VNU Journal of Science: Natural Sciences and Technology
Journal homepage: https://js.vnu.edu.vn/NST

Original Article

Sequencing Analysis of *COI* Genes Isolated from Vietnam's Endemic Insect Species (Insecta: Plecoptera) for DNA Barcoding Database

Nguyen Quang Huy[1,2], Nguyen Thi Thu Huyen[1], Ha Thi Lan Anh[1], Nguyen Thi Uyen[1], Le Duc Viet[1], Dang Thi Lan Anh[1], Ho Thi Yen[1], Tran Anh Duc[1], Nguyen Van Vinh[1], Cao Thi Kim Thu[3], Nguyen Van Sang[1,*]

[1]*VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Hanoi, Vietnam*
[2]*Key Laboratory of Enzyme and Protein Technology, 334 Nguyen Trai, Thanh Xuan, Hanoi, VietNam*
[3]*Institute of Ecology and Biological Resources, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

Received 12 October 2022
Revised 20 March 2023; Accepted 26 July 2023

Abstract: The Insecta are the most diverse group of animals on the planet. The diversity of Plecoptera aquatic insect order is used as an environmental and ecological indicator. The study of endemic species is important for confirming the national biodiversity and contributing to the determination of national sovereignty. Despite Vietnam being a biodiversity hotspot, the number of known insect species remains limited. As a result, DNA barcoding has been used as a cutting-edge technique to efficiently and rapidly classify biological species. The *Cytochrome c oxidase subunit I* (*COI*) gene is widely regarded as a standard barcode for animal classification. In this study, we amplified and sequenced a portion of the *COI* gene fragment with a molecular weight of 710 bp for eight insect species belonging to the order Plecoptera, endemic to Vietnam. Compared to the previously published sequences of other Plecoptera species, the COI gene sequence analysis results showed a high divergence of at least 9.9%. Molecular analysis of the studied species accurately demonstrated their endemism and morphological similarities. This study contributes to the development of a DNA barcoding database for endemic insects in Vietnam.

Keywords: DNA barcoding, Plecoptera, Endemic insects, Vietnam.

1. Introduction

Insects are among the most diverse groups of animals on the Earth. According to Stork (2015), the average number of insect species worldwide is approximately 5.5 million, with over 1.1 million species classified and described [1]. Owing to a lack of experts and research on the class Insecta, approximately 80% of insect species remain unidentified. This is because the traditional classification techniques based on insect morphology have certain limitations such as waste of time for finding [2], having errors by the phenotypical

* Corresponding author.
E-mail address: nvsangvnu@hus.edu.vn

https://doi.org/10.25073/2588-1140/vnunst.5513

variation [3, 4]. To address these restrictions, DNA barcoding has been identified as an efficient classification method for the majority of organisms. Hebert et al., coined the term "DNA barcode" in 2003 and established the significance of the *Cytochrome c oxidase subunit I* (*COI*) gene sequence isolated from the mitochondrial genome in differentiation among Lepidoptera species [5]. The *COI* gene is a prime candidate as a standard DNA barcode among the various mitochondrial genes (such as *12S* rRNA, *16S* rRNA, *ND*1-6), owing to several factors, including (1) the ease of designing universal primers due to the presence of conserved regions [6], (2) its sufficient length for sequencing and analysis [7], and (3) the characteristic differences it exhibits for each species [7]. Compared to nuclear genes (5.*8S* rDNA, *18S* rDNA, *28S* rDNA, *ITS* regions, elongation factor 1-α), the *COI* gene is a protein-coding region that evolves rapidly and exists in multiple copies, thereby offering significant advantages for amplification and classification purposes. Then, the *COI* gene was widely utilized as the DNA barcode in the classification of animal species, including the insect group [8]. Therefore, the application of mitochondrial *COI* gene as DNA barcodes has made significant strides in recent years, particularly in the areas of species biodiversity [9-12], description of insect development throughout its life stages [13, 14], and identification of cryptic species [15, 16].

Vietnam is considered one of the prominent biodiversity hotspots. It is home to a wide variety of flora and fauna, including numerous endemic species listed as the Living National Treasures (lntreasures.com/Vietnam). The order, Plecoptera (stoneflies) includes endemic aquatic insect species reported as an ideal and cost-effective means for the assessment of water quality [12] because of their sensitivity and poor resistance to environmental pollution. Stoneflies have a high endemism rate because they rely on ecosystems with high levels of dissolved oxygen in moving water [11, 17]. However, previous classifications of endemic insect species in Vietnam using DNA barcoding were limited. We, therefore, conducted this study to examine the *COI* gene sequences of selected endemic Plecopteran species in Vietnam. These findings contribute to developing a global network of taxonomic systems and integrating DNA barcoding in applied biodiversity, serving as a platform for scientific research.

2. Experimental

2.1. Samples

In this study, eight Plecoptera specimens from Vietnam's national parks and conservation areas were collected from the wild. These specimens were morphologically examined to determine their endemism and stored in absolute alcohol at 4-8 °C until DNA extraction. The sample data are provided below.

