*Journal of the American Mosquito Control Association,* 19(3):261–264, 2003


# SCIENTIFIC NOTE

# SUGGESTING NEW SPECIES? COMMENTS ON "EVIDENCE FOR A NEW SPECIES OF *ANOPHELES MINIMUS* FROM THE RYUKYU ARCHIPELAGO, JAPAN"

W. VAN BORTEL AND M. COOSEMANS

*Department of Parasitology, Prince Leopold Institute of Tropical Medicine, Nationalestraat 155, B-2000 Antwerpen, Belgium*

ABSTRACT. Species recognition and identification are crucial for the planning and evaluation of vector control. In recent years, considerable effort has been made to clarify the species composition of vector taxa. The danger exists, however, that new species are suggested on little evidence and that intraspecific variation is insufficiently surveyed. We discuss these topics on the basis of recent studies on the *Anopheles minimus* complex, one of the most widespread vector groups in Southeast Asia.

**KEY WORDS** *Anopheles minimus,* species complex, population genetics, Southeast Asia, Vietnam

Elucidating the taxonomic status of the major malaria vector taxa and characterizing the individual species within each complex are important to understand the complexity of the vector systems in Southeast Asia. Moreover, failure to identify mosquito species hampers the study and monitoring of vectors and, hence, complicates the follow-up and reassessment of preventive control measures. Ultimately, this failure will impede progress toward the consolidation and further improvement of the malaria situation in Southeast Asia. Consequently, several studies have been conducted to understand the species composition of the *Anopheles minimus* complex, one of the most widespread malaria vector taxa in Southeast Asia (Green et al. 1990, Sharpe 1997, Van Bortel et al. 1999).

Within the *An. minimus* complex are 2 commonly recognized species, informally designated *An. minimus* A and C (Harbach 1994). These species were defined by their electromorphs at the *Odh* enzyme locus (Green et al. 1990). Recently, Somboon et al. (2001) used morphological, cytogenetic, and hybridization evidence to suggest a new species within the *An. minimus* complex occurring on the Ryukyu archipelago, Japan. In the same paper, the authors also suggested the possible presence of up to 4 species in the *An. minimus* complex in northern Vietnam. This suggestion is in contrast to the results of our studies. To clarify the species composition of the *An. minimus* complex in Vietnam, we conducted a large-scale study in 4 villages in northern Vietnam with the use of allozyme electrophoresis (Van Bortel et al. 1999). We incorporated the *Odh* locus in our study because of its ability to distinguish *An. minimus* A and C in Thailand (Green et al. 1990). In our analyses, however, we considered the possibility of the presence of more than 2 species within the complex, which was already suspected at that time (Zahar 1996). In the morphologically identified *An. minimus* s.l. from Vietnam, we observed significant heterozygote deficiency at 5 out of 13 enzyme loci. The highest heterozygote deficiency was observed at locus *Odh,* which could be used to divide the complex into 2 isomorphic taxa between which only 0.9% putative hybrids were present. After separating the taxa, species C still showed a significant heterozygote deficiency at loci *Ldh* and *Gpi.* But this was observed in only 1 of the 10 samples analyzed, i.e., in the only village where the putative hybrids (<1%) between the taxa (*An. minimus* A and C) were found. On the basis of the complete similarities of the internal transcribed spacer 2 (ITS2) ribosomal DNA (rDNA) sequences, *An. minimus* taxa from Vietnam were identified as *An. minimus* species A and C known from Thailand (Van Bortel et al. 1999, 2000).

Somboon et al. (2001) suggested the existence of up to 4 species within the *An. minimus* complex in Vietnam on the basis of the presence of 2 novel sequences of the D3 28S rDNA region previously not found in *An. minimus* A and C from Thailand (Sharpe et al. 2000). Their suggestion is based on 1 D3 sequence of *An. minimus* from Hanoi, Vietnam, taken from Sharpe et al. (2000) and on 3 specimens originating from our study population from Hoa Binh Province, northern Vietnam. One of the 2 novel sequences, identified by us as *An. minimus* A, differed by only 1 out of 370 nucleotides (at position 72) from *An. minimus* A from Sharpe et al. (2000) (Somboon et al. 2001). Although the sequence was derived directly from a polymerase chain reaction amplification product, which reduces problems associated with errors made by *Taq* polymerase, errors cannot be completely excluded (Hillis et al. 1996). The 2nd sequence of Somboon et al. (2001), also based on 1 specimen, is more troublesome. We identified this specimen as species C

