Speciation and phylogeography of the Southeast Asian *Anopheles sundaicus* complex

Isabelle Dusfour [a,1], Johan R. Michaux [a,b], Ralph E. Harbach [c], Sylvie Manguin [a,*]

[a] *Institut de Recherche pour le Développement (IRD), UMR22 Biologie et Gestion des Populations (CBGP), Campus de Baillarguet, CS30016, Montferrier sur Lez 34988, France*
[b] *Zoogeography Research Unit, Institute of Zoology, Quai Van Beneden, 22, 4020 Liège, Belgium*
[c] *Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*

Abstract

Anopheles sundaicus s.l. is a malaria vector in coastal areas of Southeast Asia. Previous studies showed at least four distinct species within the complex. The present study investigated the phylogeography and the status of *A. sundaicus* s.l. populations from Cambodia, Thailand, Malaysia and Indonesia with regard to *A. sundaicus* s.s. from Sarawak, Malaysian Borneo and *A. epiroticus* in Vietnam and Thailand. Three lineages recovered by analyses of Cyt-b and COI (mtDNA) confirmed the presence of *A. sundaicus* s.s. in Malaysian Borneo, the distribution of *A. epiroticus* from southern Vietnam to peninsular Malaysia, and recognised a distinct form in Indonesia that is named *A. sundaicus* E. The phylogenetic and demographic analyses suggest that the three species were separated during the Early Pleistocene (1.8–0.78 Myr) and experienced bottlenecks followed by a genetic expansion in more recent times. Based on the results and knowledge of the biogeography of the area, we hypothesise that the combination of cyclical island and refugium creation was the cause of lineage isolation and bottleneck events during the Pleistocene.
© 2007 Elsevier B.V. All rights reserved.

Keywords: *Anopheles sundaicus* s.l.; Speciation; Phylogeography; Southeast Asia; Mitochondrial and nuclear markers

1. Introduction

Elucidating *Anopheles* species complexes is a prerequisite for determining which genetic units of the complex are vectors of malaria, and unravelling the ecological and ethological differences that are relevant to disease transmission. Such knowledge will improve the understanding of malaria epidemiology and will help in implementing appropriate vector control strategies. Among *Anopheles* species in Southeast Asia, *A. sundaicus* s.l. is a principal malaria vector taxon along coastal areas of the mainland and islands. Its reported coastal distribution ranges from northern India to southern Vietnam (up to the 11th parallel) and the islands of Nicobar, Andaman, Borneo, Java, Sumatra and Sulawesi (Fig. 1).

Anopheles sundaicus s.l. was suspected as early as Reid (1970) to be a species complex due to differences in behaviour and ecology, notably with immature stages occurring in brackish water or freshwater. Subsequently, approaches in genetics using different markers established the presence of a species complex in distinct geographic areas of Southeast Asia. The presence of three genetic forms (A, B and C) in Indonesia (Sumatra and Java) and in Thailand was demonstrated by cytogenetics, and confirmed by isozyme analysis (Sukowati et al., 1996, 1999) (Table 1). Form A was found in all study sites in both countries. Form B was strongly linked to freshwater, and occurred only in northern Sumatra and central Java. Form C was restricted to Asahan in northern Sumatra. Asahan was the only site where all three cytotypes occurred in sympatry (Sukowati et al., 1996, 1999). In addition, molecular approaches identified two allopatric species based on two mitochondrial markers, cytochrome oxidase I (COI) and cytochrome *b* (Cyt-b) (Dusfour et al., 2004b) (Table 1). One of these species is *A. sundaicus* s.s. from Sarawak, Malaysian Borneo, and the other is *A. epiroticus*, which occurs in continental Southeast Asia, particularly in Cambodia, peninsular Malaysia, Thailand, and Vietnam (Linton

* Corresponding author. Tel.: +33 499 623 327; fax: +33 499 623 345.
E-mail address: manguin@mpl.ird.fr (S. Manguin).
[1] Present address: Uniformed Services University of the Health Sciences (USUHS), Dept. PMB, 4301 Jones Bride Rd, Bethesda, MD 20814, USA.

Fig. 1. Map of Southeast Asia showing the location of the populations of *A. sundaicus* s.l. studied (black dots) and the known distribution of the species complex (in grey).

Table 1
Summary of *A. sundaicus* populations (population codes (PC), country and locality of origin) analysed in this paper, their species/form status and the associated references according to genetic markers

PC	Countries	Localities	Identification according to markers			
			Cytogenetics	Isozymes	COI + Cyt-b	ITS2
KKK	Cambodia	Kaoh Kong	–	–	*A. epiroticus* [4, 8]	–
VBL[a]	Vietnam	Bac Lieu	–	–	*A. epiroticus* [4, 6, 8]	*A. epiroticus* [6, 8]
VHCM	Vietnam	Ho Chi Minh	–	–	*A. epiroticus* [4, 6, 8]	*A. epiroticus* [6, 8]
TTR	Thailand	Trat	*A. sundaicus* form A [1]	–	*A. epiroticus* [4, 6, 8]	*A. epiroticus* [6, 8]
TPG	Thailand	Phangnga	*A. sundaicus* form A [1]	–	*A. epiroticus* [4, 6, 8]	*A. epiroticus* [6, 8]
TPA	Thailand	Pattani	–	–	*A. epiroticus* [4, 8]	–
MAT	Malaysia, Peninsula	Terengganu	–	–	*A. epiroticus* [4, 8]	–
MAL[b]	Malaysia, Borneo	Sarawak, Lundu	–	–	*A. sundaicus* s.s. [3, 4, 6, 8]	*A. sundaicus* s.s. [3, 6, 8]
MAM	Malaysia, Borneo	Sarawak, Miri	–	–	*A. sundaicus* s.s. [4, 6, 8]	*A. sundaicus* s.s. [6, 8]
MAG	Malaysia, Borneo	Sabah, Pulau Gaya	–	–	[8]	–
INA	Indonesia, Sumatra	Asahan	*A. sundaicus* forms A, B, C [1]	*A. sundaicus* forms A, B, C [2]	[8]	[8]
INS	Indonesia, Sumatra	South Tapanuli	*A. sundaicus* forms A, B [1]	*A. sundaicus* forms B [2]	[8]	[8]
INL	Indonesia, Sumatra	Lampung	*A. sundaicus* form A [1]	*A. sundaicus* form A [2]	[8]	[8]
INP	Indonesia, Java	Patuk	*A. sundaicus* form A, B [1]	*A. sundaicus* form A, B [2]	[8]	[8]
INJ	Indonesia, Java	Jakarta	–	–	[8]	[8]
INI	India	Nicobar Islands	*A. sundaicus* form D [5]	–	–	*A. sundaicus* species D [7]

A. sundaicus species D, not included herein, collected on Nicobar Island is also mentioned. References are [1] Sukowati et al. (1996), [2] Sukowati et al. (1999), [3] Linton et al. (2001), [4] Dusfour et al. (2004b), [5] Nanda et al. (2004), [6] Linton et al. (2005), [7] Alam et al. (2006) and [8] this study.

