Systematics

Molecular Phylogeny of Neotropical Anopheles (Nyssorhynchus) albitarsis Species Complex (Diptera: Culicidae)

RICHARD C. WILKERSON, PETER G. FOSTER, CONG LI, AND MARIA ANICE MUREB SALLUM

Walter Reed Army Institute of Research, Department of Entomology, 503 Robert Grant Avenue, Silver Spring, MD 20910-7500


ABSTRACT A phylogeny was reconstructed for four species belonging to the Neotropical Anopheles (Nyssorhynchus) albitarsis complex using partial sequences from the mitochondrial cytochrome oxidase 1 (COI) and NADH dehydrogenase 4 (ND4) genes and the ribosomal DNA ITS2 and D2 expansion region of the 28S subunit. The basis for initial characterization of each member of the complex was by correlated random amplification of polymorphic DNA-polymerase chain reaction (RAPD-PCR) markers. Analyses were carried out with and without an outgroup (An. (Nys.) argyritarsis Robineau-Desvoidy) by using maximum parsimony, maximum likelihood, and Bayesian methods. A total evidence approach without the outgroup, using separate models for “fast” (COI and ND4 position 3) and “slow” (rDNA ITS2 and D2, and COI and ND4 position 1) partitions, gave the best supported topology, showing close relationships of An. albitarsis Lynch-Arribliaga to An. albitarsis B and An. marajoara Galvão & Damasceno to An. demeofforum Rosa-Freitas. Analyses with the outgroup included showed poorer support, possibly because of a divergent outgroup, which caused one of the An. marajoara specimens to cluster with An. demeofforum in some analyses. The relationship of the above-mentioned result to a separately proposed hypothesis suggesting a fifth species in the complex is discussed.

KEY WORDS Culicidae, Anopheles albitarsis Complex, molecular phylogeny

Anopheles (Nyssorhynchus) marajoara Galvão & Damasceno, a member of the Albitarsis Complex, was recently recognized as the primary vector of malaria parasites in northeastern Amazonia, Brazil (Conn et al. 2002). Other species in the complex are An. albitarsis Lynch-Arribliaga, An. demeofforum Rosa-Freitas, and an unnamed species “B” (Kreutzer et al. 1976; Linthicum 1988; Narang et al. 1993; Rosu-Freitas and Deane 1989; Wilkerson et al. 1995a, b). From studies carried out in Rondônia State, Brazil, there is also evidence to support the importance of An. demeofforum as an important malaria vector (Klein et al. 1991a, b). Wilkerson et al. (1995a, b) separated the four largely isomorphic species by using species-specific random amplified polymorphic DNA (RAPD-PCR) markers. Their analysis relied on an empirical assumption of multiple correlated “fixed” markers to hypothesize reproductive isolation, similar to the use of correlated morphological characters for the same ends, but they did not address confounding factors inherent in RAPD markers such as possible nonhomology of comigrating bands or linkage of markers. Lehr et al. (2003), based on complete mitochondrial DNA cytochrome oxidase 1 (mtDNA COI) sequences, presented here, support a single intron loss event in this approach, suggest a fifth species, and present a COI gene tree for all five. Using specimens from the Wilkerson et al. (1995a, b) RAPD-PCR studies, Merritt et al. (2005) analyzed a portion of the white gene that contains the white fourth intron. They found that the intron was present in An. marajoara but not in the other three species (also noted by Krzywinski et al. 2001) for An. albitarsis. Phylogenetic analysis of coding sequence in the area of the fourth intron, correlated with intron loss hypotheses, resulted in strong support for a single intron loss event in this species complex and gave a topology different from that found both by Lehr et al. (2005) and from that presented here.

