BIODIVERSITAS
Volume 24, Number 4, April 2023
Pages: 1990-1998

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d240407

Genetic diversity and phylogenetic analysis of mayfly *Caenis* (Insecta: Ephemeroptera) using Cytochrome C Oxidase I (COI) and 12s rRNA genes from Thailand

NARUMON PRAKRONGRAK[1], BOONSATIEN BOONSOONG[2], MONTHIRA MONTHATONG[1,♥]

[1]Department of Biology, Faculty of Sciences, Khon Kaen University. 123 Mittraparb Road, Nai-Meuang, Meuang, Khon Kaen, 40002, Thailand. Tel.: +66-43-202531, ♥email: monmon@kku.ac.th

[2]Department of Zoology, Faculty of Sciences, Kasetsart University. 50 Ngam Wong Wan Road, Chatuchak, Bangkok 10900, Thailand

Manuscript received: 30 November 2022. Revision accepted: 4 April 2023.

Abstract. *Prakrongrak N, Boonsoong B, Monthatong M. 2023. Genetic diversity and phylogenetic analysis of mayfly* Caenis *(Insecta: Ephemeroptera) using Cytochrome C Oxidase I (COI) and 12s rRNA genes from Thailand. Biodiversitas 24: 1989-1997.* Mayflies in genus *Caenis* is one of the top fifth species richness belonged to Family Caenidae. The present study aimed to use the partial mitochondrial Cytochrome C Oxidase subunit I (COI) and 12s rRNA nucleotide sequences for molecular species identification, diversity and phylogenetic relationships of mayflies in genus *Caenis* Stephens, 1835 species from Thailand. From a total of 37 specimens, thirteen nymphs were morphologically identified into five species: *Caenis nasuta* Malzacher, *C. Picea* Kimmins, *C. ulmeriana* Malzacher, *C. cornigera* Kang and Yang and *C. longiforcipata* Malzacher. The other 24 samples were *Caenis* sp.1 to sp.5. Partial COI sequences of all samples were compared with BOLD and GenBank for species identification. However, the result showed only genus-level identification as *Caenis* with greater than 80% similarity. For species delimitation, interspecific genetic distance among these species ranged from 14.4 to 26.6% (COI) and from 7.36 to 22.4% (12s rRNA). Intraspecific divergence levels were 0.0% to 2.35%. ABGD analysis of both genes divided data into 10 groups corresponding to 10 morphospecies of *Caenis*. The phylogenetic relationships of *Caenis*, COI and 12s rRNA genes also classified each species to well supported clusters. In addition, we report *C. cornigera* from Thailand for the first time. The COI and 12s rRNA phylogenetic trees indicate a close relationship between *C. cornigera* and *C. longiforcipata*, which is supported by their similar morphology.

Keywords: 12s rRNA, COI, genetic diversity, oriental region

INTRODUCTION

Mayflies are aquatic insects that are classified as Order Ephemeroptera (McCafferty 1981). This insect group is an archaic lineage that was present at the end of the Paleozoic during the late Carboniferous or early Permian (Misof et al. 2014; Sartori and Brittain 2015). However, the Cretaceous fossil mayflies have been increasingly documented from all continents except Antarctica, due to their abundance in the period (Brandão et al. 2021). Since they are hemimetabolous insects, their life cycle undergoes incomplete metamorphosis and consists of a well-defined aquatic nymphal phase and two terrestrial phases (Almudi et al. 2019). Mayflies can be found in both lotic and lentic ecosystems of freshwater habitats all over the world, and in some cases brackish water; however, they are more abundant species-rich in the tropics and warm temperate regions. Nymphal stages develop in water containing higher levels of dissolved oxygen; therefore, they are important for freshwater ecology and qualified as bioindicators for pollution and water quality monitoring (Alhejoj et al. 2014; Scarduelli et al. 2017). Currently, approximately 40 families and more than 460 genera of mayflies have already been recognized worldwide (Jacobus et al. 2019). The small squaregilled or Caenid mayflies (Caenidae) consist of 26 genera and 221 species (Sartori and Brittain 2015). Among these, the genus *Caenis* Stephens, 1835 is cosmopolitan and one of the most diversified genera with approximately 170 species (Barber-James et al. 2013; Malzacher 2015). In 1853, Walker firstly studied caenid mayflies in the Oriental realm and reported *Caenis perpusilla* from Sri Lanka (Walker 1853). Later, Soldán (1978) described new genera and species of Caenidae from India. Kang and Yang (1994) reported the nymphal stages of seven new species of Caenidae in Taiwan. Since 2013, thirty-tree new *Caenis* species have been described from Thailand, Indonesia, India and Vietnam (Malzacher 2013; Malzacher 2015; Malzacher and Sangpradub 2017; Muthukatturaja and Balasubramanian 2020; Malzacher 2021; Malzacher and Sangpradub 2021; Muthukatturaja and Balasubramanian 2021; Srinivasan et al. 2021).

The study of *Caenis* diversity in Thailand began when Van Bruggen (1954) first reported *Caenis demoulini* from Bangkok, Thailand and described the female characters. In 2015, Malzacher revised the Oriental species of the genus *Caenis* and found four new species from Thailand: *C. nigropunctatula*, *C. ulmeriana*, *C. guttata* and *C. gephyria*. In 2021, Malzacher and Sangpradub found six new species: *C. ludovici*, *C. nasuta*, *C. longiforcipata*, *C. karenae*, *C. obtusostilata*, and *C. acutostilata*. Therefore; forty-three species of *Caenis* have been found in the Oriental region and twelve of them have been reported in Thailand.

However, Identification of some *Caenis* nymphs to species level is difficult since their morphology is diverse. Mostly, male adult characters are used for species identification in *Caenis*.

Recently, several molecular genetic tools including DNA barcoding have been studied which may potentially improve the taxonomic discrimination, identification, and genetic variation analyses of aquatic insects (Sivaramakrishnan et al. 2014). Mitochondrial DNA (mtDNA) is the most extensively investigated genomic system in insects (Cameron 2014). It has been used for a wide array of research goals including molecular systematics, population genetics, genetic variation and phylogenetic relationships (Hwang et al. 2013). In addition, mtDNA is a potentially useful genetic marker to resolve lower taxonomic levels for organisms (Allio et al. 2017). The mitochondrial Cytochrome C Oxidase subunit I (COI) gene was adopted as a standard for DNA barcoding of animals by the Consortium for the Barcode of Life (CBOL). The COI gene has been an important tool for animal species identification, delimitation and revealing cryptic diversity of organisms. The benefits of COI include its variation that empower species-level discrimination, multiple copies for Polymerase Chain Reaction (PCR) amplification from an enormous range of species and millions of DNA sequences on associated databases (Deagle et al. 2014). Another advantage of using the COI gene comes from its capacity for taxonomic identification of all life stages of organisms (Hwang et al. 2013; Suh et al. 2019). In particular, aquatic insect identification mostly relies on the morphological characteristics of adult males, but most ecological research has been performed on the nymphal stages that are difficult to identify species (Suh et al. 2019). Various studies on Ephemeroptera have used COI in order to confirm new species and support morphological data in phylogenetic analyses (Selvakumar et al. 2016; Boonsoong and Satori 2019; Boonsoong et al. 2021; Takayanagi and Yoshizawa 2021). The mitochondrial 12S and 16S rRNA genes provide a potential for molecular systematics of many animals (Yang et al. 2014), albeit not widely for Ephemeroptera. The complete mitochondrial genomes of *Caenis* sp. from China were sequenced for investigating the phylogenetic relationships among Paleopterous insects and between Caenidae related families (Cai et al. 2018; Xu et al. 2020). The circular mitochondrial genome of *Caenis* sp. contains 37 genes coding for two rRNAs, 22 tRNAs and 13 polypeptides. This study evaluated the effectiveness of two mitochondrial genes, COI and 12s rRNA, in identifying species, assessing genetic diversity, and determining phylogenetic relationships among *Caenis* nymphs collected in Thailand.

