


ACKNOWLEDGEMENTS. We thank Agricultural Experimental Farm, University of Calcutta, Banaripara for providing support in carrying this study during field as well as laboratory analysis. We acknowledge Department of Agricultural Chemistry and Soil Science, University of Calcutta for instrumental facilities as well as financial assistance.

Received 21 March 2019; revised accepted 12 September 2019

doi: 10.18520/cs/v117/i8/1364-1368

**Genetic lineage of *Zeugodacus caudatus* (Diptera: Tephritidae) detected with *mtCOI* gene analysis from India**

Chandra S. Prabhakar1,2,3,#, Anil1,7, Jaipal S. Choudhary2, Ravi S. Singh3, S. N. Ray4, Kalmesh Managanvi6, Mona Kumari5, Ashok B. Hadapad3 and Ramesh S. Hire3

1Department of Entomology, Veer Kunwar Singh College of Agriculture, Dumraon (Bihar Agricultural University, Sabour), India
2Department of Entomology, Bihar Agricultural College, Bihar Agricultural University, Sabour, Bhagalpur 813 210, India
3Department of Plant Breeding and Genetics, Bihar Agricultural College, Bihar Agricultural University, Sabour, Bhagalpur 813 210, India
4Department of Entomology, Dr Kalam Agriculture College, Kishanganj 855 107 (Bihar Agricultural University, Sabour), India
5Department of Genetics and Plant Breeding (Cotton Section), Chaudhary Charan Singh Haryana Agricultural University, Hisar 125 004, India
6Nuclear Agriculture and Biotechnology Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400 094, India

**Zeugodacus caudatus** (Fabricius) is a pest of eucarp plant species. The present study was conducted to draw the relationship among Indian *Z. caudatus* populations with the other defined genetic lineage of the species. A total of 18 individuals’ *mtCOI* gene sequences from 3 populations of India were analysed along with 34 individuals’ *mtCOI* gene sequences from Malaysia, Indonesia, Thailand and China and generated 14 haplotypes. Phylogenetic study revealed the presence of distinct genetic lineage in *Z. caudatus* populations collected from India. The genetic distance between three distinct lineages of *Z. caudatus* was 0.057, 0.055 and 0.018 between Indonesia and Malaysia, India and

*For correspondence. (e-mail: csprabhakar.ento@gmail.com)*

1368
Indonesia, and India and Malaysia, respectively, and also evident from phylogenetic analysis. Further, the mitochondrial cytochrome oxidase I (COI) gene sequences developed in this study will help detection and geographical distribution of new haplotypes and lineages of the species in future.

Keywords: Dacinae, fruit fly, haplotypes, population genetics.

TEPHRITID fruit flies are the most serious insect pests of horticultural crops causing enormous economic losses every year throughout the tropical and subtropical regions of the world1–3.

The family Tephritidae comprises over 4448 species distributed in more than 481 genera, of which 800 species belong to the subfamily Dacinae4–6. Tropical Asia, Australia and South Pacific regions are the native places of Bactrocera and Zeugodacus species of subfamily Dacinae; whereas Africa and warm temperate areas of Europe and Asia are the home of few other species of these genera. Fruit fly species of the genus Bactrocera and Zeugodacus are pests of polyphagous plant having extensive distribution and wide range of climatic adaptability with high flying capacity and population potential and cause high economic losses in horticultural crops7,8. Members of subgenus Zeugodacus of the genus Bactrocera are mostly pest of cucurbit plant species. Recently, Zeugodacus was elevated to the level of genus9. There are many pest species in the genus Zeugodacus: Zeugodacus cucurbitae (Coquillett), Zeugodacus tau (Walker), Zeugodacus scutellaris (Bezzi) and Zeugodacus caudatus (Fabricius)3,10–12. Zeugodacus caudatus (Fabricius) is presently renamed from Bactrocera maculipennis Doleschall, Bactrocera caudata (Fabricius) and Chaetodacus caudatus Fabricius3,10,11. Z. caudatus is identified with predominantly black scutum, yellow medial and lateral stripes on scutum; black line across mouth opening on the face and a black scutum, yellow medial and lateral stripes on scutum. Further, Yong et al.12,15–18 reported that Z. caudatus samples from Malaysia and Thailand were different from the northern hemisphere (Z. caudatus samples from Indonesia) with multigenic phylogenetic analysis.

