Medical and Veterinary Entomology (2003) **17**, 67–74

Molecular identification of the *Anopheles nili* group of African malaria vectors

P. KENGNE[*†], P. AWONO-AMBENE[†], C. ANTONIO-NKONDJIO[†], F. SIMARD[†] and D. FONTENILLE[*†]

[*]Laboratoire de Lutte contre les Insectes Nuisibles (LIN), Institut de Recherche pour le Développement (IRD), Montpellier, France and [†]Laboratoire de l'IRD, Organisation de lutte Contre les grandes Endémies en Afrique Centrale (OCEAC), Yaoundé, Cameroon

Abstract. Distinction between members of the *Anopheles nili* group of mosquitoes (Diptera: Culicidae), including major malaria vectors in riverside villages of tropical Africa, has been based mainly on doubtful morphological characters. Sequence variations of the ribosomal DNA second internal transcribed spacer (ITS2) and D3 28S region between morphological forms revealed four genetic patterns corresponding to typical *An. nili* (Theobald), *An. carnevalei* Brunhes *et al.*, *An. somalicus* Rivola & Holstein and the newly identified variant provisionally named Oveng form. Primers were designed based on ITS2 fixed nucleotide differences between haplotypes to develop a multiplex PCR for rapid and specific identification of each species or molecular form. Specimens of the *An. nili* group from Cameroon, Burkina Faso, Ivory Coast and Senegal were successfully identified to species, demonstrating the general applicability of this technique based on criteria described in this paper.

Key words. *Anopheles nili*, allele-specific PCR, malaria vectors, ribosomal DNA, sibling species, Afrotropical, Cameroon, Burkina Faso, Ivory Coast, Senegal.

Introduction

Anopheles (*Cellia*) mosquitoes are responsible for human malaria transmission in tropical areas of the Old World. In Africa, most of the important malaria vectors are members of species complexes, whose members are difficult and sometimes impossible to distinguish morphologically (Gillies & De Meillon, 1968; Coluzzi, 1984; Gillies & Coetzee, 1987; Fontenille & Lochouarn, 1999). These difficulties have stimulated the development of molecular tools for precise and reliable identification of sibling species. One of the most widely used regions of the genome to infer genetic variation and phylogenetic relationships is the ribosomal DNA (rDNA) cluster, a tandemly repeated multigene family. In eukaryotic organisms, every repeated unit of rDNA consists of an intergenic spacer (IGS), followed by genes coding the 18S, 5.8S and 28S rDNA. Preceeding the 18S gene is the external transcribed spacer (ETS), and surrounding the 5.8S rDNA are the internal transcribed spacers 1 and 2 (ITS1 and ITS2). This multigene family evolves cohesively within species through concerted evolution, a mechanism that tends to homogenize sequences within species while driving differentiation between species (Arnheim, 1983). The coding regions are highly conserved, even between distantly related species, while non-coding DNA usually rapidly drifts apart, even between closely related species. Thus, using primers located in the conserved rDNA regions, variable regions can be amplified from a wide range of species in absence of prior sequence information. As such, the rDNA cluster has become an increasingly popular tool in molecular entomology (Collins & Paskewitz, 1996) and, in particular, as a means to develop diagnostic tests to differentiate cryptic Anopheline species. For example, diagnostic PCR assays based on segregating sequence variation in the ITS and/or IGS are now available to identify members of the *An. gambiae* Giles, *An. quadrimaculatus* Say, *An. punctulatus* Dönitz, *An. maculipennis* Meigen and *An. funestus* Giles complexes (Scott *et al.*, 1993; Beebe & Saul, 1995; Cornel *et al.*, 1996b; Proft *et al.*, 1999;

Correspondence. Dr Pierre Kengne, LIN/IRD, 911 Avenue Agropolis, BP 5045 34032 Montpellier Cedex 1, France. E-mail: pierre.kengne@mpl.ird.fr

Hackett *et al*., 2000). Furthermore, the D3 domain, known to be the most variable (coding) region in the 28S subunit provides a powerful tool for taxonomic studies in anophelines (Torres *et al*., 2000; Koekemoer *et al*., 1999).

Anopheles nili (Theobald) is widespread across tropical Africa (Hamon & Mouchet, 1961) and appears to be the major vector of malaria in some rural forested areas of central Africa, with entomological inoculation rates reaching 100 infective bites/person/year (Carnevale *et al*., 1992). Larvae of *An. nili* are typically found in vegetation or in dense shade along the edges of streams and large rivers. The extensive morphological, ecological and ethological variations among *An. nili* populations have been reported by many authors (Gillies & De Meillon, 1968; Carnevale *et al*., 1992; Brunhes *et al*., 1999), suggesting that *An. nili* is a group of species.

To assess relevance of morphological characters as an accurate mean for classification within the group *An. nili*, we investigated sequence variation in the rDNA ITS2 and D3 domains of different morphological forms of this group. The analysis concerned the anthropophilic *An. nili* typical (T), the pale-winged *An. nili* Congo (Gillies & De Meillon, 1968) and three variants recently observed in Cameroon: *An. nili* A, *An. nili* B and *An. nili* Oveng. Our study included the recently described *An. carnevalei* (Brunhes *et al*., 1999) as well as the rare zoophilic and highly exophilic *An. somalicus* (Rivola & Holstein, 1957), characterized by slight morphological differences at the larval and pupal stages (Gillies & De Meillon, 1968; Gillies & Coetzee, 1987) while their adults resemble typical *An. nili*.