[a] Type locality of *A. epiroticus*.
[b] Type locality of *A. sundaicus* s.s.

et al., 2005). Both species were described based on morphological characters and molecular sequences of the Internal Transcribed Spacer 2 (ITS2) of rDNA and COI (Linton et al., 2001, 2005). Additionally, Cyt-b was used in the description of *A. epiroticus*. Recently, *A. sundaicus* species D from Nicobar Island, India was identified and studied using cytogenetics (Nanda et al., 2004) and ITS2 sequence (Alam et al., 2006) (Fig. 1, Table 1). Moreover, Dusfour et al. (2004b) hypothesised that the divergence and speciation of *A. sundaicus* s.s. and *A. epiroticus* might be due to vicariance events associated with an increase of sea level between the continent and Borneo during the Quaternary. However, populations throughout the distribution of the *A. sundaicus* complex have never been studied using the same markers. Hence, the taxonomy and evolutionary history of the complex have not been fully investigated.

The aim of the present work was to analyse the *A. sundaicus* complex on a wide scale using three molecular markers (COI, Cyt-b and ITS2) to assess its evolutionary history and the status of the individual species. It was studied over a wider geographic area than Dusfour et al. (2004b). Indeed, populations from Cambodia, southern Thailand, peninsular Malaysia, Sabah (Malaysia) and Indonesia (Sumatra and Java) were investigated in addition to those from Vietnam, Thailand and Sarawak that were analysed previously.

2. Materials and methods

2.1. Mosquitoes

Specimens used in this study were first identified as *A. sundaicus* s.l. using the morphological keys of Reid (1968), and the keys for the *Anopheles* of Vietnam prepared by the National Institute of Malariology, Parasitology and Entomology, Hanoi, Vietnam (IMPE, 1987). Specimens were collected from 15 localities, including 1 in Cambodia, 5 in Indonesia, 4 in Malaysia, 3 in Thailand and 2 in Vietnam (Fig. 1, Tables 1 and 2). Larval habitats in Miri (MAM, Malaysian Borneo) and South Tapanuli (INS, Sumatra) were the only two characterised as freshwater (Table 2, in bold) (Sukowati et al., 1996; Seng et al., 2001). Specimens previously identified by Dusfour et al. (2004b) and Linton et al. (2005) as *A. sundaicus* s.s. and *A. epiroticus* based on Cyt-b and COI sequences were used to unambiguously assess the identity of new populations and infer the phylogeography of the complex. Mosquitoes chromosomally identified as forms B and C (1 specimen each) from Asahan (Sumatra), kindly provided by S. Sukowati, were also sequenced. Sequences of *A. minimus* A from Hoa Binh Province, northern Vietnam, and *A. gambiae* obtained from GenBank (Beard et al., 1993) were selected as outgroup taxa. *A. gambiae* was chosen because the *A. gambiae* and *A. sundaicus* complexes are closely related taxa of the Pyretophorus Series of subgenus *Cellia*. A second species, *A. minimus* A (=*A. minimus* s.s., see Harbach et al., 2006) of the Myzomyia Series of subgenus *Cellia* was chosen because it is phylogenetically more distant from the *A. sundaicus* complex. These species were also selected as outgroup taxa because they gave the best resolution in comparison with other species that were tested.

Table 2
Population codes (PC), year of collection (*Y*), types of larval habitat (LH, B: brackish water, F: freshwater) and number of specimens of the *A. sundaicus* complex analysed for COI and Cyt-b individually and in combination, as well as for the ITS2 marker used in this study

PC	*Y*	LH	COI	Cyt-b	COI + Cyt-b	ITS2
KKK	2003	B	5	5	5	0
VBL	1999	B	7	7	6	4
VHCM	1999	B	7	7	7	2
TTR	1999	B	6	6	6	3
TPG	2000	B	5	5	5	1
TPA	2003	B	8	8	8	0
MAT	2003	B	4	6	4	0
MAL	1999	B	7	7	7	3
MAM	1999	**F**	6	6	6	3
MAG	2003	B	6	4	4	0
INA	2001	B	6	7	6	2
INS	2001	**F**	6	7	6	3
INL	2001	B	7	7	7	2
INJ	2001	B	1	1	1	1
INP	2001	B	5	5	5	2
Total			86	88	83	26

2.2. DNA extraction, amplification and sequencing

DNA was extracted from individual mosquitoes using to the procedure of Linton et al. (2001). The COI and Cyt-b partial genes were amplified according to Linton et al. (2001) and Dusfour et al. (2004b). PCR products were purified (NucleoSpin extract 2 in 1, Machery-Nagel, Düren, Germany) and directly sequenced using the established primers for these two genes. Multiple sequence alignments were achieved using BioEdit (Hall, 1999) with Clustal W 1.4 (Thompson et al., 1994). ITS2 was amplified with universal primers published in Beebe and Saul (1995). To detect intra-individual variability, each PCR product was cloned with the pGem-T easy vector kit (Promega, Madison, USA). Two to three specimens per population and two to five purified clones (QIAprep, Qiagen, Valencia, California) per individual were sequenced with T7 and Sp6 standard primers complementary of the plasmid sequence. A total of 93 sequences from 26 individuals of 11 populations throughout the distribution of the *A. sundaicus* complex were compared.

The sequences obtained during the study are available in the GenBank under the following accession numbers: AY243788–99, AY245283, AY245284, AY256954–57, AY253150–55, AY299094–120, AY299339–46 (sequences from Dusfour et al., 2004b), AY672287–404 for new haplotypes, and AY768540–43 for ITS2 variants.

2.3. Phylogenetic relationships

All phylogenetic analyses were performed for both separate and combined mtDNA markers (COI, Cyt-b), as well as for ITS2 sequences. Incongruence of the combined mtDNA markers was tested with the incongruence length difference test (ILD test) (Farris et al., 1994) within PAUP4.0b10 (Swofford, 2002). Trees were constructed unrooted as well as rooted with both *A. minimus* species A and *A. gambiae* as outgroup taxa for mtDNA, and *A. gambiae* for ITS2.

Maximum parsimony (MP) analyses were conducted within DNAMP of the PHYLIP package (Felsenstein, 1994) and maximum likelihood (ML) reconstruction within PHYML v2.1b1 (Guindon and Gascuel, 2003). The optimal model of molecular evolution for likelihood analysis was determined in Modeltest (Posada and Crandall, 1998) using PAUP4.0b10 (Swofford, 2002). Robustness was tested by bootstrap method (1000 times) within Seqboot from the Phylip package.

