Molecular characterization and genetic divergence of seven Culex mosquito (Diptera: Culicidae) species using Mt COI gene from Odisha State, India

Deepika Panda1 and Tapan Kumar Barik1,2*

Abstract
Background: Culex mosquitoes are involved in the transmission of arboviral diseases worldwide. Bio-ecology and identification of mosquitoes are of paramount importance to develop species-specific vector control strategies. Genetic-based species-specific approaches that reduce the burden of vector-borne diseases are made possible by molecular techniques.

Results: In the present study, the mitochondrial COI gene of Culex mosquitoes was used for molecular identification in addition to morpho-taxonomy. Our findings suggested the presence of important Culex mosquito vectors viz, Culex vishnui, Culex tritaeniorhynchus, Culex bitaeniorhynchus, Culex quinquefasciatus, Culex gelidus, Culex fuscocephala, and Culex fuscanus in the southern part of Odisha state, India. We examined the phylogeny and genetic diversity of the above seven different Culex populations from different geographical locations. An average intra-specific K2P distance of the COI gene was found to be 0.9%. Further, to measure the diversity of the Culex population among different geographical strains, haplotype diversity and nucleotide diversity were compared. Culex fuscanus showed high polymorphisms and mutations with high nucleotide diversity (0.013) and the Culex quinquefasciatus showed the lowest variation in Pi(0), 0.0013 in the intra-population polymorphism analysis of COI sequences. Similarly, the Haplotype diversity (Hd) found in Culex gelidus and Culex fuscocephala with the value of 0.972 and Culex quinquefasciatus (0.583) showed the lowest value of haplotype diversity. A haplotype network was constructed to establish the genealogical relationship between haplotypes. The phylogenetic tree was constructed that produces distinctive conspecific clusters in different Culex species. Population genetic study has illustrated the occurrence of genetic differentiation within the population.

Conclusions: The findings of this study contribute to greater evidence that DNA barcode sequences can be used to monitor mosquito species diversity. This study also adds valuable information about the systematics and molecular biology of seven public health important mosquito species acting as a significant vector for Japanese encephalitis in various Asian continents. This information is further used for the effective implementation of region-specific vector control strategies.

Keywords: Culex, Mosquito, Molecular phylogeny, Genetic diversity, India

*Correspondence: tkbarik@rediffmail.com
1 Department of Zoology, Berhampur University, Berhampur, Odisha 760007, India
Full list of author information is available at the end of the article

Background
Mosquitoes act as obligate intermediate hosts for several diseases that collectively represent a significant cause of human morbidity and mortality worldwide.
The family Culicidae is a large and abundant group that occurs throughout temperate and tropical regions of the world, and well beyond the Arctic Circle. Culicinae is the biggest subfamily of the Culicidae family, having 3,081 no. of species in 110 genera and clubbed into 11 tribes. Approximately 3,554 different types of mosquito species are reported globally (Harbach, 2018), comprising a total of 112 genera, of which nearly three-quarters are native to the humid tropics and subtropical region. In relation to mosquito fauna diversity, India occupied ranks fifth globally (Foley et al., 2007) and includes 393 species in 49 genera and 41 subgenera (Bhattacharyya et al., 2014) out of which the online systematic catalog of Culicidae included 356 mosquito species (Gaffigan et al., 2014).

Culex is one of the largest and most important groups with 770 species divided between 26 subgenera (Harbach, 2016) that act as prime vectors to spread the pathogens to cause diseases like West Nile fever, St. Louis encephalitis, Japanese encephalitis, equine encephalitis, acute encephalitis syndrome, filariasis, etc. Therefore, effective vector control, early detection, as well as ongoing vector surveillance are crucial in the battle against these arboviral diseases. The success of vector control activities, on the other hand, is dependent on the accurate identification of targeted mosquito species, along with the proper understanding of their biology and ecology. Besides, many taxonomists also consider combined behavioral and population biology data in identifying and classifying a species.

Mosquito vector recognition has traditionally relied on the description and study of morphological features. However, mosquito species identification has been more challenging in recent years due to a substantial decrease in the number of taxonomists and other professional experts in species identification. Furthermore, the core morphological characters of mosquito vectors are often damaged at the time of collection and storage, while some are not available in all stages of their development. Besides this, the morphological features used to distinguish adult specimens differ very little between species that only professional mosquito taxonomists can normally distinguish between them.