Experimental procedures

Mosquito collection and morphological identification

Specimens of *An. nili* s.l. were collected during surveys in four sites, Afan-Essokie (2°20′ N, 10°00′ E), Mbebe (4°00′ N, 11°02′ E), Oveng (2°24′ N, 10°22′ E) and Simbock (3°55′ N, 11°30′ E), situated in the forested area of southern Cameroon. Specimens from Tibati (6°28′ N, 12°37′ E) and from Burkina Faso, Ivory Coast and Senegal were subsequently used to test for reliability of the PCR assay. Field sampling included both adult and larval collections.

Adult females were collected by landing catches on humans outdoors and indoors from 19.00 to 06.00 hours, and by pyrethrum spraying in selected bedrooms from 16.00 to 17.00 hours, during March and November. Adult specimens were identified morphologically in the field to separate *An. nili sensu lato* from other anopheline species, using morphological keys of Gillies & De Meillon (1968), Gillies & Coetzee (1987) and Brunhes *et al*. (1999). *Anopheles nili s.l.* specimens were then classified into six morphological types: *An. nili* typical, *An. nili* A, *An. nili* B, *An. nili* Congo, *An. nili* Oveng and *An. carnevalei*, according to the size and distribution of the pale fringe spots and pales veins spots of the wing. The typical (T) *An. nili* shows two pale spots on the alar fringe (CuA and M2 nerves). Variants A and B present one additional pale spot on the alar fringe at position (R4 + 5) nerve and CuA nerve, respectively, whereas the Congo form is characterized by the presence of both (R4 + 5) and CuA additional pale spots, and pale scales on the stem of the CuA nerve. A new morphological variant was sampled from the locality of Oveng (in South-western Cameroon) and will be hereafter referred to as *An. nili* Oveng. It is distinct from the Congo form by the presence of a large white base spot affecting the costa, with the CuA largely white. Finally, *An. carnevalei* differs from *An. nili* Congo by the presence of a large white pre-basal spot affecting slightly the costa and all the basal part of R1 nerve; the CuP is largely white and white spots are present on R3 and M1 veins (Brunhes *et al*., 1999).

Anopheles somalicus specimens were morphologically identified at the larval stage, as only aquatic stages of this species bears diagnostic morphological characters (i.e. longer clypeal hairs, see Gillies & de Meillon, 1968). Larvae of *An. somalicus* were found only in Mbebe.

DNA extraction, amplification and sequencing

Genomic DNA was extracted from whole larvae preserved in alcohol, single adult mosquitoes or from adult legs alone, following a slightly modified version of the protocol of Cornel & Collins (1996), and re-suspended in sterile water. ITS2 and D3 regions of the rDNA were amplified from approximately 10 ng of template DNA in 25 µl reaction mixture containing 2.5 µL 10× reaction buffer (Qiagen, France), 1.5 mM $MgCl_2$, 200 µM each deoxynucleotide triphosphate (Eurogentec, Belgium), 0.625 U Taq DNA polymerase (Qiagen, France), 20 pmol each of ITS2a and ITS2b primers for ITS2 region, D3a and D3b primers for D3 domain.

ITS2a and ITS2b were designed from conserved position proximal to the 3′ and 5′ ends of the 5.8S and the 28S-rDNA, respectively (Collins & Paskewitz, 1996; Beebe *et al*., 1999). To amplify the variable D3 domain, D3a and D3b primers were designed in the 28S gene (Koekemoer *et al*., 1999; Sharpe *et al*., 1999). Sequences are as follows:
ITS2a 5′TGTGAACTGCAGGACACAT3′ (forward)
ITS2b 5′TATGCTTAAATTCAGGGGGT3′ (reverse)
D3a 5′GACCCGTCTTGAAACACGGA3′ (forward)
D3b 5′TCGGAAGGAACCAGCTACTA3′ (reverse)

PCR conditions included an initial denaturation step at 94°C for 3 min, followed by 35 cycles of 30 s at 94°C, 30 s at 55°C and 1 min at 72°C with a final extension step of 10 min at 72°C. After amplification, 5 µL of the PCR products were analysed by electrophoresis onto a 1.5% agarose gel containing 0.5 µg/mL ethidium bromide. Bands were revealed and photographed under UV light. After electrophoresis, the remaining PCR products were cleaned using spin columns and used for sequencing in both directions with the previous forward and reverse primers. Using CLUSTER V (Higgins & Sharp, 1988) we compared the sequences of complementary strands for each sample. Sequences were edited using the GCG package (Genetics Computer Group, 1997). Multiple alignment was performed with the PILEUP program of GCG. Genetic distances between haplotypes (Nei, 1987) were computed using MEGA 2.1 (available at http://www.megasoftware.net), under the pairwise deletion

option and using the Kimura 2-parameters correction to take into account multiple hits.

