https://doi.org/10.33307/entomon.v49i3.1250
Entomon 49(3): 295-308
Article No. ent. 49301

Molecular identification of ecologically relevant hoverflies (Diptera, Syrphidae) from eastern India

Oishik Kar, Debjani Ghosh, Arka Mukherjee, Koustav Mukherjee, Debdeep Pramanik, Saikat Sarkar, Jayita Sengupta, Atanu Naskar* and Dhriti Banerjee

Diptera Section, Zoological Survey of India, New Alipore, Kolkata 700053, West Bengal, India. Email: atanudiptera@gmail.com

ABSTRACT: Syrphid identification keys that cover all life cycle stages of the taxonomy are insufficient, and there are problems with the morphological identification of these flies. Cytochrome oxidase I (COI) is widely used for molecular identification and phylogenetic reconstruction. The study examined the effectiveness of COI in identifying 18 specimens containing 13 agriculturally important species of syrphids collected from different geo-climatic regions of West Bengal. Phylogenetic analysis was performed using Maximum Likelihood (ML) and Bayesian (BA) trees, which were almost congruent. Barcodes were generated for *Dasysyrphus orsua* and *Eristalinus polychromata* for the first time. This is the first study to use the COI for barcoding ecologically and agriculturally relevant syrphid flies from eastern India and their phylogeny. The findings contribute to the basic understanding of the diversity of syrphids across West Bengal and the molecular characterization of hoverflies, promoting their conservation and thus leading to the augmentation of crops.

KEY WORDS: Flower flies, phylogenetic analysis, COI gene, genetic divergence, barcodes

INTRODUCTION

Issues such as pollution, global warming, urbanization, and industrialization, as well as current farming procedures, are causing intense harm to classic supporting functions such as pollination (Klein *et al.,* 2007). Most plants, especially commercial crops, require pollination to reproduce. Many animals perform this ecological role. Numerous species of plants would be pushed to extinction if this service did not exist, and many current crops might be challenging to cultivate (Abrol, 2012). Pollination is thought to be responsible for up to 75 per cent of the production of food from agriculture (Klein *et al.,* 2007). During the last few years, there has been an enormous decline in the ratio of insect pollinators across the world (Rhodes, 2019). India has many endemic species and is ranked sixth among the world's 12 biodiversity hotspots (Singh and Chaturvedi, 2017). With over 6,000 known species worldwide, divided into 300 separate groups (Skevington *et al.,* 2019), the Syrphidae (hoverflies or flower flies) family of Diptera is one of the most diverse and well-known to the people (Courtney *et al.,* 2017). Adult hoverflies are essential for pollinating flowering plants (Free, 1993; Richards *et al.,* 1997). Syrphids are one of the most common families of flies, having significant potential as the ecosystem's first-line pollinator (Owen and Gilbert, 1989), especially in certain landscapes

* Author for correspondence

where hymenopteran pollination competence is at risk of depletion because they rely on endothermically produced energy (Milièiæ *et al.,* 2017). Thus, there is a need for second-line pollinators like hoverflies for these types of ecosystems. Another characteristic of the syrphid flies is that their larvae consume various foods (Knutson *et al.,* 1975). Some larvae are saprotrophs, while others are insectivores, feeding on Hemipterans (aphids, adelgids, psyllids), as well as Thysanopterans (Thrips) and other plant-sucking insects (Brunetti, 1907, 1908; Thompson and Ghorpadé 1992; Tenhumberg, 1995; Rojo *et al.,* 2003; Bugg *et al.,* 2008; Sengupta *et al.,* 2018). A few syrphid larvae can reduce weed growth (Rizza *et al.,* 1988).

Flower flies are a cosmopolitan family that can be found in most biomes, except deserts, high-altitude tundra, and Antarctica (Sutherland *et al.,* 2001). Some hoverfly species have a restricted distributional range, i.e. they are indigenous to a specific habitat or location, whilst others have a wide distribution across numerous countries (Owen and Gilbert, 1989). Hoverflies are usually distinguishable from other flies by a spurious vein or Vena spuria that runs parallel to the fourth longitudinal wing vein (Vockeroth, 1992), however, there are exceptions (e.g. syrphid flies of the genera *Graptomyza* and *Paragodon*) (Thompson, 1969). Four subfamilies are currently recognized (Eristalinae, Microdontinae, Pipizinae and Syrphinae) (Mengual *et al.,* 2015), but some authors (Speight, 1987; Thompson, 1969, 1972) have split off the basal clades of Syrphidae, recognizing a separate family Microdontidae. There are 202 genera and 96 subgenera of Syrphidae currently recognized, grouped into 13 tribes and 12 subtribes (Thompson, 1972; Vockeroth, 1992; Young *et al.,* 2016). In India, 357 species from 69 genera have been identified (Ghorpadé, 2014). Several of the species are indigenous to India.

Morphological taxonomic keys require entomological expertise to identify species, as many species have a similar appearance and are difficult to distinguish (Achint and Singh, 2021). Identification markers such as wing venation, eye color pattern, color patterns of legs, setae of thorax and abdomen, and their color for specimens are frequently degraded during storage and collection techniques such that morphological identification of syrphids is difficult and time-consuming. Another disadvantage of morphological identification is the lack of keys to all life stages, although taxonomic keys for adult syrphids are well documented. Alternative strategies to address these challenges include DNA barcoding (Hogg and Hebert, 2004). A 350-700bp of the mitochondrial cytochrome c oxidase I (cox1) (Hebert *et al.,* 2003a, b) is used as a proper approach for identifying worldwide biota (Waugh, 2007). Using a standardized DNA locus for DNA barcoding has become a popular and effective way of differentiating species (Achint and Singh, 2021; Bajaj *et al.,* 2023). For the correct identification of numerous groups to the species level (http://www.ibol.org/resources/) as well as species complexes, a brief sequence of standardized COI gene mitochondrial DNA has been employed and recognized (Tyagi *et al.,* 2017). The current study aimed to test the COI gene to correctly identify these pollinating hoverflies from the different geo-climatic zones of West Bengal with respect to morphology-based identification procedures.

MATERIALS AND METHODS

Our survey for the collection was carried out in different regions of West Bengal. It was chosen as our study due to its vastness and diversity as it includes the different geo-climatic regions, namely the hilly regions, arid region, the Gangetic plains, and coastal areas (Maity *et al.,* 2016). West Bengal is also one of the leading states in terms of agricultural crop production in our country. Hence, a COI barcode database for these crop-friendly pollinators was needed for easy identification and conservation. The study used 54 specimens from 19 species of 9 genera representing distinct syrphid subfamilies. During the years 2020–2021, 18 specimens containing 13 species were collected from various districts of West Bengal (Murshidabad, Kolkata, South 24 Parganas, Kalimpong) including the vulnerable Sagar Islands (Table 1) (Fig 1.). Their DNA sequences were submitted to GenBank, while sequences for the

remaining species were acquired from the Genbank database (NCBI). We surveyed extensively in this one year, covering all three seasons, namely pre-monsoon, monsoon, and post-monsoon, where these 13 species of hoverflies were found in all the seasons. Fly specimens were collected by sweep net collection from flowering vegetation and agricultural lands. The specimens were preserved in high-grade-ethyl alcohol (70%). After that, part of the collected specimens was dried and pinned with entomological pins, and after morphological identification with stereomicroscope, specimens were deposited in the designated repository of the National Zoological Collections (NZC), Zoological Survey of India, Kolkata. The necessary specimen photographs were obtained with a Leica stereo-iso microscope M205A, a Leica DFC 500 camera, and the Leica Application Suite LASv 3.6 software. According to Systema Dipterorum, valid species names were allocated (Evenhuis and Pape, 2022).

