

RESEARCH ARTICLE

Complete mitochondrial genome of *Zeugodacus tau* (Insecta: Tephritidae) and differentiation of *Z. tau* species complex by mitochondrial cytochrome *c* oxidase subunit I gene

Hoi-Sen Yong¹, Sze-Looi Song^{2*}, Phaik-Eem Lim^{2*}, Praphathip Eamsobhana³

1 Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia, **2** Institute of Ocean and Earth Sciences, University of Malaya, Kuala Lumpur, Malaysia, **3** Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand

* szelooi@um.edu.my (SLS); phaikeem@um.edu.my (PEL)



OPEN ACCESS

Citation: Yong H-S, Song S-L, Lim P-E, Eamsobhana P (2017) Complete mitochondrial genome of *Zeugodacus tau* (Insecta: Tephritidae) and differentiation of *Z. tau* species complex by mitochondrial cytochrome *c* oxidase subunit I gene. PLoS ONE 12(12): e0189325. <https://doi.org/10.1371/journal.pone.0189325>

Editor: Zuogang Peng, SOUTHWEST UNIVERSITY, CHINA

Received: September 17, 2017

Accepted: November 23, 2017

Published: December 7, 2017

Copyright: © 2017 Yong et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Funding for this work was provided in part by University of Malaya Research Grant H-5620009.

Competing interests: The authors have declared that no competing interests exist.

Abstract

The tephritid fruit fly *Zeugodacus tau* (Walker) is a polyphagous fruit pest of economic importance in Asia. Studies based on genetic markers indicate that it forms a species complex. We report here (1) the complete mitogenome of *Z. tau* from Malaysia and comparison with that of China as well as the mitogenome of other congeners, and (2) the relationship of *Z. tau* taxa from different geographical regions based on sequences of cytochrome *c* oxidase subunit I gene. The complete mitogenome of *Z. tau* had a total length of 15631 bp for the Malaysian specimen (ZT3) and 15835 bp for the China specimen (ZT1), with similar gene order comprising 37 genes (13 protein-coding genes—PCGs, 2 rRNA genes, and 22 tRNA genes) and a non-coding A + T-rich control region (D-loop). Based on 13 PCGs and 15 mt-genes, *Z. tau* NC_027290 (China) and *Z. tau* ZT1 (China) formed a sister group in the lineage containing also *Z. tau* ZT3 (Malaysia). Phylogenetic analysis based on partial sequences of *cox1* gene indicates that the taxa from China, Japan, Laos, Malaysia, Bangladesh, India, Sri Lanka, and *Z. tau* sp. A from Thailand belong to *Z. tau* sensu stricto. A complete *cox1* gene (or 13 PCGs or 15 mt-genes) instead of partial sequence is more appropriate for determining phylogenetic relationship.

Introduction

Zeugodacus tau (Walker) is the most common tephritid fruit fly species of the genus *Zeugodacus* found in Southeast Asia [1]. It is among the economically important species belonging to the Dacinae subfamily, occurring from Pakistan to Philippines and south to Indonesia [2]. It is a polyphagous fruit pest, infesting host fruits of the families Anacardiaceae, Cucurbitaceae, Elaeocarpaceae, Moraceae, Myrtaceae, Oxalidaceae, Rutaceae, Sapotaceae, and Solanaceae [3–7]. The adult male flies are attracted to Cue lure.

Studies based on cytogenetics, partial sequences of mitochondrial cytochrome *c* oxidase subunit I (*cox1*) gene and allozymes have revealed that *Z. tau* (previously referred to as

Bactrocera tau (Walker)) is a species complex comprising eight species (or morphs) in Thailand, with species A designated as *Z. tau* sensu stricto [8–10]. *Z. tau* A may be reliably separated from *Z. tau* B, C, D, E, F, G, and I by the heat shock protein 70 cognate gene *Bthsc1* [11].

Phylogenetic analysis using mitochondrial *cox1* gene sequences revealed that the *Z. tau* population in Himachal Pradesh (India) is closely related to *Z. tau* sp. A from Thailand [12]. The overall genetic variability in this Indian taxon is substantial, with 10 different haplotypes detected in 16 individuals. A study of 23 *Z. tau* populations (Myanmar and western Yunnan; Laos and southern Yunnan; central Yunnan; Thailand; southern China, central China and northern Vietnam; and southwestern China), based on mitochondrial NADH dehydrogenase gene (*nad1*), revealed six genetic groups corresponding to geographical characteristics, and strong genetic structure for the populations in western China, Thailand, and Laos [13]. *Z. tau* in China has also been reported to exhibit seven cytochrome *b* haplotypes (NCBI GenBank 26-JUL-2016: AY953491–AY953497).

To date, there is only a single report on the complete mitochondrial genome (mitogenome) of *Z. tau* [14]. The taxon is from Shenzhen, China. We report here the complete mitogenome of *Z. tau* from Malaysia and compare it to that of China as well as the mitogenome of other congeners. We also carry out phylogenetic analysis using *cox1* gene to determine the relationship of *Z. tau* taxa from different geographical regions.

Materials and methods

Specimen collection and mitochondrial DNA extraction

Male fruit flies of *Z. tau* were collected in Malaysia (Kuala Lumpur– 3.1390°N, 101.6869°E) and China (Zhuhai, Guangdong– 22.2710°N, 113.5767°E) by means of Cue lure according to the method of Yong *et al.* [15]. The specimens were preserved in absolute ethanol and stored in -20°C freezer until use. *Z. tau* is an insect pest. It is not endangered or protected by law. No permits are required to study this fruit fly. The extraction of mitochondrial DNA was according to the method of Yong *et al.* [16].

Library preparation, genome sequencing and analysis

Sample and library preparation (using Nextera DNA Sample Preparation Kit), genome sequencing using the Illumina MiSeq Desktop Sequencer (2 × 150 bp paired-end reads) (Illumina, USA), and genome analysis were as described in Yong *et al.* [15–16]. The mitogenome sequences have been deposited in GenBank—accession number MF966383 (ZT1) and MF966384 (ZT3).

Mitogenomes and cytochrome *c* oxidase subunit I sequences from GenBank

The complete mitogenomes of Tephritidae available from GenBank (Table 1) were used for phylogenetic comparison. Species of *Drosophila*—*D. incompta* Wheeler & Takada NC_025936 [17]; *D. melanogaster* Meigen NC_024511 (unpublished); and *D. yakuba* Burla NC_001322 [18]—were used as outgroup taxa. Representative *cox1* sequences of *Z. tau* from different geographic regions were used for reconstruction of phylogenetic tree.

Phylogenetic analysis

Alignment of nucleotide sequences and reconstruction of phylograms based on 15 mt-genes and cytochrome *c* oxidase subunit I gene sequences followed that described in Yong *et al.* [26].

Table 1. Complete mitogenomes of Tephritidae available from GenBank.

