Journal of Asia-Pacific Entomology 21 (2018) 252–257

Contents lists available at ScienceDirect

Journal of Asia-Pacific Entomology

journal homepage: www.elsevier.com/locate/jape

Complete mitochondrial genome of *Bactrocera ritsemai* (Insecta: Tephritidae) and phylogenetic relationship with its congeners and related tephritid taxa

Sze-Looi Song[a,*], Hoi-Sen Yong[b], I. Wayan Suana[c], Phaik-Eem Lim[a]

[a] Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia
[b] Institute of Biological Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Faculty of Science and Mathematics, Mataram University, Mataram, Indonesia

ARTICLE INFO

Keywords:
Bactrocera ritsemai
Mitogenome
Phylogeny
Fruit flies

ABSTRACT

Bactrocera ritsemai is a dacine fruit fly found in Indonesia. We report here the complete mitogenome of this fruit fly from Lombok, Indonesia determined by Illumina MiSeq sequencing and its phylogenetic relationship with its congeners and related tephritid taxa. The whole mitogenome of *B. ritsemai* had a total length of 15,927 bp, comprising 37 genes – 13 protein-coding genes (PCGs), 2 ribosomal ribonucleic acid (rRNA) and 22 transfer ribonucleic acid (tRNA) genes – and a control region (D-loop). Of the PCGs, 6 (*atp6*, *cob*, *cox2*, *cox3*, *nad4*, *nad4l*) had ATG start codon, 4 (*nad2*, *nad3*, *nad5*, *nad6*) had ATT, and one each had ATA (*nad1*), GTG (*atp8*) and TCG (*cox1*). Seven PCGs (*atp6*, *atp8*, *cox2*, *cox3*, *nad2*, *nad4l*, *nad6*) had TAA stop codon, 3 (*cob*, *nad3*, *nad4*) had TAG, and 3 had incomplete stop codon (*cox1* – TA; *nad1*, *nad5* – T). The TΨC-loop of tRNA was absent in *trnF* while *trnS1* lacked the DHU-loop. Phylogenetic analysis based on 15 mt-genes (13 PCGs + 2 rRNA genes) indicated *B. ritsemai* forming a sister group with *B. umbrosa* and the subgenus *Bactrocera* was monophyletic. The genera *Bactrocera* and *Zeugodacus* were monophyletic while the subfamilies Dacinae and Tephritinae were paraphyletic. A broader taxa sampling of the Tephritidae is needed to better elucidate the phylogenetics and systematics of the tribes and subfamilies of tephritid fruit flies.

Introduction

Fruit flies of the genus *Bactrocera* are of economic importance in agriculture (White and Elson-Harris, 1992). They are represented by some 651 described species (Vargas et al., 2015). Seventy-three species have been documented as economically important in the Pacific Region and many are highly polyphagous (Vargas et al., 2015).

The genus *Bactrocera* is represented by some 218 species in Southeast Asia, of which 89% are endemic to the region (Drew, 2004). Among them, *B. ritsemai* (Wayenbergh), a member of the subgenus *Bactrocera*, is found only in Indonesia (Java, Lombok and Sulawesi) (Drew and Romig, 2013; Hardy, 1983). The male is attracted to Cue lure. There is no known record of its host plant (Drew and Romig, 2013).

To date, the complete mitochondrial genomes (mitogenomes) of nine species of the subgenus *Bactrocera* – *B. arecae*, *B. carambolae*, *B. correcta*, *B. dorsalis* (including the conspecific taxa *B. invadens*, *B. papayae* and *B. philippinensis*), *B. latifrons*, *B. melastomatos*, *B. tryoni*, *B. umbrosa*, and *B. zonata* – are available in GenBank. However, there is no report on the complete mitogenome of *B. ritsemai* and its phylogenetic relationship with other taxa of the subgenus *Bactrocera* as well as its congeners of other subgenera and related tephritid taxa. In this study, we compared the gene synteny of *B. ritsemai* with other congeners and generated a robust 15-mt genes phylogeny using the complete mitogenome gene sequences. Compared to partial individual genes, mitogenomes provide more gene contents for systematics and phylogenetic analyses.

Materials and methods

Ethics statement

B. ritsemai is an insect pest. It is not endangered or protected by law. No permits are required to study this fruit fly.