The maps for this paper were created using Arc GIS® Desktop software (version 10.8) by ESRI after registering the geographical coordinates of the collection sites in Garmin GPS device (Fig. 1).

DNA extraction, Polymerase Chain Reaction (PCR), and DNA sequencing:

Genomic DNA (gDNA) was extracted from individual fly specimens using the QIAmp DNA extraction kit (Qiagen). The whole procedure was done according to the manufacturer's instructions. Voucher specimens were submitted in the Diptera Section of the ZSI, Kolkata. The amount of DNA was recorded on a Qubit Fluorometer (Life Technologies, USA), and the extracted DNA was kept at -20°C for subsequent analysis. Using primers- forward LCO-1490 (F) (GGT CAA CAA ATC ATA AAG ATA TTG G) and reverse HCO-2198 (R) (TAA ACT TCA GGG TGA CCA AAA AAT CA), roughly 20 ng genomic DNA was utilized

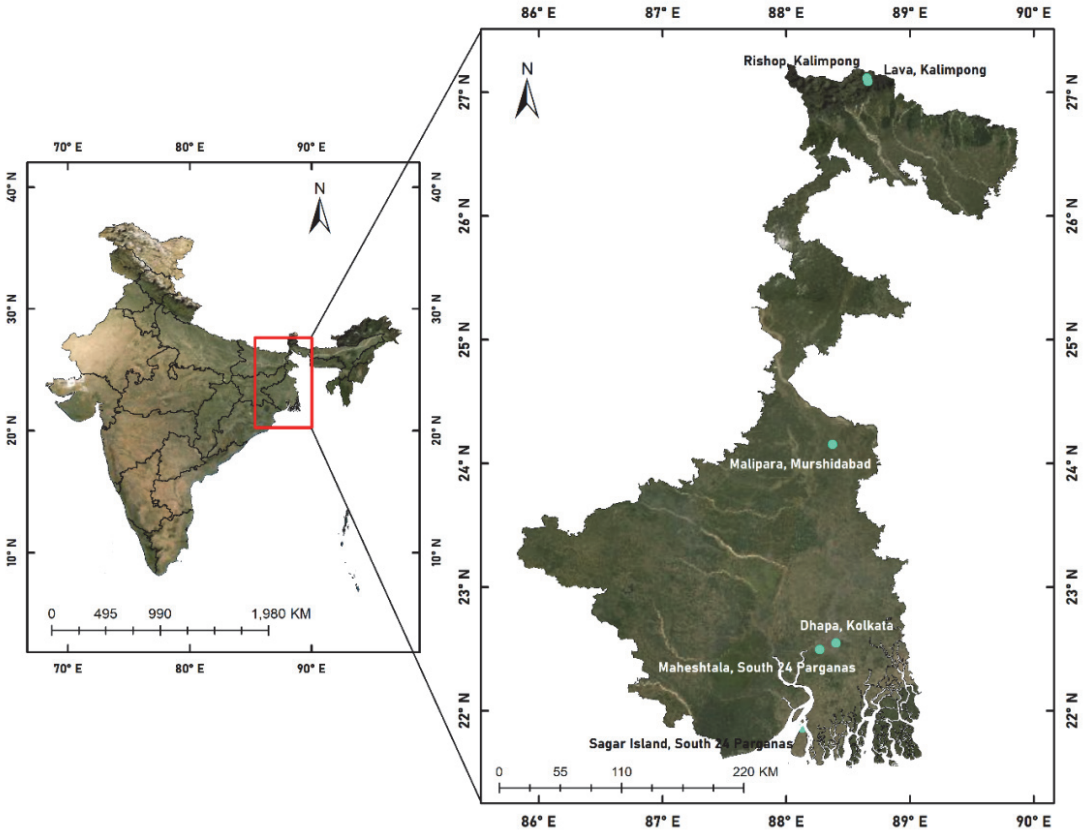

Fig 1. Sampling localities of collected syrphid flies from different geo-climatic regions of West Bengal

Table 1 Analyzed syrphid samples, with sampling locations from West Bengal, GenBank accession numbers, species names, and collection dates

No.	Locality	Accession no.	Species name	Date of collection	Coordinates
1	Murshidabad, Malipara	ON421581	*Episyrphus balteatus*	08/03/2020	24.15341 N; 88.37381 E
2	Murshidabad, Malipara	ON421583	*Ep. balteatus*	08/03/2020	24.15341 N; 88.37381 E
3	Murshidabad, Malipara	ON421584	*Episyrphus balteatus*	08/03/2020	24.15341 N; 88.37381E
4	South 24 Parganas, Maheshtala	ON210051	*Er. quinquestriatus*	21/10/2021	22.496025 N; 88.2703667 E
5	South 24 Parganas, Maheshtala	ON217545	*Er. quinquestriatus*	21/10/2021	22.496025 N; 88.2703667 E
6	South 24 Parganas, Maheshtala	ON226501	*Er. polychromata*	21/10/2021	22.496025 N; 88.2703667 E
7	Murshidabad, Malipara	ON421642	*Er. polychromata*	08/03/2020	24.15341 N; 88.37381 E
8	Kolkata, Dhapa	ON248238	*Er. arvorum*	09/11/2021	22.54769167N;88.40233611E
9	Kalimpong, Rishop	ON260958	*Eristalis cerealis*	27/11/2021	27.1123 N; 88.65324 E
10	Kalimpong, Rishop	ON248443	*Eristalis tenax*	27/11/2021	27.1123 N; 88.65324 E
11	Kalimpong, Lava	ON209555	*Ersitalis himalayensis*	18/10/2021	27.0863 N; 88.6615 E
12	Kalimpong, Lava	ON222740	*Eupeodes luniger*	18/10/2021	27.0863 N; 88.6615 E
13	Kalimpong, Lava	ON261094	*Dasysyrphus orsua*	18/10/2021	27.0863 N; 88.6615 E
14	South 24 Parganas, Sagar Island	ON422271	*Dideopsis aegrota*	05/10/2021	21.86216 N; 88.12949 E
15	Murshidabad, Malipara	ON421526	*Ischiodon scutellaris*	08/03/2020	24.15341 N; 88.37381 E
16	South 24 Parganas, Sagar Island	ON440975	*I. scutellaris*	17/02/2021	21.86216 N; 88.12949 E
17	Murshidabad, Malipara	ON421461	*Melanostoma orientale*	08/03/2020	24.15341 N; 88.37381 E
18	Murshidabad, Malipara	ON421571	*Paragus crenulatus*	08/03/2020	24.15341 N; 88.37381 E

