Variation in Number and Formation of Repeat Sequences in the rDNA ITS2 Region of Five Sibling Species in the Anopheles barbirostris Complex in Thailand

Author: Otsuka, Yasushi

Source: Journal of Insect Science, 11(137) : 1-11

Published By: Entomological Society of America

URL: https://doi.org/10.1673/031.011.13701

Variation in number and formation of repeat sequences in the rDNA ITS2 region of five sibling species in the *Anopheles barbirostris* complex in Thailand

Yasushi Otsuka

Department of Infectious Disease Control, Faculty of Medicine, Oita University, Hasama, Yufu, Oita 879-5593, Japan

Abstract

Repeat sequences of approximately 100 base pairs in length were found in the rDNA ITS2 region of *Anopheles barbirostris* van der Wulp (Diptera: Culicidae) species A1, A2, A3, A4, and *An. campestris*-like in the *An. barbirostris* complex. Variation in the number of repeats was observed among the five sibling species. Specifically, 10 repeats were observed in A1, eight in A2, A4, and *campestris*-like, and three in A3. Based on similarities in the sequences of the repeats, related repeats were classified into nine groups. Although A2, A4, and the *campestris*-like species had the same number of repeats, the ITS2 region of the three species contained different groups of repeats. Excluding the repeat sequences facilitated good alignment of the ITS2 region in the five sibling species. Phylogenetic analyses of the 95 isolines were compared with results obtained from mitochondrial genes (COI and COII). The results revealed marked differences among the five sibling species, particularly regarding the ITS2 region of A3, which was more distinct from the other four species than COI and COII. Repeat sequences in the ITS2 region of other *Anopheles* species retrieved from GenBank also were analyzed. New repeat sequences were found in *An. beklemishevi* Stegnii and Kabanova, *An. crucians* Wiedemann and *An. funestus* Giles, suggesting that the occurrence of repeat sequences in the ITS2 region are not rare in anopheline mosquitoes.

Keywords: concerted evolution, repetitive sequences, sibling species members
Correspondence: yotsuka@oita-u.ac.jp
Editor: Marcelo Jacobs-Lorena was Editor of this paper.
Received: 25 November 2010, **Accepted:** 9 June 2011

ISSN: 1536-2442 | Vol. 11, Number 137

Cite this paper as:
Otsuka Y. 2011. Variation in number and formation of repeat sequences in the rDNA ITS2 region of five sibling species in the *Anopheles barbirostris* complex in Thailand. *Journal of Insect Science* 11:137 available online: insectscience.org/11.137

Introduction

Anopheles (*Anopheles*) *barbirostris* van der Wulp and *An. campestris* Reid (Diptera: Culicidae) both belong to the Barbirostris subgroup of the Myzorhynchus series, and are natural vectors of both malaria, due to *Plasmodium vivax* Grassi and Feletti (Haemosporida: Plasmodidae), and filariasis, caused by periodic *Brugia malayi* Brug (Spirurida: Onchocercidae) in Malaysia and Indonesia (Reid 1968; Atomosoedjono et al. 1976; Kirnowardoyo 1985). In addition, these mosquitoes are also suspected vectors of malaria and/or filariasis in Thailand (Iyengar 1953; Griffith 1955), where they may be natural vectors of *P. vivax* in the Aranyaprathet district of Sa Kaeo province (Limrat et al. 2001; Apiwathnasorn et al. 2002). Sattabongkot et al. (2004) also considered that these vectors play an important role in increasing *P. vivax* infections in Thailand. In addition, the overlapped adult morphology between *An. barbirostris* and *An. campestris* has led to problems in species identification, particularly when using damaged scales of wild-caught females in the field (Harrison and Scanlon 1975). Recently, at least five species members of the *An. barbirostris* complex, namely *An. barbirostris* species A1, A2, A3, A4, and *An. campestris*-like, have been discovered in Thailand using cytogenetic and molecular markers and crossing experiments (Saeung et al. 2007, 2008; Suwannamit et al. 2009; Thongsahuan et al. 2009).

The high interspecific divergence and low degree of intraspecific divergence associated with the internal transcribed spacer 2 (ITS2) region of ribosomal DNA (rDNA) has been applied widely to distinguish between closely related *Anopheles* species, e.g., Maculipennis group (Porter and Collins 1991; Cornel et al. 1996; Proft et al. 1999) and the *An. culicifacies* complex (Goswami et al. 2005). In previous molecular analyses of the five species in the *An. barbirostris* complex, ITS2 assays were capable of distinguishing between closely related species (Saeung et al. 2007, 2008; Suwannamit et al. 2009; Thongsahuan et al. 2009). However, unlike most *Anopheles* species, in which the ITS2 region ranges from 200 to 600 base pairs (bp) (Wilkerson et al. 2004), the ITS2 region of the five species of the *An. barbirostris* complex is not only longer than 900 bp, but it also varies between these species. The primary objective of this study was to characterize the extra length and size variation of ITS2 in the *An. barbirostris* complex resulting from repeat sequences. Existence of the repeat sequences made it difficult to align the ITS2 sequences of the five species. However, by excluding the repeat sequences, clear alignment of the ITS2 region could be obtained. Based on these sequence alignments, phylogenetic analyses of the five species were conducted and compared with phylogenies that were estimated using the mitochondrial cytochrome *c* oxidase subunits I and II (COI and COII). *Anopheles fluminensis* also included three repeats (125 bp each) in the ITS2 region (Brelsfoard et al. 2006), suggesting that other *Anopheles* species may also have repeat sequences in this region. Therefore, another objective of this study was to interpret the present findings within the context of repeat sequences in the ITS2 region of other anopheline mosquitoes deposited in GenBank.

Materials and Methods

The ITS2, COI, and COII sequences of 95 isolines of *An. barbirostris* species A1 (n = 39), A2 (n = 15), A3 (n = 3), A4 (n = 3), and

An. campestris-like (n = 35) were obtained from the GenBank database (Supplementary Table). The repeat sequences of the ITS2 region were aligned using the MARNA web server, based on the primary sequence and secondary structure as a default setting (Siebert and Backofen 2005; www.bioinf.uni-freiburg.de/Software/MARNA/index.html). Phylogenetic trees were estimated from the repeat sequence data using the neighbor joining (NJ) and bootstrapping algorithms implemented in the MEGA software package version 4 (Tamura et al. 2007). Genetic distances were estimated by the Jukes-Cantor model with pairwise deletion of gaps.

