Development of Allele-Specific Single-Nucleotide Polymorphism-Based Polymerase Chain Reaction Markers in Cytochrome Oxidase I for the Differentiation of *Bactrocera papayae* and *Bactrocera carambolae* (Diptera: Tephritidae)

Author(s): Tock H. Chua, B. K. Song, and Y. V. Chong
Published By: Entomological Society of America
DOI: 10.1603/EC10169
Development of Allele-Specific Single-Nucleotide Polymorphism-Based Polymerase Chain Reaction Markers in Cytochrome Oxidase I for the Differentiation of Bactrocera papayae and Bactrocera carambolae (Diptera: Tephritidae)

TOCK H. CHUA,1,2 B. K. SONG,3 AND Y. V. CHONG3


ABSTRACT  Differentiation of Bactrocera papayae Drew & Hancock and Bactrocera carambolae Drew & Hancock (Diptera: Tephritidae) based on morphological characters has often been problematical. We describe here a single-nucleotide polymorphism (SNP)-based polymerase chain reaction (PCR) assay to differentiate between these two species. For detection of SNPs, fragments derived from each species were amplified using two primer pairs, COIF/COIR and UEA7/UEA10, sequenced, and aligned to obtain a contiguous 1,517-bp segment. Two new sets of primers were designed based on the 11 SNPs identified in the region. Results of the SNP-PCR test using any one of these species-specific primer sets indicate that these two species could be differentiated on basis of presence or absence of a band in the gel profile. We also tested the SNP-PCR primers on Bactrocera umbrosa F., Bactrocera cucurbitae Coquillett, Bactrocera latifrons Hendel, and Bactrocera tau (Walker) but did not detect any band in the gel, indicating the likelihood of a false positive for B. papayae is nil. This SNP-PCR method is efficient and useful, especially for immature life stages or when only adult body parts of the two species are available for identification, as encountered often in quarantine work.

KEY WORDS  Bactrocera papayae, B. carambolae, species differentiation, cytochrome oxidase I, single-nucleotide polymorphism-polymerase chain reaction

Fruit and vegetable production in Malaysia is perennially plagued by infestation by fruit flies mainly of the genus Bactrocera (Diptera: Tephritidae), as in many other Asian countries. There are possibly at least a hundred Malaysian Bactrocera species of which approximately half have been recorded (Chua 1998). Of these, the Bactrocera papayae Drew & Hancock; Bactrocera carambolae Drew & Hancock; the melon fly, Bactrocera cucurbitae Coquillett; Bactrocera umbrosa F.; Bactrocera latifrons Hendel, and Bactrocera caudata (F.) are major agricultural pests that can cause serious losses and increase greatly the production cost.

Of these fruit fly pests, the two most serious pests in Malaysia are undoubtedly B. papayae and B. carambolae, both belonging to the Bactrocera dorsalis complex and considered as sibling species. Both are polyphagous and infest mainly carambola, Averrhoa carambola L.; water apple, Eugenia spp.; sapodilla (chiku), Manilkara zapota L.; guava, Psidium guajava L.; mango, Mangifera indica L.; and soursop, Annona muricata L. In addition, B. papayae also infests papaya, Carica papaya L.; banana, Musa spp.; brinjal, Solanum melongena L. var. esculentum; and chili, Capsicum annuum L.

Until mid-1990s, B. papayae and B. carambolae were not known by their current names. In fact for a long time, before exhaustive taxonomy study by Drew and Hancock (1994) was undertaken, both were mistakently identified as a single species, viz. Dacus dorsalis Hendel (e.g., Tan and Jaal 1986). Sometimes they were separated into Dacus dorsalis and Dacus pedestris (Bezzi), respectively (e.g., Vijaysegaran 1983), although both these do not occur in Malaysia. Subsequently, they were referred to as Malaysian A and Malaysian B (Drew 1991) or Taxon A and Taxon B (Ooi 1991) or as sp. nr B. dorsalis (A) and (B) by White and Elson-Harris (1992).