2.4. Estimation of time of divergence

Evolution rate estimations currently applied for *Anopheles* mtDNA are either 2% per million years calculated for the genus *Drosophila* (De Salle et al., 1987) or 2.3% for mtDNA of arthropods (Brower, 1994). However, differences between rates have been observed among mitochondrial markers (Fleischer et al., 1998; Simmons and Weller, 2001). Therefore, we investigated the specific evolution rate of COI and Cyt-b separately. The evolution rate of COI was estimated among arthropods from known data (Knowlton et al., 1993; Brower, 1994; Tamura et al., 2004). The time of divergence among *A. sundaicus* s.l. clades was estimated based on the evolution rate of COI, and the time of divergence was subsequently used to calculate the evolution rate of Cyt-b, as follows. Firstly [1], the evolution rate of COI was estimated based on published divergence times and genetic divergences among some arthropod species (Knowlton et al., 1993; Brower, 1994; Tamura et al., 2004). The slope obtained by linear regression analysis (Xlstat v7.1, Addinsoft) between genetic divergence and time of divergence, and crossing axis origins, gave the estimation of rate. The analysis was supported by the coefficient of correlation (R) and a P value. Secondly [2], times of divergence among *A. sundaicus* s.l. clades were estimated based on the genetic divergence of COI calculated under the Tamura-Nei (93) model (Gamma = 0.82) within MEGA 2.1 (Kumar et al., 2001) and the evolution rate calculated in [1]. Thirdly [3], the evolution rate of Cyt-b within *A. sundaicus* s.l. was estimated based on the divergence times among *A. sundaicus* clades calculated in [2] and the associated Cyt-b genetic divergences. As for arthropod COI analysis, a linear regression (Xlstat v7.1, Addinsoft) between clade divergence times [2] and Cyt-b genetic divergences, calculated under Tamura-Nei model (MEGA 2.1), produced a slope estimating the Cyt-b evolution rate.

For this approach, the constancy of evolutionary rates between pairs of clades were subjected to the relative rate test within RRTree version 1.1 (Robinson-Rechavi and Huchon, 2000) to allow divergence times and evolution rate calculations. The program RRTree compared substitution rates between DNA sequences grouped in the phylogenetically defined lineages. The null hypothesis tested is the heterogeneity among the groups. If $P > 0.05$, the two lineages evolve at the same rate. If $P < 0.05$, the two lineages evolve at different rates.

2.5. ITS2 analysis

Polymorphic sites (S), parsimony informative mutations (P) and the mean average of nucleotide differences were evaluated using DNAsp 4.0 (Rozas et al., 2003). The number of different ITS2 sequences (ribotypes) within each specimen and overall clone set were counted.

2.6. Demographic trends

Inferences about population sizes and demographic tendencies can be obtained from genealogies through a number of different approaches. Basic diversity estimates were calculated from mitochondrial markers for each lineage. Evaluation of haplotype number and estimation of haplotype diversity (h), as well as nucleotide diversity (p) associated with their respective standard divergence (S.D.), were performed using DNAsp 4.0 (Rozas et al., 2003). The pattern of haplotype diversity (h) and nucleotide diversity (p) combinations are, according to Avise (2000), indicators of population demographic events. The mismatch distribution analysis was also used to investigate indications for demographic expansions (Slatkin, 1991; Rogers and Harpending, 1992) using DNAsp 4.0 (Rozas et al., 2003). The shape of a mismatch distribution is affected by historical demographic changes of the major lineages. If a population has undergone rapid expansion, an unimodal mismatch distribution approximating a Poisson curve is expected (Rogers and Harpending, 1992), whereas populations approaching mutation drift equilibrium are expected to produce a multimodal or 'ragged' mismatch distribution. We examined the 'raggedness' statistic (Rg) and the sum of squared deviation (SSD) to assess the statistical significance of the distribution within ARLEQUIN 2.000 (Schneider et al., 2000) under the rapid expansion model. Expansion parameters Theta initial and Tau, the date of the growth in units of mutational time were produced by the analysis. Furthermore, we calculated Fu's Fs statistic (Fu, 1997) and Fu and Li's (1993) D^* and F^* statistics as indicators of demographic expansion. The Fs statistic uses information from the haplotype distribution to test population growth. This parameter is a powerful statistic for detecting population growth (Fu, 1997). The D^* and F^* statistics were calculated for comparison with Fs. The effects of background selection can be distinguished from population growth or range expansion by examining the pattern of significance between Fs, F^* and D^* (Fu, 1997). If Fs is significant and F^* and D^* are not, then population growth or range expansion is indicated, whereas the reverse suggests selection (Fu, 1997). All these values were calculated using DNAsp 4.0 (Rozas et al., 2003).

Finally, networks were constructed using the minimum spanning method (ARLEQUIN 2.0 of Schneider et al., 2000), statistical parsimony (TCS of Clement et al., 2000) and median joining network, NETWORK 4.000 (available at www.fluxus-engineering.com) (Bandelt et al., 1999), for separate COI and Cyt-b markers.

3. Results

3.1. Phylogenetic relationships

The best fit model for ITS2 sequences was HKY. However, no phylogenetic relationships were resolved. ITS2 was not

suitable for evaluating phylogenetic relationships within *A. sundaicus* s.l. (data not shown).

The best fit model for COI, including outgroup taxa, was defined as GTR + G + I (Gamma = 0.75; P-inv = 0.49) and for Cyt-b by TIM + G (Gamma = 0.09). The best fit for COI without outgroup was defined by TrN + I + G (Gamma = 0.82; P-inv = 0.68) and for Cyt-b by TrN + I (P-inv = 0.87). Phylogenetic reconstructions of both markers, rooted or unrooted, showed three clades (data not shown). However, bootstrap supports were less than 50% except for the clade that included the Malaysian Borneo populations.

The K81uf + G + I model (Gamma = 0.78; P-inv = 0.62) was introduced into PHYML as the best fit model for analysing combined mtDNA markers by maximum likelihood (ML), using *A. minimus* species A and *A. gambiae* as outgroup taxa. The combination of COI and Cyt-b in the ML and maximum parsimony (MP) algorithms increased the robustness of the three clades, which became greater than 50% (Fig. 2a), even if the ILD test revealed incongruence (*P* = 0.003). The Borneo clade was the more robust overall, with bootstrap values (BSV) of 93.7 and 85 for ML and MP, respectively (Fig. 2a). The clade consisting of Indonesian (Sumatra and Java) populations was well supported, with BSV from 78 to 76 (Fig. 2a). Finally, The Continental clade had the weakest support, with BSV of 60.6 and 54 (Fig. 2a). Whatever the reconstruction, the basal node remains unresolved. However, mtDNA unrooted ML trees, reconstructed under the HKY + G + I model (Gamma = 0.78; P-inv = 0.73), strongly supported the three clades with higher BSV than rooted trees. Indeed, ML reconstruction exhibited BSV ranging from 86 to 99 and MP from 82 to 98 (Fig. 2b).

These three clades are allopatric. One clade is represented by populations from Vietnam, Cambodia, Thailand and peninsular Malaysia, the second by populations from Malaysian Borneo (Sarawak and Sabah) and the third by Sumatra and Java populations. Two populations from Thailand (TTR, TPG) and two from Vietnam (VBL, VHCM) identified as *A. epiroticus* (formerly *A. sundaicus* species A) (Dusfour et al., 2004b; Linton et al., 2005), as well as two populations from Sarawak (MAL, MAM) identified as *A. sundaicus* s.s. (Linton et al., 2001), were included in this study. Therefore, continental populations can be considered as *A. epiroticus* and those from Malaysian Borneo as *A. sundaicus* s.s. The third clade constituted a distinct genetic entity in Sumatra and Java (Indonesia).