MATERIALS AND METHODS

Specimen collection and identification

Caenis larvae were collected by a 500 micron D-frame dip net from streams, rivers, ponds, and lakes in Northern (N), North-Eastern (NE), Eastern (E), Western (W), Central (C) and Southern (S) Thailand during 2017 to 2020 (listed in Table 1). Specimens were stored in 95% ethanol. The morphological structures of *Caenis* larvae were identified to the lowest taxonomic level under the stereomicroscope, light microscope and Scanning Electron Microscope (SEM) using suitable taxonomical literature of *Caenis* in the last three decades (Kang and Yang 1994, 1996; Tong and Dudgeon 2002; Sangpradub and Boonsoong 2006; Jia et al. 2010; Malzacher 2013, 2015; Malzacher and Sangpradub 2021; Muthukatturaja and Balasubramanian 2021; Srinivasan et al. 2021).

DNA extraction, amplification, and sequencing

Total DNA of each specimen was extracted using the PureLink™ Genomic DNA Mini Kit (Invitrogen, Carlsbad, CA, USA), following the manufacturer instructions. When whole larvae were obtained, legs or a very small quantity of thoracic tissue was removed, and the remainder of the specimen was preserved in 95% ethanol (a voucher). Partial sequences of the mitochondrial Cytochrome C Oxidase (COI) and 12S ribosomal RNA genes were amplified by PCR. The COI gene was amplified using the universal primers LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') (Folmer et al. 1994). The universal primers 12Sai (5'-AAA CTA CGA TTA GAT ACC CTA TTA T-3') and 12Sbi (5'-AAG AGC GCG GGC GAT GTG T-3') (Simon et al. 1994) were used to amplify the DNA fragment of 12s rRNA gene. Each 50 µL PCR mixture was composed of 1x ViBuffer A (500mM KCl, 100mM Tris-HCl pH 9.1, 0.1% Triton TMX-100), 2.5 mM $MgCl_2$, 0.2 µM of each primer, 0.2 mM dNTPs, 2 units of *Pfu* DNA polymerase and 10-50 ng DNA template. PCR assays were conducted using the Prima-96™ thermal cycler (Himedia, India). The PCR conditions were as follows: one cycle of an initial denaturation at 94°C for 5 min; 35 cycles of denaturation at 94°C for 30 sec, annealing at 48°C or 50°C (COI or 12s rRNA) for 30 sec, and elongation at 72°C for 60 sec; and a final extension step at 72°C for 5 min. All PCR amplicons were verified by 1.5% agarose gel stained with FluoroVue™ (SMOBIO, Taiwan). The PCR products were purified and sequenced on ABI 3730xl DNA Analyzer (Macrogen Inc, Korea).

Data analysis

The raw sequences of COI and 12s rRNA were examined and corrected with BioEdit version 7.2.5 (Hall 1999). The COI sequences were compared with the GenBank nucleotide database by Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) and Barcode of Life Data Systems (BOLD) (Ratnasingham and Hebert 2007; 2013). Sequences were aligned using CLUSTALW via MEGA version 10 (Kumar et al. 2018) and Indel (insertion/deletion) regions were removed using SeqFIRE (Ajawatanawong et al. 2012). Intraspecific and interspecific sequence divergences based on Kimura-2-Parameter (K2P) distances were calculated using MEGA version 10. Additionally, the Automatic Barcode Gap Discovery (ABGD) method was used to delimit species by detecting a significant gap in the pairwise distance distribution generated via https://bioinfo.mnhn.fr/abi/public/abgd/

(Puillandre et al. 2012). Genetic diversity and phylogenetic relationships were obtained by Maximum Likelihood (ML) analysis using MEGA version 10 and bootstrap analyses of 1,000 replicates. *Potamanthellus edmundsi* Ulmer, 1939 (Neoephemeridae) sequences were used as the outgroup since it belongs to the Superfamily Caenoidea as does *Caenis* spp. (Caenidae). The COI and 12sRNA sequence data for all samples were deposited in GenBank.