Unfortunately, the adults of these two species are very similar morphologically, and mistakes in identification occur frequently. Differentiating these two species could be almost impossible if only immature stages or body parts are available as in quarantine work that requires definite species identification in the shortest time possible. It may take a few days to rear out the adults for identification, and this may be disadvantageous to the importers, exporters, or to the pest control personnel. In these cases, molecular identification would be an advantage, because it is more accurate and takes shorter time.
Mitochondrial DNA is used as a common marker for identification of species in fruit flies (Han and McPherson 1997, Muraji and Nakahara 2002). The mitochondrial control region also has been used as a marker to detect genetic variations in *B. dorsalis* (Nakahara et al. 2008). Recent studies on molecular identification of insects are based mostly on ITS and internal transcribed spacer (ITS) sequences by using polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP) (Salazar et al. 2002, Muraji et al. 2004, Ebina and Kenji 2006). Molecular diagnostics have also been used by insect researchers to infer phylogenetic relationships and for identification to species level (Chua et al. 2009, Smith and Bush 1997, Jannongluk et al. 2003, Orlandi et al. 2003). In general, this technique involves designing of allele-specific forward primers, by incorporating SNPs in the 3′ end of the oligonucleotides and allows amplification only of the sequences harboring the SNPs.

The current study is a continuation of the research seeking a definite and simpler way of differentiating *B. papayae* and *B. carambolae*, based on the SNPs in *B. papayae*. We describe the use of SNP-PCR technique to differentiate *B. papayae* from *B. carambolae*.

### Materials and Methods

#### Biological Samples

Adult specimens of *B. papayae*, *B. carambolae*, and other *Bactrocera* spp. were collected from different parts of Malaysia (Table 1) by using methyl eugenol and cue lure as the chemical lures. The specimens were either stored at −20°C freezer or preserved in 96% alcohol before analysis. In total, 31 adults were used in this analysis.

#### DNA Extraction, Amplification, and Sequencing

Total DNA was extracted using DNeasy tissue kit (QIAGEN, Valencia, CA) following the manufacturer’s protocol for animal tissue with slight modifications to increase DNA yield.

PCR amplification of COI markers of the two species was performed with the primer pairs COIF/COIR (Tan 2004) and UEA7/UEA10 (Lunt et al. 1996) (Table 2). The amplifications conditions for COIF/COIR were as follows: an initial denaturation at 95°C for 3 min, 40 cycles at 95°C for 40 s, 35 cycles at 60°C for 1 min, extension at 72°C for 1 min, and final extension at 72°C for 15 min. The conditions for UEA7/UEA10 were as follows: an initial denaturation at 94°C for 3 min, 35 cycles at 94°C for 1 min, 50°C for 1 min.

### Table 1. List of specimens of *Bactrocera* spp. collected in 2007–2010 and used in the study: their collection sites, preservation methods, and number of individuals

<table>
<thead>
<tr>
<th>Species</th>
<th>Collection site</th>
<th>Preservation method</th>
<th>No. individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. papayae</td>
<td>Damansara Jaya</td>
<td>96% alcohol</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Kota Kinabalu</td>
<td>96% alcohol</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Pantai Hillpark</td>
<td>96% alcohol</td>
<td>4</td>
</tr>
<tr>
<td>B. carambolae</td>
<td>Damansara Jaya</td>
<td>96% alcohol</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Pantai Hillpark</td>
<td>96% alcohol and</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>−20°C freezer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. umbrosa</td>
<td>Damansara Jaya</td>
<td>−20°C freezer</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Pantai Hillpark</td>
<td>96% alcohol</td>
<td>1</td>
</tr>
<tr>
<td>B. latifrons</td>
<td>Petaling Jaya</td>
<td>−20°C freezer</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Lukut</td>
<td>−20°C freezer</td>
<td>1</td>
</tr>
<tr>
<td>B. cucurbitae</td>
<td>Kerinchi</td>
<td>−20°C freezer</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Lukut</td>
<td>−20°C freezer</td>
<td>1</td>
</tr>
<tr>
<td>B. tau</td>
<td>Fraser’s Hill</td>
<td>−20°C freezer</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Lukut</td>
<td>−20°C freezer</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Gombak</td>
<td>96% alcohol</td>
<td>2</td>
</tr>
</tbody>
</table>

