

# New record of Carnidae (Diptera) from Taiwan and potential challenges in DNA barcode amplification due to pseudogene

Hsuan-Pu Chen<sup>‡</sup>, Fang-Tse Chan<sup>§</sup>, Shih-Feng Shiao<sup>‡</sup>, Ming-Chung Chiu<sup>‡</sup>

<sup>‡</sup> Department of Entomology, National Taiwan University, Taipei, Taiwan

<sup>§</sup> Wildlife Rescue and Research Center, Taiwan Biodiversity Research Institute, Nantou, Taiwan

Corresponding author: Ming-Chung Chiu ([mingchung310@ntu.edu.tw](mailto:mingchung310@ntu.edu.tw))

Academic editor: AJ Fleming

## Abstract

## Background

The genus *Carnus* Nitzsch, 1818 comprises small ectoparasites that feed on the blood of juvenile avians. They are characterised by dealated adults with setose abdominal intersegmental membranes. *Carnus orientalis* Maa, 1968 was previously recorded in Malaysia and the Ryukyu Islands of Japan, parasitising two owl species: *Ketupa ketupu* (Horsfield, 1821) and *Otus elegans* (Cassin, 1852). This study confirms the occurrence of *C. orientalis* in Taiwan and presents a new host record, along with *COI* barcode sequences. Additionally, the study also elucidates the difficulties posed by blood meal contamination and pseudogene amplification as confounding factors intrinsic to the molecular taxonomic delineation of *C. orientalis* via universal DNA barcoding primers.

## New information

The following new information regarding *C. orientalis* is provided in this study:

1. *Carnus orientalis* is first recorded in Taiwan, filling the gap in its East Asian distribution. This is also the first record of Carnidae from Taiwan.
2. *Otus lettia* (Hodgson, 1836) (Aves, Strigidae) is reported as a new host for *C. orientalis*, identified on a fallen fledgling.
3. Co-amplification of the host's *COI* is reported in this study using the universal PCR primer set LCO1490/HCO2198. Additionally, the amplification of a *COI*-like pseudogene using a newly-designed primer set is detected through abnormal translated amino acid sequences and the occurrence of a stop codon.
4. New specific primers for the *COI* gene of *Carnus* were designed in this study.

The new distribution and ecological data of *C. orientalis* enhance our understanding of this species. The provision of new *COI* primers is anticipated to contribute to future studies employing DNA barcoding in bird-parasitic flies.

The following new information regarding *C. orientalis* is provided in this study:

1. *Carnus orientalis* is first recorded in Taiwan, filling the gap in its East Asian distribution. This is also the first record of Carnidae from Taiwan.
2. *Otus lettia* (Hodgson, 1836) (Aves, Strigidae) is reported as a new host for *C. orientalis*, identified on a fallen fledgling.
3. Co-amplification of the host's *COI* is reported in this study using the universal PCR primer set LCO1490/HCO2198. Additionally, the amplification of a *COI*-like pseudogene using a newly-designed primer set is detected through abnormal translated amino acid sequences and the occurrence of a stop codon.
4. New specific primers for the *COI* gene of *Carnus* were designed in this study.

The new distribution and ecological data of *C. orientalis* enhance our understanding of this species. The provision of new *COI* primers is anticipated to contribute to future studies employing DNA barcoding in bird-parasitic flies.

## Keywords

DNA barcode, *COI*, pseudogene, new record, Carnidae

## Introduction

Bird flies, *Carnus* Nitzsch, 1818, constitute a small group of ectoparasite of avians belonging to the family Carnidae. Unlike most saprophagous carnids, *Carnus* feeds on the blood or skin secretions of its hosts (Brake 2011). Nearly all the known hosts of *Carnus* are tree-nesting birds (Grimaldi 1997). Adult *Carnus* remove their wings after finding a bird host and live in the nests (Grimaldi 1997, Brake 2011). With the particular behaviour of removing the wings, the dealated adult is the most remarkable diagnostic characteristic of the genus *Carnus*. Additionally, *Carnus* can be morphologically distinguished from other carnid genera by the absence of cross-vein dm-cu, reduced abdominal sternites I–V in females, physogastry in females and setose abdominal intersegmental membrane (Brake 2011).

Most of the *Carnus* are sampled on the host in their nest and, thus, are usually overlooked in regular insect sampling due to their strong association with the bird. Therefore, their biodiversity and distribution could be highly underestimated. In this study, five adult bird flies were collected in Taiwan from a fallen fledgling of collared scops owl (*Otus lettia* (Hodgson, 1836)). This finding marks the first record of Carnidae in Taiwan. Five valid species of the genus *Carnus* have been described around the world, including *C. floridensis* Grimaldi, 1997, *C. hemapterous* Nitzsch, 1818 and *C. occidentalis* Grimaldi, 1997 from the Holarctic, *C. mexicana* Grimaldi, 1997 from the Neotropical Region and *C.*

*orientalis* Maa, 1968 from the Oriental Region (Grimaldi 1997). Additionally, Grimaldi (1997) reported other two morphospecies (sp. A and sp. B) from Mexico. However, since the male specimens of these morphospecies were not collected thus making the critical diagnostic characters of male genitalia unavailable, Grimaldi did not describe them. Furthermore, DeConinck (1986) reported another undescribed *Carnus* species from the Afrotropical Region, collected from the nestlings of the kingfisher *Halcyon albiventris* (Scopoli, 1786), which remains undescribed. In East Asia, only one species, *C. orientalis*, has been reported in Malaysia (Maa 1968) and Ryukyu Islands (Iwasa et al. 2014) from the nestlings of fish owl (*Ketupa ketupu* (Horsfield, 1821)) (Strigidae) and Ryukyu scops owl (*Otus elegans* (Cassin, 1852)) (Strigidae), respectively. Taiwan, a continental island in subtropical East Asia, serves as an intermediary between the known distribution sites of this species. Therefore, the occurrence of *C. orientalis* in Taiwan is expected, based on the known distribution of this species and its hosts (Maa 1968, König et al. 1999, Iwasa et al. 2014).