Table 1. Profile of specimen collections examined in this study

Species	Collection locality (region)	Latitude	Longitude	Sample ID		GenBank accession no.	
				COI	12s rRNA	COI	12s rRNA
Caenis nasuta	Klity Creek, Kanchanaburi (W)	14°52'36.3"N	98°57'00.9"E	C21	R21	OQ476166	OQ244322
C. nasuta	Ban Song Khon stream, Loei (NE)	17°21' 21.70"N	101° 24' 37.50"E	C102	R102	OQ476167	OQ244323
C. picea	Peing Din stream, Loei (NE)	17°3' 9520"N	101°44' 8440"E	C103	R103	OQ476168	OQ244324
C. picea	Suan Hom stream, Loei (NE)	17° 2' 82.10"N	101°45' 69.90"E	C105	R105	OQ476169	OQ244325
C. ulmeriana	Huay sarong, Prachuap Khiri Khan (S)	12°06'54.3"N	99°38'00.0"E	C59	R59	OQ476173	OQ244332
C. ulmeriana	Huay sarong, Prachuap Khiri Khan (S)	12°06'54.3"N	99°38'00.0"E	C60	R60	OQ476174	OQ244333
C. ulmeriana	Ban NonPattana, Loei (NE)	17°05'54.5"N	101°28'48.9"E	C91	R91	OQ476175	OQ244334
C. ulmeriana	Ban Song Khon stream, Loei (NE)	17°21' 21.70"N	101°24' 37.50"E	C99	R99	OQ476176	OQ244335
C. ulmeriana	Ban Song Khon stream, Loei (NE)	17°21' 21.70"N	101°24' 37.50"E	C100	R100	OQ476177	OQ244336
C. cornigera	Klity Creek, Kanchanaburi (W)	14°52'36.3"N	98°57'00.9"E	C20	R20	OQ476196	OQ244355
C. cornigera	Klity Creek, Kanchanaburi (W)	14°52'36.3"N	98°57'00.9"E	C22	R22	OQ476197	OQ244356
C. longiforcipata	Pha Ta Baek, Nakhon Ratchasima (NE)	14°21'51.1"N	101°20'44.2"E	C95	R95	OQ476198	OQ244358
C. longiforcipata	Ched khot stream, Saraburi (C)	14°27'37.8"N	101°09'33.1"E	C109	R109	OQ476199	OQ244357
Caenis sp.1	Pha Ta Baek, Nakhon Ratchasima (NE)	14°21'51.1"N	101°20'44.2"E	C93	R93	OQ476170	OQ244326
Caenis sp.1	Pha Ta Baek, Nakhon Ratchasima (NE)	14°21'51.1"N	101°20'44.2"E	C94	R94	OQ476171	OQ244327
Caenis sp.1	Ched khot stream, Saraburi (C)	14°27'37.8"N	101°09'33.1"E	C98	R98	OQ476172	OQ244328
Caenis sp.2	Namtok laung waterfall, Ranong (S)	9°51'11.2"N	98°37'14.4"E	C44	R44	OQ476182	OQ244341
Caenis sp.2	Namtok laung waterfall, Ranong (S)	9°51'11.2"N	98°37'14.4"E	C45	R45	OQ476183	OQ244342
Caenis sp.2	Lue Jae Pond, Wang, Narathiwat (S)	5°54'50.2"N	101°52'24.0"E	C72	R72	OQ476184	OQ244343
Caenis sp.2	Lue Jae Pond, Wang, Narathiwat (S)	5°54'50.2"N	101°52'24.0"E	C74	R74	OQ476185	OQ244344
Caenis sp.2	Lue Jae Pond, Wang, Narathiwat (S)	5°54'50.2"N	101°52'24.0"E	C75	R75	OQ476186	OQ244345
Caenis sp.2	Lue Jae Pond, Wang, Narathiwat (S)	5°54'50.2"N	101°52'24.0"E	C76	R76	OQ476187	OQ244346
Caenis sp.3	Lam Chiang Sa, Nakhonratchasima (NE)	14° 29.826'N	101°57.252' E	C10	R10	OQ476188	OQ244347
Caenis sp.3	Bo Khlueng, Ratchaburi (W)	13°31'07.3"N	99°14'46.7"E	C27	R27	OQ476189	OQ244348
Caenis sp.3	Khlong Tong, Prachuap Khiri Khan (S)	11°16'38.3"N	99°25'43.6"E	C46	R46	OQ476190	OQ244349
Caenis sp.3	Khlong Tong, Prachuap Khiri Khan (S)	11°16'38.3"N	99°25'43.6"E	C47	R47	OQ476191	OQ244350
Caenis sp.3	Yang Chum, Prachuap Khiri Khan (S)	12°07'11.0"N	99°37'51.8"E	C61	R61	OQ476192	OQ244351
Caenis sp.3	Bo Khlueng, Ratchaburi (W)	13°31'07.3"N	99°14'46.7"E	C89	R89	OQ476193	OQ244352
Caenis sp.3	Pla Baa steam, Loei (NE)	17°23' 49.40"N	101°22' 22.60"E	C104	R104	OQ476194	OQ244353
Caenis sp.3	Khlong Huay Pee lok stream, Saraburi (C)	14°28'58.3"N	101°09'00.9"E	C107	R107	OQ476195	OQ244354
Caenis sp.4	Nam Thob Ranger station, Loei (NE)	17°15' 35"N	101° 34' 49" E	C14	R14	OQ476200	OQ244329
Caenis sp.4	Nam Thob Ranger station, Loei (NE)	17°15' 35"N	101° 34' 49" E	C17	R17	OQ476201	OQ244330
Caenis sp.4	Nam Thob Ranger station, Loei (NE)	17°15' 35"N	101° 34' 49" E	C35	R35	OQ476202	OQ244331
Caenis sp.5	Khlong Wang, Narathiwat (S)	5°56'10.8"N	101°53'03.3"E	C64	R64	OQ476178	OQ244337
Caenis sp.5	Khlong Wang, Narathiwat (S)	5°56'10.8"N	101°53'03.3"E	C65	R65	OQ476179	OQ244338
Caenis sp.5	Khlong Bala, Narathiwat (S)	5°49'10.6"N	101°51'38.8"E	C68	R68	OQ476180	OQ244339
Caenis sp.5	Khlong Bala, Narathiwat (S)	5°49'10.6"N	101°51'38.8"E	C69	R69	OQ476181	OQ244340
Potamanthellus edmundsi	Khun Korn stream, Chiang Rai (N)	19°51'54.6"N	99°38'54.4"E	-*	R39	MN186576	OQ244359

Note: N: North; NE: North-East; E: East; W: West; C: Central; and S: South. *: Using GenBank sequence accession number: MN186576

RESULTS AND DISCUSSION

Morphological identification

Based on nymphal stage morphological diagnosis; with the exception of *C. ulmeriana* which was identified from nymph and adult male, thirteen *Caenis* specimens were identified into five species which were *Caenis nasuta* Malzacher, 2021 (in Malzacher and Sangpradub 2021), *C. picea* Kimmins, 1947, *C. ulmeriana* Malzacher, 2015, *C. cornigera* Kang & Yang, 1994 and *C. longiforcipata* Malzacher, 2021 (in Malzacher and Sangpradub 2021). The other 24 samples with ambiguous species identification were sorted into featured clusters and given the epithet "sp" including sp.1 to sp.5 (Table 1 and Figure 1). The genus *Caenis* is distributed throughout the world, with 43 species found in the Oriental region, with only 12 species reported in Thailand (Malzacher 2015; Malzacher and Sangpradub 2021). *Caenis* mayflies are typically identified by male imaginal characters such as penis shape, forceps shape, apical tuft of spines on forceps, prosternal triangle, foretarsus, base of the antennal flagellum, scutellum pigmentation, styliger sclerite, lateral sclerite and apophyses (Malzacher 2015; Malzacher and Sangpradub 2021). In this work *Caenis* was studied from the nymphal stage, which limited identification to only five species according to available published documents (Kang and Yang 1994; Malzacher 2015; Malzacher and Sangpradub 2021; Srinivasan 2021).

C. nasuta and *Caenis* sp.1 were very similar in morphology. Both of them have sides of mesonotum with an anterolateral, nose-shaped process densely provided with short spatulate bristles but *Caenis* sp.1 can be distinguished from *C. nasuta* by a forefemur on the dorsal side without a transverse row of bristles. Although, *C. picea* is morphologically similar to *Caenis* sp.4 by having the same tergum I and II each with a cone shaped posteromedian process, *Caenis* sp.4 can be discriminated from *C. picea* by the presence of transverse row of bristles on the forefemur. *Caenis* sp.5 can be distinguished from other species by the following set of characters: pronotum clearly diverging with 5 setae on anterolateral margins, forefemur dorsally with transverse row bristles, sternum IX medially with slightly deep incision and mesonotum with the anterior-lateral nose-shaped process without bristle. The distinctive features of *Caenis* sp.2 collected from southern Thailand are long bristles on the dorsal, lateral and medial margin of the operculate gills, turgum VII with long acute bristles, hind margin of sternum IX rounded and mesonotum without anterolateral, nose-shaped process. *Caenis* sp.3 can be found in almost every region of Thailand. It is characterized by 5-7 bristles on the posterolateral of the operculate gill, hind margin of sternum IX rounded, turgum VII with long bristles and turgum VIII with slightly long bristles. *C. cornigera* can be distinguished from all other *Caenis* species by the following combination of characters: mesonotum with a distinct process near anterolateral corner. Middle and hind femora with stout biforked clavated setae (Kang and Yang 1994). *C. ulmeriana* can be distinguished from all other *Caenis* species by the following combination of characters: Bristles of transverse row on forefemur and dorsal side of mid femur broad, strongly spatulated. Hind claw strongly bowed. Mid-ridge on the anterior operculate gill with 5-6 spatulated bristles in the apical half or two-thirds (Malzacher and Sangpradub 2021). *C. longiforcipata* can be distinguished from all other *Caenis* species by the following combination of characters: Mesonotum with anterolateral, Nose-shaped process without bristles. Operculate gill dorsally along the very flat inner ridge basally with about 4 short blunt bristles and posteriorly with two or three long and thin ones (Malzacher and Sangpradub 2021).