Phylogenetic trees of all 95 isolines of the *An. barbirostris* complex were generated based on ITS2, COI, and COII sequences using distance, maximum likelihood (ML), and maximum parsimony (MP) methods. To analyze the ITS2 region, the repeat sequences were removed and the remaining sequences were aligned by Clustal W (Thompson et al. 1994). Combined COI and COII sequences were aligned using the sequences of *An. gambiae* Giles (NC_002084) and *An. pullus* Yamada (AY444349 and AY444350) as outgroups. The phylogenetic trees of ITS2 were left unrooted, because an outgroup with an easily aligned ITS2 region was not available and gaps were excluded from the alignments. For the distance analyses, the NJ algorithm was used, as described above, except for gap treatment. For ML, the best-fit models for the ITS2 and COI+COII sequences were selected using the hierarchical likelihood ratio test implemented in Modeltest 3.06 (Posada and Crandall 1998). The models selected were HKY+Γ for ITS2 and GTR+I+Γ for COI+COII. The ML trees for the selected models were generated by PhyML 3.0 (Guindon and Gascuel 2003). For the MP analysis, the trees were generated using the default heuristic search option in PAUP* 4.0 b10, with ten random-addition sequence replicates (Swofford 1998). Bootstrap analyses for the 1000 replicate datasets were performed for each tree to assess the statistical support for the nodes. Dot plot analyses were conducted to scan the ITS2 sequences of *Anopheles* species in GenBank for repeat sequences using the BioEdit program (Hall 1999). However, since short repeats were difficult to identify on plots, only repeat sequences longer than 30 bp were detected.

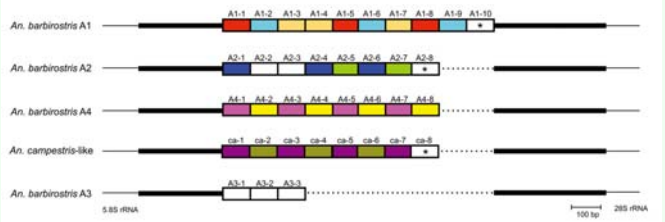

Figure 1. ITS2 characteristics of five sibling species in the *Anopheles barbirostris* complex. Each repeat unit is represented by a rectangle, and the designated name of each repeat is shown above the rectangle. Colors of the repeats correspond with the groups shown in Figure 3; white-colored repeats did not cluster with any of these groups. Despite belonging to different species, the repeats shown with an asterisk (A1-10, A2-8, and ca-8) clustered with high associated bootstrap values. The ITS2 region, except for the repeat sequences, is shown by the bold line. The 5.8 and 28S rRNA genes flanking the ITS2 region are shown by a thin line. Dots indicate gaps. High quality figures are available online.

Results

Repeat sequences in the ITS2 region of the five species in the *An. barbirostris* complex

