

RESEARCH

Open Access



Molecular characterization of Anopheline (Diptera: Culicidae) mosquitoes from eight geographical locations of Sri Lanka

Thilini C. Weeraratne¹, Sinnathambi N. Surendran², Lisa J. Reimer³, Charles S. Wondji³, M. Devika B. Perera⁴, Catherine Walton⁵ and S. H. P. Parakrama Karunaratne^{1,6*}

Abstract

Background: Genus *Anopheles* is a major mosquito group of interest in Sri Lanka as it includes vectors of malaria and its members exist as species complexes. Taxonomy of the group is mainly based on morphological features, which are not conclusive and can be easily erased while handling the specimens. A combined effort, using morphology and DNA barcoding (using the markers *cytochrome c oxidase subunit I (COI)* gene and internal transcribed spacer 2 (ITS2) region, was made during the present study to recognize anophelines collected from eight districts of Sri Lanka for the first time.

Methods: *Cytochrome c oxidase subunit I* and ITS2 regions of morphologically identified anopheline mosquitoes from Sri Lanka were sequenced. These sequences together with GenBank sequences were used in phylogenetic tree construction and molecular characterization of mosquitoes.

Results: According to morphological identification, the field-collected adult mosquitoes belonged to 15 species, i.e., *Anopheles aconitus*, *Anopheles annularis*, *Anopheles barbirostris*, *Anopheles culicifacies*, *Anopheles jamesii*, *Anopheles karwari*, *Anopheles maculatus*, *Anopheles nigerrimus*, *Anopheles pallidus*, *Anopheles peditaeniatus*, *Anopheles pseudo-jamesii*, *Anopheles subpictus*, *Anopheles tessellatus*, *Anopheles vagus*, and *Anopheles varuna*. However, analysis of 123 *COI* sequences (445 bp) (16 clades supported by strong bootstrap value in the neighbour joining tree and inter-specific distances of >3%) showed that there are 16 distinct species. Identity of the morphologically identified species, except *An. subpictus*, was comparable with the DNA barcoding results. *COI* sequence analysis showed that morphologically identified *An. subpictus* is composed of two genetic entities: *An. subpictus* species A and species B (inter-specific K2P distance 0.128). All the four haplotypes of *An. culicifacies* discovered during the present study belonged to a single species. ITS2 sequences (542 bp) were obtained for all the species except for *An. barbirostris*, *An. subpictus* species B, *An. tessellatus*, and *An. varuna*. Each of these sequences was represented by a single species-specific haplotype.

Conclusions: The present study reflects the importance and feasibility of *COI* and ITS2 genetic markers in identifying anophelines and their sibling species, and the significance of integrated systematic approach in mosquito taxonomy. Wide distribution of malaria vectors in the country perhaps indicates the potential for re-emergence of malaria in the country.

Keywords: *Anopheles*, DNA barcoding, *COI*, ITS2, Mosquitoes, Taxonomy, Sri Lanka

*Correspondence: shppk@pdn.ac.lk; shppkaru@yahoo.com

¹ Department of Zoology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka

Full list of author information is available at the end of the article

Background

Species identification is essential for many biological studies. Morphological characters and DNA barcoding are two major approaches used in species identification. The modern system of taxonomy, referred to as 'DNA barcoding', which is based on molecular techniques, has become increasingly popular as it potentially produces results with very high precision and accuracy within a short period of time, compared to traditional morphology-based taxonomy. DNA barcoding is currently used as a guide in biodiversity conservation and management programmes due to its ability to confirm the identity of organisms, to infer the presence of new/endangered/invasive/cryptic species and to estimate genetic diversity among individuals of a species [1].

Due to their importance in disease transmission, mosquitoes are one of the most intensely barcoded group of insects, although many species are still to be identified and barcoded [2]. Further, mosquito species may exist as species complexes and sibling species of a given complex often show different feeding behaviour, biting habits, vector competence, distribution, and thereby different vectorial capacities. Accurate and precise identification of mosquito species and sibling species is vital in planning vector control strategies. Since these sibling species cannot be distinguished by external morphological features, molecular taxonomy is required to obtain reliable information about the vector.

The most commonly used molecular markers/barcode regions for mosquito barcoding studies are the *cytochrome c oxidase subunit I (COI)* gene located in the mitochondrial genome (mtDNA) followed by the internal transcribed spacer 2 (ITS2), a region from the nuclear ribosomal DNA. Several workers have used *COI* as the only marker in mosquito species recognition and in investigating their molecular evolution [3–9]. However, some studies have shown that mosquito identification through *COI* alone is not always sufficient to make precise conclusions and the need for genetic examination using analysis of faster evolving regions, such as ITS2, has been emphasized [8]. ITS2 regions alone have been used in distinguishing closely related mosquito species belonging to various genera such as *Anopheles* [10], *Culex* [11] and *Aedes* [12]. Moreover, it has been indicated that more reliable information about species could be obtained if several molecular markers are used simultaneously [13, 14].

The mosquito fauna in Sri Lanka comprises 141 species belonging to 17 genera [15–17]. Out of these, 16.31% (23 species) are anophelines. An accurate and detailed species characterization of members of this genus is required. In Sri Lanka, major malaria vector species

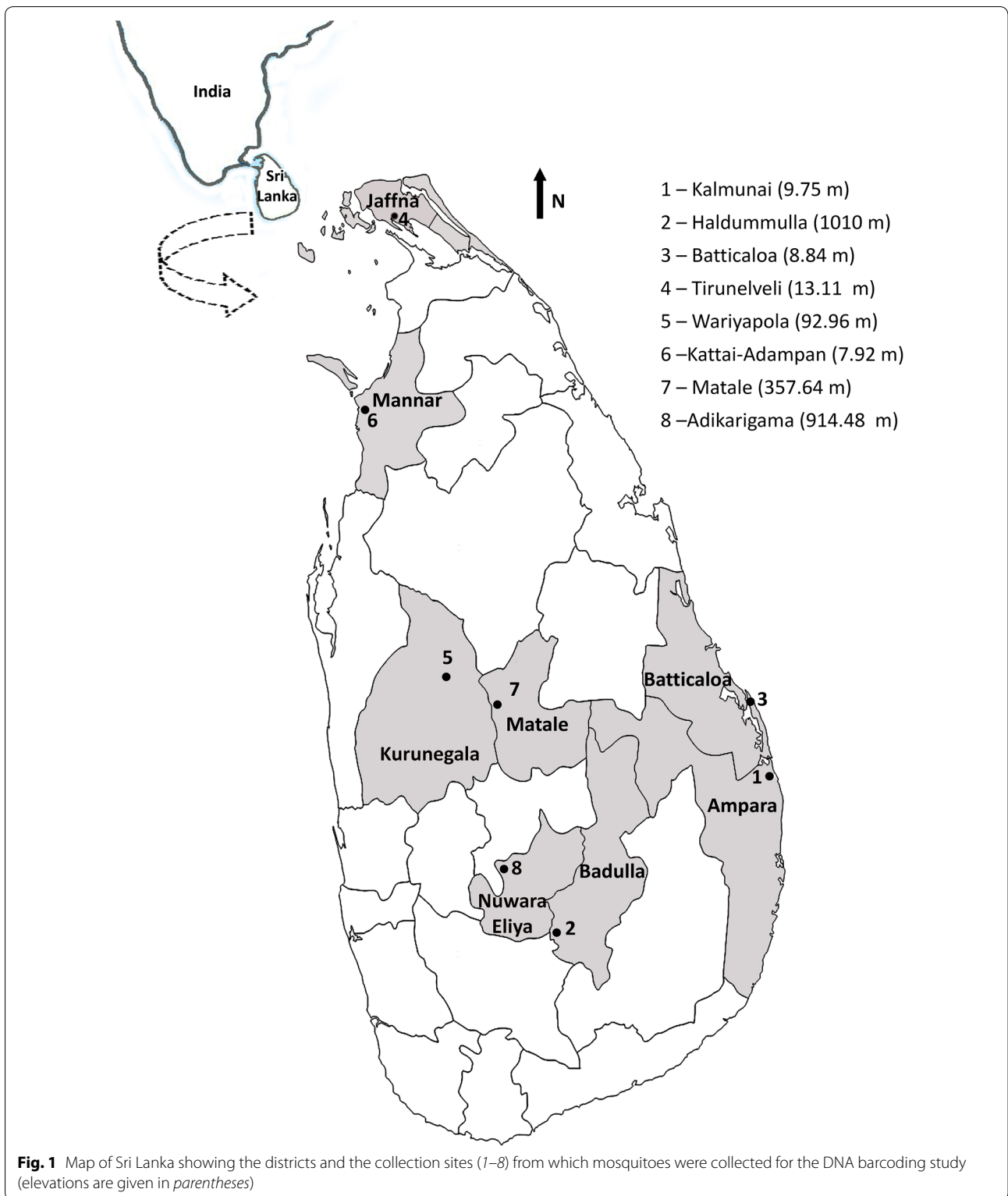
Anopheles culicifacies and *Anopheles subpictus* exist as species complexes [18–20]. Among the two *An. culicifacies* species B and E found in Sri Lanka, species E is the primary vector of malaria in the country [18]. *Anopheles subpictus* (four sibling species viz. A, B, C and D) is considered the major vector of malaria in Jaffna district and the secondary vector in other parts of the country [21]. Several studies have found higher malaria sporozoite rates in *An. subpictus* than *An. culicifacies* during the peak transmission period in Jaffna district [22, 23]. Further, *Anopheles aconitus*, *Anopheles annularis*, *Anopheles barbirostris*, *Anopheles nigerrimus*, *Anopheles pallidus*, *Anopheles peditaeniatus*, *Anopheles tessellatus*, *Anopheles vagus*, and *Anopheles varuna* are considered potential vectors of malaria [24, 25].

