes map as circular molecules, but also that the control regions likely contain large tandem low-complexity repeats, as seen in some other Auchenorrhyncha (Du et al. 2017) and many other organisms (Zhang and Hewitt 1997). An alternative explanation for these banding patterns can be heteroplasmy—the presence of more than one variant of a mitochondrial genome in the studied specimens (Ramos et al. 2013; Rebolledo-Jaramillo et al. 2014), as reported in some weevils (Boyce et al. 1989). The approximate length of the longest and strongest PCR band for *T. crinita* (the sole member of the family Tettigarctidae) was 1.7 kb, and for diverse Cicadidae the longest bands ranged from approximately 3–5 kb (Supplementary Figure S1). Assuming that these products represented the complete control region plus approximately 300 bp of the flanking sequence, we estimate the complete mitochondrial genome size of cicadas to be approximately 15.7 kb in *T. crinita*, and 17–19.5 kb for the majority of Cicadidae. These estimates are consistent with the results of a restriction digest

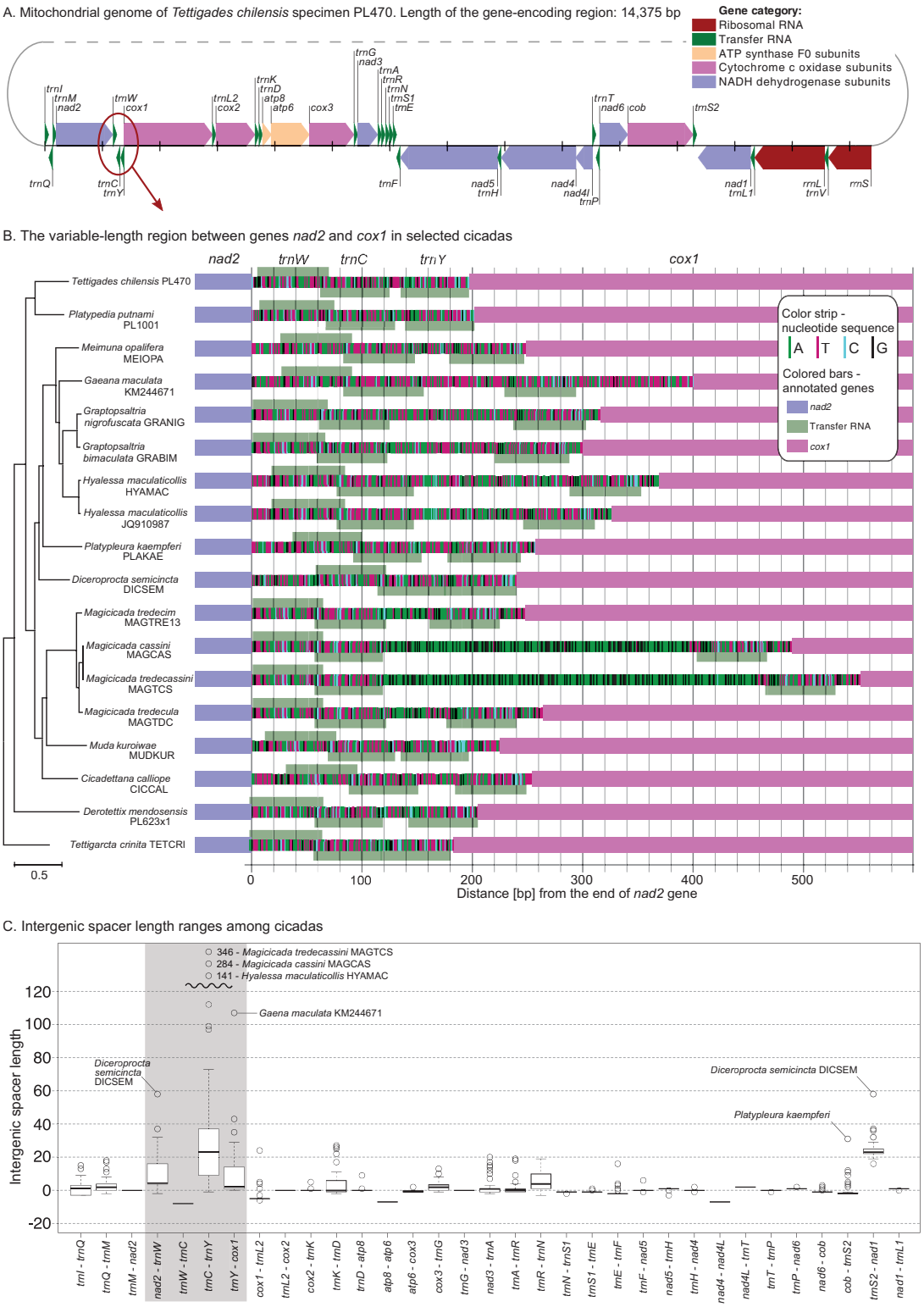

Figure 1. (**A**) The organization and orientation of genes within the gene-encoding portion of the mitochondrial genome of *Tettigades chilensis* specimen PL470, representative for the 117 studied cicadas. The genomes are circular, but the reliable sequences of the AT-rich control regions could not be obtained. (**B**) The visualization of the variable-length region between genes *nad2* and *cox1* in selected cicada specimens, including all those with the most extreme intergenic interval lengths. The nucleotide sequences of the region demonstrate the repetitive nature and base composition bias in some of the expanded intergenic regions. For maximum likelihood phylogeny reconstruction, see Figure 2; all nodes have ≥90% bootstrap support. (C) The range of lengths of intergenic spacer regions in 71 studied cicadas (one from each *Tettigades* sp., one *Magicicada tredecim*, and all specimens from other species). The variable region corresponding to panel B is shaded. Negative spacer length indicates an overlap among adjacent genes. Spacers around rRNA genes are not shown, since their lengths were used for rRNA gene delimitation.