Original sequences of ITS2 and D3 have been deposited in the EMBL database under the following accession numbers: AJ429048 AJ429053 for *An. nili* typical, AJ429049 AJ429052 for *An. nili* Oveng, AJ438689 AJ438690 for *An. somalicus* and AJ429050 AJ429051 for *An. carnevalei*.

Results

Analysis of D3 sequences

The D3 domain of the 28S rDNA was successfully amplified in all specimens belonging to the seven morphological forms of *An. nili* using D3a and D3b primers. A single band of approximately 400 bp was revealed after electrophoresis. PCR products were directly sequenced in both directions using the same primers. The length of the sequenced fragments was 392 bp in *An. nili* typical, *An. nili* A, *An. nili* B and *An. nili* Congo and *An. carnevalei*, 390 bp in *An. somalicus*, and 397 bp in *An. nili* Oveng. At least eight sequences for *An. nili* typical, four sequences for each of *An. nili* A, *An. nili* B and *An. nili* Congo, eight sequences for *An. carnevalei*, five sequences for *An. somalicus* and six sequences for *An. nili* Oveng were aligned together. Four distinct haplotypes were observed among specimens morphologically identified as *An. nili* T ($n = 8$), three of which appeared in single copy in the dataset (haplotype diversity, $h = 0.643$; nucleotide diversity, $\pi = 0.00098$). The most common haplotype (five out of eight sequences, thus representing the consensus sequence for *An. nili* typical) was the only one observed in specimens belonging to *An. nili* A, *An. nili* B and *An. nili* Congo. Similarly, one single (specific) haplotype was observed among specimens of *An. nili* Oveng, *An. carnevalei* and *An. somalicus*, respectively. Sequence data obtained from the D3 domain are given in Fig. 1 and Table 1. Genetic distances between groups were thus computed through pairwise comparisons of consensus sequences (i.e. using one sequence per group). Pairwise genetic distances between haplotypes were low, ranging from 0.005 between *An. carnevalei* and *An. somalicus* to 0.024 between *An. nili* typical and *An. carnevalei* (Table 2).

Analysis of ITS2 sequences

PCR primers used to amplify the ITS2 region anneal to the 5.8S and the 28S coding regions of all species examined so far within the genus *Anopheles* (Porter & Collins, 1991). The ITS2 amplified products varied in size approximately between 500 bp and 600 bp within the *An. nili* group (data not shown). The banding pattern of ITS2 PCR products on agarose gels revealed four molecular entities but size differences were not sufficient to discriminate between forms. Sixty-two female individuals were sequenced: 30 *An. nili* typical, 5 *An. nili* A, 7 *An. nili* B, 3 *An. nili* Congo, 9 *An. carnevalei*, 5 *An. somalicus* and 3 *An. nili* Oveng. The length of the sequenced fragments was 450 bp for *An. nili* T, A, B and Congo, 480 for *An. carnevalei*, 513 bp for *An. somalicus* and 503 for *An. nili* Oveng. Figure 2 shows an alignment of the consensus sequences obtained for each of the seven morphological forms collected in Cameroon. As observed for the D3 domain, sequence comparison between forms revealed four distinct patterns of sequence variation, although low levels of variation were observed within each group (mean nucleotide diversity within group: $\pi = 0.0057 \pm 0.0009$). Variation between haplotypes included nucleotide insertion-deletions and base substitutions. As shown in Fig. 2, *An. nili* typical and *An. nili* A, B and Congo presented highly homogenous consensus sequences. *Anopheles nili* Oveng had a microsatellite (AT) insertion at the 3′ end of the ITS2 fragment that was absent in the other forms. Pairwise genetic distances computed between consensus haplotypes are given in Table 2. Observed values were 7–20 times higher for this region than for the D3 domain and fixed differences between haplotypes provided the basis on which to develop specific PCR primers for an allele-specific PCR assay.

Allele-specific PCR assay

In order to achieve a quick, inexpensive and robust diagnostic tool for rapid identification of field-collected specimens down to their respective molecular form within the *An. nili* group, an allele-specific PCR assay (ASPCR) that combined five primers was developed. Fixed differences between haplotypes encountered in the ITS2 region were used as template to design diagnostic PCR primers (Fig. 2). The strategy we used for ITS2 allele-specific amplification is therefore similar to the approach of Walton *et al.* (1999) to distinguish between members of the *An. dirus* complex, and the approaches of Paskewitz & Collins (1990) and Scott *et al.* (1993) for sibling species identification within the *An. gambiae* complex.

