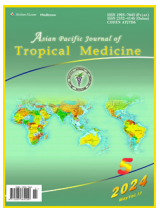

Original Article

Asian Pacific Journal of Tropical Medicine

apjtm.org

doi: 10.4103/apjtm.apjtm_790_23

Phylogenetic, phylogeographic and divergence time analysis of *Anopheles subpictus* species complex using ITS2 and COI sequences

Lihini Sandaleka Muthukumarana[1], Methsala Madurangi Wedage[1], Samanthika Rathnayake[1], Nissanka Kolitha De Silva[1,2,3]✉

[1]*Department of Zoology, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka*

[2]*Genetics and Molecular Biology Unit, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka*

[3]*Sri Lanka Institute of Biotechnology, Pitipana, Homagama, Sri Lanka*

ABSTRACT

Objective: To address the phylogenetic and phylogeographic relationship between different lineages of *Anopheles (An.) subpictus* species complex in most parts of the Asian continent by maximum utilization of Internal Transcriber Spacer 2 (ITS2) and cytochrome C oxidase I (COI) sequences deposited at the GenBank.

Methods: Seventy-five ITS2, 210 COI and 26 concatenated sequences available in the NCBI database were used. Phylogenetic analysis was performed using Bayesian likelihood trees, whereas median-joining haplotype networks and time-scale divergence trees were generated for phylogeographic analysis. Genetic diversity indices and genetic differentiation were also calculated.

Results: Two genetically divergent molecular forms of *An. subpictus* species complex corresponding to sibling species A and B are established. Species A evolved around 37-82 million years ago in Sri Lanka, India, and the Netherlands, and species B evolved around 22-79 million years ago in Sri Lanka, India, and Myanmar. Vietnam, Thailand, and Cambodia have two molecular forms: one is phylogenetically similar to species B. Other forms differ from species A and B and evolved recently in the above mentioned countries, Indonesia and the Philippines. Genetic subdivision among Sri Lanka, India, and the Netherlands is almost absent. A substantial genetic differentiation was obtained for some populations due to isolation by large geographical distances. Genetic diversity indices reveal the presence of a long-established stable mosquito population, at mutation-drift equilibrium, regardless of population fluctuations.

Conclusions: *An. subpictus* species complex consists of more than two genetically divergent molecular forms. Species A is highly divergent from the rest. Sri Lanka and India contain only species A and B.

KEYWORDS: Molecular systematics; ITS2; COI; DNA sequences; Phylogeny; Phylogeography

1. Introduction

Malaria is a fatal vector-borne disease transmitted by 70-80 species of mosquitoes of the genus *Anopheles* infected by *Plasmodium* parasites[1,2]. Despite the availability of prevention and control methods, it has resulted in the most significant mortality burden to the world population[1]. More than 3 billion people worldwide are at risk of infecting malaria[3,4], where 85 malaria-endemic countries have been reported by 2022[5]. According to the latest malaria report, there were an estimated 249 million cases with 608 000 deaths in 2022 (https://www.who.int/news-room/fact-sheets/detail/malaria). Malaria was one of the most devastating health burdens in Sri Lanka in the 1930s. However, as a result of the historical milestone associated with the malaria eradication programme and national-

Significance

The phylogenetic and phylogeographic analysis of *Anopheles subpictus* species complex was performed using Internal Transcriber Spacer 2 and cytochrome C oxidase I sequences. More than two genetically divergent forms of the species complex are present, whereas sibling species A is highly divergent from the rest. Sri Lanka and India contain only species A and B.

✉To whom correspondence may be addressed. E-mail: nissanka@sci.sjp.ac.lk

How to cite this article: Muthukumarana LS, Wedage MM, Rathnayake S, De Silva NK. Phylogenetic, phylogeographic and divergence time analysis of *Anopheles subpictus* species complex using ITS2 and COI sequences. Asian Pac J Trop Med 2024; 17(5): 214-225.

Article history: Received 15 October 2023 Revision 30 March 2024
Accepted 22 April 2024 Available online 29 April 2024

level elimination strategies conducted for several decades, Sri Lanka was officially certified as a malaria-free country on 5th September 2016 by the World Health Organization (WHO)[3,6]. Yet, there is a threat of the re-emergence of malaria due to the geographic spread of disease vectors, facilitated by international trade and tourism, climatic and ecological changes, human migration, changes in vector behavior, and drug resistance[3,7,8].

Anopheles (Cellia) subpictus Grassi 1899 belongs to the phylum Arthropoda, class Insecta, order Diptera, family Culicidae, subfamily Anophelinae and is a member of the Pyretophorous series of the subgenus *Cellia*[9,10]. It is the most abundant anopheline mosquito in most parts of the Indian subcontinent and Southeast Asia[11]. It is the primary vector of malaria in Southeast Asian countries and the secondary vector in Sri Lanka because it transmits *Plasmodium vivax* and *Plasmodium falciparum* in many parts of the country[12]. *Anopheles (An.) subpictus s.l.* (sensu lato) is distributed throughout the Oriental and Australian zones, ranging from northeastern Pakistan to India, Sri Lanka, Bangladesh, Myanmar, Indonesia, Iran, Nepal, Thailand, the coastal region of Cambodia, Vietnam, Malaysia, Maldives, Mariana Islands, Timor-Leste, Papua New Guinea, Philippines, Afghanistan, Australia, China extending at Solomon Islands[7,11,13].

The taxon *An. subpictus* is a species complex, provisionally designated previously as A, B, C, and D based on stage-specific morphometric characteristics, including the number of egg float ridges, fourth instar larval mesothoracic seta Ⅳ, pupal seta 7-Ⅰ and apical pale and pre-apical dark bands in palpi of adults and proboscis characters. The four sibling species are also confirmed through fixed paracentric inversions of the X arm of polytene chromosomes[14–17]. All four species are recorded in Sri Lanka based on morphometric characteristics and polytene chromosome banding patterns[18], but accurate identification is difficult due to the presence of cryptic taxa with overlapping morphological characters that possibly have occurred by genetic drift in small island populations[19]. The incongruity between morphological identification and molecular analysis based on rDNA studies appeared when the majority of morphologically identified *An. subpictus* species B are members of the *sundaicus* complex[20,21]. Jayatunga *et al*. also confirmed that due to independent genomic variations, the *subpictus* species complex in Sri Lanka is not identifiable using morphological characters[22].

Sibling species are reproductively distinct organisms that cannot be solely distinguished by morphological characteristics[23]. They differ in their ability to transmit pathogens and their sensitivity to commonly used insecticides. Hence, it is crucial to correctly identify the vector sibling species and their distribution through molecular assays for cumulative control strategies[1,12]. The two most common molecular markers proven in characterizing the Anophelines, especially in detecting their sibling species are Cytochrome C Oxidase I (COI) from the mitochondrial genome and Internal Transcriber Spacer 2 (ITS2) from the nuclear ribosomal DNA[24–29].

Molecular studies performed using ITS2 and COI sequences confirm that the *subpictus* complex in Sri Lanka is composed of two genetically distinct species corresponding to species A and B in India[30,31], where the two sibling species are sexually incompatible and may be undergoing speciation[8]. Speciation of a population by the reconstruction of a genotype could occur due to pre-zygotic or post-zygotic isolation[8,32]. Accordingly, if a founder population becomes isolated as small populations, they could undergo random drift, causing speciation by rapid evolution[33]. Mayr describes that two species are allopatric, if they do not occur together due to geographical separation, whereas they are sympatric if the two species co-exist under overlapped distribution areas and still acquire reproductive isolation[34]. The speciation of *An. subpictus* species complex in different geographic locations in correspondence to their diversification of evolutionary time scale remains undefined in previous studies.

Wilai *et al*. suggest that *An. subpictus* species in Thailand, Cambodia, Indonesia, Vietnam, and the Philippines are more closely related to species B than A, and species A is highly diverged and is confined to India and Sri Lanka, whereas species B is found in Sri Lanka, India and Myanmar[35]. However, all the previous studies were carried out using a limited number of sequences either generated or extracted from a database, leaving a gap between the sequences being used and the rest. The current study aimed to investigate the phylogenetic and phylogeographic relationship and the diversification of the timescale of *An. subpictus* species complex by maximum utilization of ITS2 and COI sequence data deposited at the GenBank and the derived concatenated sequences for the analysis that was the first attempt in the molecular studies of *An. subpictus* species complex.

