Am. J. Trop. Med. Hyg., 53(5), 1995, pp. 478–481

DISCRIMINATION OF ALL MEMBERS OF THE *ANOPHELES PUNCTULATUS* COMPLEX BY POLYMERASE CHAIN REACTION–RESTRICTION FRAGMENT LENGTH POLYMORPHISM ANALYSIS

NIGEL W. BEEBE AND ALLAN SAUL

Tropical Health Program, Queensland Institute of Medical Research, Brisbane, Queensland, Australia

Abstract. A method has been developed to identify the members of the *Anopheles punctulatus* complex using polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP). Members of the *An. punctulatus* complex are the most important vectors of malaria in the southwest Pacific and consist of 10 cryptic species, *An. farauti* no. 1–7, *An. punctulatus*, *An.* sp. near *punctulatus*, and *An. koliensis*. For each species, PCR amplification of the ribosomal DNA internal transcribed spacer produced a 750-basepair product. Digestion with *Msp* I and electrophoresis on a 3.0% agarose gel results in banding patterns unique to each species. Isolates of the same species from different locations gave an identical pattern. The technique is sensitive enough so that a PCR-RFLP can be generated from as little as a single mosquito leg, allowing the rest of the mosquito to be used for other important epidemiologic analyses such as determining host feeding source, and for parasite detection.

Members of the *Anopheles punctulatus* complex are the main vectors of malaria throughout Papua New Guinea, Irian Jaya (Indonesia), the Solomon Islands, and Vanuatu. Members of this complex are also important vectors of human filariasis caused by *Wuchereria bancrofti*. There are 10 known members within this group with *An. koliensis*, *An. punctulatus*, and *An. farauti* being identified initially on morphology.[1,2] Cross-mating and polytene chromosomes studies showed *An. farauti* could be separated into at least three species designated *An. farauti* no. 1, 2, and 3.[2–4] Allozyme patterns using 35 loci confirmed this classification and revealed another four isomorphic species of *An. farauti* s.s. (no. 1), *An. farauti* no. 4, 5, 6, and 7, and one species of *An. punctulatus* designated *An.* species near *punctulatus*.[5–7] Overlapping characteristics were also found between *An. farauti* no. 4, *An. koliensis*, and *An. punctulatus*, indicating that morphologic identification is inadequate.[5] Genomic DNA probes have been constructed for all 10 species and have proven to be useful for large scale identification of field-collected specimens.[8–10]

Analysis of the spacer regions in the repeated ribosomal DNA (rDNA) genes has become a common procedure when examining intraspecific variation in Diptera, and for phylogenetic analysis.[11–14] The rDNA transcriptional unit is a tandem repeat separated by a nontranscribed intergenic spacer. Each transcribed unit has two internally transcribed spacers designated ITS1, which separates the 18S and the 5.8S rDNA subunits and ITS2, which separates the 5.8S and 28S rDNA subunits. The function of these ITS regions is unclear but it is thought that they form a hairpin secondary structure that is important in the processing of the transcribed rRNA.[15] The ITS regions seem to undergo a higher mutation rate than transcribed genes and have been used to distinguish closely related species.[12] In the work reported here, the ITS2 sequence divergence has allowed a polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) tool to be developed for discriminating the 10 cryptic species in the *An. punctulatus* complex.

MATERIALS AND METHODS

Mosquito isolates. Mosquito samples were obtained from different locations in Australia, Papua New Guinea, the Solomon Islands, and Vanuatu. Mosquito isolates were either isofemale lines or were collected and identified using allozymes from different locations in the southwest Pacific.[5,6] Table 1 provides a summary of *Anopheles* mosquito isolates and capture locations. Geographic isolates of the same species were sampled at an average of 3–5 specimen per location except for *An. farauti* no. 1 from northern Queensland, which was sampled over 20 times.