Comparisons of ITS2 sequence alignments revealed that all five sibling species in the *An. barbirostris* complex contained sequences repeated tandemly, with each repeat sequence ~ 100 bp long. *An. barbirostris* species A1 had ten repeats; A2, A4, and *An. campestris*-like had eight; and A3 had three (Figure 1). The sequences flanking the repeats in the five species were conserved, suggesting that the repeats were located at the same position. Since substitutions and indels had occurred among the repeat sequences, the relationships among the repeats were examined. First, the ITS2 sequences representing each of the five species (A1, AB331555; A2, AB331551; A3,

```
       1........10........20........30........40........50........60........70........80........90........100.......110.......120.......
A1-1   GGGTGTGACTTCTGC-CCGTT--GTGCTGC--CC-GTAGCT--A-CGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTAT-GA-CGT-GGT-CC--AGTCG-GGACG-GGGTC-CAT-C-C-
A1-5   GGGTGTGTCAACTGC-CCGTT--GCGCTGC--CC-GTAGCT--A-CGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTAC-GA-CGT-GGT-TC--AGTCG-GGACG-TTGGC-CAC-C-C-
A1-8   GGGTGTGTCTACTGC-CCGTT--GCGCTGC--CC-GTAGCT--A-CGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTAC-GA-CGT-GGT-TC--AGTCG-GGACG-TGGGC-CAC-C-C-
A1-4   GGGTGT-TCAGCGGT-CTGCTGGGCGCTGC--CC-GTAGCT--A-CGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTAC-GA-CGT-GGT-TCCCAG-TA-GGACG-AGGTC-CAC-C-C-
A1-7   GGGTGTGTCAGCGGT-CTGCTGGGCGCTGC--CC-GTAGCT--A-CGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTAC-GA-CGT-GGT-TCCCAG-TA-GGACG-AGGTC-CAC-C-C-
A1-3   GGGTGTGTCAACGGA-CTGCTGGGCGCTGC--CC-GTAGCT--A-TGT--TTGGC-CAGTATTG-GCTATT--CCTCT---GGCTGC-GA-CGT-GGT-TCCCAG-TA-GGACG-GGGGC-CAT-C-C-
A4-1   GGGTGTG-CAACGGA-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGAACTTGGC-TTGTCATGAGCTGTT--CCGCACAATGCTACGGTGCGT-GGT-TCCCAG-TA-GGACG-GGGTTACAC-C-C-
A4-3   GGGTGTG-CAACGGC-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGAACTTGGC-TTGTCATGAGCTGTT--CCGCACAATGCTACGGTGCGT-GGT-TCCCAG-TA-GGACG-GGGTTACAC-C-C-
A4-5   GGGTGTG-CAACGGC-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGAACTTGGC-TTGTCATGAGCTGTT--CCGCACAATGCTACGGTGCGT-GGT-TCCCAG-TA-GGACG-GGGTTACAC-C-C-
A4-7   GGGTGTG-CAACGGC-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGAACTTGGC-TTGTCATGAGCTGTT--CCGCACAATGCTACGGTGCGT-GGT-TCCCAG-TA-GGACG-GGGTTACAC-C-C-
ca-1   GGGTGTG-CAACGGT-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGATCT-GGC-CATTCGTGAGCTGAT--CCGCACAATGCTACGGTGTGT-GGT-TCCCGG-TG-GGACG-GGGTTACAC-C-C-
ca-3   GGGTGTG-CAACGGT-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGATC-TGGC-CATTCGTGAGCAGAT--CCGCACAATGCTACGGTGTGT-GGT-TCCCGG-TG-GGACG-GGGTTACAC-C-C-
ca-5   GGGTGTG-CAACGGT-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGATCT-GGC-CATTCGTGAGCTGAT--CCGCACAATGCTACGGTGTGT-GGT-TCCCTG-TG-GGACG-GGGTTACAC-C-C-
ca-7   GGGTGTG-CAACGGA-CTGCTGGGCGCTGC--CCTGTAGCA--T-GGATCT-GGC-CATTCGTGAGCTGAT--CCGCACAATGCTACGGTGTGT-GGT-TCCCGG-TA-GGACG-GGG-TACAT-C-C-
A2-1   GGGTGTGTCAACGGA-CTGCTG--TGCTGC--CC-GTGGCG--A-AGG--TTGGC-CATACTTGAGCCTTT--CCG-A---TGCTGC-GA-CGT-GGT--CC-AG-TG-GGACG-GGGTC-CAC-C-C-
A2-4   GGGTGTGTCAACGGA-CTGCTG--TGCTGC--CC-GTGGCG--A-AGA--TTGGC-CATACTTGAGCCTTT--CCG-A---TGCTGC-GA-CGT-GGT--CC-AG-TG-GGACG-GGGTC-CAC-C-C-
A2-6   TGGTGTGTCAACGGA-CTGCTG--TGCTGC--CC-GTGGCG--A-AGA--TTGGC-CATACTTGAGCCTTT--CCG-A---TGCTGC-GA-CGT-GGT--CC-AG-TG-G-ACG-GGGTC-CAC-C-A-
A3-1   GAGTGTGTCAACGGA-CTGCTCGGTGCTGC--CC-GT---------------GGC---GAAA---GCC--------------G---C-GCG-GT-GGT--CC-AG-TG-GGACG-GAGTC-CAT-T-C-
A1-2   GGGTGTGTCAAAGGT-TTGCTG--TGCCGC--CC-GCAGCC--C-GGA--ATGGT-GCGTCA-GTGCCACT--TCGGA----GTTGC-GG-C-TCGGA-C-C-GG-TA-AGCTG-GAGGCACAT-C-C-
A1-6   GGGTGGGTCAAAGGT-TTGCTG--TGCCGC--CC-GCAGCC--C-GGA--ATGGT-GCGTCA-GTGCCACT--TCGGA----GTTGC-GG-C-TCGGA-C-C-GG-TA-AGCTG-GAGGCCCAT-C-T-
A1-9   GGGTGTGTCAATGGT-TTGCTG--TGCCGC--CC-GCAGCC--C-GGA--ATGGT-GCGTCA-GTGCCACT--TCGGA----GTTGC-GG-C-TCGGA-C-C-AG-TA-AGCTG-GGGGC-CAC-C-C-
A2-5   GGGTGTGTCAACGGT-TTGCTGTGTTCTGC--CC-GTGGTGGTA-GGA--TTGGC-CAGTCATGAGCCAGT--CCGAA----GCCGC-GA-CGT-GGG-CC--AG-TA-AGTCG-GGGGC-CAT-C-C-
A2-7   GGGTGTGTCAATGGT-TTGCTGTGTTCTGC--CC-GTGGTGTTA-AGA--TTGGC-CAGTCATGAGCCAGT--CCGAA----GCCGC-GA-CGT-GGG-CC--AG-CA-AGCTG-GGGGC-CAT-C-C-
A4-2   GGGTGTGTCTGCTGTCCCGTT--GTGCTGC--CC-GTAGAA--A-AGA--TTGGC-CATACATGAGCCATT-TCTGGA-----CTAC-GA-CGT-AGT-GC--AG-CG-GGTGG-TGGGC-CAT-C-C-
A4-4   GGGTGTGTCTGCTGTCCCGTT--GTGCTGC--CC-GTAGAA--A-AGA--TTGGC-CATACATGAGCCATT-TCTGGA-----CTAC-GA-CGT-AGT-GC--AG-CG-GGTGG-TGGGC-CAT-C-C-
A4-6   GGGTGTGTCTGCTGTCCCGTT--GTGCTGC--CC-GTAGAA--A-AGA--TTGGC-CATACATGAGCCATT-TCTGGA-----CTAC-GA-CGT-AGT-GC--AG-CG-GGTGG-TGGGC-CAT-C-C-
A4-8   GGGTGTGTCTGCTGTCCCGTT--GTGCTGC--CC-GTAGAA--A-AGA--TTGGC-CAGTCATGAGCCATT-TCTGGA-----CTAC-GA-CGT-AGT-GC--AG-CG-GGTGG-TGGGC-CAT-C-C-
ca-2   GGGTGGTTCTGCTGC-CCGCT--GTGCTGC--CC-GTAGCT--T-AGA--TTGGC-CAGTCATGAGCCAGT--CAGC----GGCTGC-GA-CGT-GGT-GC--AGTCG-GGACG-TGGTC-CAT-C-C-
ca-4   GGGTGGTTCTGCTGC-CCGCT--GTGCTGC--CC-GTAGCT--T-AGA--TTGGC-CAGTCATGAGCCAGT--CAGC----GGCTGC-GA-CGT-GGT-GC--AGTCG-GGACG-TGGTC-CAT-C-C-
ca-6   GGGTGGTTCTGCTGC-CCGCT--GTGCTGC--CC-GTAGCT--T-AGA--TTGGC-CAGTCATGAGCCAGT--CAGC----GGCTGC-GA-CGT-GGA-GC--AGTTG-GGACG-TGGAC-CAT-C-C-
A2-3   GGGTGCGTCAACGGT-CAGTTG--TGCTAG--CC-GTGGCA--A-CGT--GTGGC-CAGTCATGAGCCATG--CCG-AG---GCTGC-GT--GTTGG--G-C-AACAG-GACCG-GAGGC-CAT-C-C-
A3-3   TGGTGTGTCGTC----CAGTTC--AGTTGT-----GT-GC-----TGC-CCGTG----GGGA-TCGC-GT-----GGTCA--G--TC----AGT-GGCTG---GG-TG-GAA-----GGCGCATACTCC
A2-2   GGGTGTGTCGCCGT-C-TAGT--GTGCTTC--CC-ATGGAG--ATTGA--TTCGCGCCTTTG-GGTGGGA--TTAGC----TGCCGT-GG-TGC-GGT-TC--AC-AG-GGACG-GAGGC-CAC-C-C-
A3-2   GGGTGCGTTGTCGT-CCAGCT--TTGCTGTGGCC-ATGGAGTGGTAGATGCATG--CTTTCG-G-GT-GTGCTCTAAAC--TGCCGTGGTGC---AGT----AGG-CTCGGACG-GGGGC-CAC-C-C-
A1-10  GGGTGTGTTGTCAT-CGTGCT--GTGCTGC--CC-GTGGGGTA--AGC-CTCGGC--CTTCG-G-GCTGTGGTCTTGC---TGACGC-GG-TGT-AGT--C--GG-TACTGGTG-GGGAC-CAT-C-C-
ca-8   GGATACGTGTCCAT-CGTGCT--GTGCTGG--CC-GTGGGGCA--AGA-TTCGGT-CCATTG-G-GCTGTGGTCTTGC---TGACGC-GG-TGT-AGTGTC--AG-TGCTGTTG-GGGGC-TAT-C-C-
A2-8   GGGTGTGTTGTCAT-CGTGCC--GTGCTGC--CC-GCGGGGCGA--GC-TTCGGT-CCTTCG-GGGCCGTTGTCTTGC---TGACGC-GG-TGT-GGT--C--GG-TGCTGATGGTGGAC-CAC-C-T-
       †††††.†††..††._†††††._†.††..†__. _††††††__._†††__.††††_†......†.††††.†__††...___††††††_†._.†._.††_.†._††_††_†..††_..†††_†††_†_†_
```

Figure 2. Alignment of the 37 repeat sequences from the ITS2 region of each of the five sibling species in the *Anopheles barbirostris* complex using the MARNA web server. Dashes indicate indels. Corresponding † indicate positions at which nucleotides were conserved in more than 50% of the repeats. High quality figures are available online.