by enzyme electrophoresis (locus *Odh*), though it exhibited the highest sequence similarity with species A from Sharpe at al. (2000), from which the D3 region differed only at positions 248 and 296. However, Somboon et al. (2001) did not clearly state whether possible sources of error, e.g., mislabeling or contamination, could be ruled out. Hence, it would have been desirable to confirm the presence of these 2 novel sequences by analyzing more specimens.

Somboon et al. (2001) proposed 2 alternative explanations for the existence of the 4 sequences at the D3 locus in northern Vietnam: 1) the Vietnamese *An. minimus* A and/or C species exhibit large intraspecific variation at the D3 locus, 2) alternatively, each sequence could come from an individual of a different species. The D3 region was identified by Sharpe et al. (1999) as a diagnostic marker for *An. minimus* A and C from Thailand. Fine-tuning of diagnostic markers is necessary when specimens are tested coming from an area, in casu Vietnam, where its variation was not or insufficiently screened. Moreover, the utility of a locus for taxonomic and phylogenetic purposes depends on the ability to accurately predict whether the nature of potential variation is intra- or interspecific. The possibility of intraspecific variation at the D3 region, as alternative explanation of the observed sequence patterns, was considered unlikely by Somboon et al. (2001) because of the amount of variation observed in Vietnamese species compared with the absence of variation in Thai *An. minimus* species. The only exception to this was a different sequence pattern at the D3 region in 1 specimen (individual #157) found in Kanchanaburi Province, Thailand. Sharpe et al. (1999) suggested that this specimen could be from another species. However, they were tentative in their hypothesis because this specimen, #157, clustered with *An. minimus* C in a maximum likelihood tree based on the mitochondrial cytochrome II oxidase (Sharpe et al. 2000). Consequently, intraspecific variation at the D3 region seems still a likely explanation for the observed sequence patterns. Moreover, genetic variation is not a fortiori evenly distributed throughout the geographic range of a species. For example, genetic diversity of the *Anopheles gambiae* s.s. population from eastern Kenya was lower than that of western Kenya (Lehmann et al. 1998, Donnelly et al. 2001), whereas in *Aedes aegypti,* 1 out of 10 populations from the northeastern coast of Mexico showed a reduced heterozygosity (Gorrochotegui-Escalante et al. 2000). Whether this is the case for the members of the *An. minimus* complex needs more investigation. However, the 2 *An. minimus* species, A and C, recognized by Green et al. (1990) in Thailand were monomorphic for the *Odh* locus, whereas *An. minimus* C from Vietnam was polymorphic for the same locus (Van Bortel et al. 1999). Hence, in *An. minimus* C, variation at the *Odh* locus was not evenly distributed among populations from different localities.

Somboon et al. (2001) concluded, on the basis of the low bootstrap values, that the neighbor-joining tree of the D3 sequences contained only a limited amount of phylogenetic information. Phylogenies derived from a single gene do not necessarily reflect the species phylogeny (Avise and Ball 1990). The probability of a gene tree having the same topology as the species tree is low when divergence times between species are short and when the effective population sizes are large (Pamilo and Nei 1988). Both these conditions apply to the *An. minimus* complex (Sharpe et al. 2000, Van Bortel, unpublished data). Hence, at this stage it is premature to decide on whether the observed variation at the D3 region reflects intra- or interspecific patterns. Therefore, the presence of a deviating sequence pattern does not suffice to suggest a new species, for this requires more information from different genes. For example, for almost 20 years, 5 chromosomal forms in *An. gambiae* s.s. have been known. Extensive research is still going on to explore the reproductive status of these chromosomal forms, examining the variation among 12 genes located throughout the *An. gambiae* s.s. genome. However, the need for broader geographic sampling of *An. gambiae* s.s is still recognized (Black and Lanzaro 2001). Consequently, species can be recognized only by combining information from multiple characters (allele frequencies at multiple loci, different traits) of which inter- and intraspecific variation is sufficiently surveyed. This is especially the case for closely related species (Tabachnick and Black 1995, Black and Munstermann 1996).