### Table 2. Oligonucleotides used for DNA amplification and sequencing of the *Bactrocera* spp.

<table>
<thead>
<tr>
<th>Primer Set</th>
<th>Primer</th>
<th>Sequence</th>
<th>Expected size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set 1</td>
<td>COIF</td>
<td>5’-TACAATTTATGCCTAAACCTTACGCC-3’</td>
<td>1,300</td>
</tr>
<tr>
<td></td>
<td>COIR</td>
<td>5’-CATTCAAGTTGTGTAAGGATGC-3’</td>
<td>Tan (2004)</td>
</tr>
<tr>
<td>Set 2</td>
<td>UEA7</td>
<td>5’-TACATGTTGAAATACCTTCAAGC-3’</td>
<td>700</td>
</tr>
<tr>
<td></td>
<td>UEA10</td>
<td>5’-TCTCAATGGACTAAATCTGGCATTATTA-3’</td>
<td>Lunt et al. (1996)</td>
</tr>
<tr>
<td>Set 3</td>
<td>BPSN1PF</td>
<td>5’-GGGGGACCTTCAAGTGC-3’</td>
<td>1,030</td>
</tr>
<tr>
<td></td>
<td>BPSN1PR</td>
<td>5’-GACTGCGTGTGATAGGTAAC-3’</td>
<td></td>
</tr>
<tr>
<td>Set 4</td>
<td>BPSN2PF</td>
<td>5’-GGGGGACCTTCAAGTGC-3’</td>
<td>1,070</td>
</tr>
<tr>
<td></td>
<td>BPSN2PR</td>
<td>5’-TTCAGCTGGGAGGATATTG-3’</td>
<td></td>
</tr>
</tbody>
</table>
extension at 72°C for 1 min, and final extension at 72°C for 30 min.

The PCR products were electrophoresed on a 1% SeaKem LE agarose gel. The DNA fragment sizes were estimated by comparing with 1-kb commercial markers (Promega, Madison, WI). The band corresponding to the target PCR product was excised and purified using QIAquick gel extraction kit (QIAGEN). All samples were then sent to First BASE Laboratories Sdn. Bhd. (Seri Kembangan, Malaysia) for sequencing.

**SNP Differentiation.** Fragments amplified by using two primer pairs COIF/R and UEA7/10 were sequenced, and the sequences obtained to align a contiguous 1,517-bp segment. The derived longer nucleotide sequence was examined for SNPs that would be useful for distinguishing between the two species. Using *B. papayae* COI sequence as reference template, two primer sets were manually designed with the SNPs incorporated into the 3′ end of the forward and reverse primer. These primer sets were labeled as BPSNP1 F'/BPSNP1R and BPSNP2 F'/BPSNP2R.

**PCR amplification** using these SNP-based primers was then performed for *B. papayae* and *B. carambolae* under the following conditions: 25 μl of reaction mix containing 10 mM Tris-HCl, pH 8.0, 50 mM KCl, 2.5 mM MgCl₂, 0.2 μM each primer, 0.25 mM dNTP mix, and 1 U of TaqDNA polymerase (Promega). DNA amplifications were carried out in a thermal cycler (MJ Mini, Bio-Rad Laboratories, Hercules, CA) with an initial denaturation at 95°C for 5 min followed by 35 cycles of denaturation at 95°C for 40 s; primer annealing at 60°C for 1 min; and primer extension at 72°C for 1 min, with a final extension at 72°C for 15 min. Presence and size of amplified products in PCR reactions were analyzed by gel electrophoresis on 2% agarose gel in 1× TAE buffer. PCR amplification and gel electrophoresis were repeated one to three times for each individual of *B. papayae* and *B. carambolae*.

Both the SNP-PCR primer sets also were tested on other economically important *Bactrocera* spp., viz., *B. umbrosa, B. cucurbitae, B. latifrons,* and *Bactrocera tau* (Walker).

**Results**

From examining the derived 1,517-bp sequence of the COI region, 53 polymorphic sites between the two *Bactrocera* species were revealed (Fig. 1). None of the identified SNPs would cause amino acid substitution. Using *B. papayae* COI sequence as reference template, two pairs of SNP-PCR primer were successfully designed to amplify internal fragments of COI gene (Table 2). Among these oligonucleotide pairs, two were flanked by the same forward primer BPSNP1F (also labeled as BPSNP2F).