Taxon	Accession no	Reference
<i>Bactrocera arecae</i> (Hardy & Adachi)	NC_028327	[15]
<i>Bactrocera carambolae</i> Drew & Hancock	NC_009772	Unpublished
<i>Bactrocera correcta</i> (Bezzi)	NC_018787	Unpublished
<i>Bactrocera dorsalis</i> (Hendel)	NC_008748	Unpublished
<i>Bactrocera dorsalis</i> (Hendel) (= <i>papayae</i> Drew & Hancock)	NC_009770	Unpublished
<i>Bactrocera dorsalis</i> (Hendel) (= <i>philippinensis</i> Drew & Hancock)	NC_009771	Unpublished
<i>Bactrocera dorsalis</i> (Hendel) (= <i>invadens</i> Drew, Tsura & White)	NC_031388	[19]
<i>Bactrocera latifrons</i> (Hendel)	NC_029466	[20]
<i>Bactrocera melastomatos</i> Drew & Hancock	NC_029467	[20]
<i>Bactrocera ritsemai</i> (Weyenbergh)	MF668132	Unpublished
<i>Bactrocera tryoni</i> (Froggatt)	NC_014611	[21]
<i>Bactrocera umbrosa</i> (Fabricius)	NC_029468	[20]
<i>Bactrocera zonata</i> (Saunders)	NC_027725	[22]
<i>Bactrocera (Daculus) oleae</i> (Rossi)	NC_005333	[23]
<i>Bactrocera (Tetradacus) minax</i> (Enderlein)	NC_014402	Unpublished
<i>Zeugodacus caudatus</i> (Fabricius) Malaysia	KT625491	[24]
<i>Zeugodacus caudatus</i> (Fabricius) Indonesia	KT625492	[24]
<i>Zeugodacus cucurbitae</i> (Coquillett)	NC_027254	Unpublished
<i>Zeugodacus depressus</i> Shiraki	KY131831	[25]
<i>Zeugodacus diaphorus</i> (Hendel)	NC_028347	[26]
<i>Zeugodacus scutellatus</i> (Hendel)	NC_027254	Unpublished
<i>Zeugodacus tau</i> (Walker)	NC_027290	[14]
<i>Ceratitis capitata</i> (Wiedemann)	NC_000857	[27]
<i>Ceratitis fasciventris</i> (Bezzi)	KY436396	[28]
<i>Dacus longicornis</i> Wiedemann	NC_032690	[29]
<i>Anastrepha fraterculus</i> (Wiedemann)	NC_034912	[30]
<i>Procecidochares utilis</i> Stone	NC_020463	Unpublished

<https://doi.org/10.1371/journal.pone.0189325.t001>

Results

Mitogenome features

The raw/final sequencing reads produced by next-generation sequencing on Illumina MiSeq Sequencer were 3286014/3205571 for *Z. tau* ZT3 (Malaysia) and 3191750/3113181 for *Z. tau* ZT1 (China).

The complete mitogenome of *Z. tau* had a total length of 15631 bp for the Malaysian specimen (ZT3) and 15835 bp for the China specimen (ZT1), with similar gene order comprising 37 genes (13 protein-coding genes—PCGs, 2 rRNA genes, and 22 tRNA genes) and a non-coding A + T-rich control region (D-loop) (Table 2, Fig 1, S1 and S2 Tables). The control region was flanked by *rrnS* and *trnI* genes respectively, with 745 bp in *Z. tau* ZT3 and 946 bp in *Z. tau* ZT1. It contained a long polyT-stretch of 14 bp in *Z. tau* ZT3 and 19 bp in *Z. tau* ZT1. It also contained in both taxa a long poly A-stretch (20 bp) after 'ATAGA' motif.

There were 16 intergenic regions with spacing sequence and 9 regions with overlaps in both *Z. tau* ZT3 and *Z. tau* ZT1. The region between *trnR* and *trnN* genes in both taxa was separated by the largest sequence of 34 bp. This sequence had clear stem-loop structures.

Z. tau ZT3 and *Z. tau* ZT1 had identical start/stop codons for the 13 PCGs (Table 2, S1 and S2 Tables). Of the start codons, the commonest was ATG (in 6 PCGs—*cox2*, *atp6*, *cox3*, *nad4*, *nad4l*, *cob*), followed by four ATT (*nad2*, *atp8*, *nad5*, *nad6*), two ATA (*nad3*, *nad1*) and one

Table 2. Gene order and features of mitochondrial genome of *Zeugodacus tau*. NC_027290 (China), ZT1 (China), ZT3 (Malaysia).

Gene	Size (bp)	Size (bp)	Size (bp)	Intergenic sequence (bp)	Start/stop codons
	NC_027290	ZT1	ZT3	NC_027290:ZT1:ZT3	NC_027290:ZT1:ZT3
<i>trnI(gat)</i>	66	66	66	-3:-3:-3	
<i>trnQ(ttg)</i>	69	69	69	8:8:8	
<i>trnM(cat)</i>	69	69	69		
<i>nad2</i>	1023	1023	1023	9:9:9	All: ATT/TAA
<i>trnW(tca)</i>	68	68	68	-9:-8:-8	
<i>trnC(gca)</i>	66	66	63	-1:1:1	
<i>trnY(gta)</i>	67	67	67	-2:-2:-2	
<i>cox1</i>	1534	1534	1534		All: TCG/T
<i>trnL2(taa)</i>	66	66	66	4:4:4	
<i>cox2</i>	690	690	690	5:5:5	All: ATG/TAA
<i>trnK(ctt)</i>	71	71	71		
<i>trnD(gtc)</i>	67	67	67		
<i>atp8</i>	162	162	162	-7:-7:-7	All: ATT/TAA
<i>atp6</i>	678	678	678	-1:-1:-1	All: ATG/TAA
<i>cox3</i>	789	789	789	6:6:6	All: ATG/TAA
<i>trnG(tcc)</i>	65	65	65	-3:-3:-3	
<i>nad3</i>	357	357	357	4:4:4	All: ATA/TAA
<i>trnA(tgc)</i>	66	66	4	4:4:4	
<i>trnR(tcg)</i>	64	64	64	34:34:34	
<i>trnN(gtt)</i>	65	65	65		
<i>trnS1(gct)</i>	68	68	68		
<i>trnE(ttc)</i>	68	68	68	18:18:18	
<i>trnF(gaa)</i>	66	66	66		
<i>nad5</i>	1720	1720	1720	15:15:15	All: ATT/T
<i>trnH(gtg)</i>	65	65	65	3:3:3	
<i>nad4</i>	1341	1341	1341	-7:-7:-7	All: ATG/TAA
<i>nad4l</i>	297	297	297	2:2:2	All: ATG/TAA
<i>trnT(tgt)</i>	65	65	65		
<i>trnP(tgg)</i>	66	66	66	2:2:2	
<i>nad6</i>	525	525	525	-1:-1:-1	All: ATT/TAA
<i>cob</i>	1137	1137	1137	-2:-2:-2	All: ATG/TAG
<i>trnS2(tga)</i>	67	67	67	-65:15:15	
<i>nad1</i>	1020	940	940	10:10:10	ATA/TAA:ATA/T:ATA/T
<i>trnL1(tag)</i>	65	65	65		
<i>rrnL</i>	1327	1327	1327		
<i>trnV(tac)</i>	72	72	72		
<i>rrnS</i>	792	792	792		
Control region	801	946	745		
Total size	15687	15835	15631		

<https://doi.org/10.1371/journal.pone.0189325.t002>

TCG (*cox1*). Nine PCGs had a TAA stop codon (*nad2*, *cox2*, *atp8*, *atp6*, *cox3*, *nad3*, *nad4*, *nad4l*, *nad6*), one had TAG (*cob*), and three had truncated T stop codon (*cox1*, *nad5*, *nad1*).