Table 1. List of specimens collected in this study

Nº	**Specimen voucher**	**Morphological species**	**Collected area**
1	PLE-DNA.01	*Cerconychia sapa* Stark & Sivec, 2007	Lao Cai, Bac Kan, Cao Bang, Son La
2	PLE-DNA.03	*Flavoperla hmong*Stark & Sivec, 2008	Cao Bang, Lao Cai, Son La, Thua Thien - Hue
3	PLE-DNA.04	*Togoperla thinhi* Cao & Bae, 2010	Thua Thien - Hue, Quang Nam, Cao Bang
4	PLE-DNA.05	*Neoperla zonata* Stark & Sivec, 2008	Cao Bang, Lao Cai, Son La
5	PLE-DNA.06	*Sinacroneuria biocellata* Stark & Sivec, 2008	Cao Bang, Lao Cai
6	PLE-DNA.08	*Cryptoperla karen* Stark, 1989	Vinh Phuc, Cao Bang, Son La
7	PLE-DNA.09	*Peltoperlopsis malickyi* Stark & Sivec, 1999	Vinh Phuc, Cao Bang
8	PLE-DNA.10	*Neoperla yentu* Cao & Bae, 2007	Quang Ninh, Cao Bang, Son La

2.2. Total DNA Isolation

Muscular tissues (legs) for DNA isolation were obtained from the collected insects. DNA extraction was performed using the phenol-chloroform method [18] as follows: i) the sample was prepared and cut into small pieces. The alcohol was allowed to evaporate at room temperature, and the samples were transferred to a newly labeled microcentrifuge tube; ii) The samples were lysed with lysis buffer (0.5% SDS, 10 mM Tris HCl pH 8.0, 100 mM EDTA pH 8.0, and 50 mM NaCl) supplemented with proteinase K (2 mg/mL) and incubated at 56 °C overnight; and iii) The lysed solution was treated with a mixture of phenol: chloroform: isoamyl alcohol (25:24:1) and then transferred upper layer to a newly labeled microcentrifuge tube. (4) DNA was precipitated with absolute ethanol, and (5) DNA samples were collected and stored at -20 °C for subsequent experiments.

2.3. Primer Design and Amplification

The primers used in this study were designed based on the *COI* gene sequences of insect mitochondria [6, 19]. The selected primers were used to amplify a 710 bp fragment at the 3′ end of the *COI* gene.

The designed primer pair (dg-LCO1490/dg-HCO2198) was a modified form of the universal primer pair (LCO1490 / HCO2198). The sequences of the primer pairs used in the study are presented in Table 2.

Table 2. Primer sequences used for *COI* gene fragment amplification

Primer name	**Direction**	**Sequence (5'-3')**	**Size (bp)**	**Reference**
LCO1490	Forward	GGTCAACAAATCATAAAGATATTGG	710	Folmer *et al.* [6]
HCO2198	Reverse	TAAACTTCAGGGTGACCAAAAAATCA		
dg-LCO1490	Forward	TCTCTACCAAYCAYAARGAYATTGG	710	This study
dg-HCO2198	Reverse	TATACYTCTGGRTGGCCRAARAAYCA		

The polymerase chain reaction (PCR) mix included 5 µL HF buffer 5X (100 mM Tris-HCl, pH 8.5, 300 mM KCl, 100 mM $(NH_4)_2SO_4$, 7.5 mM $MgCl_2$), 1 µL forward primer (10 µM), 1 µL reverse primer (10 µM), 0.5 µL dNTPs mix (10 mM), 0.25 µL Phusion Hotstart II DNA polymerase enzyme (Thermo Scientific, Waltham, Massachusetts, USA) (2 U/µL) and 1.5 µL DNA template extracted from the insect samples. The final volume was adjusted to 25 µL with nuclease-free water. The thermal cycling conditions were as follows: initial denaturation at 95 ºC for 30 s, 45 cycles of denaturation at 95 ºC for 30 s, annealing at 45-50 ºC for 45 s, elongation at 72 ºC for 30 s, and final elongation at 72 ºC for 5 min. The PCR products were run on a 1% agarose gel prepared in TAE buffer at 100 V for 25 min and stained with Redsafe (iNtRON). After electrophoresis, the gel was exposed to UV light and imaged using an Alpha Imager system.

2.4. Sequencing and Analysis

The purified PCR products were outsourced to the 1st Base company (Malaysia) for bidirectional sequencing using the Sanger method. The top and bottom sequences were read and trimmed using BioEdit 7.2.5 software [20]. The sequences were compared and assembled for consensus sequence in ".FASTA" or ".txt" format. All sequences were aligned using the ClustalW method to determine the gaps before being subjected to phylogenetic analysis.