to amplify about 700 base pairs from the 5' end of the mitochondrial cytochrome c oxidase subunit I (COI) gene (Folmer *et al.,* 1994). PCR was carried out in a 50µl total reaction volume comprising 20 Pico moles of each primer, 100 mM KCl, 20 mM Tris–HCl (pH 8.0), 1 mM DTT, 0.1 mM EDTA, 2.0 mM MgCl2, 0.25 mM of each dNTP, primer cocktail, and 1U of Taq polymerase (Takara BIO Inc., Japan) with the following cycling parameters: 5 min at 94^0C; followed by 40 cycles of 30 s at 94^0C, 40 s at 53^0C, 1 min at 72^0C and final extension for5 min at 72^0C. To confirm the amplicon size, the amplified products were seen in a 1 per cent agarose gel, stained with SYBR@safe DNA gel dye, and imaged on a safe gel imager (Invitrogen). The QIAquick Gel Extraction Kit (Qiagen) was used to purify the PCR-amplified products according to the manufacturer's instructions. For cycle sequencing, about 15 ng of purified PCR product was utilized. Cycle sequencing was performed on an ABI thermal cycler using the BigDye®Terminator ver. 3.1Cycle Sequencing Kit (Applied

Biosystems, Inc) with both forward and reverse PCR primers using the following parameters: 96°C for 1 min, then 25 cycles of 96°C for 10 s, 50°C for 5 s, and a final extension at 60°C for 1 min15 s. After cycle sequencing, the products were cleaned with the BigDye X-terminator kit (Applied Biosystems Inc.) and placed into an ABI 3730 capillary Genetic analyzer at the Zoological Survey of India sequencing laboratory (Banerjee *et al.,* 2015; Tyagi *et al.,* 2017; Achint and Singh, 2021).

Sequence analysis and dataset formation: MEGA X was used to manually edit the sequences from each specimen (Kumar *et al.,* 2018). All sequences were matched to identical reported sequences in the NCBI database utilizing the BLAST (https://blast.ncbi.nlm.nih.gov) algorithm (Chakraborty *et al.,* 2019). The ORF finder of NCBI (https://www.ncbi.nlm.nih.gov/orfnder/gorf.html) is used to examine the accurate amino acid codes devoid of any stop codon or indels (insertion or deletions). Each sequence was uploaded to GenBank library, and unique accession numbers were issued to each one (Table 1). A total of 54 specimens from 19 species belonging to 9 genera from different syrphid subfamilies were included in this investigation (both Indian and global sequences). The 9 genera include *Ischidon* containing 1 species, *Dasysyrphus* containing 3 species, *Eupeodes* containing 1 species, *Episyrphus* containing 1 species, *Dideopsis* containing 1 species, *Melanostoma* containing 1 species, *Paragus* containing 2 species, *Eristalis* containing 4 species, and *Eristalinus* containing 5 species. A member of the putative Syrphidae sister-group, Pipunculidae, was added as outgroup (Ståhls *et al.,* 2003). 36 sample sequences from 55 taxa were downloaded from NCBI, Genbank, including the outgroup. Those 36 samples includes *Ischiodon scutellaris* KY845775 (Pakistan), MK771152 (Bangladesh), KY846329 (Pakistan); *Dasysyrphus amalopis* JX828010 (Canada), JX828112 (Canada); *Dasysyrphus pauxillus* MZ610653 (Finland), MZ629684 (Finland); *Eupeodes luniger* KT959887 (Finland), KY834510 (Pakistan), MW077802 (France); *Episyrphus balteatus* OL765264 (India), MN973969 (India), *Dideopsis aegrota* MW473976 (Canada); *Melanostoma orientale* KY839783 (Pakistan), KY837293 (Pakistan), KT175592 (India); *Paragus serratus* MG194422 (India), KY837201 (Pakistan); *Paragus crenulatus* JN298750 (Canada), JF872389 (Canada); *Eristalis himalayensis* OL442159 (India); *Eristalis tenax* OL441830 (India), MN967351 (India), MN967352 (India); *Eristalis cerealis* OK465106 (India), OK287112 (India); *Eristalis arbustorum* JN269860 (Canada), MN868856 (Portugal); *Eristalinus arvorum* MK751019 (Germany), MK751022 (Germany), MK751021 (Germany); *Eristalinus aeneus* MW473968 (Canada); *Syrphidae sp.* KY841659 (Pakistan); *Eristalinus pari*a OK655827 (India), OK444104 (India); *Eristalinus sp.* MK771154 (Bangladesh); Pipunculidae KR506987 (Canada) (outgroup).

Genetic divergence and cluster analysis:

Initially, sequences were aligned (multiple sequence alignment) in MEGA X software via the ClustalW algorithm (Kumar *et al.,* 2018). To avoid any form of coherent outcomes, the dataset is constructed to be 663 base pairs long. The genetic divergence between and within taxonomic groups was estimated in MEGAX using the Kimura-2-parameter (Kimura, 1980; Kumar *et al.*, 2018). The best-fit nucleotide substitution model was determined using JModelTest v2.1.10 (Darriba *et al.,* 2012) through CIPRES server (Miller *et al.,* 2010) to discover a suitable evolutionary model for the syrphid flies dataset based on the Bayesian Information Criterion (BIC). Models with the lowest BIC scores (-5650.248) were considered to describe the substitution pattern the best (Nei and Kumar, 2001). The GTR+G+I model was selected for the syrphid COI dataset. MEGA X was used to investigate nucleotide substitution and nucleotide composition data (Kumar *et al.,* 2018).

The COI dataset has been used to construct the phylogenetic trees based on the Maximum Likelihood (ML) algorithm (Fig. 2). The dataset was designed and analyzed in IQ-TREE on XSEDE (2.1.2v) (Nguyen *et al.,* 2015; Minh *et al.,* 2020) via the CIPRES website (Miller *et al.,* 2010), employing 1,000 bootstrapping tests and default parameter settings (Siriwut *et al.,* 2021). The

FigTree v1.4.4 software (http://tree.bio.ed.ac.uk/software/fgtree/) was used to edit the resultant files. This offered a graphic depiction of the specimen's sequencing divergence. The Bayesian (BA) tree (Fig. 3) was generated in Mr. Bayes v3.2.7a with nst=6 for the GTR+G+I model utilizing metropolis-coupled Markov Chain Monte Carlo (MCMC) and run for 5,000,000 generations with 25 per cent burn-in and trees saved every 100 generations. The posterior probability was used to determine branch support (PP). The web-based iTOL v6 program (https://itol.embl.de/) was used to create a tree from the produced files, which aided visual display. Haplotypes calculations were done in DnaSP v5.10 (Librado and Rozas, 2009).

The PTP model (Zhang *et al.,* 2013) was utilized for species delimitation, which defines species based on the number of substitutions in the phylogenetic tree changing. For the species delimitation study, the BA tree file in Newick format was submitted to the bPTP server (https://species.h-its.org/html). The robustness of species delimitations is estimated using Bayesian support values. A higher bootstrap value at the node indicates that the terminal node is more certain to belong to a specific species. The PTP analysis was run for 500,000 MCMC (Markov Chain Monte Carlo) generations with a thinning value of 100 and a burn-in of 25 per cent, and the outgroup was excluded.