2. Subjects and methods

2.1. Data collection

A total of 75 ITS2 sequences (Supplementary Table 1) and 210 COI sequences (Supplementary Table 2) of *An. subpictus* species complex available up to date at the NCBI public data repository (www.ncbi.nlm.nih.gov) were retrieved. Table 1 depicts the number of ITS2 and COI sequences used from each country. *Aedes aegypti* was selected as the outgroup for the phylogenetic study. Accordingly, three ITS2 and COI sequences were retrieved separately (Supplementary Table 3).

2.2. Sequence alignment

Clustal W in MEGA v 5.2 software was used for sequence alignment and trimming to obtain a consistent region[27,36]. In the ITS2 sequence alignment, 9 nucleotide sequences containing hypervariable regions were excluded out of 75 due to misalignment:

Table 1. Number of ITS2 and COI sequences of *Anopheles subpictus* species complex from each country.

Country	ITS2		COI	
	No. of sequences	Percentage (%)	No. of sequences	Percentage (%)
Sri Lanka	27	36.0	32	15.2
India	22	29.3	14	6.7
Papua New Guinea	1	1.3	-	-
Vietnam	3	4.0	39	18.6
Cambodia	2	2.7	10	4.8
Indonesia	6	8.0	16	7.6
Thailand	12	16.0	18	8.6
Myanmar	1	1.3	37	17.6
Netherlands	1	1.3	1	0.5
Philippines	-	-	43	20.5
Total	75	100.0	210	100.0

EF192277, GQ870335, HQ703002, JQ845942, KJ437453, KY000689, MG976758, MT366210, MW078487. Accordingly, 66 sequences of length 509 bp were obtained. In the COI sequence alignment, repeated similar sequences of each country, low quality and hyper-variable sequences were excluded to obtain a filtered data set for better tree visualization. Accordingly, 128 sequences of length 471 bp were obtained. A concatenated sequence alignment was generated by joining ITS2 sequences to the trimmed end of COI sequences available from the same original species from the retrieved data (Supplementary Table 4). A total of 25 concatenated sequences of length 1 089 bp were maintained by eliminating the mismatching sequence of specimen code: 11608.

Since the proper designation of sibling species of ITS2 and COI sequences of *An. subpictus* species complex was not mentioned in the data repository at NCBI, the most recent study by Wilai *et al*[35] was used to identify the corresponding sibling species A and B in the phylogenetic, phylogeographic and haplotype network of the current study.

2.3. Phylogenetic analysis

Phylogenetic trees were constructed for ITS2, COI, and concatenated sequences separately. The FASTA file formats of trimmed sequence alignments were converted to nexus (.nex) and phylip (.phy) file formats by Geneious Prime v 2021.2.2.0 software. The PartitionFinder v1.1.1 software was used to obtain the best scheme models of evolution for the three separate tree constructions (Table 2). MrBayes v 3.2.5_win 64 software was used to perform the Bayesian analysis to reconstruct the phylogeny for posterior probability analyses with two simultaneous runs until $f < 0.01$. FigTree v 1.4.2 was used to edit and visualize the tree topology[37,38].

2.4. Construction of haplotype networks

A total of 62 ITS2 sequences of 455 bp and 204 COI sequences of 393 bp were filtered from the originally retrieved data, excluding the mismatching sequences with R and Y values. DNA Sequence Polymorphism (DnaSP) v 5.10 software, Network v 10.2 software and Arlequin v 3.5.2.2 software were used to define sequence sets, to determine the spatial distribution of haplotypes through median-joining, and to generate the haplotype frequencies, respectively[31,39].

2.5. Determination of genetic diversity

Genetic diversity of both ITS2 and COI sequences of *An. subpictus* species complex for the total number of sites (excluding sites with gaps) was obtained in terms of the number of haplotypes, haplotype diversity, nucleotide diversity and neutrality tests (Tajima's D, Fu and Li's D, Fu and Li's F, Fu's F stat) using DnaSP v 5.1.0 software by executing the FAS alignment file[8].

Table 2. Best scheme models of evolution for ITS2, COI and concatenated gene subsets.

Gene subset	Fragment size	Subset partitions	Coding positions	Best fitting model
ITS2	509 bp	ITS2_pos1	1-509	K80+G
		ITS2_pos2	2-509	K80+G
		ITS2_pos3	3-509	K80+G
COI	471 bp	COI_pos1	1-471	SYM+I
		COI_pos2	2-471	F81
		COI_pos3	3-471	HKY+G
Concatenated sequence: COI	482 bp	COI_pos1	1-482	HKY+I
		COI_pos2	2-482	K80+I
		COI_pos3	3-482	F81
Concatenated sequence: ITS2	607 bp	ITS2_pos1	483-1 089	K80+G
		ITS2_pos2	484-1 089	K80+G
		ITS2_pos3	485-1 089	K80+G

2.6. Determination of genetic differentiation

Genetic differentiation of both ITS2 and COI sequences of *An. subpictus* species complex was determined using Arlequin v 3.5.2.2 software based on fixation index (F_{ST}) values[40].

2.7. Phylogeographic analysis

Phylogeographic trees were constructed for ITS2 (66 sequences of 509 bp), COI (128 sequences of 471 bp) and concatenated (25 sequences of 1 089 bp) sequences separately. HKY was used as the substitution model utilizing the Yule process in BEAUti v 1.8.2 software[37,38]. A Bayesian Markov Chain Monte Carlo approach (MCMC) was performed in BEAST v 1.8.2 software to infer the topology and node ages. The diversification age of the most recent common ancestor (MRCA) of the *Anopheles* genus at 83.23 million years ago with 95% credibility ranging from 54.33 million years ago to 115.88 million years ago was used to calibrate the node ages[41]. Tree Annotator v 1.8.2 software and FigTree v 1.4.2 software were used to edit and visualize node branches[37,41,42].

2.8. Ethical approval

Ethical approval was not required as this study involved the use of already published data.

3. Results

3.1. Phylogenetic analysis

The phylogenetic tree of ITS2 sequences (Supplementary Figure 1) constructed using Bayesian Inference (BI) reveals the separation of *An. subpictus* individuals into two clades with posterior probability (pp) 100%. The derived clade is divided into four main clades (A, B, C, D), whereas the basal clade (E) depicts a significantly different genetic structural evolution from the rest. This clade E corresponds to the *An. subpictus* sibling species A encompassing sequences limited to Sri Lanka and India. The clade A includes Sri Lankan, Indian and Myanmar species separated from the rest (pp=94%) and corresponds to the *An. subpictus* species B. The clade B belongs to species from Cambodia, Vietnam, Thailand, and Indonesia. *An. subpictus* species in Papua New Guinea has evolved with a separate genetic structure (pp=82%) in clade C. Clade D includes individuals from Vietnam, Cambodia, and Thailand (pp=99%).

The phylogenetic tree of COI sequences constructed using BI (Supplementary Figure 2) reveals the separation of *An. subpictus* individuals into two clades with a posterior probability (pp) of 100%. The derived clade further divides into two separate clades (pp=99%). The top clade consists of four clades, A, B, C and D, with pp of 100% from which A, B and C evolved as sister clades. The bottom clade is divided into four sub-clades: E, F, G and H. The clade A and C includes only Thailand species. The clade B includes two Thailand species, mostly from Vietnam. The clade D includes species from Myanmar and Sri Lanka and corresponds to *An. subpictus* species B. The clade E consists of a mixture of species from Thailand, Cambodia, and Vietnam. The clades F and G have evolved with species from Indonesia, particularly South Sulawesi and Java, respectively. The clade H consists only of the Philippines species. The basal clade I includes Sri Lankan, Indian species along with one Netherland species corresponding to *An. subpictus* species A. The Thailand species are distributed in 4 clades: A, B, C and E. Indonesian species from South Sulawesi and Flores in clade F has occurred in a sister relationship to clade E. However, the Indonesian species from Java in clade G shows a polyphyletic relationship to clade F.