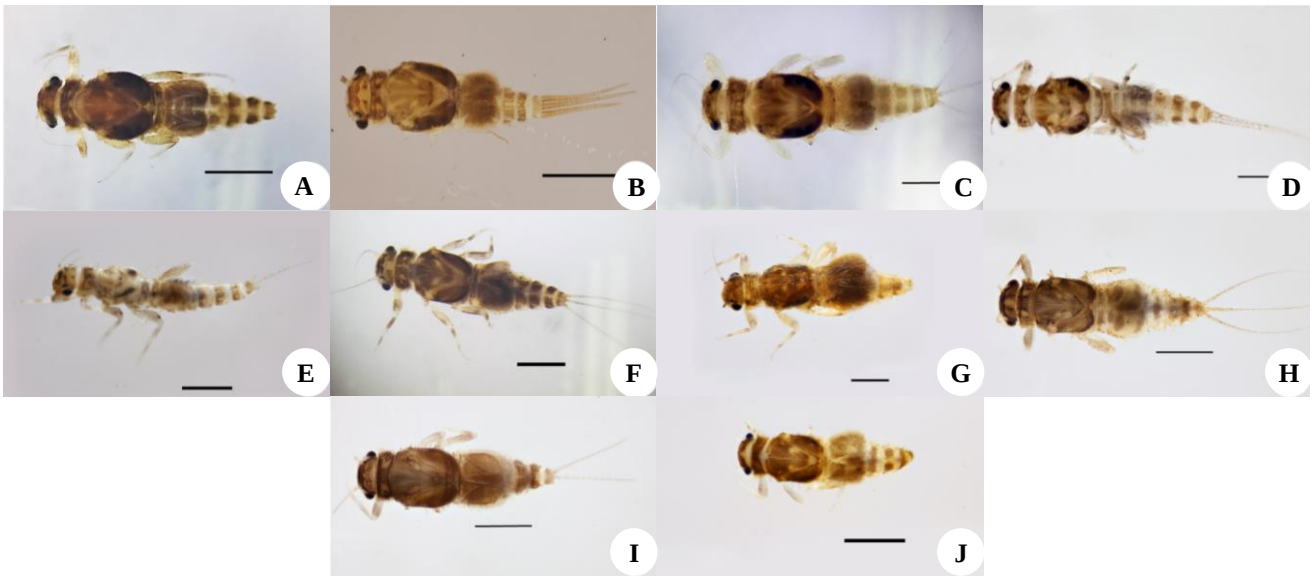

Figure 1. The nymphal characters of *Caenis* spp. in this study. A. *Caenis nasuta*; B. *C. picea*; C. *C. ulmeriana*; D. *C. Cornigera*; E. *C. longiforcipata*; F. *Caenis* sp.1; G. *Caenis* sp.2; H. *Caenis* sp.3; I. *Caenis* sp.4; J. *Caenis* sp.5 (Scale bar = 1 mm)

Malzacher and Sangpradub (2021) described six new species of *Caenis* from Thailand: *C. ludovici*, *C. nasuta*, *C. longiforcipata*, *C. karenae*, *C. obtusostilata* and *C. acutostilata*. Our study also found *C. nasuta* and *C. longiforcipata* as well but from different regions. One sample of *C. nasuta* was found in Kanchanaburi province, western Thailand and two specimens of *C. longiforcipata* were collected from Narathiwat province, southern Thailand. While these two species were previously reported from Northeastern Thailand, our results preliminarily indicate that they are widely distributed.

Moreover, the specimens collected from Khlong Wang, Narathiwat represent the first report of *C. cornigera* from Thailand. *C. cornigera* can be distinguished from *C. longiforcipata* by the setation of the operculate gill (Malzacher and Sangpradub 2021). The species was first described by Kang and Yang in 1994 from Taiwan based on the nymphal stage (Kang and Yang 1994) but there has been no previous reports of this species from other areas of the Oriental region including Thailand.

Molecular identification using COI by BOLD and GenBank

The COI amplicons from all 37 *Caenis* specimens were successfully amplified and sequenced. The 653 bp fragments of the COI were compared with BOLD and GenBank for species identification. Eleven samples matched two unidentified species recorded in Genbank and BOLD: *Caenis* sp.3 matched *Caenis* sp. (Genbank accession number ON704656.1) with 95% identity and *Caenis* sp.4 matched *Caenis* sp. (BOLD:AAL8077 and Genbank accession number HQ979423.1) with 94-95% identity. The other samples did not match the species level records when searching against both databases. However, there was no misidentification at the genus level in BOLD and GenBank. All our query sequences corresponded to the genus *Caenis* with greater than 80% similarity. According to the latest information, BOLD contains barcode sequences for 45,791 Ephemeroptera, 1,791 Caenidae and 1,261 *Caenis* covering formally described 27 *Caenis* species (accessed on 7th December 2022). BOLD and GenBank are the main public repositories of DNA barcode sequences (Meiklejohn et al. 2019), but the information for *Caenis* species identification is still inadequate. Pentinsaari et al. (2020) revealed that BOLD and GenBank demonstrated similar performance in identification and the limitations of both platforms to identify to species-level imposed by the lack of a polished species-level taxonomy for many organisms. The GenBank accession numbers of COI and 12sRNA sequences from this study were shown in Table 1.

Genetic divergence

Partial sequences of the COI and 12s rRNA genes were obtained for all 38 samples, including out-group species sample *Potamanthellus edmundsi*. The COI and 12s rRNA sequence alignments were 653 bp and 359 bp in length, respectively. Both genes had no indel (insertion/deletion) content. The average nucleotide composition of COI was 36.4% T, 19.9% C, 24.5% A and 19.2% G. The mean overall nucleotide frequencies observed for 12s rRNA sequences were 32.5% T, 11.8% C, 38.7% A and 17.0% G. For both genes, A + T (COI: 60.9%, 12s rRNA: 71.2%) were present in a substantially larger percentage than G + C (COI: 39.1%, 12s rRNA: 28.8%), as is typical for insects (Cameron 2014; Zhou and Yang 2022).

