Mitochondrial Markers for Molecular Identification of *Aedes* Mosquitoes (Diptera: Culicidae) Involved in Transmission of Arboviral Disease in West Africa

SHELLEY COOK,[1] MAWLOUTH DIALLO,[2] AMADOU A. SALL,[2] ALAN COOPER,[1] AND EDWARD C. HOLMES[1]

J. Med. Entomol. 42(1): 19–28 (2005)

ABSTRACT Correct classification of the insect vector is central to the study of arboviral disease. A simple molecular method for identification of the main vectors of the mosquito-borne viruses, dengue, yellow fever, and Rift Valley fever in Senegal, West Africa, was developed. We present a system in which the five mosquito species (Diptera: Culicidae) responsible for the majority of flaviviral disease transmission in Senegal can be reliably identified using small amounts of DNA coextracted during flaviviral screening procedures, via an easy amplification of the mitochondrial gene cytochrome oxidase c subunit I or II (COI or COII, respectively). We observed that despite very similar morphology, the two cryptic disease vector species *Aedes furcifer* Edwards and *Aedes taylori* Edwards are highly divergent at the molecular level. This sequence variation was used as a basis for the development of a polymerase chain reaction-restriction fragment-length polymorphism system for the differentiation of the two species. We also present the first investigation of the phylogeny of the culicine mosquitoes based on all COI and COII sequences currently available. There seems to be very low intraspecific variation in both genes, whereas interspecific variation is high. As a consequence, COI and COII are ideal candidates for the molecular identification of disease vectors to species level, whereas deeper divergences remain equivocal by using these genes. This system provides a new technique for the accurate identification of culicine disease vectors in West Africa and provides a basis for the expansion of such methods into the study of a range of diseases.

KEY WORDS *Aedes*, mitochondrial DNA, PCR-restriction fragment-length polymorphism

VECTOR-BORNE DISEASES SUCH AS malaria, filariasis, and arboviruses have been of long-standing interest for entomologists because they represent major public health problems worldwide. For this reason, mosquito vectors have been well described and morphological keys are available for a majority of species for identification and taxonomic purposes (Munstermann 1995). Although these morphological keys remain essential, they have several limitations including that 1) some mosquito species are barely distinguishable morphologically and sometimes not at all, 2) identification requires highly trained entomologists, and 3) the time and conditions necessary for identification are not always compatible with the process of pathogen isolation. The advent of molecular biology not only allowed these limitations to be overcome but also led to the expansion of identification protocols by providing molecular tools to distinguish "sibling" or morphologically identical species. This has greatly improved our understanding of the transmission and epidemiology of vector-borne diseases. For example, unraveling the interspecific relationships of the *Anopheles gambiae* complex, the primary vector for malaria in Africa, by molecular methods, including early genomic hybridization work (Gale and Crampton 1987) and polymerase chain reaction (PCR) on rDNA intergenic spacer regions (Paskewitz and Collins 1990, Paskewitz et al. 1993), led to the identification of seven species within the complex and allowed the relevant modification of control strategies. Likewise, Reno et al. (2000) used molecular tools to differentiate the culicine mosquitoes *Aedes triseriatus* Say and *Aedes hendersoni* Cockerell at the interspecific level (Reno et al. 2000).

Thus far, studies aimed at the molecular identification of arboviral vectors are limited in comparison with malaria research, despite the medical importance of diseases such as dengue, yellow fever, and Rift Valley fever, which are among the most important viral diseases transmitted to humans by mosquitoes. Dengue virus (DENV; family Flaviviridae) is prevalent in >100 countries and 2.5 million people live in dengue-endemic areas (Guzman and Kouri 2002).

[1] Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom.
[2] Institut Pasteur de Dakar, 36, Avenue Pasteur, Dakar, Senegal.

0022-2585/05/0019–0028$04.00/0 © 2005 Entomological Society of America

Yellow fever virus (YFV; family Flaviviridae) affects 200,000 persons annually (Monath 2001), with a case fatality rate of ≈20% (Monath and Heinz 1996). Finally, Rift Valley fever virus (RVFV; family Bunyaviridae) infects humans, cattle, and sheep and threatens the economy of sub-Saharan Africa and the Arabic peninsula during large epizootics. The main vector of both DENV and YFV is *Aedes aegypti* L. However, the morphological identification of domestic versus sylvatic forms of this species remains controversial, whereas classification of *Aedes vexans* Meigen, one of the probable vectors of RVFV in West Africa (Fontenille et al. 1998) as a species complex is still under discussion due to difficulties in morphologically distinguishing the two candidate subspecies, *Aedes vexans vexans* versus *Aedes vexans arabiensis* Patton (Edwards 1941, White 1975)

In Senegal, YFV and DENV transmission are mediated by *Ae. aegypti*, *Aedes luteocephalus* Newstead, *Aedes taylori* Edwards, and *Aedes furcifer* Edwards (Diallo et al. 2003). The latter two species are morphologically very similar and identification can be very tedious, even for the specialist. Indeed, despite morphological keys proposed by Ferrara et al. (1983), the identification of the two species remains difficult and is subject to many errors. Some studies have even treated them as a single species complex, *Ae. furcifer-taylori* (Inoue 2003), although evidence suggests that the two vector species may play different roles in disease transmission (Diallo et al. 2003). Similarly, *Ae. luteocephalus* is morphologically very close to other species belonging to the *Stegomyia* subgenus such as *Aedes africanus* Theobald, *Aedes neoafricans* Theobald, and *Aedes opock* Corbet & Van Someren, which are present in many YFV and DENV study sites. Furthermore, although RVFV is of major economic and veterinary importance in Senegal and its bordering countries, the taxonomic status of its main vector, *Ae. vexans* is still questionable. Beyond RVFV, *Ae. vexans* also is implicated in the transmission of West Nile virus (WNV) and Wesselbron virus in West Africa. Recently, serious concern has been raised regarding the worldwide distribution of *Ae. vexans* (Horsfall 1972) due to its potential role in the colonization of new geographic areas by RVFV and WNV, as highlighted during the epidemic/epizootic in Saudi Arabia and New York, respectively (Jupp et al. 2002, Miller et al. 2002).

It is therefore clear that a molecular method for the identification of *Ae. vexans*, *Ae. luteocephalus*, and the two cryptic species *Ae. taylori* and *Ae. furcifer*, would greatly aid the understanding and foster further studies of arboviral disease transmission and control in West Africa. In particular, the classification of the latter two species currently relies solely on morphological keys. For that reason, the purpose of this study was to develop, evaluate, and validate a method based on molecular and phylogenetic tools for the identification of *Aedes* mosquitoes responsible for arboviral disease transmission in West Africa.