A universal forward primer (ANU) that anneals to the 5′ end of the ITS2 of all *An. nili* s.l. specimens was designed,

```
1    TTCGGAAAAAGG**AC**CAGCTACTAGATGGTTCGATTGGTCTTTCGCCCCTATGCCCAATTCTGACAATCGATTTGCACGTC   80
81   AGAATTGCTTCGGTCCTCCATCAGGGTTTCCCCTGACTTCAACCTGATCAGGCATAGTTCACCATCTTTCGGGTCGCATC  160
161  CTACGCACTCGGGGGATGCCCGCTGGGTGGCGCACGTGTGACCGCACGCCAGCCCGTACCGGGGCACCCTGGGATGGAGG  240
241  GAGGCGTCCGTGGCTTGCGCCAAGCGCGCCCCCGTAATCCCGCGACGAAACCGTCTCGAGTTGTCTGCGCCTGTGGGGTT  320
321  CTC**TCTCGC**G**GTATA**CTGGG**G**CG**CAAG**C**GCCCCAAC**AAC**C**T**GG**CCCATTGGCTCGCGCGTAAGATAGACTTC          392
```

Bold characters: polymorphic site; underlined characters: boundaries of an insertion in non-TFO specimens.

Fig. 1. D3 sequence from *An. nili* typical form.

Table 1. Variable sites in *c*. 400 bp of the D3 PCR product. Bases are numbered relative to *An. nili* typical form (TFO). Dots represent identity to *An. nili* TFO; dashes represent indels.

		Base number																																											
Haplotype	*n*	13	14	324	325	326	327	328	329	331	331a	333	334	335	341	344	345	346	347	349	350	351	352	353	354	355	356	357	357a	357b	357c	358	359	361	362a	362b	362c	362d	362e	362f	362g	362h	362i	*
TFO	8	A	C	T	C	T	C	T	C	G	-		A	T	A	G	C	A	G	C	G	C	C	C	C	A	A	C	A-	-	-	A	C	T	-	-	-	-	-	-	-	-	-	
CFO	4	.	.	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	-	-	-	-	-	-	-	-	-	
AFO	4	.	.	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	-	-	-	-	-	-	-	-	-	
BFO	4	.	.	.	.	.	.	.	.	.	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	-	-	-	-	-	-	-	-	-	
SOM	5	G	A	-	-	-	-	-	-	T	C	G	.	G	T	G	.	C	C	-	-	-	-	G	G	G	.	.	-	G	G	G	T	C	C	-	-	-	-	C	C	A	C	
CAR	8	G	A	-	-	-	-	-	-	.	-	-	.	G	T	G	.	C	C	A	.	-	-	G	G	.	T	G	-	-	G	G	T	C	C	-	C	C	A	C	G	C	C	
OVE	6	-	A	-	-	-	-	-	-	.	-	G	C	.	.	G	G	C	C	A	.	A	.	G	G	.	T	G	C	G	G	G	T	C	C	G	C	C	C	C	A	A	C	

TFO, *An. nili* typical form; CFO, *An. nili* Congo form; AFO, *An. nili* A form; BFO, *An .nili* B form; SOM, *An. somalicus*; CAR, *An. carnevalei*; OVE, *An. nili* Oveng form.

together with four reverse primers ANO, ANC, ANS and ANT, specific to the *An. nili* Oveng (ANO), *An. carnevalei* (ANC), *An. somalicus* (ANS) and to the typical (T) *An. nili*, respectively, including variants A, B and Congo (ANT). Primer sequences were defined so that at least four nucleotides in their 3′ end would prevent hybridization to the alternative haplotypes and guarantee specificity of the assay. Sequences were as follows:

Primer ANU: 5′GATGCACACATTCTTGAGTG CC3′
Primer ANO: 5′AGCACGGTCACCTACGGTTCT CC3′
Primer ANC: 5′CTGGTGGGGTTCTTCTCTTCT CG3′
Primer ANT: 5′TGGCTGCTTCTCGTGGCGCG3′
Primer ANS: 5′ATGCACCAGGGGGTTTGGG CC3′

The Tm was 52°C, 56°C, 55°C, 56°C and 56°C, respectively. Resulting PCR products differ from one another by at least 30 bp, so that they are easily separated on regular agarose gels. The size of the diagnostic band is 188 bp for the typical (T) *An. nili*, 357 bp for *An. nili* Oveng, 408 bp for *An. carnevalei*, and 329 bp for *An. somalicus* (Fig. 3). Hybrid specimens were mimicked by mixing DNA from all combinations prior to amplification. Two bands were obtained in all cases at the expected size (Fig. 4).

Various PCR conditions were tested in order to optimize amplification. PCR mixture consisted of 1.5 mM $MgCl_2$, 200 μM each dNTP (Eurogentec, Belgium), 2.5 μL 10× Taq buffer, 0.625 U Taq polymerase (Qiagen, France) and 10 ng of template DNA in 25 μL final reaction volume. The amount of each primer used in the PCR assay was 40 pmol for ANU, and 10 pmol each for ANT, ANO, ANC and ANS. PCR conditions included an initial denaturation step at 94°C for 5 min, followed by 30 cycles of 30 s at 94°C, 30 s at 63°C and 1 min at 72°C with a final extension step of 10 min at 72°C. The amplified fragments were separated by electrophoresis on a 2% agarose gel.