In Sri Lanka, mosquito identification has mainly relied on taxonomic keys based on external morphological features of adult and larval stages. However, a few studies have attempted to use DNA barcodes to describe a few species, i.e., *An. culicifacies* [26, 27], *An. subpictus* [18], *An. barbirostris* [28], and *Anopheles sundaicus* [29], from limited geographic locations. The country urgently needs a well planned molecular and morphology-based taxonomy to characterize its anopheline population. The present study aims to provide both morphological and molecular taxonomic (DNA barcoding) information about *Anopheles* mosquitoes collected from different geographical areas of Sri Lanka. The importance of species identification using two molecular markers *COI* and *ITS2* is also discussed.

Methods

Study sites

Mosquitoes were collected from eight study sites located in eight districts in Sri Lanka: Kalmunai in Ampara district, Haldummulla in Badulla district, Batticaloa in Batticaloa district, Tirunelveli in Jaffna district, Wariyapola in Kurunegala district, Matale in Matale district, Kattai-Adampan in Mannar district, and Adikarigama in Nuwara-Eliya district (Fig. 1). These sites are located in five different climatic zones, i.e., Adikarigama in up-country wet zone (>900 m elevation, >2500 mm rainfall); Matale in mid-country wet zone (300–900 m elevation, >2500 mm rainfall); Haldummulla in up-country intermediate zone (>900 m elevation, 1750–2500 mm rainfall); Wariyapola in low-country intermediate zone (0–300 m elevation, 1750–2500 mm rainfall); Kalmunai, Batticaloa, Tirunelveli, Katti-Adampan in low-country dry zone (0–300 m elevation, <1750 mm rainfall, with a distinct dry period) (Fig. 1). Habitat types of each study site are presented in Additional file 1.



Mosquito collection

Adult mosquitoes were collected using cattle-baited traps and backpack aspirators. Dried specimens were used in

morphological and molecular identifications. Morphological identifications to species or complex level were done using standard taxonomic keys [16, 30]. Voucher

samples of specimens from each species were mounted using standard techniques for a reference collection at the Department of Zoology, Faculty of Science, University of Peradeniya, Sri Lanka. The remainder were dried and stored for molecular characterization.

DNA extraction, polymerase chain reaction (PCR) and sequencing

Genomic DNA was extracted from head and thoracic regions of each individual using nexttec™ DNA Isolation Kits (Nexttec Biotechnologies GmbH, Leverkusen, Germany), according to the manufacturer's protocol. DNA was extracted from a maximum of ten individuals from each morphologically identified species, depending on the availability.

A region of the *COI* gene was amplified using forward primer C1-J-1718 (5'-GGAG GATTGGAAATTGATTAGTTCC-3') and reverse primer C1-N-2191 (5'-CCCG-GTAAATTAATAAATAAATTC-3') [31] and the ITS2 was amplified using forward primer 5'-ATCACTCG-GCTCATGGATCG-3' and reverse primer 5'-ATGCT-TAAATTTAGGGGTAGTC-3' [32]. Each amplification was performed in 15 µl that included 1 µl DNA template, 1.5 µl 10× KAPA buffer A, 0.12 µl KAPA taq, 0.12 µl 2.5 mM dNTPmix, 0.75 µl 50 mM MgCl₂, 0.51 µl each primer (10 mmol), and 10.49 µl ddH₂O. The PCR parameters were 95 °C for 5 min and 35 cycles of 94 °C for 30 s, 51 °C (for *COI*)/55 °C (for ITS2) for 40 s and 72 °C for 45 s, followed by a final extension step of 72 °C for 10 min. PCR products were run in 1.5% agarose gel stained with Medori green and visualized in a gel imaging system.

PCR products showing positive clear bands were purified using QIAquick® PCR Purification kits according to the manufacturers' protocol. A maximum of three PCR positive samples of each species from each district were sequenced bidirectionally at Source Bioscience, Nottingham, UK.

DNA sequence analysis

The trace files/chromatograms of *COI* and ITS2 sequences were manually edited using BioEdit software. Sequences of low quality were excluded during data analysis. Consensus sequences were aligned using Clustal W in BioEdit software [33]. Once the alignment was completed, species identification was confirmed by comparison to publicly available sequence data in GenBank using BLAST [34] and the barcode of life database (BOLD) interface [35]. The number of parsimony informative sites, number of variable sites, number of haplotypes, haplotype diversity, and GC content were analysed using the DNA Sequences Polymorphism software (dnaSP) (Version 5.1.10). Mega version 6.0 was used to calculate intra-specific and inter-specific pairwise sequence

divergence using the Kimura-2 parameter distance model [36]. Neighbour joining (NJ) phylogenetic trees of *COI* and ITS2 sequences were constructed in MEGA 6.0 using Kimura-2 parameter distances. Branch support of NJ trees was assessed by bootstrapping with 1000 replicates. Codon positions included 1st + 2nd + 3rd + non-coding regions. All the haplotype sequences of *COI* and generated ITS2 sequences were deposited in GenBank and the accession numbers are given in Additional file 2.

Results

According to morphological identification, the wild mosquito samples collected from all eight study sites belonged to 15 anopheline species: *Anopheles aconitus*, *An. annularis*, *An. barbirostris*, *An. culicifacies*, *An. jamesii*, *An. karwari*, *An. maculatus*, *An. nigerrimus*, *An. pallidus*, *An. peditaeniatus*, *An. pseudojamesi*, *An. subpictus*, *An. tessellatus*, *An. vagus*, and *An. varuna*. The distributions of species among the study sites are given in Table 1. *Anopheles peditaeniatus* and *An. vagus* were the most widespread, occurring at seven out of eight study sites. *Anopheles peditaeniatus* was absent from Adikarigama in Nuwara-Eliya district site while *An. vagus* was absent from Kattai-Adampan in Mannar district. *Anopheles aconitus* and *Anopheles pseudojamesi* were present only in Haldummulla (Badulla district) while *An. maculatus* and *An. tessellatus* were recorded only from Adikarigama (Nuwara-Eliya district) and Wariyapola (Kurunegala district), respectively (Table 1). Anopheline adults were collected using two different collection methods: cattle-baited traps and backpack aspirators. Cattle-baited traps were positive for all the above species whereas backpack aspirators were positive only for *An. culicifacies*, *An. jamesii*, *An. varuna*, and *An. subpictus*.