The primary objectives of this study are to characterise the collected *Carnus* specimens in Taiwan through morphological descriptions and DNA barcoding. Additionally, we share our experiences addressing issues related to host blood contamination and pseudogene interference during the amplification of the cytochrome c oxidase I (*COI*) gene as the DNA barcode. To date, *COI* barcoding has been a crucial tool for the identification of animals based on genetic similarity and finds wide applications in various fields (Hebert et al. 2003a, Hebert et al. 2003b, Hebert et al. 2004, Hajibabaei et al. 2006, Janzen et al. 2009). However, the use of universal PCR primer sets could lead to the co-amplification of the non-target genes in the blood meal of blood-sucking insects (Reeves et al. 2018) and pseudogenes, which are non-functional copies of mitochondrial sequences incorporated into the nuclear genome (Song et al. 2008, Leite 2012). Both scenarios render the universal PCR primer sets ineffective for amplifying the barcoding genes and both have the potential to occur during the molecular identification of *Carnus* species. In addition to host blood contamination, typically during the barcoding amplification in blood-sucking insects, pseudogenes are likely to be prevalent in dipteran genomes (Harrison et al. 2003). These challenges in amplifying the barcoding genes will likely result in a limited representation of DNA barcode information in public databases. Given that these issues occurred in this study, newly-designed primer sets, explicitly targeting the *COI* partial sequences (COI-5P) for *Carnus*, are also introduced through a comparative analysis with sequences from the host blood meal and pseudogene.

## **Materials and methods**

### **Morphological examination**

Morphological terminology follows the conventions established by Grimaldi (1997) and Brake (2000). The measuring method is shown in Suppl. material 1. The chaetotaxy was counted from the left side of specimens. Specimens were examined and measured using a LEICA S8APO microscope (Leica Microsystems, Germany) equipped with an XFCAM autofocus CCD (Jet measurements, Taiwan). Photographs were captured with a LEICA

DMC5400 in conjunction with a LEICA Z16 APO, utilising the auto-stacking system LAS V4.13 (Leica Microsystems, Germany). The line drawings were made by Procreate (Savage Interactive, Australia). All figures were edited and arranged into figure plates by Adobe Illustrator CC and Photoshop CC (Adobe Systems Inc., San Jose, CA, USA). The five specimens used in this study have been deposited in the National Museum of Natural Science (NMNS), Taichung, Taiwan.

## Molecular data

Total genomic DNA was extracted by the DNeasy Blood and Tissue Kit (Qiagen, Düsseldorf, Germany) following the non-destructive protocol (Cruaud et al. 2019) from samples with two treatments: the whole body (n = 2; CM02 and CM03) and body with abdomen removed for reducing the host blood contamination (n = 2; CM01 and CM04). Partial sequences of one mitochondrial cytochrome c oxidase I (*COI*) were amplified by PCR. The primer sets and conditions used are shown in Table 1. Each PCR was conducted in a 15 µl volume containing 4.3 µl sterile distilled water, 0.6 µl of each primer (10 µM), 7.5 µl of GoTaq® Green Master Mix (Promega, Madison, WI, USA) and 2.0 µl of DNA template. All PCR products were purified and sequenced by Tri-I Biotech Inc. (Taipei, Taiwan). Sequences were edited using Codoncode Aligner version 10.0.2 (CodonCode Corporation, Dedham, MA, USA), aligned and translated into amino acids to check for stop codons in MEGA11 (Tamura et al. 2021). Position numbering starts from the first amino acid or nucleotide site of the gene (reference genomes: *Phyllomyza* sp. (Diptera, Milichiidae) and *Drosophila melanogaster* (Diptera, Drosophilidae)) (Table 2).

## Process of specific primer design

Using the universal primer set LCO1490/HCO2198 (Folmer et al. 1994), low-quality sequences or the host's *COI* sequences from the *COI*-5P region were amplified in all four *Carnus* samples. Therefore, a *Carnus*-specific forward primer, C1-J-1571\_Carnus, was designed with its binding site located 56 base pairs downstream of the binding site of LCO1490, based on a comparison of three complete *COI* sequences from GenBank (*Otus lettia*, *Phyllomyza* sp., *Drosophila melanogaster*) and four *COI*-5P sequences of *Carnus* species (*C. hemapterus*, *Carnus* sp. CA1, and *Carnus* sp. CA2) from BOLD systems (Ratnasingham and Hebert 2007). The use of primer set C1-J-1571\_Carnus/HCO2198 led to the amplification of the *COI*-like pseudogene, so specific internal primers were designed to amplify the real *COI* (*COIF*\_Carnus and *COIR*\_Carnus) and *COI*-like pseudogene (*COIF*\_pseudo\_Carnus and *COIR*\_pseudo\_Carnus) of *Carnus*. Detailed information on the sampled sequences is provided in Table 2 and the relative binding sites of the newly-designed primers are illustrated in Suppl. material 2.

## Molecular phylogeny and species delimitation

A total of 12 *COI*-5P sequences, including 11 from *Carnus* spp. and one from the sister genus of *Carnus*, *Meoneura triangularis* (Diptera, Carnidae) (Buck 2006) as the outgroup, were used from BOLD systems for phylogeny reconstruction and species

delimitation analysis (Table 2). The Maximum Likelihood (ML) phylogenetic trees were reconstructed by the programme IQ-TREE 1.6.12 (Nguyen et al. 2015) through the web server W-IQ-TREE (Trifinopoulos et al. 2016) (available at <http://iqtree.cibiv.univie.ac.at/>) and rooted by the outgroup *Meoneura triangularis*. Nodal support was assessed with ultrafast bootstrap approximation (UFBoot2) (Hoang et al. 2017) and SH-like approximate likelihood ratio test (SH-aLRT) (Guindon et al. 2010), based on the default parameter setting in the ML method. The nodes with SH-aLRT  $\geq 80\%$  and UFBoot  $\geq 95\%$  were considered strongly supported.

Two DNA-based methods, assembling species by Automatic Partitioning (ASAP) (Puillandre et al. 2020) and Bayesian-based Poisson Tree Processes (bPTP) (Zhang et al. 2013), were conducted for species delimitation, based on the *COI* dataset. ASAP and bPTP were conducted respectively with the *COI* dataset and the unrooted *COI* gene tree. The analyses were performed at web interfaces (ASAP: <https://bioinfo.mnhn.fr/abi/public/asap/> and bPTP: <https://species.h-its.org/>) with default parameter settings. Pairwise K2P nucleotide genetic distances (K2P distances) were calculated using MEGA11 to evaluate the degree of interspecific genetic diversity.