Validation of the assay on field-collected specimens

To test for the reliability of the ASPCR assay, a total of 226 wild *An. nili* s.l. adult females collected in five localities in South Cameroon were investigated. PCR reactions using the five primers ran on individual specimens produced the expected pattern of size variation. The results confirmed that all the 176 specimens morphologically identified as *An. nili* typical, *An. nili* A, *An. nili* B, or *An. nili* Congo corresponded to typical (T) molecular form. The 30 specimens identified as *An. carnevalei* based on morphological characters unambiguously showed an *An. carnevalei* molecular pattern and all of the 20 specimens identified as *An. nili* Oveng belonged to the Oveng molecular form. Field-collected *An. somalicus* larvae ($n = 15$) and emerging adults ($n = 10$) of *An. somalicus* were correctly identified. It is noteworthy that none of the adult specimens *An. nili s.l.* tested showed the *An. somalicus*-specific PCR product. Moreover, five specimens

Table 2. Pairwise genetic distance between consensus haplotypes of the rDNA ITS2 and D3 regions from south Cameroon populations of the *An. nili* group. Alignment gaps were excluded from the analysis under the pairwise-deletion option and the Kimura 2-parameters correction was applied to take into account multiple hits. Figures above the dashed line are ITS2, figures below the dashed line are D3.

	An. nili typical	*An. nili* A	*An. nili* B	*An. nili* Congo	*An. nili* Oveng	*An. carnevalet*	*An somalicus*
An. nili typical	–	0.002	0.002	0	0.143	0.146	0.144
An. nili A	0	–	0	0.002	0.143	0.143	0.144
An. nili B	0	0	–	0.002	0.143	0.143	0.144
An. nili Congo	0	0	0	–	0.143	0.146	0.144
An. nili Oveng	0.019	0.019	0.019	0.019	–	0.207	0.206
An. carnevalei	0.024	0.024	0.024	0.024	0.010	–	0.143
An. somalicus	0.019	0.019	0.019	0.019	0.013	0.005	–

identified as *An. nili* collected in Burkina Faso, Ivory Coast and Senegal, were tested. All of them belonged to *An. nili* typical (T) molecular form. Sensitivity of the assay therefore was 100% and no hybrid pattern between members of the *An. nili* group was observed.

In tropical Africa, members of the *An. nili* group are usually sympatric with *An. funestus*, members of the *An. gambiae* complex and *An. moucheti* Evans among others. In the conditions of our technique, no PCR product was obtained with 10 specimens of each of these other sympatric species.

Discussion

Distinction between members of the *An. nili* group is presently based mainly on morphological characters. However, these criteria are often difficult to apply because of a number of biological and/or technical issues (e.g. shared overlapping characters, inadequate sampling and preservation of specimens), as for example in Cameroon where seven different morphological forms were observed. Hence, the taxonomic status of each of the reported morphological variants of *An. nili s.l.* remained largely subjective. Precise knowledge of the biology, and thus of the specific importance of members of the group as malaria vectors, has been limited by the absence of reliable diagnostic characters. Our study of sequence polymorphism of the ITS2 and D3 domains of rDNA has overcome this limitation by developing an allele-specific PCR assay which can serve as a reproducible and standardized diagnostic tool for distinguishing between members of the *An. nili* group. Similar approaches have provided powerful diagnostic tools for the study of other anopheline species complexes, such as *An. maculipennis* (Porter & Collins, 1991), *An. gambiae* (Paskewitz *et al*., 1993), *An. quadrimaculatus* (Cornel *et al*., 1996) and *An. funestus* (Koekemoer *et al*., 2002).

Sequence variation in both the rDNA ITS2 and D3 domains among the seven morphological forms of *An. nili s.l.* collected in Cameroon defined four genetic clusters. The first cluster included specimens morphologically identified as typical *An. nili*, variants A and B and the Congo form, whereas the three others included specimens identified as *An. carnevalei*, *An. nili* Oveng and *An. somalicus*, which consistently showed unique ('specific') haplotypes. These results are in partial agreement with morphological features, and suggest that the *An. nili* group may comprise at least four reproductively distinct biological species. Indeed, genetic distances between ITS2 consensus haplotypes of each of these molecular forms were in the range 0.11–0.25, values higher than expected between conspecific populations (Avise, 1994), and similar to those observed within the *An. funestus* group (Hackett *et al*., 2000; Koekemoer *et al*., 2002), between members of the North American *An. quadrimaculatus* complex (Cornel *et al*., 1996) and between the isomorphic South American species *An. trinkae* and *An. dunhami* (Lounibos *et al*., 1998). To date, however, specimens of the different molecular forms of the *An. nili* group have only been found in allopatry, thus preventing further interpretation of their taxonomic status. Putative areas of sympatry are actually being screened and multilocus genetic analyses are under way, using isozymes as well as recently developed microsatellite markers, to ascertain the hypothesis of speciation within the *An. nili* group.