Taking this into consideration, several researchers have sought to come up with new approaches to optimize and improve the method of species discrimination, making it more easily accessible to non-experts. In addition, sophisticated genetic and biotechnological technologies should be used to obtain a more accurate and complete picture of mosquito diversity in the study region (Attaullah et al., 2021). A significant breakthrough in DNA-sequencing technology has encouraged researchers to investigate more biodiversity in identifying and cataloguing organisms using a convenient and productive DNA analysis tool termed DNA barcoding. This tool has been used as a platform that not only improved access to taxonomic information but also reinforced collaborations for all those interested in biodiversity recording, learning, and restoration. In recent times, this DNA barcoding approach employed to compare and differentiate closely related species by amplifying and sequencing a small conserved region of DNA. This technique has been utilized in various taxonomic investigations involving dipteran taxa (Pramual et al., 2011; Stahls et al., 2009) including mosquito vectors (Gonzalez et al., 2010; Laboudi et al., 2011; Ruiz et al., 2010; Ruiz-Lopez et al., 2012). Molecular taxonomic studies on local strains of Canada, India, Pakistan, Persian Gulf region and China (Ashfaq et al., 2014; Azari-Hamidian et al., 2010; Cywinska et al., 2006; Kumar et al., 2007; Wang et al., 2012) were carried out, showing the use of DNA barcoding technique in Culicidae family.

For DNA barcoding analysis, a short DNA segment that is universally available in the target lineages and has enough sequence variance to differentiate species and is being used for DNA barcoding (Hebert et al., 2003). A mitochondrial cytochrome oxidase subunit I gene (mt COI), can be used to distinguish vertebrate and invertebrate individual species (Folmer et al., 1994). Barcoding with the COI gene is also suggested as a standard for cryptic taxa discovery, the association of different life stages of the same species, and wildlife conservation genetics, in addition to the identification of recognized and new species (Trivedi et al., 2016). The phylogenetic signal of COI tends to be stronger than that of the other mitochondrial genes (Strüder-Kypke & Lynn, 2010).

Integrating conventional taxonomy with molecular DNA barcoding techniques could be an important approach for assessing insect vectors and their genetic variations. Further, Genetic variation and population structure analysis are significant features of population genetics in the mosquito vector. Comparative analyses of these factors can therefore aid in elucidating the factors that influence their population levels. Several factors, such as climate change, ecological conditions, natural barriers, anthropogenic activity, and migration can influence the genetic diversity of species. Therefore, understanding the diversity, distribution of mosquito species and their population genetic study is essential for the control of mosquito vectors and mosquito-borne diseases.

Odisha state of India contributes about 261 acute encephalitis syndrome (AES) cases which are about 4.5% of total cases in India and reported one death case in 2021. Likewise, 18 Japanese encephalitis cases, i.e., about 2.4% of total cases in India (NVBDCP, 2021) in 2021. Out of thirty districts of this state, twenty districts
are reported as endemic for lymphatic filariasis (NVB-DCP, 2021). Earlier, Hazra and Dash (1998) reported 21 species belonging to the Culicinae subfamily from the coastal region of Odisha state. Rajavel et al. (2005a, 2005b) reported 74 species belonging to 12 genera from Jeypore Hill tracks. Similarly, Dash and Hazra (2011) reported 22 species under 6 genera from the Puri and Khurda districts of Odisha. 24 Anopheles species were reported from the Ganjam district (Dash, 2014). Relatively higher humidity and availability of stagnant water bodies, throughout the year in southern districts, provide suitable breeding sites for mosquitoes. Therefore, the mosquito faunal diversity in this region is fascinating. Various changes in ecological environments caused by massive deforestation, periodic cyclones, and widespread use of insecticides have occurred over the last 3 decades. However, there is no sufficient recent data on the Culex mosquito diversity of the Southern districts of Odisha state and therefore, we have considered this region as our study area. DNA barcodes and genetic variation analysis of Culex mosquito species found in the studied region will aid in the detection and surveillance of mosquito vectors, assisting in the control of mosquito-borne disease outbreaks. The main objective of this study was to use both morpho-taxonomy and molecular taxonomy to assess the Culex faunal diversity of different regions of the southern districts in Odisha state. Furthermore, using a single mitochondrial COI marker gene, the efficiency of DNA barcoding in quickly identifying and analyzing genetic variation among seven Culex mosquito species was examined.

Methods
Collection and rearing of mosquitoes
Different developmental stages of the mosquito species were collected from various breeding sites in the Southern districts of Odisha state (Fig. 1 and Table 1). Preferences of study sites were given to those areas having earlier records of epidemic conditions. Mosquitoes were collected through various standard methods and reared in an insectary available at the Applied Entomology Laboratory; Post Graduate Department of Zoology, Berhampur University. The immature stages of mosquitoes were reared to adults in the insectary with standard protocol. The adult mosquitoes were used for morpho-taxonomy followed by molecular taxonomy.

Morphological identification
Morphology-based identification of adult mosquitoes was made as per the identification keys described by
Barraud (1934), Reuben et al. (1994) and Tyagi et al. (2015). After morphological identification, the mosquito samples were vouchered and stored in the laboratory. Further, to confirm the morphology-based identification, genomic DNA isolation and gene amplification was performed for molecular taxonomy.