3.2. Evolution rates and divergence times

The relative rate test results performed in RRTree (Robinson-Rechavi and Huchon, 2000) among COI sequences of *Drosophila* species, and also among COI and Cyt-b sequences within *A. sundaicus* s.l., rejected the hypothesis of heterogeneity of evolution rates (RRTree, *P* > 0.05). These results allowed us to use genetic distances to estimate times of divergence among *Drosophila* species and among *A. sundaicus* s.l. clades (Table 3).

Regression analysis gave a slope that approximates the evolution rate of COI within arthropods at 2.31% of divergence
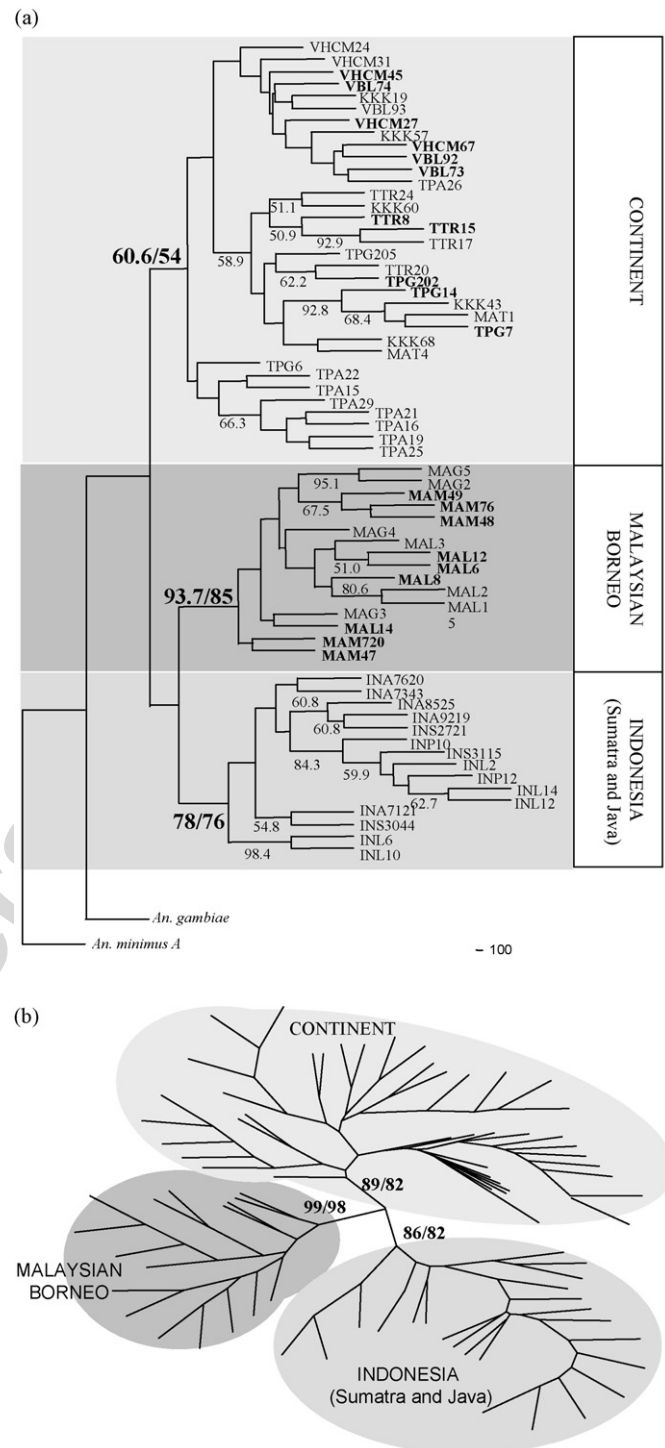

Fig. 2. ML rooted (a) and unrooted (b) trees based on COI and Cyt-b sequences after 1000 bootstraps within PHYML (Guindon and Gascuel, 2003). Support values are for ML reconstruction based on the K81uf + G + I model (rooted tree) and the HKY + G + I (unrooted tree) (left values). The additional support values correspond to MP reconstruction (right values) within Phylip (Felsenstein, 1994). Specimens in bold were studied by Dusfour et al. (2004b), the others are new to this study. Grey zones indicate the three clades that denote species.

per million years (*R* = 0.97). Therefore, based on COI genetic distances, the given averages of divergence time are approximately 0.8 (±0.2) million years (Myr) between the Malaysian Borneo and Indonesian populations, and 1.0 (±0.3)

Table 3
P values of the relative rate test between clades for each mtDNA marker implemented within RRtest, threshold = 0.05

	A. epiroticus/A. sundaicus s.s.	A. sundaicus s.s./Indonesia	A. epiroticus/Indonesia
COI			
RRtree *p*-value	0.26	0.90	0.34
D (S.E.)	0.024 (0.006)	0.018 (0.005)	0.024 (0.006)
TD (S.D.)	1.0 (0.3)	0.8 (0.2)	1.0 (0.3)
Cyt-b			
RRtest *P*-value	0.44	0.55	0.84
D (S.E.)	0.021 (0.006)	0.018 (0.005)	0.016 (0.005)

Tamura-Nei (93) net genetic distances (*D*) and standard errors (S.E.) are indicated and were used to calculate, based on 2.31% divergence per million years, the time of divergence (TD) in million years and its interval (S.D.) for COI.

Myr between the Continental and each other population (Table 3), which fall into the early Pleistocene (1.8–0.78 Myr). The evolution rate of Cyt-b was estimated at 1.82% of divergence per million years ($R = 0.95$).

3.3. ITS2 sequence analysis

Seventy-two base pairs (12.4%) of the 578-bp ITS2 sequence were polymorphic, containing 1.7% (10) of parsimonious informative sites. Overall mean nucleotide diversity was 0.0046 and the number of ribotypes was 50 on 93 sequences. Mean inter-population and intra-population diversities were low, with values equal to 0.001 and 0.004, respectively. The average number of nucleotide differences was 2.680. Among the 10 parsimonious sites, three positions (413, 472 and 532) were observed for more than three clones. Position 532 exhibited a deletion. Because these mutations were consistent with the findings of Linton et al. (2005), and were reproducible between forward and reverse sequences, we excluded possible errors due to the Taq polymerase during DNA amplification. Therefore, based on these three major mutations, we defined four variants, denoted I–IV. At the nucleotidic positions 413, 472 and 532, Variant I showed T, G, C; Variant II presented C, T, C; Variant III exhibited C, G, C; and Variant IV showed C, T, deletion.