study of mitogenomes from the *Magicicada* "decim" clade: the estimated sizes of the 2 genome "types" were 19.6 ± 0.5 kb and 20.0 ± 0.5 kb (Martin and Simon 1990). This range is well above the typical size of an insect mitogenome reported as complete (NCBI, accessed on 31 May 2018), but it is by no means exceptional. For example, the complete mitogenome of *D. melanogaster* (Genbank accession U37541) is 19 517 bp, and in some weevils, mitochondrial genomes were reported to be over 30 kb (Boyce et al. 1989). The 33 mitochondrial genomes of Auchenorrhyncha that have been deposited in NCBI databases as complete (and this is not obviously incorrect) range in size from 15 131 to 16 626 bp, with the exception of *Nilaparvata lugens* (JX880069) at 17 619 bp. However, sequencing of the control regions is often challenging, and in a large proportion of insect mitogenomes in public databases the region is missing or only partial (Cameron 2014a), sometimes even when the genomes are annotated as complete and circular.

Cicada Mitogenomes Have Relatively High AT Content

We find the AT contents of the gene-encoding region of cicada mitogenomes ranges from 72.3% in *Mendozana platypleura* to 80.5% in *Terpnosia vacua* and shows some phylogenetic signal (Supplementary Table S1; Figure 2). Mitochondrial genomes of other organisms are known to vary considerably in nucleotide content (Smith 2012). Typically, the value is substantially higher than 50% AT, but there are significant differences among animal clades with deuterostomes (chordates and echinoderms) typically less AT-biased than protostomes (e.g., the AT% of fish mitogenomes is consistently low) (Smith 2012). In insect mitogenomes annotated as complete, AT% ranges from high 80s in Hymenoptera to mid- and lower 60s in termites, stoneflies, and isolated species from other groups (although the absence of the control region in many NCBI records complicates comparisons). From early sequences of selected insect mitochondrial genes, it was suggested that there was a trend toward increasing AT bias from basal nodes to shallower nodes in the insect phylogeny (Simon et al. 1994). However, whole mitogenomes that have accumulated since that time suggest that this generalization does not hold up due to large variation in AT-content within and among orders and families (NCBI). For example, in beetles, AT-contents of the coding region ranges from 66% to 80% (Sheffield et al. 2008). Substanial variability in AT content is also true of the family Cicadidae. With the overall AT content of about 83% (conservatively assuming that the trimmed control region is 3 kb long and contains 95% A+T), *T. vacua* may be among the top few percent of all insects, while *M. platypleura* at about 75% A+T would be close to the median.

Variation in mitochondrial AT content is thought to result from a variety of processes, including an often strong bias toward "C" and "A" on one DNA strand resulting from varying mutational pressures during replication and transcription (GC-skew and AT-skew; Perna and Kocher 1995; Hassanin et al. 2005; Bernt et al. 2013a; Chong and Mueller 2013a), and perhaps an overall mutational pressure towards A+T in bacterial genomes (Hershberg and Petrov 2010; Hildebrand et al. 2010). Among factors thought to drive the overall AT-GC% change in mitogenomes are the rate of metabolism, generation time, and lifestyle (Martin 1995; Smith 2012). In mammals, GC% is positively correlated with the generation time, suggesting that increased levels of natural selection in long-lived species may mitigate the mutational tendency of mitochondrial genomes to increase the contents of A+T (Min and Hickey 2008). Cicada species vary considerably in generation times but data is available for only a few species in our set (Campbell et al. 2015). Also, genome rearrangements may influence the nucleotide distribution within the genomes (Hassanin et al. 2005), but we have not observed rearrangements in cicadas.

No Clear Roles of tRNA Anticodon Sequence Changes or Sequence Variants

An interesting observation was an anticodon sequence change within a tRNA gene, which has happened at least 3 times in cicadas. In *T. crinita*, the sole representative in our study of the family Tettigarctidae, anticodon on the gene *trnA* changed from "TGC" to "CGC," relative to the ancestral state represented, for example, by *D. melanogaster* and all Cicadidae. Also, in 2 *Kikihia* species, anticodon on the gene *trnS1* changed from "GCT" to "TCT," and in related *Amphipsalta zelandica*, to "ACT." These were all silent changes that did not affect codon pairings: tRNAs with the modified anticodon encode the same amino acid—alanine (codon: GCN) and serine (codon: AGN), respectively. The *trnS1* anticodon changes from "GCT" to "TCT" have been reported from several insect groups, including many beetles, hymenopterans, and lice (Cameron et al. 2007; Sheffield et al. 2008; Kaltenpoth et al. 2012), while other anticodon sequence changes are relatively less common. Their functional significance is unknown.