The concatenated tree (Figure 1) reveals the separation of *An. subpictus* individuals into two clades with a pp of 100%. With the same pp, the derived clade further separates into two clades. The upper clade is separated into four clades, and the lower clade has separated into three clades. The clade A includes Sri Lankan and Myanmar species corresponding to sibling species B. The clade B consists of species from Cambodia, Thailand, and Vietnam. Both C and D clades include Thailand species. Clade E includes Vietnam and Cambodia. Clade F, containing Indonesian species from South Sulawesi and Flores, shows a sister group relationship to clade E. Clade G includes Thailand species. The basal clade H, including Sri Lankan species, has significantly separated from the rest, evolved with a different genetic structure, and corresponds to sibling species A.

3.2. Construction of haplotype network

Haplotype network was obtained using 62 ITS2 sequences of *An. subpictus* species complex (Figure 2), which reveals the presence of 29 haplotypes under two significant haplogroups. Haplogroup A consists only of Indian and Sri Lankan species, and B includes a mixture of individuals from India, Sri Lanka, Papua New Guinea, Vietnam, Cambodia, Indonesia, Thailand, and Myanmar. The haplotype frequencies and haplotype densities (Hd) of each population are listed in Supplementary Table 8. Haplotypes 15 and 12 are observed as the regional ancestors of haplogroups A and B, respectively. A common ancestor is not clearly visible in this network. H_10, H_12, H_13, and H_15 are recorded as shared haplotypes, while the rest are private haplotypes.

Haplotype network obtained using COI sequences of *An. subpictus*

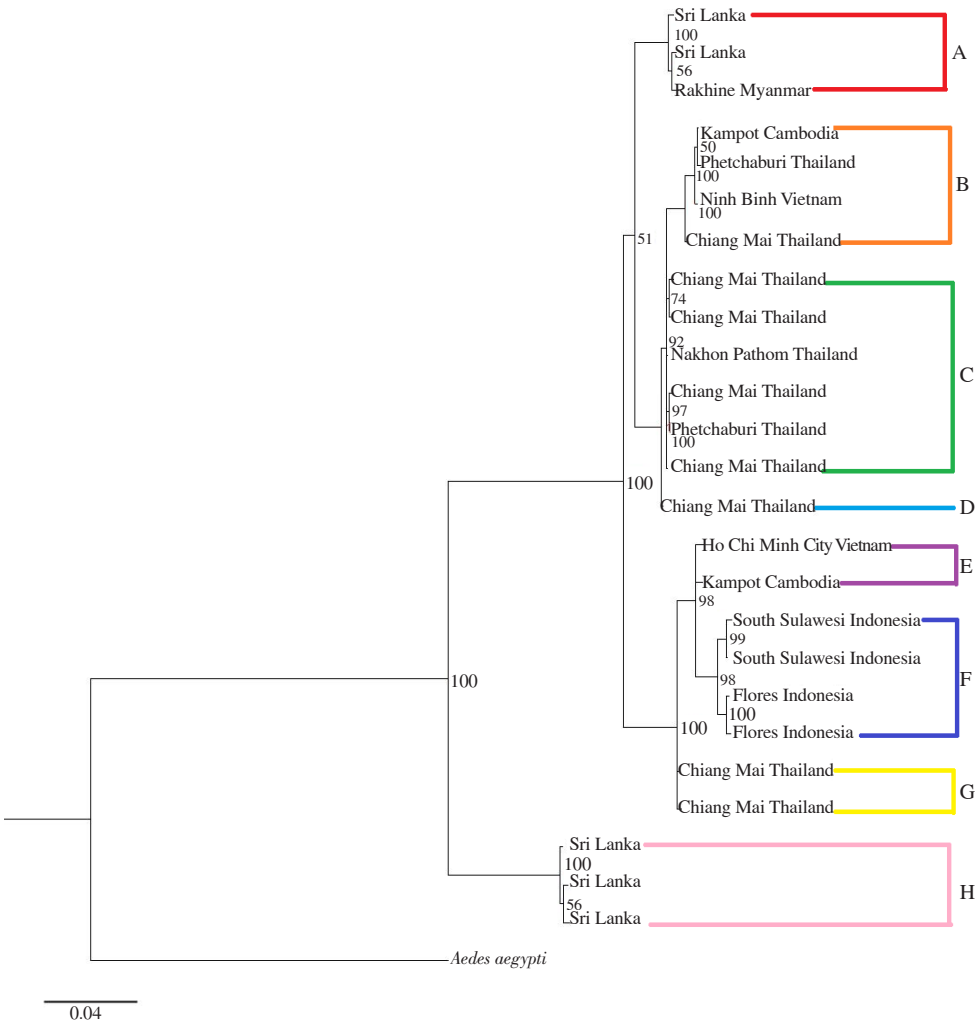

Figure 1. Phylogenetic analysis based on consensus ITS2 and COI sequences of *Anopheles subpictus* species using Bayesian likelihood tree generated by MrBayes v 3.2.5_win 64 software. 25 sequences, 1 089 characteristics, 420 000 generations, 2 mcmc runs. Numbers associated with nodes are Bayesian posterior probabilities (%) above 50. Clades are marked from A-H. GenBank accession numbers of concatenated sequences are listed in Supplementary Table 7.

species (Figure 3) reveals the presence of seven significant haplogroups. In each haplogroup (A-G), the regional ancestors are identified as H_10, H_40, H_113, H_4, H_75, H_43, and H_12, respectively. Haplogroups A, E and F consist only of private haplotypes found in Thailand, the Philippines, and Indonesia. Haplogroup B includes shared haplotypes (H_11, H_38, H_40) from Vietnam, Cambodia, and Thailand. Haplogroup C consists of most private haplotypes from Sri Lanka and Myanmar, a shared haplotype (H_60) of the two countries and a single private haplotype from India. Haplogroup D consists mainly of private haplotypes from Sri Lanka and India, with two shared haplotypes (H_1, H_4) from Sri Lanka, India, and the Netherlands. Haplogroup G includes two shared haplotypes of Thailand, Cambodia, and Vietnam. The haplotype frequencies and haplotype densities (Hd) of each population are listed in Supplementary Table 9. The median vector 29 could be expressed as the common ancestor in this network.

3.3. Determination of genetic diversity

Genetic diversity of *An. subpictus* species complex calculated using ITS2 and COI sequences (Table 3) indicate averagely high haplotype and nucleotide diversity. Nucleotide diversity and average number of nucleotide differences are greater for ITS2 sequences than COI. Tajima's D value is negative for ITS2 sequences and positive for COI sequences with non-significant *P* values. Fu and Li's D and F values are negative for both sequences, whereas only Fu and Li's D values

Table 3. Genetic diversity indices including sample size (x), number of variable sites (s), haplotypes (h), haplotype diversity (Hd), nucleotide diversity (π), the average number of nucleotide differences (k), Tajima's D, Fu and Li's D, Fu and Li's F, Fu's Fs statistics calculated using ITS2 and COI sequences of *Anopheles subpictus* species complex.

Region	x	s	h	Hd	π	k	Tajima's D	Fu and Li's D	Fu and Li's F	Fu's Fs statistics
ITS2	62	157	29	0.875	0.087 20	35.403 49	-0.194 33**	-3.062 23*	-2.314 01**	11.847
COI	204	114	116	0.986	0.076 22	29.953 93	0.458 81**	-1.509 93**	-0.677 11**	-39.359

*Significant $P<0.05$; **non-significant $P>0.10$.