Mosquito DNA extraction. Individual mosquitoes were crushed in 1.5-ml microcentrifuge tubes (Eppendorf, Hamburg, Germany) containing 100 μl of lysis buffer (0.2 M NaCl, 10 mM Tris-HCl pH 8, 25 mM EDTA, 0.5% sodium dodecyl sulfate) containing 1.0 mg/ml of proteinase K and then incubated at 55°C for 2 hr prior to being extracted twice with 50 μl of phenol and 50 μl of chloroform:isoamyl alcohol (24:1). The upper aqueous layer was transferred to a new tube and the DNA was precipitated by adding 50 μl of 7.5 M ammonium acetate and 300 μl of ice-cold absolute ethanol. The tubes were then placed at −70°C for 15 min, microfuged at 4°C for 15 min, and then washed in 500 μl of ice-cold 70% ethanol. The pellet was dried and reconstituted in 50 μl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing RNase (100 μg/ml).

Primer selection and design. The primer designated ITS2A was designed as a 19-mer from the 5.8S rDNA of *Drosophila melanogaster* (5′TGTGAACTGCAGGACACAT) and the primer ITS2B was designed from common invertebrate sequences at the 5′ end of the 28S rDNA (5′TATGCTTAAATTCAGGGGGT). The oligonucleotide primers were constructed on an Applied Biosystems (Foster City, CA) 394 DNA/RNA Synthesizer.

Amplification of ITS2. All PCRs were carried out in 0.5-ml microfuge tubes in a 25-μl volume using a Minicycler PTC-150 (MJ Research Inc., Watertown, MA). The final PCR mixture contained 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5% Triton X-100, 1.0 mM $MgCl_2$, 0.2 mM of each deoxynucleotide triphosphate, 50 μM of each primer, 10% dimethylsulfoxide (DMSO), and 2.5 units of *Taq* polymerase. The template was either purified DNA (1–10 ng), 1 μl of allozyme triturate (reconstituted in 20 μl of double-distilled water),[5,7] or from a single leg placed in the PCR. Cy-

TABLE 1

Anopheles mosquito isolates previously identified by allozyme analysis that were used in the polymerase chain reaction–restriction fragment length polymorphism analysis, including localities and coordinate points where known

Species	Location*	Coordinate points
An. farauti no. 1	North Queensland (Aus)	145°00'E 14°47'S
	Sepik region (PNG)	144°04'E 3°50'S
	Espirito Santo (Vanuatu)	Unknown
	Mamara (Solomon Islands)	159°53'E 9°24'S
An. farauti no. 2	Northern Territory (Aus)	132°32'E 12°46'S
	North Queensland (Aus)	145°51'E 18°13'S
An. farauti no. 3	Northern Territory (Aus)	136°08'E 12°30'S
	Northern Territory (Aus)	130°05'E 13°57'S
An. farauti no. 4	Gonoa (PNG)	145°40'E 5°22'S
An. farauti no. 5	Ketarabo (PNG)	145°20'E 6°01'S
An. farauti no. 6	Hiwanda (PNG)	145°20'E 6°01'S
An. farauti no. 7	Mamara (Solomon Islands)	159°53'E 9°24'S
An. koliensis	Sepik region (PNG)	142°20'E 4°20'S
	Mebat (PNG)	145°47'E 5°04'S
	Maraga (PNG)	145°44'E 5°22'S
An. punctulatus	Sepik region (PNG)	142°40'E 4°20'S
	Hudini (PNG)	145°45'E 5°17'S
	Mebat (PNG)	145°47'E 5°04'S
An. sp. near *punctulatus*	Rentoul River (PNG) A	142°38'E 6°22'S
	Rentoul River (PNG) B	142°30'E 6°24'S

* Aus = Australia; PNG = Papua New Guinea.

cling involved an initial denaturation at 94°C for 5 min prior to the addition of *Taq* enzyme and an oil overlay, and then 35 cycles at 94°C for 1 min, 51°C for 1 min, and 72°C for 2 min using minimum transition times.