Materials and Methods

Taxa. The taxa investigated during the study are listed in Table 1. Seventeen *Aedes* voucher specimens were collected (see Fig. 1 for locations) and morphologically identified in Senegal by a resident expert entomologist. Two *Ae. aegypti* voucher samples from laboratory colonies in the United Kingdom and the United States also were used to develop and assess the method. In addition, one *Culex* specimen was included as a control. Samples are listed in Table 1A. After development of the cytochrome oxidase c subunit I or II (COI and COII, respectively) hypothesized phylogenetic "backbones" for the study, 42 test samples previously identified morphologically in Senegal, were tested blind in Oxford by using the newly developed assay. These are listed in Table 1B. All specimens were placed on ice when captured and stored at −80°C. In addition, outgroup species comprising three members of the subfamily Anophelinae (namely, *Chagasia bathana* Dyar, *Bironella gracilis* Theobald, and *An. gambiae*), a member of the Toxorhynchitinae (namely, *Tx. Amboinensis* Doleschall) and a member of the Culicini tribe (namely, *Culex tarsalis* Coquillett) also were included in the analysis.

DNA Extraction, Amplification, and Sequencing. Nucleic acids from mosquito heads were extracted using Promega RNAgents (Southampton, United Kingdom), according to manufacturer's instructions, to a total elution volume of 20 μl. One microliter of a 1:10 extraction dilution was used in PCR reactions. The remainder of the extraction and mosquito bodies were kept for further analysis and viral screening.

The COI gene was amplified using the primers UEA3 (5′-TATRGCWTTYCCWCGAATAAATAA-3′) (Lunt et al. 1996) and Fly10 (5′-ASTGCACTA-ATCTGCCATATTAG-3′) (Sallum et al. 2002). COII sequences were amplified using the modified primers SCTL2-J-3037 (5′-ATGGCAGATTAGTGCAATGA-3′) (Liu and Beckenbach 1992), TK-N-3785 (5′-GTT-TAAGAGACCAGTACTTG-3′) (Liu and Beckenbach 1992) and modified SC1-TD-N-3862 (5′-TTTAGTTT-GACATACTAAKGTTAT-3′) (Liu and Beckenbach 1992).

Amplifications were performed in 25-μl volumes. Each PCR master mix was prepared in a flow hood and consisted of the following per reaction tube: 2.5 μl of reaction buffer [750 mM Tris-HCl, pH 8.8, 200 mM $(NH_4)SO_4$, 0.1% (vol:vol) Tween 20) (ABGene, Surrey, United Kingdom), 1 μl of magnesium chloride (25 mM, ABGene), 0.25 μl of dNTPs (100 mM each, ABGene), 0.35 μl of forward primer (10 μM, Invitrogen, Paisley, United Kingdom), 0.35 μl of reverse primer (10 μM, Invitrogen), 0.25 μl of *Taq* polymerase (5 U/μl Thermoprime Plus DNA polymerase, ABGene), 17.3 μl of distilled, deionized water (ddH_2O), and 1 μl of the 1:10 dilution of the mosquito extraction.

Reaction conditions for the COI gene consisted of 2 min at 94°C, 35 cycles of 20 s at 94°C, 20 s at 50°C, and 60 s at 72°C, followed by final extension time of 7 min at 72°C. For the COII gene conditions were identical except for the annealing temperature, which was 52°C.

Table 1. Sequences used for the study

Voucher	Individual	Origin	COI	COII
		A		
1	*Ae. furcifer V1*	Kedougou	AY645210	AY645272
2	*Ae. furcifer V2*	Kedougou	AY645211	AY645273
3	*Ae. furcifer V3*	Kedougou	AY645212	AY645274
4	*Ae. taylori V1*	Kedougou	AY645226	AY645288
5	*Ae. taylori V2*	Kedougou	AY645227	AY645289
6	*Ae. taylori V3*	Kedougou	AY645228	AY645290
7	*Ae. taylori V4*	Kedougou	AY645229	AY645291
8	*Ae. luteocephalus V1*	Kedougou	AY645234	AY645296
9	*Ae. luteocephalus V2*	Kedougou	AY645235	AY645297
10	*Ae. luteocephalus V3*	Kedougou	AY645236	AY645298
11	*Ae. luteocephalus V4*	Kedougou	AY645237	AY645299
12	*Culex decens V1*	Kedougou	AY645241	AY645303
13	*Ae. furcifer V4*	Kedougou	AY645213	AY645275
14	*Ae. furcifer V5*	Kedougou	AY645214	AY645276
15	*Ae. furcifer V6*	Barkedji	AY645215	AY645277
16	*Ae. vexans V1*	Barkedji	AY645242	AY645304
17	*Ae. vexans V2*	Barkedji	AY645243	AY645305
18	*Ae. vexans V3*	Barkedji	AY645244	AY645306
19	*Ae. aegypti V1*	Laboratory Colony UK, Oxford	AY645248	AY645310
20	*Ae. aegypti V2*	Laboratory Colony USA, Texas	AY645249	AY645311
		B		
Test	Individual	Origin	COI	COII
1	*Ae. aegypti T1*	Ndione	AY645250	AY645312
2	*Ae. aegypti T2*	Ndione	AY645251	AY645313
3	*Ae. aegypti T3*	Ngoye	AY645252	AY645314
4	*Ae. aegypti T4*	Ngoye	AY645253	AY645315
5	*Ae. aegypti T5*	Ngoye	AY645254	AY645316
6	*Ae. aegypti T6*	Nguel	AY645255	AY645317
7	*Ae. aegypti T7*	Nguel	AY645256	AY645318
8	*Ae. aegypti T8*	Nguel	AY645257	AY645319
9	*Ae. aegypti T9*	Ngascope	AY645258	AY645320
10	*Ae. aegypti T10*	Ngascope	AY645259	AY645321
11	*Ae. aegypti T11*	Ngascope	AY645260	AY645322
12	*Ae. aegypti T12*	Touba Diacksao	AY645261	AY645323
13	*Ae. aegypti T13*	Touba Diacksao	AY645262	AY645324
14	*Ae. aegypti T14*	Touba Diacksao	AY645263	AY645325
15	*Ae. aegypti T15*	Thiakhat	AY645264	AY645326
16	*Ae. aegypti T16*	Thiakhat	AY645265	AY645327
17	*Ae. aegypti T17*	Thiakhat	AY645266	AY645328
18	*Ae. aegypti T18*	Kédougou	AY645267	AY645329
19	*Ae. aegypti T19*	Kédougou	AY645268	AY645330
20	*Ae. aegypti T20*	Kédougou	AY645269	AY645331
21	*Ae. furcifer T1*	Darou Salam Lewe	AY645216	AY645278
22	*Ae. furcifer T2*	Darou Salam Lewe	AY645217	AY645279
23	*Ae. furcifer T3*	Darou Salam Lewe	AY645218	AY645280
24	*Ae. furcifer T4*	Belel Doki	AY645219	AY645281
25	*Ae. furcifer T5*	Belel Doki	AY645220	AY645282
26	*Ae. taylori T1*	Kédougou	AY645230	AY645292
27	*Ae. taylori T2*	Kédougou	AY645231	AY645293
28	*Ae. taylori T3*	Kédougou	AY645232	AY645294
29	*Ae. luteocephalus T1*	Kédougou	AY645238	AY645300
30	*Ae. luteocephalus T2*	Kédougou	AY645239	AY645301
31	*Ae. luteocephalus T3*	Kédougou	AY645240	AY645302
32	*Ae. aegypti T21*	Barkedji	AY645270	AY645332
33	*Ae. aegypti T22*	Barkedji	AY645271	AY645333
34	*Ae. furcifer T6*	Barkedji	AY645221	AY645283
35	*Ae. furcifer T7*	Barkedji	AY645222	AY645284
36	*Ae. furcifer T8*	Barkedji	AY645223	AY645285
37	*Ae. vexans T1*	Barkedji	AY645245	AY645307
38	*Ae. vexans T2*	Barkedji	AY645246	AY645308
39	*Ae. vexans T3*	Barkedji	AY645247	AY645309
40	*Ae. furcifer T9*	Touba Diacksao	AY645224	AY645286
41	*Ae. taylori T4*	Kédougou	AY645233	AY645295
42	*Ae. furcifer T10*	Kédougou	AY645225	AY645287