AB362232; A4, AB373939; *An. campestris*-like, AB331563) were selected according to frequency of occurrence in each species. In this way, the 37 repeat sequences extracted from the ITS2 sequences of the five species were named; A1-1 to A1-10 (10 repeats), A2-1 to A2-8 (8 repeats), A3-1 to A3-3 (3 repeats), A4-1 to A4-8 (8 repeats) in the *An. barbirostris* species, and ca-1 to ca-8 (8 repeats) in the *campestris*-like species (Figure 1). The repeats were composed of inverted sequences, suggesting that they either formed secondary structures by themselves or in conjunction with other repeats. Most of the repeats had conserved sequences at both ends; GGGTG at the 5.8S rRNA-end and CA(C/T)CC at the 28S rRNA-end.

Alignments based on primary sequence and secondary structure were performed (Figure 2), and an NJ tree was constructed based on the alignments (Figure 3). In order to examine the structural characteristics of the repeats in each species, sequences separated by genetic distances of less than 0.1 were classified as the same group. Nine groups were formed, and seven repeats remained ungrouped (Figure 3). The groups of repeats were correlated with species. The three ungrouped repeats, A1-10, A2-8, and ca-8, however, were clustered in the NJ tree with a high bootstrap value. These repeats were located closest to the 28S rRNA. Species A4 and *campestris*-like had eight similar repeat sequences consisting of tandem units that comprised one repeat from groups 3 and 8, and one from groups 4 and 9, respectively. The NJ tree demonstrated that group 3 and 8 were the most closely related to groups 4 and 9, respectively. However, the *campestris*-like species had ca-8, which was not found among any of the groups in the A4 and *campestris*-like species. A2 also included eight repeats, which had a different structure from those of the A4 and *campestris*-like species. A1 had 10 repeats belonging to groups 1, 2, and 6. A3 had three repeats, two of which were short in length; A3-1 and A3-3 were only 74 bp and 81 bp long, respectively, while the other 35 repeats were 95-112 bp in length. The three

repeats of A3 had diverged considerably from one another, and did not belong to any group. In the NJ tree, A3-1 was placed closely to group 5, which included A2-1. In addition, A3-2 and A3-3 were grouped closest to A2-2 and A2-3, respectively, indicating that the array of the three repeats, A3-1, A3-2, and A3-3 was distantly related to A2-1, A2-2, and A2-3.

Phylogenetic relationship between the five sibling species of the *An. barbirostris* complex

Phylogenetic trees of the 95 isolines were generated based on the alignment of ITS2 sequences (Figure 4A). Each species was clustered in all of the trees and very little intraspecific variation was observed. A3 was placed separately from the other four species. Phylogenetic trees also were constructed based on the combined sequences of the mitochondrial COI and COII genes (Figure 4B). With the exception of the A1 isoline (APbB27) in the ML analysis, all species were separated. However, the relationships between A2, A4, and *campestris*-like were not clearly estimated. Marked differences between the trees of the two regions were observed. First, A3 was placed more closely to the other four species in the trees of the mitochondrial genes than in those of ITS2. Second, A2 was related closely to A1 in ITS2, but clustered with A4 and *campestris*-like in the mitochondrial genes.

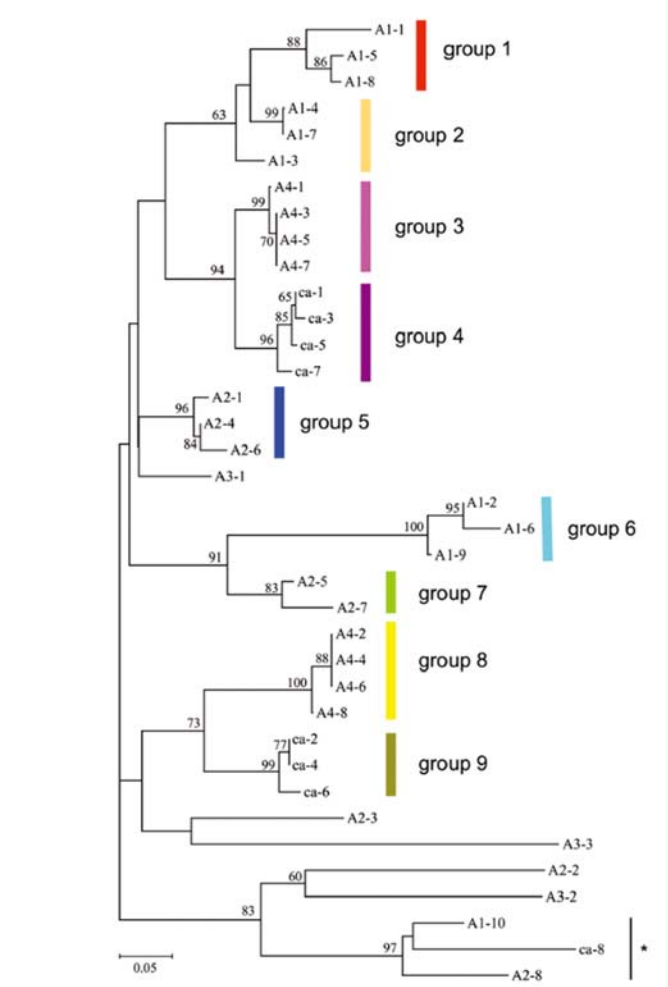

Figure 3. Neighbor-joining tree, constructed using data from the 37 repeat sequences of the ITS2 region in each of the five sibling species in the *Anopheles barbirostris* complex. Numbers above branches indicate the bootstrap value (> 50%). Groups were formed based on pairwise genetic distances of less than 0.1 between repeats, with some exceptions. Although the distance between A1-1 and A1-5 was 0.109, A1-1 was grouped with A1-5 and A1-8 because this inclusion was supported by a high bootstrap value. The distances between the three repeats (ca-1, ca-5, and ca-7) in group 4 and those in group 3 were less than 0.1. However, these two groups were grouped separately based on high bootstrap values. Groups are indicated by color bars to the right of the tree. The colors correspond to those used for the groups of repeats in Figure 1. High quality figures are available online.