A total of 123 *COI* sequences were obtained for all 15 anopheline species and 79 *COI* haplotypes were identified. *COI* sequences with a fragment size of 445 bp were used in the phylogenetic analysis. ITS2 sequences were obtained only for 12 species and each of those species had only a single haplotype (12 species-specific ITS2 haplotypes). The ITS2 sequences generated for *An. barbirostris*, *An. tessellatus* and *An. varuna* were not included in the analysis since they were not in good quality. The length of ITS2 fragment varied from 398 bp in *An. maculatus* to 506 bp in *An. nigerrimus*. Fragment sizes obtained from all the species are given in Additional file 2 together with the GenBank accession numbers where the sequences were deposited. The sequence match between these haplotypes and the closest publicly available sequence was always 99–100%. The NJ tree constructed using the 12 ITS2 sequences is shown in Fig. 2. The NJ tree formed 12 distinct clades, each representing single species (supported by 98–100% bootstrap values).

Table 1 Statistical parameters of the COI sequences and the distribution of species in the study sites

Species	n	h	Study sites	Mean pairwise distance	Pairwise distance range	Haplotype diversity
<i>An. aconitus</i>	2	2	Haldummulla	0.003 ± 0.003	–	1.000 ± 0.500
<i>An. annularis</i>	3	3	Haldummulla, Tirunelveli	0.010 ± 0.004	0.012 to 0.009	1.000 ± 0.272
<i>An. barbirostris</i>	5	1	Haldummulla, Wariyapola, KattaiAdampan	0.000 ± 0.000	–	0.000 ± 0.000
<i>An. culicifacies</i>	5	4	Haldummulla, Tirunelveli, Adikarigama	0.012 ± 0.004	0.000 to 0.018	0.900 ± 0.161
<i>An. jamesii</i>	13	7	Haldummulla, Tirunelveli, Wariyapola, KattaiAdampan, Matale	0.001 ± 0.001	0.000 to 0.006	0.731 ± 0.133
<i>An. karwari</i>	2	1	Wariyapola, Batticaloa	–	–	0.000 ± 0.000
<i>An. maculatus</i>	3	3	Adikarigama	0.012 ± 0.005	0.006 to 0.017	1.000 ± 0.272
<i>An. nigerrimus</i>	6	5	Tirunelveli, KattaiAdampan, Matale	0.009 ± 0.003	0.003 to 0.018	0.933 ± 0.122
<i>An. pallidus</i>	16	4	Kalmunai, Haldummulla, Batticaloa, Tirunelveli, Wariyapola	0.002 ± 0.001	0.000 to 0.006	0.525 ± 0.137
<i>An. peditaeniatus</i>	19	10	Kalmunai, Haldummulla, Batticaloa, Tirunelveli, Wariyapola, KattaiAdampan, Matale	0.007 ± 0.002	0.000 to 0.023	0.883 ± 0.056
<i>An. pseudojamesi</i>	2	1	Haldummulla	0.000 ± 0.000	–	0.000 ± 0.000
<i>An. subpictus</i> species A	24	16	Kalmunai, Haldummulla, Batticaloa, Wariyapola, KattaiAdampan	0.006 ± 0.002	0.000 to 0.018	0.938 ± 0.001
<i>An. subpictus</i> species B	2	2	Tirunelveli	0.002 ± 0.002	–	1.000 ± 0.250
<i>An. tessellatus</i>	3	3	Wariyapola	0.006 ± 0.003	–	1.000 ± 0.272
<i>An. vagus</i>	13	12	Kalmunaim, Haldummulla, Batticaloa, Tirunelveli, Wariyapola, Matale, Adikarigama	0.009 ± 0.003	0.000 to 0.017	0.987 ± 0.035
<i>An. varuna</i>	5	5	Tirunelveli, Wariyapola	0.014 ± 0.004	0.009 to 0.029	1.000 ± 0.126

Intra-specific distances were calculated using Kimura 2-parameter distance algorithm

n total number of COI sequences, *h* haplotype number

Cytochrome c oxidase subunit I sequences were AT rich and with an overall G+C content of 0.315. According to polymorphism analysis, there were 149 variable bases (33.9%) of which 130 were parsimony informative and only 17 were singleton mutations. Out of all the nucleotide variations, 78.5% was observed at the third codon position with 117 variable sites and eight singleton mutations. The first position had 25 variable sites and five singleton mutations accounting for 16.8% of all the variable sites. The rest of the variation (4.7%) was observed at the second codon position.

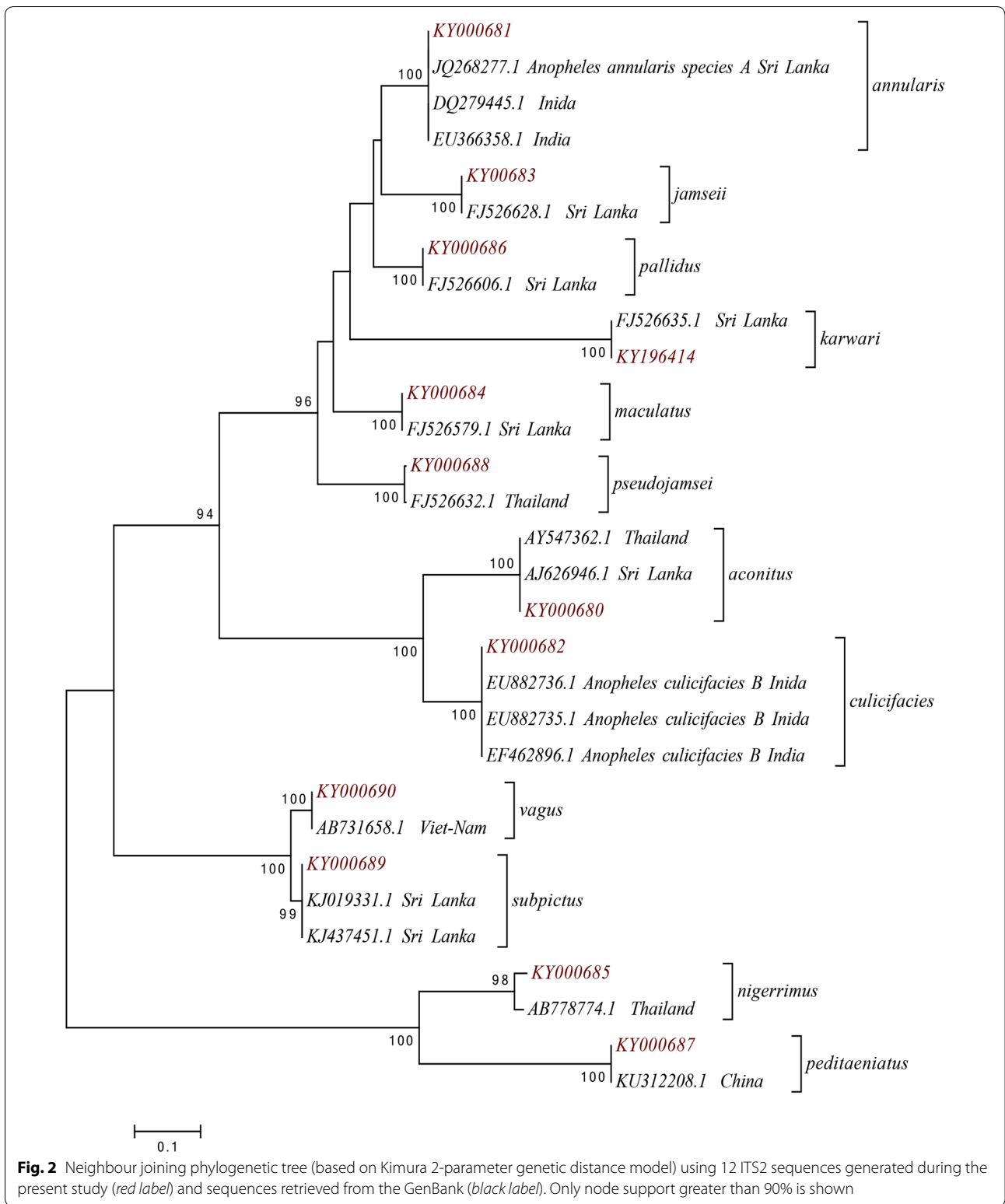
All the 15 morphologically identified *Anopheles* species were grouped into 16 distinct clades in the NJ tree constructed using COI sequences. All of these clades were strongly supported by 99% bootstrap value except *An. peditaeniatus*, which was supported by a 97% bootstrap value. For 15 species, the specimens of each species clustered into a single clade but the specimens of *An. subpictus* fell into two separate clades, as shown in Figs. 3 and 4.