RESULTS AND DISCUSSION

DNA sequences:

The dataset includes 663bp of the cox1 gene of ecologically and agriculturally essential syrphids of 19 species under 9 genera namely *Ischidon*, *Dasysyrphus*, *Eupeodes*, *Episyrphus*, *Dideopsis*, *Melanostoma*, *Paragus*, *Eristalis*, and *Eristalinus*. The conserved, variable, and parsimoniously informative sites for the studied species were examined. The data set contains 55 sequences with 663 base pairs, 299 distinct patterns (variable sites), 201 parsimony-informative sites, 34 singleton variable sites, and 428 constant (conserved) sites. Hence, this study demonstrates that the COI gene is highly conserved.

Base composition and nucleotide substitution:

Both the nucleotide sequence and the specific nucleotide percentage were examined in this study since both characteristics are significant for evaluating variation among different species in MEGA X (Kumar *et al.,* 2018). The nucleotide base composition of the sequenced fly species collected by us showed that ON421583 (*Episyrphus balteatus*) and ON421581 (*Episyrphus balteatus*) have the highest (AT) percentage (70.6%), while ON260958 (*Eristalis cerealis*) had highest (GC) percentage (31.7%). On the other hand, ON421642 (*Eristalinus polychromata*) had the lowest (AT) percentage (67.9%) and ON421583 (*Episyrphus balteatus*) and ON421581 (*Episyrphus balteatus*) have lowest (GC) percentage (29.4%) (Table 2). The average nucleotide frequencies are 30.53 (A), 38.64 (T/U), 14.72 (C), and 16.11 per cent (G). This clearly reveals that in the nucleotide sequences (A + T) content was higher than (G + C). Thus it proves that insect mtDNA has a higher (A + T) frequency (Lunt and Hyman, 1997).

Evolutionary analyses and divergences were examined in MEGA X (Kumar *et al.,* 2018). To visualize the characterization of genetic variations of different species, sequences were downloaded from GenBank in FASTA format (Achint and Singh, 2021). Transition/transversion rate ratios were k1 = 2.002 (purines) and k2 = 1.915 (pyrimidines). The overall transition/transversion bias was R = 0.833, {where R = [A*G*k1 + T*C*k2]/[(A+G) * (T+C)]} as calculated by the Maximum Composite Likelihood method in MEGA X. This result shows that the rate of transitions was higher than the rate of transversions. Nucleotide Substitution patterns were calculated (Table 3). Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated (complete deletion option) (Tamura *et al.,* 2004; Kumar *et al.*, 2018; Ghosh *et al.,* 2022).

Haplotypes and haplotype diversity: Further research indicated the presence of 16 haplotypes in 18 sampled syrphid fly specimens, representing a significantly high degree of haplotype diversity. Most of samples showed unique haplotypes (Table 4).

Sequence divergence: To ensure that the COI gene is reliable for identification, intraspecific divergences should not exceed 3 per cent and interspecific divergences should be greater than 3 per cent (Wells and Sperling, 2001). A method of DNA barcode analysis for species discrimination has been developed that contrasts intraspecific and interspecific genetic divergence (bar-coding gap) (Hebert *et al.,* 2004; Ghosh *et al.,* 2022). For species recognition in insects, a 3 per cent interspecific genetic divergence limit was advised to close the barcoding gap (Hebert *et al.,* 2003a). Later, it was proposed that intraspecific and interspecific genetic divergences vary between taxa, and that a universal cutoff does not exist (Meyer and Paulay, 2005). Intraspecific divergence was found to be 0.00-5.00 per cent. The highest intraspecific divergence was seen in *Eristalinus polychromata* (5%) and *Eristalis himalayensis* (5%) while lowest intraspecific divergence was shown by *Ischiodon scutellaris* (0%), *Eristalinus arvorum* (0%), *Eristalinus paria* (0%), *Eristalinus arbustorum* (0%), *Eristalis tenax* (0%), *Dideopsis aegrota* (0%), *Episyrphus balteatus* (0%), *Paragus crenulatus* (0%), *Melanostoma orientale* (0%), *Eupeodes luniger* (0%), *Dasysyrphus pauxillus* (0%), and *Dasysyrphus amalopis* (0%). The intraspecific divergence for *Eristalinus quinquestriatus* was 1 and *Eristalis cerealis* was 2 per cent. The intraspecific genetic distance of *Eristalinus polychromata* and *Eristalis himalayensis* showed a value of more than 3 per cent, indicating a possible overlap of intra and interspecific divergences. The interspecific divergence ranged from 0.8 to 19.6 per cent. The lowest interspecific divergence (0.8%) was recorded between *Dasysyrphus pauxillus* and *Dasysyrphus amalopis* (Meyer and Paulay, 2005). Interspecific divergences less than 3 per cent in Diptera have previously been considered as evidence of species complexes or cryptic species. Additional evidence regarding the delimitation of *Dasysyrphus amalopis* and *Dasysyrphus pauxillus* is needed to be investigated, including genes other than COI (Banerjee *et al.,* 2015).

Phylogenetic analysis:

In the Maximum Likelihood (ML) tree (Fig. 2), the two main clades correspond to the subfamilies Eristalinae and Syrphinae. The first clade of the Eristalinae subfamily contained species of the genera *Eristalinus* and *Eristalis*, while in the second clade of subfamily Syrphinae, species of the genera *Paragus*, *Melanostoma*, *Dideopsis*, *Episyrphus*, *Eupeodes*, *Dasysyrphus*, and *Ischiodon* were present. The intraspecific genetic distance of *Eristalinus quinquestriatus* was less than 3 per cent (1%) and the global sequence MK771154 *Eristalinus sp.* formed the same clade with ON210051 *Eristalinus quinquestriatus* and ON217545 *Eristalinus quinquestriatus*. MK77154 is a likely sequence of Bangladesh sample of *Eristalinus quinquestriatus* probably. Sequence ON260958 fell in the same clade as OK465106 *Eristalis cerealis* and OK287112 *Eristalis cerealis*, both from India. The bootstrap value of the clade was 100 per cent. We conclude that ON260958 is *Eristalis cerealis* although morphological identification of the specimen was complicated by its poor condition. Clades with >90% bootstrap support are considered as strongly supported. The genera *Ischiodon, Melanostoma* and *Paragus* are strongly supported with 100 per cent bootstrap support. Some species developed their unique conspecific cluster due to geographical differences in the collected samples.

In the Bayesian (BA) tree (Fig. 3), it is evident that the two subfamilies distinguished from one another with a very high support value of 0.99 at the deep node, showing complete congruence with the ML tree and the observations done by other researchers (Mengual *et al.,* 2015; Mengual, 2020; Moran *et al.*, 2022) done on Palaearctic and Nearctic regions respectively. The deep node value supporting each of the subfamilies show Eristalinae to be monophyletic (0.99) but Syrphinae to be a non-monophyletic clade (0.716) which is supported by the work done by Mengual *et al.* (2008).