### Table 1 Collection of *Culex* mosquitoes from various breeding sites at different developmental stages

| Sl. No | Name of the mosquito species | Stage of Mosquitoes collected | Month of collection | No. of sample collected | Type of habitat | Name of the collection site/place | Co-ordinates | Distance from the study centre (km) |
|--------|------------------------------|-------------------------------|--------------------|------------------------|----------------|----------------------------------|--------------|-----------------------------------|
| 1      | *Cx. quinquefasciatus*      | Adult                          | January            | 7                      | Indoor         | Berhampur (Ganjam)               | 19.3150 N 84.7941 E | 12 |
| 2      | *Cx. quinquefasciatus*      | Adult                          | January            | 3                      | Indoor         | Balipadar (Ganjam)               | 19.7370 N 84.7074 E | 68 |
| 3      | *Cx. quinquefasciatus*      | Larvae                         | July               | 5                      | Outdoor        | Balugaon (Khurda)                | 20.1783 N 85.1129 E | 72 |
| 4      | *Cx. vishnui*               | Adult                          | July               | 2                      | Outdoor        | Balugaon (Khurda)                | 20.1783 N 85.1129 E | 72 |
| 5      | *Cx. quinquefasciatus*      | Adult                          | August             | 3                      | Indoor         | Koraput                          | 18.8135 N 82.7123 E | 327 |
| 6      | *Cx. vishnui*               | Adult                          | August             | 2                      | Outdoor        | Koraput                          | 18.8135 N 82.7123 E | 327 |
| 7      | *Cx. gellidus*              | Adult                          | September          | 1                      | Outdoor        | Jeypore (Koraput)                | 18.8606 N 82.5510 E | 351 |
| 8      | *Cx. tritaeniorhynchus*     | Adult                          | October            | 1                      | Outdoor        | Sunabeda (Koraput)               | 18.7299 N 82.8317 E | 327 |
| 9      | *Cx. quinquefasciatus*      | Adult                          | October            | 4                      | Outdoor        | Sunabeda (Koraput)               | 18.7299 N 82.8317 E | 327 |
| 10     | *Cx. vishnui*               | Adult                          | November           | 3                      | Outdoor        | Sonapur (Ganjam)                 | 19.1133 N 84.7753 E | 40  |
| 11     | *Cx. bitaeniorhynchus*      | Adult                          | November           | 1                      | Outdoor        | Sonapur (Ganjam)                 | 19.1133 N 84.7753 E | 40  |
| 12     | *Cx. gellidus*              | Adult                          | November           | 2                      | Outdoor        | Berhampur (Ganjam)               | 19.3150 N 84.7941 E | 12  |
| 13     | *Cx. tritaeniorhynchus*     | Adult                          | December           | 3                      | Outdoor        | Berhampur (Ganjam)               | 19.3150 N 84.7941 E | 12  |
| 14     | *Cx. fuscanus*              | Adult                          | December           | 5                      | Outdoor        | Berhampur (Ganjam)               | 19.3150 N 84.7941 E | 12  |
| 15     | *Cx. fuscocephala*          | Adult                          | December           | 8                      | Outdoor        | Bhanjibihar (Ganjam)             | 19.3017 N 84.8830 E | 0   |
| 16     | *Cx. bitaeniorhynchus*      | Adult                          | December           | 2                      | Outdoor        | Bhanjibihar (Ganjam)             | 19.3017 N 84.8830 E | 0   |
| 17     | *Cx. vishnui*               | Larvae                         | January            | 5                      | Outdoor        | Rayagada                         | 19.1712 N 83.4163 E | 218 |
| 18     | *Cx. tritaeniorhynchus*     | Larvae                         | January            | 4                      | Outdoor        | Rayagada                         | 19.1712 N 83.4163 E | 218 |
| 19     | *Cx. quinquefasciatus*      | Larvae                         | January            | 7                      | Outdoor        | Paralakhemundi (Gajapati)         | 18.7783 N 84.0936 E | 125 |
| 20     | *Cx. vishnui*               | Larvae                         | January            | 5                      | Outdoor        | Paralakhemundi (Gajapati)         | 18.7783 N 84.0936 E | 125 |
| 21     | *Cx. tritaeniorhynchus*     | Adult                          | February           | 3                      | Indoor         | Chatrapur (Ganjam)               | 19.3597 N 84.9887 E | 17  |
| 22     | *Culex gellidus*            | Adult                          | February           | 2                      | Outdoor        | Chatrapur (Ganjam)               | 19.3597 N 84.9887 E | 17  |