Anopheles subgenus Nyssorhynchus includes 33 species (Harbach 2004), including the two most important vectors in the New World tropics: An. darlingi Root (Linthicum 1988) and An. albimanus Wiedemann (Faran 1980). The subgenus is divided into three sections based on morphological characters: the Argyritarsis Section, which includes the Albitarsis...
# Molecular Phylogeny of Neotropical Anopheles (Nyssorhynchus) albitaris Species Complex (Diptera: Culicidae)

## 1. REPORT DATE
NOV 2005

## 2. REPORT TYPE

## 3. DATES COVERED
00-00-2005 to 00-00-2005

## 4. TITLE AND SUBTITLE
Molecular Phylogeny of Neotropical Anopheles (Nyssorhynchus) albitaris Species Complex (Diptera: Culicidae)

## 6. AUTHOR(S)

## 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Walter Reed Army Institute of Research, Department of Entomology, Silver Spring, MD, 20910

## 8. PERFORMING ORGANIZATION REPORT NUMBER

## 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

## 10. SPONSOR/MONITOR’S ACRONYM(S)

## 11. SPONSOR/MONITOR’S REPORT NUMBER(S)

## 12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

## 13. SUPPLEMENTARY NOTES

## 14. ABSTRACT
see report

## 15. SUBJECT TERMS

## 16. SECURITY CLASSIFICATION OF:
<table>
<thead>
<tr>
<th>a. REPORT</th>
<th>b. ABSTRACT</th>
<th>c. THIS PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>unclassified</td>
<td>unclassified</td>
<td>unclassified</td>
</tr>
</tbody>
</table>

## 17. LIMITATION OF ABSTRACT
Same as Report (SAR)

## 18. NUMBER OF PAGES
8

## 19. NAME OF RESPONSIBLE PERSON

---

Standard Form 298 (Rev. 8-98) 
Prepared by ANSI Z39-18
Complex and *An. darlingi*; the Albinanus Section, which includes *An. albimanus*; and the little known Myzorhynchella Section (Peyton et al. 1992). Relationships among the 33 included species are not well resolved, but it is known that the Argyritarsis and Albinanus Sections are paraphyletic relative to each other because one putative clade contains both *An. darlingi* and *An. albimanus*, suggesting a possible evolutionary link to vector capacity (Conn 1998, Sallum et al. 2000). Cryptic species are common in *An. (Nyssorhynchus)* and in *Anopheles* in general, and most groups that are closely studied yield new taxa (Rosa-Freitas et al. 1998), with *An. darlingi* being an apparent exception (Mangun et al. 1999). Considering the medical importance of this complex and our general lack of knowledge regarding the relationships within subgenus *Nyssorhynchus*, we undertook this study to corroborate results produced by RAPDs and to investigate how the Albitarsis Complex species are related to each other. We report here a molecular phylogenetic analysis of the four species, initially separated by RAPDs, by using two ribosomal DNA (rDNA) sequences, internal transcribed spacer two (ITS2) and the D2 expansion of the 28S subunit (D2), and partial sequence from two mitochondrial genes, NADH dehydrogenase 4 (ND4) and COI.

**Materials and Methods**

**Source and Identification of Specimens.** Morphological characters from Linthicum (1998) were used for identification of *An. (Nyssorhynchus)* s.l. and *An. (Nyssorhynchus)* Robineau-Desvoidy, the outgroup species. The ingroup specimens also were used by Wilkerson et al. (1995a, b) for their studies (Table 1). These were identified to species using species-specific RAPF markers as described by them. For DNA analysis, we used individuals from progeny broods preserved in 100% ethyl alcohol maintained at −70°C. A portion of each brood was retained for morphological study and includes individually reared pin-pointed adults with associated pupal and fourth instar exuviae that are deposited in the Smithsonian Institution, National Museum of Natural History (NMNH) and the Faculdade de Saúde Pública, Universidade de São Paulo (FSP-USP). DNA voucher specimens are deposited in NMNH.

**Laboratory Methods.** DNA was extracted from individual mosquitoes of each species as described in Wilkerson et al. (1993). Portions of the mitochondrial COI and the ND4 genes and the nuclear rDNA, 28S D2 expansion region, and the ITS2 were amplified and sequenced for at least three individuals of each species. The region of each gene that was sequenced, sequences, and positions of the primers used in this study are in Table 2.