Variant I was previously described in *A. epiroticus* on the continent (Linton et al., 2005) and the Variant III was found in *A. sundaicus* D (Alam et al., 2006) on Andaman and Nicobar islands of India. Eight specimens among the continental populations exhibited the Variant I pattern, and two others had Variants I and II as intra-individual variations. Combinations of Variants I and III, II and III, and I, II and III were found, respectively, in two individuals, seven individuals and one individual from Sumatra and Java. Variant IV, exhibiting the deletion, was unique to Malaysian Borneo populations. Therefore, Indonesian populations showed 10 of 10 individuals with intra-individual variations, and continental populations exhibited 2 of 10, whereas Borneo populations had none.

3.4. Demographic trends

Nucleotide diversity (p) calculated for separate mitochondrial markers was low for Cyt-b, ranging from 0.0023 to 0.0057, and low to moderate for COI, from 0.0077 to 0.0146

(Cognato et al., 2003) (Table 4). Haplotype diversity values (h) were low for Cyt-b, from 0.501 to 0.899, and moderate to high for COI, from 0.927 to 0.994 (Table 4). According to Avise (2000), the patterns of Cyt-b revealed that all three lineages underwent recent bottlenecks. COI patterns with higher h and p than Cyt-b patterns represented ancient or transient loss of diversity convergent to stability. Mismatch distribution showed rapid expansion patterns for all three lineages, which were supported by a non-significant raggedness index (Rg) and SSD (Table 4, bold values). Expansion parameters (Theta and Tau), estimated under the expansion model, are indicated in Table 4. This pattern was confirmed by the comparison of indices obtained from Fs, D^* and F^* for the two genes. Indeed, Fs was significant whereas D^* and F^* were not.

Networks of 33 haplotypes for Cyt-b exhibited a separation of five and six mutational steps from *A. epiroticus* (continent) to the Indonesian lineage and to *A. sundaicus* s.s. (Malaysian Borneo), respectively (Fig. 3a). Indonesian and *A. sundaicus* s.s. lineages were characterised by a low number of haplotypes that were organised in a star-like topology. *Anopheles epiroticus* was represented by one major haplotype (number 1), but it was more differentiated in having the highest number of haplotypes (21 on 33). All three networks permitted us to conclude that lineages underwent strong bottleneck events. On the contrary, networks based on the 63 haplotypes from COI (Fig. 3b) revealed minor segregation of lineage haplotypes. They were more diversified than Cyt-b with a larger number of haplotypes and equivalent number of mutational steps between and within clades (Fig. 3b). A maximum of five mutational steps between the *A. sundaicus* s.s. and *A. epiroticus* lineages and two steps between the *A. sundaicus* s.s. and Indonesian lineages were observed. The cores of the *A. epiroticus* and Indonesian lineages are represented by one and two major haplotypes, respectively. The *A. sundaicus* s.s. lineage did not have a representative haplotype. The absence of a star-like topology also suggested that bottleneck events are ancient events and lineages might tend to stabilisation, as for diversity parameter comparisons.

3.5. Assignment of cytotyped individuals to mtDNA lineages

The inclusion of cytotyped individuals from Asahan (Sumatra) in the molecular analysis provided an opportunity

Table 4
Parameters of mitochondrial DNA diversity, mismatch analyses and mutation neutrality tests for each mtDNA marker and each lineage

	N	H	p S.D.(p)	h S.D.(h)	SSD (P_{SSD})	Rg (P_{Rg})	Theta initial	Tau	Fs (P_{Fs})	D^* (P_{D^*})	F^* (P_{F^*})
Cyt-b											
A. sundaicus s.s.	17	6	0.0027 (0.0005)	0.743 (0.089)	**0.0105 (0.42)**	**0.0659 (0.66)**	**0**	1.338	−1.68 (0.00)	**−1.4183 (>0. 10)**	**−1.4436 (>0. 10)**
A. epiroticus	44	21	0.0057 (0.0004)	0.899 (0.036)	**0.0133 (0.06)**	**0.0454 (0.20)**	**0**	2.788	−13.55 (0.00)	**−2.0474 (>0.05)**	**−2.0718 (>0.05)**
Indonesia	27	6	0.0023 (0.0006)	0.501 (0.112)	**0.0195 (0.56)**	**0.1400 (0.68)**	0.846	0.275	−1.44 (0.00)	**−0.5255 (>0. 10)**	**−0.9231 (>0. 10)**
Total	88	33									
COI											
A. sundaicus s.s.	19	18	0.0092 (0.0007)	0.994 (0.019)	**0.0049 (0.43)**	**0.0330 (0.41)**	1.910	6.215	−18.03 (0.00)	**−0.7330 (>0.10)**	**−0.8614 (>0.10)**
A. epiroticus	42	30	0.0146 (0.0011)	0.977 (0.012)	**0.0041 (0.64)**	**0.0128 (0.60)**	2.118	4.318	−16.25 (0.00)	**−1.8040 (>0. 10)**	**−1.9036 (>0.10)**
Indonesia	25	15	0.0077 (0.0010)	0.927 (0.034)	**0.0022 (0.91)**	**0.0150 (0.94)**	1.031	2.416	−7.04 (0.00)	**−2.495 (>0.05)**	**−2.594 (>0.05)**
Total	86	63									

The number of specimens (*N*), number of haplotypes (*H*), nucleotide diversity (*p*) and its standard deviation (S.D.(*p*)), haplotype diversity (*h*) and its standard deviation (S.D.(*h*)), sum of squared deviation (SSD) and its probability P_{SSD}, the raggedness index Rg and its probability P_{Rg}, Theta initial, Tau as well as Fs, D^* and F^* statistics and their associated levels of significance (P_{Fs}, P_{D^*}, P_{F^*}) are indicated. Non-significant values are in bold.
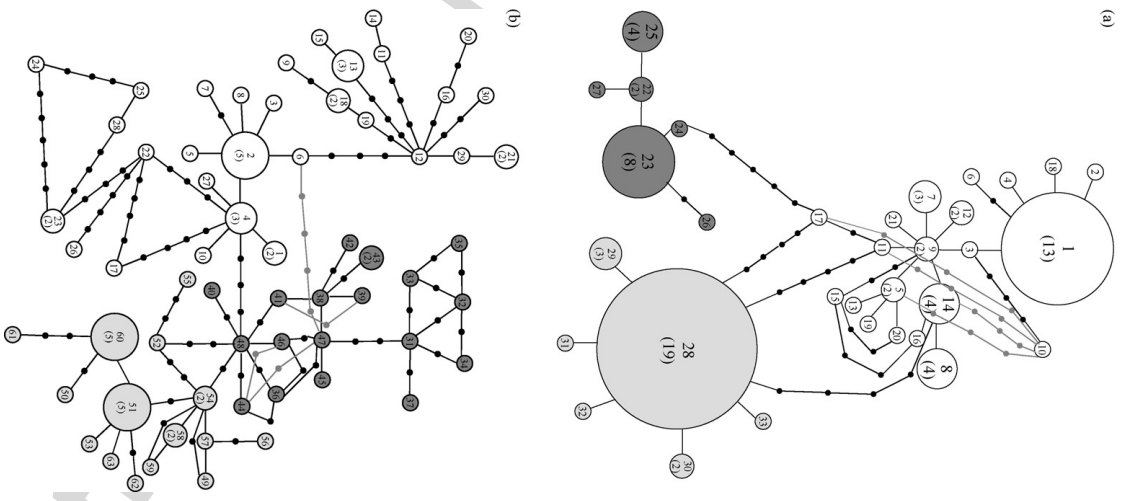