Precise resolution on species identification and characterization of different complex species/forms is often missing through morphological differences17. It can be enhanced by integration of molecular biology, i.e. DNA barcode using mtCOI gene sequences17. Phylogeographic structures and intra- and inter-specific relationships of different insect species including species of genus Bactrocera and Zeugodacus have been widely determined through robust evolutionary mitochondrial DNA based markers (COI gene sequences)16–22. The present study was conducted to examine the genetic variability and lineage of Z. caudatus species present in India using mtCOI gene sequencing and to determine their association in the context of divergent genetic lineages (cryptic species/sibling species) present in Z. caudatus populations.

Adult males of Z. caudatus were collected from three distant locations in India, viz. Mumbai (Maharashtra), Ranchi (Jharkhand) and Bhagalpur (Bihar) during 2012–2015. The individual of Z. caudatus adult fly was preserved in a separate vial with 800 μl of 95% ethyl alcohol at –25°C until genomic DNA isolation. Fruit fly specimens were identified as Z. caudatus based on available literature1,23 and also the identity of specimens as Z. caudatus was confirmed by the fruit fly taxonomist M. L. Agarwal (Department of Entomology, Dr Rajendra Prasad, Central Agricultural University, Pusa, Bihar).

DNA (genomic) of individual Z. caudatus was isolated from the preserved specimens using the cetyl trimethyl ammonium bromide (CTAB) method24. Each fly was crushed into fine powder individually in a sterile mortar and pestle. The material was then transferred to a microcentrifuge tube containing 300 μl of pre-heated (60°C) CTAB buffer made up of 1 M tris HCl pH 8.0, 5% CTAB, 5 M NaCl, 0.5 M EDTA pH 8.0 and 4% β-mercaptopethanol. The preparation was then incubated at 60°C for 30 min with gentle mixing at regular intervals. After incubation, 300 μl of premixed solution of chloroform and isoamyl alcohol (24:1) were added, mixed properly by inverting the tube several times and then the tube were centrifuged at 8000 rpm for 10 min to remove the aqueous phase. The aqueous phase of the content in tube was transferred to a new microcentrifuge tube and 150 μl pre-chilled isopropanol was added. Microcentrifuge tubes were kept for 20–30 min at –20°C for the precipitation of genomic DNA. Then, the tubes were centrifuged at 12,000 rpm for 12 min and the supernatant was decanted. The DNA pellet present in tubes was washed with 70% pre-chilled ethyl alcohol, kept for drying and then dissolved in 30 μl of Tris EDTA buffer (10 mM Tris HCl pH 8.0 and 1 mM EDTA) and stored at –20°C.

Primer pair UEA7 5′ TACAGTGGAAATAGACGTT-GATAC 3′ (forward) and UEA10 5′ TCCAATGCACTA-TCTGCGATTTAA 3′ (reverse) for the amplification of mitochondrial COI gene developed by Lunt et al.25 were
used to perform the polymerase chain reaction (PCR). The DNA amplification was carried out in 0.2 ml micro tubes with 20 μl reaction volume containing 2 μl 10× reaction buffer with 25 mM MgCl₂, 0.5 μl 10 mM dNTPs, 1 μl 20 pmol of each primer, 0.2 μl 5U/μl Taq polymerase (all manufactured by HI-MEDIA India, Mumbai) and 2 μl 10 ng of DNA template. Amplification was carried out in Flexigene 9700 thermal cycler (QIAGEN India) with an initial DNA denaturation for 3 min at 94°C, followed by 35 cycles at 94°C for 1 min, 50°C for 1 min, 72°C for 1 min and a final step of elongation at 72°C for 30 min. The amplification of targeted DNA fragment was confirmed with separation of PCR product in 2% (w/v) agarose gel using DNA electrophoresis with TAE buffer (1 mM EDTA, 40 mM Tris-acetate). Freeze-dried PCR products of mtCOI gene were custom sequenced (ABI PRISM 310™ Genetic Analyser, Applied Biosystems, USA) using the same primers pair (Xceliris Labs Limited, India and Scigenome, Kochi, India). Unique mtCOI gene sequences of Z. caudatus were deposited in GenBank with accession no. KT989670-KT989677 (Ranchi, India), KU041685-KU041688, MF038800–MF038805 (Mumbai, India). Descriptive statistics, viz. haplotype diversity (Hd), number of haplotypes (H), average number of nucleotide difference (K) and nucleotide diversity (π) were calculated using Dnasp version 5.0 software.

A median-joining haplotype network was constructed to depict the evolutionary and geographical relationships among haplotypes using NETWORK version 4.6 software. Z. caudatus mitochondrial haplotypes were colour coded with the country of origin of specimens to know the geographical genetic relationship among the populations.