## Taxon treatment

### *Carnus orientalis* Maa, 1968

- ZooBank [0987F4A7-D043-4907-B518-0FD8E01BDDC1](https://zoobank.org/0987F4A7-D043-4907-B518-0FD8E01BDDC1)

#### Nomenclature

*Carnus orientalis* Maa, 1968: 33; holotype, ♀; Type locality: Malaysia, Selangor, Rantau Panjang. Deposited at Bishop Museum, Hawaii, USA.

#### Material

- a. scientificName: *Carnus orientalis* Maa, 1968; originalNameUsage: *Carnus orientalis* Maa, 1968; namePublishedIn: Maa, Tsing-Chao. 1968. A new *Carnus* from Malaya (Diptera: Milichiidae). *Pacific Insects* 10(1): 33-36.; namePublishedInID: urn:lsid:zoobank.org:pub:6BF9831E-2AE1-49C4-BDD2-523E7BAD9FF8; taxonomicStatus: accepted; taxonID: urn:lsid:zoobank.org:act:0987F4A7-D043-4907-B518-0FD8E01BDDC1; higherClassification: Animalia; Arthropoda; Insecta; Diptera; Carnidae; Carnus; Carnus orientalis; kingdom: Animalia; phylum: Arthropoda; class: Insecta; order: Diptera; family: Carnidae; taxonRank: species; vernacularName: 東方鳥蠅, ミナミトリチスイコバエ; nomenclaturalCode: ICZN; genus: *Carnus*; specificEpithet: *orientalis*; scientificNameAuthorship: T. C. Maa; higherGeography: East Asia; Taiwan; Nantou; Xinyi Township; Ziqiang Village, Wusonglun; continent: Asia; islandGroup: Taiwan; island: Taiwan; country: Taiwan; countryCode: TW; stateProvince: Nantou; municipality: Xinyi Township; locality: Ziqiang Village, Wusonglun; decimalLatitude: 23.677539; decimalLongitude: 120.847545; eventDate: 2023/05/03; year: 2023; month: 5; day: 3; individualCount: 5; sex: 1 female, 4 males; lifeStage: adult; preparations: whole animal (ETOH), DNA extracted; occurrenceRemarks: collected on rescued individual of *Otus letitia*; recordedBy: Ming-Jung Chan; disposition: in collection; associatedSequences:

GenBank: PP192111, PP192112, PP192113, PP192114, PP199482, PP199483, PP199484, PP199485, identifiedBy: Shih-Tsai Yang, Hsuan-Pu Chen; dateIdentified: 2023; identificationReferences: Iwasa et al. 2014; type: PhysicalObject; institutionCode: National Museum of Natural Science (NMNS), Taichung, Taiwan; basisOfRecord: PreservedSpecimen; occurrenceID: 24F24E3A-C843-5F4E-BC68-01E02025B726

## Description

The description was based on five Taiwanese specimens (four males and one female) (Figs 1, 2, 3).

**Male.** Head polished and smooth, 1.88–2.27 ( $2.05 \pm 0.16$ ) $\times$  as wide as long; eye (Fig. 2A and C) suboval, 0.38–0.57 ( $0.49 \pm 0.08$ ) $\times$  as dorsal wide as height, 0.71–0.95 ( $0.85 \pm 0.11$ ) $\times$  as lateral wide as height; interocular width 0.55–0.63 ( $0.57 \pm 0.04$ ) $\times$  as wide as head width (Fig. 2A); minimum height of gena 0.29–0.33 ( $0.31 \pm 0.02$ ) $\times$  eye height; postgenal width 0.21–0.24 ( $0.23 \pm 0.01$ ) $\times$  eye height, 0.22–0.33 ( $0.27 \pm 0.05$ ) $\times$  of lateral eye width (Fig. 2C); antenna with arista long and pubescent (Fig. 2B); palpus clavate (Fig. 2C); proboscis with labial theca longer than wide, bulbous and strongly sclerotised, labellum small with single ring of long setae at base (Fig. 2C); Chaetotaxy (Fig. 2) with interfrontal setae two, medioclinate; frontal setae two, medioclinate; orbital setae two, lateroclinate; ocellar setae two, anteroclinate; postocellar seta absent; supra-antennal seta one, anteroclinate; medial vertical seta one, medioclinate; lateral vertical seta one, postero-lateroclinate; paraverticilar seta one, medioclinate; postvertical seta absent; oral vibrissa one, with two setae between antenna and oral vibrissa; genal seta three.

Thorax (Fig. 1A and B) polished and coriaceous; mesoscutum 1.00–1.08 ( $1.03 \pm 0.04$ ) $\times$  as wide as long; scutellum wide and short; wing shaded with remnants of wing base; haltere clavate; Chaetotaxy with postpronotal seta one; presutural seta one; notopleural setae two; postsutral supra-alar seta one; postalar seta one; dorsocentral seta one; prescutellar seta one; scutellar setae two, with one basal and one apical seta. Fore coxa enlarged, mid- and hind coxa small; Legs with fore and hind femora enlarged, fore femora with two or three posteroventral setae (Fig. 1A).

Abdomen (Fig. 1A and B) with intersegmental membrane setose, setae strong and growing from strongly sclerotised spots; tergites I–V rectangular, covered with one row of strong and long seta apically, width-length ratio of tergites I–IV 5.50–7.71 ( $6.41 \pm 1.08$ ), 3.64–4.38 ( $3.94 \pm 0.33$ ), 3.43–4.00 ( $3.82 \pm 0.26$ ), 2.39–2.73 ( $2.51 \pm 0.16$ ); sternite present, square or rectangular.