As shown in Table 2, the mean intraspecific K2P genetic distances of COI ranged from 0% to 2.35% with an average of 0.90%, while 12s rRNA K2P distances were between 0% and 2.0% with an average of 0.52%. The maximum mean intraspecific K2P divergences were 2.35% and 2.0% in *C. nasuta* for COI and 12s rRNA, respectively. In this study, two specimens of *C. nasuta* were collected from the west and northeast Thailand. The results suggested that geographic distance between two sampling sites may affect genetic variation within the species. Isolation by geographic distance is a primary and reasonable expectation for how genetic differentiation is structured at a geographic scale (Sexton et al. 2014). The mean interspecific genetic distance for COI ranged from 14.4% to 26.2% with a mean of 20.35%, while 12s rRNA ranged from 7.36% to 22.4% with a mean of 16.99% (as shown in Tables 3 and 4). The minimum interspecific genetic K2P divergence was 14.4% between *C. ulmeriana* and *Caenis* sp.1 for COI and 7.4% between *C. cornigera* and *C. longiforcipata* for 12s rRNA. The maximum intraspecific genetic divergences from both genes are considerably less than the minimum interspecific genetic divergences. There was clear evidence of a barcoding gap indicating that all *Caenis* species are genetically different from each other. Both COI and 12s rRNA genes gave low intraspecific and high enough interspecific values to be able to differentiate. The 12s rRNA is highly conserved in insects and used for the study of genetic diversity in phyla (Gerber et al. 2001). Yang et al. (2014) mentioned that short sequences (approximately 100 bp) of 12s rRNA might hinder species identification due to the existence of multiple similar sequences in closely related species but the longer sequences (approximately 430 bp) of 12s rRNA contains multiple polymorphisms that enable interspecies and intraspecific identification. Our study analyzed 359 bp of 12s rRNA from *Caenis* spp. and found that the 12s rRNA gene was also able to differentiate species and gave similar results of intraspecific and intraspecific genetic divergences when the COI gene was used.

Table 2. Mean intrapecific Kimura-2-Parameter nucleotide divergence (%) of COI and 12s rRNA genes

Species	COI	12s rRNA
Caenis nasuta	2.35	2.00
C. picea	1.06	0.00
C. ulmeriana	1.29	0.50
C cornigera	0.00	0.00
C. longiforcipata	0.00	0.00
Caenis sp.1	0.40	0.56
Caenis sp.2	2.07	1.03
Caenis sp.3	1.36	0.35
Caenis sp.4	0.30	0.56
Caenis sp.5	0.15	0.14
Mean	0.90	0.51

The Automatic Barcode Gap Discovery (ABGD) method revealed a barcoding gap between 4% and 13% (COI) and between 3% and 7% (12s rRNA) pairwise distances (Figure 2). The outcomes for both gene analyses agreed with one another. Initial partition of the sequence data at each value of the prior intraspecific divergence (P) divided data of the *Caenis* samples into ten groups that corresponded to the ten *Caenis* species. Recursive partition analysis with prior maximal distance P=7.74e-03 divided dataset for ten groups.

Phylogenetic tree construction

Through the Maximum Likelihood (ML) method, the phylogenetic trees constructed from COI and 12s rRNA genes were slightly different as shown in Figures 3 and 4. Both phylogenetic trees revealed that the 37 *Caenis* mayfly samples were grouped into monophyletic species and divided into individual species in the same clades but were unable to indicate the deeper evolutionary history. This is due to the sequences being conserved and the samples being too closely related to each other. However, the trees of the COI gene showed that *Caenis* sp.3 was a sibling to *Caenis* sp.2 while the tree of the 12s rRNA gene showed that *Caenis* sp.3 was sibling to *Caenis ulmeriana*. Because protein-coding genes like COI are typically under strong selective pressure to maintain their function, which can result in a higher rate of base substitutions and approximately three times higher rate of molecular evolution compared to non-coding regions like 12S or 16S rRNA (Knowlton and Weigt 1998).

Table 3. Mean interspecific Kimura-2-Parameter nucleotide divergence (%) of COI gene for 37 specimens belonging to 11 different species

		1	2	3	4	5	6	7	8	9	10	*P. edmundsi*
1.	*C. nasuta*											
2.	*C. picea*	16.4										
3.	*C. ulmeriana*	19.0	19.7									
4.	*Caenis* sp.5	19.5	18.0	17.5								
5.	*C. cornigera*	22.4	21.6	20.1	19.8							
6.	*C. longiforcipata*	22.6	26.2	19.5	21.1	16.1						
7.	*Caenis* sp.1	15.9	16.2	14.4	16.8	21.2	21.1					
8.	*Caenis* sp.2	21.6	24.1	21.7	21.9	21.2	25.2	24.1				
9.	*Caenis* sp.3	19.4	20.1	15.5	16.3	19.2	20.4	18.6	17.2			
10.	*Caenis* sp.4	23.3	20.1	22.5	21.9	21.7	24.5	21.6	24.7	24.0		
11.	*P. edmundsi*	21.7	19.4	17.7	18.5	22.4	22.8	17.9	24.7	18.1	25.8	

Table 4. Mean interspecific Kimura-2-Parameter nucleotide divergence (%) of 12s rRNA gene for 37 specimens belonging to 11 different species

		1	2	3	4	5	6	7	8	9	10	*P. edmundsi*
1.	*C. nasuta*											
2.	*C. picea*	11.6										
3.	*C. ulmeriana*	16.9	21.2									
4.	*Caenis* sp.5	18.0	16.1	19.2								
5.	*C. cornigera*	21.9	20.0	20.3	17.3							
6.	*C. longiforcipata*	21.0	21.1	17.4	15.5	7.36						
7.	*Caenis* sp.1	8.96	11.0	18.8	17.4	20.3	22.4					
8.	*Caenis* sp.2	16.7	18.3	18.3	16.4	15.0	13.0	15.9				
9.	*Caenis* sp.3	14.6	17.5	8.86	15.7	17.7	15.2	16.4	14.6			
10.	*Caenis* sp.4	15.9	17.3	19.7	20.4	20.5	20.3	15.9	16.0	19.3		
11.	*P. edmundsi*	33.9	37.0	34.0	31.1	32.0	33.8	32.4	33.1	32.3	35.2	