Pair-wise distance measurements between individual sequences of Z. caudatus were performed using the p genetic distance model implemented in MEGA 6.0 software. p genetic distance model was chosen for these analyses with GenBank sequences which specifically addressed the level of divergence at the mtCOI gene of Z. caudatus. The phylogenetic relationships were established between different lineages of Z. caudatus with minimum evolution, maximum likelihood and neighbour-joining methods using Kimura-2 parameters as estimate of genetic divergence with Zeugodacus cucurbitae (HQ378218) and Zeugodacus tau (HQ378233) mtCOI sequences as outgroup species using MEGA 6.0 software. Branch truthness of phylogenetic trees constructed with different methods was assessed with bootstrap test (1000 replications).

All 52 mtCOI gene sequences were aligned and used in the genetic analysis of Z. caudatus comprising seven populations from five countries of Asia (Table 1). The annotated final length of mtCOI gene sequences was 637 bp. The averaged mtCOI gene sequences’ nucleotide composition was 35.7% T, 30.6% A, 13.3% G and 20.4% C with 45 parsimony informative sites, 5 singleton sites and a total of 50 variable sites. Basic descriptive genetic diversity results obtained from populations of five countries of Z. caudatus are presented in Table 1. The haplotype numbers per population (H) ranged from 1 to 10. Haplotype diversity (Hd) and nucleotide (π) diversity ranged from 0 to 1.0 and 0 to 0.00785 respectively. Highest Hd was found in Z. caudatus population of Thailand (1.0) followed by India (0.90196).

A total of 14 unique haplotypes from 52 individual sequences of 7 populations were identified from 3 geographic regions of Asia, viz. India, Malaysia and Indonesia (Table 1). Six haplotypes were shared by at least 2 individuals of Z. caudatus from the same or different populations and 8 haplotypes composed of only single individual sequences from 14 identified haplotypes. The most common haplotype was H11 followed by H12, H2 comprised of 25, 7 and 5 individuals respectively, from different populations of Z. caudatus. Haplotype 11 was shared by 3 populations of Malaysia, China and Thailand whereas haplotype H12 was found in only in Indonesia.

### Table 1

Sample size (n), number of unique haplotypes (H), haplotype diversity (Hd), nucleotide diversity (π), average number of nucleotide differences (K), number of variable sites (V), uncorrected average pairwise distances between samples (P) and per cent nucleotide composition for mtCOI gene for different populations of Zeugodacus caudatus

<table>
<thead>
<tr>
<th>Countries</th>
<th>n</th>
<th>H</th>
<th>Hd</th>
<th>K</th>
<th>V</th>
<th>P ± SE</th>
<th>%T</th>
<th>%C</th>
<th>%A</th>
<th>%G</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>18</td>
<td>10</td>
<td>0.90196</td>
<td>0.00477</td>
<td>3.03922</td>
<td>15</td>
<td>0.005 ± 0.001</td>
<td>35.8</td>
<td>20.3</td>
<td>30.4</td>
<td>13.5</td>
</tr>
<tr>
<td>Malaysia*</td>
<td>24</td>
<td>2</td>
<td>0.08333</td>
<td>0.00013</td>
<td>0.00013</td>
<td>1</td>
<td>0.000 ± 0.000</td>
<td>35.8</td>
<td>20.4</td>
<td>30.8</td>
<td>13.0</td>
</tr>
<tr>
<td>Indonesia*</td>
<td>7</td>
<td>1</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0</td>
<td>0.000 ± 0.000</td>
<td>34.9</td>
<td>20.9</td>
<td>30.3</td>
<td>14.0</td>
</tr>
<tr>
<td>China*</td>
<td>1</td>
<td>1</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>n/c</td>
<td>n/c</td>
<td>35.8</td>
<td>20.4</td>
<td>30.8</td>
<td>13.0</td>
</tr>
<tr>
<td>Thailand*</td>
<td>2</td>
<td>2</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>5</td>
<td>0.008 ± 0.003</td>
<td>35.8</td>
<td>20.4</td>
<td>30.7</td>
<td>13.1</td>
</tr>
</tbody>
</table>

*mtCOI gene sequences obtained from GenBank, NCBI; n/c, not calculated.
population of *Z. caudatus*. Haplotype 2 composed of five individuals shared by two populations from India, i.e. Mumbai and Bhagalpur. When we compared the endemic/private haplotypes formation in each population of different regions, populations with a higher degree of endemic/private haplotypes formation were from India (10 haplotypes from 18 individuals) and Malaysia (2 haplotypes from 24 individuals). Only one haplotype was detected from seven individuals of *Z. caudatus* from Indonesia.