The nucleotide compositions of the mitochondrial whole genome, protein-coding genes, rRNA genes and control region of *Z. tau* ZT3 and *Z. tau* ZT1 are summarized in S3 and S4 Tables. Both were A+T rich as expected for mitochondrial genomes. The A + T content for PCGs was lowest in *cox3* (64.8% for *Z. tau* ZT3, and 64.6% for *Z. tau* ZT1) and highest in

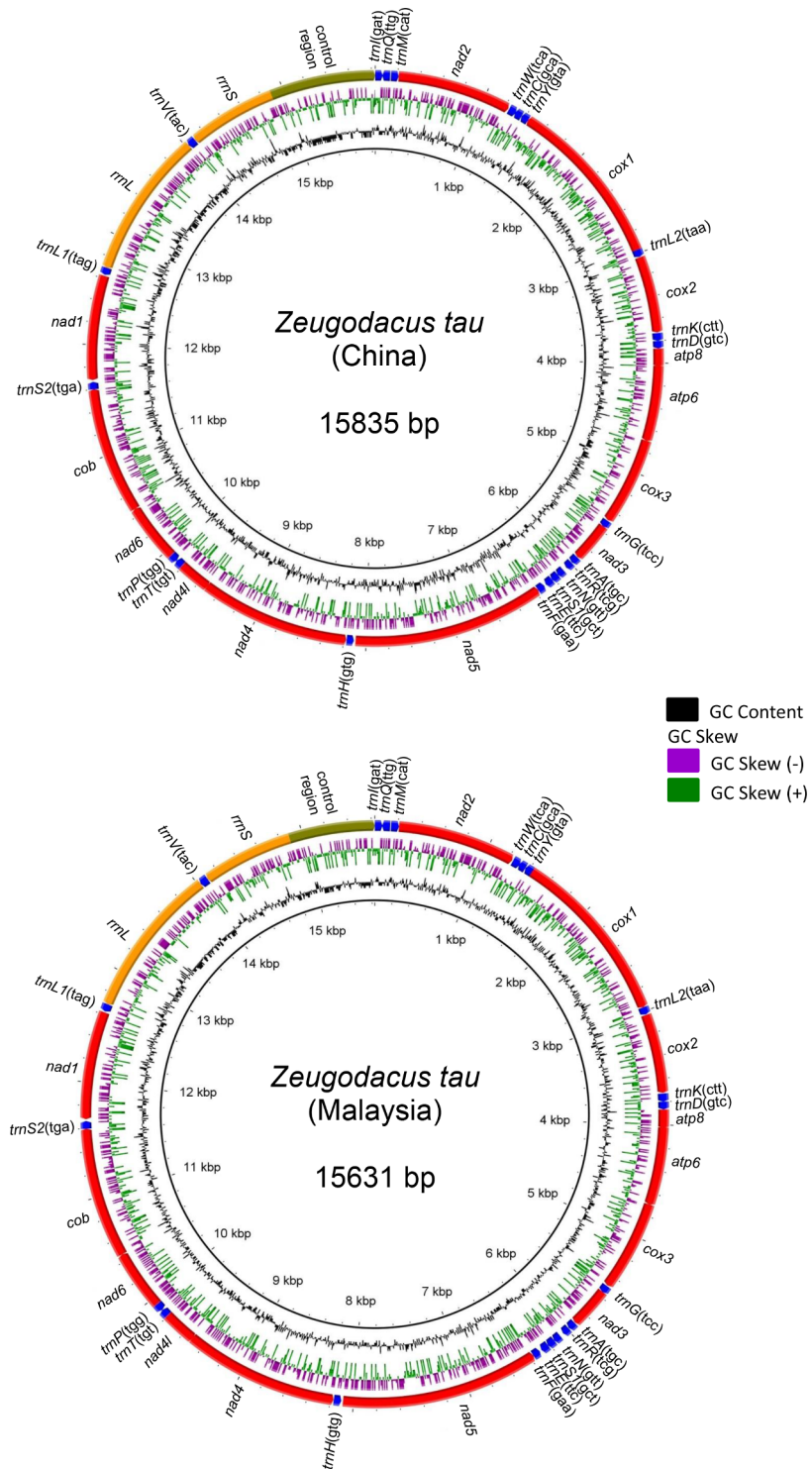


Fig 1. Complete mitogenomes of *Zeugodacus tau* ZT3 and *Z. tau* ZT1 with BRIG visualization showing the protein-coding genes, rRNA and tRNA genes. GC skew is shown on the outer surface of the ring whereas GC content is shown on the inner surface. The anticodon of each tRNAs is shown in bracket.

<https://doi.org/10.1371/journal.pone.0189325.g001>

nad4l (79.8% for *Z. tau* ZT3, and 80.1% for *Z. tau* ZT1). The A + T content of the non-coding control region was 83.5% for *Z. tau* ZT3 and 85.0% for *Z. tau* ZT1. For the two ribosomal operons, *rrnL* had a higher A + T content than *rrnS* (79.7% vs 74.7% for *Z. tau* ZT3, and 79.6% vs 74.6% for *Z. tau* ZT1). The GC skew content which included the whole genome, PCGs, rRNA genes and control region in the two taxa was negative indicating a bias toward the use of Cs over Gs. Although the AT skewness value was positive for the whole genome, rRNA genes and control region, it was variable in the individual PCGs.

As in other insects, the mitogenomes of *Z. tau* ZT3 and *Z. tau* ZT1 had three main tRNA clusters: (1) I-Q-M; (2) W-C-Y; and (3) A-R-N-S1-E-F (Fig 1). The cloverleaf structure for the respective tRNAs was similar in *Z. tau* ZT3 and *Z. tau* ZT1. The TψC-loop was absent in *trnF* while *trnS1* lacked the DHU-loop (S1 and S2 Figs).

Phylogenetic relationship and genetic divergence

Fig 2 depicts the molecular phylogeny of *Z. tau* in relation to other congeners and other taxa of the Tephritidae based on 15 mt-genes (13 PCGs + 2 rRNA genes). The phylogram based on 13 PCGs was congruent with that based on 15 mt-genes. Most of the nodes were well-supported.