**PCR reactions were carried out in a total volume of 50 µl by using standard protocols (Falumbi 1996).** PCR temperature profiles to obtain the above-mentioned sequence were initial denaturation at 95°C for 3 min followed by 40 cycles of denaturation at 94°C for 40 s, annealing at 56°C for 40 s (ITS2 and COI) or 52°C (D2 and ND4); and extension at 72°C for 1 min and final extension at 72°C for 10 min. For sequencing, PCR products were purified using polyethylene glycol (PEG) precipitation (20% PEG 8000 and 2.5 M NaCl). Sequencing reactions were carried out directly on both strands of DNA by using ABI Big Dye chemistry (Applied Biosystems, Foster City, CA), and the sequences were generated with an ABI 377 automated sequencer. The sequences were analyzed and questionable base calls resolved using Sequencer 3.0 (Gene Codes Corp., Ann Arbor, MI). Sequences were initially aligned using ClustalX, ver-

---

**Table 1. List of species and source of specimens used in this study**

<table>
<thead>
<tr>
<th>Species</th>
<th>Tree reference no</th>
<th>Collection no.</th>
<th>Locality of specimen</th>
<th>Coordinates</th>
<th>GenBank accession no. (AY prefix)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ingroup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>A2</td>
<td>BR511(3)</td>
<td>Parauá, near Coaraze</td>
<td>24°04' 5.54&quot; 15' W</td>
<td>ITSS2, D2, ND4, COI</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>A3</td>
<td>BR400(38)</td>
<td>São Paulo, 6 km SW Registro</td>
<td>24°38.8' 47°31.1' W</td>
<td>846322, 846323, 846330, 846307</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>A4</td>
<td>BR500(16)</td>
<td>São Paulo, 6 km SW Registro</td>
<td>24°38.8' 47°31.1' W</td>
<td>846322, 846323, 846330, 846307</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>B5</td>
<td>BR517(4)</td>
<td>Bahia, Itaquara</td>
<td>13°26' 53.99&quot; 56' W</td>
<td>846324, 846325, 846343, 846310</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>B6</td>
<td>BR2005(1)</td>
<td>Bahia, Itaquara</td>
<td>13°26' 53.99&quot; 56' W</td>
<td>846324, 846325, 846343, 846310</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>B7</td>
<td>BR5001(45)</td>
<td>São Paulo, 6 km SW Registro</td>
<td>24°38.8' 47°31.1' W</td>
<td>846324, 846325, 846330, 846310</td>
</tr>
<tr>
<td><em>An. albimanus</em></td>
<td>B8</td>
<td>BR200513(5)</td>
<td>Bahia, Itaquara</td>
<td>13°26' 53.99&quot; 56' W</td>
<td>846324, 846325, 846343, 846310</td>
</tr>
<tr>
<td><em>An. margarita</em></td>
<td>C9</td>
<td>BR2010(9)</td>
<td>Mato Grosso, Prêto do Avevedo</td>
<td>10°27' 5.54&quot; 54' W</td>
<td>846339, 846341, 846331, 846335</td>
</tr>
<tr>
<td><em>An. margarita</em></td>
<td>C10</td>
<td>BR305(12)</td>
<td>Amazonas, Manaus</td>
<td>2°53' 5.60&quot; 15' W</td>
<td>846339, 846341, 846335, 846310</td>
</tr>
<tr>
<td><em>An. margarita</em></td>
<td>C11</td>
<td>HBS44(83)</td>
<td>Mato Grosso, Prêto do Avevedo</td>
<td>10°27' 5.54&quot; 54' W</td>
<td>846339, 846341, 846335, 846310</td>
</tr>
<tr>
<td><em>An. margarita</em></td>
<td>C12</td>
<td>BR/BO01(3)</td>
<td>Pará, Ilha de Marajó</td>
<td>1°00' 5.49&quot; 30' W</td>
<td>846328, 846330, 846310</td>
</tr>
<tr>
<td><em>An. darlingi</em></td>
<td>D15</td>
<td>BR/BO007(11)</td>
<td>Rondônia, Guajará-Mirim</td>
<td>10°50' 6.65&quot; 20' W</td>
<td>846322, 846334, 846331, 846331</td>
</tr>
<tr>
<td><em>An. darlingi</em></td>
<td>D16</td>
<td>BR/BO007(8)</td>
<td>Rondônia, Ariquemes</td>
<td>9°55' 6.34&quot; 04' W</td>
<td>846330, 846335, 846330, 846330</td>
</tr>
<tr>
<td><em>An. darlingi</em></td>
<td>D17</td>
<td>BR/BO007(17), D17</td>
<td>Rondônia, Guajará-Mirim</td>
<td>10°50' 6.65&quot; 20' W</td>
<td>846330, 846335, 846330, 846330</td>
</tr>
<tr>
<td>Outgroup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>An. argyritarsis</em></td>
<td>Anargyritarsis 18</td>
<td>VZ13-100</td>
<td>Venezuela, Merida, Rd to Jaji</td>
<td>7°37.34&quot; N 72°25.62&quot; W</td>
<td>846553, 846537, 846524, 846524</td>
</tr>
<tr>
<td><em>An. argyritarsis</em></td>
<td>Anargyritarsis 19</td>
<td>VZ13-4</td>
<td>Venezuela, Merida, Rd to Jaji</td>
<td>7°37.34&quot; N 72°25.62&quot; W</td>
<td>846553, 846539, 846555, 846561</td>
</tr>
<tr>
<td><em>An. argyritarsis</em></td>
<td>Anargyritarsis 20</td>
<td>BR10-112</td>
<td>Coara, Ubajara</td>
<td>3°51.27&quot; S 40°51.25&quot; W</td>
<td>846554, 846539, 846536, 846532</td>
</tr>
</tbody>
</table>