The *mtCOI* median-joining network generated using NETWORK program displayed a simple genealogy of *Z. caudatus* for the entire population set of three major populations, i.e. India (Mumbai, Bhagalpur and Ranchi), Malaysia and Indonesia (Figure 1). Indian populations of *Z. caudatus* were found as connecting link between the *Z. caudatus* population of Malaysia and Indonesia (Figure 1). Most of the haplotypes detected in the present study were region-specific, i.e. haplotypes H1-H10 were from the Indian region of *Z. caudatus*. Whereas H11-H13 were from Malaysia, Thailand and China. Haplotype H14 was from Indonesian population of *Z. caudatus*.

Phylogenetic relationships were constructed with sequences available in GenBank and the *Z. caudatus* sequences generated in this study. Most of the Indian specimens of *Z. caudatus* form separate clades in the phylogenetic tree from other *Z. caudatus* lineages described earlier from Malaysia and Indonesia with high bootstrap support value (99) reconstructed with three different methods (Figures 2–4). The $p$ genetic distance among isolates of *Z. caudatus* from different countries varies from 0.004 to 0.059 (Table 2). Highest $p$ genetic distances were found between the populations of Indonesia and other countries isolates of *Z. caudatus*. Also, the $p$ genetic distance between different lineages of *Z. caudatus* varied from 0.018 to 0.059 (Table 3).

*Z. caudatus* is a pest of cucurbit plants infesting mostly flowers and immature fruits of the plant. The species from genus *Bactrocera* and *Zeugodacus* of the family Tephritidae formed many species complexes. Morphological characteristics sometimes have proven to be of limited help, particularly in identifying the sibling/cryptic species present in the species complex. An earlier study on the genetic relationship of *Bactrocera* species from India to other countries showed similarity with the reported sibling/cryptic species from other parts of the world. The present study reports the presence of distinct genetic lineage of *Z. caudatus* in India on the basis of *mtCOI* gene sequences. Genetically, *Z. caudatus* forms two different lineages present separately in two different hemispheres of the world. Indian populations of *Z. caudatus* form a separate clade in the phylogenetic reconstruction by different methods with earlier studied population of *Z. caudatus* from Indonesia and Malaysia including China and Thailand with *mtCOI* gene sequences and robust bootstrap support. *mtCOI* gene has already been in use to study and analyse the genetic diversity and resolution of cryptic species complex.
Figure 2. Phylogenetic tree constructed with minimum evolution method\cite{29} for the lineages present in *Zeugodacus caudatus* populations. Values near nodes present the bootstrap support value. Sequence data of *Zeugodacus tau* and *Zeugodacus cucurbitae* are used as out groups.

Figure 3. Phylogenetic tree constructed with maximum likelihood method based on Kimura-2 parameter model\cite{31} for the lineages present in *Zeugodacus caudatus* populations. Values near nodes present the bootstrap support value. Sequence data of *Zeugodacus tau* and *Zeugodacus cucurbitae* are used as out groups.
Figure 4. Phylogenetic tree constructed using neighbor-joining method\textsuperscript{30} with Kimura-2 parameters\textsuperscript{31} for the lineages present in Zeugodacus caudatus populations. Values near nodes present the boot strap support value. Sequence data of Zeugodacus tau and Zeugodacus cucurbitae are used as out groups.

Table 2. Estimates of evolutionary divergence ($p$ distance) between Zeugodacus caudatus populations from various geographical regions

<table>
<thead>
<tr>
<th>Populations of $B$. caudata</th>
<th>India</th>
<th>Malaysia</th>
<th>China</th>
<th>Thailand</th>
<th>Indonesia</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>0.0048</td>
<td>0.005</td>
<td>0.005</td>
<td>0.004</td>
<td>0.009</td>
</tr>
<tr>
<td>Malaysia</td>
<td>0.018</td>
<td>0.0001</td>
<td>0.000</td>
<td>0.002</td>
<td>0.009</td>
</tr>
<tr>
<td>China</td>
<td>0.018</td>
<td>0.000</td>
<td>n/c</td>
<td>0.002</td>
<td>0.009</td>
</tr>
<tr>
<td>Thailand</td>
<td>0.015</td>
<td>0.004</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>0.056</td>
<td>0.056</td>
<td>0.057</td>
<td>0.059</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Within group $p$ distance are shown in diagonal with bold letter. $p$ distance between group is shown below the diagonal. Standard error estimate(s) are shown above the diagonal.