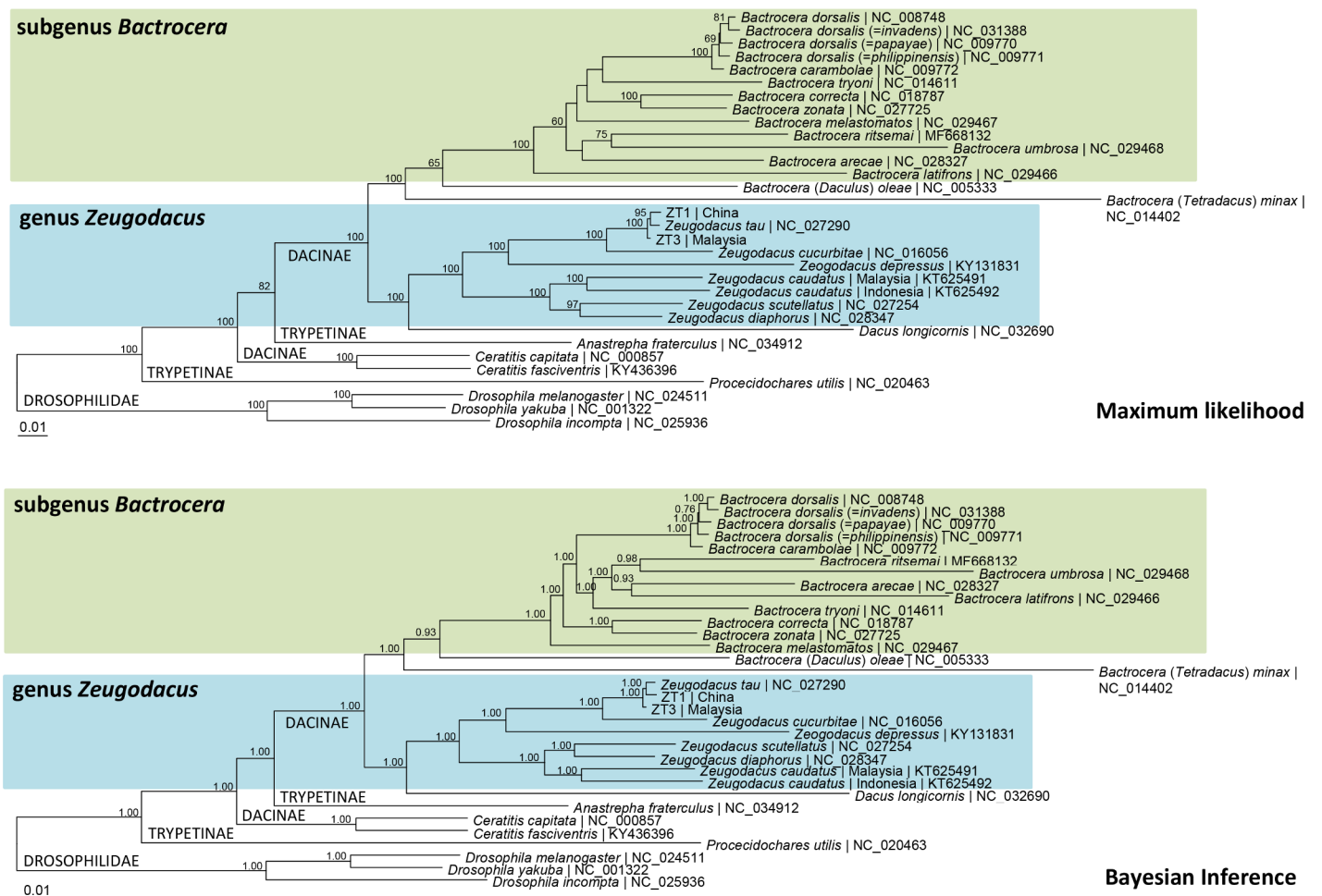


Fig 2. Maximum likelihood and Bayesian inference tree based on 15 mt-genes (13 PCGs and 2 rRNA genes) of the whole mitogenome of *Zeugodacus tau* and other Tephritid fruit flies with Drosophilidae as outgroup. Numeric values at the nodes are ML bootstrap or Bayesian posterior probabilities. The total nucleotide sequences of 15 mt-genes was 13,377 bp with AIC model = GTR+Gamma and BIC model = SYM+Gamma.

<https://doi.org/10.1371/journal.pone.0189325.g002>

Z. tau NC_027290 (China) and *Z. tau* ZT1 (China) formed a sister group in the lineage containing also *Z. tau* ZT3 (Malaysia). The genus *Zeugodacus* was monophyletic and formed a clade with *Dacus longicornis*.

The phylogenetic relationship of some of the component taxa of genus *Bactrocera* was not congruent between ML and BI analyses (Fig 2). For example, ML analysis indicated *B. melastomatos* to be a member of the *B. dorsalis* complex, but in BI analysis it was basal to the other taxa of subgenus *Bactrocera*. Nonetheless, the genus *Bactrocera* was monophyletic.

Phylogenetic analysis based on partial *cox1* sequences from bp 50–700 indicated that the *Z. tau* taxa from China, Bangladesh, India (Meghalaya, north of Bangladesh) and Malaysia formed a clade with several haplotypes (Fig 3). The uncorrected genetic distance ranged from 'p' = 0 to 'p' = 0.72% (S5 Table).

Based on the partial *cox1* sequence from bp 900–1500, the *Z. tau* taxa from India, Sri Lanka, Malaysia, Laos, China and Japan formed a clade with *Z. tau* sp. A from Thailand (Fig 4), with uncorrected genetic distance ranging from 'p' = 0% to 'p' = 1.39% (S6 Table). This clade was distinctly different from *Z. tau* sp. B, C, D, E, F, G, and I from Thailand, with uncorrected genetic distance ranging from 'p' = 9.03% to 'p' = 14.06% (S6 Table).

Haplotype diversity and nucleotide diversity

Twelve haplotypes were revealed in the present 18 *cox1* sequences (from bp 50–700) of *Z. tau* from four geographical regions (China, Malaysia, Bangladesh and India) (Fig 5). A common haplotype was found in China (3 sequences), Bangladesh (2 sequences) and India (1 sequence). The haplotype/gene diversity was 0.8954 ± 0.0653 , and the nucleotide diversity was 0.0033 ± 0.0022 .

Sixteen haplotypes were revealed in the 22 *cox1* sequences (from bp 900–1500) of *Z. tau* sensu stricto from six geographical regions (China, Laos, Malaysia, India, Sri Lanka, and Thailand sp. A) (Fig 6). A common haplotype was found in China (2 sequences), Japan (1 sequence), India (1 sequence) and Sri Lanka (1 sequence). Another haplotype was common to Malaysia (2 sequences) and India (1 sequence). The haplotype/gene diversity was 0.9437 ± 0.0372 , and the nucleotide diversity was 0.0056 ± 0.0034 . *Z. tau* sp. B, C, D, E, F, G and I from Thailand formed a distinct cluster from *Z. tau* sensu stricto, and each was represented by a distinct haplotype. The haplotypes of *Z. tau* F and *Z. tau* B had a small difference of 4 bp.



Fig 3. Bayesian inference and maximum likelihood tree based on partial sequence from bp 50–700 of mitochondrial *cox1* gene of *Zeugodacus tau* with *Bactrocera dorsalis* and *B. carambolae* as outgroup. Numeric values at the nodes are Bayesian posterior probabilities/ML bootstrap.