We looked for sequence variants (polymorphic sites and insertions/deletions, or indels) by mapping reads to finished, consensus genomes. We found that a substantial proportion of reads contained sequence variants, and that some of these changes disrupted open reading frames of protein-coding genes. We compared read alignments against reference genomic sequences for 2 populations for which 6 or more specimens were sequenced, and had an average read coverage of at least 10× (*Tettigades undata* and *Magicicada tredecim*; Supplementary Figure S2). We found that within the majority of genomes, several nucleotide positions contained substitutions or deletions in at least 10% of mapped reads. Many of these sites were variable in multiple specimens. It is possible that these patterns are due to heteroplasmy (Ramos et al. 2013; Rebolledo-Jaramillo et al. 2014). However, the interpretation of these patterns is challenging because of the possibility of mitochondrial genome introgression events into the nuclear genome, or numts (Bensasson et al. 2001; Hazkani-Covo et al. 2010), as well as sequencing artifacts. We know that numts are common in some species of *Tettigades* (Łukasik et al. 2018).

Phylogenies Confirm the Monophyly of Main Cicada Subfamilies, and Suggest a New One

Our maximum likelihood (ML) phylogenetic tree (Figure 2) agrees with phylogenies generated from broader samplings of species, based on either morphology (Moulds 2005) or concatenations of mitochondrial and nuclear gene sequences obtained using the Sanger method (Marshall et al. 2018). Our Bayesian phylogenetic tree is consistent with our ML phylogeny (Figure 2, Supplementary Figure S3). Based on these phylogenetic analyses, *T. crinita* (Tettigarctidae) represents a clade that is sister to all Cicadidae. The sampled Cicadidae form 3 distinct, well-supported clades corresponding to the 3 recognized, established subfamilies: Tibicininae, Cicadettinae, and Cicadinae. An interesting exception is *Derotettix mendosensis*, which appears to represent a clade within Cicadidae that is sister to the monophyletic clade comprising these 3 subfamilies. While the genus *Derotettix* is currently assigned, based on prior morphological comparisons, to tribe Parnisini within Cicadettinae (Marshall et al. 2018), our data suggest its taxonomic position may need to be revisited. At lower taxonomic levels, our analyses agree with phylogenetic reconstructions that were based on other, partially overlapping sample sets and various genes (Sota et al. 2013; Marshall et al. 2016; Łukasik et al. 2018). Phylogenetic analyses also reveal substantial genetic variation