As mentioned above, owing to the limitations of morphological identification methods, very few data are available on biology of the larvae and adults, on the distribution and on the exact role in malaria transmission of the different members of the *An. nili* group. The multiplex PCR that we have developed is simple, sensitive, inexpensive and usable in a laboratory having basic molecular equipment. It can be used on all stages of development, on both sexes (although hybrid males may be impossible to identify if all rDNA repeats were located on the X chromosome, as is the case for *An. gambiae*, see Collins *et al*., 1989) and correct amplification has been obtained from dry as well as alcohol preserved specimens. Moreover, as frequently highlighted as an advantage of PCR-based methods, only a small quantity of biological material (for example a single leg) would contain sufficient template DNA to allow amplification, which leaves the possibility of using the remaining body of the mosquito, even alive, for other studies. Additionally, the technique allows simultaneous testing of a great number of specimens, as usually required during epidemiological investigations. *Anopheles nili* s.l. mosquitoes from Senegal, Burkina Faso, Côte d'Ivoire and Cameroon were successfully identified using this protocol, suggesting that the technique

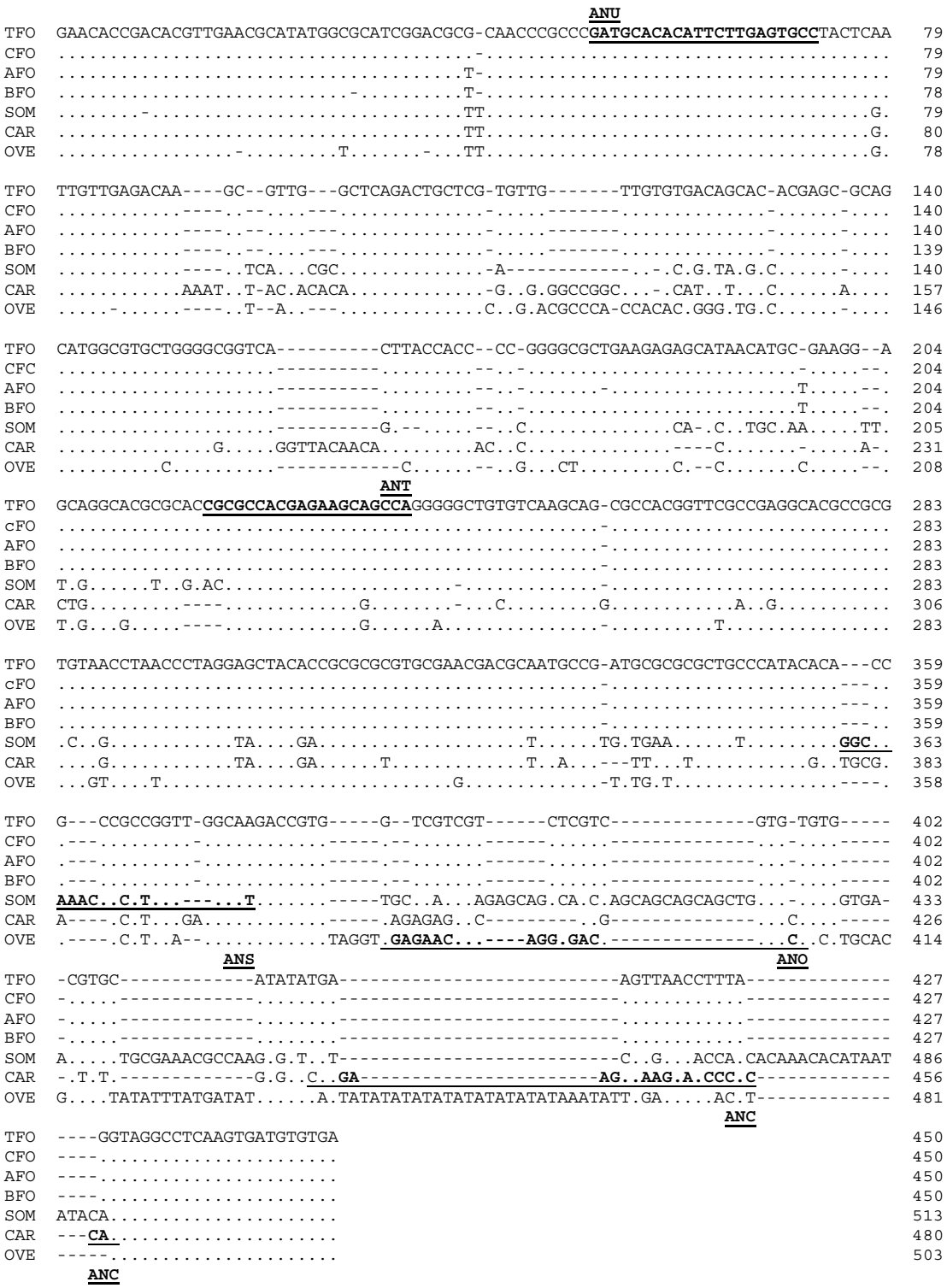

Fig. 2. Alignment of ITS2 consensus sequence from isolates representing each member of the *An. nili* group. The sequences in the alignment represent *An. nili* typical form (TFO), Congo form (CFO), A form (AFO), B form (BFO), *An. somalicus* (SOM), *An. carnevalei* (CAR), and Oveng form (OVE). The dots (·) indicate that the sequence at that point is the same as in *An. nili* Congo. Dash (–) indicates alignment space. Areas from which the primer where designed are underlined. The primers are as follows: ANU (universal primer), ANT (*An. nili* typical, Congo, A and B), ANS (*An. somalicus*), ANO (*An. nili* Oveng) and ANC (*An. carnevalei*).