Table 1. Continued

Species	COI	Origin	COII	Origin
		C		
Ch. bathana	AF417726	Equador, Teputini	AF417762	Equador, Teputini
An. gambiae	AF417706	LC Gambia, CDC, G3 strain	AF417742	LC Gambia, CDC, G3 strain
Bironella gracilis	AF417725	Papua New Guinea, Sepikarka	AF417761	Papua New Guinea, Sepikarka
Tx. amboinensis	AF417727	LC Puerto Rico, CDC	AF417763	LC Puerto Rico, CDC
Culex tarsalis	AF425847	LC USA, Queens University	AF425847	LC USA, Queens University
Ae. cinereus	AF253027	France, Savoie	N/A	N/A
Ae. rusticus	AF253038	France, Savoie	N/A	N/A
Ae. sticticus	AF253039	France, Savoie	N/A	N/A
Ae. sticticus	AF253040	France, Savoie	N/A	N/A
Ae. mariae	AF253032	France, Cap Martin	N/A	N/A
Ae. mariae	AF253033	France, Collioure	N/A	N/A
Ae. caspius	AF253024	France, Alsace	N/A	N/A
Ae. caspius	AF253025	France, Lorraine	N/A	N/A
Ae. geniculatus	AF253031	France, Savoie	N/A	N/A
Ae. detritus	AF253030	France, Camargue	N/A	N/A
Ae. detritus	AF253029	France, Camargue	N/A	N/A
Ae. detritus	AF253028	Spain, Northern	N/A	N/A
Ae. pullatus	AF253034	France: Oisans massif	N/A	N/A
Ae. cantans	AF253023	France: Savoie	N/A	N/A
Ae. cataphylla	AF253026	France: Savoie	N/A	N/A
Ae. punctor	AF253035	France: Grandes Rousses massif	N/A	N/A
Ae. punctor	AF253037	France: Taillefer massif	N/A	N/A
Ae. punctor	AF253036	France: Grandes Rousses massif	N/A	N/A
Ae. vexans	AF253041	France, Savoie	N/A	N/A
Ae. albopictus	AF253022	LC Italy	AF324898	China, Zhjiang
Ae. albopictus	AY100671	China, Kaili	N/A	N/A
Ae. albopictus	AY100670	China, Beijing	N/A	N/A
Ae. albopictus	AY100669	China, Dafang	N/A	N/A
Ae. albopictus	AY100668	China, Zunyi	N/A	N/A
Ae. albopictus	AY100667	China, Mawei	N/A	N/A
Ae. albopictus	AY100666	China, Xinxiang	N/A	N/A
Ae. aegypti	AF425846	LC, USA, Queens University	AF425846	LC, USA, Queens University
Ae. aegypti	AF253021	LC Strain Bora Bora	N/A	N/A
Ae. aegypti	AF380835	LC Strain Moyo-R	N/A	N/A
Ae. aegypti	AF390098	LC Strain RED	N/A	N/A
Ae. aegypti	AY056597	LC Strain Formosus	N/A	N/A
Ae. aegypti	AY056596	LC Strain Liverpool	N/A	N/A
Ae. ashworthi	N/A	N/A	AF042706	Australia, Augusta, WA
Ae. ashworthi	N/A	N/A	AF042705	Australia, Augusta, WA
Ae. ashworthi	N/A	N/A	AF042704	Australia, Margaret River, WA
Ae. ashworthi	N/A	N/A	AF042703	Australia, Rottnest Island, WA
Ae. ashworthi	N/A	N/A	AF042702	Australia, Rottnest Island, WA
Ae. ashworthi	N/A	N/A	AF042701	Australia, Rottnest Island, WA
Ae. ashworthi	N/A	N/A	AF042700	Australia, Rottnest Island, WA
Ae. australis	N/A	N/A	AF042682	Australia, Sydney, NSW
Ae. australis	N/A	N/A	AF042696	Australia, Walkerville, VIC
Ae. australis	N/A	N/A	AF042695	Australia, Walkerville, VIC
Ae. australis	N/A	N/A	AF042694	Australia, Walkerville, VIC
Ae. australis	N/A	N/A	AF042693	Australia, Lowhead, TAS
Ae. australis	N/A	N/A	AF042692	Australia, Devonport, TAS
Ae. austarlis	N/A	N/A	AF042691	Australia, Marrawah, TAS
Ae. australis	N/A	N/A	AF042690	Australia, Kingston, TAS
Ae. australis	N/A	N/A	AF042689	New Zealand, Purakanui
Ae. australis	N/A	N/A	AF042688	New Zealand, Purakanui
Ae. australis	N/A	N/A	AF042687	New Zealand, Purakanui
Ae. australis	N/A	N/A	AF042686	New Zealand, Purakanui
Ae. australis	N/A	N/A	AF042685	New Zealand, Purakanui
Ae. australis	N/A	N/A	AF042684	Australia, Sydney, NSW
Ae. australis	N/A	N/A	AF042683	Australia, Sydney, NSW
Ae. wardangensis	N/A	N/A	AF042698	Australia, Wardang Island, WA
Ae. wardangensis	N/A	N/A	AF042699	Australia, Wardang Island, WA
Ae. wardangensis	N/A	N/A	AF042697	Australia, Wardang Island, WA