Product digestion and visualization. A 5-μl aliquot of the PCR mixture was added to water, 2.5 μl of 10× *Msp* I buffer and 10× bovine serum albumin (10 mg/ml) and 1 μl of *Msp* I restriction endonuclease (20 units; New England Biolabs, Beverly, MA) to give a total volume of 20 μl, and the sample was incubated at 37°C for 2 hr. Ten microliters of the digested product was run on a 3% agarose gel (NuSieve GTG; FMC BioProducts, Rockland, ME) containing 0.5 μg/ml of ethidium bromide and visualized at 312 nm on an ultraviolet transilluminator (International Biotechnologies, Inc., New Haven, CT).

RESULTS

After 35 cycles, the PCR amplification products from each species were visualized on a 0.8% agarose gel. All species gave a strong band at 750 basepairs (bp) except *An. farauti* no. 4, which was slightly larger. The amplification product from a single leg placed in the 25-μl reaction was less intense than purified DNA, but still readily observed.

Species-specific banding patterns were generated from an *Msp* I digestion using one-fifth of the PCR mixture and could be visualized on a 3.0% agarose gel containing ethidium bromide. As shown in Figure 1, a 3% agarose gel can separate the *Msp* I-digested ITS2 PCR products for all 10 sibling species. The RFLPs generated were in the range of 500 bp to less than 50 bp. All species had a characteristic RFLP in this range. The bands were well separated with approximately four resolved bands per species.

Isolates of the same species from different geographic locations showed identical RFLPs. Figure 2 shows the intraspecific consistency of the five cryptic species.

DISCUSSION

Sibling species for the *An. punctulatus* complex were readily distinguished using a PCR-RFLP analysis based on the ITS2 region of the rDNA. The mosquito samples did not require a particular storage condition because air-dried samples contain ample template to generate a PCR product whether the DNA was extracted or a segment of the mosquito was used. This would allow the rest of the mosquito to be used for other purposes such as host blood meal source identification or human parasite detection.

The rDNA ITS2 was selected because of its rapid evolution compared with the coding region. In *Drosophila*, it was observed that 60% of the ITS spacer sequence was free to diverge based on the fact that 40% of the sequence could be aligned. These conserved sequences, especially 3′ to the 5.8S rDNA (2S and 5S rDNA in *Drosophila*) and 5′ to the 28S rDNA can pair and form stable stem loops that were also found in distantly related *Drosophila* species.[11] Due to this strong secondary structure, 10% DMSO was used in the amplification reaction to destabilize the DNA secondary structure. This concentration of DMSO is detrimental to the *Taq* polymerase so that more enzyme was required (2.5 units).

It is now apparent that there is sequence variation in this region between separate species, although not within geographic isolates of the same species. The most geographically dispersed species in this complex is *An. farauti* no. 1. The Australian population of *An. farauti* no. 1 was sampled more than 20 times while the other populations were sampled between three and five times. The RFLP result was identical in all samples. It was also fortunate that discrimination required only one 4-bp restriction enzyme (*Msp* I) (CCGG). On average, a 4-bp restriction enzyme should cut every 256 bases (4^4) assuming random basepair arrangements, and therefore cut three times generating four bands. Only *An. farauti* no. 4 showed any obvious difference, con-