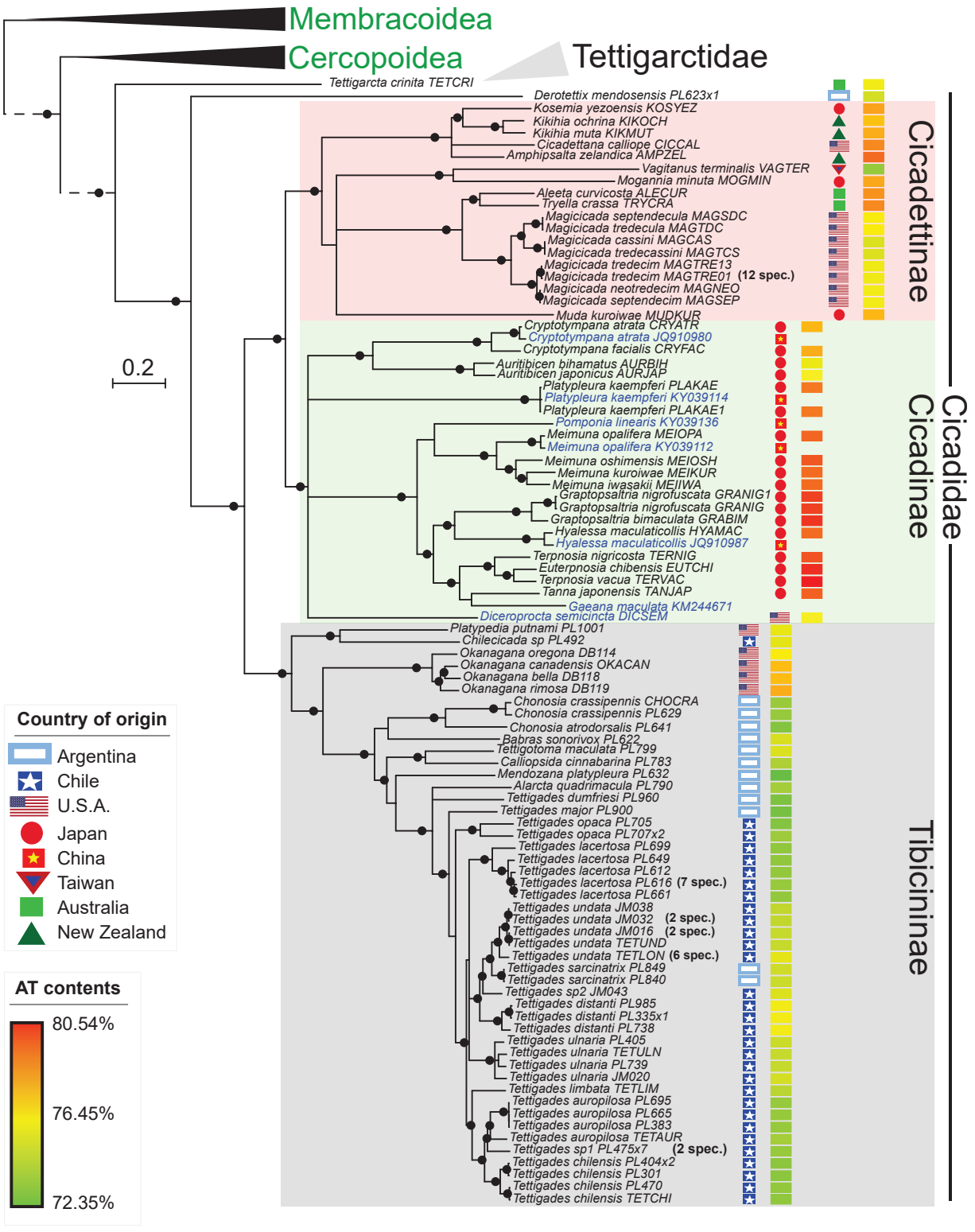

Figure 2. The maximum likelihood tree based on 37 concatenated mitochondrial genes for representative specimens from 93 cicada populations. Nodes with bootstrap support <70% were collapsed. Nodes with bootstrap support values of 100%, as well as independently calculated Bayesian posterior probabilities of 1.0, are represented by black dots. In cases when more than one specimen per population was sequenced, data for only one is shown. Symbols adjacent to specimen names indicate the country of collection, as well as the AT contents of the gene-encoding region. Sequences from GenBank, for which read alignments have not been inspected, are typed in blue font; for some of them the gene-encoding region was not complete, preventing us from calculating AT%, and in one case the collection location was missing.

among cicada specimens that were classified to a single species based on morphology but were sampled at different locations, including *Tettigades lacertosa*, *T. undata*, and *T. ulnaria* (Łukasik et al. 2018). The relationships among these divergent populations and borders among *Tettigades* species need to be systematically revisited. However, our sampling bias toward certain cicada genera, the absence of many important cicada clades in our dataset, and the known limitations of relying solely on mitochondrial sequence data for phylogenetic inference prevent us from making definitive taxonomic conclusions.

Rates of Evolution Vary among the Cicada Clades

We find rapid changes in the relative rates of mitochondrial gene evolution across the cicada phylogeny (Figure 3A). We observed a substantial slow down at the base of Tibicininae (Figure 3A) with some lineages maintaining slower rates (e.g., *Platypedia putnami* + *Chilecicada sp.*) while the majority of the lineages return to an average level. We also observed both slow-downs and accelerations within the clade consisting of subfamilies Cicadinae and Cicadettinae (Figure 3A). Specifically, some taxa maintain an accelerated rate (e.g., *Mogannia minuta* + *V. terminalis*) while others decelerate (e.g., *Diceroprocta semicincta*).

To estimate rates of evolution within more densely sampled genera, we reconstructed Bayesian phylogenies for the *Magicicada* and *Tettigades* datasets. Within *Magicicada*, we observe a slight increase in rates at the base of the "decim" species group, followed by a decrease in *M. septendecim* + *M. neotredecim*. We observe an overall rate increase in the "cassini" species group compared with a deceleration in the "decula" species group (Figure 3B). The genus *Magicicada* is known for massive,