Figure 5 shows the phylogenetic tree constructed using all the 18 *An. subpictus* COI haplotypes obtained during the current study and the sequences retrieved from GenBank. Separation of *An. subpictus* haplotypes into two distinct clusters indicates the presence of two different genetic

entities of *An. subpictus* in Sri Lanka. The two haplotypes of *An. subpictus* from Jaffna (KX644182 and KX644183) formed a separate clade together with *An. subpictus* species B from India (DQ310147.1 and DQ310149.1) sharing 99% sequence similarity according to BLAST search results. All the other 16 *An. subpictus* specimens from other areas clustered into one group (bootstrap value of 100%) with *An. subpictus* species A sequences from Sri Lanka (HQ609132.1, HQ609129.1) and India (DQ310146.1).

Out of these 16, six species, i.e., *An. aconitus*, *An. annularis*, *An. maculatus*, *An. subpictus* species B, *An. tessellatus*, and *An. varuna*, showed a haplotype diversity of 1.000. The lowest haplotype diversity was among *An. pallidus* specimens (0.525 ± 0.137). No polymorphism was shown by *An. barbirostris*, *An. karwari* and *An. pseudojamesi* (Table 1).

The overall mean Kimura-2 parameter genetic distance between these 14 *Anopheles* species was 0.106 ± 0.011. The intra-specific K2P genetic distances was <2%, as given in Table 1. The maximum mean intra-specific distance was for *An. varuna* (1.4%) followed by *An. culicifacies* and *An. maculatus* (1.2%), while *An. jamesii* had the minimum intra-specific K2P distance of 0.1%.





(See figure on previous page.)

Fig. 3 Neighbour joining phylogenetic tree (based on Kimura 2-parameter genetic distance model) of *COI* sequences of all 79 haplotypes of anopheline species collected from Sri Lanka during the study (red label) and the sequences retrieved from the GenBank (black label). *Armigeres subalbatus* and *Culex tritaeniorhynchus* were used as the out-groups. Only node support greater than 90% is shown. The study site/s from which each haplotype was collected are given in parentheses (*Ampa* Ampara, *Badu* Badulla, *Batt* Batticaloa, *Jaff* Jaffna, *Kuru* Kurunegala, *Mann* Mannar, *Mata* Matale, *NE* Nuwara-Eliya)

The inter-specific K2P genetic distances between all the species were more than 2%. The highest K2P value was between *An. jamesii* and *An. barbirostris* (17.2%) and least was between *An. peditaeniatus* and *An. nigerrimus* 7.2%. The inter-specific distances between *An. subpictus* species A and species B was 12.8% (Table 2).

Translation of nucleotide data to amino acids resulted in an alignment of 144 amino acids. The 95th amino acid has changed from isoleucine to valine (or I arrow V) in all specimens of *An. peditaeniatus*, *An. barbirostris*, *An. nigerrimus*, *An. tessellatus*, *An. vagus*, and *An. subpictus* species A. All the other species shared identical AA amino acid sequences.

Discussion

The present study provides morphological and molecular characterization of anopheline species from eight districts across different climatic zones in Sri Lanka. Morphology-based taxonomy recognized 15 *Anopheles* species while DNA barcoding was able to recover two sibling species within the morphologically identified *An. subpictus* specimens, bringing the total species count to 16.

According to the morphological identification results, 65.22% of morphologically identified *Anopheles* species present in Sri Lanka (15 species out of 23) were found in these eight study sites. Although Sri Lanka is a small island (65,525 sq km), a sharp increase in elevation can be seen from sea level to the centre of the country (highest elevation 2524 m). Elevation has been identified as one of the major factors that influences the vegetation pattern of a given region as it affects climatic factors. Western, southern and central regions situated in the wet and intermediate zones experience heavy showers for a long period of the year with a short dry period. Northern and eastern areas (dry zone) receive relatively low rains and undergo a prolonged dry season. Climate affects the prevalence and abundance of mosquito populations in different geographical areas of the country.

Prevalence of Sri Lankan mosquito species, including anophelines, has shown a significant seasonal and site-related density variation and weather has been identified as the most responsible macro-ecological factor [37–39]. Further, mosquito abundance declines during the dry season compared to the monsoon season with high rainfall [39]. During the present study, *An. peditaeniatus* and

An. vagus, two potential vectors of malaria were common to almost all the study sites. These two are paddy field breeders and paddy fields were common in all the study sites [39, 40]. Absence of *An. peditaeniatus* from Adikarigama site and *An. vagus* from Kattai-Adampan site may be due to seasonal variations at the time of mosquito collection.

Alternatively, *An. aconitus*, *An. maculatus*, *An. pseudojamesi*, and *An. tessellatus* were the least common mosquitoes, and each was restricted to a single locality. *An. maculatus* is considered as a stream breeder adapted to survive mainly at high altitudes near hilly or mountainous areas [41]. A previous mosquito survey conducted in Kandy district in the up-country wet zone of Sri Lanka reported *An. maculatus* as the dominant species [42]. *An. aconitus* is also considered a species that breeds at high altitudes [43]. In the present study, both these species were found at study sites with highest elevations: *An. maculatus* from Adikarigama (614.48 m) and *An. aconitus* from Haldummulla (1010 m). *An. pseudojamesi* and *An. tessellatus* are rare and least abundant anophelines in Sri Lanka [44]. It has been reported that *An. tessellatus* commonly breeds in locations similar to Wariyapola study site, which had ample sunlight with less vegetation [45]. The presence of the major and potential vectors of malaria throughout the country regardless of environmental conditions shows the potential threat of malaria re-emerging in the country where it has at present been eliminated.

As mosquitoes are vectors of many human diseases, accurate identification is essential in implementing vector control programmes [6, 46]. During the current study mitochondrial gene *COI* and nuclear ribosomal region ITS2 were used to confirm the identity of morphologically identified species. Presence of only a single ITS2 haplotype for each species reflects the importance of use of ITS2 in barcoding studies as the ITS2 sequences are species specific [47, 48].

As was previously mentioned, all the species were confirmed by comparing the sequences with the sequences already available in the GenBank and BOLD system and all the sequences had a maximum identity value between 99 and 100%. NJ trees using K2P substitution model are the most widely constructed phylogenetic trees used in many barcoding studies, especially on mosquitoes, to obtain more accurate recognition of the targeted group