Repeats in the ITS2 region of other *Anopheles* species

ITS2 sequences of more than 120 *Anopheles* species deposited in GenBank were surveyed for the presence of repeat sequences by dot plot analyses. Consequently, *An. funestus* Giles (AF062512), *An. beklemishevi* Stegnii and Kabanova (AY593958), and *An. crucians* complex (species A, AY245553; species B, AY386963) were found to have repeats. *An. funestus* and *An. beklemishevi* had two repeats. *An. crucians* species B had three types of repeats with different sequence patterns. The number of repeats in each type was two, five, and three, respectively. The two repeats found in *An. crucians* species A had sequence similarities to the type found in *An. crucians* species B (Figure 5). These repeat

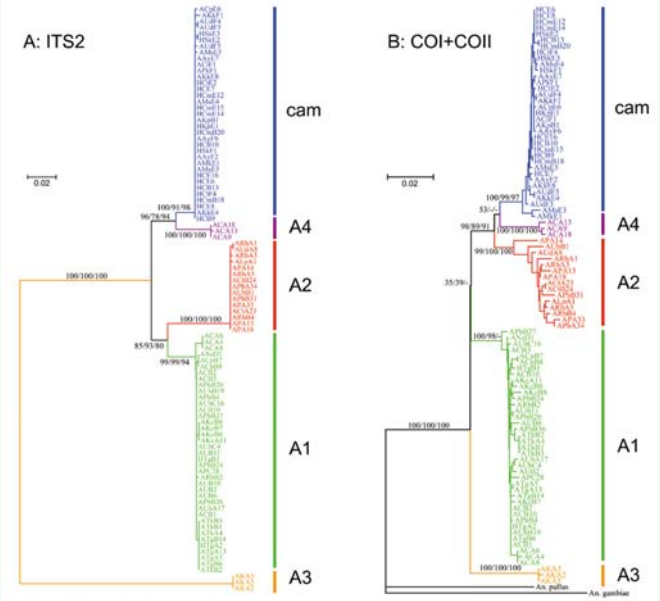

Figure 4. NJ trees of the 95 isolines of the five sibling species in the *Anopheles barbirostris* complex based on analysis of (A) the ITS2 region, with the repeat sequences excluded, and (B) combined sequences of COI and COII. Values at tree nodes indicate the level of support estimated by bootstrap replicates for distance, ML and MP methods, respectively, with dashes indicating a lack of support for the analysis. Colors indicate the isolines obtained in each species. High quality figures are available online.

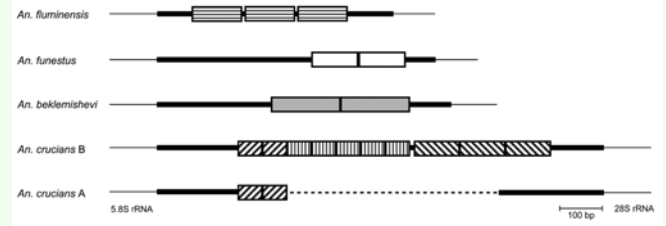

Figure 5. Schematic diagram of the ITS2 region in *Anopheles fluminensis* (DQ328638), *Anopheles funestus* (AF062512), *Anopheles beklemishevi* (AY593958), and *Anopheles crucians* B (AY386963) and A (AY245553). Repeats are shown by rectangles, with different shading patterns indicating no sequence similarity. The ITS2 region, excluding the repeat sequences, is shown by a bold line. The 5.8 and 28S rRNA genes flanking the ITS2 are shown by a thin line. Dots indicate gaps. High quality figures are available online.

sequences found in the ITS2 region of *Anopheles* species were 40-200 bp in length, and no sequence similarity was observed between species except for repeats found in the same species complex. Blast searches in GenBank did not reveal any DNA sequences having significant homology with the repeats. Apart from the repeat sequences in the *An. barbirostris* complex, secondary structures and inverted terminal sequences were not detected in the repeat sequences.

Discussion

At 938 to 1729 bp, the ITS2 region of the five species in the *An. barbirostris* complex is extremely large. Sequence comparisons indicated that the unusual length of ITS2 was due primarily to sequences repeated tandemly. Considerable variation was observed in the number and formation of the repeat regions in the five sibling species of the *An. barbirostris* complex. The repeats could be classified into groups based on the observed variation in the sequences between species. With the exception of *An. barbirostris* A3, the other four species had repeats from the same group, implying that concerted evolution occurred between repeats of the ITS2 region in each species. ITS2 is part of the rDNA cistron, which is also repeated tandemly and undergoes concerted evolution (Brown et al. 1972; Arnheim et al. 1980; Tautz et al. 1987; Porter and Collins 1991). Thus, the repeat sequences in ITS2 are subjected to concerted evolution, at the level of being repeats themselves and as tandem repeats of rDNA. In concerted evolution, sequences of repetitive elements are homogenized, which has occurred in rDNA. However, the repeats in the ITS2 region of *An. barbirostris* could be classified into groups based on sequence differences. In addition, several repeat sequences belonging to different groups were arrayed in turn. For example, the ITS2 region of *An. barbirostris* species A4 consisted of four tandemly arrayed units consisting of a repeat from group 3 and one from group 8. It appears that concerted evolution occurs as a unit of two repeats in *An. barbirostris* species A2, A4, and *An. campestris*-like, and as one of three repeats in *An. barbirostris* species A1. Furthermore, some repeats of an ITS2 region do not appear subjected to concerted evolution. The three repeat sequences of *An.*

barbirostris species A3 (i.e. A3-1, A3-2, and A3-3) were highly divergent, while A1-10, A2-8, and ca-8, all of which belong to different species, were related more closely than other groups of repeats in the same species. Molecular mechanisms, such as unequal crossing over, gene conversion, and transition have been proposed to explain the concerted evolution in rDNA (Smith 1976; Dover 1982, 2002; Eickbush and Eickbush 2007). Indeed, it is possible that different mechanisms operate on concerted evolution of repeats at these levels, i.e., at the level of the repeat itself and at the general level of the rDNA.