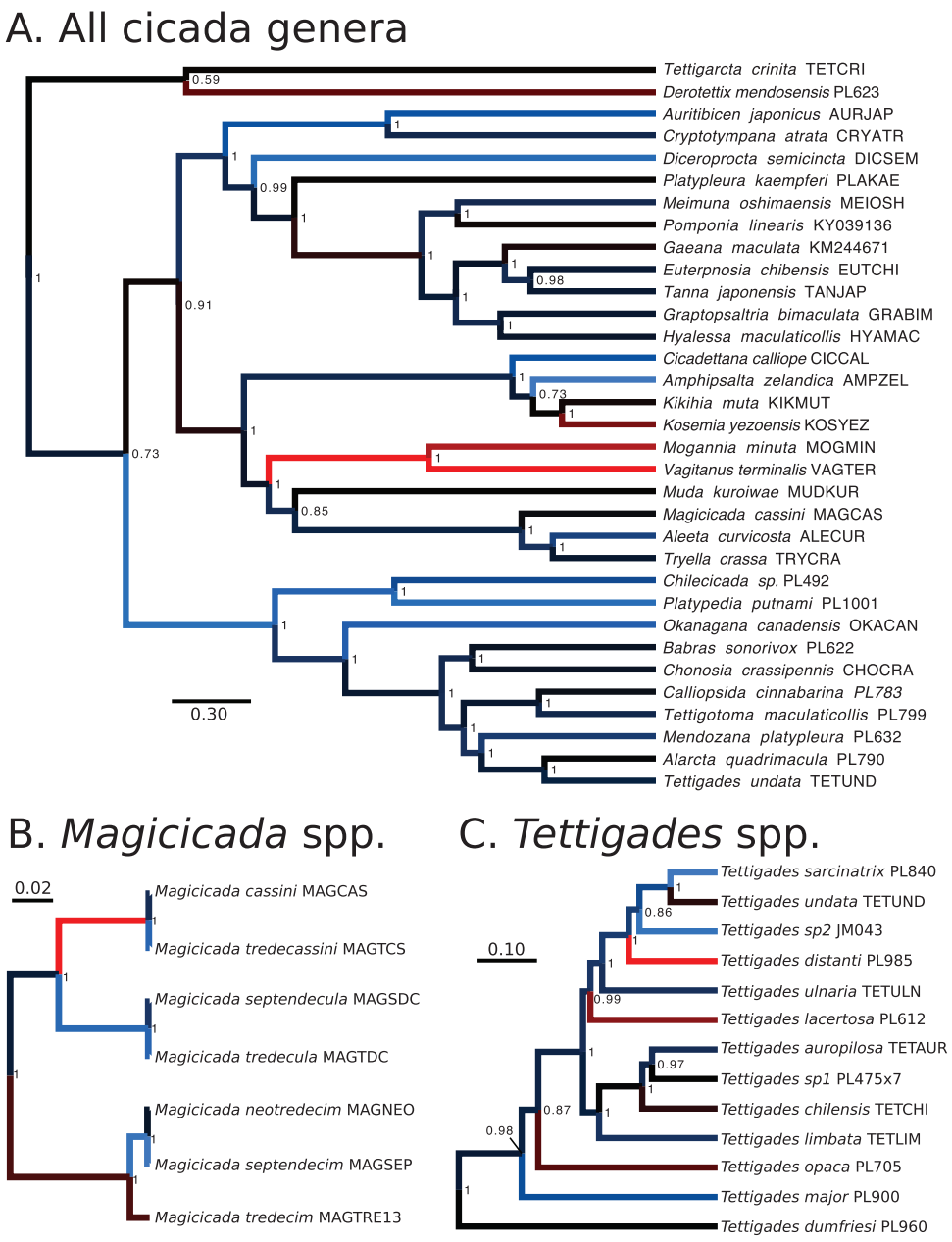

Figure 3. Bayesian phylogenetic tree reconstruction of evolutionary rates using all 13 mitochondrial protein coding genes for (**A**) all cicada genera, (**B**) *Magicicada* species, and (**C**) *Tettigades* species. Numbers represent node support (BPP). Colors reflect relative rates of evolution, where black represents a relative rate of one and shades of red represent intermediate rate increases and shades of blue represent intermediate decreases in rates along branches.

synchronized emergences that occur every 13 or 17 years, and that typically consist of several species from across the species groups (Williams and Simon 1995). Whereas some ecological differences among the *Magicicada* species groups are known, it is unclear how they may influence the evolutionary rates. Within *Tettigades*, we observe more significant variation in rates across lineages. Specifically, several species show increases in relative rates of evolution including *Tettigades opaca*, *Tettigades lacertosa*, and *Tettigades distanti*, while other species, including *Tettigades major, Tettigades sp2*, and *Tettigades sarcinatrix,* show decrease (Figure 3C). There is little information on the general biology of the genus *Tettigades*, but notably, *T. lacertosa* and *T. distanti* specimens represent populations where their specialized, maternally transmitted, nutritional endosymbiont *Hodgkinia cicadicola* have undergone substantial structural changes ("splits") relatively recently (Łukasik et al. 2018). Interestingly, in the cicada genera comparison, the slowly evolving *D. semicincta* harbors *Hodgkinia* that has not undergone structural rearrangements (Van Leuven et al. 2014), and 2 of the fastest-evolving clades in the all-cicada comparison, *V. terminalis* and *Kosemia yezoensis*, harbor degenerated *Hodgkinia* (Matsuura et al. 2018) (Figure 3A). On the other hand, the genus *Magicicada*, hosting the most fragmented *Hodgkinia* genomes known (Campbell et al. 2017), does not stand out from among other cicadas in rates of evolution (Figure 3A). Also, the Japanese cicada clades where *Hodgkinia* has been replaced by *Ophiocordyceps* fungi (Matsuura et al. 2018) did not experience consistent rate of evolution changes: in some cases it has accelerated (*Mogannia*), in other cases decelerated (*Amphipsalta*), but often lineages maintained rates similar to the underlying rate across the cicada phylogeny. In most other cicada species in our dataset, the content and structure of the symbioses await description. However, taken together, these results suggest that there is significant variation in rates of evolution of mitochondrial genes across the cicada phylogeny, only some of which may correlate with endosymbiont complexity. Similarly variable rates of mitochondrial gene evolution within insect taxa have been reported in other insects (e.g., Pons et al. 2010). Future work will reveal the likely causes, and how other maternally transmitted genomes (nuclear and those of the endosymbionts) may have been affected.

Cicada Mitochondrial Genes Are under Strong Purifying Selection

Estimates of ω (the ratio of nonsynonymous and synonymous rates of substitution, or *d*N/*d*S) were used to infer the strength and direction of selection across mitochondrial genes and reveal strong levels of purifying selection (ω << 1) across all of the datasets examined (Figure 4). We observe that the *cox1* gene is consistently under the strongest level of purifying selection, which is consistent with results from other studies of selection on mitochondrial genes (Simon et al. 1994; Shen et al. 2009; Chong and Mueller 2013b). In contrast to *cox1*, we find that mitochondrial genes *atp8* and *nad2* consistently show elevated ω values, suggesting that these 2 genes are under relaxed levels of selection relative to other mitochondrial genes. Based on all protein coding genes, we observe higher ω values in *Tettigades* compared with the other 2 datasets suggesting that this group may experience increased differences in substitution rates, which is congruent with results from the BEAST2 analysis examining relative rates of evolution. Collectively, these results highlight the role of both gene specific and lineage specific substitution rates in driving mitochondrial genome evolution in cicadas. Changes in relative substitution rates may reflect differences in population size or changes in biological factors related to metabolic rate, generation time, or parasitic lifestyles (Dowton and Austin 1995; Woolfit and Bromham 2005; Thomas et al. 2010; Chong and Mueller 2013b), even between species within the same genus.