Genitalia with epandrium (Fig. 3A and B) strongly sclerotised and rounded, with one long and robust seta dorsally, one dorso-laterally, one posteriorly and three shorter strong setae around the cerci; cercus (Fig. 3B) weakly sclerotised and convex posteriorly, with short setae dorsally; surstylus (Fig. 3B) broad trapezoidal, 1.78–2.00 ( $1.94 \pm 0.11$ ) $\times$  as long as wide, inner-curved with short setae and several small tooth-like projections apically; decasternum (Fig. 3B) laterally elongated plate with ventral projection medially, 2.44–2.75 ( $2.58 \pm 0.14$ ) $\times$  as wide as long, 0.89–1.00 ( $0.94 \pm$

0.06)× as long as maximum width of surstylus; strongly sclerotised and rounded ventrally; hypandrium (Fig. 3C and D) U-shape, opened and anteriorly protruded dorsally and connected ventrally; aedeagus (Fig. 3A) short, bulbous and membranous; aedeagal apodeme (Fig. 3A and C) clavate apically; paraphysis (Fig. 3A and C) long and slender.

Colouration (Fig. 1A and B) blackish-brown, except the eyes reddish-brown; basal and apical part of tibia, tarsus, abdominal intersegmental membrane yellowish-white.

**Female.** Similar to males, except abdomen physogastric (Fig. 1C and D), tergites narrow (Fig. 1D) and sternites I–V reduced. Measurements with head 1.82× as wide as long, eye 0.86× as lateral wide as high, 0.50× as dorsal wide as high; minimum height of gena 0.27× of eye height; interocular width 0.55× as head width; postgena 0.18 as high as eye, 0.21× lateral eye width; mesoscutum 1.16× as wide as length; tergites I–IV with wide-length ratios 8.71, 4.08, 3.77, 2.06; ovipositor (Fig. 1D) long, weakly sclerotised. Ovipositor (Fig. 1C and D) yellowish-white.

### Diagnosis

Combining the re-description in the current study and diagnoses proposed by Grimaldi (1997) and Iwasa et al. (2014), this species can be distinguished from other congeners by the combination of the following characters: Gena narrow, with its minimum height 0.25–0.33 of eye height; female abdominal tergites large (Fig. 1D vs. Grimaldi 1997: 14); surstylus broad trapezoidal, with several small tooth-like projections apically; paraphysis slender; decasternum with height 0.7–1.0 of maximum surstylus width.

### Distribution

Malaysia (Maa 1968), Japan (Ryukyu Islands) (Iwasa et al. 2014), and Taiwan (this study).

### Biology

*Carnus orientalis* is considered a blood-sucking ectoparasite of owls (Strigidae: *Ketupa ketupu*, *Otus elegans* and *Otus lettia*) (Maa 1968, Iwasa et al. 2014, this study). A video (<https://youtu.be/7KmhYKcGmg>) of the living individuals collected in this study shows host blood in the flies' guts, providing evidence of the blood-sucking nature of this species. The dealated adults can be found in the owl nests or on their fledging. The congeners *C. hemapterus* and *C. occidentalis* have wider host range spanning several bird families and different hosts can be recorded even within close localities (Grimaldi 1997). Therefore, the possibility that *C. orientalis* has a broader host range beyond owls cannot be omitted.

## Taxon discussion

The re-description, based on Taiwanese specimens, mostly fits the re-description of Iwasa et al. (2014). Still, morphological variations were found between these re-descriptions, including: (1) minimum height of gena 0.27–0.33 of eye height in Taiwanese specimens, but 0.25–0.29 in Iwasa et al. (2014); (2) interfrontal setae two in Taiwanese specimens, but one in Iwasa et al. (2014). The interfrontal seta observed in the present study is short and somehow inconspicuous. According to the illustration of Iwasa et al. (2014) (page 485, fig. 1), they could misinterpret the suprantennal seta as the interfrontal seta.

## Analysis

### DNA barcode and *COI*-like pseudogene

Compared with the morphologically similar species *C. hemapterus*, the differentiation between the two species at the *COI*-5P barcode can be translated into 0.0888 of K2P genetic distance (Table 3).

Using the universal primer set LCO1490/HCO2198 resulted in co-amplification of the blood meal of the host. Additionally, the use of the primer set C1-J-1571\_Carnus/HCO2198 resulted in the amplification of a *COI*-like pseudogene, where a two-base pair deletion occurred at nucleotide position 420–421 when aligned with the real *COI* of other *Carnus* species. The amino acid sequences of *COI*-like pseudogene were found to exceed the intrageneric variation of *Carnus COI*, with amino acid position Nos. 42 (P vs. G), 47 (R vs. G), 51 (N vs. I), 82 (S vs. P), 88 (L vs. P), 140 (deletion vs. S), 141 (deletion vs. V), 145 (S vs. I), 176 (C vs. F), 185 (F vs. S) and 216 (L vs. S). A stop codon at amino acid position 242 also occurred when translated to amino acids (see Suppl. material 3).

### Molecular-based species delimitation

The analytical dataset consists of 16 *COI*-5P sequences, including five ingroup Operational Taxonomic Units (OTUs). It intentionally includes four *COI*-like pseudogene sequences to illustrate the misleading effect of the *COI*-like pseudogene. This dataset is characterized by 681 bp in total length, 624.4 bp in average length, 47.1% of GC content, 136 bp of variable sites and 96 bp of parsimony informative sites. The Maximum Likelihood phylogenetic tree (Fig. 4) indicates that each sampled OTU and the *COI*-like pseudogene form strongly-supported monophyletic groups (SH-aLRT > 80% and UFBoot > 95%). The intrageneric relationships of sampled *Carnus* are resolved as (*C. hemapterus*, ((*Carnus* sp. CA1, *Carnus* sp. CA2), (*C. orientalis*, (*Carnus* sp. BD, *C. orientalis* pseudogene)))). The oriental samples (*C. orientalis*, *Carnus* sp. BD) plus the *COI*-like pseudogene of *C. orientalis* form a strongly-supported monophyletic group (SH-aLRT/UFBoot = 98/96).



For molecular-based species delimitations, both ASAP and bPTP results are in congruence, identifying five *Carnus* OTUs as independent species and treating the *COI*-like pseudogene as a separate one (Fig. 4). The results are congruent with the monophyly of each OTU in *COI*-based phylogenetic tree (Fig. 4). K2P-distances of *COI* amongst the sampled OTUs and the *COI*-like pseudogene are mostly larger than 0.04, but only that of *C. orientalis* and *Carnus* sp. BD is 0.0247 (Table 3), which is lower than the interspecific threshold distance (0.035) applied in many insect groups (Hebert et al. 2004).