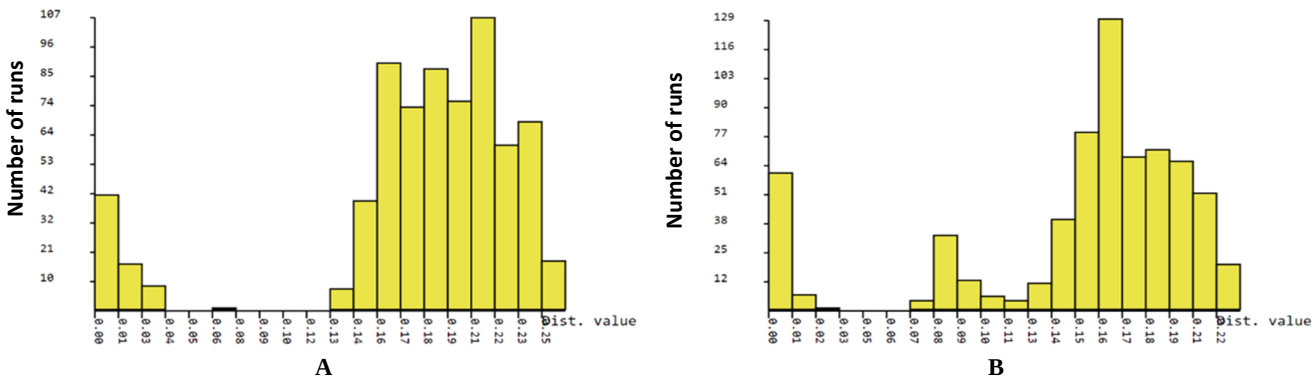

Figure 2. Results of Automatic Barcode Gap Discovery (ABGD) partition analysis for COI (A) and 12s rRNA (B) sequences of *Caenis* mayfly (nbr: number of runs)

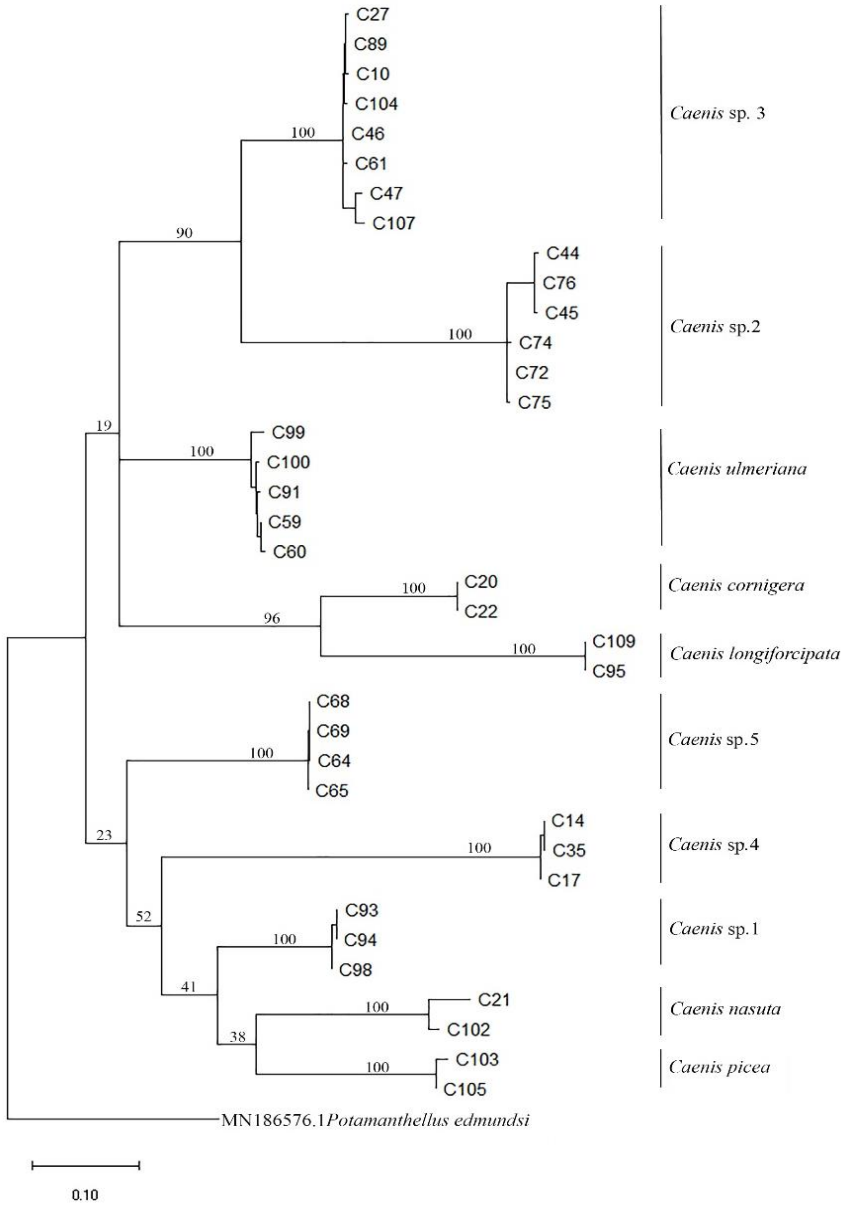

Figure 3. Phylogenetic tree of COI sequences generated from 37 *Caenis* specimens based on Maximum Likelihood (ML) analysis (the best model: T92+G+I, I=0.5918, G=1.1812, -lnl= 4287.83). Numbers at nodes indicate bootstrap values

The results were also in agreement with the morphological identification of all specimens. Each indeterminate species (sp.1 to sp.5) was also clustered in the same clade according to their given numbers. *C. cornigera* and *C. longiforcipata* were placed at the same node in both trees confirming their close relative species not only in morphology but also in genetics. The results proved that COI and 12s rRNA genes could be used in phylogenetic relationship studies of the family Caenidae. Mitochondrial based markers have been embraced and progressively utilized as molecular markers for phylogenetic studies and the mtDNA COI gene is especially valuable for determining intra and interspecific levels and can be utilized for phylogenetic analyses at both the species and genus levels because the marker has a well-defined barcoding gap (Matumba et al. 2020).

In conclusion, *Caenis* mayfly nymphs were collected from different regions in Thailand. They were classified into five species based on nymphal morphological identification. Using COI as barcoding on BOLD and GenBank, all specimens could be correctly identified at the genus level. We reported herein the comparative use of COI and 12s rRNA genes to investigate *Caenis* genetic diversity. Both genes gave similar results in both intra and interspecific divergences and in phylogenetic construction although the COI seems to present greater diversity values. This may be due to the longer nucleotide sequences of COI compared to 12s rRNA.