2.5. Phylogenetic Analysis

The final dataset of the *COI* gene contained eight studied and twenty reference sequences

obtained from the National Center for Biotechnology Information (NCBI). The sequences were aligned and subjected to phylogenetic analysis. A phylogenetic tree was constructed by MEGA7 [21] using the Maximum Likelihood method. To find the best DNA/Protein Model for your *COI* sequences we have used the DNA/Protein Model (ML) tool [21] from MEGA7 software. The obtained result suggested that the GTR+G+I model with the lowest the Bayesian Information Criterion score (BIC) is the most suitable model for phylogenetic analysis [22]. The accuracy of the phylogenetic tree was assessed by bootstrapping with 1000 replicates [23]. Genetic distances were performed using the uncorrected distance (*p*-distance) approach based on the number of base substitutions per site of the dataset after removing the gaps and missing data.

3. Results and Discussion

3.1. Total DNA Isolation

We analyzed the obtained DNA samples by agarose gel electrophoresis, and the results are shown in Figure 1. The genomic DNA extracted from eight insect samples presented bright and unbroken bands with an individual molecular weight greater than 10 kb.

This indicated a successful extraction of good-quality total DNA from the samples. The DNA samples were then used to clone the *COI* genes using universal and designed primers (Table 1).

Figure 1. Agarose gel electrophoresis of total DNA extracted from eight insect samples.

Mk: Marker 1 kb plus (Cleaver); The name of the studied species were signed as Ple01 (*Cerconychia sapa*), Ple02 (*Flavoperla hmong*), Ple03 (*Togoperla thinhi*), Ple04 (*Neoperla zonata*), Ple05 (*Sinacroneuria biocellata*), Ple06 (*Cryptoperla karen*), Ple07 (*Peltoperlopsis malickyi*), Ple08 (*Neoperla yentu*).

3.2. PCR Amplification of COI Gene

The *COI* fragments from eight DNA samples were amplified using two separate PCR reactions with two pairs of primers LCO1490/HCO2198 and dg-LCO1490/dg-HCO2198. As shown in Figure 2, the PCRs using primer pair LCO1490 / HCO2198 showed that a uniform bright band was observed for each of the six samples including Ple01, Ple02, Ple04, Ple05, Ple06, and Ple07. This indicated that the PCRs using the primer pair LCO1490/ HCO2198 successfully amplified the *COI* gene fragment in these samples (Figure 2). In contrast, no band appeared for Ple03 and Ple08, suggesting that the gene amplification failed. We, therefore, attempted to amplify the gene fragment in these samples with the designed primer pair dg-LCO1490/dg-HCO2198 (Figure 3). In Figure 3, the lanes loaded with the samples, Ple03 and Ple08 presented bright bands of approximately 710 bp. This result confirmed the successful amplification of the *COI* gene fragments of Ple03 and Ple08 using the primer pair dg-LCO1490/dg-HCO2198. The PCR products with a molecular weight of approximately 710 bp were purified and subjected to sequencing.

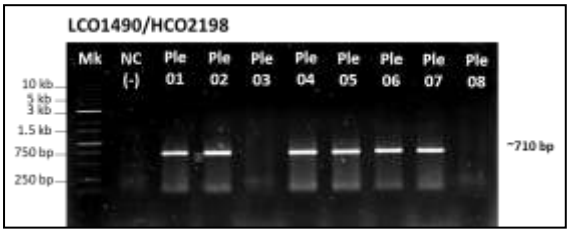

Figure 2. Agarose gel electrophoresis of desirable PCR products of 6/8 Plecoptera specimens.

PCR with primer pair LCO1490/ HCO2198. Mk: Marker 1 kb plus (Cleaver); NC(-): negative control without the DNA template; The name of the studied species were signed as Ple01 (*Cerconychia sapa*), Ple02 (*Flavoperla hmong*), Ple03 (*Togoperla thinhi*),

Ple04 (*Neoperla zonata*), Ple05 (*Sinacroneuria biocellata*), Ple06 (*Cryptoperla karen*), Ple07 (*Peltoperlopsis malickyi*), Ple08 (*Neoperla yentu*).

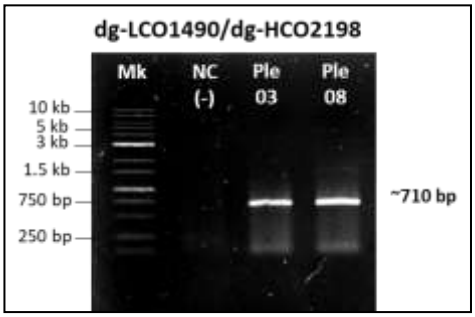

Figure 3. Agarose gel electrophoresis of desirable PCR products of 2/8 Plecoptera specimens.

PCR with primer pair dg-LCO1490/dg-HCO2198. Mk: Marker 1 kb plus (Cleaver); NC(-): negative control without the DNAtemplate; The names of the studied species were signed as Ple03 (*Togoperla thinhi*), Ple08 (*Neoperla yentu*).

3.3. Sequencing Analysis

Bioinformatics software was used to analyze the *COI* gene sequences. The analysis revealed that the received sequences were successfully amplified and had a length ranging between 676 bp and 700 bp (Table 3). All sequences were deposited in NCBI GenBank under the accession numbers OP160244 (Ple01), OP160246 (Ple02), OP160247 (Ple03), OP160248 (Ple04), OP160249 (Ple05), OP160250 (Ple06), OP160251 (Ple07), OP160252 (Ple08). The ClustalW method was used to align the sequences. Figure 4 displays the results of the multiple sequence alignment for the *COI* gene of the studied species, illustrating the similarity and variation between sequences.