There are five described tribes under the subfamily Syrphinae: Bacchini, Melanostomini, Paragini, Syrphini, and Toxomerini. The studied sample set

Table 2 The nucleotide base composition of the sequenced fly species using MEGAX

Species	Position	T	C	A	G	AT (%)	GC (%)
ON440975 *Ischiodon scutellaris*	1st	49.5	1.0	48.1	1.5	69.9	30.1
	2nd	26.6	15.5	27.5	30.4		
	3rd	43.0	25.6	15.0	16.4		
ON421526 *I. scutellaris*	1st	49.4	1.3	48.8	0.6	69.4	30.5
	2nd	28.3	14.5	27.7	29.6		
	3rd	41.5	28.3	12.6	17.6		
ON261094 *Dasysyrphus orsua*	1st	49.1	1.8	49.1	0.0	69.4	30.6
	2nd	27.7	16.9	26.5	28.9		
	3rd	42.2	27.1	13.9	16.9		
ON222740 *Eupeodes luniger*	1st	54.5	1.9	42.7	0.9	69.6	30.4
	2nd	27.4	15.6	26.4	30.7		
	3rd	43.4	25.0	14.6	17.0		
ON421583 *Episyrphus balteatus*	1st	50.7	0.9	48.3	0.0	70.6	29.4
	2nd	28.9	14.7	25.6	30.8		
	3rd	43.6	25.1	14.7	16.6		
ON421584 *Ep. balteatus*	1st	48.8	0.6	50.6	0.0	70.4	29.6
	2nd	31.2	14.7	24.7	29.4		
	3rd	42.4	26.5	13.5	17.6		
ON421581 *Ep. balteatus*	1st	49.1	0.6	50.3	0.0	70.6	29.4
	2nd	31.4	14.8	24.9	29.0		
	3rd	42.6	26.6	13.6	17.2		
ON422271 *Dideopsis aegrota*	1st	54.2	0.0	45.8	0.0	70.4	29.5
	2nd	28.3	13.3	29.2	29.2		
	3rd	41.2	26.1	12.6	20.2		
ON421461 *Melanostoma orientale*	1st	49.7	2.7	47.0	0.7	68.9	31.1
	2nd	26.8	16.8	28.9	27.5		
	3rd	43.0	28.9	11.4	16.8		
ON421571 *Paragus crenulatus*	1st	47.1	0.8	51.3	0.8	69.4	30.6
	2nd	26.1	16.0	29.4	28.6		
	3rd	40.7	27.1	13.6	18.6		
ON209555 *Eristalis himalayensis*	1st	46.4	5.3	47.4	1.0	69.1	30.9
	2nd	27.3	15.3	28.7	28.7		
	3rd	42.6	25.4	14.8	17.2		
ON248443 *Er. tenax*	1st	42.2	6.2	51.7	0.0	69.1	30.8
	2nd	27.8	15.1	28.3	28.8		
	3rd	42.9	25.0	14.6	17.5		
ON260958 *Er. cerealis*	1st	42.6	5.9	51.0	0.5	68.3	31.7
	2nd	26.0	14.7	28.9	30.4		
	3rd	43.9	24.9	12.7	18.5		
ON248238 *Er. arvorum*	1st	44.9	5.3	49.8	0.0	69.8	30.3
	2nd	29.0	14.5	27.5	29.0		
	3rd	43.0	25.1	15.0	16.9		

ON421642 *Er. polychromata*	1st 2nd 3rd	40.6 28.2 41.8	6.5 17.1 27.1	52.4 27.1 13.5	0.6 27.6 17.6	67.9	32.2
ON226501 *Er. polychromata*	1st 2nd 3rd	46.2 29.6 41.4	3.0 16.0 27.2	50.9 27.2 13.6	0.0 27.2 17.8	69.7	30.4
ON217545 *Er. quinquestriatus*	1st 2nd 3rd	45.6 30.2 41.4	2.4 15.4 27.2	51.5 27.2 13.6	0.6 27.2 17.8	69.9	30.2
ON210051 *Er. quinquestriatus*	1st 2nd 3rd	46.9 27.4 42.9	2.4 15.6 25.5	49.8 27.4 14.6	0.9 29.7 17.0	69.7	30.4

(All frequencies are given in percentage)

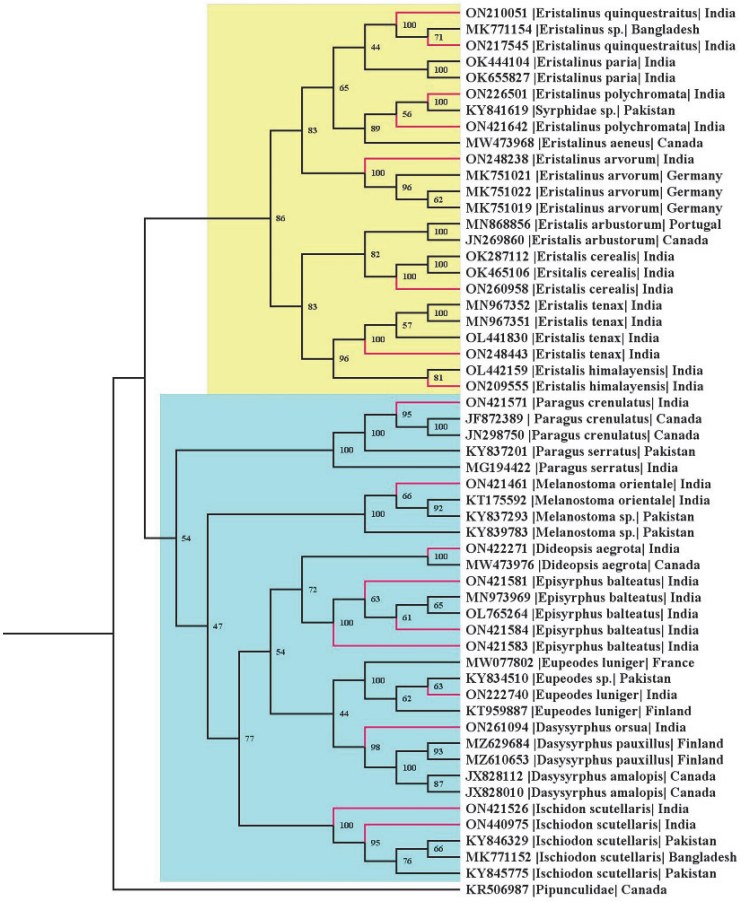

Fig. 2 Phylogenetic Tree (Maximum-likelihood) from collected syrphid COI (pink branch) dataset. IQ-TREE on XSEDE (v2.1.2) generated 1000 bootstrapped ML (GTR + G+I) tree of syrphid flies based on COI gene. Numbers indicate bootstrap values from ML analysis. Pipunculidae (Diptera) was used as an out-group