**Genomic DNA isolation and quantification**

DNA was isolated by the Bender Buffer method (Collins et al., 1987) from the morphologically identified adult mosquito. Qualitative, as well as quantitative measurements of isolated DNA, were carried out by 1% Agarose Gel Electrophoresis and by a Nanodrop spectrophotometer (Thermo Scientific, USA), respectively.
Gene amplification and sequencing
The extracted DNA from all the specimens was subjected to PCR amplification using universal primers as suggested by Folmer et al. (1994), targeting the mitochondrial COI gene. The primers used for COI gene amplification are:

- Forward: LCO1490 5’-GGTCAAAAAATCATAAAG ATATTGGG-3’
- Reverse: HCO2198 5’-TAAACTTCAGGGTGACCA AAAAAATCAA-3’

The reaction mixture of the COI gene amplification consisting of 1X PCR buffer, 0.5 uTaq DNA, 2.5 mM MgCl₂, 200 µM dNTPs, 10 pmol of each primer, 100 pmol template DNA, total dilution was made up to 25 µl. The thermo-cycling program for the COI gene consisted of 95 °C for 5 min; 35 cycles of 95 °C for 30 s, 51 °C for 30 s and 72 °C for 1 min. A final extension step of 7 min at 72 °C was added after cycling. The amplicons were resolved in 1.5% agarose gel using 1x Tris-Acetate-EDTA (TAE) buffer. The PCR products were sequenced commercially by the Sanger method using ABI Prism 3730XL Big Dye Terminator V3.1 cycle sequencer (Applied Biosystem, USA).

Sequence data analysis
The trace files of the generated COI sequences were trimmed and assembled using Geneious prime 2020.1.1 (www.geneious.com) software and sequences of low quality were excluded during data analysis. The generated nucleotide sequences from each species were compared with those publicly available barcode sequences on NCBI using the BLASTn tool. The final assembled good-quality sequences were submitted to NCBI to get the accession numbers. A total of 68 Culex COI nucleotide sequences (including the sequences generated in the present study) of all the studied species were retrieved from the GenBank database and used for data analysis. To understand the evolutionary relationships among the studied Culex mosquito species, the generated sequences were subjected to phylogenetic analysis.

Genetic divergence
The Clustal W algorithm was implemented in the software package MEGA X (Kumar et al., 2018) for the multiple sequence alignment (MSA). The transition/transversion (ts/tv) bias (R) and intra-specific and inter-specific pairwise sequence divergence was inferred using the Tamura-3 parameter (T92) and Kimura 2 parameter model (K2P) (Kimura, 1980), respectively. Furthermore, 68 no. of the nucleotide sequences of the studied Culex species were used to determine the intra-population polymorphism using DnaSp v6.12.03, and the number of variable sites, nucleotide diversity, haplotype diversity, parsimony informative sites were also analyzed.

Haplotype network
Haplotype networks of all the studied Culex mosquito species were constructed based on the number of COI nucleotide differences, using the Median-joining network algorithm in Popart software to determine the genealogical inter-relationship between haplotypes.

Phylogenetic analysis
The Maximum-Likelihood (ML) tree of all the studied mt COI genes was constructed in MEGA X. The support for nodes in ML analysis was assessed by bootstrapping with 1000 replicates. The substitution pattern is best described by the model with the lowest Bayesian Information Criterion (BIC) values (Nei & Kumar, 2000). The T92+G model (Tamura 3 Parameter model with Gamma distribution parameter) was adopted for all the three studied gene sequences. 1st + 2nd + 3rd + non-coding codon positions were included in the study and all the positions containing gaps and missing data were eliminated. The phylogenetic tree analysis includes 68 Culex COI nucleotide sequences (both the generated during the current study and retrieved sequences from the database). A sequence from Armigeres subalbatus was used as the out-group.

Results
Morpho-taxonomy based identification
Both mature and immature stages of mosquito species were collected and immature stages were reared up to adults for identification. However, the adult mosquitoes were identified directly based on the morpho-taxonomy and further confirmed by molecular taxonomy using DNA barcoding. Here, we identified seven different Culex mosquito species such as Culex vishnui, Culex tritaeniorhynchus, Culex bitaeniorhynchus, Culex quinquefasciatus, Culex gelidus, Culex fuscoccephala, and Culex fuscatus (Which is a homotypic synonym of Lutiza fuscana).

Computational analysis of sequence data
A total of 7 generated COI sequences were used in the current study from seven Culex mosquito species. The generated sequences were found to be AT-rich with an average of 69.2% and an average GC content of 30.8 % (Table 2). Further, none of the generated sequences contain insertion, deletion, or stop codon, which is supporting the origin of the mitochondrial gene. The transition and transversion bias (R) of seven Culex mosquito
species were found to be 0.81. The no. of transition between C&T was equal to A&G (i.e., 16.97). The number of transversion between A&T in the sequences was higher (i.e., 8.91) as compared to G & C (i.e., 3.90), which might be due to high A+T content in mitochondrial nucleotide sequences.