Distinct genetic lineage has already been reported from $Z$. tau and Zeugodacus ascita based on mtCOI gene sequences\textsuperscript{18,34}. In the present study, we designated here the three distinct genetic lineages of $Z$. caudatus as lineage 1: Indonesia type, lineage 2: Malaysia type and lineage 3: India type. Lineage 2, i.e. Malaysia type was also comprised of $B$. caudatus populations from Thailand, China and one individual from India (Ranchi). The presence of distinct genetic lineage in Indian population of $Z$. caudatus was also confirmed by $p$ value of genetic distances generated for the whole population in this study. The $p$ genetic distance value of $Z$. caudatus of India and Indonesia was 0.056, whereas India and Malaysia, including China and Thailand, varied from 0.015 to 0.018. The calculated $p$ distance among a population of $Z$. caudatus is similar to the earlier studied population of $Z$. caudatus\textsuperscript{13,14}. This can also be reproduced with the mitochondrial haplotype network construction among individuals of $Z$. caudatus. Median-joining network clearly showed the presence of different genetic lineage in $Z$. caudatus population from India (Figure 1). The Indian population of $Z$. caudatus shared 10 haplotypes out of total 14 haplotypes detected in the global population analysis. Earlier, Lim et al.\textsuperscript{13} detected only two mtCOI gene haplotypes in $Z$. caudatus population collected from Malaysia and Indonesia.
Further, two more mtCOI gene haplotypes were reported by Yong et al. in Z. caudatus population studied from Indonesia, Malaysia, China and Thailand. Yong et al. also reported that Z. caudatus lineages present in the northern and southern hemispheres are distinct from each other. The present analysis also supports Yong et al. on the presence of distinct lineages of Z. caudatus in the northern and southern hemisphere of the world, as no one haplotype was found similar to the haplotype present in the population of Z. caudatus from the southern hemisphere (Indonesia). Similar reports on the presence of distinct lineages, species complex and cryptic species/sibling species in the same species of fruit fly with mtCOI gene sequences have also been established earlier. Overall 8 species/subspecies have been reported in the species complex of Z. tau around the world. With the result of present study and support from the earlier studies, we suggest presence of three sub-species/type in Z. caudatus species complex localized in different parts of the world and can be recognized as lineage 1: Indonesia type, lineage 2: Malaysia type and lineage 3: India type. Formation of distinct genetic lineages may be the result of isolation of different geographical population of species long back after origin. The notion is supported by the biology of the insect as adults of these flies breed on the flower of cucurbit plants and movement of larvae with flowers to long distances is not possible even through human-mediated transportation. Earlier, many fruit fly species which infested fruits of the host plants originated at one place and are currently distributed in many countries of the world without forming genetically distinct lineage has been explained that the large distance movement of the species was through human-mediated activities as fruits were used to carry by human during travelling to different places.

In conclusion, Z. caudatus populations present in India are genetically different from the Z. caudatus populations present in other parts of the world, thus revealing the presence of sibling species/cryptic species in Z. caudatus. The present study supports earlier reports on Z. caudatus that more distinct lineages may be present in other regions of the world which need to be studied. Also, the data generated in this study would be supportive in the future study on Z. caudatus species complex.

### Table 3. Estimates of evolutionary divergence (p distance) between lineages of Zeugodacus caudatus

<table>
<thead>
<tr>
<th>Lineages of B. caudata</th>
<th>Lineage 1</th>
<th>Lineage 2</th>
<th>Lineage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lineage 1 (Indonesia Type)</td>
<td>0.000 ± 0.000</td>
<td>0.056 ± 0.008</td>
<td>0.053 ± 0.008</td>
</tr>
<tr>
<td>Lineage 2 (Malaysia Type)</td>
<td>0.057 ± 0.009</td>
<td>0.001 ± 0.000</td>
<td>0.015 ± 0.005</td>
</tr>
<tr>
<td>Lineage 3 (India Type)</td>
<td>0.055 ± 0.005</td>
<td>0.018 ± 0.009</td>
<td>0.004 ± 0.001</td>
</tr>
</tbody>
</table>

Uncorrected p distance with standard error between lineages are shown below the diagonal. Corrected p distance with standard error between lineages are shown above the diagonal. The analysis involved 52 nucleotide sequences. Within lineage p distance is shown in diagonal with bold letter.


Received 8 August 2018; revised accepted 22 July 2019

doi: 10.18520/cs/v117/i8/1368-1375