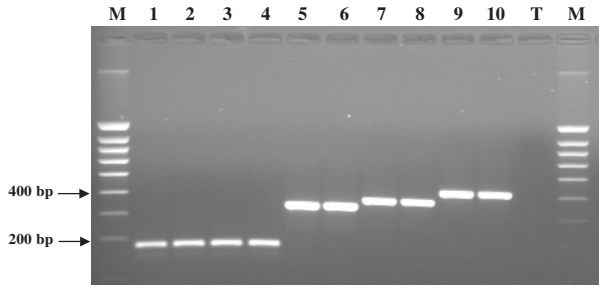

Fig. 3. Photograph of 2% agarose gel showing the amplification products from a single PCR containing the universal primer and the four specific primers. Lane M is 100-bp DNA size marker, lanes 1 = *An. nili* typical, 2 = *An. nili* congo, 3 = *An. nili* A, 4 = *An. nili* B (188 bp), 5–6 = *An. somalicus* (329 bp), 7–8 = *An. nili* Oveng (357 bp) and 9–10 = *An. carnevalei* (408 bp). Lane T is a negative control.

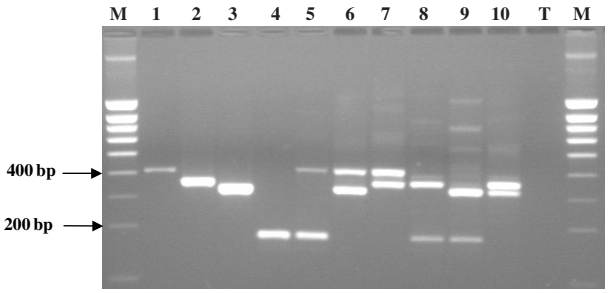

Fig. 4. Ethidium bromide-stained 2% agarose gel showing the PCR products from different molecular form in *An. nili* group and mixtures of DNA to simulate hybrids. Lane M: 100 bp DNA size marker, 1 = *An. carnevalei*; 2 = *An. nili* Oveng; 3 = *An. somalicus*; 4 = *An. nili* typical; 5 = *An. carnevalei* plus *An. nili* typical; 6 = *An. carnevalei* plus *An. somalicus*; 7 = *An. carnevalei* plus *An. nili* Oveng; 8 = *An. nili* Oveng plus *An. nili* typical; 9 = *An. somalicus* plus *An. nili* typical; 10 = *An. nili* Oveng plus *An. somalicus*. Lane T is a negative control. All reaction conditions and the size of PCR products were as described in the text.

has a continent-wide application. Routine application of this new diagnostic test to complement morphological identification (not to be neglected), should help to increase our knowledge of malaria transmission dynamics in areas with vector populations of the *An. nili* group, and contribute to better targeting and implementation of specific and selective strategies for malaria vector control in Africa (Hougard *et al*., 2002).

Acknowledgements

We thank Dr N. Pasteur and the Institut des Sciences de l'Evolution (UMR 5554) of Université Montpellier-2 for their help at several points in this study. We thank the following for supplying specimens: I. Dia (IRD) Senegal, A. Diabate and T. Balde (Centre Muraz) Burkina Faso, F. Chandre and M. Akre (CPR) Ivory Coast. The research was funded by the French Ministry of Research, Programmes PAL + (Grant to P. Kengne) and WHO TDR-REG N°990831 (Grant to P. Awono-Ambene).

References

Arnheim, N. (1983) Concerted Evolution of Multigene families. *Evolution of Genes and Proteins* (ed. by M. Nei and R. K. Koehm), pp. 38–62. Sinauer Associates Inc., Massachussetts.

Avise, J.C. (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Beebe, N.W., Ellis, J.T., Cooper, R.D. & Saul, A. (1999) DNA sequence analysis of the ribosomal DNA ITS2 region for the *Anopheles punctulatus* group of mosquitoes. *Insect Molecular Biology*, **8**, 381–390.

Beebe, N.W. & Saul, A. (1995) Discrimination of all members of the *Anopheles punctulatus* complex by polymerase chain reaction-restriction fragment length polymorphism analysis. *American Journal of Tropical Medicine and Hygiene*, **53**, 478–481.

Brunhes, J., Le Goff, G. & Geoffroy, B. (1999) Afro-tropical anopheline mosquitoes. III. Description of three new species: *Anopheles carnevalei* sp. nov., *An. hervyi* sp. nov. and *An. dualaensis* sp. nov. and resurrection of *An. rageaui* Mattingly and Adam. *Journal of the American Mosquito Control Association*, **15**, 552–558.

Carnevale, P., Le Goff, G., Toto, J.C. & Robert, V. (1992) *Anopheles nili* as the main vector of human malaria in villages of southern Cameroon. *Medical and Veterinary Entomology*, **6**, 135–138.

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

Collins, F.H., Paskewitz, S.M. & Finnerty, V. (1989) Ribosomal RNA genes of *the Anopheles gambiae* species complex. *Advances in Diseases Vector Research*, **6**, 1–28.