The pattern of genetic divergence

The genetic divergence among the selected species of Culex mosquitoes was determined and all the species displayed discriminative estimations of inter and intra-specific divergence. The intra-specific K2P distance of the COI gene was < 2% (0.02), with an average of 0.9% (0.009) as mentioned in Table 3. The maximum mean intra-specific distance was observed in Culex fuscanus (1.5%) followed by Culex fuscoccephala (1.2%), while Culex quinquefasciatus had a minimum intra-specific K2P distance of 0.002 (0.2%). The inter-specific K2P genetic distances of the COI gene among all the species were more than 2% (0.02), with an average of 8.1% (0.081). The highest K2P distance was observed between Culex gelidus and Culex fuscocephala (10.3%) and the least was between sister species Culex tritaeniorhynchus and Culex vishnui (5.6%) (Table 3).

Haplotype network analysis

For haplotype network analysis, the median-joining network method was performed based on the similarity and differences in the COI gene sequences of all the seven Culex species (Fig. 2). Each sphere in the network diagram represents distinct haplotypes and this number is consistent with the result obtained from the haplotype numbers shown in Table 4. The haplotype network of Cx. quinquefasciatus and Cx. bitaeniorhynchus show less no. of the sphere which was represented by four and three haplotypes, respectively. All other Culex mosquito species (such as Cx. fuscoccephala, Cx. gelidus, Cx. tritaeniorhynchus, Cx. vishnui, Cx. fuscanus) show a complex

| SL. No | Species            | COI Accession No | Sequence length (bp) | % of AT content | % of GC content |
|--------|--------------------|------------------|----------------------|----------------|----------------|
| 1      | Culex vishnui      | MT876103         | 706                  | 69.3           | 30.7           |
| 2      | Culex tritaeniorhynchus | MT878243       | 633                  | 68.1           | 31.9           |
| 3      | Culex quinquefasciatus | MT919714       | 656                  | 68.7           | 31.3           |
| 4      | Culex fuscoccephala    | MT919696       | 706                  | 68.7           | 31.3           |
| 5      | Culex fuscanus       | MW045493        | 598                  | 69.9           | 30.1           |
| 6      | Culex gelidus        | MT919713        | 658                  | 69.5           | 30.5           |
| 7      | Culex bitaeniorhynchus| MT919715       | 602                  | 70.1           | 29.9           |

Table 3 Inter-specific (below the diagonal) and the average intra-specific distances (along the diagonal in bold) of COI sequences

| Culex vishnui | Culex tritaeniorhynchus | Culex quinquefasciatus | Culex fuscoccephala | Culex fuscanus | Culex gelidus | Culex bitaeniorhynchus |
|---------------|-------------------------|------------------------|---------------------|----------------|---------------|-------------------------|
| Culex vishnui | 0.007                   |                        |                     |                |               |                         |
| Culex tritaeniorhynchus | 0.056               | 0.011              |                     |                |               |                         |
| Culex quinquefasciatus | 0.084             | 0.072               | 0.002              |                |               |                         |
| Culex fuscoccephala   | 0.079            | 0.089               | 0.079              | 0.012          |               |                         |
| Culex fuscanus        | 0.072            | 0.091               | 0.089              | 0.087          | 0.015         |                         |
| Culex gelidus         | 0.089            | 0.078               | 0.096              | 0.103          | 0.077         | 0.006                   |
| Culex bitaeniorhynchus| 0.077          | 0.075               | 0.086              | 0.075          | 0.074         | 0.010                   |
pattern of haplotype network. There is no noticeable connection between specific haplotypes and geographic locations among the population of these species, indicating high sequence variation among the haplotypes. The lines in the diagram are proportional to the mutation and the cross lines indicate the rate of mutation (Fig. 2).

**Phylogenetic analysis**

All the seven identified *Culex* mosquitoes were clubbed into distinct conspecific clades in the ML tree using COI sequences (Fig. 3). It constitutes seven distinctive conspecific clusters, each comprised of the haplotypes of the respective species. All of these clades were strongly supported by a minimum of 98% bootstrap value. In the ML tree the highest log-likelihood value $= -3845.61$. The nucleotide sequence of *Armigeres subalbatus* was taken as the root of the phylogenetic tree.