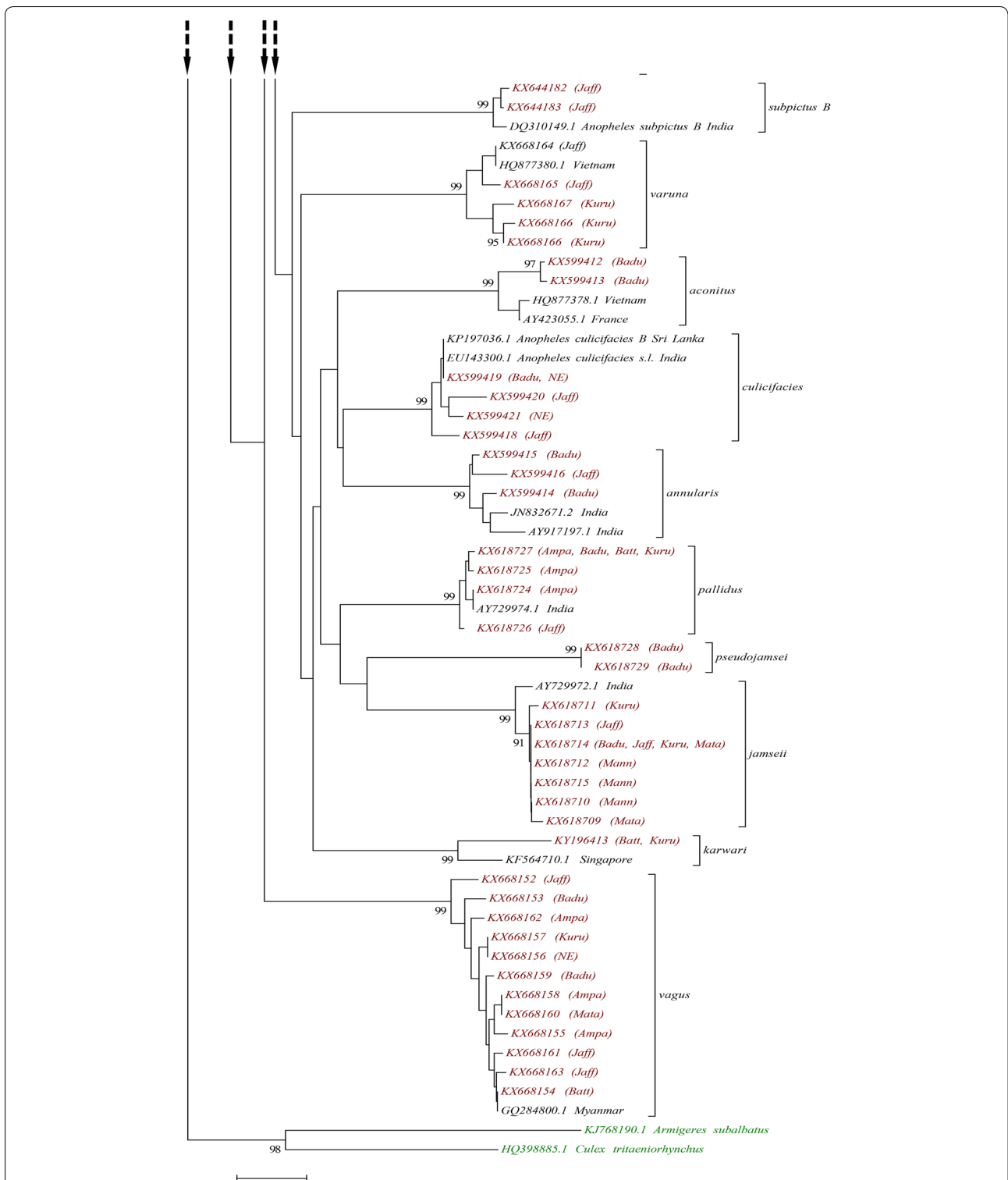
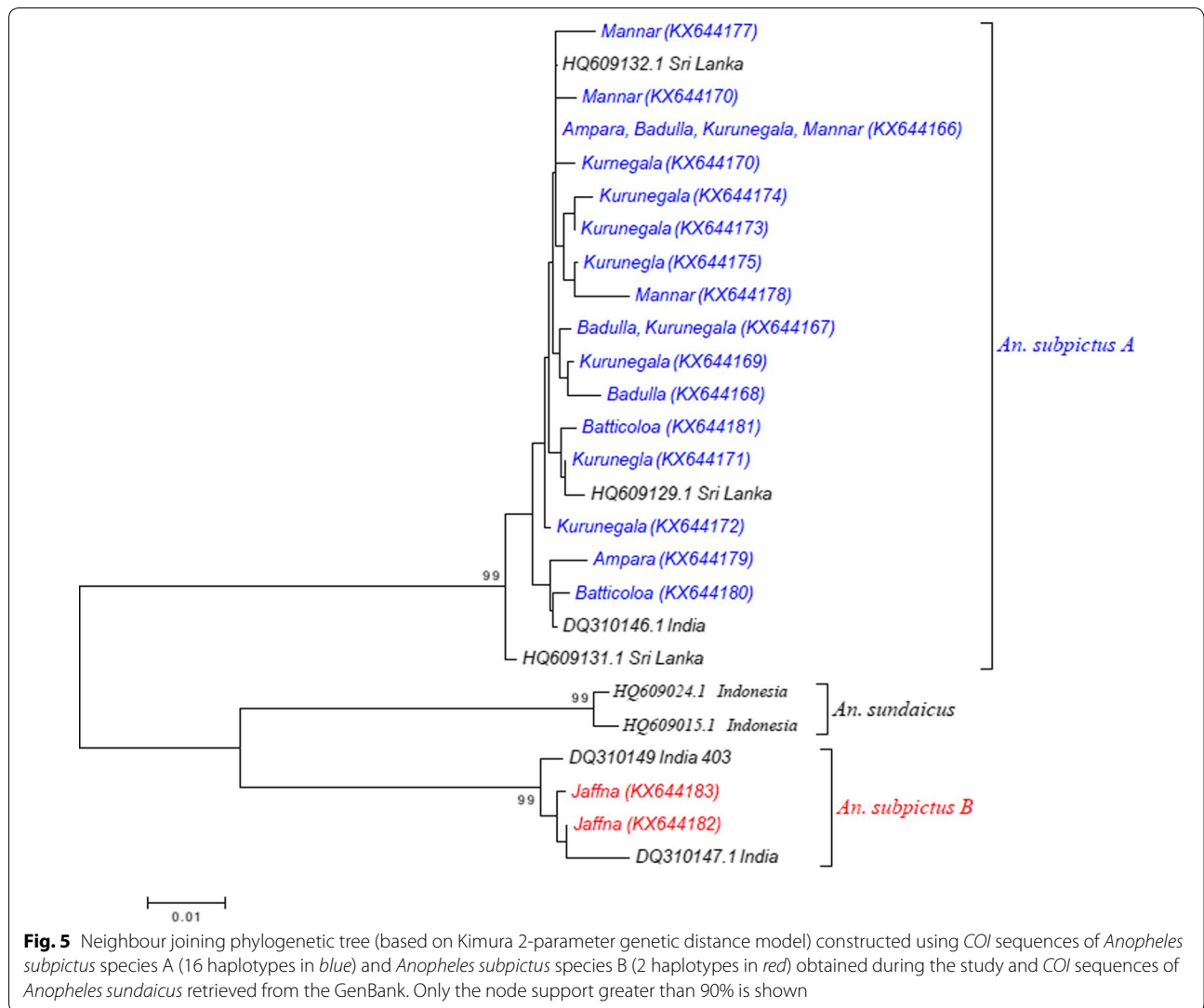


Fig. 4 Neighbour joining phylogenetic tree (based on Kimura 2-parameter genetic distance model) of *COI* sequences of all 79 haplotypes of anopheline species collected from Sri Lanka during the study (red label) and the sequences retrieved from the GenBank (black label). *Armigeres subalbatus* and *Culex tritaeniorhynchus* were used as the out-groups. Only node support greater than 90% is shown. The study site/s from which each haplotype was collected are given in parentheses (Ampa Ampara, Badu Badulla, Batt Batticoloa, Jaff Jaffna, Kuru Kurunegala, Mann Mannar, Mata Matala, NE Nuwara-Eliya)



[3, 6, 9, 49]. Accordingly, the same methodology was applied during this study to increase the reliability of the results. Based on *COI* sequence similarity, individuals that were morphologically grouped into species, except *An. subpictus*, were grouped into 14 distinct clades, with strong bootstrap support in the NJ phylogenetic tree confirming the results of morphological identification.

Based on genetic distance estimates for *COI* sequences, 2–3% sequence divergence or pairwise distance between species (inter-specific divergence) has been considered the threshold to differentiate two species [50, 51]. This has been used as the cutoff in many mosquito studies to distinguish closely related mosquito species [3, 6, 9]. The inter-specific K2P distances of the *COI* sequences of all the species exceeds 3% (0.03), verifying that they precisely belong to different species. This indicates the

significance of the *COI* region as a marker in identifying mosquitoes at species level.