Fig. 3. Minimum-spanning networks of haplotypes from mtDNA markers ((a) Cyt-b; (b) COI) implemented in ARLEQUIN 2.0. The identification number of each haplotype is indicated and the size of circles is proportional to the number of sequences (in brackets). Each clade is indicated as follows: white for *A. epiroticus*, light grey for Indonesia (Sumatra and Java populations) and dark grey for *A. sundaicus* s.s. Black and grey full circles and lines represent mutational steps.

to compare the cytogenetic identification of two forms (B and C) (Sukowati et al., 1996, 1999) with molecular profiles. However, the small number of specimens limited the interpretation. Inclusion of sequences for forms B and C in

the mtDNA phylogeny indicated they both belong to the Indonesian clade (data not shown). In addition, these forms exhibited ITS2 Variants I, II and III, as only were found for Sumatra and Java populations.

4. Discussion

4.1. mtDNA versus ITS2

Ribosomal DNA is known to exhibit intra-individual variations in rDNA for species within the genus *Anopheles* (Onyabe and Conn, 1999; Beebe et al., 2001; Whang et al., 2002), and is currently used to differentiate species based on fixed, interspecific mutations (Collins and Paskewitz, 1996). Malaysian Borneo populations of the *A. sundaicus* complex exhibited a fixed deletion in position 532 of the ITS2 sequence. This result agrees with the findings of Linton et al. (2005), and confirms, in combination with the mtDNA phylogenetic reconstruction, the presence of *A. sundaicus* s.s. in Malaysian Borneo. However, ITS2 was not able to resolve phylogenetic relationships within the *A. sundaicus* complex and was unable to distinguish, by fixed mutations, the two other mtDNA clades. In addition, specimens from Continental and Indonesian clades shared ITS2 variants and presented intra-individual variations. These results may suggest a possible introgression event (Tang et al., 1996), which was not revealed by mtDNA. Whereas the Malaysian Borneo clade remained isolated, a genetic admixture with unequal nuclear and mitochondrial gene flow during co-existence of the two lineages is suggested for the Continental and Indonesian clades (della Torre et al., 1997; Garcia-Paris et al., 2003; Gentile et al., 2002; Besansky et al., 2003).

4.2. Phylogeographical pattern

Phylogenetic reconstructions of *A. sundaicus* s.l. popula-tions based on mtDNA allowed us to identify three lineages. These lineages are allopatric, distributed in three distinct geographic areas. *Anopheles sundaicus* s.s. occurs in Malaysian Borneo; *A. epiroticus* is distributed on the continent from peninsular Malaysia northward to Thailand and eastward to Cambodia and Vietnam, and the third clade includes populations from Sumatra and Java, Indonesia. Based on COI estimation, separation of these lineages occurred in the Early Pleistocene (1.8–0.78 Myr). The *A. sundaicus* s.s. lineage is strongly differentiated from the others by both ITS2 and mtDNA. Based on Cyt-b demographic patterns, all three lineages experienced genetic bottlenecks and have undergone rapid expansions. In contrast, COI analysis suggests that the lineages tend, through genetic expansion, toward a demo-graphic stabilisation. As Cyt-b had a lower evolution rate (1.82%) than COI (2.31%), we can hypothesise that the demographic pattern exhibited by Cyt-b is antecedent to the one shown by COI. Therefore, all three lineages underwent genetic bottlenecks followed by a rapid diversification.

The populations examined in this study are distributed on the emergent part of the continental shelf that includes the continent and the islands of Borneo, Sumatra, Java and Bali. During the Pleistocene epoch (1.8 Myr–11,000 years), cyclical glacial periods impacted on the geography and climate of this shelf. During glacial periods, when sea level was low, the shelf was widely exposed and arid conditions occurred. However, persistent rain-forest refugia were recently hypothesised to have existed in northern Sumatra, northeastern Borneo and peninsular Malaysia (Brandon-Jones, 1996; Gathorne-Hardy et al., 2002; Gorog et al., 2004). During interglacial periods, when sea level was high, the continent and different islands were isolated, experiencing a widespread tropical climate. Based on the current tropical distribution of the *A. sundaicus* complex, we hypothesise that the creation of both islands and refugia affected the history of the complex and its current distribution. During the glacial/arid period, lineages might have been restricted to tropical refugia undergoing genetic bottlenecks. Then, during the interglacial/tropical period, each lineage was able to colonise either islands or continent, allowing genetic diversification. The Malaysian Borneo lineage was restricted to Borneo Island during the high sea-level periods (interglacial periods), and also in the Borneo refugium, which was separated by a large arid area and a significant distance from other refugia during ice ages. This isolation might allow, under genetic drift and bottleneck mechanisms, a stronger genetic differentiation of the Borneo lineage than for the two other lineages. In contrast, the proximity of the Indonesian and peninsular Malaysian refugia probably favoured secondary contacts during low sea-level periods. The Continental and Indonesian lineages were most likely only separated during the period of highest sea level.

The genetic markers used in this study might reflect the most recent genetic bottleneck/expansion step during the history of the *A. sundaicus* complex, which occurred since the last glacial maximum (20,000 years). However, this phenomenon has been repeated cyclically coincident with climate cycles, reinforcing the divergence among the three clades and the suspected phenomenon. These cyclical bottlenecks could have also erased or modified ancestral polymorphism. Such lack of information might explain the inability to resolve basal relationships among clades, as well as the incongruence between Cyt-b and COI markers indicated by the ILD test (Andreev et al., 1998; Gandolfi et al., 2001; Gorog et al., 2004).

4.3. Discrepancy between ecological hypothesis, chromosomal investigations and molecular analyses

Seng et al. (2001) suggested that the different ecological patterns of *A. sundaicus* s.l. reflected the existence of a species complex. However, the present work did not reveal distinctions between populations from fresh and brackish water. The populations from freshwater sites at South Tapanuli (INS, Sumatra) and Miri (MAM, Sarawak) were not genetically distinct from those of brackish water sites, in agreement with our previous findings (Dusfour et al., 2004b). These ecological differences reflect adaptation to salinity ranging from strongly brackish to freshwater (Nguyen et al., 1993; Dusfour et al., 2004a; Alam et al., 2006).

Combined analyses of forms B and C showed an absence of congruence between cytogenetic/allozyme and molecular data.