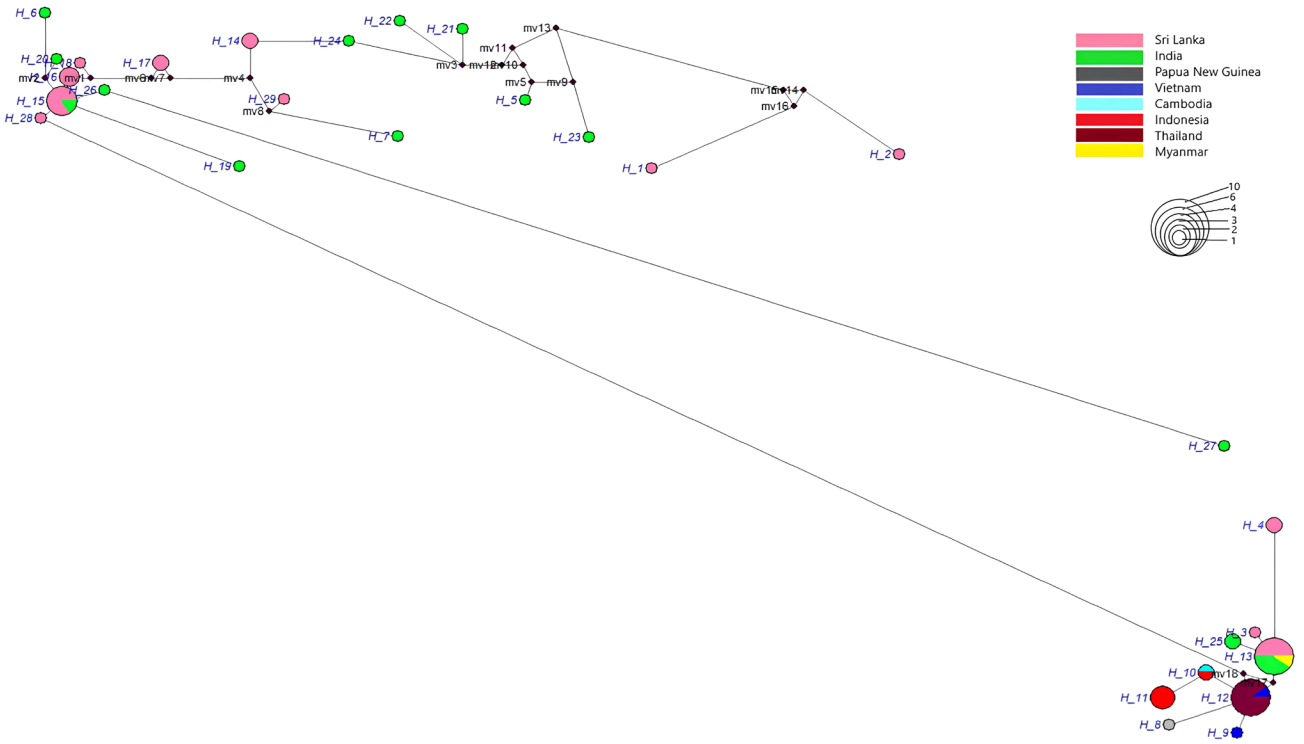

Figure 2. Median-joining network of ITS2 sequences of *Anopheles subpictus* species generated using DnaSP v5 software and Network 10.2 software. The colour code indicates different geographic areas. The scale indicates the relative frequency of sequences belonging to each haplotype. Haplogroups are named as A and B. Black colour dots indicate the median vectors.

for the ITS2 sequence are significant. Fu's Fs statistics have resulted in positive and negative values for the ITS2 and COI, respectively.

The total G+C content is 0.579 at coding positions of 406.00 sites of ITS2 sequences. The total G+C content at coding positions of 393.00 sites of COI sequences is 0.336. Total number of mutations in 157 variable sites of ITS2 sequences is 176. In COI sequences, a total of 154 mutations were observed in 114 variable sites.

3.4. Determination of genetic differentiation

Table 4 depicts Pairwise fixation (F_{ST}) values between populations of *An. subpictus* species complex in different countries calculated using ITS2 sequences. Generally, F_{ST} values are ranged as follows: 0-0.05: small, 0.005-0.15: moderate, 0.15-0.25: great, >0.25: huge genetic differentiation. The negative F_{ST} values are considered as zero, indicating no genetic subdivision between Sri Lanka and India. Almost all other countries show greater genetic differentiation among populations, resulting in a maximum value (1.000 00) for Cambodia-Papua New Guinea, Indonesia-Papua New Guinea, Indonesia-Cambodia, Myanmar-Papua New Guinea, Myanmar-Cambodia, and Myanmar-Indonesia.

Table 5 depicts Pairwise fixation (F_{ST}) values between populations of *An. subpictus* species complex in different countries calculated using COI sequences. There seems to be no genetic subdivision between India-Netherlands, and Sri Lanka-Netherlands due to the negative F_{ST} values which are effectively considered zero. The highest genetic differentiation among populations is observed in Myanmar-India, Myanmar-Netherlands, and Philippines-Netherlands.

3.5. Phylogeographic analysis

The phylogeographic tree constructed using ITS2 sequences of *An. subpictus* species is shown in Supplementary Figure 3. Accordingly,

Table 4. Pairwise fixation index (F_{ST}) values between populations of different countries using ITS2 sequences of *Anopheles subpictus* species complex.

Country	Sri Lanka	India	Papua New Guinea	Vietnam	Cambodia	Indonesia	Thailand	Myanmar
Sri Lanka	0.000 00							
India	-0.027 52	0.000 00						
Papua New Guinea	0.278 15	0.137 16	0.000 00					
Vietnam	0.376 20	0.286 00	0.714 29	0.000 00				
Cambodia	0.262 94	0.126 12	1.000 00	0.500 00	0.000 00			
Indonesia	0.470 14	0.403 59	1.000 00	0.913 98	1.000 00	0.000 00		
Thailand	0.515 53	0.469 60	0.904 76	0.293 89	0.777 78	0.906 27	0.000 00	
Myanmar	0.236 75	0.095 73	1.000 00	0.333 33	1.000 00	1.000 00	0.739 13	0.000 00

Matrix of significant F_{ST} *P* values significance level=0.05.

Table 5. Pairwise fixation index (F_{ST}) values between populations of different countries using COI sequences of *Anopheles subpictus* species complex.

Country	India	Netherland	Sri Lanka	Thailand	Indonesia	Cambodia	Myanmar	Philippines	Vietnam
India	0.000 00								
Netherland	-0.812 70	0.000 00							
Sri Lanka	0.130 60	-0.305 50	0.000 00						
Thailand	0.676 50	0.633 53	0.434 82	0.000 00					
Indonesia	0.738 80	0.730 93	0.499 96	0.524 38	0.000 00				
Cambodia	0.637 55	0.503 02	0.412 94	0.200 47	0.348 80	0.000 00			
Myanmar	0.900 05	0.946 80	0.551 29	0.760 91	0.834 21	0.767 44	0.000 00		
Philippines	0.886 19	0.924 65	0.662 14	0.715 97	0.712 64	0.690 62	0.909 93	0.000 00	
Vietnam	0.652 75	0.613 67	0.474 85	0.151 16	0.456 24	0.047 17	0.686 30	0.635 08	0.000 00

Matrix of significant F_{ST} *P* values significance level=0.05.