Specimen collection

Male fruit flies of *B. ritsemai* were collected by H.-S. Yong and I. Wayan Suana in Karandangan Forest, Lombok, Indonesia on 6 November 2015 by means of Cue lure according to the method of Yong

* Corresponding author.
E-mail address: szelooi@um.edu.my (S.-L. Song).

https://doi.org/10.1016/j.aspen.2017.12.009
Received 17 August 2017; Received in revised form 20 October 2017; Accepted 28 December 2017
Available online 29 December 2017

et al. (2015). The specimens were preserved in 95% absolute ethanol and stored in − 20 °C freezer until use.

Mitochondria isolation and DNA extraction

A small piece of the alcohol-preserved tissue was pressed onto a C-fold paper towel to remove excess ethanol before homogenisation. The mitochondria were isolated by standard differential centrifugation method (White and Densmore, 1992) and the mtDNA was extracted using Mitochondrial DNA Isolation Kit (Abnova, Taiwan) following the manufacturer's instructions.

Library preparation and genome sequencing

The purified mtDNA was quantified using Qubit dsDNA High Sensitivity Assay Kit (Life Technologies, USA) and normalized to a final concentration of 50 ng. Library was prepared using Nextera DNA Sample Preparation Kit (Illumina, USA) following the manufacturer's protocols. Size estimation of the library was performed on a 2100 Bioanalyzer using High Sensitivity DNA Analysis Kit (Agilent Technologies). The library was quantified with Qubit 2.0 Fluorometer (Life Technologies, USA). The library was sequenced using the Illumina MiSeq Dekstop Sequencer (2 × 150 bp paired-end reads) (Illumina, USA).

Sequence and genome analysis

Raw sequence reads were extracted from the Illumina MiSeq system in FASTQ format and the quality of sequences was evaluated using the FastQC software (Andrews, 2010). All the ambiguous nucleotides and reads with an average quality value lower than Q20 were excluded from further analysis. De novo assembly was performed using the CLC Genomic Workbench v.8.0.1 (https://www.qiagenbioinformatics.com/). Contigs greater than 15 kbp were subjected to BLAST alignment against the nucleotide database at National Center for Biotechnology Information (NCBI). Contigs with hits to mitochondrial genes or genomes were identified and extracted from CLC Genomic Workbench.

Mitogenome identification, annotation and visualization

A single contig which blasted as mitochondrial sequence was manually examined for repeats at the beginning and end of the sequence to establish a circular mtDNA. The mitogenome was then annotated by manual validation of the coding regions using the NCBI ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder), with the genetic code option of "Invertebrate Mitochondrial", minimal ORF length of 75 nucleotides and ORF start codon of "ATG" and alternative start codon. The sequin file generated from MITOS was edited and submitted to NCBI according to ORF Finder result. The circular mitogenome was visualized with Blast Ring Image Generator (BRIG) (Alikhan et al., 2011). The mitogenome sequence has been deposited in GenBank – accession number MF668132.

Mitogenomes from GenBank

The mitogenomes of Tephritidae available from GenBank – *Bactrocera arecae* NC_028327 (Yong et al., 2015); *Bactrocera carambolae* NC_009772 (unpublished); *Bactrocera correcta* NC_018787 (unpublished); *Bactrocera dorsalis* complex [*B. dorsalis* NC_008748 (unpublished); *B. papayae* NC_009770 (unpublished); *B. philippinensis* NC_009771 (unpublished); *B. invadens* NC_031388 (Zhang et al., 2016a,b)]; *Bactrocera latifrons* NC_029466 (Yong et al., 2016a); *Bactrocera melastomatos* NC_029467 (Yong et al., 2016a); *Bactrocera tryoni* NC_014611 (Nardi et al., 2010); *Bactrocera umbrosa* NC_029468 (Yong et al., 2016a); *Bactrocera zonata* NC_027725 (Choudhary et al., 2015); *Bactrocera* (*Daculus*) *oleae* NC_005333 (Nardi et al., 2003); *B.* (*Tetradacus*) *minax* NC_014402 (unpublished); *Zeugodacus caudatus* Malaysia KT625491 (Yong et al., 2016b); *Zeugodacus caudatus* Indonesia KT625492 (Yong et al., 2016b); *Zeugodacus cucurbitae* NC_016056 (unpublished); *Zeugodacus depressus* KY131831 (Jeong et al., 2017); *Zeugodacus diaphorus* NC_028347 (Zhang et al., 2016a,b); *Zeugodacus scutellatus* NC_027254 (unpublished); *Zeugodacus tau* NC_027290 (Tan et al., 2016); *Ceratitis capitata* NC_000857 (Spanos et al., 2000); *Ceratitis fasciventris* KY436396 (Drosopoulou et al., 2017); *Dacus longicornis* NC_032690 (Jiang et al., 2016); *Anastrepha fraterculus* NC_034912 (Isaza et al., 2017); *Procecidochares utilis* NC_020463 (unpublished) – were used for phylogenetic comparison. Species of *Drosophila* – *D. incompta* NC_025936 (De Re et al., 2014); *D. melanogaster* NC_024511 (unpublished); *D. yakuba* NC_001322 (Clary and Wolstenholme, 1985) – were used as outgroup taxa.