## Discussion

The current discovery of *C. orientalis* in Taiwan aligns with the expectation proposed by Maa (1968) that this species could be widespread in the Oriental Region. In addition to filling the distribution gap in Taiwan, *Carnus* sp. BD from Bangladesh, which exhibits a low *COI* genetic distance (0.0247) with *C. orientalis*, indicates the potential extension of the distribution of *C. orientalis* to the Indian subcontinent. The newly-discovered host record, *O. lettia*, and previous host records, support the host specificity of *C. orientalis* to owls. Providing *COI* barcodes could contribute to the future study of *Carnus*, particularly in the Oriental Region. Our encounter with the interference of host blood meals and pseudogenes when using universal primer sets may elucidate the lack of barcode information in *Carnus*. Identifying pseudogenes and the introduction of newly-designed primers could serve as potential strategies to address similar issues.

Previous records have reported that dealated *C. hemapterus* adults can frequently infest 2–3 weeks old downy-haired nestlings (Bequaert 1942, Capelle and Whitworth 1973, Kirkpatrick and Colvin 1989, Iwasa et al. 2008). Similarly, *C. orientalis* has only been reported on the juveniles of fish owls (Maa 1968) and nestlings of Ryukyu scops owls aged 12–14 days old (Iwasa et al. 2014). Since they are less frequently reported from brooding birds, *Carnus* is more likely to disperse independently rather than being harboured by avian hosts (Bequaert 1942, Capelle and Whitworth 1973, Grimaldi 1997, Iwasa et al. 2014). The low collecting frequency of *Carnus* from the adult birds may also result in the underestimated distribution and biodiversity since surveys of ectoparasites from wild avian hosts mainly rely on adult birds (e.g. Hamstra and Badyaev (2009)). *Otus lettia* inhabits low-altitude environments, which highly overlap with urban areas in Taiwan (König et al. 1999, Cheng 2004). Despite *O. lettia* is not uncommon in Taiwan, this case provides a rare case to examine the infestation of *Carnus* on a fallen *O. lettia* fledgling.

In this study, amplifying host DNA from the contamination of blood meals by the universal primer set accidentally provides a chance to speculate on the host use, as that has been applied in previous studies (Alcaide et al. 2009, Reeves et al. 2018). The blood meals in the fly's gut also support the trophic relationship between our samples and the collared scops owl. However, the mixed DNA samples also pose a challenge for DNA barcoding and increase the cost by changing or redesigning primers.

Challenges in DNA barcoding are not only caused by the host blood contamination, but also by a pseudogene. Mitochondrial pseudogenes have been identified in several insect species (e.g. Bensasson et al. (2001), Pamilo et al. (2007), Hazkani-Covo et al. (2010)) and can mislead the results of *COI*-based species delimitation due to the similarity in nucleotide sequences (Song et al. 2008, Leite 2012). This study's results of molecular phylogeny and DNA-based species delimitation considered the *COI*-like pseudogene as a closed species of the actual *COI*-5P sequences of *C. orientalis*. Such a result could lead to confusion in the comparison of DNA sequences. Pseudogenes can be characterised by several properties, including double peaks in chromatograms, additional stop codons, different evolutionary rates and radical differences in expected or contradictory topology (Leite 2012). In this study, two abnormal deletions were found in the DNA sequences. Moreover, amino acid sequences translated from the aligned *COI*-like pseudogene sequences are distinct from all other examined *COI*-5P sequences and a stop codon was found. As pseudogenes often vary in function as they are not necessarily encoded for proteins (Sorenson and Quinn 1998), the deletions and stop codons could be randomly evolved and not necessarily found amongst all the pseudogenes. On the contrary, the contradictory topologies in phylogenetic trees reconstructed by DNA and amino acids are more likely to be a common property shared by pseudogenes. Therefore, checking the translated amino acid sequences might be necessary to reduce the risk caused by pseudogenes in DNA barcoding.

Subtle morphological variations were observed between the populations in Taiwan and the Ryukyu Islands. However, since no distinct differences were found in the male genitalia—the crucial diagnostic character in this genus—and given the expected broad host range of this species, we regard these variations as intraspecific. To understand the species boundaries of *Carnus* and uncover potential cryptic diversity, broader sampling and comparison of *COI* barcodes, along with other molecular markers, are necessary to facilitate the taxonomy of these small parasitic flies.

## Acknowledgements

We want to thank Shih-Tsai Yang (Animal and Plant Health Inspection Agency, Ministry of Agriculture, Executive Yuan, Taiwan) for supporting the identification; Siou-Huei Chen for providing the video of the living *Carnus orientalis*; and Chao-Sheng Liao (WildOne Wildlife Conservation Association) for providing the information regarding the specimens. This work was supported by grants from the National Science and Technology Council (NSTC 112-2313-B-002-056 to Ming-Chung Chiu; NSTC 111-2621-B-002-002 to Shih-Feng Shiao).

## Author contributions

HPC conceptualised and designed the study, conducted morphological observations and descriptions, performed molecular experiments and analyses and drafted the manuscript. FTC contributed to specimen sampling and provided sample information. SFS

contributed to manuscript drafting and provided funding and resources for the study. MCC co-conceptualised and designed the study, conducted data analysis, contributed to manuscript drafting and provided funding and resources for the study. All authors have read and approved the final manuscript.