Somboon et al. (2001) argued that the 4 specimens from Vietnam originated from discrete reproductive populations because they were all homozygous for the D3 fragment. Beside the fact that it is not clear how homozygosity of the D3 fragment was determined, being homozygote per se does not give any indication of belonging to a particular reproductive population. Species are evolutionary lineages, i.e., collections of allelic pathways among interbreeding individuals. The lineage boundaries that define species arise from the forces that create reproductive communities (Avise and Wollenberg 1997, Templeton 1998). A generally applicable and efficient approach to define reproductively isolated gene pools is a population genetic analysis based on allozyme electrophoresis, DNA microsatellites, or other codominant molecular markers that allow the estimation of allele frequencies. Population genetic structure can then be inferred by the Wright's *F*-statistics ($F_{it}$, $F_{st}$, and $F_{is}$). The few specimens Somboon et al. (2001) used to suggest new species are not suitable to generate these statistics. Consequently, the use of these data to make inference about reproductive isolation of the samples is troublesome, even if one expects that novel sequences should appear in heterozygous form with the com-

mon allele. Moreover, because rDNA is liable to concerted evolution and gene conversion, heterozygotes of rDNA may get lost (Hillis et al.1991, Tripet et al. 2001).

To support their interpretation, Somboon et al. (2001) referred to the significant heterozygote deficiency we observed in species C of Vietnam (Van Bortel et al. 1999), which they interpreted as indicating assortative mating. In other words, they considered what is called species C in Vietnam to be a mixture of different species. As mentioned above, this Hardy–Weinberg disequilibrium occurred only in the population where species A and C hybridized (Van Bortel et al. 1999). In hybrid zones, different evolutionary forces may be operating. On one hand, positive assortative mating seems a common feature. On the other hand, novel genotypes might arise in hybrid zones. Moreover, introgression can increase the genetic variation of the recipient taxon (Howard 1998), whereas selection and associations between loci in hybrid zones can be different compared with nonhybrid zones (Barton and Hewitt 1989). So nonrandom association of gametes in the parental population is not unlikely, which might explain the Hardy–Weinberg disequilibrium observed at loci *Ldh* and *Gpi*. Finally, the Hardy–Weinberg disequilibria at these loci were not observed again in subsequent samples of *An. minimus* C collected in the same village in 1998–99 (Van Bortel, unpublished).

In conclusion, although Somboon et al. (2001) only tentatively suggested the possible presence of up to 4 species in the *An. minimus* complex in northern Vietnam, they found it the most likely explanation for the sequence patterns at the D3 locus. However, in view of the rationale presented above, we think that even this tentative conclusion is insufficiently supported by the currently available data. To understand the differentiation between populations across geographic ranges, the structuring of genetic polymorphisms must be analyzed. Yet, the occurrence of polymorphisms is not an a priori indication of the existence of different taxa. Species can be recognized only in a broader context of inter- and intraspecific variation at different geographic scales. Studying inter- and intraspecific diversity in population structure and traits will also be useful in understanding the complexity of the vector system.

We are grateful to T. Backeljau for the fruitful discussions. The research of *Anopheles minimus* in Vietnam, on which this discussion is based, is financed by the Belgian Development Co-operation (DGIS) and by the INCO-DC project ERBIC18CT970211.

## REFERENCES CITED

Avise JC, Ball RMJ. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surv Evol Biol* 7:45–67.

Avise JC, Wollenberg K. 1997. Phylogenetics and the origin of species. *Proc Natl Acad Sci USA* 94:7748–7755.

Barton NH, Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature* 341:497–502.

Black WC IV, Lanzaro GC. 2001. Distribution of genetic variation among chromosomal forms of *Anopheles gambiae s.s.*: introgressive hybridization, adaptive inversions, or recent reproductive isolation? *Insect Mol Biol* 10:3–7

Black WC IV, Munstermann LE. 1996. Molecular taxonomy and systematics of arthropod vectors. In: Beaty RJ, Marquart WC, eds. *The biology of disease vectors* Boulder, CO: Univ. Press of Colorado. p 438–470.