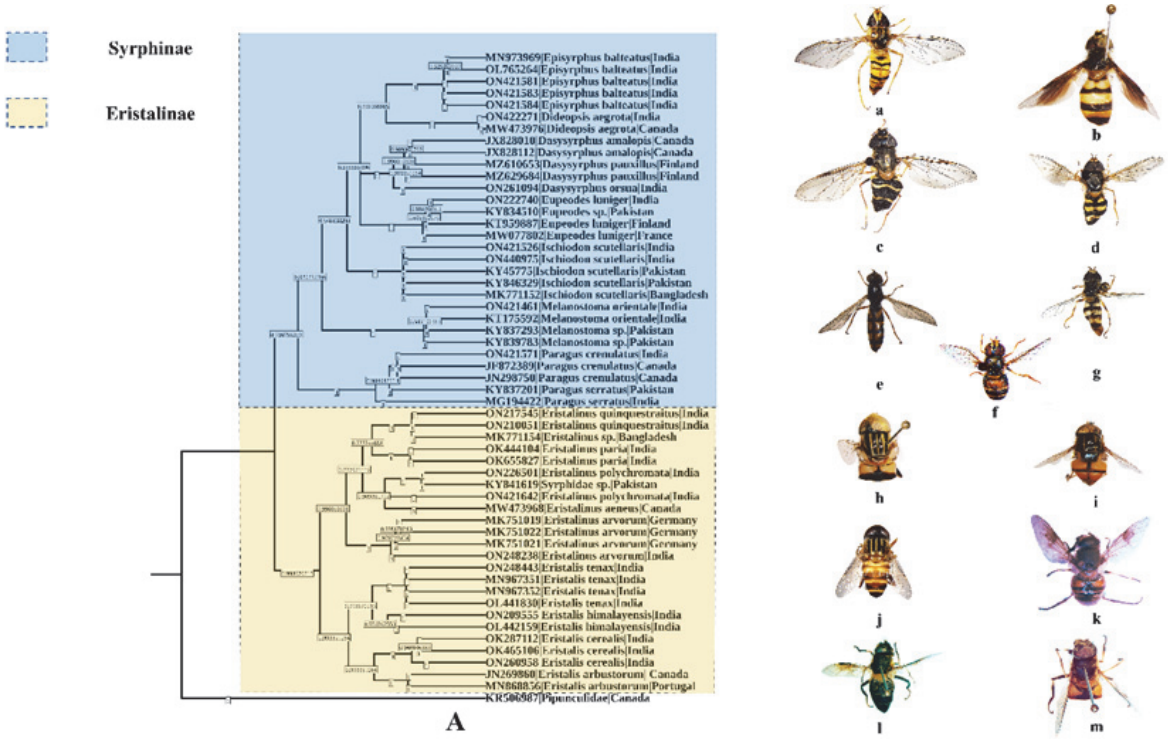

Fig. 3 The Bayesian (BA) Tree of 54 hoverflies sequences with posterior probabilities from Bayesian analysis. Pipunculidae has been used as outgroup. The different species are shown in this figure namely- a. *Episyrphus balteatus*, b. *Dideopsis aegrota*, c. *Dasysyrphus orsua*, d. *Eupeodes luniger*, e. *Melanostoma orientale*, and f. *Paragus crenulatus*. g. *Ischiodon scutellaris*, h. *Eristalinus quinquestriatus*, i. *Eristalinus polychromata*, j. *Eristalinus arvorum*, k. *Eristalis tenax*, l. *Eristalis himalayensis,* m. *Eristalis cerealis*

contains taxon sampling under three of these tribes: Melanostomini, Paragini, and Syrphini.

Melanostomini comprises one genus in our dataset with *Melanostoma orientale* that forms a sister clade with the tribe Paragini which is a monotypic taxon with only one genus *Paragus*. Under the tribe, Syrphini, *Episyrphus*, and *Dideopsis* form a sister clade with 0.8 BI (Bayesian Inference) support, whereas, *Dasysyrphus* and *Eupeodes* form a large polytomy with the previous group. The lower value at the support branch in the deep nodes creates confusion about the exact relationship of the said groups from an oriental perspective.

The three species of the sample set under the genus *Dasysyrphus* are resolved together, showing prominent monophyly, although the genus shows significant variation in male genitalia and larval character (Mengual *et al.,* 2008).

Under the subfamily Eristalinae, there are multiple tribes, and our studied taxons fall under the tribe Eristalini and subtribe Eristalina. In the BA tree, the two genera under consideration form definitive monophyly with a branch support of 0.99, whereas

Table 3. Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

	A	T	G	C
A	-	9.77	3.72	**8.16**
T	7.72	-	**7.13**	4.07
G	7.72	**18.7**	-	4.07
C	**15.45**	9.77	3.72	-

Each entry shows the probability of substitution (r) from one base (row) to another base (column).
For simplicity, the sum of r values is made equal to 100.
Rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics.

each of the genera also forms monophyly at the species level, having a branch support of 0.996 for the genus *Eristalinus* and 0.993 for genus *Eristalis*. At the species level, *Eristalis cerealis* and *Eristalis arbustorum* form monophyly with each other and are of the closest proximation. On the other hand, *Eristalis tenax* and *Eristalis himalayensis* form another monophyly with a support of 0.9. The other sister clade shows subclades formed by *Eristalinus quinquestriatus* and *Eristalinus paria* and *Eristalinus polychromata* and *Eristalinus aeneus*. The previous shows non-monophyly at the recent divergence, whereas the latter shows monophyly. The whole subclade forms a monophyletic clade with *Eristalis arvorum*.

At the species level, several taxa show branching and variation with low BI values, like *Episyrphus balteatus*, *Eupeodes luniger*, and *Melanostoma orientale*, so it is hard to assign a definitive relationship between these taxa and their conspecifics. According to the bPTP analysis, there are 42 putative species in the study set, which can be due to the varied geographical distribution and the presence of haplotypes.

The COI gene can be used to identify Indian syrphid flies and to determine their phylogeny. Because of its variety, the barcoding gene COI is an excellent marker for phylogenetic study and confirming geographical population distribution patterns. It is crucial to identify and segregate the data of the agriculturally important syrphid flies to understand their effect in the agricultural and food industries. We anticipate that the barcode data provided by this research may aid in rapidly identifying these syrphid flies in both their adult and larval stages. This is the first-ever study that provided the barcoding of the syrphid flies from West Bengal, along with the first-ever barcoding of the COI gene of *Dasysyrphus orsua* and *Eristalinus polychromata*. It is critical to accurately quantify the pollinator diversity to uphold existing variety and prevent further reduction in their population, and barcoding them allows us to identify them even if skilled taxonomists are not present. Correctly identifying these pollinating insects is critical in developing methods for conserving them and augmentation programs in crops to increase production. Hence, this study and the combination of additional molecular markers and morphological and ecological data would be helpful in better characterization and understanding them in the study site on a larger scale in the future.