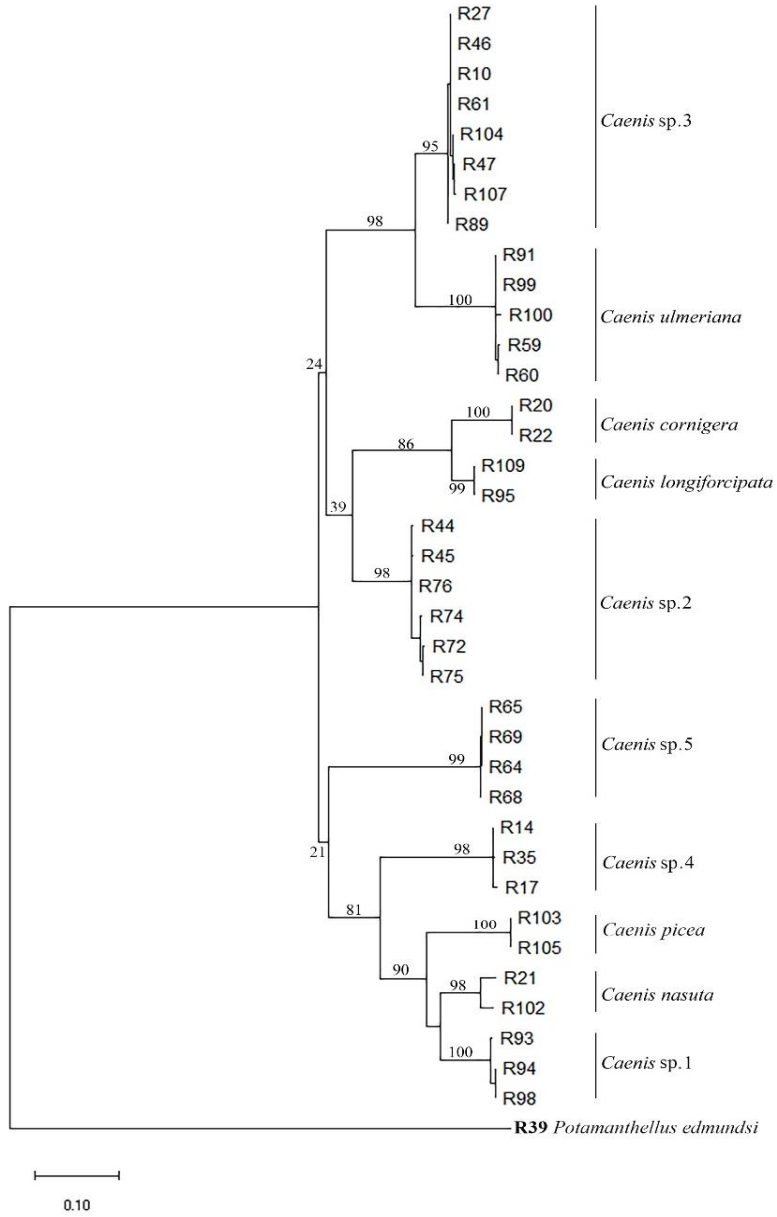

Figure 4. Phylogenetic tree of 12s rRNA sequences generated from 37 *Caenis* specimens based on Maximum Likelihood (ML) analysis (the best model: T92+G+I, I=0.2540, G=0.4952, -lnl= 2060.43). Numbers at nodes indicate bootstrap values

ACKNOWLEDGEMENTS

This work was partially financially supported by the Center of Excellence on Biodiversity (BDC), Office of Higher Education Commission (BDC-PG-160020) and Buriram Rajabhat University. We are grateful to Assoc. Prof. Dr. Narumon Sangpradub and members of Wet Lab., Department of Biology, Faculty of Science, Khon Kaen University for their valuable suggestions.

REFERENCES

Ajawatanawong P, Atkinson GC, Watson-Haigh NS, MacKenzie B, Baldauf SL. 2012. SeqFIRE: A web application for automated extraction of indel regions and conserved blocks from protein multiple sequence alignments. Nucleic Acids Res 40: W340-347. DOI: 10.1093/nar/gks561.

Alhejoj I, Salameh E, Bandel K. 2014. Mayflies (order Ephemeroptera): An effective indicator of water bodies conditions in Jordan. Intl J Sci Res Environ Sci 2 (10): 361-370. DOI: 10.12983/ijsres-2014-p0361-0370.

Allio R, Donega S, Galtier N, Nabholz B. 2017. Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Mol Biol Evol 34 (11): 2762-2772. DOI: 10.1093/molbev/msx197.

Almudi I, Martín-Blanco CA, García-Fernandez IM, Lopez-Catalina A, Davie K, Aerts S, Casares F. 2019. Establishment of the mayfly *Cloeon dipterum* as a new model system to investigate insect evolution. EvoDevo 10 (6): 1-10. DOI: 10.1186/s13227-019-0120-y.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215 (3): 403-410. DOI: 10.1016/S0022-2836(05)80360-2.

Barber-James H, Sartori M, Gattolliat J, Webb J. 2013. World checklist of freshwater Ephemeroptera species. In: Bánki O, Roskov Y, Döring M, Ower G, Vandepitte L, et al., Catalogue of Life Checklist. DOI: 10.48580/dfp3-3cn.

Boonsoong B, Sartori M. 2019. Review and integrative taxonomy of the genus *Prosopistoma* Latreille, 1833 (Ephemeroptera, Prosopistomatidae)

in Thailand, with description of a new species. Zookeys 825: 123-144. DOI: 10.3897/zookeys.825.32443.

Boonsoong B, Auychinda C, Sartori M, Khanyom N. 2021. First record of *Paegniodes* Eaton, 1881 (Ephemeroptera, Heptageniidae) from Thailand with description of a new species. Zookeys 1036: 153-170. DOI: 10.3897/zookeys.1036.64880.

Brandão NCA, Bittencourt JS, Calor AR, Mendes M, Langer MC. 2021. The Ephemeroptera (Hexapoda, Insecta) from the Lower Cretaceous Crato Formation (NE Brazil): A new genus and species, and reassessment of *Costalimella zucchii* Zamboni, 2001 and *Cratogenites corradiniae* Martins-Neto, 1996. Cretac Res 127: 104923. DOI: 10.1016/j.cretres.2021.104923.

Cai YY, Gao YJ, Zhang LP, Yu DN, Storey KB, Zhang JY. 2018. The mitochondrial genome of *Caenis* sp. (Ephemeroptera: Caenidae) and the phylogeny of Ephemeroptera in Pterygota. Mitochondrial DNA B Resour 3 (2): 577-579. DOI: 10.1080/23802359.2018.1467239.

Cameron SL. 2014. Insect mitochondrial genomics: Implications for evolution and phylogeny. Annu Rev Entomol 59: 95-117. DOI: 10.1146/annurev-ento-011613-162007.

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P. 2014. DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. Biol Lett 10 (9): 20140562. DOI: 10.1098/rsbl.2014.0562.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol 3 (5): 294-299.

Gerber AS, Loggins R, Kumar S, Dowling TE. 2001. Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes?. Annu Rev Genet 35: 539-566. DOI: 10.1146/annurev.genet.35.102401.091106.

Hall TA. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT. Nucleic Acids Symp Ser 41: 95-98

Hwang JM, Yoon TJ, Suh KI, Bae YJ. 2013. Molecular phylogeny evidence of altitudinal distribution and habitat adaptation in Korean *Ephemera* species (Ephemeroptera: Ephemeridae). Entomol Res 43 (1): 40-46. DOI: 10.1111/1748-5967.12008.

Jacobus LM, Macadam CR, Sartori M. 2019. Mayflies (Ephemeroptera) and their contributions to ecosystem services. Insects 10 (6): 170. DOI: 10.3390/insects10060170.

Kang SC, Yang CT. 1994. Caenidae of Taiwan (Ephemeroptera). Chin J Entomol 14: 93-113. DOI: 10.6660/TESFE.1994009.

Knowlton N, Weigt LA. 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc Biol Sci 265 (1412): 2257. DOI: 10.1098/rspb.1998.0568.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 35 (6): 1547-1549. DOI: 10.1093/molbev/msy096.