All localities are in Brazil unless otherwise indicated.
Table 2. Sequences of COI, ND1, D2, and ITS2 primers used in this study

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence (5'→3')</th>
<th>Position</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fly5IF(COI)</td>
<td>GGATTATAGGATTTATATTCT</td>
<td>842–861*</td>
<td>Sallum et al. (2002)</td>
</tr>
<tr>
<td>Fly10F(COI)</td>
<td>GCAAATAACGAAAATCTTCT</td>
<td>1773–1792*</td>
<td>Sallum et al. (2003)</td>
</tr>
<tr>
<td>ND4F</td>
<td>CCCAGCTACAAAATATCACC</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>ND4R</td>
<td>GCGAGGCTTTTTAAAGATTACG</td>
<td>This study</td>
<td></td>
</tr>
<tr>
<td>D2F</td>
<td>AGTCCCTCCTCGTACTGTC</td>
<td>588–597</td>
<td>Sallum et al. (2002)</td>
</tr>
<tr>
<td>D2R</td>
<td>CTCTGTCGTTGCTTCAACGC</td>
<td>821–840</td>
<td>Sallum et al. (2002)</td>
</tr>
<tr>
<td>ITS2F</td>
<td>TC5GAACCTGACGGACATGAA</td>
<td>3' end 5'5S</td>
<td>Cornell et al. (1996)</td>
</tr>
<tr>
<td>ITS2R</td>
<td>ATCCGAAAATTGACCCGCTAGTC</td>
<td>3' end 28S</td>
<td>Cornell et al. (1996)</td>
</tr>
</tbody>
</table>

* Nucleotide position relative to COI and COII sequence of An. quadrimaculatus (NC000575), except for Fly10 whose position is relative to COI gene in Drosophila yakuba.

sion 1.8 (Thompson et al. 1997) and then compared and visually aligned using Se-Al version 2.0a9 (Sequence Alignment Editor, A. Rambaut, University of Oxford) or MacClade (Maddison and Maddison 2000). GenBank accession numbers for all sequences are in Table 1.

Phylogenetic Analyses. Unweighted parsimony analyses were done using PAUP 4.0b10 (Swofford 2004) by using the heuristic search option with TBR branch-swapping with 1000 random-taxon-addition replicates. Parsimony bootstrapping (Felsenstein 1985) used 1000 pseudoreplicates, with 10 random-taxon-addition replicates per pseudoreplicate. Parsimony-uninformative characters and the hyper-variable sites were excluded from all the analyses.

For maximum likelihood (ML), we used PAUP 4.0b10 (Swofford 2004). Starting models were chosen with ModelTest (Posada and Crandall 1998) by using the Akaike Information Criterion (AIC). The resulting tree was saved and also used to test site-specific models. Some maximum likelihood determinations and data manipulations were done using p4 (Foster 2000). GenBank accession numbers for all sequences are in Table 1.

In total, 1,846 sites were included in the analysis (Table 3). Because the outgroup was very divergent, we suspected that it might cause us to choose inappropriate models or be responsible for long branch effects. To test this, the analyses were conducted with and without the outgroup. To better fit models to the data, we reasoned that the available genes could be separated into two groupings based on apparent relative substitution rates, “fast” and “slow”, with the
Phylogenetic Analysis with the Outgroup Excluded. A summary of branch supports for ML and Bayesian analysis for fast partition, slow partition, and combined data are given in Table 4. The best resolution was obtained with combined data (Fig. 1).