## References

- Alcaide M, Rico C, Ruiz S, Soriguer R, Muñoz J, Figuerola J (2009) Disentangling vector-borne transmission networks: A universal DNA barcoding method to identify vertebrate hosts from Arthropod bloodmeals. PLOS One 4 (9): e7092. <https://doi.org/10.1371/journal.pone.0007092>
- Bensasson D, Zhang X, Hartl DL, Hewitt GM (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology & Evolution 16: 314-321. [https://doi.org/10.1016/S0169-5347\(01\)02151-6](https://doi.org/10.1016/S0169-5347(01)02151-6)
- Bequaert J (1942) *Carnus hemapterus* Nitzsch, an ectoparasitic fly of birds, new to America (Diptera). Bulletin of the Brooklyn Entomological Society 37: 140-149.
- Brake I (2000) Phylogenetic systematics of the Milichiidae (Diptera, Schizophora). Entomologica Scandinavica, Supplements 57: 1-120.
- Brake I (2011) World catalog of the family Carnidae (Diptera, Schizophora). Myia 12 (113): e169.
- Buck M (2006) A new family and genus of acalypterate flies from the Neotropical region, with a phylogenetic analysis of Carnoidea family relationships (Diptera, Schizophora). Systematic Entomology 31: 377-404. <https://doi.org/10.1111/j.1365-3113.2006.00328.x>
- Capelle KJ, Whitworth TL (1973) The distribution and avian hosts of *Carnus hemapterus* (Diptera: Milichiidae) in North America. Journal of Medical Entomology 10: 525-526. <https://doi.org/10.1093/jmedent/10.5.525>
- Cheng YR (2004) Observations of breeding of the collared scops owl (*Otus lettia*) in winter on the campus of National Taiwan University, Taipei. Raptor Research of Taiwan 3: 1-7. <https://doi.org/10.30121/RRT.200409.0001>
- Cruaud A, Nidelet S, Arnal P, Weber A, Fusu L, Gumovsky A, Huber J, Polaszek A, Rasplus J (2019) Optimized DNA extraction and library preparation for minute arthropods: Application to target enrichment in chalcid wasps used for biocontrol. Molecular Ecology Resources 19: 702-710. <https://doi.org/10.1111/1755-0998.13006>
- DeConinck E (1986) Consideration of the host-parasite relationship and systematics of the first African *Carnus*. In: Darvas B, Papp L (Eds) Abstracts of the First International Congress of Dipterology, Budapest, 17-24 Aug., 1986. the First International Congress of Dipterology, Budapest. Hungarian Academy of Sciences, Budapest
- Folmer O, Black M, W H, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3 (5): 294-299.
- Grimaldi D (1997) The bird flies, genus *Carnus*: Species revision, generic relationships, and a fossil *Meoneura* in amber (Diptera: Carnidae). American Museum Novitates 3190: 1-30.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the

- performance of PhyML 3.0. *Systematic Biology* 59 (3): 307-321. <https://doi.org/10.1093/sysbio/syq010>
- Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PD (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences* 103: 968-971. <https://doi.org/10.1073/pnas.0510466103>
  - Hamstra TL, Badyaev AV (2009) Comprehensive investigation of ectoparasite community and abundance across life history stages of avian host. *Journal of Zoology* 278: 91-99. <https://doi.org/10.1111/j.1469-7998.2008.00547.x>
  - Harrison PM, Milburn D, Zhang Z, Bertone P, Gerstein M (2003) Identification of pseudogenes in the *Drosophila melanogaster* genome. *Nucleic Acids Research* 31: 1033-1037. <https://doi.org/10.1093/nar/gkg169>
  - Hazkani-Covo E, Zeller RM, Martin W (2010) Molecular poltergeists: Mitochondrial DNA copies in sequenced nuclear genomes. *PLoS Genetics* 6: 1-11. <https://doi.org/10.1371/journal.pgen.1000834>
  - Hebert PD, Ratnasingham S, Dewaard JR (2003a) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B* 270: S96-S99. <https://doi.org/10.1098/rsbl.2003.0025>
  - Hebert PD, Cywinska A, Ball SL, Dewaard JR (2003b) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* 270: 313-321. <https://doi.org/10.1098/rspb.2002.2218>
  - Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences* 101: 14812-14817. <https://doi.org/10.1073/pnas.0406166101>
  - Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2017) UFboot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* 35 (2): 518-522. <https://doi.org/10.1093/molbev/msx281>
  - Iwasa M, Mori S, Furuta O, Komatsu T, Iida T, Nakamori J, Kataoka N (2008) A bird-parasitic fly, *Carnus hemapterus* Nitzsch (Diptera, Carnidae) in Japan: avian hosts, infestations, and a case of human dermatitis caused by adult. *Medical Entomology and Zoology* 59: 47-53. <https://doi.org/10.7601/mez.59.47>
  - Iwasa M, Sakamoto H, Asahi K (2014) Discovery of a bird-parasitic fly, *Carnus orientalis* (Diptera: Carnidae), in Japan, with bionomic remarks and a key to *Carnus* species. *Journal of Medical Entomology* 51 (2): 484-488. <https://doi.org/10.1603/ME13087>
  - Janzen DH, Hallwachs W, Blandin P, Burns JM, et al. (2009) Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Molecular Ecology Resources* 9: 1-26. <https://doi.org/10.1111/j.1755-0998.2009.02628.x>
  - Kirkpatrick CE, Colvin BA (1989) Ectoparasitic fly *Carnus hemapterus* in a nesting population of common barn-owls (Strigiformes: Tytonidae). *Journal of Medical Entomology* 26: 109-112. <https://doi.org/10.1093/jmedent/26.2.109>
  - König C, Weick F, Becking JH (1999) *Owls: a guide to the owls of the world*. Yale University Press, Mountfield, 462 pp.
  - Leite LA (2012) Mitochondrial pseudogenes in insect DNA barcoding: Differing points of view on the same issue. *Biota Neotropica* 12: 301-308. <https://doi.org/10.1590/S1676-06032012000300029>
  - Maa TC (1968) A new *Carnus* from Malaya (Diptera: Milichiidae). *Pacific Insects* 10: 33-36.

- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32 (1): 268-274. <https://doi.org/10.1093/molbev/msu300>
- Pamilo P, Viljakainen L, Vihavainen A (2007) Exceptionally high density of NUMTs in the honeybee genome. *Molecular Biology and Evolution* 24: 1340-1346. <https://doi.org/10.1093/molbev/msm055>
- Puillandre N, Brouillet S, Achaz G (2020) ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources* 21 (2): 609-620. <https://doi.org/10.1111/1755-0998.13281>
- Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes* 7: 355-364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x>
- Reeves LE, Gillett-Kaufman JL, Kawahara AY, Kaufman PE (2018) Barcoding blood meals: New vertebrate-specific primer sets for assigning taxonomic identities to host DNA from mosquito blood meals. *PLOS Neglected Tropical Diseases* 12 (8): e0006767. <https://doi.org/10.1371/journal.pntd.0006767>
- Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences* 105 (36): 13486-13491. <https://doi.org/10.1073/pnas.0803076105>
- Sorenson MD, Quinn TW (1998) Numts: a challenge for avian systematics and population biology. *The Auk* 115: 214-221. <https://doi.org/10.2307/4089130>
- Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* 38 (7): 3022-3027. <https://doi.org/10.1093/molbev/msab120>
- Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* 44 (W1): W232-W235. <https://doi.org/10.1093/nar/gkw256>
- Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29 (22): 2869-2876. <https://doi.org/10.1093/bioinformatics/btt499>