Neither ITS2 nor COI confirmed the genetic distinction of forms B and C from Asahan (Sumatra). This is not the first time that such incongruence has been noted in *Anopheles* species (Besansky, 1999). The presence of chromosomal inversions was correlated with the salinity of habitats where larvae of forms B and C were collected (Sukowati et al., 1996, 1999). Chromosome inversions in members of the *A. gambiae* complex are known to be related to environmental adaptation (White, 1973; Coluzzi, 1982; Toure et al., 1998), but not necessarily linked to speciation (Rieseberg, 2001). As is the case for *A. funestus* (see Besansky, 1999), the adaptation of *A. sundaicus* s.l. at a local scale is recent. Chromosome rearrangements would show this ecological adaptation, whereas genes would still indicate current gene flow. Incongruence between isozyme and DNA data analyses within Diptera has rarely been reported (Reyes and Ochando, 1998). However, intrinsic properties and extrinsic factors are known to influence the ability of markers to reflect population history. Therefore, isozymes could have also been subjected to differential evolution correlated with environmental constraints.

4.4. Taxonomic status

Two species of the *A. sundaicus* complex, *A. sundaicus* s.s. and *A. epiroticus*, have been described based on morphology and molecular markers (Linton et al., 2001, 2005). *Anopheles epiroticus* occurs along the continental coast from southern Vietnam (below the 11th parallel) to peninsular Malaysia and *A. sundaicus* s.s. is currently only known on Malaysian Borneo. The Indonesian lineage has not been formally named. Based on mtDNA, this taxon is distinct from both *A. sundaicus* s.s. and *A. epiroticus*. Phylogenetic differentiation related to equivalent genetic divergence, such as those observed between members of *Anopheles* species complexes (Foley et al., 1998), suggests the existence of possible new species (Cracraft, 1983).

The Indonesian lineage does not correspond specifically to either form B or C of Sukowati et al. (1996, 1999) in Sumatra and its status in relation to form D (Nanda et al., 2004; Alam et al., 2006) on Nicobar Island is uncertain. Form D was recognised through cytogenetics and ITS2 analysis. Its ITS2 sequence corresponds to Variant III described above, but no further investigation based on mtDNA has been done. Therefore, we prefer to recognise the Indonesian lineage as *A. sundaicus* E.

With the development of a PCR assay that can differentiate *A. sundaicus* s.s., *A. epiroticus* and *A. sundaicus* E (Dusfour et al., 2007), complementary study sites in peninsular Malaysia, Sumatra and the Malacca Strait, where a zone of sympatry may occur, will be of great interest for further study of these incipient species and their possible introgression.

Acknowledgements

References

Alam, M.T., Das, M.K., Ansari, M.A., Sharma, Y.D., 2006. Molecular identification of *Anopheles* (*Cellia*) *sundaicus* from the Andaman and Nicobar islands of India. Acta Trop. 97, 10–18.

Andreev, D., Breilid, H., Kirkendall, L., Brun, L.O., ffrench-Constant, R.H., 1998. Lack of nucleotide variability in a beetle pest with extreme inbreeding. Insect Mol. Biol. 7, 197–200.

Avise, J.C., 2000. Phylogeography. The History and Formation of Species. Harvard University Press, Cambridge, MA, 464 pp.

Bandelt, H.J., Forster, P., Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Beard, C.B., Hamm, D.M., Collins, F.H., 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Mol. Biol. 2, 103–124.

Beebe, N.W., Saul, A., 1995. Discrimination of all members of the *Anopheles punctulatus* complex by polymerase chain reaction-restriction fragment length polymorphism analysis. Am. J. Trop. Med. Hyg. 53, 478–481.

Beebe, N.W., Maung, J., van den Hurk, A.F., Ellis, J.T., Cooper, R.D., 2001. Ribosomal DNA spacer genotypes of the *Anopheles bancroftii* group (Diptera: Culicidae) from Australia and Papua New Guinea. Insect Mol. Biol. 10, 407–413.

Besansky, N.J., 1999. Complexities in the analysis of cryptic taxa within the genus *Anopheles*. Parassitologia 41, 97–100.

Besansky, N.J., Krzywinski, J., Lehmann, T., Simard, F., Kern, M., Mukabayire, O., Fontenille, D., Toure, Y., Sagnon, N., 2003. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. Proc. Natl. Acad. Sci. U.S.A. 100, 10818–10823.

Brandon-Jones, D., 1996. The Asian Colobinae (Mammalia: Cercopithecidae) as indicators of quaternary climatic change. Biol. J. Linn. Soc. Lond. 59, 327–350.

Brower, A.V., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. U.S.A. 91, 6491–6495.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657–1659.

Cognato, A.I., Harlin, A.D., Fisher, M.L., 2003. Genetic structure among pinion pine beetle populations (Scolytinae: *Ips confusus*). Environ. Entomol. 32, 1262–1270.

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

Coluzzi, M., 1982. Spatial distribution of chromosomal inversions and speciation in anopheline mosquitoes. Prog. Clin. Biol. Res. 96, 143–153.

Cracraft, J., 1983. Species concepts and speciation analysis. Curr. Ornith. 1, 159–187.

De Salle, R., Friedman, T., Prager, E.M., Wilson, A.C., 1987. Tempo and mode of sequence evolution in mitochondrial of Hawaiian *Drosophila*. J. Mol. Evol. 26, 157–164.

della Torre, A., Merzagora, L., Powell, J.R., Coluzzi, M., 1997. Selective introgression of paracentric inversions between two sibling species of the *Anopheles gambiae* complex. Genetics 146, 239–244.

Dusfour, I., Harbach, R.E., Manguin, S., 2004a. Bionomics and systematics of the oriental *Anopheles sundaicus* complex in relation to malaria and vector control. Am. J. Trop. Med. Hyg. 71, 518–524.

Dusfour, I., Linton, Y.M., Cohuet, A., Harbach, R.E., Baimai, V., Trung, H.D., Chang, M.S., Matusop, A., Manguin, S., 2004b. Molecular evidence of speciation between island and continental populations of *Anopheles* (*Cellia*) *sundaicus* Rodenwaldt (Diptera: Culicidae), a principal malaria vector in Southeast Asia. J. Med. Entomol. 41, 287–295.

Dusfour, I., Blondeau, J., Harbach, R.E., Vythilingham, I., Baimai, V., Trung, H.D., Sochanta, T., Bangs, M., Manguin, S., 2007. PCR identification of the three members of the *Anopheles sundaicus* (Diptera: Culicidae) complex, malaria vectors in Southeast Asia. J. Med. Entomol. 44, in press.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of congruence. Cladistics 10, 315–319.

Felsenstein, J., 1994. PHYLIP Computer Program, Version by Felsenstein J. University of Washington, Seattle, Washington.

Fleischer, R.C., McIntosh, C.E., Tarr, C.L., 1998. Evolution on a volcanic conveyor belt: using phylogeographic reconstructions and K–Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Mol. Ecol. 7, 533–545.

Foley, D.H., Bryan, J.H., Yeates, D., Saul, A., 1998. Evolution and systematics of *Anopheles*: insights from a molecular phylogeny of Australasian mosquitoes. Mol. Phyl. Evol. 9, 262–275.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Fu, Y.X., Li, W.H., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.