Taxonomy as follows: subfamily Anophelinae: *Ch. bathana*, *Bironella* (*Bironella*) *gracilis*, *An.* (*Cellia*) *gambiae*. subfamily Toxorhynchitinae: *Toxorhynchites* (*Toxorhynchites*) *amboinensis*. subfamily Culicinae: *Ae.* (*Diceromyia*) *furcifer*, *Ae.* (*Diceromyia*) *taylori*, *Ae.* (*Stegomyia*) *luteocephalus*, *Ae.* (*Stegomyia*) *aegypti*, *Ae.* (*Stegomyia*) *albopictus*, *Ae.* (*Ochlerotatus**) *caspius*, *Ae.* (*Ochlerotatus**) *geniculatus*, *Ae.* (*Ochlerotatus**) *detritus*, *Ae.* (*Ochlerotatus**) *pullatus*, *Ae.* (*Ochlerotatus**) *cantans*, *Ae.* (*Ochlerotatus**) *cataphylla*, *Ae.* (*Ochlerotatus**) *punctor*, *Ae.* (*Ochlerotatus**) *mariae*, *Ae.* (*Ochlerotatus**) *rusticus*, *Ae.* (*Ochlerotatus**) *sticticus*, *Ae.* (*Aedimorphus*) *vexans*, *Ae.* (*Halaedes**) *ashworthi*, *Ae.* (*Halaedes**) *australis*, *Ae.* (*Halaedes**) *wardangensis*, *Ae.* (*Aedes*) *cinereus*, *Cx.* (*Culex*) *decens*, *Cx.* (*Culex*) *tarsalis*. *, generic rank transferred from *Aedes* to *Ochlerotatus* genus. (A) Voucher samples sequenced for this study. Voucher samples 1–18 were field-collected in Senegal. (B) Test samples field-collected in Senegal and sequenced for this study. (C) Genbank sequences. COI, Genbank accession no. for COI sequence, COII GenBank accession no. for COII sequence, LC, laboratory colony; N/A, no sequence available; NSW, New South Wales; TAS, Tasmania; VIC, Victoria; WA, Western Australia.

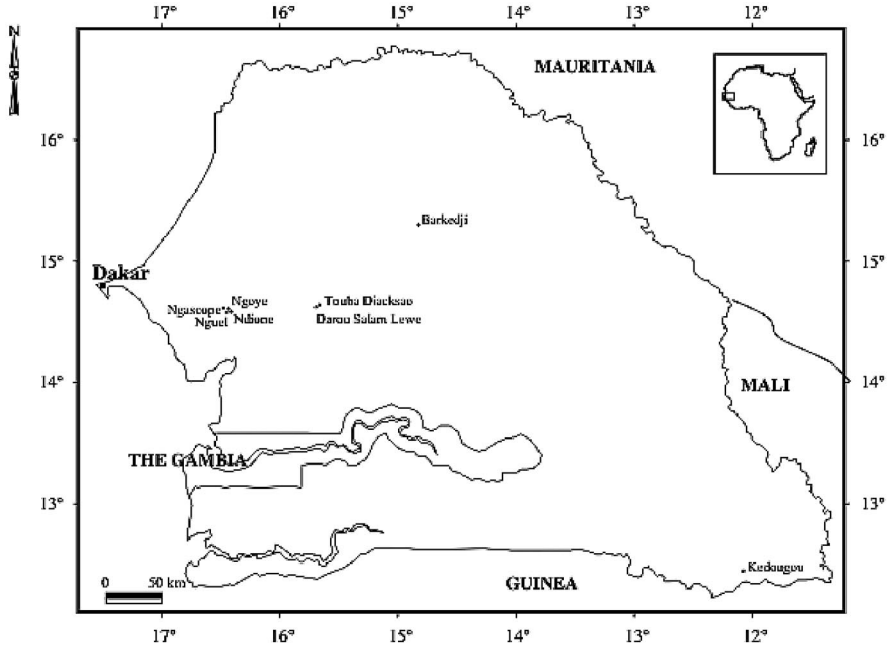

Fig. 1. Map showing sampling locations. Latitude and longitude of each site is Touba Diacksao (14° 38′01″, 15° 39′36″), Darou Salam Lewe (14° 38′01″, 15° 39′31″), Ngascope (14° 37′13″, 16° 27′45″), Nguel (14° 36′55″, 16° 28′21″), Ngoye (14° 38′11″, 16° 25′54″), Ndione (14° 36′55″, 16° 24′52″).

PCR products were visually analyzed by electrophoresis through ethidium bromide-stained 2% agarose gels under UV light. Products were purified using the Qiaquick Spin PCR purification kit (QIAGEN, Crawley, United Kingdom). Amplicons were sequenced on both strands via an automated sequencer ABI 377 (Applied Biosystems, Foster City, CA).

Sequences were deposited in GenBank with the accession numbers listed in Table 1A and B. Table 1C lists all the COI and COII sequences currently available for *Aedes* mosquitoes on GenBank, with which our derived sequences were compared.

PCR-Restriction Fragment-Length Polymorphism Work. Restriction enzyme digestions were performed in 20-μl volume. Each digest master mix consisted of the following per reaction tube: 1 μl of NE buffer 4 (50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol, pH 7.9, at 25°C, New England Biolabs, Hitchin, United Kingdom), 10,000 U of enzyme (10×, New England Biolabs), 8 μl of ddH_2O, and 10 μl of PCR product. The reaction was incubated at 37°C for 1 h. Digestion products were analyzed by electrophoresis through ethidium bromide-stained 2% agarose gels under UV light.