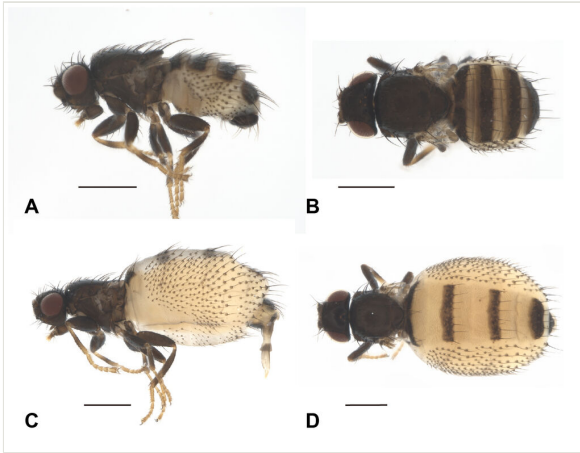


Figure 1.

Habitus of *Carnus orientalis* from Taiwan. **A** male in lateral view; **B** ditto in dorsal view; **C** female in lateral view; **D** ditto in dorsal view. Scale = 0.5 mm. Photographed by Hsuan-Pu Chen.

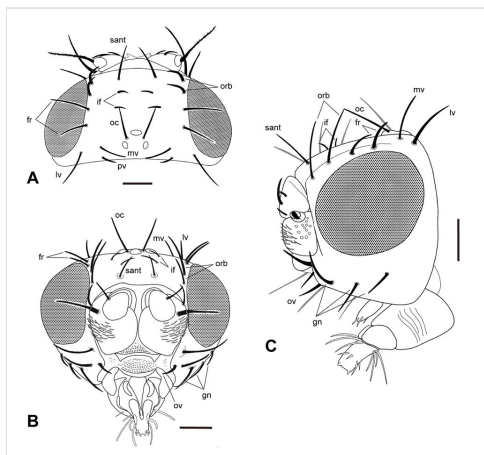


Figure 2.

Head of *Carnus orientalis* from Taiwan. **A** head in dorsal view; **B** anterior view; **C** lateral view. Abbreviations: **fr**, frontal seta; **gn**, genal seta; **if**, inner frontal seta; **lv**, lateral vertical seta; **mv**, medial vertical seta; **oc**, ocellar seta; **orb**, orbital seta; **ov**, oral vibrissa; **sant**, suprantennal seta. Scale = 0.1 mm. Illustrated by Hsuan-Pu Chen.

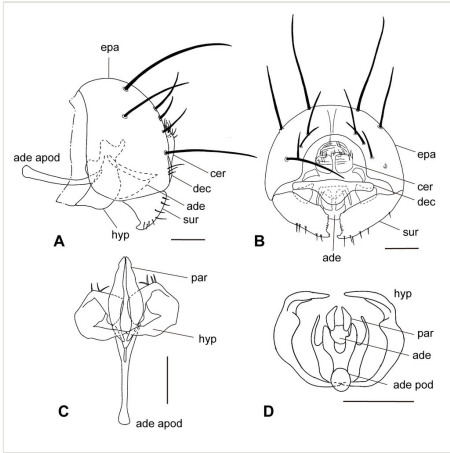


Figure 3.

Male genitalia of *Carnus orientalis* from Taiwan. **A** male genitalia in lateral view; **B** ditto in posterior view; **C** hypandrium and phallic complex in ventral view; **D** ditto in anterior view. Abbreviations: **ade**, aedeagus; **ade apod**, aedeagal apodeme; **cer**, cercus; **dec**, decasternum; **epa**, epandrium; **hyp**, hypandrium; **par**, paraphysis; **sur**, surstylus. Scale = 0.05 mm. Illustrated by Hsuan-Pu Chen.



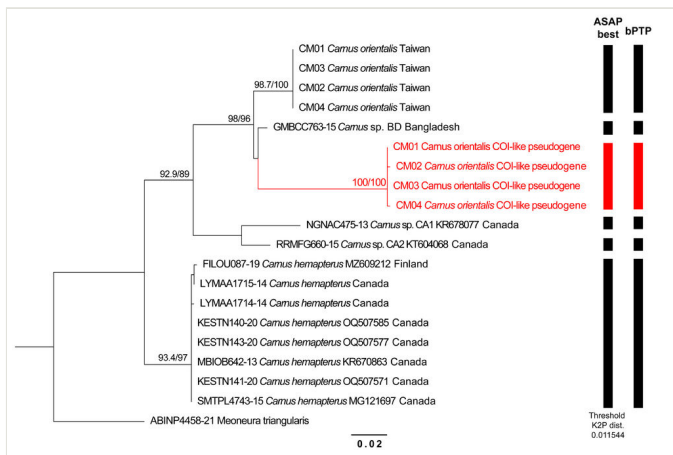


Figure 4.

Maximum-Likelihood phylogenetic tree of the *Camus* species, based on the *COI* and *COI*-like pseudogene dataset (681 bp, GTR+F+I) with species delimitation analyses (vertical bars on the right side of the tree). Branch lengths of the phylogenetic tree are proportional to the inferred nucleotide substitutions. Node numbers represent 'SH-aLRT/UFBoot' values in percent (%). The red clade indicates the *COI*-like pseudogene.

Table 1.

PCR primers used in this study.