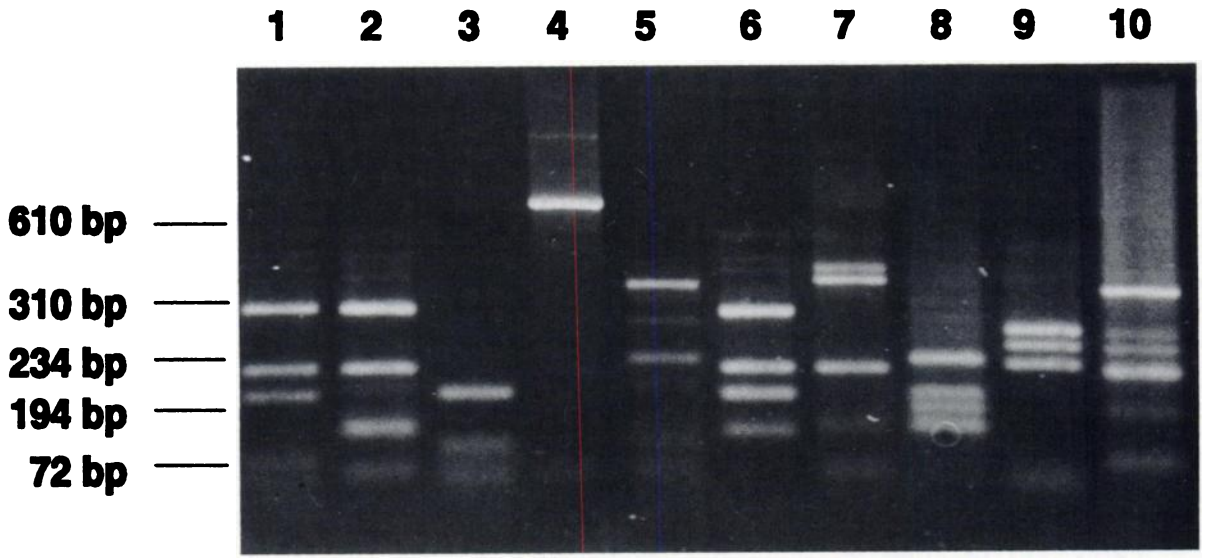

FIGURE 1. Internal transcribed spacer 2 amplification products from each species of the *Anopheles punctulatus* complex digested with *Msp* I and run on a 3% agarose gel. Lane 1, *An. farauti* no. 1; 2, *An. farauti* no. 2 ; 3, *An. farauti* no. 3; 4, *An. farauti* no. 4; 5, *An. farauti* no. 5; 6, *An. farauti* no. 6; 7, *An. farauti* no. 7; 8, *An. koliensis;* 9, *An. punctulatus;* 10, *An.* sp. near *punctulatus.* bp = basepairs.

taining no restriction sites for *Msp* I and a slightly larger amplification product.

The sub-banding present in Figure 1 (lanes 5 and 10) may represent either a coamplified artifact that seems to occasionally appear in the RFLP analysis or heterogeneity in some copies of the ITS2.

The ITS2 length in the *An. punctulatus* complex is larger than in the *An. maculipennis* complex ITS2 region (440 bp), the *An. gambiae* complex (426 bp), and the *An. nuneztovari* cryptic species in South America (363–369 bp). These are the only anopheline ITS2 regions that have been reported.[12-14] We have also found this region to be larger in other local anophelines, as well as in *Anopheles* mosquitoes from the Philippines (unpublished data).

Allozyme and PCR-RFLP analysis of the same field specimens from different locations in the southwest Pacific gave identical species verification, providing strong evidence supporting the species identification based on allozyme frequencies by Foley and others.[5-7]

Several methods are now available for the routine identification of cryptic members of the *An. punctulatus* complex. Allozyme analysis has been the primary method for describing new species in this complex. However, specimens must be fresh or stored frozen to prevent protein degradation; moreover, difficulties in storage arise when working in the field. Also, large numbers of loci are required for statistical significance and electrophoretic standards must be used.[5,7]

In contrast, DNA probes are ideal for large scale identification and specimens can be air-dried or stored in alcohol. Hybridization may require overnight incubation to achieve a yes/no answer for a single species; furthermore, as the number of members in the complex increases, the need to reprobe for extra species also increases. In Papua New Guin-