Data Analysis. SeqMan II (DNASTAR Inc., Madison, WI) was used to combine reverse and forward sequences from each sample, and final data sets were compiled using Se-Al (available at http://evolve.zoo.ox.ac.uk/software/Se-Al/main.html). These sequences were combined with those currently available from GenBank as detailed in Table 1C.

For both the COI and COII data sets, maximum likelihood (ML) phylogenetic trees were estimated using the GTR+Γ+I model of nucleotide substitution. Accordingly, parameter values for the general time-reversible (GTR) substitution matrix, in which the substitutions A↔C, A↔G, A↔T, C↔G, C↔T, and G↔T are allowed to occur at a different rate (Rodríguez et al. 1990), as well as the base composition, the gamma distribution of among-site rate variation (Γ) and the proportion of invariant sites (I) were estimated from the data (parameter values are available from the authors on request). To assess the robustness of particular phylogenetic groupings, a bootstrap resampling analysis was undertaken using 1000 replicate neighbor-joining (NJ) trees by using the ML substitution matrix described above. All phylogenetic analyses were performed using the PAUP* package (Swofford 2003).

Results

PCR-Restriction Fragment-Length Polymorphism Results. COI and COII sequences from all *Ae. furcifer* and *Ae. taylori* samples were examined for restriction enzyme sites by using NEBCutter version 2.0 (Vincze et al. 2003). It was predicted that the enzyme MfeI would cut the COI amplicon from *Ae. taylori* once, whereas this site was predicted to be absent from *Ae. furcifer.* PCR-restriction fragment-length polymorphism results for COI PCR products digestion with MfeI are shown in Fig. 2. Based on species-specific

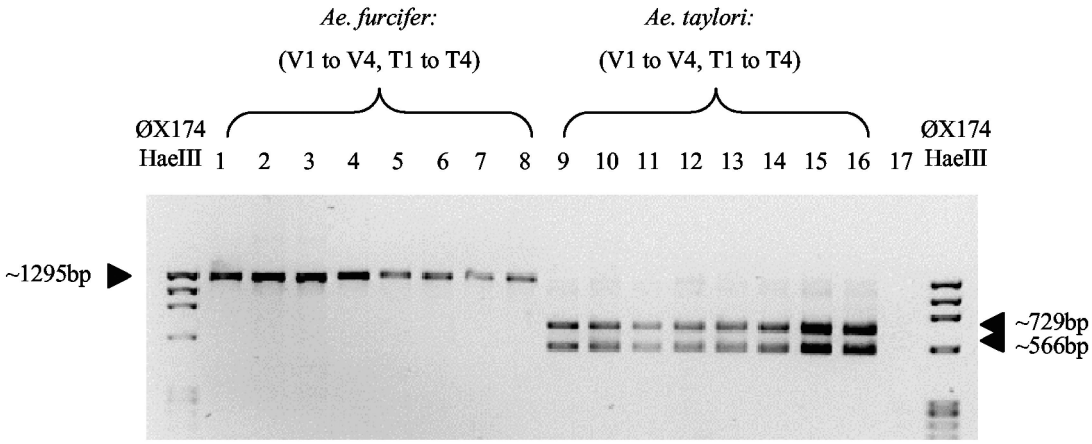

Fig. 2. Diagnostic PCR-restriction fragment-length polymorphism (MfeI Digest) for the cryptic species *Ae. furcifer* and *Ae. taylori.* (1–4) *Ae. furcifer* V1 to V4, (5–8) *Ae. furcifer* T1 to T4, (9–12) *Ae. taylori* V1 to V4, (13–16) *Ae. taylori* T1 to T4, (17) negative control. 2% agarose gel, PhiX174 Hae ladder.

sequences, PCR-restriction fragment-length polymorphism patterns were consistent with predicted results, for all voucher and sample individuals. Hence, for all 16 *Ae furcifer* mosquitoes (i.e., all voucher and test samples), COI amplicons remained intact. In contrast, in all eight COI PCR products from *Ae. taylori* mosquitoes (four voucher and four test samples), digestion occurred at the MfeI restriction site, with a characteristic double band when viewed on an agarose gel.

Phylogenetic Analysis of COI and COII Sequences. In total, COI and COII sequences were amplified from 62 Senegalese samples. Sequences of ≈1,295 bp for COI and ≈792 to ≈843 bp (TK-N-3785 or SC-TD-N-3862 as reverse primer, respectively) for COII were obtained. This covered the most of the COI gene and the entire COII gene for each specimen. No insertions or deletions were identified within the aligned sequences. In addition, the alignment of derived COI and COII sequences was unambiguous because there was no length variation in the amplified gene region

Derived COI and COII sequences were combined with those currently available on GenBank and trimmed to produce the final data sets, which were 902 and 603 bp in length, respectively. These correspond to positions 2040–2942 and 3054–3657 of *Anopheles quadrimaculatus* Say (GenBank accession no. NC_000875). Within the COI data set, 353 positions were variable, whereas for COII, 267 sites were variable. In common with other insect mitochondrial genes, both the COI and COII data sets exhibited a relatively high proportion of A+T nucleotides (30 and 40% of total nucleotide composition for COI and 35 and 40% for COII, respectively). The sequences also show a typical pattern of variability, with third codon positions containing the highest number of variable sites (85% in each gene). Maximum uncorrected interspecific pairwise distances for COI were 17.7% between *An. gambiae* and *Ae. furcifer* V6. In the case of COII, *Ch. bathana* and *Ae. taylori* T1 differed by 20.1%. Intraspecific uncorrected pairwise distances were <2% in the majority of cases and were zero in a number of cases, including all the *Ae. vexans* sequences.