<https://doi.org/10.1371/journal.pone.0189325.g003>

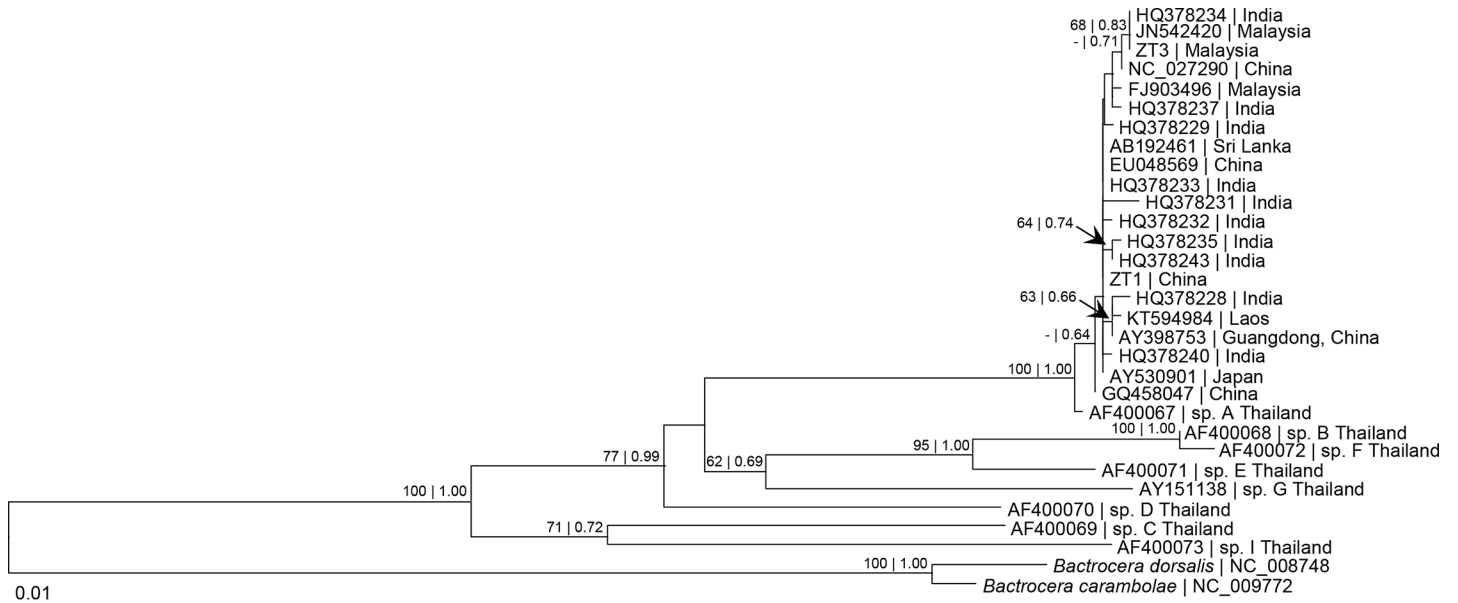
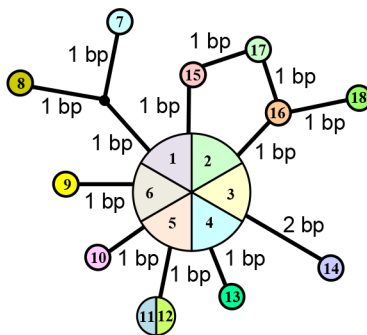


Fig 4. Bayesian inference and maximum likelihood tree based on partial sequence from bp 900–1500 of mitochondrial *cox1* gene of *Zeugodacus tau* with *Bactrocera dorsalis* and *B. carambolae* as outgroup. Numeric values at the nodes are Bayesian posterior probabilities/ML bootstrap.

<https://doi.org/10.1371/journal.pone.0189325.g004>

Discussion

The genus *Zeugodacus* is represented by 52 named and some 19 unnamed species [31]. To date, the complete mitogenome has been reported for seven taxa—*Z. caudatus* Malaysia, *Z. caudatus* Indonesia, *Z. cucurbitae*, *Z. depressus*, *Z. diaphorus*, *Z. scutellatus* and *Z. tau* (China).



Zeugodacus tau

- | | |
|------------------------------|--------------------------------|
| 1. GQ154158 Bangladesh | 10. KF660134 China |
| 2. KT175576 India | 11. GQ154157 Bangladesh |
| 3. JX266424 Fujian, China | 12. KM023421 Bangladesh |
| 4. KF660196 China | 13. GQ154160 China |
| 5. JX266425 Yunnan, China | 14. ZT1 China |
| 6. GQ154159 Bangladesh | 15. JX266422 Zhejiang, China |
| 7. JX266423 Guizhou, China | 16. KJ833950 China |
| 8. KF660183 China | 17. ZT3 Malaysia |
| 9. NC_027290 China | 18. GQ154161 Malaysia |

Fig 5. Haplotype network of *Zeugodacus tau* based on cytochrome c oxidase subunit I (*cox1*) sequences (from bp 50–700) generated by NETWORK software. Circles represent haplotypes and numbers within the circle represent individuals sharing the specific haplotype.

<https://doi.org/10.1371/journal.pone.0189325.g005>

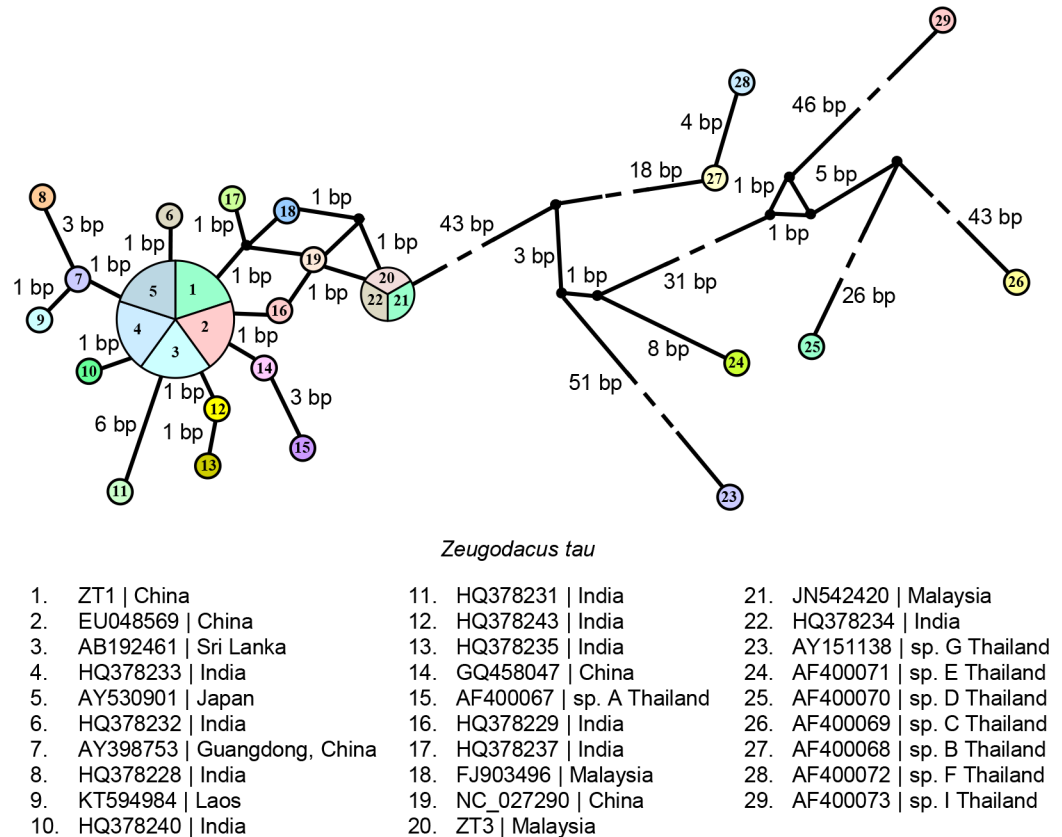


Fig 6. Haplotype network of *Zeugodacus tau* based on cytochrome *c* oxidase subunit I (*cox1*) sequences (from bp 900–1500) generated by NETWORK software. Circles represent haplotypes and numbers within the circle represent individuals sharing the specific haplotype.

<https://doi.org/10.1371/journal.pone.0189325.g006>

Molecular studies indicate that *Z. caudatus* Malaysia and *Z. caudatus* Indonesia are sibling species [24,32], and *Z. tau* in Thailand consists of eight species [8–10].