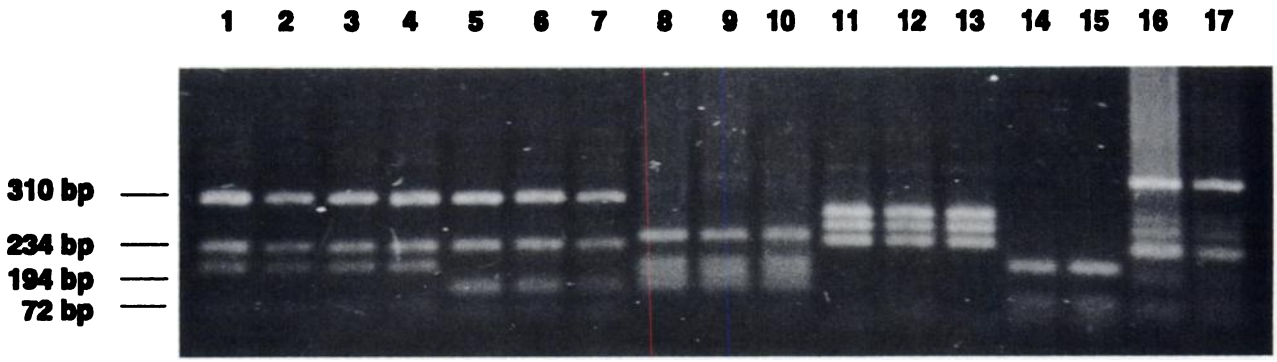

FIGURE 2. Internal transcribed spacer region 2 amplification products from known geographic isolates of of the *Anopheles punctulatus* complex. Lanes 1–4, *An. farauti* no. 1 (1 = North Queensland, Australia; 2 = Sepik Region, Papua New Guinea [PNG]; 3 = Solomon Islands; 4 = Vanuatu); lanes 5–7, *An. farauti* no. 2 (5 = Northern Territory, Australia; 6 = Russel River, Queensland; 7 = north Queensland, Australia); lanes 8–10, *An. koilensis* (8 = Sepik Region, PNG; 9 = Gonoa village, PNG; 10 = Maraga village, PNG); lanes 11–13, *An. punctulatus* (11 = Sepik Region, PNG; 12 = Western Province, PNG; 13 = Irian Jaya, Indonesia); lanes 14 and 15, *An. farauti* no. 3 (14 = Northern Territory, Australia; 15 = Queensland, Australia); lanes 16 and 17, *An.* sp. near *punctulatus* from separate locations around the Rentoul River in PNG (see Table 1 for coordinates). bp = basepairs.

ea, it is common to find at least three different species at one location. For example, around Hudini village near Madang, four species were identified using allozyme analysis.[5] Sequential probing for each species will increase the identification time substantially and limits are imposed by the stability of sample DNA surviving multiple rounds of probing and stripping.

A particular difficulty arises with the DNA probing technique when a mosquito fails to give a strong hybridization signal. In this case, there is no option but to screen with all possible probes. If the sample remains negative, it is difficult to distinguish between a potential new species or a technical failure such as poor sample preparation.

The PCR-RFLP described here provides a third technique. This is particularly important for samples that have failed to hybridize in the squash blot or for comparing field material with reference specimens. Since the primers used are in conserved regions, unrecognized species have a high probability of giving a PCR product, and the difficulties of interpreting negative results in DNA hybridization can now be resolved. This is important to the *An. punctulatus* complex in which new species are still being identified. The PCR-RFLP analysis is accurate, reliable, has minimal storage prerequisites, and a potential for rapid discrimination of species type. Unlike DNA probes, there is a limit to the number of specimens that can be examined by the PCR-RFLP analysis, yet it has the advantage of typing for all species simultaneously. The DNA probing and PCR-RFLP techniques are complementary. We expect hybridization will remain the method of choice for screening large samples and the PCR-RFLP for resolving anomalies or smaller sample sizes. To maximize the advantages, we have modified our own DNA probing procedure by homogenizing mosquitoes in a 96-well plate (using a custom made 96-well homogenizer/applicator) in preference to the standard squash blot to retain a portion of each mosquito for later PCR-RFLP analysis, if required. How applicable this technique will be to other anopheline complexes remains to be determined. The rDNA ITS2 region of five members of the *An. gambiae* complex has been sequenced and there is much less variation;[14] thus, the RFLP approach used for the *An. punctulatus* complex would not work. Primary sequence data indicates the PCR-RFLP approach is displaying more than just restriction sites since some of the observed polymorphism is due to insertions and deletions rather than polymorphisms in restriction sites. Thus, the PCR-RFLP technique is sampling a substantial amount of total sequence information.