According to phylogenetic tree analysis, all the morphologically identified *An. subpictus* samples comprise two genetic entities as they separate into two different clades supported by 99% bootstrap value (a clade containing samples from Jaffna and a clade including samples from all the other localities). Further, the K2P distances calculated between these two genetically different *An. subpictus* samples were >3%. This verifies that the *An. subpictus* samples, which were morphologically identified as a single species, belong to two different species. Further, the present analysis revealed that all the samples from Jaffna are *An. subpictus* species B while all the other samples are *An. subpictus* species A. *An. subpictus* species B has been reported as the most dominant *Anopheles*

Table 2 Inter-specific (below the diagonal) and the mean intra-specific distances (along the diagonal) of COI sequences

	<i>aco</i>	<i>ann</i>	<i>bar</i>	<i>cul</i>	<i>jam</i>	<i>kar</i>	<i>mac</i>	<i>nig</i>	<i>pal</i>	<i>ped</i>	<i>psej</i>	<i>subA</i>	<i>subB</i>	<i>tes</i>	<i>vag</i>	<i>var</i>
<i>aco</i>	0.003															
<i>ann</i>	0.114	0.010														
<i>bar</i>	0.143	0.134	0.000													
<i>cul</i>	0.088	0.077	0.120	0.012												
<i>jam</i>	0.122	0.123	0.172	0.116	0.001											
<i>kar</i>	0.142	0.114	0.126	0.115	0.137	0.000										
<i>mac</i>	0.134	0.131	0.127	0.114	0.109	0.126	0.012									
<i>nig</i>	0.123	0.140	0.122	0.114	0.140	0.142	0.115	0.009								
<i>pal</i>	0.092	0.080	0.125	0.089	0.106	0.097	0.108	0.115	0.002							
<i>ped</i>	0.127	0.117	0.119	0.111	0.126	0.128	0.118	0.072	0.102	0.007						
<i>psej</i>	0.147	0.101	0.171	0.120	0.107	0.143	0.128	0.140	0.088	0.143	0.000					
<i>subA</i>	0.139	0.124	0.133	0.112	0.129	0.144	0.120	0.135	0.119	0.097	0.151	0.002				
<i>subB</i>	0.130	0.118	0.133	0.112	0.137	0.135	0.139	0.139	0.132	0.130	0.141	0.128	0.006			
<i>tes</i>	0.131	0.134	0.093	0.101	0.153	0.135	0.122	0.114	0.132	0.098	0.159	0.116	0.145	0.006		
<i>vag</i>	0.155	0.135	0.134	0.149	0.139	0.126	0.132	0.126	0.138	0.111	0.155	0.134	0.130	0.123	0.009	
<i>var</i>	0.113	0.111	0.135	0.090	0.116	0.131	0.148	0.141	0.107	0.111	0.143	0.125	0.117	0.138	0.125	0.014

Distances were calculated using Kimura 2-parameter distance algorithm

aco An. aconitus, *ann* An. annularis, *bar* An. barbirostris, *cul* An. culicifacies, *jam* An. jamesii, *kar* An. karwari, *mac* An. maculatus, *nig* An. nigerrimus, *pal* An. pallidus, *ped* An. peditaeniatus, *psej* An. pseudojamesi, *subA* An. subpictus species A, *subB* An. subpictus species B, *tes* An. tessellatus, *vag* An. vagus, *var* An. varuna

species in both coastal (brackish water) and inland (fresh and nearly saline water) of Jaffna district [23] and it is more prevalent in coastal areas [18, 52, 53]. According to the results of this study, *An. subpictus* species B was found only from Jaffna district. This might be due to the high abundance of this species in the district and presence of water bodies with brackish and saline water. *An. subpictus* from other areas were *An. subpictus* species A. This species is known to breed mainly in inland freshwater bodies [52, 53]. Hence, it was found from study sites within the country and also from Batticaloa and Kalmunai with coastal areas. Higher prevalence of *An. subpictus* species A than species B in these areas might be the reason behind this observation. The DNA barcoding study by Kumar et al. [9], reported a 11.3% difference between these two sibling species in India and it was slightly higher (12.8%) in the present study. A molecular (using *COI* and *ITS2*) and morphology (identified based on the number of egg ridges)-based study showed that, although there are four sibling species of *An. subpictus* (A, B, C, D) according to egg morphology, the Sri Lankan subpictus complex actually comprise two genetically distinct species: *An. subpictus* species A and species B [18]. The phylogenetic tree analysis of this study showed that *An. subpictus* species B is more closely related to *An. sundaiensis* than *An. subpictus* species A, as observed in previous studies [18]. This indicates the considerable genetic distance between these two species and supports the above statement. More importantly the current study reveals

the importance of DNA barcoding in identifying sibling species/closely related species that may have different vectorial capacities.

Anopheles culicifacies species E is the primary vector of malaria in Sri Lanka and *An. culicifacies* species B is a non- or poor vector [18]. All the four haplotypes of *An. culicifacies* discovered during the present study belong to a single species. Although several studies have attempted to use DNA barcodes *COI*, *COII*, *ITS2* and *D3* in characterizing *An. culicifacies* species complex in Sri Lanka, none could accurately identify sibling species of this vector [27]. This may be due to the unavailability of all the sibling species in their collections or due to the inefficiency of the molecular-based approaches in recognizing sibling species of *An. culicifacies* species complex.

A study limited only to Jaffna district revealed the presence of *An. annularis* species A in Sri Lanka [53]. The present study shows no inter-specific variation among *An. annularis* collected from two different climatic zones: dry zone (Jaffna) and wet zone (Badulla). This confirms the presence of *An. annularis* species A in Sri Lanka. A phylogenetic analysis of medically important mosquitoes in India has observed the presence of the same substitutions and indels in the *ITS2* sequences of Indian *An. annularis* species A and *An. annularis* specimens from Sri Lanka, further confirming the sibling status of *An. annularis* in Sri Lanka [4]. Moreover, *An. annularis* species A is a malaria vector species in India, making it a very important vector species that needs attention [53].

The intra-specific divergence obtained during the study varied between 0 and 1.4%, showing that the *COI* divergence between the haplotypes of a single species is lower than the threshold value. DNA barcoding studies conducted in Pakistan and China reported 0–2.4 and 0–1.67% divergence ranges, respectively, similar to observations of the current study [5, 7, 54]. An intra-specific divergence of 1.84% has reported among the haplotypes of *An. pallidus* in India [9]. The highest intra-specific divergence was observed for *An. varuna* (1.4%) which was based on a comparison of a single sequence obtained from each of two distant localities: Wariyapola in Kurunegala and Tirunelveli in Jaffna district. The two study sites are located in two climatic zones and the Wariyapola site is at a higher elevation (low-country intermediate zone) than the Tirunelveli site (dry zone) (≈ 80 m difference). A study in India has observed very high intra-specific divergence of 4.9% among *Anopheles stephensi* samples and variation in the climate of the areas was found as the main reason for this high intra-specific divergence [4]. Further, Murugan et al. [4] suggest increasing the sample size and expanding the geographical area of collection to clarify this answer. *Anopheles varuna* was collected only from two study sites located at different elevations and climatic zones and sequences obtained from a relatively small sample would have been the reason behind this high intra-specific divergence, which needs further investigation. The K2P inter-specific diversity ranged between 7.2 and 17.2% in the current study and is comparable to other mosquito barcoding studies conducted in Pakistan (2.3–17.8%) [5], China (2.3–21.8%) [7], India (5.87–25.65%) [9], and Japan (4.1–12.5%) [54]. Although the effect is not known, the prominent amino acid mutation I95V, which was present in six species, has been reported from *Anopheles gambiae* as well [13].