The maximum likelihood trees for COI and COII are shown in Figs. 3 and 4, respectively. In the case of the COI tree, each clade in those species for which more than one sequence was available contained a single *Aedes* species. Furthermore, all these clades were supported by high (78–100%) bootstrap values, with the exceptions of the *Ae. punctor* Kirby and *Ae. caspius* Pallas groups. Bootstrap values for *Ae. taylori* and *Ae. furcifer* were notably high, at 100% for both groupings. For COII, the *Aedes* genus again contained one clade per species tested, generally with higher bootstrap values than COI (82–100%) with the exception of a lineage represented by a single *Aedes albopictus* Skuse specimen (53%). In several cases, sequences for a given species originate from wild mosquito populations or laboratory colonies from a range of countries, for example *Ae. aegypti* COI and COII, *Ae. vexans* COI. Significantly, despite being drawn from widely dispersed geographic locations, these lineages still group strongly according to species. In contrast, phylogenetic structure at the subfamily level for both genes was either absent or of limited support, reflected by bootstrap values of <50%. This aspect is dealt with in a separate study of the entire Culicidae family (S.C. et al., unpublished data)

Validation and Accuracy of Molecular Identification System. The phylogenetic trees shown in Figs. 3 and 4 comprise analyses using all sequences obtained throughout the study. However, an equivalent tree topology was obtained during the preliminary stage of the study based on the 20 voucher specimens alone. The second stage of the study comprised the sequencing of the 42 blind "test" samples, and their insertion onto this backbone phylogeny. Test samples were then assigned to species according to which group

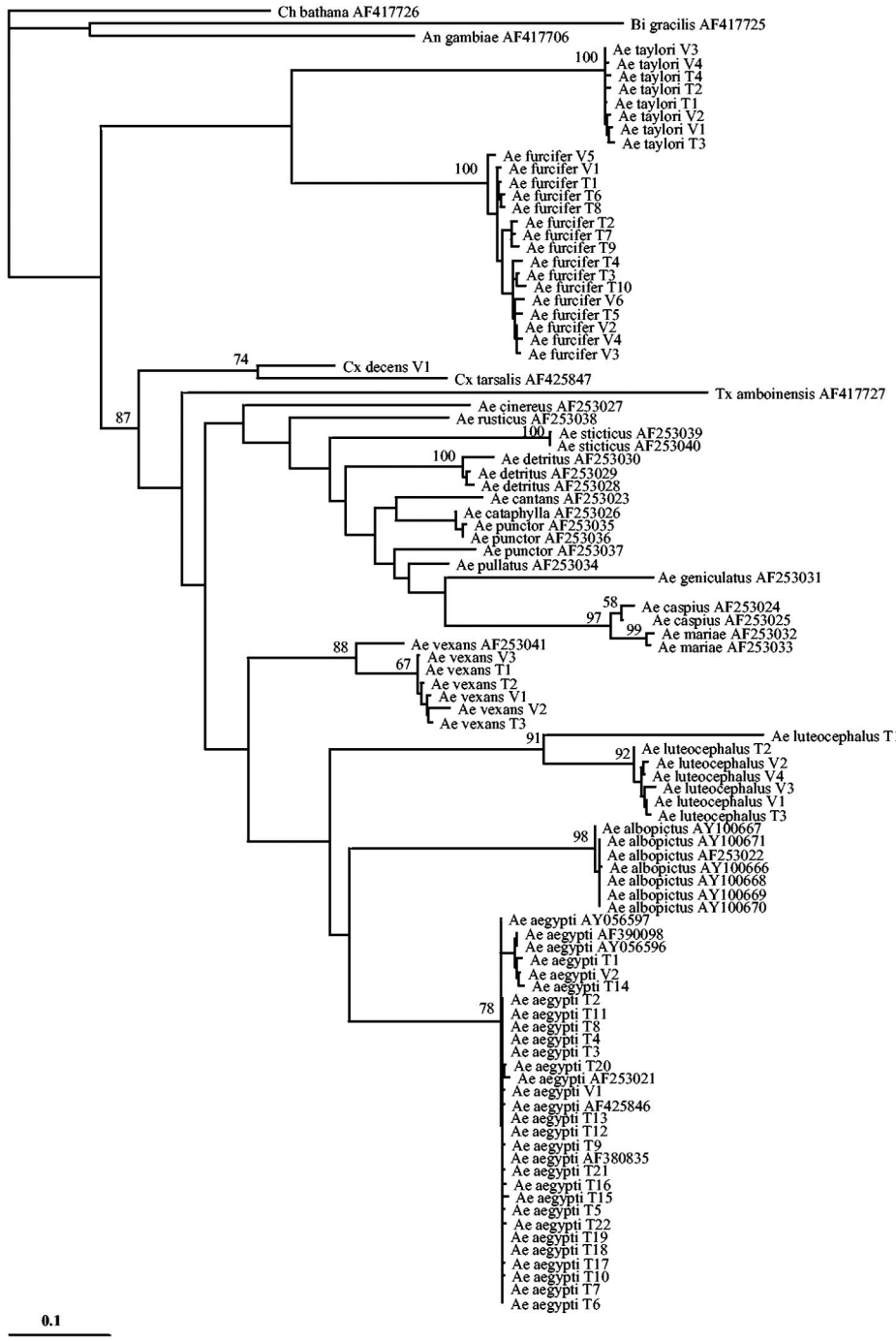

Fig. 3. Maximum likelihood phylogenetic tree of COI sequences among the Culicidae under the GTR+Γ+I model of nucleotide substitution. Rooted using *Ch. bathana*. Bootstrap proportions >50% are shown.

they were most closely related to. All test samples were identified correctly according to prior morphological analysis in Senegal. Hence, the system proved 100% in agreement with morphological identification.

Discussion

Mitochondrial genes have been a popular source of data for numerous insect molecular phylogeny studies