Table 3. Sequence information of samples used in this study

Nº.	**Sequence code**	**Specimen voucher**	**Size (bp)**	**Accession number**
1	Ple01	Ple-DNA.01	700	OP160244
2	Ple02	Ple-DNA.03	700	OP160246
3	Ple03	Ple-DNA.04	680	OP160247
4	Ple04	Ple-DNA.05	686	OP160248
5	Ple05	Ple-DNA.06	689	OP160249
6	Ple06	Ple-DNA.08	689	OP160250
7	Ple07	Ple-DNA.09	689	OP160251
8	Ple08	Ple-DNA.10	676	OP160252

3.4. Phylogeny Analysis

Initially, we conducted BLAST searches of the sequence data against the NCBI GenBank database to prepare for constructing the phylogenetic tree. The outcomes of these searches are shown in Table 4.

The phylogenetic analysis was performed using *COI* gene sequences of 8 endemic Plecoptera species of Vietnam: Ple01 (*Cerconychia sapa*), Ple02 (*Flavoperla hmong*), Ple03 (*Togoperla thinhi*), Ple04 (*Neoperla zonata*), Ple05 (*Sinacroneuria biocellata*), Ple06 (*Cryptoperla karen*), Ple07 (*Peltoperlopsis malickyi*), and Ple08 (*Neoperla yentu*); 27 GenBank references of genera, *Cerconychia, Styloperla, Cryptoperla, Peltoperlopsis, Flavoperla, Sinacroneuria, Togoperla, Calineuria, Dinocras, Yugus, Suwallia, Capnia* and *Isogenoides*, and three species *Ephemera orientalis, Parafronurus youi* and *Isonychia ignota* designated as outgroups.

The phylogenetic tree (Figure 5) illustrates the relationships between the stonefly species selected as candidates for the study.