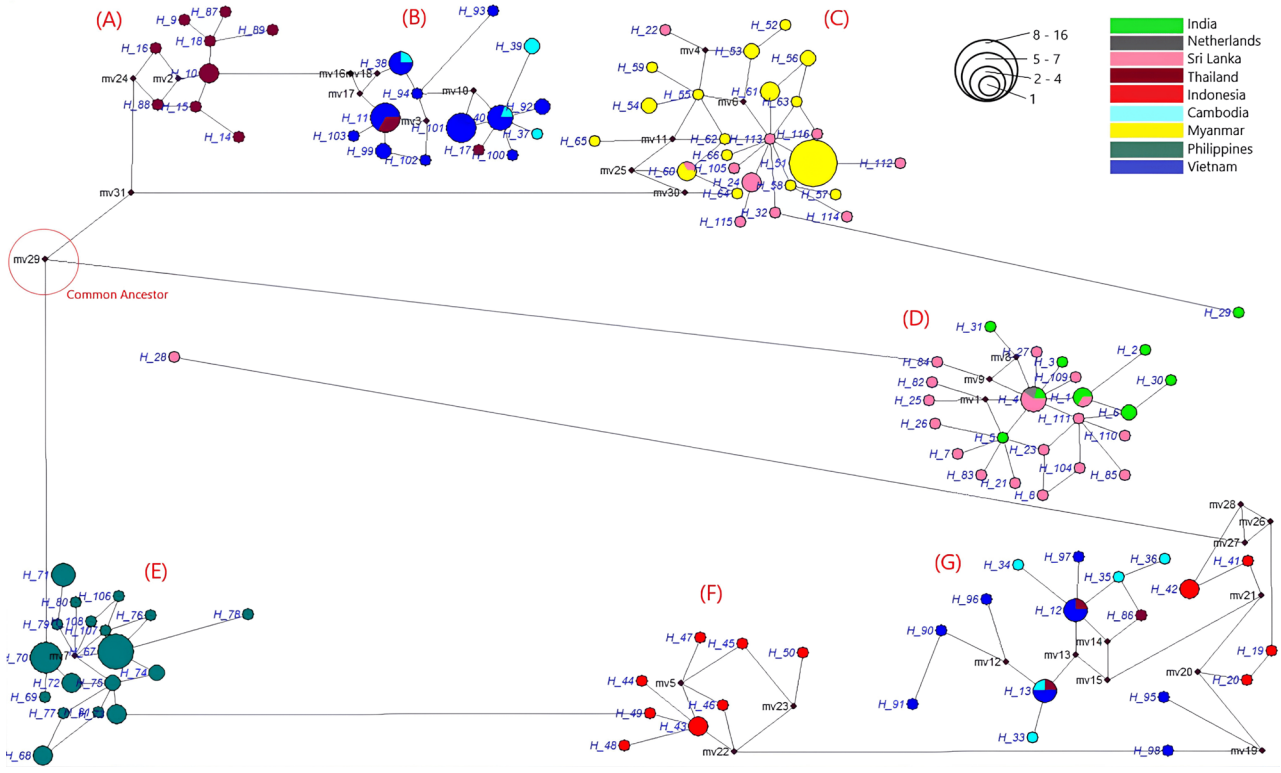

Figure 3. Median joining network of COI sequences of *Anopheles subpictus* species generated using DnaSP v5 software and Network 10.2 software. The colour code indicates different geographic areas. The scale indicates the relative frequency of sequences belonging to each haplotype. Haplogroups are named from A-G. Black colour dots indicate the median vectors. Common ancestor is circled in red.

the MRCA of *An. subpictus* species ages back to 83.18 million years ago with a 95% highest posterior density (HPD), which belongs to the early Cretaceous period. The tree depicts a significant separation of the species into 2 clades. The upper clade A (corresponding to species A) diverged 57.36 million years ago from India and Sri Lanka during the Paleocene period, which later diverged further during the Eocene, Oligocene, and Miocene periods. The bottom clade evolved 49.96 million years ago during the Eocene period and later diverged into two separate clades. The first clade evolved 43.14 million years ago, forming clade B (corresponding to species B), including species from India, Myanmar, and Sri Lanka. The second clade evolved 33.91 million years ago, where individuals from Papua New Guinea evolved separately, forming clade C, and individuals from Thailand, Vietnam, and Cambodia forming clade D. Clade E is the latest to have diverged in the late Oligocene period in Indonesia, Vietnam, Cambodia, and Thailand.

Supplementary Figure 4 shows the phylogeographic tree constructed using COI sequences of *An. subpictus* species. The MRCA of *An. subpictus* species dates back to 83.2 million years ago with 95% HPD in the early Cretaceous period. It has diverged into two separate clades in the Paleogene period. The upper clade has evolved 63.48 million years ago, forming a separate clade A that consists of individuals from India, Sri Lanka, and the Netherlands corresponding to species A (Supplementary Figure 4). In 57.99 million years ago, the lower clade has diverged again into two separate clades. Clade B (corresponding to species B) evolved from Sri Lanka and Myanmar, 28.08 million years ago in the Oligocene period. Clade C evolved 36.89 million years ago from Thailand and Vietnam. The bottom clade diverged 58.71 million years ago in the Paleocene period into 2 main clades, which further diverged later,

forming three separate clades. Clade D evolved separately from Indonesia in the Miocene period. Clade E evolved in the Oligocene period 28.85 million years ago, consisting of individuals only from the Philippines. A mixture of individuals from Thailand, Cambodia, Vietnam, and Indonesia evolved together 32.41 million years ago forming clade F.

The phylogeographic tree of *An. subpictus* species constructed using concatenated sequences of COI and ITS2 is shown in Figure 4. It proves that its MRCA ages back to 83.21 million years ago with a 95% HPD, which evolved in the early Cretaceous period. According to this analysis, it has been separated into two clades where the basal clade G (corresponding to species A) has significantly diverged from the rest, including species from Sri Lanka. The derived clade evolved 60.42 million years ago into two clades where each has further diverged. Clade A and B were separated 31.39 million years ago into clades, where clade A consisted only of Thailand individuals and clade B involved individuals from Vietnam, Thailand, and Cambodia. The lower clade has diverged 53.76 million years ago into four separate clades. Clade C (corresponding to species B) evolved from Sri Lanka and Myanmar 17.11 million years ago in the Miocene period. Clade D evolved 17.11 million years ago in the Miocene period from Thailand. Clade E includes individuals from Vietnam and Cambodia evolved together 11.92 million years ago. Indonesian species have separately evolved, with their MRCA aging back to 13.54 million years ago.

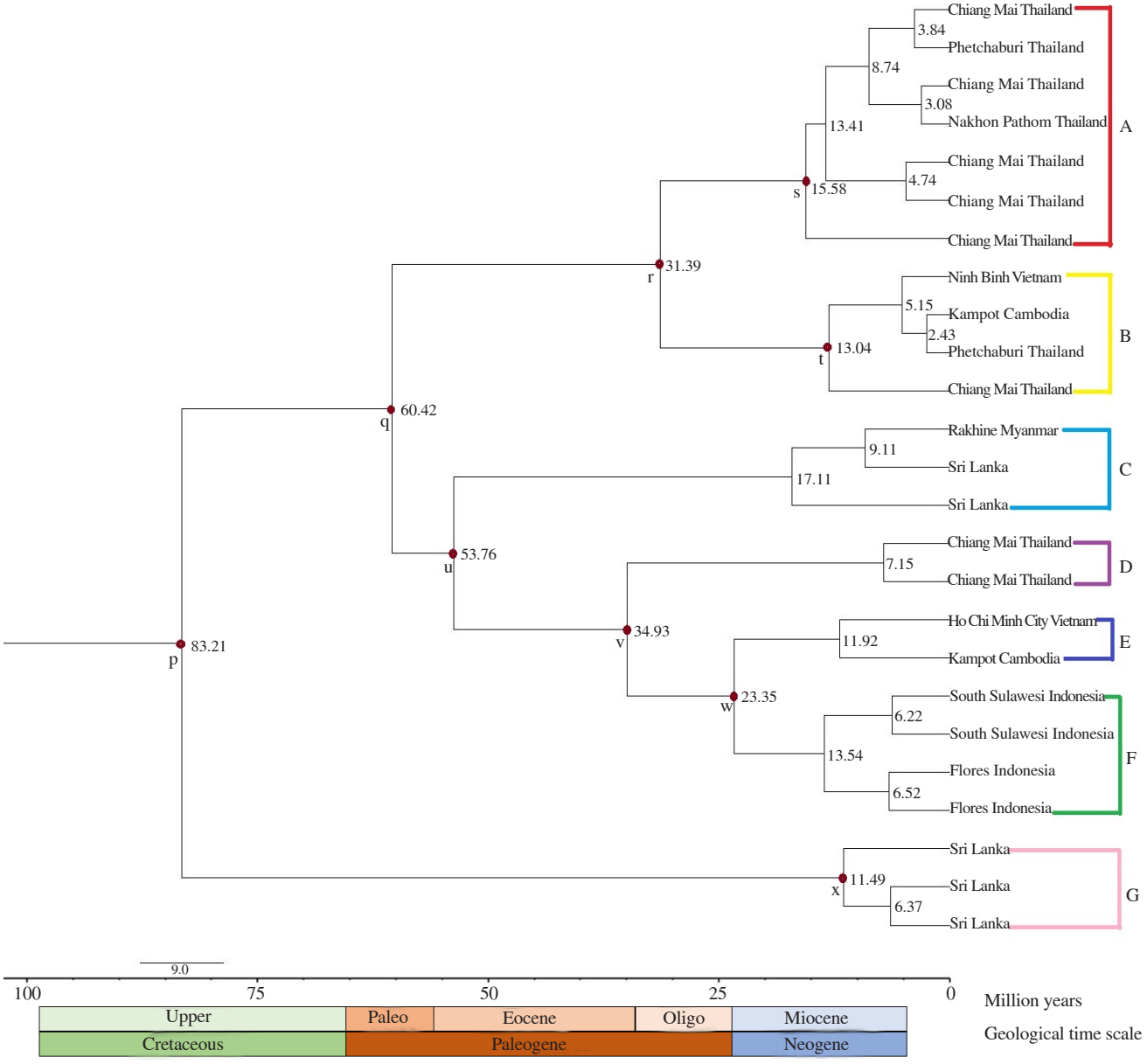

Figure 4. Evolutionary timescale for *Anopheles subpictus* species using ITS2 and COI sequences generated by BEAST v1.8.2 software. Numbers associated with nodes indicate the average divergence time estimated in million years. Geological time scale includes upper Cretaceous, Paleocene, Eocene, Oligocene, and Miocene periods. Posterior probability values, mean values of diversification times and 95% highest posterior density (HPD) of each node are indicated in Table 6.