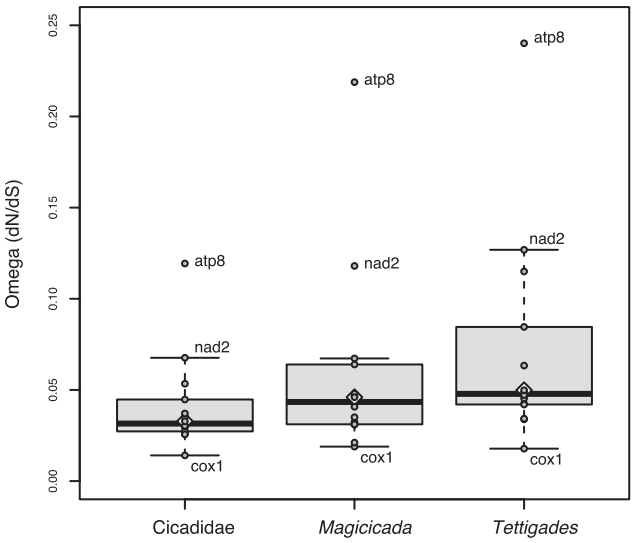

Figure 4. Estimates of ω (nonsynonymous and synonymous substitution rate ratio, *d*N/*d*S), across all cicada species, *Magicicada* species, and *Tettigades* species show strong levels of purifying selection on mitochondrial protein-coding genes. Elevated ω values for mitochondrial genes *atp8* and *nad2* reflect relaxed selection relative to other mitochondrial genes, while *cox1* is consistently under the strongest level of purifying selection across all datasets. White diamonds represents ω values for the concatenated dataset of all 13 mitochondrial coding-genes.

Conclusions

We show that cicada mitochondrial genomes are remarkably stable. The 117 specimens in our study, representing much of the known cicada diversity (Marshall et al. 2018), retain the ancestral genome organization and gene set (Cameron 2014b). The patterns we observe in some species, such as substantial increases in the length of *trnC-trnY* intergenic region in *M. cassini* and *M. tredecassini* relative to their sister clade, might suggest an onset of genomic instability or differences in selective constraints on mitochondrial genomes. But unlike some other organisms whose mitogenomes have fragmented into mini-circles or undergone other dramatic alterations (Smith and Keeling 2015), it is not clear whether these changes will lead to dramatic genomic rearrangements. The mitogenomic stability we show here contrasts sharply with the patterns reported in the genomes of the cicadas' maternally transmitted, rapidly evolving endosymbiont *Hodgkinia,* which in most cicada clades studied to date has undergone major structural changes (Van Leuven et al. 2014; Campbell et al. 2017; Łukasik et al. 2018) or has been lost altogether and replaced by a fungus (Matsuura et al. 2018).

We also observe variation in the rate of mitogenome sequence evolution across the entire cicada phylogeny. Overall, we find signatures of strong purifying selection on mitochondrial genes, though some genes show slightly relaxed levels of selection (e.g., *atp8* and *nad2*). These results suggest there is substantial variation in rate of evolution both across lineages and between genes, which highlight the dynamic nature of mitochondrial genome evolution across the cicada phylogeny.

Supplementary Material

Supplementary material is available at *Journal of Heredity* online.

Funding

This study was supported by National Science Foundation (IOS-1256680, IOS-1553529 and DEB-1655891); National Aeronautics and Space Administration Astrobiology Institute Award (NNA15BB04A); National Geographic Society (9760-15); American Genetics Association EECG Research Award 2015; and the University of Hawai'i at Mānoa startup funds to RAC from.

Author Contributions

P.Ł. and J.P.M. conceived this study. P.Ł., K.N., and all subsequent authors participated in specimen collection, identification, genomic library preparation, and/or sequencing. P.Ł. assembled, curated, and annotated the genomes. P.Ł. and R.A.C. analyzed and illustrated the data. P.Ł., R.A.C., C.S. and J.P.M. wrote the article.

Data Availability

All genomes were deposited in GenBank, under accessions MG674192–MG674196 and MG737715–MG737816. Accessions for individual genomes are listed in Supplementary Table S1.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, *et al*. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19:455–477.

Bensasson D, Zhang D, Hartl DL, Hewitt GM. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol*. 16:314–321.

Bernt M, Braband A, Schierwater B, Stadler PF. 2013a. Genetic aspects of mitochondrial genome evolution. *Mol Phylogenet Evol*. 69:328–338.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF. 2013b. MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Mol Phylogenet Evol*. 69:313–319.

Boore JL. 1999. Animal mitochondrial genomes. *Nucleic Acids Res*. 27:1767–1780.

Boore JL, Brown WM. 1998. Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. *Curr Opin Genet Dev*. 8:668–674.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 10:e1003537.

Boyce TM, Zwick ME, Aquadro CF. 1989. Mitochondrial DNA in the bark weevils: size, structure and heteroplasmy. *Genetics*. 123:825–836.

Brandley MC, Schmitz A, Reeder TW. 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. *Syst Biol*. 54:373–390.

Burger G, Gray MW, Lang BF. 2003. Mitochondrial genomes: anything goes. *Trends Genet*. 19:709–716.

Cameron SL. 2014a. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Syst Entomol*. 39:400–411.

Cameron SL. 2014b. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu Rev Entomol*. 59:95–117.

Cameron SL, Johnson KP, Whiting MF. 2007. The mitochondrial genome of the screamer louse *Bothriometopus* (phthiraptera: ischnocera): effects of extensive gene rearrangements on the evolution of the genome. *J Mol Evol*. 65:589–604.