According to morphology-based taxonomy, anophelines can be categorized into five groups: Pyrethrophorus (*An. subpictus*, *An. vagus*), Myzorhynchus (*An. barbrostris*, *An. nigerrimus*, *An. peditaeniatus*), Myzomyia (*An. aconitus*, *An. culicifacies*, *An. varuna*), Neocellia (*An. annularis*, *An. jamesii*, *An. karwari*, *An. maculatus*, *An. pallidus*, *An. pseudojamesi*), and Neomyzomyia (*An. tessellatus*) [30]. The NJ tree for ITS2 agrees with this grouping system as it forms four separate, strongly supported clades (96–100% bootstrap value), each representing a morphology-based group. Although the *COI* phylogenetic tree is different from the ITS2 tree and does not totally agree with the morphology-based grouping system, true phylogenetic relationships need further investigation.

Only the undamaged specimens were used for morphological identifications and these were subsequently

subjected to DNA barcoding. It was evident that both systems can be successfully used in species identification although subspecies-level identification could only be achieved with barcoding. Therefore, molecular characterization is a good method for both species and subspecies-level identification even with damaged specimens.

This is the first record of using both traditional and DNA barcoding to characterize several *Anopheles* species in Sri Lanka and it demonstrates the usefulness of DNA barcoding in confirming the identity of sibling species and closely related species. Further, as ITS2 confirmed *COI* findings, these two markers could be successfully used in characterizing Sri Lankan anophelines.

Conclusions

Genetic markers *COI* and ITS2 can be successfully used in characterizing anophelines of Sri Lanka, especially in recognizing the members of the *An. subpictus* species complex. Molecular characterization data can be reliably used in situations where morphological identification fails due to damaged morphological characters. The *COI* and ITS2 sequences generated could be used as reference sequences in future mosquito identification studies. Since DNA barcoding needs advanced laboratory facilities, the use of this method in routine malaria control programmes may be restricted unless the morphological characters are uncertain or identification to subspecies level is required. Presence of the major and potential vectors of malaria in different geographical locations perhaps indicates the potential for re-emergence of malaria in the country where malaria has recently been eliminated.

Additional files

Additional file 1. Anopheline species present in each study site with habitat type descriptions.

Additional file 2. GenBank accession numbers of *COI* and ITS2 sequences obtained during the current study and the publicly available sequence that showed highest similarity (>96% similarity) to these sequences. The fragment length of the ITS2 sequences of each species generated during the present study is also given.

Abbreviations

COI: cytochrome c oxidase subunit I; ITS2: gene and internal transcribed spacer 2 region; BOLD: barcode of life database; K2P: Kumura-2 parameter; NJ: neighbor joining.

Authors' contributions

SHPPK and SNS conceptualized and designed the work. TCW and MDBP did mosquito collections and morphological identifications. TCW, LSR and CSW conducted molecular laboratory work. TCW, CW, SNS, and SHPPK analysed the data. All contributed to the discussion and writing the manuscript. All authors read and approved the final manuscript.

Author details

¹ Department of Zoology, Faculty of Science, University of Peradeniya, Peradeniya, Sri Lanka. ² Department of Zoology, Faculty of Science, University

of Jaffna, Jaffna, Sri Lanka. ³ Liverpool School of Tropical Medicine, Liverpool, UK. ⁴ Regional Malaria Office, Kurunegala, Sri Lanka. ⁵ School of Earth and Environment, Faculty of Science and Engineering, University of Manchester, Manchester, UK. ⁶ National Institute of Fundamental Studies, Hantana, Kandy, Sri Lanka.

Acknowledgements

Field assistance by Mr. M. A. C. M. Mansoor (Anti-Malaria Campaign, Kurunegala, Sri Lanka) and the laboratory technical assistance by Mrs Helen Irving (Liverpool School of Tropical Medicine) is gratefully acknowledged.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets generated during the current study are available in the GenBank repository: <http://www.ncbi.nlm.nih.gov/genbank/>. All data generated during this study are included in this article (Additional file 2).

Funding

Commonwealth Association for the Split-Site Scholarship (2014) granted to TCW (Reference No. LKCN-2014-263) to conduct molecular laboratory work at Liverpool School of Tropical Medicine, UK for a one-year period. International Research Centre (InRC), University of Peradeniya, Sri Lanka (Grant No. InRC/RG/13/21 for funding field and laboratory work in Sri Lanka). Royal Society, UK Travel Grant to SHPPK to visit University of Manchester (Commonwealth Science Conference Follow-on Grants 2015).

Publisher's Note

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

Received: 13 February 2017 Accepted: 25 May 2017

Published online: 02 June 2017

References

- Kress WJ, Erickson DL. DNA barcodes: methods and protocols. *Methods Mol Biol.* 2012;858:3–8.
- Ondrejicka DA, Locke SA, Morey K, Borisenko AV, Hanner RH. Status and prospects of DNA barcoding in medically important parasites and vectors. *Trends Parasitol.* 2014;30:582–91.
- Guany F, Alten B, Simsek F, Aldemir A, Linton Y. Barcoding Turkish *Culex* mosquitoes to facilitate arbovirus vector incrimination studies reveals hidden diversity and new potential vectors. *Acta Trop.* 2015;143:112–20.
- 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.* 2015;115:107–21.
- Ashfaq M, Hebert PDN, Mirza JH, Khan AM, Zafar Y, Mirza MS. Analyzing mosquito (Diptera:Culicidae) diversity in Pakistan by DNA barcoding. *PLoS ONE.* 2014;9:e97268.
- Chan A, Chiang L, Hapuarachchi HC, Tan C, Pang S, Lee R, et al. DNA barcoding: complementing morphological identification of mosquito species in Singapore. *Parasit Vectors.* 2014;7:569.
- Wang G, Li C, Guo X, Xing D, Dong Y, Wang Z, et al. Identifying the main mosquito species in China based on DNA barcoding. *PLoS ONE.* 2012;7:e47051.
- Cywinska A, Hunter FF, Hebert PDN. Identifying Canadian mosquito species through DNA barcodes. *Med Vet Entomol.* 2006;20:413–24.
- Kumar NP, Rajavel AR, Natarajan R, Jambuligam P. DNA barcodes can distinguish species of Indian mosquitoes (Diptera: Culicidae). *J Med Entomol.* 2007;44:1–7.
- Marrelli MT, Sallum MAM, Marinotti O. The second internal transcribed spacer of nuclear ribosomal DNA as a tool for Latin American anopheline taxonomy—a critical review. *Mem Inst Oswaldo Cruz.* 2006;101:817–32.
- Toma T, Miyagi I, Crabtree MB, Miller BR. Identification of *Culex vishnui* subgroup (Diptera: Culicidae) mosquitoes from the Ryukyu Archipelago, Japan: development of a species-diagnostic polymerase chain reaction assay based on sequence variation in ribosomal DNA spacers. *J Med Entomol.* 2000;37:554–8.
- Beebe NW, Whelan PI, Van den Hurk AF, Ritchie SA, Corcoran S, Cooper RD. A polymerase chain reaction based diagnostic to identify larvae and eggs of container mosquito species from the Australian region. *J Med Entomol.* 2007;44:376–80.
- Garros C, Ralph EH, Manguin S. Systematics and biogeographical implications of the phylogenetic relationships between members of the *Funestus* and *Minimus* groups of *Anopheles* (Diptera: Culicidae). *J Med Entomol.* 2005;42:7–18.
- Sharpe RG, Harbach RE, Butlin RK. Molecular variation and phylogeny of members of the *Minimus* group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). *Syst Entomol.* 2000;25:263–72.
- Gunathilaka N, Fernando T, Hapugoda M, Abeyewickreme W, Wickremasinghe R. Revised morphological identification key to the larval anopheline (Diptera: Culicidae) of Sri Lanka. *Asian Pac J Trop Biomed.* 2014;4:222–7.
- Amerasinghe FP. Illustrated keys to the genera of mosquitoes (Diptera: Culicidae) in Sri Lanka. *J Nat Sci Council Sri Lanka.* 1995;23:183–221.
- Amerasinghe FP, Munasingha NB, Kulasekera VL. Some new records of mosquitoes occurring in Sri Lanka. *Mosq Syst.* 1987;19:162–6.
- Surendran NS, 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.
- Surendran SN, Abhayawardana TA, De Silva BG, Ramasamy R, Ramasamy MS. *Anopheles culicifacies* Y-chromosome dimorphism indicates sibling species (B and E) with different malaria vector potential in Sri Lanka. *Med Vet Entomol.* 2000;14:437–40.
- Abhayawardana TA, Wijesuriya SR, Dilrukshi RK. *Anopheles subpictus* complex: distribution of sibling species in Sri Lanka. *Indian J Malariol.* 1996;33:53–60.
- Surendran SN, Jude PJ, Weeraratne TC, Karunaratne SHPP, Ramasamy R. Variations in susceptibility to common insecticides and resistance mechanisms among morphologically identified sibling species of the malaria vector *Anopheles subpictus* in Sri Lanka. *Parasit Vectors.* 2012;5:34.
- Thavara C, Rajendram GF. A survey of anopheline mosquitoes in Jaffna peninsula. *Proc Jaffna Sci Assoc.* 1995;4:20.
- Kannathasan S, Antonyrajan A, Srikrishnaraj KA, Karunaratne SHPP, Karunaweera ND, Surendran SN. Studies on prevalence of *Anopheline* species and community perception of malaria in Jaffna district, Sri Lanka. *J Vector Borne Dis.* 2008;45:231–9.
- Amerasinghe PH, Amerasinghe FP, Konradsen F, Fonseka KT, Wirtz RA. Malaria vectors in a traditional dry zone village in Sri Lanka. *Am J Trop Med Hyg.* 1999;60:421–9.
- Herath PRJ, Abeywardena T, Padmalal UKGK. A study of the role of different indigenous anopheline species in the transmission of human malaria in Sri Lanka. *Proc Sri Lanka Assoc Adv Sci.* 1983;39:6.
- Surendran SN, Hawkes NJ, Hemingway SAJ, Ramasamy R. Molecular studies of *Anopheles culicifacies* (Diptera: Culicidae) in Sri Lanka: Sibling species B and E show sequence identity at multiple loci. *Eur J Entomol.* 2006;103:233–7.
- Surendran SN, Truelove N, Sarma DK, Ramasamy R, Gajapathy K, Peiris LBS, et al. Karyotypic assignment of Sri Lankan *Anopheles culicifacies* species B and E does not correlate with *cytochrome oxidase subunit I* and microsatellite genotypes. *Parasit Vectors.* 2015;8:327.
- Gajapathy K, Jude PJ, Goodacre SL, Peiris LBS, Ramasamy R, Surendran SN. Molecular characterization of the malaria vector *Anopheles barbirostris* vander Wulp in Sri Lanka. *Parasit Vectors.* 2014;7:348.
- Surendran NS, Singh OP, Jude PJ, Ramasamy R. Genetic evidence for malaria vectors of the *Anopheles sudaicus* complex in Sri Lanka with morphological characteristics attributed to *Anopheles subpictus* species B. *Malar J.* 2010;9:343.
- Amerasinghe FP. A Guide to the identification of the *Anopheline* mosquitoes (Diptera: Culicidae) of Sri Lanka. I. Adult females. *Cey J Sci (Biological Science).* 1990;21:1–16.
- Simon C, Frati F, Beckenbach A, Crepsi B, Liu H, Flook K. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am.* 1994;87:651–701.