Table 6. Posterior probability, mean values of diversification times and 95% HPD of nodes of ITS2 and COI concatenated sequence divergence time scale.

Node	Posterior probability	Mean (million years)	95% HPD (million years)
p	1	83.21	81.35-85.23
q	0.34	60.42	37.32-83.47
r	0.94	31.39	10.04-56.80
s	0.68	15.58	3.22-33.72
t	1	13.04	1.91-28.55
u	0.39	53.76	29.24-81.32
v	0.99	34.93	12.26-60.78
w	0.52	23.35	7.40-42.97
x	1	11.49	1.30-26.59

HPD: highest posterior density.

4. Discussion

Identifying fixed molecular forms or isomorphic species and their distribution in different geographical locations is crucial for epidemiological significance. Scientists have shifted from morphological techniques towards molecular approaches as a more accessible and successful technique to identify sibling species of a species complex[13,43]. Studies have revealed the importance and accountability of utilizing COI and ITS2 sequences to study the sibling species status of Anophelinae[13,28]. Especially ITS2 is the most variable region of nuclear DNA and COI is highly informative in insect groups[36]. Yet, as Singh and Vashist[44] mentioned, a combined approach such as COI-ITS2 can be used to identify anopheline species accurately. To the best of our knowledge, this is the first study to use these two markers as concatenated sequences to investigate the phylogenetic relationship, phylogeographic distribution, and diversification time of sibling species of *An. subpictus* species complex by maximum utilization of data in the GenBank at NCBI up to date.

According to the Bayesian Inference, all 3 phylogenetic trees derived using ITS2, COI and concatenated sequences revealed the clear separation of *An. subpictus* population into two genetically distinct clades with 100% posterior probability. The basal clade includes only Sri Lankan species in the concatenated analysis (Figure 1) and Sri Lankan and Indian species in the ITS2 analysis (Supplementary Figure 1), while one individual from Schiphol-Netherlands appeared other than Sri Lanka and India in the COI analysis (Supplementary Figure 2), which is a new finding. The basal clades described above in all 3 trees correspond to the *An. subpictus* species A[35]. According to this study, sibling species B is found only in Sri Lanka, India, and Myanmar. Also, Sri Lankan and Indian species are not observed in any other clade of the phylogenetic analysis. This confirms the presence of only two molecular forms of *An. subpictus* species complex corresponding to sibling species A and B in both Sri Lanka and India, where species A is highly diverged from the rest of the species.

Previous studies suggest that sibling species A and B found in sympatry in Southern Asia are genetically isolated, whereas species B is mainly found in the coastal areas[13,16]. Species A in this study includes individuals from Puducherry, Tamil Nadu, Haryana, Punjab, Mogra, Goa and Orissa of India; Pasyala, Monaragala, Unichchai and Suthumalai of Sri Lanka; Schiphol of Netherlands. Species B includes individuals from Goa of India; Kallady, Kilinochchi, Sarasalai, Suthumalai, Hambantota coastal areas of Sri Lanka; and Rakhine of Myanmar. It clearly shows that certain geographical regions like Goa of India and Suthumalai of Sri Lanka are shared between both species A and B. Thus, this study disagrees with the strict renaming of species A and B as inland and coastal species. As Surendran *et al.*[30] mentioned, it is incorrect to strictly classify based on habitat of subpictus species B, since it is found in both inland and coastal regions.

The *An. subpictus* species in other Asian countries, including Vietnam, Cambodia, Thailand, Indonesia, Papua New Guinea, and the Philippines, are phylogenetically closer to species B than A. Both COI and concatenated analyses revealed that Thai species fall into 4 clades, where 3 clades (A, B, and C in Supplementary Figure 2 and B, C, and D in Figure 1) of Thai species in Chiang Mai, Nakhon Pathom, Phetchaburi, Ratchaburi are closer to Vietnam species in Ninh Binh, and Bac Lieu. They show a sister relationship to sibling species B. The remaining clade (E in Supplementary Figure 2 and G in Figure 1) of Thai species in Chiang Mai and Phang Nga are closer to Vietnam species in Ho Chi Minh City, Cambodian species in Kampot, and Indonesian species in South Sulawesi and Flores. They are more diverged and show a paraphyletic relationship to sibling species B. Hence, this study suggests the presence of two diverged molecular forms (where one is similar to sibling species B) of *An. subpictus* species in Thailand and Vietnam rather than three, as mentioned by Wilai *et al.*[35]. Philippines species falls completely into a separate clade different from both A and B.

Haplotype networks also support the above finding that sibling species A is highly diverged from the rest, and is only found in Sri Lanka, India (haplogroup A in Figure 2), and the Netherlands (haplogroup D in Figure 3). Generally, significant differentiation of the ITS2 marker indicates discrete species. However, recently diverged species may show little or no divergence at this marker, resulting in similar ITS2 sequences but highly divergent COI sequences[35]. Thus, the haplotype network of the COI study (Figure 3) has become more informative than the ITS2 study (Figure 2). It confirms that *An. subpictus* species in Vietnam, Cambodia, and Thailand belongs to two different haplogroups (B and G in Figure 3) that even contain shared haplotypes. The highest number of private haplotypes was recorded in the Philippines, with the least mutation distance (Supplementary Table 9). It indicates a large effective population size and restricted gene flow from that area due to a population adapted to local conditions with no intraspecific differentiation giving rise to sympatric speciation[45].

High Hd values were obtained for all geographical locations in both ITS2 (Supplementary Table 8) and COI (Supplementary Table 9) analyses. Neutrality tests carried out for both sequences (Table 3) also resulted in high haplotype diversity (Hd > 0.8), indicating a large effective population size and extensive genetic diversity of mosquitoes, with high gene flow among local populations due to favourable environmental factors for breeding, irrespective of the mosquito control campaigns[46]. Also, high nucleotide diversity (π > 0.08) and a high average number of nucleotide differences (k > 29) imply the presence of a long-established mosquito population that has not undergone recent bottlenecks.

Negative Tajima's D and positive Fu's F values of ITS2 analysis signify allele deficiency and the presence of rare alleles at low frequencies. Positive Tajima's D and negative Fu's F values of COI analysis indicate the excess of multiple alleles and low-frequency haplotypes[47,48]. However, the condition of the null hypothesis being true due to non-significant Tajima's D values revealed that populations evolve neutrally without any purifying selection, population expansion, and selective sweeps[21]. Thus, the study suggests that *An. subpictus* species complex distributed in the Southeast Asian countries might have undergone older bottlenecks due to habitat fragmentation, intense control efforts, and disappearance of vertebrates[45], but currently is at mutation-drift equilibrium, maintaining a large stable population regardless of the population fluctuations[46].

Genetic differentiation between countries was calculated using pairwise F_{ST} values of ITS2 (Table 4) and COI (Table 5) sequences. Both analyses show that genetic subdivision among the countries: Sri Lanka, India, and Netherlands is very low or absent due to the least F_{ST} values resulting in similar allele frequencies[49]. However, a substantial genetic differentiation is observed among the countries: Cambodia-Papua New Guinea, Indonesia-Papua New Guinea, Indonesia-Cambodia, Myanmar-Papua New Guinea, Myanmar-Cambodia, and Myanmar-Indonesia, probably due to isolation by large geographical distance resulting in different allele frequencies from spatially limited gene flow[50].