Donnelly MJ, Licht MC, Lehmann T. 2001. Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Mol Biol Evol* 18:1353–1364.

Gorrochotegui-Escalante N, Munoz ML, Fernandez-Salas I, Beaty GJ, Black WC IV. 2000. Genetic isolation by distance among *Aedes aegypti* populations along the northeastern coast of Mexico. *Am J Trop Med Hyg* 62: 200–209.

Green CA, Gass RF, Munstermann LE. 1990. Population-genetic evidence for two species in *Anopheles minimus* in Thailand. *Med Vet Entomol* 4:25–34.

Harbach RE. 1994. Review of the internal classification of the genus *Anopheles* (Diptera: Culicidae): the foundation for comparative systematics and phylogenetic research. *Bull Ent Res* 84:331–342.

Hillis DM, Mable BK, Larson A, Davis SK, Zimmer EA. 1996. Nucleic acids IV: sequencing and cloning. In: Hillis DM, Moritz C, Mable BK, eds. *Molecular systematics* Sunderland, MA: Sinauer Associates. p 321–381.

Hillis DM, Moritz C, Porter CA, Baker RG. 1991. Evidence for biased conversion in concerted evolution of ribosomal DNA. *Proc R Soc Lond B* 251:308–310.

Howard DJ. 1998. Unanswered questions and future directions in the study of speciation. In: Howard DJ, Berlocher SH, eds. *Endless forms* Oxford, United Kingdom: Oxford Univ. Press. p 439–452.

Lehmann T, Hawley WA, Grebert H, Collins FH. 1998. The effective population size of *Anopheles gambiae* in Kenya: implications for population structure. *Mol Biol Evol* 15:264–276.

Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. *Mol Biol Evol* 5:568–583.

Sharpe RG. 1997. The status of cryptic species within *Anopheles minimus*. Ph.D. thesis. School of Biology, The University of Leeds, Leeds, United Kingdom.

Sharpe RG, Harbach RE, Butlin RK. 2000. Molecular variation and phylogeny of members of the minimus group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). *Syst Entomol* 25:263–272.

Sharpe RG, Hims MM, Harbach RE, Butlin RK. 1999. Two PCR based methods for identification of species of the *Anopheles minimus* group: allele specific amplification and single strand conformation polymorphism. *Med Vet Entomol* 13:265–273.

Somboon P, Walton C, Sharpe RG, Higa Y, Tuno N, Tsuda Y, Takagi M. 2001. Evidence for a new sibling species of *Anopheles minimus* from the Ryukyu Archipelago, Japan. *J Am Mosq Control Assoc* 17:98–113.

Tabachnick WJ, Black IV WC. 1995. Making a case for molecular population genetic studies of arthropod vectors. *Parasitol Today* 11:27–30.

Templeton AR. 1998. Species and speciation—geography, population structure, ecology, and gene trees. In: Howard DJ and Berlocher SH, eds. *Endless forms* Oxford, United Kingdom: Oxford Univ. Press. p 32–43.

Tripet F, Toure YT, Taylor CE, Norris DE, Dolo G, Lanzaro GC. 2001. DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Mol Ecol* 10: 1725–1732.

Van Bortel W, Trung HD, Manh ND, Roelants P, Verlé P, Coosemans M. 1999. Identification of two species within the *Anopheles minimus* complex in northern Vietnam and their behavioural divergences. *Trop Med Int Health* 4:257–265.

Van Bortel W, Trung HD, Roelants P, Harbach RE, Backeljau T, Coosemans M. 2000. Molecular identification of *Anopheles minimus* s.l. beyond distinguishing the members of the complex. *Insect Mol Biol* 9:335–340.

Zahar AR. 1996. *Vector bionomics in the epidemiology and control of malaria* Part III. *The WHO South-East Asia region and the Western Pacific region* Volume II. Leading literature-general review 1970–1994. Document CTD/MAL/96.1. Geneva, Switzerland: World Health Organization.