32. Paskewitz SM, Collins FH. Use of the polymerase chain reaction to identify mosquito species of the *Anopheles gambiae* complex. *Med Vet Entomol*. 1990;4:367–73.
33. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22:4673–80.
34. NCBI National Center for Biotechnology Information. <http://www.blast.ncbi.nlm.nih.gov>.
35. BOLDSYSTEMS. <http://www.boldsystems.org>.
36. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol*. 1980;16:111–20.
37. Amerasinghe FP, Ariyasena TG. Survey of adult mosquitoes (Diptera: Culicidae) during irrigation development in the Mahaweli project, Sri Lanka. *J Med Entomol*. 1991;28:387–93.
38. Amerasinghe FP, Ariyasena TG. Larval survey of surface water-breeding mosquitoes during irrigation development in the Mahaweli Project, Sri Lanka. *J Med Entomol*. 1990;27:789–820.
39. Amerasinghe FP, Munasingha NB. A predevelopment mosquito survey in the Mahaweli development project area, Sri Lanka: adults. *J Med Entomol*. 1988;25:276–85.
40. Van der Hoke W, Amerasinghe FP, Konradsen F, Amerasinghe PH. Characteristics of malaria vector breeding habitats in Sri Lanka: relevance for environmental management. *Southeast Asian J Trop Med Public Health*. 1998;29:168–72.
41. Sinka ME, Bangs MJ, Manguin S, Chareonviriyaphap T, Patil AP, Temperley WH, et al. The dominant *Anopheles* vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic précis. *Parasit Vectors*. 2011;4:89.
42. Kusumawathie PHD, Fernando WP. *Anopheles maculatus* (Theobald) and *Anopheles elegans* (James) breeding in water storage containers in Kandy, Sri Lanka. *Ceylan J Med Sci*. 2002;45:71–6.
43. Stoops CA, Rusmiarto S, Susapto D, Munif A, Andris H, Barbara KA, et al. Bionomics of *Anopheles* spp. (Diptera: Culicidae) in a malaria endemic region of Sukabumi, West Java, Indonesia. *J Vector Ecol*. 2009;34:200–7.
44. Gunathilaka N, Fernando T, Hapugoda M, Wickremasinghe R, Wijeyerathne P, Abeywickreme W. *Anopheles culicifacies* breeding in polluted water bodies in Trincomalee district of Sri Lanka. *Malar J*. 2013;12:285.
45. Rahman WA, Adanan CR, Abu Hassan A. Species composition of adult *Anopheles* populations and their breeding habitats in Hulu Peak district, peninsular Malasia. *Southeast Asian J Trop Med Public Health*. 2002;33:547–50.
46. Higa Y, Toma T, Tsuda Y, Miyagi I. A multiplex PCR-based molecular identification of five morphologically related, medically important subgenus *Stegomyia* mosquitoes from the genus *Aedes* (Diptera: Culicidae) found in the Ryukyu Archipelago, Japan. *Jpn J Infect Dis*. 2010;63:312–6.
47. Prakash A, Walton C, Bhattacharyya DR, Mohapatra PK, Mahanta J. Characterization of ITS2 rDNA of *Anopheles philippinensis* and *Anopheles nivipes* (Diptera: Culicidae) from north-east India. *Southeast Asian J Trop Med Public Health*. 2006;37:1134–8.
48. Walton C, Handley M, Kuvangkadilok C, Collins FH, Harbach RE, Baimai V, et al. Identification of five species of the *Anopheles dirus* complex from Thailand using allele-specific polymerase chain reaction. *Med Vet Entomol*. 1999;13:24–32.
49. Rozo-Lopez P, Mengual X. Mosquito species (Diptera, Culicidae) in three ecosystems from the Colombian Andes: identification through DNA barcoding and adult morphology. *ZooKeys*. 2015;513:39–64.
50. Hebert PDN, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes. *Proc Biol Sci*. 2003;270:313–21.
51. Hebert PDN, Ratnasingham S, de Waard JR. Barcoding animal life: cytochrome *c oxidase subunit 1* divergences among closely related species. *Proc Biol Sci*. 2003;270:96–9.
52. Jude PJ, Ramasamy R, Surendran SN. Bionomic aspects of the *Anopheles subpictus* species complex in Sri Lanka. *J Insect Sci*. 2014;14:97.
53. Surendran SN, Gajapathy K, Kumaran V, Tharmatha T, Jude PJ, Ramasamy R. Molecular evidence for the presence of malaria vector species A in the *Anopheles annularis* complex in Sri Lanka. *Parasit Vectors*. 2011;4:239.
54. Taira K, Toma T, Tamashiro M, Miyagi I. DNA barcoding for identification of mosquitoes (Diptera: Culicidae) from Ryukyu Archipelago, Japan. *Med Entomol Zool*. 2012;2012(63):289–306.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