The Culicidae fossil records are rare, and only two *Anopheles* records are available where the oldest fossil record is *Anopheles (Nyssorhynchus) dominicans* from the Late Eocene (33.9-41.3 million years ago), and the most recent is *An. rottensis* from the Late Oligocene (13.8-33.9 million years ago)[37,51]. However, the usage of these records is problematic due to the dating techniques applied[51]. Therefore, the diversification age of MRCA of the *Anopheles* genus, which is 83.23 million years ago, was used to calibrate the phylogeographic trees[41]. The link between anophelines and the malaria parasites is still controversial, and it is thought that the vector-parasite relationship is a recent one, supporting the hypothesis that mosquito susceptibility to malaria infection is a primitive characteristic[37].

The phylogeographic analyses of this study support the above dating of the *Anopheles* genus, giving 81.23-85.16 million years ago in the ITS2 study (Supplementary Figure 3), 83.3-85.22 million years ago in the COI (Supplementary Figure 4) study, and 81.35-85.23 million years ago in the concatenated study (Figure 4) respectively with 95% HPD. It also reveals that *An. subpictus* species A evolved earlier during the Paleocene period, ranging from 37.60-81.86 million years ago in the ITS2 study (Supplementary Figure 3), and 45.37-82.46 million years ago in the COI study (Supplementary Figure 4), whereas species B evolved later during the Eocene period ranging from 22.12-67.60 million years ago in the ITS2 study (Supplementary Figure 3), and 39.32-79.50 million years ago in the COI study (Supplementary Figure 4). Consideration of the concatenated study is not that reliable in determining the ages of sibling species due to the lesser number of sequences.

Overall, the study confirms the presence of two genetically divergent molecular forms of *An. subpictus* species complex, where species A evolved around 37-82 million years ago in Sri Lanka, India, and Netherlands and species B evolved around 22-79 million years ago in Sri Lanka, India, and Myanmar. Sri Lanka and India have only species A and B. Vietnam, Thailand, and Cambodia have two molecular forms, and one is phylogenetically similar to species B. Other molecular form of *An. subpictus* species present in the above countries and in Indonesia and the Philippines, which evolved recently, is more likely to be another molecular form of *An. subpictus* different from species A and B.

Further molecular studies are favoured using ITS2 sequences of *An. subpictus* species from countries like the Philippines, where COI sequences are already analyzed. Then, a more informative concatenated tree using both sequences would be possible. Besides that, it is recommended that a multi-locus analysis using several nuclear and mitochondrial DNA in addition to ITS2 and COI be carried out, which will help to confirm the finding.

In conclusion, the importance of understanding the genetic diversity, phylogenetic and phylogeographic relationships of *An. subpictus* species complex using ITS2 and COI sequences was focused. *An. subpictus* species complex consists of more than two genetically divergent molecular forms. Species A evolved earlier and is highly divergent from the rest. Sri Lanka and India contain only species A and B. Vietnam, Thailand, and Cambodia have two molecular forms, and one is phylogenetically similar to species B. Indonesia and the Philippines contain a different molecular form. High genetic and haplotype diversity values indicate a large effective population size. The genetic differentiation of populations increases with the isolation by geographical distance. *An. subpictus* species complex distributed in Southeast Asian countries might have undergone older bottlenecks, but currently it is at mutation-drift equilibrium, maintaining a large stable population regardless of population fluctuations. Correct identification of sibling species

of a particular mosquito species is crucial to implementing effective vector control strategies.

Conflict of interest statement

All authors declared no conflict of interest.

Acknowledgements

All the members of the Centre for Biotechnology and Genetics and Molecular Biology Unit at the University of Sri Jayewardenepura, Sri Lanka were greatly acknowledged for the assistance rendered during this study.

Funding

The authors received no extramural funding for the study.

Authors' contributions

L.S.M. did the literature search, data acquisition and manuscript preparation. N.K.D. provided the concept, supervised the project, and reviewed the manuscript. Both L.S.M. and M.M.W. conducted experimental studies and statistical analysis. Both L.S.M. and S.R. conducted the data analysis. Both L.S.M. and N.K.D. edited the manuscript. L.S.M., M.M.W. and N.K.D. designed the research. All L.S.M., M.M.W., S.R. and N.K.D. defined the intellectual content and guaranteed the study.

References

[1] Manguin S, Boëte C. Global impact of mosquito biodiversity, human vector-borne diseases and environmental change. In: *The importance of biological interactions in the study of biodiversity*. London: IntechOpen; 2011, p. 27-50.

[2] Owino EA. Why Kenya should worry about *Anopheles stephensi*. *Asian Pac J Trop Med* 2023; **16**(3): 99-101.

[3] Karunaweera ND, Galappaththy GN, Wirth DF. On the road to eliminate malaria in Sri Lanka: Lessons from history, challenges, gaps in knowledge and research needs. *Malar J* 2014; **13**(59): 1-10.

[4] Harischandra IN, Dassanayake RS, De Silva BGDNK. Mitochondrial-COII sequence polymorphism reflects spatial genetic clustering of *Anopheles culicifacies* sibling species E in Sri Lanka. *J Vector Borne Dis* 2020; **57**(2): 139-146.

[5] Venkatesan P. The 2023 WHO world malaria report. *Lancet Microbe* 2024; **5**(3): e214.

[6] Simac JN, Badar S, Farber J, Brako MO, Giudice-Jimenez RL, Raspa S, et al. Malaria elimination in Sri Lanka. *J Health Spec* 2017; **5**(2): 60-65.

[7] Ibáñez-Justicia A, Smitz N, Den Hartog W, Van de Vossenberg B, De Wolf K, Deblauwe I, et al. Detection of exotic mosquito species (Diptera: Culicidae) at international airports in Europe. *Int J Environ Res Public Health* 2020; **17**(10): 3450.

[8] Jayatunga DPW, Harischandra IN, Chandrasekharan NV, De Silva BGDNK. Compensatory base changes reveal sexual incompatibility among members of the *Anopheles subpictus* Sensu Lato (Diptera: Culicidae) species complex in Sri Lanka. *Life* 2021; **11**(3): 211.

[9] Gunathilaka PADHN. Annotated checklist and review of the mosquito species (Diptera: Culicidae) in Sri Lanka. *J Insect Biodivers* 2018; **7**(3): 38-50.

[10] Rueda LM, Debboun M. Taxonomy, identification, and biology of mosquitoes. In: *Mosquitoes, communities, and public health in Texas*. New York: Academic Press; 2019, p. 3-7.

[11] Chhilar JS, Chaudhry S. Phylogenetic analysis of *Anopheles (Cellia) subpictus* Grassi using rDNA-ITS2 sequence. *Proc Zool Soc* 2012; **65**(1): 1-10.

[12] Surendran S, Ramasamy R. The *Anopheles culicifacies* and *Anopheles subpictus* species complexes in Sri Lanka and their implications for malaria control in the country. *Trop Med Health* 2010; **38**(1): 1-11.

[13] Sindhania A, Das MK, Sharma G, Surendran SN, Kaushal BR, Lohani HP, et al. Molecular forms of *Anopheles subpictus* and *Anopheles sundaicus* in the Indian subcontinent. *Malar J* 2020; **19**(417): 1-17.

[14] Suguna SK, Rathinam KG, Rajavel AR, Da VDH. Morphological and chromosomal descriptions of new species in the *Anopheles subpictus* complex. *Med Vet Entomol* 1994; **8**(1): 88-94.

[15] Subbarao SK. *Anopheline species complexes in South-East Asia*. New Delhi: WHO Regional Office for South-East Aisa; 1998, p. 1-69.

[16] Jude PJ, Ramasamy R, Surendran SN. Bionomic aspects of the *Anopheles subpictus* species complex in Sri Lanka. *J Insect Sci* 2014; **14**(1): 97.

[17] Singh SP. Morphotaxonomical studies to identify the member of the *Anopheles subpictus* Grassi (Diptera: Culicidae) species complex in villages of District Mewat Haryana State, India. *Int Interdiscip Res J* 2014; **4**(1): 239-244.

[18] Abhayawardana TA, Wijesuriya SR, Dilrukshi RK. *Anopheles subpictus* complex: Distribution of sibling species in Sri Lanka. *Indian J Malariol* 1996; **33**(2): 53-60.

[19] Cooper RD, Edstein MD, Frances SP, Beebe NW. Malaria vectors of Timor-Leste. *Malar J* 2010; **9**(40): 1-11.