**Table 4** Analysis of intra-population polymorphism among different *Culex* mosquito species

| Species                  | *Culex vishnui* | *Culex tritaeniorhynchus* | *Culex quinquefasciatus* | *Culex fuscocephala* | *Culex fuscanus* | *Culex gelidus* | *Culex bitaeniorhynchus* |
|--------------------------|-----------------|----------------------------|--------------------------|----------------------|-----------------|-----------------|--------------------------|
| No. of haplotypes        | 6               | 9                          | 4                        | 8                    | 7               | 8               | 3                        |
| Polymorphic sites        | 7               | 13                         | 3                        | 13                   | 27              | 10              | 7                        |
| No. of mutation          | 7               | 14                         | 3                        | 13                   | 29              | 10              | 7                        |
| Singleton variable site  | 3               | 10                         | 3                        | 8                    | 22              | 8               | 1                        |
| Parsimony informative site| 4               | 3                          | 0                        | 5                    | 5               | 2               | 6                        |
| Nucleotide differences ($k$) | 2.472     | 2.807                      | 0.666                    | 3.722                | 5.963           | 2.833           | 4                        |
| Nucleotide diversity ($P_i$)  | 0.0058   | 0.0082                     | 0.0013                   | 0.0089               | 0.013           | 0.0044          | 0.01                     |
| Haplotype diversity ($Hd$) | 0.855     | 0.936                      | 0.583                    | 0.972                | 0.818           | 0.972           | 0.7                      |

(See figure on next page.)

**Fig. 2** Haplotype network (Median-joining network) generated using COI sequences in PopArt software for A *Culex vishnui*, B *Culex tritaeniorhynchus*, C *Culex quinquefasciatus*, D *Culex fuscocephala*, E *Culex fuscanus*, F *Culex gelidus*, and G *Culex bitaeniorhynchus* from different geographical regions. Each sphere represented a different group of identical haplotypes; the size describes the number of haplotypes clustered together in unique groups.

(See figure on next page.)

**Fig. 3** Phylogenetic tree of *Culex* mosquito species based on COI sequences. The Maximum-Likelihood tree was constructed from the alignment of COI gene sequences in MEGA X software. The numbers displayed on branches are the bootstrap support obtained through 1000 replications. GenBank sequences are shown with accession numbers. Species along with the NCBI accession no. within the rectangular shape indicates the generated sequences. Sequence from *Armigeres subalbatus* was used as the out-group.
Fig. 3 (See legend on previous page.)
Discussion

Culex mosquito species are the principal vectors of pathogens that causes Lymphatic filariasis, West Nile fever, St. Louis encephalitis, Japanese encephalitis, etc. Culex mosquitoes are also reported as the potential vectors of Rift Valley fever in Africa (Kenawy et al., 2018) and are also suspected to be a vector of the Zika virus in Brazil (Benelli & Romano, 2017; Guedes et al., 2017; Guo et al., 2016). In 2016, a JE virus outbreak was reported from the southern district (Malkangiri) of Odisha state and the JE virus was detected from the Cx. vishnui mosquito vector from the affected villages of that area (Sahu et al., 2018).

Thus, correct identification, distribution, and biogeography of the species are paramount important for conducting programs for the control and prevention of pathogenic diseases that are transmitted by the mosquito vectors, as it allows us to concentrate only on specific mosquito species those spread certain diseases (Murugan et al., 2016). Even though morpho-taxonomy is regarded as the gold standard method for discrimination of mosquito species, it seems to be quite difficult the identification of field-collected mosquitoes as they may be lost some of their important identifying features during handling. For this reason, there is a need for an alternative technique for identification, especially in the case of cryptic and ambiguous species. DNA barcoding technique is proven to be a reliable system of identification. Molecular data are also widely used for producing molecular phylogenetic, phylogeographic, population genetic, and species identification studies. Genetic variation within the species provides knowledge about the origin and migration of the species and contributes support to vector surveillance and vector control strategy. Molecular phylogeny is further utilized for analysis of gene duplication, estimating rates of diversification, polymorphism, recombination, population dynamics, and inferring organismal phylogenies by combining them with other data sources. Earlier studies have evidenced the use of the mitochondrial COI marker to find more biodiversity and increase species richness as compared to conventional taxonomic tools by uncovering undescribed and cryptic species (Hebert et al., 2003; Schmidt et al., 2015; Wilson et al., 2017) and also used to infer the phylogeny of various dipteran taxa within the genera of Aedes, Anopheles and Culex mosquitoes (Ashfaq et al., 2014; Chan-Chable et al., 2019; Panda et al., 2021; Weeraratne et al., 2018). We employed DNA barcoding of a single mt COI gene sequence for seven different species of Culex mosquito in the present study to re-confirm the species identified by morpho-taxonomy. We observed that the composition of generated COI sequences was AT-rich which is in agreement with the findings of Cywinska et al. (2006) and Rivera and Currie (2009) which indicates a general consistency across the dipteran taxa. The current study compared the produced sequences to sequences from various locations in India and other globally available sequences in the database. Based on COI sequence sets, the pairwise genetic divergence of 2–3% has been considered as the threshold value to discriminate between two different species (Hebert et al., 2003) and even for the two closely related mosquito species (Chan et al., 2014; Guany et al., 2015; Kumar et al., 2007).