Gandolfi, A., Bonilaui, P., Rossi, V., Menossi, P., 2001. Intraindividual and intraspecies variability of ITS1 sequences in the ancient asexual *Darwinula stevensoni* (Crustacea: Ostracoda). Heredity 87, 449–455.

Garcia-Paris, M., Alcobendas, M., Buckley, D., Wake, D.B., 2003. Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (*Salamandra*) inferred from discordance of genetic and morphological traits. Evolution 57, 129–143.

Gathorne-Hardy, F.J., Syaukani, R.G., Davies, R.G., Eggleton, P., Jones, D.T., 2002. Quaternary refugia in south-east Asia: using termites (Isoptera) as indicators. Biol. J. Linn. Soc. Lond. 75, 453–466.

Gentile, G., Della Torre, A., Maegga, B., Powell, J.R., Caccone, A., 2002. Genetic differentiation in the African malaria vector, *Anopheles gambiae* s.s., and the problem of taxonomic status. Genetics 161, 1561–1578.

Gorog, A.J., Sinaga, M.H., Engstrom, M.D., 2004. Vicariance or dispersal? Historical biogeography of three Sunda shelf murine rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteadi*). Biol. J. Linn. Soc. Lond. 91, 91–109.

Guindon, S., Gascuel, O., 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Harbach, R.E., Parson, E., Chen, B., Butlin, R.K., 2006. *Anopheles* (*Cellia*) *minimus* Theobald (Diptera: Culicidae): Neotype designation, characterization, and systematics. Proc. Entomol. Soc. Wash. 108, 198–209.

IMPE, 1987. Keys to *Anopheles* in Vietnam (Adults–Pupae–Larvae). Department of Entomology, Institute of Malariology, Parasitology and Entomology, Hanoi, Vietnam.

Knowlton, N., Weigt, L.A., Solorzano, L.A., Mills, D.K., Bermingham, E., 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260, 1629–1632.

Kumar, S., Tamura, K., Jacobsen, I.B., Masatoshi, N., 2001. MEGA molecular evolutionary genetics analysis version 2.1. Bioinformatics 17, 1244–1245.

Linton, Y.M., Harbach, R.E., Seng, C.M., Anthony, T.G., Matusop, A., 2001. Morphological and molecular identity of *Anopheles* (*Cellia*) *sundaicus* (Diptera: Culicidae), the nominotypical member of a malaria vector species complex in Southeast Asia. Syst. Entomol. 26, 357–366.

Linton, Y.M., Dusfour, I., Howard, T.M., Ruiz, L.F., Duc Manh, N., Ho Dinh, T., Sochanta, T., Coosemans, M., Harbach, R.E., 2005. *Anopheles* (*Cellia*) *epiroticus*, a new malaria vector species in the Southeast Asian Sundaicus Complex. Bull. Entomol. Res. 95, 329–339.

Nanda, N., Das, M.K., Wattal, S., Adak, T., Subbarao, S.K., 2004. Cytogenetic characterization of *Anopheles sundaicus* (Diptera: Culicidae) population from Car Nicobar Island. India. Ann. Entomol. Soc. Am. 97, 171–176.

Nguyen, T.A., Le Quy, R., Vu Thi, H., Nguyen Bieh, L., 1993. Entomo-epidemiological studies of malaria in the coastal zone of Ho Chi Minh City, 1990–1992. Santé 3, 464–473.

Onyabe, D.Y., Conn, J.E., 1999. Intragenomic heterogeneity of a ribosomal DNA spacer (ITS2) varies regionally in the neotropical malaria vector *Anopheles nuneztovari* (Diptera: Culicidae). Insect Mol. Biol. 8, 435–442.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Reid, J.A., 1968. Anopheline Mosquitoes of Malaya and Borneo. Studies of the Institute of Medical Research Malaya. Government of Malaysia, Kuala Lumpur, Malaysia, 520 pp.

Reid, J.A., 1970. Systematics of malaria vectors—anopheline systematics and malaria control, with special reference to Southeast Asia. Misc. Publ. Entomol. Soc. Am. 7, 52–62.

Reyes, A., Ochando, M.D., 1998. Genetic differentiation in Spanish populations of *Ceratitis capitata* as revealed by abundant soluble protein analysis. Genetica 104, 59–66.

Rieseberg, L.H., 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16, 351–358.

Robinson-Rechavi, M., Huchon, D., 2000. RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics 16, 296–297.

Rogers, A.R., Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. Mol. Biol. Evol. 9, 552–569.

Rozas, J., Sanchez-Delbarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Schneider, S., Roessli, D., Excoffier, L., 2000. ARLEQUIN Version 2.000 A Software for Population Genetics Data Analysis Computer Program, Version by Schneider, S., Roessli, D., Excoffier, L. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva.

Seng, C.M., Linton, Y.M., Harbach, R.E., 2001. The *Anopheles sundaicus* species complex—a preliminary study of larval ecology from two sites in Sarawak, Malaysia. Mekong Malaria Forum 8, 91–95.

Simmons, R.B., Weller, S.J., 2001. Utility and evolution of cytochrome *b* in insects. Mol. Phylogenet. Evol. 20, 196–210.

Slatkin, M., 1991. Inbreeding coefficients and coalescence times. Genet. Res. 58, 167–175.

Sukowati, S., Baimai, V., Andris, H., 1996. Sex chromosome variation in natural populations of the *Anopheles sundaicus* complex from Thailand and Indonesia. Mosq. Borne Dis. Bull. 13, 8–13.

Sukowati, S., Baimai, V., Harun, S., Dasuki, Y., Andris, H., Efriwati, M., 1999. Isozyme evidence for three sibling species in the *Anopheles sundaicus* complex from Indonesia. Med. Vet. Entomol. 13, 408–414.

Swofford, D.L., 2002. PAUP: Phylogenetic Analysis Using Parsimony (and Others Methods) Version 4.0b10. D. L., Sinauer Associates, Sunderland, MA.

Tamura, K., Subramanian, S., Kumar, S., 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol. Biol. Evol. 21, 36–44.

Tang, J., Toe, L., Back, C., Unnasch, T.R., 1996. Intra-specific heterogeneity of the rDNA internal transcribed spacer in the *Simulium damnosum* (Diptera: Simuliidae) complex. Mol. Biol. Evol. 13, 244–252.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Toure, Y.T., Petrarca, V., Traore, S.F., Coulibaly, A., Maiga, H.M., Sankare, O., Sow, M., Di Deco, M.A., Coluzzi, M., 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the *Anopheles gambiae* complex in Mali, West Africa. Parasitologia 40, 477–511.

Whang, I.J., Jung, J., Park, J.K., Min, G.S., Kim, W., 2002. Intragenomic length variation of the ribosomal DNA intergenic spacer in a malaria vector, *Anopheles sinensis*. Mol. Cells 14, 158–162.

White, M.J.D., 1973. Animal Cytology and Evolution. Cambridge University Press, Cambridge, 375 pp.