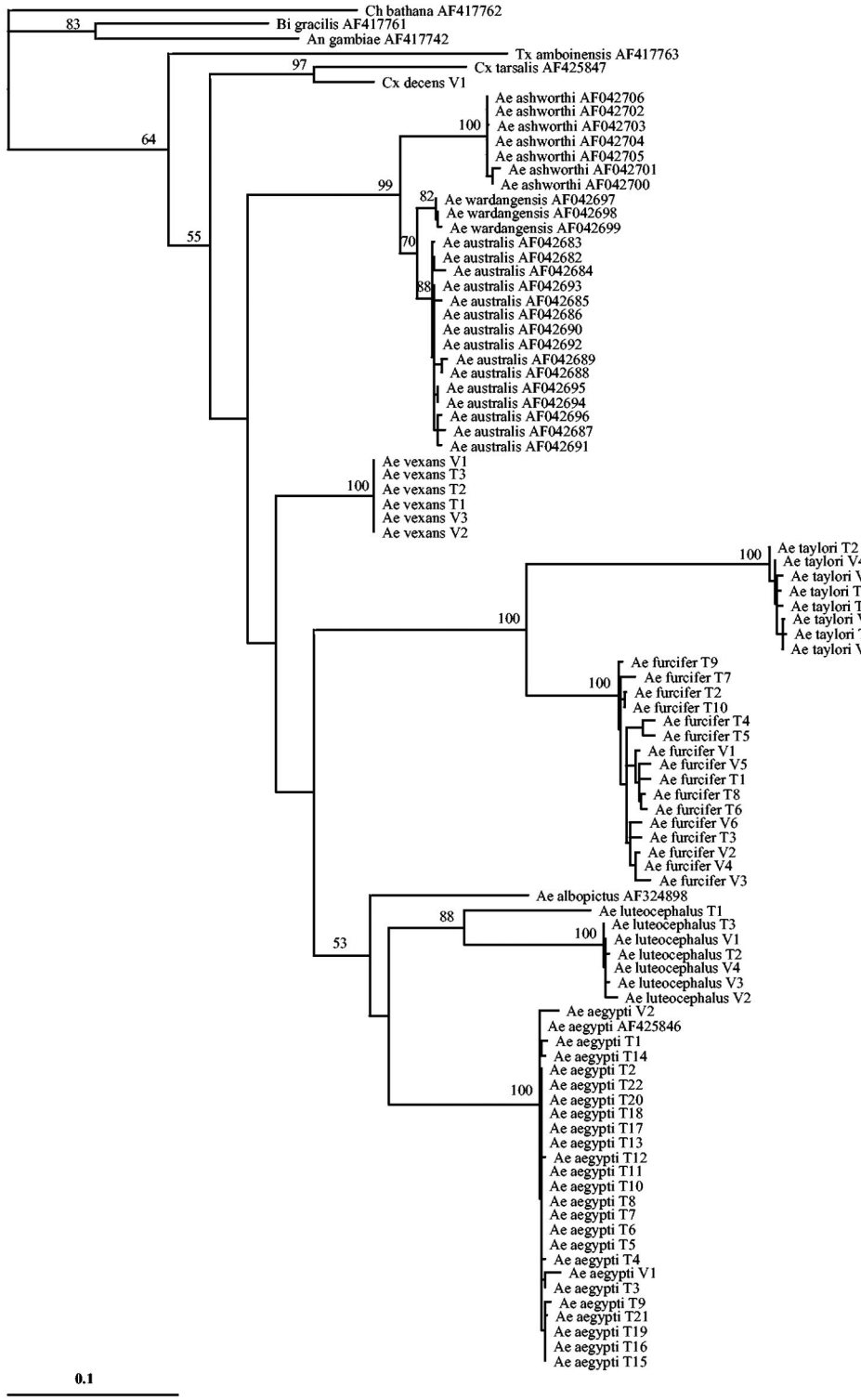

Fig. 4. Maximum likelihood phylogenetic tree of COII sequences among the Culicidae under the GTR+Γ+I model of nucleotide substitution. Rooted using *Ch. bathana*. Bootstrap proportions >50% are shown.

(Avise et al. 1987, Harrison 1989, Simon 1991, Liu and Bechenbach 1992) and more specifically for the Culicidae (Mitchell et al. 2002). Critically, mitochondrial genes are present in multiple copies and so are easier to amplify than single-copy nuclear genes from a given extract. This is a particularly important consideration given that flaviviral assay work may produce only trace amounts of DNA, coextracted with RNA from mosquito heads (the rest of the insect body being placed into cold storage in case of a requirement for later virus isolation work). In addition, mitochondrial genes contain no introns and are maternally inherited, thus recombination, paralogy, and heterozygosity are less of a problem for phylogenetic analysis (Kryswinski and Besansky 2003). Finally, mitochondrial genes evolve faster than protein-coding nuclear genes in insects (DeSalle et al. 1987, Simon et al. 1994) and hence seem particularly well suited to the study of closely related taxa such as the *Aedes* species, as opposed to the single-copy nuclear genes that have been used for the study of subfamily relationships in the Culicidae (Besansky and Fahey 1997, Kryswinski et al. 2001).

Overall, the results presented here reveal a pattern of low intraspecific variation and high interspecific variation for both COI and COII in the *Aedes* species examined. As such, both genes provide a strong basis for the species-level identification of common flaviviral vectors in Senegal. Both COI and COII generate results in agreement with each other with respect to species-level groupings and support for the five main *Aedes* lineages found in Senegal was extremely high, even in those cases where samples from widely dispersed locations were included (i.e., *Ae. vexans* and *Ae. aegypti*). This confirms the robust nature of COI and COII in distinguishing between these species. However, because the availability of sequences for COI and COII differs with respect to species, it is difficult to draw firm conclusions regarding the interspecies relationships of the culicine mosquitoes from this study alone. Given that support for the COII tree is slightly higher in general, and that the COI phylogeny suggests a potential but unlikely paraphyly of *Aedes* with respect to *Culex*, future species-level studies may best be conducted using the COII gene. Divergences at the subfamily or genus level tended to have low bootstrap values. This may reflect a rapid diversification of lineages or the high rate of evolution in insect mitochondrial DNA (mtDNA) such that there are large numbers of multiple substitutions per site (Lewis et al. 1995, Bernasconi et al. 2000). As such, these results corroborate earlier studies in suggesting that a nuclear gene may be more suitable for estimates of deeper divergences (Besansky and Fahey 1997).

Molecular divergence among cryptic species is a common occurrence within the Culicidae, albeit generally at the subspecies level, and rDNA-based PCR assays have been developed based on molecular divergence between many of the world's *Anopheles* malaria vectors (Collins and Paskewitz 1996). Similarly, this study has shown that two cryptic members of the culicine subfamily, *Ae. taylori* and *Ae. furcifer*, are exceptionally divergent at the molecular level, with a degree of interspecific variation comparable with that seen between the other *Aedes* species present in the region. The COI PCR-restriction fragment-length polymorphism system developed here provides a fast and easy method to accurately distinguish these species without the need for amplicon sequencing, potentially allowing a high throughput sample analysis. Nucleotide sequences at the MfeI digest site seem to be highly conserved within species. Together with the low intra- versus interspecies diversity in the *Aedes* COI gene in general, this approach seems to be a useful and reliable system.

Because the aim of the study was to assess the suitability of mitochondrial genes for species-level identification, the current data set provides a useful reference as a backbone for further work. With respect to future development, species other than the five addressed here are likely to be collected for disease research in Senegal. Due to the absence of an appropriate reference sequence, it is conceivable that a novel species' sequence may group with the most closely matched backbone sample. However, as demonstrated by the *Ae. aegypti* group, even when sequences from a given species are drawn from locations dispersed on a worldwide scale, there is a relatively low level of intraspecific variation and COI and COII sequences still group together with strong bootstrap support. This substantiates the high utility of mtDNA for interspecific distinction.