Phylogenetic analysis

Nucleotide sequences of the 15 mt-genes (13 PCGs + 2 rRNA genes) were separately aligned using MAFFT version 7 (Katoh and Standley, 2013) and subsequently edited and trimmed to equal lengths using BioEdit v.7.0.5.3 (Hall, 1999). The best-fit nucleotide substitution models for maximum likelihood (ML) using the corrected Akaike Information Criterion (Akaike, 1973) and Bayesian (BI) analyses using the Bayesian Information Criterion (Schwarz, 1978) were determined by Kakusan v.3 (Tanabe, 2007). Phylograms of 15 mt-genes were constructed using TreeFinder (Jobb et al., 2004) with bootstrap values (BP) generated via 1000 ML bootstrap replicates. Bayesian analyses were conducted using the Markov chain Monte Carlo (MCMC) method via Mr. Bayes v.3.1.2 (Huelsenbeck and Ronquist, 2001), with two independent runs of 2×10^6 generations with four chains, and with trees sampled every 200th generation. Likelihood values for all post-analysis trees and parameters were evaluated for convergence and burn-in using the "sump" command in MrBayes and the computer program Tracer v.1.5 (http://tree.bio.ed.ac.uk/software/tracer/). The first 200 trees from each run were discarded as burn-in (where the likelihood values were stabilized prior to the burn-in), and the remaining trees were used for the construction of a 50% majority-rule consensus tree. Phylogenetic trees were viewed and edited by FigTree v.1.4 (Rambaut, 2012). The total nucleotide sequences of concatenated 15 mt-genes was 13,465 bp with AIC model = GTR + Gamma and BIC model = SYM + Gamma (nst = 6).

Results and discussion

Mitogenome features

The complete mitochondrial genome of *B. ritsemai* had a total length of 15,927 bp, comprising 37 genes – 13 protein-coding genes (PCGs), 2 ribosomal ribonucleic acid (rRNA) and 22 transfer ribonucleic acid (tRNA) genes – and a control region (D-loop) (Table 1, Fig. 1). The gene order conforms to other tephritid mitogenomes (Drosopoulou et al., 2017; Isaza et al., 2017; Jeong et al., 2017; Jiang et al., 2016; Yong et al., 2015, 2016a, 2016b).

There were 14 intergenic regions with spacing sequence ranging from 2 to 84 bp, and 9 regions with overlapping sequence ranging from − 1 to − 8 bp (Table 1). The largest intergenic sequence (84 bp) was between *trnQ* and *trnM*. Large intergenic sequence between *trnQ* and *trnM* is present in other taxa of the subgenus *Bactrocera* (Table S1), for example, 55 bp in *B. arecae* (Yong et al., 2015), 94 bp in *B. latifrons*, 82 bp in *B. melastomatos* and 79 bp in *B. umbrosa* (Yong et al., 2016a). In contrast, the largest intergenic sequence in the genus *Zeugodacus* (previously designated as a subgenus of genus *Bactrocera*) was located between *trnR* and *trnN* (Table S1). It is noteworthy that the two subgenera *Daculus* and *Tetradacus* of genus *Bactrocera* did not possess the largest intergenic space between *trnQ* and *trnM* (Table S1). Sequences with 30, 33 and 84 bases in *B. ritsemai* had clear stem-loop structures as