PCR primers						
Targets	Primers	Sequences (5'-3')			References	
COI	LCO1490 (F)	GGTCAACAATCATAAAGATATTGG			Folmer et al. (1994)	
	HCO2198 (R)	TAAACTTCAGGGTGACCAAAAAATCA				
	C1-J-1571_Carnus (F)	GGAATAGTTGGAACCTCYTTAAGAATTC			This study	
	COIF_Carnus (F)	TCCACCTTCTTTAACACTTTTA				
	COIR_Carnus (R)	AATTACAGCTCCTAAAAT				
COI-like pseudogene	COIF_pseudo_Carnus (F)	AACTGTATATCCACCTCTATCA			This study	
	COIR_pseudo_Carnus (R)	GAAAACTAGCTAAATCAAG				
PCR conditions amplifying for 35 cycles						
Primer sets (F/R)	Primary denaturation	Denaturation	Annealing	Extension	Additional extension	Products
LCO1490/HCO2198	95°C, 4 min	95°C, 30 sec	47°C, 1 min	72°C, 45 sec	72°C, 10 mins	Bird COI-5P ca. 680 bp
LCO1490/COIR_Carnus			46°C, 1 min			Carnus COI-5P ca. 440 bp
C1-J-1571/COIR_Carnus						Carnus COI-5P ca. 385 bp
COIF_Carnus/HCO2198			48°C, 1 min			Carnus COI-5P ca. 390 bp
C1-J-1571_Carnus/HCO2198	95°C, 4 min	95°C, 30 sec	47°C, 1 min	72°C, 45 sec	72°C, 10 mins	Carnus COI-like pseudogene ca. 610 bp
C1-J-1571_Carnus/COIR_pseudo_Carnus			48°C, 1 min			Carnus COI-like pseudogene ca. 340 bp
COIF_pseudo_Carnus/HCO2198						Carnus COI-like pseudogene ca. 325 bp

Table 2.

Sample information of the *COI* and *COI*-like pseudogene sequences included in this study.

Taxon name	Locality	Voucher	Accession number	Source	Region
<i>Otus lettia</i>	China	PB2019-211-1	MW364567	GenBank	complete gene
<i>Phyllomyza</i> sp.	China	none	OP612805	GenBank	complete gene
<i>Drosophila melanogaster</i>	USA	SAMN02803731	NC_024511	GenBank	complete gene
<i>Meoneura triangularis</i>	Canada	BIOUG66093-G12	ABINP4458-21	BOLD systems	<i>COI</i> -5P
<i>Camus hemapterus</i>	Canada	BIOUG08418-E02	MBIOB642-13	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	CCDB-21327-A04	LYMAA1714-14	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	BIOUG21954-H08	SMTPL4743-15	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	BIOUG62752-E07	KESTN140-20	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	BIOUG62752-E08	KESTN141-20	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	BIOUG62752-E10	KESTN143-20	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Finland	PP0087	FILOU087-19	BOLD systems	<i>COI</i> -5P
<i>C. hemapterus</i>	Canada	CCDB-21327-A05	LYMAA1715-14	BOLD systems	<i>COI</i> -5P
<i>Camus</i> sp. BD	Bangladesh	BIOUG22253-G07	GMBCC763-15	BOLD systems	<i>COI</i> -5P
<i>Camus</i> sp. CA1	Canada	BIOUG07296-C02	NGNAC475-13	BOLD systems	<i>COI</i> -5P
<i>Camus</i> sp. CA2	Canada	BIOUG22722-A04	RRMFG660-15	BOLD systems	<i>COI</i> -5P
<i>C. orientalis</i>	Taiwan	CM01	PP192111	This study	<i>COI</i> -5P
<i>C. orientalis</i>	Taiwan	CM02	PP192112	This study	<i>COI</i> -5P
<i>C. orientalis</i>	Taiwan	CM03	PP192113	This study	<i>COI</i> -5P
<i>C. orientalis</i>	Taiwan	CM04	PP192114	This study	<i>COI</i> -5P
<i>C. orientalis</i>	Taiwan	CM01	PP199482	This study	<i>COI</i> -like pseudogene
<i>C. orientalis</i>	Taiwan	CM02	PP199483	This study	<i>COI</i> -like pseudogene
<i>C. orientalis</i>	Taiwan	CM03	PP199484	This study	<i>COI</i> -like pseudogene

<i>C. orientalis</i>	Taiwan	CM04	PP199485	This study	<i>COI</i> -like pseudogene
<i>O. lettia</i> (bloodmeal)	Taiwan	CM02	PP188559	This study	<i>COI</i> -5P
<i>O. lettia</i> (bloodmeal)	Taiwan	CM03	PP188560	This study	<i>COI</i> -5P

Table 3.

Mean pairwise K2P distances and estimated variance of *COI* sequences of *Carnus* species. The values in the lower left of the matrix represent the K2P distances and the upper right represents the variances estimated by the Bootstrap method in 100 replications

	<i>Carnus</i> sp. BD	<i>C.</i> <i>hemapterus</i>	<i>Carnus</i> sp. CA1	<i>Carnus</i> sp. CA2	<i>C.</i> <i>orientalis</i>	<i>C. orientalis</i> pseudogene
<i>Carnus</i> sp. BD		0.0112	0.0116	0.0112	0.0064	0.0121
<i>C. hemapterus</i>	0.0725		0.0117	0.0104	0.0115	0.0131
<i>Carnus</i> sp. CA1	0.0842	0.0722			0.0124	0.0167
<i>Carnus</i> sp. CA2	0.0699	0.0816	0.0447		0.0113	0.0140
<i>C. orientalis</i>	0.0247	0.0888	0.0903	0.0771		0.0126
<i>C. orientalis</i> pseudogene	0.0706	0.1046	0.1248	0.1118	0.0837	

## Supplementary materials

### Suppl. material 1: The method of measurement used in this study

**Authors:** Hsuan-Pu Chen

**Data type:** images

**Brief description:** The supplementary figure illustrates the measurement methods for *Carnus* used in this study. The captions for the supplementary figure are as follows: A head in dorsal view; B head in lateral view; C habitus in lateral view; D habitus in dorsal view; E deca sternum (upper) and surstylus (lower).

[Download file](#) (653.60 kb)

### Suppl. material 2: The relative positions of the primers designed and used in this study

**Authors:** Hsuan-Pu Chen

**Data type:** images

**Brief description:** The supplementary figure illustrates the relative binding sites of the used and newly-designed primers in this study.

[Download file](#) (212.12 kb)

### Suppl. material 3: Compressed Fasta files containing the *COI* datasets analysed in this study

**Authors:** Hsuan-Pu Chen

**Data type:** sequences

**Brief description:** The supplementary compressed file includes two Fasta files of the *COI* datasets analysed in this study: one with and one without the reference full-length *COI* sequence of *Drosophila melanogaster* for position numbering.

[Download file](#) (3.08 kb)