In the future, the rDNA ITS2 region of the *An. punctulatus* complex may facilitate alternative assays, such as sequence-specific PCR primers or oligonucleotide probes. The sequence variation and size of the ITS2 region are unlike any other anopheline studied and pose interesting questions on the time of divergence and evolution of the *An. punctulatus* complex.

Acknowledgments: We thank Des Foley and Major Robert Cooper for advice and for supplying material used in this work.

Financial support: This work was supported by the National Health and Medical Research Council of Australia.

Authors' address: Nigel W. Beebe and Allan Saul, Tropical Health Program, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane, Queensland, Australia 4029.

REFERENCES

1. Belkin JN, 1962. *Mosquitoes of the South Pacific.* Volume 1. Berkeley: University of California Press.
2. Bryan JH, 1973. Studies on the *Anopheles punctulatus* complex. II. Hybridisation of the member species. *Trans R Soc Trop Hyg 67:* 70–84.
3. Mahon RJ, Miethke PM, 1982. *Anopheles farauti* no. 3, a hitherto unrecognised biological species of mosquito within the taxon *An. farauti* Laveran (Diptera: Culicidae). *Trans R Soc Trop Med Hyg 76:* 8–12.
4. Mahon RJ, 1983. Identification of the three sibling species of *Anopheles farauti* Laveran by the banding patterns on their polytene chromosomes. *J Aust Entomol Soc 22:* 31–34.
5. Foley DH, Paru R, Dagoro H, Bryan JH, 1993. Allozyme analysis reveals six species within the *Anopheles punctulatus* complex of mosquitoes in Papua New Guinea. *Med Vet Entomol 7:* 37–48.
6. Foley DH, Meek SR, Bryan J, 1994. The *Anopheles punctulatus* group of mosquitoes in the Solomon Islands and Vanuatu surveyed by allozyme electrophoresis. *Med Vet Entomol 8:* 340–350.
7. Foley DH, Cooper RD, Bryan J, 1995. Allozyme analysis reveals a new species within the *Anopheles punctulatus* complex in Western Province, Papua New Guinea. *J Am Mosq Control Assoc 11:* 122–127.
8. Cooper L, Cooper RD, Burkot TR, 1992. The *Anopheles punctulatus* complex: DNA probes for identifying the Australian species using isotopic, chromogenic, and chemiluminescence detection systems. *Exp Parasitol 73:* 27–35.
9. Booth DR, Mahon RJ, Sriprakash KS, 1991. DNA probes to identify members of the *Anopheles farauti* complex. *Med Vet Entomol 5:* 447–454.
10. Beebe N, Foley DH, Saul A, Cooper L, Bryan JH, Burkot TR, 1994. DNA probes for identifying members of the *An. punctulatus* complex in Papua New Guinea. *Am J Med Hyg 50:* 229–234.
11. Shlotterer C, Hauser M-T, von Haeseler A, Tautz D, 1994. Comparative evolutionary analysis of rDNA ITS regions in *Drosophila. Mol Biol Evol 11:* 513–522.
12. Porter CH, Collins FH, 1991. Species-diagnostic differences in a ribosomal DNA internal transcribed spacer from sibling species *Anopheles freeborni* and *Anopheles hermsi* (Diptera: Culicidae). *Am J Trop Med Hyg 45:* 271–279.
13. Paskewitz SM, Weeson DM, Collins FH, 1993. The internal transcribed spacers of ribosomal DNA in five members of the *An. gambiae* complex. *Insect Mol Biol 2:* 247–257.
14. Fritz GN, Conn J, Cockburn A, Seawright J, 1994. Sequence analysis of ribosomal DNA internal transcribed spacer 2 from populations of *Anopheles nunezyovari* (Diptera: Culicidae). *Mol Biol Evol 11:* 406–416.
15. van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raue HA, Planta RJ, 1992. Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. *J Mol Biol 233:* 899–910.