Brief mention will first be made here of the separate analyses of fast (mtDNA position 3) and slow (mtDNA position 1 plus rDNA) partitions (Table 4). Using the mtDNA position 3 partition (fast) only, the topologies of the ML and Bayesian analyses were the same, and showed resolution of An. albitarsis and An. albitalis B as separate groups, and good resolution of the (An. marajoara, An. dementorum) group from the others (designated "A," "B," "D," and "C," respectively in Figs. 1 and 2). However, sequences of An. marajoara and An. dementorum were not recovered into two separate nonexclusive clades (de Queiroz 1998) because An. marajoara (C10) and An. dementorum (D17) clustered together in a poorly supported clade that was closer to An. marajoara than to An. dementorum. Using the slow partition by itself (mtDNA position 1 and rDNA), there was more ambiguity, and the topology for the ML analysis differed from the Bayesian analysis. Support for separate groups was generally poor, with the highest support for the branch separating the entire An. marajoara sequence-group from the rest of the tree.

Parsimony analysis of the combined rDNA and mtDNA data sets generated six most passitionally trees (MPTs) (not shown). The strict consensus tree generated from those six MPTs recovered three well-supported groups: An. albitarsis, An. albitalis B, and (An. marajoara + An. dementorum). The latter also were recovered as separate groups but with less support: An. dementorum with 72% and An. marajoara with 91% bootstrap support.

ML and Bayesian analyses were carried out using the mtDNA position 3 in 1 partition, and mtDNA position 1 plus rDNA data in another partition. For the unpartitioned data, ModelTest suggested the TVM + I model; however, the TVM + SS, a site-specific model based on the two partitions, gave a better likelihood, showing an increase of 87.6 log units, so the TVM + SS model was used in all ML analyses. The ML topology, including the nonparametric bootstrap support values, is shown in Fig. 1. For Bayesian analysis, the data were partitioned in the same way. The general time-revers-

**Table 4.** ML bootstrap support and posterior probability (PP) for relationships within the An. albitarsis complex, when the outgroup was both included and excluded.

<table>
<thead>
<tr>
<th>Phylogenetic group</th>
<th>COI + ND4 pos 3 with</th>
<th>COI + ND4 pos 3 without</th>
<th>ITS2 + D2+ (COI + ND4 pos 1) with</th>
<th>ITS2 + D2+ (COI + ND4 pos 1) without</th>
<th>rDNA + mtDNA with</th>
<th>rDNA + mtDNA without</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ML PP</td>
<td>ML PP</td>
<td>ML PP</td>
<td>ML PP</td>
<td>ML PP</td>
<td>ML PP</td>
</tr>
<tr>
<td>An. albitarsis</td>
<td>0.69 0.93</td>
<td>0.96 1.00</td>
<td>0.53 0.60</td>
<td>0.76 0.65</td>
<td>0.92 0.89</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>An. albitalis B</td>
<td>0.96 0.91</td>
<td>0.99 1.00</td>
<td>0.77 0.81</td>
<td>0.83 0.95</td>
<td>0.92 0.96</td>
<td>1.00 1.00</td>
</tr>
<tr>
<td>An. dementorum</td>
<td>0.88 0.91</td>
<td>0.92 0.99</td>
<td>0.75 0.78</td>
<td>0.89 0.97</td>
<td>0.83 0.85</td>
<td>0.83 0.76</td>
</tr>
<tr>
<td>An. marajoara</td>
<td>0.51 0.59</td>
<td>0.89 1.00</td>
<td>0.25 0.52</td>
<td>0.57 0.63</td>
<td>0.94 0.94</td>
<td>0.95 0.95</td>
</tr>
<tr>
<td>An. marajoara (C10)</td>
<td>0.55 0.45</td>
<td>0.63 0.56</td>
<td>0.25 0.40</td>
<td>0.76 0.63</td>
<td>0.86 0.66</td>
<td>1.00 1.00</td>
</tr>
</tbody>
</table>