In the ITS2 region of *Anopheles* species, simple tandem repeats (2-5 bp in length) are observed frequently (Fritz et al. 1994; Cornel et al. 1996; Xu and Qu 1997; Jariyapan et al. 2005; Li and Wilkerson 2007). Variations in the type and number of simple repeats have been used to distinguish between species of the *An. dirus* complex (Xu and Qu 1997) and *An. albitarsis* complex (Li and Wilkerson 2007), as well as populations of *An. nuneztovari* Gabaldón (Fritz et al. 1994). These simple repeats are thought to arise partly due to slipped-strand mispairing (Levinson and Gutman 1987). In this study, it is shown that longer repeat sequences (> 30 bp) are included in the five sibling species of the *An. barbirostris* complex, and also in another four *Anopheles* species. These findings indicate that although the generation mechanism is unknown, the occurrence of repeats in ITS2 is not rare in *Anopheles* mosquitoes. For example, variation was observed in the number of repeats between the species A and B of *An. crucians* complex, and the other species of *An. crucians* complex were found to have no repeats in the ITS2 region (Wilkerson et al. 2004). Similarly, repeat sequences in the ITS2 region were found in *An. funestus* and *An. beklemishevi*, though several closely related species in the same group had no repeats in the region (Garros et al. 2004; Kampen 2005). While some long ITS2 sequences from *Anopheles* species have been deposited in the GenBank database, they do not contain repeat sequences. If the repeat sequences in the ITS2 region accumulated mutations, then repeat detection would prove difficult. It is thus possible that long ITS2 regions, that apparently lack repeat sequences, may have contained repeats previously, and the accumulated mutations resulted in some of these sequences becoming relics. Indeed, the three repeats in *An. barbirostris* species A3 are so divergent that finding any similarities between them is difficult without comparing them against the other four species in the *An. barbirostris* complex. The core secondary structure of ITS2 is conserved in a wide variety of eukaryotic taxa (Schultz et al. 2005). However, no such conservation of secondary structure has been observed in *Anopheles* mosquitoes to date. This may be one reason why the ITS2 region in some *Anopheles* species contains repeat sequences, which explains the observed size variation in related species.

Exclusion of the repeat sequences from the ITS2 region of the five species of *An. barbirostris* complex facilitated good alignment. The resulting phylogenetic trees showed that A3 was highly distinct from the other four sibling species, which implies that A3 may not be related to the other four species as proposed by Paredes-Esquivel et al. (2009). However, A3 had three repeat sequences that were homologous to the sequences found in the ITS2 region of another member of the *An. barbirostris* complex. In addition, the repeats in all five sibling species were inserted at the same position, implying

that the ITS2 region of these species had a common origin. Furthermore, phylogenetic analysis of the mitochondrial COI and COII sequence data placed A1 almost equally between A3 and the other species, indicating a lack of support for the separation of A3 from the other species obtained in the ITS2 region data. Although the phylogenetic status of A3 remains unclear, future analyses of sequence data in other species of the Barbirostris subgroup will further understanding of the taxonomic position of A3 in this species complex.

Acknowledgements

The author thanks Professor Wej Choochote, Dr. Atiporn Saeung, Dr. Sittiporn Suwannamit, and Mr. Sorawat Thongsahuan from the Department of Parasitology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand for their valuable assistance. In addition, Professor Emeritus Hiroyuki Takaoka from the Institute of Biological Sciences, University of Malaya and former Head of the Department of Infectious Disease Control, Faculty of Medicine, Oita University, for his various comments on this manuscript and interest in this research.

References

Apiwathnasorn C, Prommongkol S, Samung Y, Limrat D, Rojruthai B. 2002. Potential for *Anopheles campestris* (Diptera: Culicidae) to transmit malaria parasites in Pa Rai subdistrict (Aranyaprathet, Sa Kaeo Province), Thailand. *Journal of Medical Entomology* 39(4): 583-586.

Arnheim N, Krystal M, Schmickel R, Wilson G, Ryder O, Zimmer E. 1980. Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proceedings of the National Academy of Sciences USA* 77(12): 7323-7327.

Atomosoedjono S, Van Peenen PF, Putrali J. 1976. *Anopheles barbirostris* (Van der Wulp) still an efficient vector of *Brugia malayi* in Central Sulawesi (Celebes), Indonesia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 70(3): 259.

Brelsfoard CL, Fritz GN, Rodriguez R. 2006. Sequence analysis of the rDNA internal transcribed spacer 2 and polymerase chain reaction identification of *Anopheles fluminensis* (Diptera: Culicidae: *Anopheles*) in Bolivia. *Journal of Medical Entomology* 43(3): 460-466.

Brown DD, Wensink PC, Jordan E. 1972. A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes. *Journal of Molecular Biology* 63(1): 57-73.

Cornel AJ, Porter CH, Collins FH. 1996. Polymerase chain reaction species diagnostic assay for *Anopheles quadrimaculatus* cryptic species (Diptera: Culicidae) based on ribosomal DNA ITS2 sequences. *Journal of Medical Entomology* 33(1): 109-116.

Dover G. 1982. Molecular drive: a cohesive mode of species evolution. *Nature* 299(5879): 111-117.

Dover G. 2002. Molecular drive. *Trends in Genetics* 18(11): 587-589.

Eickbush TH, Eickbush DG. 2007. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 175(2): 477-485.

Fritz GN, Conn J, Cockburn A, Seawright J. 1994. Sequence analysis of the ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae). *Molecular Biology and Evolution* 11(3): 406-416.

Garros C, Koekemoer LL, Kamau L, Awolola TS, Van Bortel W, Coetzee M, Coosemans M, Manguin S. 2004. Restriction fragment length polymorphism method for the identification of major African and Asian malaria vectors within the *Anopheles funestus* and *An. minimus* groups. *American Journal of Tropical Medicine and Hygiene* 70(3): 260-265.

Goswami G, Raghavendra K, Nanda N, Gakhar SK, Subbarao SK. 2005. PCR-RFLP of mitochondrial cytochrome oxidase subunit II and ITS2 of ribosomal DNA: markers for the identification of members of the *Anopheles culicifacies* complex (Diptera: Culicidae). *Acta Tropica* 95(2): 92-99.

Griffith ME. 1955. A note on *Anopheles minimus* Theobald as a malaria vector in Thailand. *VI International Congress of Microbiology* 5: 565-567.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52(5): 696-704.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95-98.

Harrison BA, Scanlon JE. 1975. Medical entomology studies - II. The subgenus *Anopheles* in Thailand (Diptera: Culicidae). *Contributions of the American Entomological Institute* 12(1): 1-307.

Iyengar MOT. 1953. Filariasis in Thailand. *Bulletin of the World Health Organization* 9: 731-766.

Jariyapan N, Choochote W, Junkum A, Jitpakdi A, Komalamisra N, Bates PA, Crampton JM. 2005. Sequence analyses of three nuclear ribosomal loci and a mitochondrial locus in cytologically different forms of Thai *Anopheles aconitus* mosquitoes. *Southeast Asian Journal of Tropical Medicine and Public Health* 36(5): 1162-1173.

Kampen H. 2005. The ITS2 ribosomal DNA of *Anopheles beklemishevi* and further remarks on the phylogenetic relationships within the *Anopheles maculipennis* group of species (Diptera: Culicidae). *Parasitology Research* 97(2): 118-128.

Kirnowardoyo S. 1985. Status of *Anopheles* malaria vectors in Indonesia. *Southeast Asian Journal of Tropical Medicine and Public Health* 16(1): 129-132.

Levinson G, Gutman GA. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* 4(3): 203-221.

Li C, Wilkerson RC. 2007. Intragenomic rDNA ITS2 variation in the neotropical *Anopheles* (*Nyssorhynchus*) *albitarsis* complex (Diptera: Culicidae). *Journal of Heredity* 98(1): 51-59.