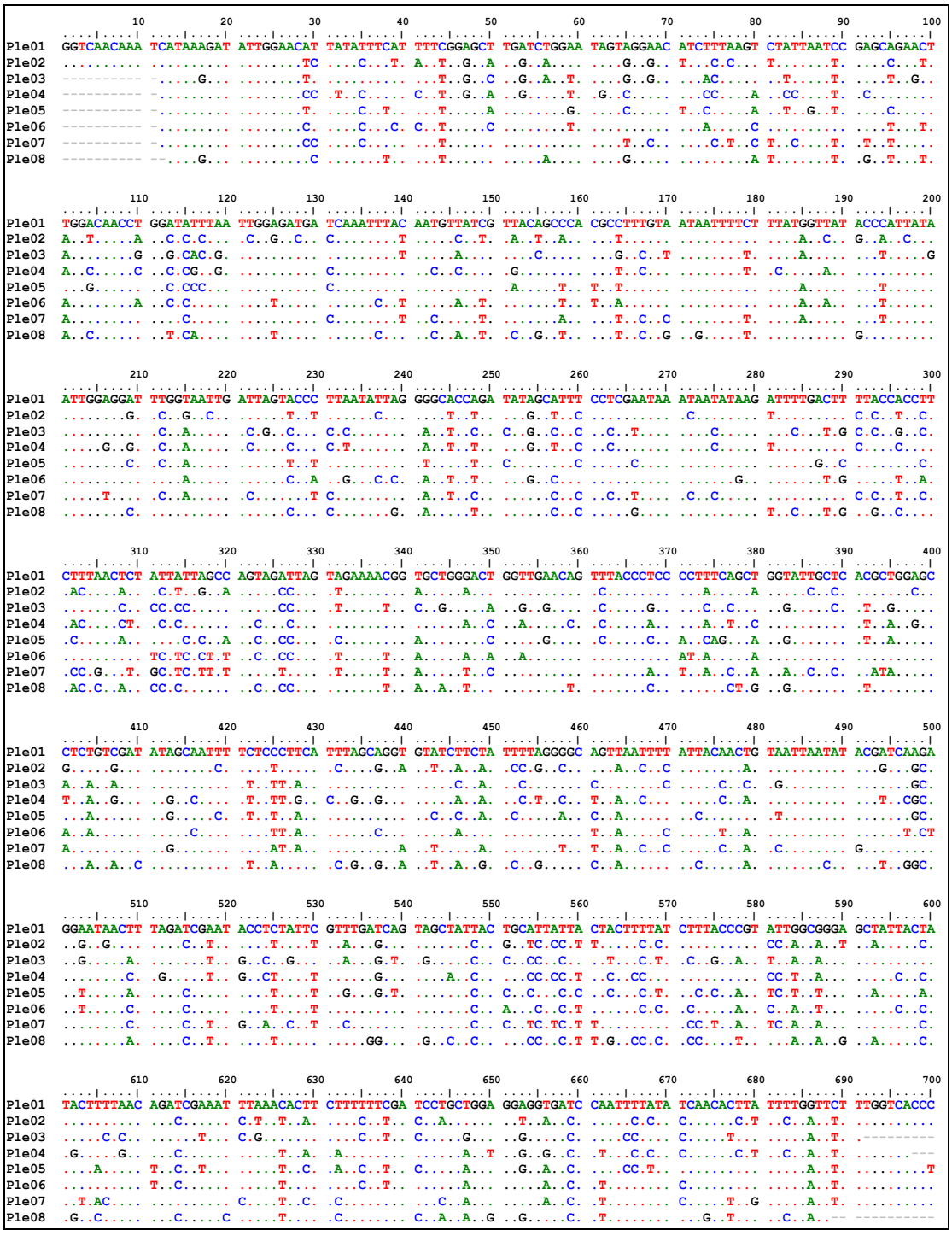

Figure 4. Sequence alignment for the *COI* gene of eight species belonging to the Plecoptera order. The length of eight *COI* gene sequences of studied samples was 676–700 bp. The names of the *COI* gene sequences were signed as Ple01 (*Cerconychia sapa*), Ple02 (*Flavoperla hmong*), Ple03 (*Togoperla thinhi*), Ple04 (*Neoperla zonata*), Ple05 (*Sinacroneuria biocellata*), Ple06 (*Cryptoperla karen*), Ple07 (*Peltoperlopsis malickyi*), Ple08 (*Neoperla yentu*). Similar bases are represented by dots and the sites of variation are depicted.

Table 4. The "BLAST search" data of eight *COI* gene sequences from Plecoptera endemic species in Vietnam against the NCBI GenBank database

Nº.	**Samples in study**	***COI* gene fragment (bp)**	**"BLAST search" include morphological genus**			**"BLAST search" exclude morphological genus**		
			Species name	*Similarity score (% Identity)*	*(Number of identical sites/ query size)*	*Species name*	*Similarity score (% Identity)*	*(Number of identical sites/ query size)*
1	*Cerconychia sapa* (Ple01)	700	*Cerconychia flectospina* (MF100783)	89.86%	620/690 (90%)	*Styloperla spinicercia* (KX845569)	87.25%	609/698 (87%)
2	*Flavoperla hmong* (Ple02)	700	*Flavoperla sp.* (MN419916)	86.33%	600/695 (86%)	*Flavoperla hatakeyamae* (NC_057436)	85.10%	594/698 (85%)
3	*Togoperla thinhi* (Ple03)	680	*Togoperla brevispinis* (AB770034)	81.19%	505/622 (81%)	*Capnia zijinshana* (KX094942)	81.82%	558/682 (82%)
4	*Neoperla zonata* (Ple04)	686	*Neoperla sp.* (KX091859)	84.84%	582/686 (85%)	*Calineuria stigmatica* (MG677941)	81.91%	557/680 (82%)
5	*Sinacroneuria biocellata* (Ple05)	686	*Sinacroneuria dabieshana* (MK492253)	85.67%	592/691 (86%)	*Flavoperla sp.* (MN419916)	83.38%	577/692 (83%)
6	*Cryptoperla karen* (Ple06)	689	*Cryptoperla kawasawai* (NC_065830)	86.00%	596/693 (86%)	*Capnia yunnana* (NC_064508)	85.78%	591/689 (86%)
7	*Peltoperlopsis malickyi* (Ple07)	689	*Peltoperlopsis sagittata* (NC_065831)	85.51%	590/690 (86%)	-	-	-
8	*Neoperla yentu* (Ple08)	689	*Neoperla sabang* (MZ224447)	80.06%	542/677 (80%)	*Styloperla spinicercia* (KX845569)	81.51%	551/676 (82%)

Table 5. Genetic distances inferred from the partial sequences of the *COI* genes between species selected for this study

Name	**Ple01**	**Ple02**	**Ple03**	**Ple04**	**Ple05**	**Ple06**	**Ple07**	**Ple08**
Cerconychia sapa (Ple01) (OP160244)								
Flavoperla hmong (Ple02) (OP160246)	0.208							
Togoperla thinhi (Ple03) (OP160247)	0.205	0.224						
Neoperla zonata (Ple04) (OP160248)	0.210	0.202	0.210					
Sinacroneuria biocellata (Ple05) (OP160249)	0.185	0.198	0.197	0.216				
Cryptoperla karen (Ple06) (OP160250)	0.163	0.187	0.197	0.205	0.189			
Peltoperlopsis malickyi (Ple07) (OP160251)	0.203	0.200	0.198	0.208	0.224	0.187		
Neoperla yentu (Ple08) (OP160252)	0.203	0.211	0.223	0.202	0.218	0.203	0.231	

Four major clusters corresponded to each studied family: Styloperlidae, Peloperlidae, Perlidae, and Perlodidae. However, the results of the genetic distance analysis (Tables 5 and 6) indicate that the *COI* gene was highly variable between the species within each genus. Thus, the molecular phylogenetic tree and genetic distance matrix were used to estimate the evolutionary divergence between sequences and calculate the origin of the species in the evolutionary tree.