Campbell MA, Łukasik P, Simon C, McCutcheon JP. 2017. Idiosyncratic genome degradation in a bacterial endosymbiont of periodical cicadas. *Curr Biol*. 27:3568.e3–3575.e3.

Campbell MA, Van Leuven JT, Meister RC, Carey KM, Simon C, McCutcheon JP. 2015. Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont Hodgkinia. *Proc Natl Acad Sci USA*. 112:10192–10199.

Chong RA, Mueller RL. 2013a. Evolution along the mutation gradient in the dynamic mitochondrial genome of salamanders. *Genome Biol Evol*. 5:1652–1660.

Chong RA, Mueller RL. 2013b. Low metabolic rates in salamanders are correlated with weak selective constraints on mitochondrial genes. *Evolution*. 67:894–899.

Cornuet JM, Garnery L, Solignac M. 1991. Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. *Genetics*. 128:393–403.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 9:772.

Dowton M, Austin AD. 1995. Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the Hymenoptera. *J Mol Evol*. 41:958–965.

Dowton M, Cameron SL, Dowavic JI, Austin AD, Whiting MF. 2009. Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. *Mol Biol Evol*. 26:1607–1617.

Du Y, Zhang C, Dietrich CH, Zhang Y, Dai W. 2017. Characterization of the complete mitochondrial genomes of *Maiestas dorsalis* and *Japananus hyalinus* (Hemiptera: Cicadellidae) and comparison with other Membracoidea. *Sci Rep*. 7:14197.

Embley TM, Martin W. 2006. Eukaryotic evolution, changes and challenges. *Nature*. 440:623–630.

Good JM, Hird S, Reid N, Demboski JR, Steppan SJ, Martin-Nims TR, Sullivan J. 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. *Mol Ecol*. 17:1313–1327.

Gray MW. 2012. Mitochondrial evolution. *Cold Spring Harb Perspect Biol*. 4:a011403.

Hassanin A, Léger N, Deutsch J. 2005. Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inferences. *Syst Biol*. 54:277–298.

Hazkani-Covo E, Zeller RM, Martin W. 2010. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet*. 6:e1000834.

Hershberg R, Petrov DA. 2010. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genet*. 6:e1001115.

Hildebrand F, Meyer A, Eyre-Walker A. 2010. Evidence of selection upon genomic GC-content in bacteria. *PLoS Genet*. 6:e1001107.

Kaltenpoth M, Showers Corneli P, Dunn DM, Weiss RB, Strohm E, Seger J. 2012. Accelerated evolution of mitochondrial but not nuclear genomes of Hymenoptera: new evidence from crabronid wasps. *PLoS One*. 7:e32826.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30:772–780.

Kircher M, Sawyer S, Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res*. 40:e3.

Lanfear R, Calcott B, Ho SY, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol*. 29:1695–1701.

Lavrov DV. 2014. Mitochondrial genomes in invertebrate animals. In: Bell E, editor. *Molecular life sciences: an encyclopedic reference*. New York (NY): Springer New York. p. 1–8.

Li H, Leavengood JM, Chapman EG, Burkhardt D, Song F, Jiang P, Liu J, Zhou X, Cai W. 2017. Mitochondrial phylogenomics of Hemiptera reveals adaptive innovations driving the diversification of true bugs. *Proc R Soc B Biol Sci*. 284:20171223.

Łukasik P, Nazario K, Van Leuven JT, Campbell MA, Meyer M, Michalik A, Pessacq P, Simon C, Veloso C, McCutcheon JP. 2018. Multiple origins of interdependent endosymbiotic complexes in a genus of cicadas. *Proc Natl Acad Sci USA*. 115:E226–E235.

Ma C, Yang P, Jiang F, Chapuis MP, Shali Y, Sword GA, Kang L. 2012. Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. *Mol Ecol*. 21:4344–4358.

Marshall DC, Hill KB, Cooley JR, Simon C. 2011. Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: lessons from New Zealand cicadas (genus *Kikihia*). *Syst Biol*. 60:482–502.

Marshall DC, Hill KB, Moulds M, Vanderpool D, Cooley JR, Mohagan AB, Simon C. 2016. Inflation of molecular clock rates and dates: molecular phylogenetics, biogeography, and diversification of a global cicada radiation from Australasia (Hemiptera: Cicadidae: Cicadettini). *Syst Biol*. 65:16–34.

Marshall DC, Moulds M, Hill KBR, Price BW, Wade EJ, Owen CL, Goemans G, Marathe K, Sarkar V, Cooley JR, *et al*. 2018. A molecular phylogeny of the cicadas (Hemiptera: Cicadidae) with a review of tribe and subfamily classification. *Zootaxa*. 4424:1–64.

Martin AP. 1995. Metabolic rate and directional nucleotide substitution in animal mitochondrial DNA. *Mol Biol Evol*. 12:1124–1131.

Martin A, Simon C. 1990. Differing levels of among-population divergence in the mitochondrial DNA of periodical cicadas related to historical biogeography. *Evolution*. 44:1066–1080.

Matsuura Y, Moriyama M, Łukasik P, Vanderpool D, Tanahashi M, Meng XY, McCutcheon JP, Fukatsu T. 2018. Recurrent symbiont recruitment from fungal parasites in cicadas. *Proc Natl Acad Sci USA*. 115:E5970–E5979

Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*. 2010:pdb.prot5448.