[20] Surendran SN, Singh OP, Jude PJ, Ramasamy R. Genetic evidence for malaria vectors of the *Anopheles sundaicus* complex in Sri Lanka with morphological characteristics attributed to *Anopheles subpictus* species B. *Malar J* 2010; **9**(343): 1-9.

[21] Sivabalakrishnan K, Kanapathy G, Jeyadas TTP, Surendran SN. Population genetic structure of *Anopheles subpictus* species B using COII and Cytb markers. *J JSA* 2020; **2**(1): 38-47.

[22] Jayatunga DPW, Harischandra IN, Chandrasekharan NV, De Silva NK. Alterations and interchange of morphometric characters in different life cycle stages with reference to genomic variations of *Anopheles subpictus*

(Diptera; Culicidae) sibling species complex in Sri Lanka. *Insects* 2018; **9**(3): 89.

[23]Hill SM, Crampton JM. DNA-based methods for the identification of insect vectors. *Ann Trop Med Parasitol* 1994; **88**(3): 227-250.

[24]Sallum MAM, Schultz TR, Foster PG, Aronstein K, Wirtz RA, Wilkerson RC. Phylogeny of Anophelinae (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences. *Syst Entomol* 2002; **27**(3): 361-382.

[25]Zomuanpuii R, Ringngheti L, Brindha S, Gurusubramanian G, Kumar NS. ITS2 characterization and *Anopheles* species identification of the subgenus *Cellia*. *Acta Trop* 2013; **125**(3): 309-319.

[26]Murugan K, Vadivalagan C, Karthika P, Panneerselvam C, Paulpandi M, Subramaniam J, et al. DNA barcoding and molecular evolution of mosquito vectors of medical and veterinary importance. *Parasitol Res* 2016; **115**: 107-121.

[27]Khoshdel-nezamiha F, Vatandoost H, Oshaghi MA, Azari-hamidian S, Mianroodi RA, Dabiri F, et al. Molecular characterization of mosquitoes (Diptera: Culicidae) in Northwestern Iran by using rDNA-ITS2. *Jpn J Infect Dis* 2016; **69**(4): 319-322.

[28]Weeraratne TC, Surendran SN, Reimer LJ, Wondji CS, Perera MD, Walton C, et al. Molecular characterization of *Anopheline* (Diptera: Culicidae) mosquitoes from eight geographical locations of Sri Lanka. *Malar J* 2017; **16**(234): 1-14.

[29]Singh D, Rai MK, Vashist D. Molecular studies on Indian mosquitoes (Diptera: Culicidae): A review. *Int J Entomol Res* 2018; **3**(2): 36-103.

[30]Surendran SN, Sarma DK, Jude PJ, Kemppainen P, Kanthakumaran N, Gajapathy K, et al. Molecular characterization and identification of members of the *Anopheles subpictus* complex in Sri Lanka. *Malar J* 2013; **12**(304): 1-11.

[31]Weeraratne TC, Surendran SN, Walton C, Karunaratne SHPP. Genetic diversity and population structure of malaria vector mosquitoes: *Anopheles subpictus*, *Anopheles peditaeniatus*, and *Anopheles vagus* in five districts of Sri Lanka. *Malar J* 2018; **17**(271): 1-8.

[32]Coyne JA. The evolutionary genetics of speciation. *Philos Trans R Soc B Biol Sci* 1998; **353**(1366): 287-305.

[33]Provine WB. Ernst Mayr: Genetics and speciation. *Genetics* 2004; **167**(3): 1041-1046.

[34]Mayr E. Origin and history of some terms in systematic and evolutionary biology. *Syst Zool* 1978; **27**(1): 83-88.

[35]Wilai P, Ali RS, Saingamsook J, Saeung A, Junkum A, Walton C, et al. Integrated systematics of *Anopheles subpictus* (Diptera: Culicidae) in the Oriental Region, with emphasis on forms in Thailand and Sulawesi, Indonesia. *Acta Trop* 2020; **208**: 105503.

[36]Karimian F, Oshaghi MA, Sedaghat MM, Waterhouse RM, Vatandoost H, Hanafi-Boid AA, et al. Phylogenetic analysis of the Oriental-Palearctic-Afrotropical members of *Anopheles* (Culicidae: Diptera) based on nuclear rDNA and mitochondrial DNA characteristics. *Jpn J Infect Dis* 2014; **67**(5): 361-367.

[37]Moreno M, Marinotti O, Krzywinski J, Tadei WP, James AA, Achee NL, et al. Complete mtDNA genomes of *Anopheles darlingi* and an approach to anopheline divergence time. *Malar J* 2010; **9**(1): 1-13.

[38]Lorenz C, Alves JMP, Foster PG, Suesdek L, Sallum MAM. Phylogeny and temporal diversification of mosquitoes (Diptera: Culicidae) with an emphasis on the Neotropical fauna. *Syst Entomol* 2021; **46**(4): 798-811.

[39]Depaquit J, Lienard E, Verzeaux-Griffon A, Ferte H, Bounamous A, Gantier JC, et al. Molecular homogeneity in diverse geographical populations of *Phlebotomus papatasi* (Diptera, Psychodidae) inferred from ND4 mtDNA and ITS2 rDNA. Epidemiological consequences. *Infect Genet Evol* 2008; **8**(2): 159-170.

[40]Rangel-Gamboa L, Martinez-Hernandez F, Maravilla P, Arenas-Guzman R, Flisser A. Update of phylogenetic and genetic diversity of *Sporothrix schenckii* sensu lato. *Med Mycol J* 2015; **54**(3): 248-255.

[41]Martinez-Villegas L, Assis-Geraldo J, Koerich LB, Collier TC, Lee Y, Main BJ, et al. Characterization of the complete mitogenome of *Anopheles aquasalis*, and phylogenetic divergences among *Anopheles* from diverse geographic zones. *PLoS One* 2019; **14**(9): 1-22.

[42]Reidenbach KR, Cook S, Bertone MA, Harbach RE, Wiegmann BM, Besansky NJ. Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evol Biol* 2009; **9**(298): 1-14.

[43]Harbach RE. The classification of genus *Anopheles* (Diptera: Culicidae): A working hypothesis of phylogenetic relationships. *Bull Entomol Res* 2004; **94**(6): 537-553.

[44]Singh D, Vashist D. DNA barcoding as a tool supporting species identification and molecular phylogeny for Indian *Anopheles* (Diptera: Culicidae: Anophelinae). *J Entomol Res* 2020; **44**(3): 409-416.

[45]Mayoke A, Muya SM, Bateta R, Mireji PO, Okoth SO, Onyoyo SG, et al. Genetic diversity and phylogenetic relationships of tsetse flies of the palpalis group in Congo Brazzaville based on mitochondrial *cox1* gene sequences. *Parasit Vectors* 2020; **13**(253): 1-16.

[46]Dharmarathne HA, Weerasena OV, Perera KL, Galhena G. Genetic characterization of *Aedes aegypti* (Diptera: Culicidae) in Sri Lanka based on *COI* gene. *J Vector Borne Dis* 2020; **57**(2): 157.

[47]Dogan I, Dogan N. Statistical tests for neutrality: Review. *Turkiye Klin J Biostat* 2017; **9**(2): 167-174.

[48]Ebrahimi S, Bordbar A, Parvizi P. Genetic dynamics in the sand fly (Diptera: Psychodidae) nuclear and mitochondrial genotypes: Evidence for vector adaptation at the border of Iran with Iraq. *Parasit Vectors* 2016; **9**(319): 1-13.

[49]Jacquot MP. Genomic diversity of pathogenic bacteria of the *Borrelia burgdorferi* species complex: Evolution and molecular epdemiology. Doctoral dissertation. Universite Blaise Pascal-Clermont-Ferrad Ⅱ University; 2014.

[50]Jensen JL, Bohonak AJ, Kelley ST. Isolation by distance, web service. *BMC Genet* 2005; **6**(1): 1-6.

[51]Freitas LA, Russo CAM, Voloch CM, Mutaquiha OCF, Marques LP, Schrago CG. Diversification of the genus *Anopheles* and a neotropical clade from the Late Cretaceous. *PLoS One* 2015; **10**(8): 1-12.

Publisher's note

Edited by Zhang Q, Lei Y, Pan Y