Table 6. Genetic distances inferred from the partial sequences of the *COI* genes between species selected for this study and GenBank references

Reference species	**Ple01**	**Ple02**	**Ple03**	**Ple04**	**Ple05**	**Ple06**	**Ple07**	**Ple08**
Cerconychia flectospina (MF100783)	0.099	0.174	0.205	0.218	0.191	0.173	0.174	0.194
Styloperla spinicercia (NC 034809)	0.134	0.208	0.205	0.199	0.186	0.173	0.215	0.195
Styloperla spinicercia (KX845569)	0.134	0.208	0.205	0.199	0.186	0.173	0.215	0.195
Cryptoperla kawasawai (AB746173)	0.189	0.197	0.195	0.210	0.204	0.145	0.176	0.202
Peltoperlopsis sagittata (NC 065831)	0.186	0.220	0.213	0.207	0.192	0.158	0.148	0.228
Peltoperlopsis cebuano (MK387068)	0.207	0.202	0.225	0.217	0.195	0.171	0.189	0.213
Isogenoides colubrinus (HQ578977)	0.213	0.199	0.192	0.200	0.204	0.192	0.199	0.212
Isogenoides zionensis (HQ961197)	0.213	0.182	0.212	0.221	0.208	0.192	0.204	0.220
Isogenoides frontalis (GU711831)	0.213	0.200	0.191	0.199	0.204	0.189	0.205	0.212
Suwallia errata (MF198253)	0.200	0.212	0.187	0.194	0.210	0.191	0.207	0.220
Yugus bulbosus (HQ568928)	0.226	0.236	0.223	0.220	0.221	0.215	0.205	0.228
Yugus arinus (HQ961194)	0.213	0.216	0.203	0.200	0.205	0.200	0.197	0.218
Flavoperla sp. (MK905206)	0.189	0.145	0.181	0.192	0.173	0.171	0.192	0.191
Flavoperla hatakeyamae (MN821010)	0.194	0.156	0.208	0.207	0.187	0.191	0.197	0.197
Sinacroneuria dabieshana (MK492253)	0.197	0.189	0.197	0.200	0.142	0.189	0.210	0.220
Togoperla brevispinis (AB770034)	0.210	0.199	0.191	0.238	0.192	0.208	0.226	0.220
Togoperla sp. (KM409708)	0.210	0.215	0.186	0.202	0.218	0.202	0.199	0.217
Togoperla limbata (AB770119)	0.202	0.194	0.195	0.204	0.181	0.186	0.217	0.218
Calineuria stigmatica (MG677941)	0.235	0.221	0.220	0.189	0.205	0.212	0.230	0.254
Dinocras cephalotes (NC 022843)	0.236	0.220	0.178	0.205	0.210	0.208	0.217	0.230
Neoperla sp. (KX091859)	0.205	0.208	0.213	0.156	0.207	0.204	0.221	0.197
Neoperla sabang (MZ224447)	0.217	0.218	0.210	0.182	0.210	0.212	0.210	0.208
Capnia zijinshana (KX094942)	0.179	0.207	0.194	0.208	0.217	0.169	0.191	0.207
Capnia yunnana (NC 064508)	0.161	0.189	0.195	0.207	0.194	0.148	0.195	0.195
Ephemera orientalis (NC 012645)	0.202	0.199	0.228	0.233	0.207	0.186	0.207	0.217
Parafronurus youi (NC 011359)	0.189	0.225	0.235	0.239	0.205	0.210	0.213	0.235
Isonychia ignota (HM143892)	0.218	0.226	0.238	0.254	0.235	0.233	0.248	0.256

The genetic distance between the stoneflies selected for this study and those belonging to different families found in the GenBank database (Tables 5 and 6) indicated that *Cerconychia sapa* (Ple01) (OP160244) was the closest to *Cerconychia flectospina* (MF100783) belonging to the same genus, with a similarity value of 90.1%. Similarly, *Flavoperla hmong* (Ple02) (OP160246), *Neoperla zonata* (Ple04) (OP160248), *Sinacroneuria biocellata* (Ple05) (OP160249), *Cryptoperla karen* (Ple06) (OP160250), and *Peltoperlopsis malickyi* (Ple07) (OP160251) were also closely related to the species of the same genus, with similarity values of 85.5%, 84.4%, 85.8%, and 85.5%, and 85.2% respectively. As shown in Figure 5, both species *Neoperla zonata* (Ple04) (OP160248) and *Neoperla yentu* (Ple08) (OP160252) belonged to the genus *Neoperla* grouped into a monophyletic clade in the phylogeny, whereas *Neoperla yentu* (Ple08) (OP160252) was the closest to *Flavoperla sp*. (MK905206), the species of another genus in the family Perlidae with a high divergence value of 19.1%.

On the other hand, the phylogeny also indicated that the classification of the species within the genus *Togoperla* under the family Perlidae was not plain. *Togoperla thinhi* (Ple03) (OP160247) grouped in the genus *Dinocras* showed a divergence value of 17.8% with *Dinocras cephalotes* (NC_022843). *Togoperla sp*. (KM409708) and *Togoperla limbata* (AB770119) revealed a divergent branching from the genus *Dinocras* to the genus *Sinacroneuria*. This group of cryptic species was, therefore, difficult to distinguish.