The gene order of *Z. tau* mitogenome conforms to other *Zeugodacus* and other tephritid mitogenomes [15,20,24,25,28–30]. The mitogenome of *Z. tau* ZT3 (Malaysia) is shorter than that of *Z. tau* ZT1 (China) and *Z. tau* NC_027290 (China), while *Z. tau* ZT1 is longer than *Z. tau* NC_027290 (Table 2). The difference in the total size of the mitogenome is due mainly to the length of the control region— 745 bp for *Z. tau* ZT3, 946 bp for *Z. tau* ZT1 and 801 bp for *Z. tau* NC_027290 (Table 2).

There are differences in the spacing/overlap sequence in some intergenic regions among *Z. tau* mitogenomes: -1 bp in *Z. tau* NC_027290 between *trnC* and *trnY* versus 1 bp in *Z. tau* ZT3 and *Z. tau* ZT1 (Table 2).

The difference in size of the *nad1* gene (1020 bp in *Z. tau* NC_027290, and 940 bp in *Z. tau* ZT1 and *Z. tau* ZT3) and the stop codon (TAA in NC_027290 and incomplete T in ZT1 and ZT3) can be attributed to annotation of the intergenic space between *trnS2* and *nad1* genes (overlap of 65 bp in NC_027290 and spacing sequence of 15 bp in ZT3 and ZT1); this intergenic space is 15 bp in most of the *Zeugodacus* taxa. Incomplete stop codons have been reported in other taxa of tephritid fruit flies [15,20,24]. The incomplete stop codons can be converted to TAA by post-translational polyadenylation [33].

A long poly-A stretch of 20 bp is present in the control region after 'ATAGA' motif in the Malaysian and China taxa of *Z. tau*. In addition, a long poly-T stretch is present in the control

region of *Z. tau* ZT3 (14 bp) and *Z. tau* ZT1 (19 bp); this poly-T stretch is not present in *Z. tau* NC_027290.

In both *Z. tau* ZT3 and *Z. tau* ZT1, the T Ψ C-loop was absent in *trnF* while *trnS1* lacked the DHU-loop (S1 and S2 Figs). The T Ψ C-loop and DHU-loop of tRNA act as special recognition site during protein biosynthesis or translation [34–36]. It has been reported that misacylation of tRNA can affect the survivability of an organism [36]. However, deviant tRNA secondary structures are frequent in Arthropoda [37].

The mitochondrial *cox1* gene has been commonly used for differentiation of various taxa of *Z. tau* [9,12,38–46]. In the present study based on partial sequences of *cox1* gene (Figs 3–6), the *Z. tau* taxa showed several haplotypes. The taxa from China, Japan, Laos, Malaysia, Bangladesh, India, and Sri Lanka were genetically similar to *Z. tau* sp. A from Thailand, with 'p' = 0–1.39% (S5 and S6 Tables). As Fuzhou (Foochow), Fujian, China is the type locality of *Z. tau*, the taxa from various geographical regions that grouped with those from China can be designated as *Z. tau* sensu stricto. Although many taxa had been included for comparison, none were similar to any of the *Z. tau* sp. B, C, D, E, F, G, and I reported from Thailand. Among the *Z. tau* taxa from Thailand, the genetic distance between *Z. tau* F and *Z. tau* B was 'p' = 0.69% (S6 Table) with haplotype difference of 4 bp (Fig 6), indicating that these two taxa may be conspecific.

In the present study based on partial *cox1* sequences, *Z. tau* ZT1 (China) and *Z. tau* NC_027290 (China) were not closely related to each other compared to *Z. tau* ZT3 (Malaysia) (Figs 3 and 4). This differs from their closer relationship based on complete *cox1* gene (S3 Fig) and 15 mt-genes (Fig 2). A complete *cox1* gene (or 13 PCGs or 15 mt-genes) instead of partial sequence is therefore more appropriate for determining phylogenetic relationship.

At the higher-level phylogeny, the phylogenetic analysis indicated that *Anastrepha fraterculus* (Tribe Toxotrypanini) of subfamily Trypetinae was grouped with Tribe Dacini of Dacinae, while *Procecidochares utilis* (Tribe Ceceidocharini) of subfamily Trypetinae was basal to the clade containing Dacinae and *A. fraterculus* (Fig 2). This discrepancy may be due to insufficient taxon sampling. A broader taxa sampling, particularly Trypetinae, is needed to better elucidate the higher-level phylogeny of the tribes and subfamilies of Tephritidae.

In summary, we have successfully sequenced the complete mitogenome of *Z. tau* from Malaysia and China and confirmed that they were conspecific. Based on partial *cox1* sequences, the taxa from China, Japan, Laos, Malaysia, Bangladesh, India, Sri Lanka, and *Z. tau* sp. A from Thailand are conspecific and belong to *Z. tau* sensu stricto. The mitogenome will prove useful for studies on phylogenetics and systematics of fruit flies of the *Z. tau* species complex and other taxa of Tephritidae.

Supporting information

S1 Fig. Cloverleaf structure of the 22 inferred tRNAs in the mitogenome of *Zeugodacus tau* ZT3 (Malaysia). The cloverleaf structure for *trnF* lacked the T Ψ C-loop, and *trnS1* lacked the DHU-loop.
(TIF)

S2 Fig. Cloverleaf structure of the 22 inferred tRNAs in the mitogenome of *Zeugodacus tau* ZT1 (China). The cloverleaf structure for *trnF* lacked the T Ψ C-loop, and *trnS1* lacked the DHU-loop.
(TIF)

S3 Fig. Bayesian inference and maximum likelihood tree based on complete sequence of mitochondrial *cox1* gene of *Zeugodacus tau* with *Bactrocera dorsalis* and *B. carambolae* as

outgroup. Numeric values at the nodes are Bayesian posterior probabilities/ML bootstrap. (TIF)

S1 Table. Characteristics of the mitochondrial genome of *Zeugodacus tau* ZT3 (Malaysia). The anticodon of each tRNAs is shown in bracket. J (+) or N (-) indicates gene directions. (DOCX)

S2 Table. Characteristics of the mitochondrial genome of *Zeugodacus tau* ZT1 (China). The anticodon of each tRNAs is shown in bracket. J (+) or N (-) indicates gene directions. (DOCX)

S3 Table. Nucleotide composition of whole mitogenome, protein-coding genes, rRNA genes and control region of *Zeugodacus tau* ZT3 (Malaysia). (DOCX)

S4 Table. Nucleotide composition of whole mitogenome, protein-coding genes, rRNA genes and control region of *Zeugodacus tau* ZT1 (China). (DOCX)

S5 Table. Uncorrected genetic distance (%) between pairs of *Zeugodacus tau* taxa with *Bactrocera dorsalis* and *B. carambolae* as outgroup taxa based on partial sequence from bp 50–700 of mitochondrial *cox1* gene. (DOCX)

S6 Table. Uncorrected genetic distance (%) between pairs of *Zeugodacus tau* taxa with *Bactrocera dorsalis* and *B. carambolae* as outgroup taxa based on partial sequence from bp 900–1500 of mitochondrial *cox1* gene. (DOCX)

Acknowledgments

We would like to thank our institutions for providing various support and facilities.

Author Contributions

Conceptualization: Hoi-Sen Yong, Sze-Looi Song.

Data curation: Hoi-Sen Yong, Sze-Looi Song.