Fig. 1. Results of combined data. The data were placed in two partitions as described in the text, consisting of mtDNA position 3 from COI and ND4 in partition 1, and mtDNA position 1 and ribosomal sequences D2 and ITS2 in the other. The maximum likelihood bootstrap tree and the Bayesian consensus tree had identical topologies, as shown, with support (ML bootstrap/Bayesian posterior probability) shown for the major groups. The ML analysis used the TVM + SS model in PAUP. The Bayesian analysis used a site-specific model, where a GTR + G model was applied to the position 3 mtDNA data partition and GTR + plnv model to position 1 mtDNA data plus rDNA data partition. A, An. albitarsis; B, An. albitalis B; C, An. marajoara; D, An. dementorum.
An. albitarsis (C11), An. marajoara (C12), and An. marajoara (C9) formed a better supported group (72% bootstrap value). Bootstrap support for the clade (An. albitarsis, An. albitarsis B) is moderate (93%) and for An. deaneorum is low (81%).

For the combined data, ModelTest suggested the GTR + I + G model. However, the GTR + SS model, where the data were partitioned into mitochondrial position 3 partition, and mitochondrial position 1 plus ribosomal partitions, had a better likelihood, with an increase of 86 log units, and so the GTR + SS model was used for ML analysis in PAUP. Using ML two major groups were recovered [(An. deaneorum, An. marajoara), (An. albitarsis, An. albitarsis B) outgroup]. Support for placement of the outgroup is low (51%), but the groups consisting of An. albitarsis and An. albitarsis B are moderately and strongly supported (92 and 99%, respectively) (Table 4). For the Bayesian analysis, a site-specific model was used, but a test was made in p4 using maximum likelihood to determine whether individual rate matrices and compositions in the two partitions were better than using a single overall rate matrix and composition. The increase in the log likelihood owing to a separate rate matrix and composition was 205, which is highly significant, and so this strategy was used in the Bayesian analysis. The settings were as described above for combined data without the outgroup. That includes using gamma model for the mitochondrial position 3 partition, and a pIvar model for the ribosomal partition. The support for relationships among the sequences of each ingroup taxon was generally lower when the outgroup was included in both ML and Bayesian analyses than when the outgroup was excluded (Table 4).

To estimate the distribution of likely root positions, the combined post-burn-in MCMC samples from two runs were reanalyzed to obtain a consensus tree based on retained root information (Fig. 2). The largest number of input trees (667/2,000) had the outgroup attached on the branch separating the An. albitarsis B group from the rest of the tree. However, many other input trees had the outgroup attached on the branches leading to the An. marajoara, An. deaneorum clade, and many other input trees had the outgroup attached on the branch leading to the An. albitarsis clade. Also, many trees had the outgroup attached on the branches leading to sequences C10 (An. marajoara) and A4 (An. albitarsis), which we interpret to be caused by long branch effects and disregard. Few trees had the root attached along the branches leading to either the An. marajoara or An. deaneorum clades separately, showing little evidence for placement of the root on this part of the tree. However, note that the ML bootstrap tree has this root placement (not shown). Pending further observations we conclude that the rooting is as shown or on the branches leading to the (An. marajoara, An. deaneorum) clade or to the An. albitarsis clade. Stated another way, we find that the largest number of sampled trees had the outgroup attach.

![Diagram of phylogenetic relationships](image)

**Fig. 2.** Outgroup attachment distribution. From post-burn-in samples from two MCMC runs using the combined data including the outgroup, a consensus tree was made such that the position of the attachment point of the outgroup was preserved. Of the 2000 samples, most (677) had the out group attach as shown, at the base of the An. albitarsis B clade. However, during the MCMC the outgroup attachment position varied widely, spending its time on the branches in proportion to the numbers shown. A, An. albitarsis B, An. albitarsis B; C, An. marajoara; D, An. deaneorum.

Possible (GTR) rate matrix and composition parameters were unlinked between the two partitions. The applicability of this model was confirmed by ML in p4, which showed an increase of 137 log units by allowing free rates and compositions in the two partitions, at a cost of eight parameters, compared with having the same rate matrix and composition in both partitions. As suggested by ModelTest, gamma-distributed among-site rate variation was applied to the mitochondrial position 3 partition, and a pIvar model was applied to the mitochondrial plus ribosomal partitions. The Bayesian tree has the same topology as the ML Bootstrap tree. Bayesian posterior probabilities are shown in Fig. 1. Three groups (An. albitarsis, An. albitarsis B, and An. marajoara) were recovered with high support, and a fourth group, An. deaneorum, was recovered with somewhat less support. The support for the split between (An. albitarsis, An. albitarsis B) and (An. marajoara, An. deaneorum) was also high.