Malzacher P. 2013. Caenidae from East Kalimantan, Borneo (Insecta: Ephemeroptera). With a discussion on phylogeny of the new tribe Clypeocaenini, subfamily Caeninae. Stuttg Beitr Naturkd A 6: 21-55.

Malzacher P. 2015. Revision of the oriental species of the genus *Caenis* Stephens (Insecta: Ephemeroptera: Caenidae). Stuttg Beitr Naturkd A 8: 27-47.

Malzacher P, Sangpradub N. 2017. Revision of the tribe Caenoculini (Insecta: Ephemeroptera: Caenidae) and its position within the Brachycercinae. Stuttg Beitr Naturkd A 10: 1-17. DOI: 10.18476/sbna.v10.a1.

Malzacher P, Sangpradub N. 2021. New mayfly species of *Caenis* and *Kalimaenis* from Thailand and descriptions of two new genera of the subfamily Caeninae (Ephemeroptera: Caenidae). Stuttg Beitr Naturkd A 3: 1-33. DOI: 10.18476/insy.v03.a1.

Malzacher P. 2021. *Mekongellina*, a new mayfly genus of the subfamily Brachycercinae (Ephemeroptera: Caenidae) from Thailand. Stuttg Beitr Naturkd A 2: 1-9. DOI: 10.18476/insy.v02.a1.

Matumba TG, Oliver J, Barker NP, McQuaid CD, Teske PR. 2020. Intraspecific mitochondrial gene variation can be as low as that of nuclear rRNA. F1000Res 9: 339. DOI: 10.12688/f1000research.23635.2.

Meiklejohn KA, Damaso, N, Robertson JM. 2019. Assessment of BOLD and GenBank - Their accuracy and reliability for the identification of biological materials. PLoS One 14 (6): e0217084. DOI: 10.1371/journal.pone.0217084.

Misof B, Liu S, Meusemann K, et al. 2014. Phylogenomics resolves the timing and pattern of insect evolution. Science 346 (6210): 763-767. DOI: 10.1126/science.1257570.

Muthukatturaja M, Balasubramanian C, Murugan A. 2020. Two new species of Clypeocaenis (Ephemeroptera: Caenidae) from Western Ghats, Southern India. Zootaxa 4722 (4): 361-370. DOI: 10.11646/zootaxa.4722.4.5.

Muthukatturaja M, Balasubramanian C. 2021. A new species of *Caenis* Stephens, 1835 (Ephemeroptera: Caenidae) from Tamil Nadu, Southern India. Zootaxa 4980 (2): 366372. DOI: 10.11646/zootaxa.4980.2.7.

Pentinsaari M, Ratnasingham S, Miller SE, Hebert PDN. 2020. BOLD and GenBank revisited - Do identification errors arise in the lab or in the sequence libraries? PLoS One 15 (4): e0231814. DOI: 10.1371/journal.pone.0231814.

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Mol Ecol 21 (8): 1864-1877. DOI: 10.1111/j.1365-294X.2011.05239.x.

Ratnasingham S, Hebert PDN. 2007. BOLD: The barcode of life data system (www.barcodinglife.org). Mol Ecol Notes 7 (3): 355-364. DOI: 10.1111/j.1471-8286.2007.01678.x.

Ratnasingham S, Hebert PDN. 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. PLoS One 8 (7): e66213. DOI: 10.1371/journal.pone.0066213.

Sangpradub N, Boonsoong B. 2006. Identification of Freshwater Invertebrates of the Mekong River and its Tributaries. Mekong River Commission. Vientiane, Lao PDR.

Sartori M, Brittain JE. 2015. Order Ephemeroptera. In: Thorp JH, Rogers DC (eds). Thorp and Covich's Freshwater Invertebrates (4th ed). Academic Press, New York. DOI: 10.1016/B978-0-12-385026-3.00034-6.

Scarduelli L, Giacchini R, Parenti P, Migliorati S, Di Brisco AM, Vighi M. 2017. Natural variability of biochemical biomarkers in the macro-zoobenthos: Dependence on life stage and environmental factors. Environ Toxicol Chem 36 (11): 3158-3167. DOI: 10.1002/etc.3893.

Selvakumar C, Sivaramakrishnan KG, Janarthanan S. 2016. DNA barcoding of mayflies (Insecta: Ephemeroptera) from South India. Mitochondrial DNA B Resour 1 (1): 651-655. DOI: 10.1080/23802359.2016.1219623.

Sexton JP, Hangartner SB, Hoffmann AA. 2014. Genetic isolation by environment or distance: Which pattern of gene flow is most common? Evolution 68 (1): 1-15. DOI: 10.1111/evo.12258.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87 (6): 651-701. DOI: 10.1093/aesa/87.6.651.

Sivaramakrishnan KG, Janarthanan S, Selvakumar C, Arumugam M. 2014. Aquatic insect conservation: A molecular genetic approach. Conserv Genet Resour 6: 849-855. DOI: 10.1007/s12686-014-0250-4.

Soldán T. 1978. New genera and species of Caenidae (Ephemeroptera) from Iran, India and Australia. Acta Entomol Bohemoslov 75: 119-129.

Srinivasan P, Sivaruban T, Barathy S, Malzacher P, Isack R. 2021. A new charismatic *Caenis* Stephens, 1835 (Ephemeroptera: Caenidae) from Southern India. Zootaxa 4926 (1): 105-116. DOI: 10.11646/zootaxa.4926.1.7.

Suh KI, Hwang JM, Bae YJ, Kang JH. 2019. Comprehensive DNA barcodes for species identification and discovery of cryptic diversity in mayfly larvae from South Korea: Implications for freshwater ecosystem biomonitoring. Entomol Res 49 (1): 46-54. DOI: 10.1111/1748-5967.12334.

Takayanagi T, Yoshizawa K. 2021. The first record of *Caenis rivulorum* (Ephemeroptera: Caenidae) from Japan. Biodivers Data J 9: e67413. DOI: 10.3897/BDJ.9.e67413.

Tong XL, Dudgeon D. 2002. Three new species of the genus *Caenis* from Hong Kong, China (Ephemeroptera: Caenidae). Zool Res 23 (3): 232-238.

Van Bruggen, AC. 1954. *Caenis demoulini*, a new Ephemeropteron from Thailand. Zool Meded 23 (2): 11-15.

Xu XD, Jia YY, Dai XY, Ma JL, Storey KB, Zhang JY, Yu DN. 2020. The mitochondrial genome of *Caenis* sp. (Ephemeroptera: Caenidae) from Fujian and the phylogeny of Caenidae within Ephemeroptera. Mitochondrial DNA B Resour 5 (1): 192-193. DOI: 10.1080/23802359.2019.1698986.

Yang L, Tan Z, Wang D, Xue L, Guan MX, Huang T, Li R. 2014. Species identification through mitochondrial rRNA genetic analysis. Sci Rep 4: 4089. DOI: 10.1038/srep04089.

Zhou J, Yang DY. 2022. Mitochondrial genomes provide new phylogenetic and evolutionary insights into Psilidae (Diptera: Brachycera). Insects 13 (6): 518. DOI: 10.3390/insects13060518.