Min XJ, Hickey DA. 2008. An evolutionary footprint of age-related natural selection in mitochondrial DNA. *J Mol Evol*. 67:412–417.

Moulds M. 2005. An appraisal of the higher classification of cicadas (Hemiptera: Cicadoidea) with special reference to the Australian fauna. *Rec Aust Mus*. 57:375–446.

Moulds MS. 2018. Cicada fossils (Cicadoidea: Tettigarctidae and Cicadidae) with a review of the named fossilised Cicadidae. *Zootaxa*. 4438:443–470.

Negrisolo E, Babbucci M, Patarnello T. 2011. The mitochondrial genome of the ascalaphid owlfly *Libelloides macaronius* and comparative evolutionary mitochondriomics of neuropterid insects. *BMC Genomics*. 12:221.

Nylander JAA. 2004. *MrModeltest v2. Program distributed by the author*. Uppsala: Evolutionary Biology Centre, Uppsala University.

Ojala D, Montoya J, Attardi G. 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature*. 290:470–474.

Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol*. 41:353–358.

Ponce M, Infante C, Jiménez-Cantizano RM, Pérez L, Manchado M. 2008. Complete mitochondrial genome of the blackspot seabream, *Pagellus bogaraveo* (Perciformes: Sparidae), with high levels of length heteroplasmy in the WANCY region. *Gene*. 409:44–52.

Pons J, Ribera I, Bertranpetit J, Balke M. 2010. Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. *Mol Phylogenet Evol*. 56:796–807.

Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014. Tracer v1.6. Available from: http://beast.bio.ed.ac.uk/Tracer

Ramos A, Santos C, Mateiu L, Gonzalez MDM, Alvarez L, Azevedo L, Amorim A, Aluja MP. 2013. Frequency and pattern of heteroplasmy in the complete human mitochondrial genome. *PLoS One*. 8:e74636.

Rand DM, Haney RA, Fry AJ. 2004. Cytonuclear coevolution: the genomics of cooperation. *Trends Ecol Evol*. 19:645–653.

R Core Team. 2018. *R: a language and environment for statistical computing*. Vienna (Austria): R Foundation for Statistical Computing.

Rebolledo-Jaramillo B, Su MSW, Stoler N, McElhoe JA, Dickins B, Blankenberg D, Korneliussen TS, Chiaromonte F, Nielsen R, Holland MM, *et al*. 2014. Maternal age effect and severe germ-line bottleneck in the inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA*. 111:15474–15479.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61:539–542.

Saito S, Tamura K, Aotsuka T. 2005. Replication origin of mitochondrial DNA in insects. *Genetics*. 171:1695–1705.

Seutin G, Lang BF, Mindell DP, Morais R. 1994. Evolution of the WANCY region in amniote mitochondrial DNA. *Mol Biol Evol*. 11:329–340.

Shao R, Kirkness EF, Barker SC. 2009. The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. *Genome Res*. 19:904–912.

Sheffield NC, Song H, Cameron SL, Whiting MF. 2008. A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Mol Biol Evol*. 25:2499–2509.

Shen YY, Shi P, Sun YB, Zhang YP. 2009. Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. *Genome Res*. 19:1760–1765.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu Rev Ecol Evol Syst*. 37:545–579.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am*. 87:651–701.

Simon S, Hadrys H. 2013. A comparative analysis of complete mitochondrial genomes among Hexapoda. *Mol Phylogenet Evol*. 69:393–403.

Smith DR. 2012. Updating our view of organelle genome nucleotide landscape. *Front Genet*. 3:175.

Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proc Natl Acad Sci USA*. 112:10177–10184.

Song N, Cai W, Li H. 2017. Deep-level phylogeny of Cicadomorpha inferred from mitochondrial genomes sequenced by NGS. *Sci Rep*. 7:10429.

Sota T, Yamamoto S, Cooley JR, Hill KBR, Simon C, Yoshimura J. 2013. Independent divergence of 13- and 17-y life cycles among three periodical cicada lineages. *Proc Natl Acad Sci USA*. 110:6919–6924.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30:1312–1313.

Stewart JB, Beckenbach AT. 2009. Characterization of mature mitochondrial transcripts in *Drosophila*, and the implications for the tRNA punctuation model in arthropods. *Gene*. 445:49–57.

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

Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010. A generation time effect on the rate of molecular evolution in invertebrates. *Mol Biol Evol*. 27:1173–1180.

Toews DP, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol Ecol*. 21:3907–3930.

Van Leuven JT, Meister RC, Simon C, McCutcheon JP. 2014. Sympatric speciation in a bacterial endosymbiont results in two genomes with the functionality of one. *Cell*. 158:1270–1280.

Warren DL, Geneva AJ, Lanfear R. 2017. RWTY (R We There Yet): an R package for examining convergence of Bayesian phylogenetic analyses. *Mol Biol Evol*. 34:1016–1020.

Williams KS, Simon C. 1995. The ecology, behavior, and evolution of periodical cicadas. *Annu Rev Entomol*. 40:269–295.

Woolfit M, Bromham L. 2005. Population size and molecular evolution on islands. *Proc R Soc B Biol Sci*. 272:2277–2282.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24:1586–1591.

Zhang D-X, Hewitt GM. 1997. Insect mitochondrial control region: a review of its structure, evolution and usefulness in evolutionary studies. *Biochem Syst Ecol*. 25:99–120.

Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate illumina paired-end reAd mergeR. *Bioinformatics*. 30:614–620.