Limrat D, Rojruthai B, Apiwathnasorn C, Samung Y, Prommongkol S. 2001. *Anopheles barbirostris*/*campestris* as a probable vector of malaria in Aranyaprathet, Sa Kaeo

Province. *Southeast Asian Journal of Tropical Medicine and Public Health* 32(4): 739-744.

Paredes-Esquivel C, Donnelly MJ, Harbach RE, Townson H. 2009. A molecular phylogeny of mosquitoes in the *Anopheles barbirostris* Subgroup reveals cryptic species: implications for identification of disease vectors. *Molecular Phylogenetics and Evolution* 50(1): 141-151.

Porter CH, Collins FH. 1991. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from the sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). *American Journal of Tropical Medicine and Hygiene* 45(2): 271-279.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics* 14(9): 817-818.

Proft J, Maier WA, Kampen H. 1999. Identification of six sibling species of the *Anopheles maculipennis* complex (Diptera: Culicidae) by a polymerase chain reaction assay. *Parasitology Research* 85(10): 837-843.

Reid JA. 1968. Anopheline mosquitoes of Malaya and Borneo. *Studies from the Institute for Medical Research Malaya* 31: 1-520.

Saeung A, Baimai V, Otsuka Y, Rattanarithikul R, Somboon P, Junkum A, Tuetun B, Takaoka H, Choochote W. 2008. Molecular and cytogenetic evidence of three sibling species of the *Anopheles barbirostris* Form A (Diptera: Culicidae) in Thailand. *Parasitology Research* 102(3): 499-507.

Saeung A, Otsuka Y, Baimai V, Somboon P, Pitasawat B, Tuetun B, Junkum A, Takaoka H, Choochote W. 2007. Cytogenetic and molecular evidence for two species in the *Anopheles barbirostris* complex (Diptera: Culicidae) in Thailand. *Parasitology Research* 101(5): 1337-1344.

Sattabongkot J, Tsuboi T, Zollner GE, Sirichaisinthop J, Cui L. 2004. *Plasmodium vivax* transmission: chances for control? *Trends in Parasitology* 20(4): 192-198.

Schultz J, Maisel S, Gerlach D, Müller T, Wolf M. 2005. A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA* 11(4): 361-364.

Siebert S, Backofen R. 2005. MARNA: multiple alignment and consensus structure prediction of RNAs based on sequence structure comparisons. *Bioinformatics* 21(16): 3352-3359.

Smith GP. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* 191(4227): 528-535.

Suwannamit S, Baimai V, Otsuka Y, Saeung A, Thongsahuan S, Tuetun B, Apiwathnasorn C, Jariyapan N, Somboon P, Takaoka H, Choochote W. 2009. Cytogenetic and molecular evidence for an additional new species within the taxon *Anopheles barbirostris* (Diptera: Culicidae) in Thailand. *Parasitology Research* 104(4): 905-918.

Swofford DL. 1998. PAUP*, Phylogenetic Analysis Using Parsimony (*and other methods) version 4. Sinauer Associates.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0.

Molecular Biology and Evolution 24(8): 1596-1599.

Tautz D, Tautz C, Webb D, Dover GA. 1987. Evolutionary divergence of promoters and spacers in the rDNA family of four *Drosophila* species. Implications for molecular coevolution in multigene families. *Journal of Molecular Biology* 195(3): 525-542.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22(22): 4673-4680.

Thongsahuan S, Baimai V, Otsuka Y, Saeung A, Tuetun B, Jariyapan N, Suwannamit S, Somboon P, Jitpakdi A, Takaoka H, Choochote W. 2009. Karyotypic variation and geographic distribution of *Anopheles campestris*-like (Diptera: Culicidae) in Thailand. *Memórias do Instituto Oswaldo Cruz* 104(4): 558-566.

Wilkerson RC, Reinert JF, Li C. 2004. Ribosomal DNA ITS2 sequences differentiate six species in the *Anopheles crucians* complex (Diptera: Culicidae). *Journal of Medical Entomology* 41(3): 392-401.

Xu JN, Qu FY. 1997. Ribosomal DNA difference between species A and D of the *Anopheles dirus* complex of mosquitoes from China. *Medical and Veterinary Entomology* 11(2): 134-138.

Supplementary table. A list of isolines used in this study with their GenBank accession numbers and localities in Thailand.