Coluzzi, M. (1984) Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. *Bulletin of World Health Organisation*, **62**, 107–113.

Cornel, A.J. & Collins, F.H. (1996) PCR of the ribosomal DNA intergenic spacer regions as a method for identifying mosquitoes in the *Anopheles gambiae* complex. *Methods in Molecular Biology*, **50**, 321–332.

Cornel, A.J., Porter, C.H. & Collins, F.H. (1996) Polymerase chain reaction species diagnostic assay *for Anopheles quadrimaculatus* cryptic species (Diptera: Culicidae) based on ribosomal DNA ITS2 sequences. *Journal of Medical Entomology*, **33**, 109–116.

Fontenille, D. & Lochouarn, L. (1999) The complexity of the malaria vectorial system in Africa. *Parassitologia*, **41**, 267–271.

Genetics Computer Group (1997) *Program Manual for the GCG Package,* v.9.1. Genetics Computer Group, Madison, Wisconsin.

Gillies, M.T. & Coetzee, M. (1987) *A Supplement to the Anophelinae of Africa South of the Sahara*. The South African Institute for Medical Research, Johannesburg.

Gillies, M.T. & De Meillon, B. (1968) *The Anophelinae of Africa South of the Sahara*. The South African Institute for Medical Research, Johannesburg.

Hackett, B.J., Gimnig, J., Guelbeogo, W., Costantini, C., Koekemoer, L.L., Coetzee, M., Collins, F.H. & Besansky, N.J. (2000) Ribosomal DNA internal transcribed spacer (ITS2) sequences differentiate *Anopheles funestus* and *An. rivulorum*, and uncover a cryptic taxon. *Insect Molecular Biology*, **9**, 369–374.

Hamon, J. & Mouchet, J. (1961) Les vecteurs secondaires du paludisme humain en Afrique. *Medecine Tropicale*, **221**, 643–660.

Higgins, D.G. & Sharp, P.M. (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene*, **15**, 237–244.

Hougard, J.M., Fontenille, D., Chandre, F., Darriet, F., Carnevale, P. & Guillet, P. (2002) Combating malaria vectors in Africa: current directions of research. *Trends in Parasitology*, **18**, 283–286.

Koekemoer, L.L., Lochouarn, L., Hunt, R.H. & Coetzee, M. (1999) Single-strand conformation polymorphism analysis for identification of four members of the *Anopheles funestus* (Diptera: Culicidae) group. *Journal of Medical Entomology*, **36**, 125–130.

Koekemoer, L.L., Weeto, M.M., Kamau, L., Hunt, R.H. & Coetzee, M. (2002) A cocktail polymerase chain reaction (PCR) assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *American Journal of Tropical Medicine and Hygiene*, **66**, 804–811.

Lounibos, L.P., Wilkerson, R.C., Conn, J.E., Hribar, L.J., Fritz, G.N. & Donoff-Burg, J.A. (1998) Morphological, molecular and chromosomal discrimination of cryptic *Anopheles* (Nyssorhynchus) (Diptera: Culicidae) from South America. *Journal of Medical Entomology*, **35**, 830–838.

Nei, M. (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Paskewitz, S.M. & Collins, F.H. (1990) Use of the polymerase chain reaction to identify mosquito species of the *Anopheles gambiae* complex. *Medical and Veterinary Entomology*, **4**, 367–373.

Paskewitz, S.M., Wesson, D.M. & Collins, F.H. (1993) The internal transcribed spacers of ribosomal DNA in five members of the *Anopheles gambiae* species complex. *Insect Molecular Biology*, **2**, 247–257.

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

Proft, J., Maier, W.A. & Kempen, H. (1999) Identification of six sibling species of the *Anopheles maculipennis* complex (Diptera: Culicidae) by a polymerase chain reaction assay. *Parasitology*, **85**, 837–843.

Rivola, E. & Holstein, M.H. (1957) Note sur une varieté d'*Anopheles nili* Theo. *Bulletin de la Société de Pathologie Exotique*, **50**, 382–387.

Scott, J.A., Brogdon, W.G. & Collins, F.H. (1993) Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chains reaction. *American Journal of Tropical Medicine and Hygiene*, **49**, 520–529.

Sharpe, R.G., Hims, M.M., Harbach, R.E. & Butlin, R.K. (1999) PCR-based methods for identification of species of the *Anopheles minimus* group: allele-specific amplification and single-strand conformation polymorphism. *Medical and Veterinary Entomology*, **13**, 265–273.

Torres, E.P., Foley, D.H. & Saul, A. (2000) Ribosomal DNA sequence markers differentiate two species of the Anopheles maculatus (Diptera: Culicidae) complex in the Philippines. *Journal of Medical Entomology*, **37**, 933–937.

Walton, C., Handley, J.M., Kuvangkadilok, C., Collins, F.H., Harbach, R.E., Baimai, V. & Butlin, R.K. (1999) Identification of five species of the *Anopheles dirus* complex from Thailand, using allele-specific polymerase chain reaction. *Medical and Veterinary Entomology*, **13**, 24–32.

Accepted 5 November 2002