The transition and transversion bias (R) were determined to understand the pattern of DNA sequence evolution, genetic differences estimation, and analysis of genomic composition as well as mutational positions. In the present study, the transition occurred at a higher frequency than the rate of a transversion, in the mt COI nucleotide sequences, which was also earlier observed by Fitch and by many other scientists (Fitch, 1967; Lyons & Lauring, 2017). Further, the R-value of the COI gene sequences was greater than 0.5, (threshold value) which indicates there is no such bias in transitional and transversional substitution and both are equally probable.

Genetic divergence among the studied haplotype of seven Culex mosquito species shows discriminative inter and intra-specific divergence. Here, our study shows, that the COI sequence differences among the inter-species were on an average almost 9 times higher than the average intra-specific differences. Further, the inter-specific K2P distances exceed 0.02 (2%) for the COI sequences of all the Culex species confirming that they belong to different mosquito species. The inter-specific K2P genetic divergence was found to be 5.6–10.3% and is comparable to the other earlier barcoding studies on mosquitoes of India (5.87–25.65%) (Kumar et al., 2007), Sri Lanka (7.2–17.2%) (Weeraratne et al., 2017), Pakistan (2.3–17.8%) (Ashfaq et al., 2014), China (2.3–21.8%) (Wang et al., 2012). Likewise, the K2P intra-specific divergence found in this study varied from 0.2 to 1.5% which is similar to the findings in China (0–1.67%), Sri Lanka (0–1.4%), Pakistan (0–2.4%) (Ashfaq et al., 2014; Wang et al., 2012; Weeraratne et al., 2017).

Nucleotide diversity (π) and haplotype diversity (Hd) are two major indicators to calculate the diversity of species populations among different geographical strains. Here, in this study, Culex fuscatus exhibited high polymorphic sites and mutations with high nucleotide diversity π (0.013) in their COI sequences, indicating high genetic variation in this population. On the other hand, the lowest variation in Pi was observed in Cx. quinquefasciatus (0.0013), indicating low genetic diversity of this population, possibly due to population expansion or selective sweep. Further, the complex pattern of the haplotype network of COI gene sequences in most of the species indicates that there is less association between
certain haplotypes and geographic locations. In addition, a high value of haplotype diversity and a certain degree of genetic divergence were observed in these species.

Furthermore, to confirm the morphological identification and to obtain the taxonomic position of the selected taxa, a phylogenetic analysis was carried out. The Maximum-Likelihood tree is a widely constructed phylogenetic tree, especially in mosquito species for illustrating the evolutionary relationship among the targeted groups (Chan et al., 2014; Guany et al., 2015). In the current study, the mt COI-based analysis strongly supports the positioning of all the seven species by forming a distinct conspecific cluster with their respective group. Further, Cx. vishnui and Cx. tritaeniorynchus sequences were positioned neighbor to each other in the constructed phylogenetic tree, indicating that both come under the same subgroup of *Culex vishnui*. Similar phylogenetic positioning among the *Culex vishnui* subgroup was inferred using ITS1 and COI genes by Karthika et al. (2018) and Toma et al. (2000).

Conclusions
The present study describes the *Culex* mosquito diversity of the Southern districts of Odisha state, India. Our study proved the potency of the DNA barcoding technique for mosquito species identification in addition to the morpho-taxonomy. This study also adds valuable information about the systematics and molecular biology of seven public important mosquito species acting as a significant vector for Japanese encephalitis in various Asian continents. The generated mt COI sequences could be used as reference nucleotide sequences of the respective haplotypes in future mosquito identification studies and will facilitate the conspecific comparison to reveal the appropriate reason for high intra-specific divergence. Furthermore, different climatic conditions and geographical barriers can be responsible for the genetic variations in a species. This information is further used for the effective implementation of region-specific vector control strategies.

Abbreviations
mt COI: Mitochondrial cytochrome oxidase subunit I gene; Cx: *Culex*; ML: Maximum-Likelihood; K: P. Kimura 2 parameter model.

Acknowledgments
This research work was supported by Dept. of Science & Technology (DST), Govt. of Odisha under Biju Patnaik Research Fellowship (BPRF) scheme.

Author contributions
PD carried out the field and laboratory studies. BTK designed and critically reviewed the manuscript. Both authors read and approved the final manuscript.

Funding
This research was funded by Dept. of Science & Technology (DST), Govt. of Odisha under Biju Patnaik Research Fellowship (BPRF) scheme (Grant No. 202861).

Availability of data and materials
All the data generated and analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declared no potential competing interest with respect to the research, authorship, and/or publication of this article.

Author details
1 Department of Zoology, Berhampur University, Berhampur, Odisha 760007, India. 2 Department of Biotechnology, Berhampur University, Berhampur, Odisha 760007, India.