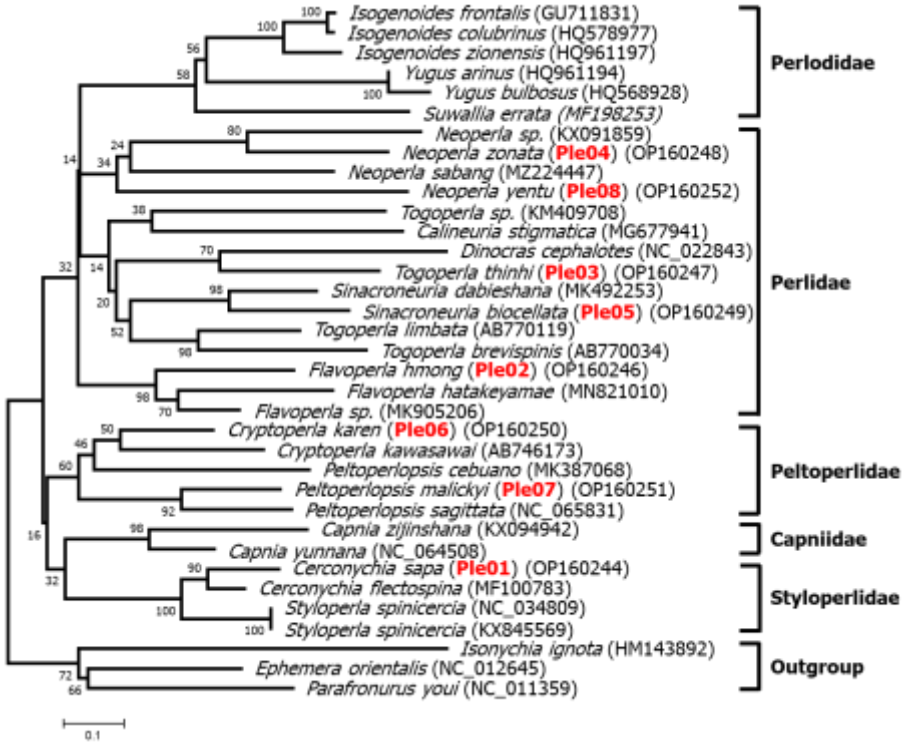

Figure 5. Phylogenetic tree of the Plecoptera order.

The phylogenetic tree based on the *COI* gene of stonefly species from the order Plecoptera was built in MEGA7 using the Maximum Likelihood method. The analysis involved 35 nucleotide sequences and all positions containing gaps and missing data were eliminated; a total of 615 positions were included in the final dataset. Bootstrap values with 1000 replicates are indicated alongside each branch. The scale bar indicates the number of nucleotide changes per site.

In summary, the *COI* gene sequence of the Plecoptera species in Vietnam showed a high divergence between 9.9% and 25.6% from the *COI* sequences of other Plecoptera species (Table 6), which were consistent with the data of endemic species. In the case of the cryptic species, we need further genomic and morphological information to clarify these genetic relationships.

4. Conclusion

In this study, we successfully amplified and sequenced the *COI* gene fragments of eight endemic Plecopteran species in Vietnam with accession numbers registered in Genbank OP160244 and OP160246–OP160252. These *COI* sequences of the Plecoptera species in Vietnam showed a high divergence between 9.9% and 25.6% from the sequences of other Plecoptera species. The phylogeny based on *COI* gene sequences presented that most of our endemic plecopteran species branched with close species in the same genera, except *Togopela thinhi* species grouped in a cryptic cluster of *Togopela* genus. Our research has shown that the analyzed *COI* sequences effectively and accurately represent the endemism exhibited by the studied samples. Therefore, DNA barcoding based on *COI* sequences can be exploited to determine genetic connections among insects and identify endemic species. This study aids in the development of a DNA barcoding database specifically for endemic species found in Vietnam.

Acknowledgements

This study was supported by the Vietnam Ministry of Science and Technology, under project number ĐTĐL.CN-65/19.

References

[1] N. E. Stork, How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth?, Annual Review of Entomology, Vol. 63, 2018, pp. 31-45, http://10.1146/annurev-ento-020117-043348.

[2] K. J. Gaston, Taxonomy of Taxonomists, Nature, Vol. 356, 1992, pp. 281-282, http:// 10.1038/356281a0.

[3] A. Moczek, Phenotypic Plasticity and Diversity in Insects, Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences. Vol. 365, 2010, pp. 593-603, http:// 10.1098/rstb.2009.0263.

[4] J. H. Xiao, N. X. Wang, Y. W. Li, R. W. Murphy, D. G. Wan, L. M. Niu, H. Y. Hu, Y. G. Fu, D. W. Huang, Molecular Approaches to Identify Cryptic Species and Polymorphic Species within a Complex Community of Fig Wasps, PLOS ONE, Vol. 5, 2010, pp. e15067, http://10.1371/journal.pone.0015067.