Formal analysis: Hoi-Sen Yong, Sze-Looi Song.

Funding acquisition: Hoi-Sen Yong.

Investigation: Hoi-Sen Yong, Sze-Looi Song.

Methodology: Hoi-Sen Yong, Sze-Looi Song.

Project administration: Hoi-Sen Yong, Sze-Looi Song.

Resources: Hoi-Sen Yong, Sze-Looi Song, Phaik-Eem Lim, Praphathip Eamsobhana.

Software: Hoi-Sen Yong, Sze-Looi Song.

Supervision: Hoi-Sen Yong.

Validation: Hoi-Sen Yong, Sze-Looi Song, Phaik-Eem Lim, Praphathip Eamsobhana.

Visualization: Hoi-Sen Yong, Sze-Looi Song.

Writing – original draft: Hoi-Sen Yong, Sze-Looi Song.

Writing – review & editing: Hoi-Sen Yong, Sze-Looi Song, Phaik-Eem Lim, Praphathip Eamsobhana.

References

1. Hardy D (1973) The fruit flies (Tephritidae–Diptera) of Thailand and bordering countries. Pacific Insects Monograph 31: 1–353.
2. Vargas RI, Piñero JC, Leblanc L (2015) An overview of pest species of *Bactrocera* fruit flies (Diptera: Tephritidae) and the integration of biopesticides with other biological approaches for their management with a focus on the Pacific Region. *Insects* 6: 297–318. <https://doi.org/10.3390/insects6020297> PMID: 26463186
3. White IM, Elson-Harris MM (1992) *Fruit flies of economic significance: their identification and bionomics*. CAB International, Wallingford, UK.
4. Carroll LE, Norrbom AL, Dallwitz MJ, Thompson FC (2004 onwards). Pest fruit flies of the world—larvae. Version: 8th December 2006. <http://deltaintkey.com> accessed 31 August 2017.
5. Kitthawee S, Dujardin JP (2010) The geometric approach to explore the *Bactrocera tau* complex (Diptera: Tephritidae) in Thailand. *Zoology* 113: 243–249. <https://doi.org/10.1016/j.zool.2009.12.002> PMID: 20817492
6. Khan M, Tahira-Binte-Rashid, Howlader AJ (2011) Comparative host susceptibility, oviposition, and colour preference of two polyphagous tephritids: *Bactrocera cucurbitae* (Coq.) and *Bactrocera tau* (Walker). *Research Journal of Agriculture and Biological Sciences* 7(3): 343–349.
7. Yong H-S (2014) *Trichosnathes wallichiana* (Cucurbitaceae): a new host fruit of *Bactrocera tau* (Insecta, Tephritidae). *Journal of Science and Technology in the Tropics* 10(2): 95–98.
8. Baimai V, Phinchongsakuldit J, Sumrandee C, Tigvattanont S (2000) Cytological evidence for a complex of species within the taxon *Bactrocera tau* (Diptera: Tephritidae) in Thailand. *Biological Journal of Linnean Society* 69: 399–409.
9. Jamnongluk W, Baimai V, Kittayapon P (2003) Molecular phylogeny of tephritid fruit flies in the *Bactrocera tau* complex using the mitochondrial *COI* sequences. *Genome* 46(1): 112–118. <https://doi.org/10.1139/g02-113> PMID: 12669803
10. Saelee A, Tigvattanont S, Baimai V (2006) Allozyme electrophoretic evidence for a complex of species within the *Bactrocera tau* group (Diptera: Tephritidae) in Thailand. *Songklanakarin Journal of Science and Technology* 28(2): 249–259.
11. Thanaphum S, Thaenkhram U (2003) Relationships of Forms within the *Bactrocera tau* (Walker) (Diptera: Tephritidae) Taxon Based on *Heat Shock Protein 70* Cognate Sequences. *Annals of the Entomological Society of America* 96(1): 44–53.
12. Prabhakar CS, Sood P, Mehta PK, Sharma PN (2013) Population genetic structure of the pumpkin fruit fly, *Bactrocera tau* (Walker) (Diptera: Tephritidae) in Himachal Pradesh, India. *Biochemical Systematics and Ecology* 51: 291–296.
13. Shi W, Kerdelhué C, Ye H (2014) Genetic structure and colonization history of the fruit fly *Bactrocera tau* (Diptera: Tephritidae) in China and Southeast Asia. *Journal of Economic Entomology* 107(3): 1256–1265.
14. Tan M, Zhang R, Xiang C, Zhou X (2016) The complete mitochondrial genome of the pumpkin fruit fly, *Bactrocera tau* (Diptera: Tephritidae). *Mitochondrial DNA Part A DNA Mapping, Sequencing, and Analysis* 27: 2502–2503.
15. Yong H-S, Song S-L, Lim P-E, Chan K-G, Chow W-L, Eamsobhana P (2015) Complete mitochondrial genome of *Bactrocera arecae* (Insecta: Tephritidae) by next-generation sequencing and molecular phylogeny of Dacini tribe. *Scientific Reports* 5: 15155. <https://doi.org/10.1038/srep15155> PMID: 26472633
16. Yong H-S, Song S-L, Lim P-E, Eamsobhana P, Suana IW (2016) Differentiating sibling species of *Zeugodacus caudatus* (Insecta: Tephritidae) by complete mitochondrial genome. *Genetica* 144: 513–521. <https://doi.org/10.1007/s10709-016-9919-9> PMID: 27502829
17. De Re FC, Wallau GL, Robe LJ, Loreto EL (2014) Characterization of the complete mitochondrial genome of flower-breeding *Drosophila incompta* (Diptera, Drosophilidae). *Genetica* 142: 525–535. <https://doi.org/10.1007/s10709-014-9799-9> PMID: 25416157
18. Clary DO, Wolstenholme DR (1985) The ribosomal RNA genes of *Drosophila* mitochondrial DNA. *Nucleic Acids Research* 13: 4029–4045. PMID: 2989784
19. Zhang L-J, Jiang L-H, Wei C-Y, Liu R-S, Liu X-L, Li J-G, Xue H-J (2016) The status of *Bactrocera invadens* Drew, Tsuruta & White (Diptera: Tephritidae) inferred from complete mitochondrial genome analysis. *Mitochondrial DNA B Resources* 1: 680–681.