Phylogenetic Analysis with the Outgroup Included.

Here, we note results for the combined data and address the question of where the outgroup attaches to the ingroup (tree topologies not shown). Parsimony analysis (not shown) of the combined rDNA and mtDNA data sets generated 12 MPTs. The strict consensus tree generated from those 12 MPTs recovered one major group consisting of (An. albitarsis + An. albitarsis B), and two minor groups, one formed by An. deaneorum and the other by An. marajoara except An. marajoara (C10). The relationship between An. marajoara and An. deaneorum is very poorly supported (<49%), as is the group consisting of An. marajoara (<33%). Within the An. marajoara clade, An. marajoara (C11), An. marajoara (C12), and An. marajoara (C9) formed a better supported group (72% bootstrap value). Bootstrap support for the clade (An. albitarsis, An. albitarsis B) is moderate (93%) and for An. deaneorum is low (81%).

Discussion

In a search for the best evolutionary hypothesis for the Albitaris Complex, we used partial sequences of two mitochondrial genes (COI and ND4), and two ribosomal DNA fragments (ITS2 and D2 expansion region of the 28S subunit) and compared maximum parsimony, maximum likelihood, and Bayesian analyses with several combinations of data partitions. Individual genes failed to give well-resolved trees, possibly because of the low number of variable sites. Also, to optimize our results we analyzed different data partition combinations, finally settling on two partitions: mtDNA position 3 alone because it was more variable and presumably faster evolving, and mtDNA position 1 plus rDNA because they were less variable and presumably slower evolving. In addition, because the outgroup was highly divergent, we tested for long branch effects by carrying out all analyses with and without the outgroup.

The strongest support for the evolutionary relationships among the four species tested was retrieved when all four genes were combined and partitioned as described above. Analyses excluding the outgroup, presumably more independent of long branch effects, offer our best hypothesis for the ingroup topology (Fig. 1). In summary, four major evolutionary lineages were recovered. Two groups, An. albitaris/An. albitaris B and An. marajoara/An. deanorum, were usually recovered, but not in all analyses. The latter group includes two species that are important vectors of human Plasmodium in localities in the Amazonas region of Brazil (Klein et al. 1991a, b; Conn et al. 2002). This suggests a possible phylogenetic link associated with the ability to transmit human malaria parasites. Similar conclusions have been made for An. albimarginatus and An. darlingi (Conn 1998).

When the mtDNA position 3 partition was analyzed alone, An. deanorum specimen D17 and An. marajoara specimen C10 clustered together sister to the remaining An. marajoara individuals. This result suggests the possibility of incomplete lineage sorting, introgression (Donnelly et al. 2004), or even the existence of an additional taxon.

Our analysis of the distribution of possible roots used sampled trees from MCMC runs, including the outgroup. It showed that An. albitaris and An. albitaris B are in an ambiguous position with relation to the root but that (An. marajoara, An. deanorum) generally formed a clade. Therefore, we think that the root is either as shown in Fig. 2 or as evidenced by the large number of MCMC sampled trees rooting there, either at the base of the An. albitaris B clade, or at the base of the (An. marajoara, An. deanorum) clade. A more definitive answer to this question requires more data.

Lehr et al. (2005) using the entire COI gene sequence data of 29 individuals of An. albitaris, An. albitaris B, An. deanorum, and An. marajoara recovered results similar to those generated in the current study with the mtDNA position 3 data partition, in that there was nonexclusivity of the An. deanorum clade with respect to An. marajoara. They showed the exclusivity of the sequences of An. albitaris, An. albitaris B, and An. marajoara and a sister group relationship of these two taxa. They also found four individuals of An. marajoara that fell outside the remaining sequences of the Albitaris Complex (in Bayesian topology), which they suggest represents a fifth species. The individuals that were used to generate these sequences were collected in Roraima State, Boa Vista, Brazil, and Venezuela. Interestingly, Lehr et al. (2005) also recovered a nonexclusive clade consisting of individuals of An. marajoara and An. deanorum. A similar grouping was found when we analyzed the mtDNA data partition for the current study (An. marajoara specimen C10 clustered with An. deanorum specimen D17). Lack of exclusivity of sequences of An. marajoara and An. deanorum are similar to our results and are also suggestive of ancestral introgression or perhaps a recent speciation event that could not be detected by partial sequences of the mitochondrial genes COI and ND4.