The amplified PCR product was the same as the expected size (1,030 and 1,070 bp for primer pair set BPSNP1 and BPSNP2, respectively; Fig. 2), which was estimated on the basis of *B. papayae* COI genomic sequence. As expected, both of the primer sets did not amplify for the COI regions in *B. carambolae*. Similarly, both the primer pair sets did not amplify for the COI regions in *B. umbrosa, B. cucurbitae, B. latifrons,* and *B. tau* because no band was detected in the gel.

**Discussion**

This study demonstrates the use of species-specific COI primers that can be used for differentiating the two major Malaysian *Bactrocera* pests. Currently, identification by quarantine officers using morphological characters can be time-consuming and sometimes technically difficult for these closely related species. Diagnostic morphological characters for immature stages are also lacking for most tephritids (White and Elson-Harris 1992). Although a PCR-RFLP protocol developed by Chua et al. (2010) using a combination of UEA primers and Rsal restriction enzyme seemed to be promising to separate *B. carambolae* from *B. papayae*, it remains uncertain whether the resulted banding profiles of the COI digested with Rsal for the *Bactrocera* spp. would be consistent when analysis is performed across large number of samples. The lack of correspondence between the in silico digestion of the COI sequences and PCR-RFLP results, as mentioned by the authors, could indicate a possible heteroplasmic makeup of cell’s mitochondria in the *Bactrocera* species. Moreover, an extra step of endonuclease cleavage would mean extra time for determining the species identity. This makes the method less advantageous in fulfilling requirements of timeliness and sensitivity, which are principal considerations in quarantine work.

That these SNP-PCR primers did not amplify the COI region of other *Bactrocera* spp. (*B. umbrosa, B. cucurbitae, B. latifrons,* and *B. tau*) would not reduce, but rather enhance the usefulness of the primers. This would indicate the probability of a false positive for *B. papayae* would be almost nil as far as the Malaysian economically important *Bactrocera* spp. are concerned. The adults of these species are so different morphologically that they can be identified easily, and that their host ranges are so different that problems on determining the immature stages would not arise. Nevertheless, if the need arises, the PCR-RFLP method of identifying these species is available (Chua et al. 2010). Because *B. papayae* and *B. latifrons* have been recorded from chili, the immature stages of these two species also can be separated using the SNP-PCR primers as only *B. papayae* will produce a band in the electrophoresis gel.

Although the specimens of *B. papayae* and *B. carambolae* tested here were collected from a limited geographical range (Peninsular Malaysia as well as East Malaysia in the Island of Borneo), we have no reasons to believe that the SNP-PCR primers would not work for specimens from other parts of Southeast Asia. Furthermore, our previous work (Chua et al. 2010) on blasting COI sequences of specimens of these species from various geographical regions had indicated very minor variation.

Our present work has successfully converted COI DNA barcodes to candidate SNPs linked to the gene, leading to development of two SNP-PCR markers for
distinguishing *B. papayae* and *B. carambolae*. By having the SNPs-contained PCR primer sets, differential identification of such species would remove the latter step of restriction enzymes digestion to obtain the restriction fragment polymorphism profiles for separating *B. papayae* and *B. carambolae*. The limited length of sequence (1,517 bp) and low GC content within the targeted COI segment (34.3%) would limit the choice of SNPs and primer sequences. An SNPs survey of Moorhead et al. (2003) on development of SNP-PCR...
markers to differentiate between lineages of *Listeria monocytogenes*, a pathogen of significant public health, also has been suggested a similar limitation in multiple assay. Further research studying the extended COI or other genetic regions may be helpful in obtaining more SNPs information and increase choices of primer sequences.

The molecular technique described here is useful for identifying the two *Bactrocera* species accurately. As a means of genetic marker, the SNP-based PCR is simpler, more sensitive, rapid, reliable, and cost-effective alternative to the PCR-RFLP described by Chua et al. (2009). It is also convenient for any small quarantine laboratories at national borders or airports because all steps consist of simple amplification and involve only widely used PCR equipment. The SNP-PCR marker would be a useful tool in quarantine work involving only widely used PCR equipment. The SNP-PCR marker would be a useful tool in quarantine work or without having to rear the larvae and wait for adult emergence.

Acknowledgments

We thank Ministry of Science and Technology for funding the research (E Science Fund 05-02-10-SF0015) and the University for providing general facilities for conducting the research.

References Cited


Received 9 May 2010; accepted 19 August 2010.