		GenBank accesison number			
Species	Isoline	ITS2	COI	COII	Province collected †
An. barbirostris A1	ACB1	AB331555 *	AB331574	AB331593	Chiang Mai
An. barbirostris A1	ACB2	AB331556	AB331575	AB331594	Chiang Mai
An. barbirostris A1	ACB3	AB331557	AB331576	AB331595	Chiang Mai
An. barbirostris A1	ACA4	AB362229	AB362235	AB362241	Chiang Mai
An. barbirostris A1	ACA6	AB362230	AB362236	AB362242	Chiang Mai
An. barbirostris A1	ACA8	AB362231	AB362237	AB362243	Chiang Mai
An. barbirostris A1	ACB10	AB435957	AB435996	AB436035	Chiang Mai
An. barbirostris A1	ATkB1	AB435959	AB435998	AB436037	Tak
An. barbirostris A1	ATkB2	AB435960	AB435999	AB436038	Tak
An. barbirostris A1	ATkB3	AB435961	AB436000	AB436039	Tak
An. barbirostris A1	ATkA4	AB435962	AB436001	AB436040	Tak
An. barbirostris A1	AUbA17	AB435964	AB436003	AB436042	Ubon Ratchathani
An. barbirostris A1	AUB2	AB331558	AB331577	AB331596	Ubon Ratchathani
An. barbirostris A1	AUB6	AB331559	AB331578	AB331597	Ubon Ratchathani
An. barbirostris A1	AUB10	AB331560	AB331579	AB331598	Ubon Ratchathani
An. barbirostris A1	AUB11	AB331561	AB331580	AB331599	Ubon Ratchathani
An. barbirostris A1	AUbB19	AB435965	AB436004	AB436043	Ubon Ratchathani
An. barbirostris A1	AUbC4	AB435966	AB436005	AB436044	Ubon Ratchathani
An. barbirostris A1	AUbC16	AB435967	AB436006	AB436045	Ubon Ratchathani
An. barbirostris A1	AKcA11	AB435969	AB436008	AB436047	Kanchanaburi
An. barbirostris A1	AKcB6	AB435970	AB436009	AB436048	Kanchanaburi
An. barbirostris A1	AKcB7	AB435971	AB436010	AB436049	Kanchanaburi
An. barbirostris A1	AKcB8	AB435972	AB436011	AB436050	Kanchanaburi
An. barbirostris A1	ARbB2	AB435973	AB436012	AB436051	Ratchaburi
An. barbirostris A1	APbB4	AB435978	AB436017	AB436056	Phetchaburi
An. barbirostris A1	APbB20	AB435979	AB436018	AB436057	Phetchaburi
An. barbirostris A1	APbB24	AB435980	AB436019	AB436058	Phetchaburi
An. barbirostris A1	APbB27	AB435981	AB436020	AB436059	Phetchaburi
An. barbirostris A1	APbB36	AB435982	AB436021	AB436060	Phetchaburi
An. barbirostris A1	APC28	AB331562	AB331581	AB331600	Phetchaburi
An. barbirostris A1	ACpB7	AB435987	AB436026	AB436065	Chumphon
An. barbirostris A1	ACpB8	AB435988	AB436027	AB436066	Chumphon
An. barbirostris A1	ANsD1	AB435989	AB436028	AB436067	Nakhon Si Thammarat
An. barbirostris A1	HTgA2	AB435990	AB436029	AB436068	Trang
An. barbirostris A1	HTgB1	AB435991	AB436030	AB436069	Trang
An. barbirostris A1	ATgA5	AB435992	AB436031	AB436070	Trang
An. barbirostris A1	ATgB6	AB435993	AB436032	AB436071	Trang
An. barbirostris A1	ATgA13	AB435994	AB436033	AB436072	Trang
An. barbirostris A1	ATgB14	AB435995	AB436034	AB436073	Trang
An. barbirostris A2	ALpA1	AB435958	AB435997	AB436036	Lampang
An. barbirostris A2	AUdA8	AB435963	AB436002	AB436041	Udon Thani
An. barbirostris A2	AUbB1	AB435968	AB436007	AB436046	Ubon Ratchathani
An. barbirostris A2	ARbA1	AB435974	AB436013	AB436052	Ratchaburi
An. barbirostris A2	ARbA3	AB435975	AB436014	AB436053	Ratchaburi
An. barbirostris A2	ARbB4	AB435976	AB436015	AB436054	Ratchaburi
An. barbirostris A2	ARbA5	AB435977	AB436016	AB436055	Ratchaburi
An. barbirostris A2	APA13	AB331551 *	AB331570	AB331589	Phetchaburi
An. barbirostris A2	APA14	AB331552	AB331571	AB331590	Phetchaburi
An. barbirostris A2	APA18	AB331553	AB331572	AB331591	Phetchaburi
An. barbirostris A2	APbB31	AB435983	AB436022	AB436061	Phetchaburi
An. barbirostris A2	APA33	AB331554	AB331573	AB331592	Phetchaburi
An. barbirostris A2	APbA34	AB435984	AB436023	AB436062	Phetchaburi
An. barbirostris A2	ACtA23	AB435985	AB436024	AB436063	Chanthaburi
An. barbirostris A2	ACtB24	AB435986	AB436025	AB436064	Chanthaburi
An. barbirostris A3	AKA2	AB362232 *	AB362238	AB362244	Kanchanaburi
An. barbirostris A3	AKA3	AB362233	AB362239	AB362245	Kanchanaburi
An. barbirostris A3	AKA5	AB362234	AB362240	AB362246	Kanchanaburi
An. barbirostris A4	ACA9	AB373939 *	AB373942	AB373945	Chiang Mai
An. barbirostris A4	ACA13	AB373940	AB373943	AB373946	Chiang Mai
An. barbirostris A4	ACA18	AB373941	AB373944	AB373947	Chiang Mai
An. campestris-like	HCB9	AB331563 *	AB331582	AB331601	Chiang Mai
An. campestris-like	HCB10	AB331564	AB331583	AB331602	Chiang Mai
An. campestris-like	HCB13	AB331565	AB331584	AB331603	Chiang Mai
An. campestris-like	HCE6	AB331566	AB331585	AB331604	Chiang Mai
An. campestris-like	HCE7	AB331567	AB331586	AB331605	Chiang Mai
An. campestris-like	HCE8	AB331568	AB331587	AB331606	Chiang Mai
An. campestris-like	HCE16	AB331569	AB331588	AB331607	Chiang Mai
An. campestris-like	HCmE12	AB436074	AB436102	AB436130	Chiang Mai
An. campestris-like	HCmE14	AB436075	AB436103	AB436131	Chiang Mai
An. campestris-like	HCmE15	AB436076	AB436104	AB436132	Chiang Mai
An. campestris-like	HCmB18	AB436077	AB436105	AB436133	Chiang Mai
An. campestris-like	HCmB20	AB436078	AB436106	AB436134	Chiang Mai
An. campestris-like	AKpB1	AB436079	AB436107	AB436135	Kamphaeng Phet
An. campestris-like	HKpE1	AB436080	AB436108	AB436136	Kamphaeng Phet
An. campestris-like	AAyF2	AB436081	AB436109	AB436137	Phra Nakhon Si Ayutthaya
An. campestris-like	AAyF6	AB436082	AB436110	AB436138	Phra Nakhon Si Ayutthaya
An. campestris-like	AAyE7	AB436083	AB436111	AB436139	Phra Nakhon Si Ayutthaya
An. campestris-like	AUdF3	AB436084	AB436112	AB436140	Udon Thani
An. campestris-like	AUdF4	AB436085	AB436113	AB436141	Udon Thani
An. campestris-like	AUdF5	AB436086	AB436114	AB436142	Udon Thani
An. campestris-like	AKkF1	AB436087	AB436115	AB436143	Khon Kaen
An. campestris-like	AKkE4	AB436088	AB436116	AB436144	Khon Kaen
An. campestris-like	AKkE8	AB436089	AB436117	AB436145	Khon Kaen
An. campestris-like	AMsE3	AB436090	AB436118	AB436146	Maha Sarakham
An. campestris-like	AMsE4	AB436091	AB436119	AB436147	Maha Sarakham
An. campestris-like	AMsE5	AB436092	AB436120	AB436148	Maha Sarakham
An. campestris-like	AMkE1	AB436093	AB436121	AB436149	Mukdahan
An. campestris-like	ACiF1	AB436094	AB436122	AB436150	Chaiyaphum
An. campestris-like	HSkF1	AB436095	AB436123	AB436151	Sa Kaeo
An. campestris-like	HSkE2	AB436096	AB436124	AB436152	Sa Kaeo
An. campestris-like	HSkE3	AB436097	AB436125	AB436153	Sa Kaeo
An. campestris-like	HCtE2	AB436098	AB436126	AB436154	Chanthaburi
An. campestris-like	HCtF4	AB436099	AB436127	AB436155	Chanthaburi
An. campestris-like	APkF1	AB436100	AB436128	AB436156	Prachuap Khiri Khan
An. campestris-like	ACpE6	AB436101	AB436129	AB436157	Chumphon

* The accession numbers marked with an asterisk were used as the representative sequences to analyze the relationship of repeat sequences. † All isolines were collected in Thailand.