Table 4. Number of specimens and mitochondrial haplotypes in collected syrphid flies

No.	Species	specimens	haplotypes
1	*Ischiodon scutellaris*	2	2
2	*Dasysyrphus orsua*	1	1
3	*Eupeodes luniger*	1	1
4	*Episyrphus balteatus*	3	1
5	*Dideopsis aegrota*	1	1
6	*Melanostoma orientale*	1	1
7	*Paragus crenulatus*	1	1
8	*Ersitalis himalayensis*	1	1
9	*Eristalis tenax*	1	1
10	*Eristalis cerealis.*	1	1
11	*Eristalinus arvorum*	1	1
12	*Eristalinus polychromata*	2	2
13	*Eristalinus quinquestriatus*	2	2

ACKNOWLEDGMENTS

The authors are thankful to the Director, Zoological Survey of India, MOEFCC, Govt. of India for her enthusiasm, moral support, and providing necessary facilities. The authors are grateful to Dr. Vikas Kumar, Scientist–E, and Dr. Kamoud Tyagi, Scientist-D of ZSI for giving them the facilities to work on molecular systematics. This work was supported by the Zoological Survey of India (ZSI), Ministry of Environment, Forest & Climate Change (MOEFCC) under Grant Number: 282/2022. There are no conflicting interests stated by the authors.

REFERENCES

Abrol D.P. (2012) Pollination biology: Biodiversity Conservation and Agricultural Production Vol. 792. Springer. xxix+792pp. doi:10.1007/978-94-007-1942-2.

Achint R. and Singh D. (2021) Application of COI gene for identification of some economically and

forensically important muscid flies of India (Diptera: Muscidae). International Journal of Tropical Insect Science 41(4): 3023–3029. doi:10.1007/s42690-021-00494-8.

Bajaj K., Chhuneja P.K., Mohindru B. and Singh J. (2023) Molecular Characterization of Pollinators in Cotton Ecosystem. Indian Journal of Entomology 1–6. doi:10.55446/IJE.2023.1073.

Banerjee D., Kumar V., Maity A., Ghosh B., Tyagi K., Singha D., Kundu S., Laskar B.A., Naskar A. and Rath S. (2015) Identification through DNA barcoding of Tabanidae (Diptera) vectors of surra disease in India. Acta Tropica 150: 52–58. doi:10.1016/j.actatropica.2015.06.023.

Brunetti E. (1907) Notes on Oriental Syrphidae. Part I. Records of the Indian Museum 1: 379–380.

Brunetti E. (1908). IX Notes on Oriental Syrphidae with descriptions of new species. Part 1. Records of the Indian Museum 2: 49–96.

Bugg R.L., Colfer R.G., Chaney W.E., Smith H.A. and Cannon J. (2008) Flower Flies (Syrphidae) and Other Biological Control Agents for Aphids in Vegetable Crops. University of California, Agriculture and Natural Resources. 25pp. doi:10.3733/ucanr.8285.

Chakraborty R., Singha D., Kumar V., Pakrashi A., Kundu S., Chandra K., Patnaik S. and Tyagi K. (2019) DNA barcoding of selected *Scirtothrips* species (Thysanoptera) from India. Mitochondrial DNA Part B 4(2): 2710–2714. doi:10.1080/23802359.2019.1644547.

Courtney G.W., Pape T., Skevington J.H. and Sinclair B.J. (2017) Biodiversity of Diptera. In: Insect Biodiversity: Science and Society Vol. 2 (Eds. Foottit R.G. and Adler P.H.), Blackwell Publishing, Oxford. pp229-278. doi:10.1002/9781118945568.ch9.

Darriba D., Taboada G.L., Doallo R. and Posada D. (2012) jModelTest 2: More models, new heuristics and parallel computing. Nature Methods 9(8): 772. doi:10.1038/nmeth.2109.

Evenhuis N.L. and Pape T. (2022) Systema Dipterorum, Version 3.7. http://www.diptera.org/ (Accessed on 11 November, 2022).

Folmer O., Black M., Hoeh W., Lutz R. and 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.

Free J.B. (1993) Insect pollination of crops. Insect Pollination of Crops. 2nd Edition. Academic Press, London. 684 pp.

Ghorpadé K. (2014) An updated Check-list of the Hover-flies (Diptera-Syrphidae) recorded in the Indian subcontinent. Colemania 44: 1–30.

Ghosh D., Kar O., Pramanik D., Mukherjee A., Sarkar S., Mukherjee K., Naskar A. and Banerjee D. (2022) Molecular identification and characterization of Muscid flies (Diptera: Muscidae) of medico-veterinary importance from the Gangetic plains of Eastern India. International Journal of Tropical Insect Science 48: 3759–3769. doi:10.1007/s42690-022-00900-9.

Hebert P.D.N., Cywinska A., Ball S.L. and DeWaard J.R. (2003a) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270(1512): 313–321.

Hebert P.D.N., Ratnasingham S. and DeWaard J.R. (2003b) Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society B: Biological Sciences 270(1): S96–S99. doi:10.1098/rsbl.2003.0025.

Hebert P.D.N., Stoeckle M.Y., Zemlak T.S. and Francis C.M. (2004) Identification of Birds through DNA Barcodes. PLOS Biology 2(10): 312. doi:10.1371/journal.pbio.0020312.

Hogg I. and Hebert P. (2004) Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. Canadian Journal of Zoology 82. doi: 10.1139/Z04-041.

Kimura M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.

Klein A.M., Vaissière B.E., Cane J.H., Steffan-Dewenter I., Cunningham S.A., Kremen C. and Tscharntke T. (2007) Importance of pollinators in changing landscapes for world crops. Proceedings of the Royal Society B: Biological Sciences 274(1608): 303–313. doi:10.1098/rspb.2006.3721.

Knutson L., Thompson F. and Vockeroth J. (1975) Family Syrphidae. In: *A catalog of the Diptera of the oriental region* Vol. II (Eds. Delfinado M.D. and Hardy D.E.), The University Press of Hawaii, Honolulu. pp307–374.

Kumar S., Stecher G., Li M., Knyaz C. and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics

Analysis across Computing Platforms. Molecular Biology and Evolution 35(6): 1547–1549. doi.10.1093/molbev/msy096.

Librado P. and Rozas J. (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. doi:10.1093/bioinformatics/btp187.

Lunt D.H. and Hyman B.C. (1997) Animal mitochondrial DNA recombination. Nature 387(6630): 247. doi:10.1038/387247a0.

Maity A., Naskar A., Hazra S., Sengupta J., Parui P., Homechaudhuri S. and Banerjee D. (2016) New distributional records of Tabanidae (Insecta: Diptera) from different geo-climatic regions of West Bengal, India. Journal of Entomology and Zoology Studies 4: 1291–1298.

Mengual X. (2020) Phylogenetic relationships of the bacchine flower flies (Diptera: Syrphidae) based on molecular characters, with a description of a new species of *Melanostoma* (Schiner, 1860). Contributions to Zoology 89(2): 210–244. doi:10.1163/18759866-20191410.

Mengual X., Ståhls G. and Rojo S. (2008) First phylogeny of predatory flower flies (Diptera, Syrphidae, Syrphinae) using mitochondrial COI and nuclear 28S rRNA genes: Conflict and congruence with the current tribal classification. Cladistics 24(4): 543–562. doi:10.1111/j.1096-0031.2008.00200.x.

Mengual X., Ståhls G. and Rojo S. (2015) Phylogenetic relationships and taxonomic ranking of pipizine flower flies (Diptera: Syrphidae) with implications for the evolution of aphidophagy. Cladistics 31(5): 491–508. doi:10.1111/cla.12105.