[5] P. Hebert, A. Cywinska, S. L. Ball, J. Dewaard, Biological Identification Through DNA Barcodes, Proceedings of the Royal Society of London B, Vol. 270, 2003, pp. 313-321, http:// 10.1098/rspb.2002.2218.

[6] O. F. Folmer, M. B. Black, W. R. Hoeh, R. V. Lutz, R. C. Vrijenhoek, DNA Primers for Amplification of Mitochondrial Cytochrome c Oxidase Subunit I from Diverse Metazoan Invertebrates, Molecular Marine Biology and Biotechnology, Vol. 3, 1994, pp. 294-299.

[7] P. Hebert, S. Ratnasingham, J. Dewaard, Barcoding Animal Life: Cytochrome c Oxidase Subunit I Divergences among Closely Related Species, Proceedings, Biological Sciences/The Royal Society, Vol. 270, Suppl 1, 2003, pp. S96-99, http://10.1098/rsbl.2003.0025.

[8] S. Ahmed, DNA Barcoding in Plants and Animals: A Critical Review, 2022.

[9] M. Ashfaq, S. Akhtar, M. A. Rafi, S. Mansoor, P. D. N. Hebert, Mapping Global Biodiversity Connections with DNA Barcodes: Lepidoptera of Pakistan, Plos One, Vol. 12, 2017, pp. e0174749, http:// 10.1371/journal.pone.0174749.

[10] R. Floyd, J. Wilson, P. Hebert, DNA Barcodes and Insect Biodiversity, 2009, pp. 417-431.

[11] R. Fochetti, J. M. T. D. Figueroa, Global Diversity of Stoneflies (Plecoptera; Insecta) in Freshwater, in Freshwater Animal Diversity Assessment, Springer Netherlands: Dordrecht, 2008, pp. 365-377.

[12] S. Ab Hamid, C. Rawi, Application of Aquatic Insects (Ephemeroptera, Plecoptera, And Trichoptera) In Water Quality Assessment of Malaysian Headwater, Tropical Life Sciences Research, Vol. 28, 2017, pp. 143-162, http:// 10.21315/tlsr2017.28.2.11.

[13] K. Meiklejohn, J. Wallman, M. Dowton, DNA Barcoding Identifies all Immature Life Stages of a Forensically Important Flesh Fly (Diptera: Sarcophagidae), Journal of Forensic Sciences, Vol. 58, 2013, pp. 184-187, http:// 10.1111/j.1556-4029.2012.02220.x.

[14] R. Vivero, E. Bejarano, L. Estrada, F. Flórez, E. O. Gomez, Y. Aparicio, C. Torres Gutierrez, S. Uribe, C. Muskus, DNA Barcode for Identification of Immature Stages of Sand Flies (Diptera: Psychodidae) Collected from Natural Breeding Sites, Zootaxa, Vol. 4277, 2017, pp. 228-236, http:// 10.11646/zootaxa.4277.2.3.

[15] T. Korshunova, B. Picton, G. Furfaro, P. Mariottini, M. Pontes, J. Prkić, K. Fletcher, K. Malmberg, K. Lundin, A. Martynov, Multilevel Fine-Scale Diversity Challenges the Cryptic Species Concept, Scientific Reports, Vol. 9, 2019, pp. 6732-6754, http:// 10.1038/s41598-019-42297-5.

[16] I. D. Paterson, R. Mangan, D. Downie, J. Coetzee, M. Hill, A. Burke, P. Downey, T. Henry, S. Compton, Two in One: Cryptic Species Discovered in Biological Control Agent Populations using Molecular Data and Crossbreeding Experiments, Ecology and Evolution, Vol. 6, 2016, pp. 6139-6154, http:// 10.1002/ece3.2297.

[17] D. Hlebec, I. Sivec, M. Podnar, M. Kučinić, DNA Barcoding for Biodiversity Assessment: Croatian Stoneflies (Insecta: Plecoptera), J. Peer, Vol. 10, 2022, pp. e13213, http:// 10.7717/peerj.13213.

[18] J. Sambrook, D. W. Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: New York, 2001.

[19] C. Simon, T. R. Buckley, F. Frati, J. B. Stewart, A. T. Beckenbach, Incorporating Molecular Evolution into Phylogenetic Analysis, and a New Compilation of Conserved Polymerase Chain Reaction Primers for Animal Mitochondrial DNA, Annual Review of Ecology, Evolution, and Systematics, Vol. 37, 2006, pp. 545-579, http://10.1146/annurev.ecolsys.37.091305.110018.

[20] T. A. Hall, BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT, Nucleic Acids Symposium Series, Vol. 41, 1999, pp. 95-98.

[21] S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Molecular Biology and Evolution, Vol. 33, 2016, pp. 1870-1874, http://10.1093/molbev/msw054.

[22] M. Nei, S. Kumar, Molecular Evolution and Phylogenetics, Oxford University Press: Oxford, 2000.

[23] J. Felsenstein, Confidence Limits on Phylogenies: An Approach using the Bootstrap, Evolution, Vol. 39, 1985, pp. 783-791, http //10.1111/j.1558-5646.1985.tb00420.x.