Received: 31 January 2022 Accepted: 10 July 2022 Published online: 25 July 2022

References
Ashfaq, M., Hebert, P. D. N., Mirza, J. H., Khan, A. M., Zafar, Y., & Mirza, M. S. (2014). Analyzing mosquito (Diptera: Culicidae) diversity in Pakistan by DNA barcoding. *PLoS ONE*, 9(5), e97268. https://doi.org/10.1371/journal.pone.0097268

Attallah, M., Gil, S., Bibi, D., Andaleeb, A., Ilahi, I., Siraj, M., Ahmad, M., Ullah, I., Ali, M., Ahmad, S., & Ullah, Z. (2021). Diversity, distribution and relative abundance of the mosquito fauna (Diptera: Culicidae) of Malakand and Dir Lower, Pakistan. *Brazilian Journal of Biology*, 81, e247374. https://doi.org/10.1590/1519-6984.247374

Azari-Hamidian, S., Linton, Y. M., Abai, M. R., Ladonni, H., Oshaghi, M. A., Hanafi-Bojd, A. A., Moosa-Kazemi, H., Shabkhiz, H., Pakari, A., & Harbach, R. E. (2010). Mosquito (Diptera: Culicidae) fauna of the Iranian islands in the Persian Gulf. *Journal of Natural History*, 44, 913–925. https://doi.org/10.1080/00222930903437338

Barraud, P. J. (1934). *The fauna of British India, Including Ceylon and Burma. Diptera Vol. 5. Family Culicidae. Tribes Megahemini and Culicini*. Taylor & Francis.

Benelli, G., & Romano, D. (2017). Mosquito vectors of Zika virus. *Entomologia Generalis*, 36(4), 309–318.

Bhattacharya, D. R., Rajavel, A. R., Mohapatra, P. K., Jambulingam, P., Mahanta, J., & Prakash, A. (2014). Faunal richness and the checklist of Indian mosquitoes (Diptera: Culicidae): Check List, 106, 1342–1358.

Chan, A., Chiang, L., Hapuarachchi, H. C., Tan, C., Pang, S., Lee, R., Lee, K. N., & Lam-phua, S. (2014). DNA barcoding: Complementing morphological identification of mosquito species in Singapore. *Parasite Vectors*, 7, 569.

Chan-Chable, R. J., MartinezArce, A., MisAvila, P. C., & OrtegaMorales, A. I. (2019). DNA barcodes and evidence of cryptic diversity of anthropophagous mosquitoes in Quintana Roo, Mexico. *Ecology and Evolution*, 9, 4692–4705. https://doi.org/10.1002/ece3.5073

Collins, F. H., Mendez, M. A., Rasmussen, M. O., Mejheffy, P. C., Besansky, N. J., & Finnerty, V. (1987). A Ribosomal RNA gene probe differentiates member species of the *Anopheles gambie* complex. *American Journal of Tropical Medicine and Hygiene*, 37, 37–41. https://doi.org/10.4269/ajtmh.1987.37.37

Cywinski, A., Hunter, F. F., & Hebert, P. D. N. (2006). Identifying Canadian mosquito species through DNA barcodes. *Medical Veterinary Entomology*, 20, 413–424. https://doi.org/10.1111/j.1365-2915.2006.00653.x
Tyagi, B. K., Munirathinam, A., & Venkatesh, A. (2015). A catalogue of Indian mosquitoes. *International Journal of Mosquito Research, 50*(625), 50–97.

Wang, G., Li, C. X., Guo, X. X., Xing, D., Dong, Y. D., Wang, Z. M., Liu, M. D., Zheng, Z., Zhang, H. D., Zhu, X. J., Wu, Z. M., & Zhao, T. Y. (2012). Identifying the main mosquito species in China based on DNA barcoding. *PLoS ONE, 7*, e47051.

Weeraratne, T. C., Surendran, S. N., & Parakrama Karunaratne, S. H. P. (2018). DNA barcoding of morphologically characterized mosquitoes belonging to the subfamily Culicinae from Sri Lanka. *Parasites Vectors, 11*, 266. [https://doi.org/10.1186/s13071-018-2810-z](https://doi.org/10.1186/s13071-018-2810-z)

Weeraratne, T. C., Surendran, S. N., Reimer Lisa, J., Wondji Charles, S., Perera, M., Devika, B., Walton, C., & Parakrama Karunaratne, S. H. P. (2017). Molecular characterization of Anopheline (Diptera: Culicidae) mosquitoes from eight geographical locations of Sri Lanka. *Malaria Journal, 16*, 234. [https://doi.org/10.1186/s12936-017-1876-y](https://doi.org/10.1186/s12936-017-1876-y)

Wilson, J. J., Sing, K. W., Floyd, R. M., & Hebert, P. D. (2017). DNA Barcodes and insect biodiversity. In R. G. Footh & P. H. Adler (Eds.), *Insect biodiversity: Science and society*, Vol. 1 (Vol. 17, pp. 575–592). Blackwell Publishing.

**Publisher’s Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.