20. Yong H-S, Song S-L, Lim P-E, Eamsobhana P, Suana IW (2016) Complete mitochondrial genome of three *Bactrocera* fruit flies of subgenus *Bactrocera* (Diptera: Tephritidae) and their phylogenetic implications. *PLoS ONE* 11: e0148201. <https://doi.org/10.1371/journal.pone.0148201> PMID: 26840430
21. Nardi F, Carapelli A, Boore JL, Roderick GK, Dallai R, Frati F (2010) Domestication of olive fly through a multi-regional host shift to cultivated olives: Comparative dating using complete mitochondrial genomes. *Molecular Phylogenetics and Evolution* 57: 678–686. <https://doi.org/10.1016/j.ympev.2010.08.008> PMID: 20723608
22. Choudhary JS, Naaz N, Prabhakar CS, Rao MS, Das B (2015) The mitochondrial genome of the peach fruit fly, *Bactrocera zonata* (Saunders) (Diptera: Tephritidae): Complete DNA sequence, genome organization, and phylogenetic analysis with other tephritids using next generation DNA sequencing. *Gene* 569: 191–202. <https://doi.org/10.1016/j.gene.2015.05.066> PMID: 26031235
23. Nardi F, Carapelli A, Dallai R, Frati F (2003) The mitochondrial genome of the olive fly *Bactrocera oleae*: two haplotypes from distant geographical locations. *Insect Molecular Biology* 12: 605–611. PMID: 14986921
24. Yong H-S, Song S-L, Lim P-E, Eamsobhana P, Suana IW (2016) Differentiating sibling species of *Zeugodacus caudatus* (Insecta: Tephritidae) by complete mitochondrial genome. *Genetica* 144: 513–521. <https://doi.org/10.1007/s10709-016-9919-9> PMID: 27502829
25. Jeong SY, Kim MJ, Kim JS, Kim I (2017) Complete mitochondrial genome of the pumpkin fruit fly, *Bactrocera depressa* (Diptera: Tephritidae). *Mitochondrial DNA Part B Resources* 2: 85–87.
26. Zhang KJ, Liu L, Rong X, Zhang GH, Liu H, Liu YH (2016) The complete mitochondrial genome of *Bactrocera diaphora* (Diptera: Tephritidae). *Mitochondrial DNA Part A DNA Mapping, Sequencing, and Analysis* 27: 4314–4315.
27. Spanos L, Koutroumbas G, Kotsyfakis M, Louis C (2000) The mitochondrial genome of the mediterranean fruit fly, *Ceratitis capitata*. *Insect Molecular Biology* 9: 139–144. PMID: 10762421
28. Drosopoulou E, Pantelidou C, Gariou-Papalexiou A, Augustinos AA, Chartomatsidou T, Kyritsis GA, et al. (2017) The chromosomes and the mitogenome of *Ceratitis fasciventris* (Diptera: Tephritidae): two genetic approaches towards the *Ceratitis* FAR species complex resolution. *Scientific Reports* 7: 4877. <https://doi.org/10.1038/s41598-017-05132-3> PMID: 28687799
29. Jiang F, Pan X, Li X, Yu Y, Zhang J, Jiang H, et al. (2016) The first complete mitochondrial genome of *Dacus longicornis* (Diptera: Tephritidae) using next-generation sequencing and mitochondrial genome phylogeny of Dacini tribe. *Scientific Reports* 6: 36426. <https://doi.org/10.1038/srep36426> PMID: 27812024
30. Isaza JP, Alzate JF, Canal NA (2017) Complete mitochondrial genome of the Andean morphotype of *Anastrepha fraterculus* (Wiedemann) (Diptera: Tephritidae). *Mitochondrial DNA Part B Resources* 2: 210–211.
31. De Meyer M, Delatte H, Mwatawala M, Quilici S, Vayssières J-F, Virgilio M (2015) A review of the current knowledge on *Zeugodacus cucurbitae* (Coquillett) (Diptera, Tephritidae) in Africa, with a list of species included in *Zeugodacus*. In: De Meyer M, Clarke AR, Vera MT, Hendrichs J (Eds) *Resolution of Cryptic Species Complexes of Tephritid Pests to Enhance SIT Application and Facilitate International Trade*. *ZooKeys* 540: 539–557.
32. Yong H-S, Lim P-E, Tan J, Song S-L, Suana IW, Eamsobhana P (2015) Multigene phylogeography of *Bactrocera caudata* (Insecta: Tephritidae): distinct genetic lineages in northern and southern hemispheres. *PLoS ONE* 10(6): e0129455. <https://doi.org/10.1371/journal.pone.0129455> PMID: 26090853
33. Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290: 470–474. PMID: 7219536
34. Agirrezabala X, Frank J (2009) Elongation in translation as a dynamic interaction among the ribosome, tRNA, and elongation factors EF-G and EF-Tu. *Quarterly Reviews of Biophysics* 42: 159–200. <https://doi.org/10.1017/S0033583509990060> PMID: 20025795
35. Chan CW, Chetnani B, Mondragón A (2013) Structure and function of the T-loop structural motif in non-coding RNAs. *Wiley Interdisciplinary Reviews RNA* 4: 507–522. <https://doi.org/10.1002/wrna.1175> PMID: 23754657
36. Hendrickson TL (2001) Recognizing the D-loop of transfer RNAs. *Proceedings of the National Academy of Sciences* 98: 13473–13475.
37. Jühling F, Pütz J, Bernt M, Donath A, Middendorf M, Florentz C, Stadler PF (2012) Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. *Nucleic Acids Research* 40: 2833–2845. <https://doi.org/10.1093/nar/gkr1131> PMID: 22139921
38. Chua TH, Chong YV, Lim SH (2010) Species determination of Malaysian *Bactrocera* pests using PCR-RFLP analyses (Diptera: Tephritidae). *Pest Management Science* 66(4): 379–384. <https://doi.org/10.1002/ps.1886> PMID: 19946858

39. Hu J, Zhang JL, Nardi F, Run J, Zhang RJ (2008) Population genetic structure of the melon fly, *Bactrocera cucurbitae* (Diptera: Tephritidae), from China and Southeast Asia. *Genetica* 134: 319–324. <https://doi.org/10.1007/s10709-007-9239-1> PMID: 18188668
40. Jiang F, Li ZH, Deng YL, Wu JJ, Liu RS, Buahom N (2013) Rapid diagnosis of the economically important fruit fly, *Bactrocera correcta* (Diptera: Tephritidae) based on a species-specific barcoding cytochrome oxidase I marker. *Bulletin of Entomological Research* 103(3): 363–371. <https://doi.org/10.1017/S0007485312000806> PMID: 23458744
41. Jiang F, Jin Q, Liang L, Zhang AB, Li ZH (2014) Existence of species complex largely reduced barcoding success for invasive species of Tephritidae: a case study in *Bactrocera* spp. *Molecular Ecology Resources* 14(6): 1114–1128. <https://doi.org/10.1111/1755-0998.12259> PMID: 24690386
42. Leblanc L, San Jose M, Barr N, Rubinoff D (2015) A phylogenetic assessment of the polyphyletic nature and intraspecific color polymorphism in the *Bactrocera dorsalis* complex (Diptera, Tephritidae). *Zookeys* 540: 339–367.
43. Lim PE, Tan J, Suana IW, Eamsobhana P, Yong HS (2012) Distinct genetic lineages of *Bactrocera caudata* (Insecta: Tephritidae) revealed by COI and 16S DNA sequences. *PLoS ONE* 7(5): e37276. <https://doi.org/10.1371/journal.pone.0037276> PMID: 22615962
44. Virgilio M, Backeljau T, Nevado B, De Meyer M (2010) Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics* 11: 206. <https://doi.org/10.1186/1471-2105-11-206> PMID: 20420717
45. Virgilio M, Jordaens K, Verwimp C, White IM, De Meyer M. (2015) Higher phylogeny of frugivorous flies (Diptera, Tephritidae, Dacini): Localised partition conflicts and a novel generic Classification. *Molecular Phylogenetics and Evolution* 85: 171–179. <https://doi.org/10.1016/j.ympev.2015.01.007> PMID: 25681676
46. Zhang B, Liu YH, Wu WX, Wang ZL (2010) Molecular phylogeny of *Bactrocera* species (Diptera: Tephritidae: Dacini) inferred from mitochondrial sequences of 16S rDNA and COI sequences. *Florida Entomologist* 93(3): 369–377.