Meyer C.P. and Paulay G. (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biology 3(12): e422. doi:10.1371/journal.pbio.0030422.

Milièiæ M., Vujiæ A., Jurca T. and Cardoso P. (2017) Designating conservation priorities for Southeast European hoverflies (Diptera: Syrphidae) based on species distribution models and species vulnerability. Insect Conservation and Diversity 10(4): 354–366. doi:10.1111/icad.12232.

Miller M.A., Pfeiffer W. and Schwartz T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE) 1–8. doi:10.1109/GCE.2010.5676129.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A. and Lanfear R. (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37(5): 1530–1534. doi:10.1093/molbev/msaa015.

Moran K.M., Skevington J.H., Kelso S., Mengual X., Jordaens K., Young A.D., Ståhls G., Mutin V., Bot S., van Zuijen M., Ichige K., van Steenis J., Hauser M. and van Steenis W. (2022) A multigene phylogeny of the eristaline flower flies (Diptera: Syrphidae), with emphasis on the subtribe Criorhinina. Zoological Journal of the Linnean Society 194(1): 120–135. doi:10.1093/zoolinnean/zlab006.

Nei M. and Kumar S. (2001) Molecular Evolution and Phylogenetics. Heredity 86(3): 385–385. doi:10.1046/j.1365-2540.2001.0923a.x.

Nguyen L.T., Schmidt H.A., von Haeseler A. and Minh B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32(1): 268–274. doi:10.1093/molbev/msu300.

Owen J. and Gilbert F.S. (1989) On the Abundance of Hoverflies (Syrphidae). Oikos 55(2): 183–193. doi:10.2307/3565422.

Rhodes C.J. (2019) Are insect species imperilled? Critical factors and prevailing evidence for a potential global loss of the entomofauna: A current commentary. Science Progress 102(2): 181–196. doi:10.1177/0036850419854291.

Richards A.J., Proctor M., Yeo P. and Lack A. (1997) The natural history of pollination. Harper Collins New Naturalist, London. Archives of Natural History 24(2): 307–307. doi:10.3366/anh.1997.24.2.307.

Rizza A., Campobasso G., Dunn P.H. and Stazi M. (1988) *Cheilosia corydon* (Diptera: Syrphidae), a Candidate for the Biological Control of Musk Thistle in North America. Annals of the Entomological Society of America 81(2): 225–232. doi:10.1093/aesa/81.2.225.

Rojo S., Gilbert F., Marcos-García M.A., Nafria J.M.N. and Mier Durante M.P. (2003) A World Review of Predatory Hoverflies (Diptera, Syrphidae: Syrphinae) and their Prey. Centro Iberoamericano de la Biodiversidad. 319pp.

Sengupta J., Naskar A., Maity A., Homechaudhuri S. and Banerjee D. (2018) Distributional scenario of hover flies (Diptera: Syrphidae) from the state of West Bengal. Munis Entomology & Zoology 13(2): 447–457.

Singh J.S. and Chaturvedi R. (2017) Diversity of Ecosystem Types in India: A Review. Proceedings of the Indian National Science Academy 83(2): 569–594. doi:10.16943/ptinsa/2017/41287.

Siriwut W., Jeratthitikul E., Panha S., Chanabun R., Ngor P.B. and Sutcharit C. (2021) Evidence of cryptic diversity in freshwater *Macrobrachium* prawns from Indochinese riverine systems revealed by DNA barcode, species delimitation and phylogenetic approaches. PLOS ONE 16(6): e0252546. doi:10.1371/journal.pone.0252546.

Skevington J., Young A., Locke M. and Moran K. (2019) New Syrphidae (Diptera) of North-eastern North America. Biodiversity Data Journal 7: e36673. doi:10.3897/BDJ.7.e36673.

Speight M.C.D. (1987) External morphology of adult Syrphidae (Diptera). Tijdschrift voor Entomologie 130: 141–175.

Ståhls G., Hippa H., Rotheray G., Muona J. and Gilbert F. (2003) Phylogeny of Syrphidae (Diptera) inferred from combined analysis of molecular and morphological characters. Systematic Entomology 28(4): 433–450. doi:10.1046/j.1365-3113.2003.00225.x.

Sutherland J., Sullivan M. and Poppy G. (2001) Distribution and abundance of aphidophagous hoverflies (Diptera: Syrphidae) in wildflower patches and field margin habitats. Agricultural and Forest Entomology 3(1): 57–64. doi:10.1046/j.1461-9563.2001.00090.x.

Tamura K., Nei M. and Kumar S. (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences 101(30): 11030–11035.

Tenhumberg B. (1995) Syrphids as natural enemies of cereal aphids in Germany: Aspects of their biology and efficacy in different years and regions. Agriculture, Ecosystems & Environment 52(1): 39–43. doi:10.1016/0167-8809(94)09007-T.

Thompson F.C. (1969) A New Genus of Microdontine Flies (Diptera: Syrphidae) With Notes on the Placement of the Subfamily. Psyche 76(1): 74–85. doi:10.1155/1969/62102.

Thompson F.C. (1972) A contribution to a generic revision of the Neotropical Milesinae (Diptera: Syrphidae). Arquivos de Zoologia 23(2): 73-215.

Thompson F.C. and Ghorpadé K. (1992) A new coffee aphid predator, with notes on other Oriental species of *Paragus* (Diptera: Syrphidae). Colemania 5: 1-24.

Tyagi K., Kumar V., Singha D., Chandra K., Laskar B.A., Kundu S., Chakraborty R. and Chatterjee S. (2017) DNA Barcoding studies on Thrips in India: Cryptic species and Species complexes. Scientific Reports 7(1): 4898. doi:10.1038/s41598-017-05112-7.

Vockeroth J. (1992) The flower flies of the subfamily Syrphinae of Canada, Alaska and Greenland: Diptera: Syrphidae. Agriculture Canada, Ottawa. 456pp.

Waugh J. (2007) DNA barcoding in animal species: Progress, potential and pitfalls. BioEssays 29(2): 188–197. doi:10.1002/bies.20529.

Wells J.D. and Sperling F.A. (2001) DNA-based identification of forensically important Chrysomyinae (Diptera: Calliphoridae). Forensic Science International 120(1–2): 110–115. doi:10.1016/s0379-0738(01)00414-5.

Young A.D., Lemmon A.R., Skevington J.H., Mengual X., Ståhls G., Reemer M., Jordaens K., Kelso S., Lemmon E.M., Hauser M., Meyer M.D., Misof B. and Wiegmann B.M. (2016) Anchored enrichment dataset for true flies (order Diptera) reveals insights into the phylogeny of flower flies (family Syrphidae). BMC evolutionary biology 16: 1–13. doi:10.1186/s12862-016-0714-0.

Zhang J., Kapli P., Pavlidis P. and Stamatakis A. (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29(22): 2869–2876. doi:10.1093/bioinformatics/btt499.

(Received April 04, 2024; revised ms accepted June 15, 2024; published September 30, 2024)