More generally, the mosquito identification system developed here may be very useful not only for attempts to identify disease vectors postviral assay but also in cases where an entomologist experienced in the use of morphological keys for the identification of mosquitoes is not available. Thus, mosquitoes that generate a positive viral assay result could be identified retrospectively using small amounts of the original RNA or DNA extract. In addition, further population-level studies may allow the identification of subspecies that differ according to their role in disease transmission, yielding novel information regarding genetic variation of the vector species and potential differences in the vector capacity of cryptic complexes or subspecies. Finally, because the system can be applied to all life stages of the Culicidae it should prove useful for a wide range of studies involving the Culicidae beyond flaviviral research, for example, insect ecology and control strategies.

Acknowledgments

We thank Morten Andreassen and Ernie Gould from the University of Oxford and the Centre of Ecology and Hydrology, Oxford, respectively, for providing the *Ae. aegypti* vouchers from the United Kingdom and United States. We also thank Ibrahima Dia and Yamar Ba for help in collecting mosquitoes in the field and the global positioning system data. This work was funded by the Wellcome Trust.

References Cited

Avise, J. A., J. Arnold, R. M. Ball, E. Mermingham, J. E. Neigei, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489–522.

Bernasconi, M. V., C. Valsangiacomo, J. C. Piffaretti, and P. I. Ward. 2000. Phylogenetic relationships among Muscoidea (Diptera: Calyptratae) based on mitochondrial DNA sequences. Insect Mol. Biol. 9: 67–74.

Besansky, N. J., and G. T. Fahey. 1997. Utility of the white gene in estimating phylogenetic relationships among mosquitoes (Diptera: Culicidae). Mol. Biol. Evol. 14: 442–454.

Collins, F. H., and S. M. Paskewitz. 1996. A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species. Insect Mol. Biol. 5: 1–9.

DeSalle, R., T. Freeman, E. M. Prager, and A. C. Wilson. 1987. Tempo and mode of sequence evolution in mitochodrial DNA of Hawaiian *Drosophila.* J. Mol. Evol. 26: 157–164.

Diallo, M., Y. Ba, A. A. Sall, O. M. Diop, J. A. Ndione, M. Mondo, L. Girault, and C. Mathiot. 2003. Amplification of the sylvatic cycle of dengue virus type 2, Senegal, 1999–2000: entomologic findings and epidemiologic considerations. Emerg. Infect. Dis. 9: 362–366.

Edwards, F. W. 1941. Mosquitoes of the Ethiopian region III. Culicine adults and pupae. British Museum, London, United Kingdom.

Fontenille, D., M. Traore-Lamizana, M. Diallo, J. Thonnon, J. P. Digoutte, and H. G. Zeller. 1998. New vectors of Rift Valley Fever in West Africa. Emerg. Infect. Dis. 4: 289–293.

Gale, K. R., and J. M. Crampton. 1987. DNA probes for species identification of mosquitoes in the *Anopheles gambiae* complex. Med. Vet. Entomol. 1: 127–136.

Guzman, M. G., and G. Kouri. 2002. Dengue: an update. Lancet Infect Dis. 2: 33–42.

Harrison, C. J. 1989. Animal mtDNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4: 6–11.

Horsfall, W. R. 1972. Mosquitoes: their bionomics and relation to disease. Hafner, New York.

Inoue. 2003. Distribution of three arbovirus antibodies among monkeys (*Macaca fascicularis*) in the Philippines. J. Med. Primatol. 32: 89–94.

Kryswinski, J., and N. J. Besansky. 2003. Molecular systematics of *Anopheles:* from subgenera to subpopulations. Annu. Rev. Entomol. 48: 111–139.

Kryswinski, J., R. C. Wilkerson, and N. J. Besansky. 2001. Toward understanding Anophelinae (Diptera: Culicidae) phylogeny: Insights from nuclear single copy genes and the weight of evidence. Syst. Biol. 50: 540–556.

Lewis, D. L., C. L. Farr, and L. S. Kaguni. 1995. *Drosophila melanogaster* mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. Insect Mol. Biol. 4: 263–278.

Liu, H., and A. T. Bechenbach. 1992. Evolution of the mitochondrial cytochrome oxidase II gene among ten orders of insects. Mol. Phylogenet. Evol. 1: 41–52.

Mitchell, A., F.A.H. Sperling, and D. Hickey. 2002. Higher-level phylogeny of mosquitoes (Diptera: Culicidae): mtDNA data support a derived placement for *Toxorhynchites.* Insect Syst. Evol. 33: 163–174.

Monath, T. M. 2001. Yellow fever: an update. Lancet Infect. Dis. 1: 11–20.

Monath, T. P., and F. X. Heinz. 1996. Flaviviruses, pp. 961–1034. *In* B. N. Fields [ed.], Virology, vol. 1. Lipincott-Raven, New York.

Munstermann, L. E. 1995. Mosquito systematics: current status, new trends, associated complications. J Vector Ecol. 20: 129–138.

Paskewitz, S. M., and F. H. Collins. 1990. Use of the polymerase chain reaction to identify mosquito species of the *Anopheles gambiae* complex. Med. Vet. Entomol. 4: 367–373.

Paskewitz, S. M., K. Ng, M. Coetzee, and R. H. Hunt. 1993. Evaluation of the polymerase chain-reaction method for identifying members of the *Anopheles gambiae* (Diptera, Culicidae) complex in southern Africa. J. Med. Entomol. 30: 953–957.

Reno, H. A., M. H. Vodkin, and R. J. Novak. 2000. Differentiation of *Aedes triseriatus* (Say) from *Aedes hendersoni* Cockerell (Diptera: Culicidae) by restriction fragment length polymorphisms of amplified ribosomal DNA. Am. J. Trop. Med. Hyg. 62: 193–199.

Rodríguez, F., J. L. Oliver, A. Marín, and J. R. Medina. 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142: 485–501.

Simon, C. 1991. Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA, pp. 33–71. *In* G. M. Hewitt, A.W.B. Johnston, and J.P.W. Young [eds.], Molecular techniques in taxonomy. Springer, Berlin, Germany.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87: 651–701.

Swofford, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer, Sunderland, MA.

Vincze, T., J. Posfai, and R. J. Roberts. 2003. NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Res. 31: 3688–3691.

White, G. B. 1975. Notes on a catalogue of Culicidae of the Ethiopian region. Mosq. Syst. 7: 303–344.

Received 13 July 2004; accepted 18 September 2004.