The current study was based on conclusions about species boundaries reached using fixed RAPD markers (Wilkerson et al. 1995a, b). Our results corroborate the RAPD evidence that indicates four putative species: An. albitaris, An. albitaris B, An. marajoara, and An. deanorum. No additional taxa were detected. This is not surprising given that the same genetic material was used in both studies. The existence of a fifth species as reported by Lehr et al. (2005) in Boa Vista, Brazil, was not directly tested by us using sequence as described in this study. However, we assumed the Boa Vista specimens to be An. marajoara based on comparison to one or two diagnostic RAPD markers (data not shown). The possibility of a fifth species is supported by independent data sets: Kreutzer et al. (1976) (chromosomes), Rosa-Freitas et al. (1990) (isozymes), and possibly Narang et al. (1993) (allozymes, mtDNA restriction fragment length polymorphisms). However, in support of a hypothesis for the existence of An. marajoara as a single widespread species is an extensive data set of rDNA ITS2 sequence from throughout its range, and taxon-specific PCR primers based on that sequence (Li and Wilkerson 2005). If in fact there is a fifth member of this complex, the RAPD results should be revisited because assumptions about the wide distribution of An. marajoara (Venezuela to southern São Paulo State) were based on the existence of seven RAPD markers that were found in nearly all individuals tested from all parts of its putative range (Wilkerson et al. 1995a, b).
The topology of the gene tree reported by Merritt et al. (2005), who used coding sequence of a portion of the white gene containing its fourth intron, varied significantly from that found by us and Lehr et al. (2005). They found good statistical support for a single loss of the fourth intron in the species complex (present in An. marajoara, absent in the other species), but weak evidence that An. marajoara is basal relative to [An. albitarsis B, (An. albitarisis, An. deaneorum)]. The alternative topology placed An. marajoara sister to An. albitarsis B. There was high support for the sister relationship of An. albitarisis and An. deaneorum. This is in contrast to our results that give high support for a close relationship between An. albitarisis and An. albitarsis B and between An. marajoara and An. deaneorum.

The above-mentioned conflicting results will certainly require additional data to resolve. A recent report by Besansky et al. (2003) addressed the issue of conflicting data sets in the resolution of species boundaries and phylogenetic relationships in the An. gambiae complex that may be germane in solving the issue of An. marajoara, and resolving the phylogenetic relationships of the Albitarsis Complex. They found evidence supporting both introgression and reproductive isolation as well as different tree topologies, depending on which sequence was sampled. They concluded that adoption of a "total evidence" approach for phylogenetic analysis of closely related species runs a risk of recovering a highly supported wrong answer and suggested that at the level of closely related species, it would be better to do a careful locus-by-locus assessment of sequence divergence rather than just adopt a total evidence approach. They were able to use various genes on all the chromosomes, inside and out of inversions, as well as mitochondrial genes, for their conclusions. The approach of Besansky et al. (2003) provides a model for future research into the phylogenetic relationships of the Albitarsis Complex.

Acknowledgments

We greatly appreciate the support of those who helped collect and rear many of the specimens used in this study, including Terry Klein and Eric Milstrey who were then members of the United States Army Medical Research Unit in Rio de Janeiro. But we especially thank Jose Bento Lima without whose extra efforts many of samples would not have been available to us. We also thank Lee Weigl at the Smithsonian Institution, Laboratory of Analytical Biology, for use of laboratory and sequencing facilities and Jan Conn and Meg Lehr for thoughtful review of a draft of this manuscript. This research was performed under a Memorandum of Understanding between the Walter Reed Army Institute of Research and the Smithsonian Institution, with institutional support provided by both organizations. This work was supported by National Institutes of Health (AI) R0154130 to J.C.

References Cited


Li, C., and R. C. Wilkerson. 2005. Identification of Anophel­
es (Nyssorhynchus) albitarsis complex species (Diptera:


Received 15 February 2005; accepted 28 July 2005.