Molecular divergence in the *timeless* and *cpr* genes among three sympatric cryptic species of the *Anopheles triannulatus* complex

Teresa Fernandes Silva-do-Nascimento¹, Luisa Damazio Rona Pitaluga², Ricardo Lourenço-de-Oliveira³

¹Laboratório de Transmissores de Hematozoários ²Laboratório de Biologia Molecular de Insetos, Instituto Oswaldo Cruz-Fiocruz, Av. Brasil 4365, 21045-900 Rio de Janeiro, RJ, Brasil
³Universidade Federal do Rio de Janeiro/Polo de Xerém, Estrada de Xerém 27, Duque de Caxias, RJ, Brasil

Anopheles triannulatus s.l. is a malaria vector with a wide geographic distribution, ranging from Argentina-Nicaragua and Trinidad. Here we analysed sequences of two genes, *timeless* and *cpr*, to assess the genetic variability and divergence among three sympatric cryptic species of this complex from Salobra, central-western Brazil. The *timeless* gene sequences did not conclusively differentiate Anopheles halophylus and *An. triannulatus* species “C”. However, a partial separation has been observed between these species and *An. triannulatus* s.s. Importantly, the analysis of the *cpr* gene sequences revealed fixed differences, no shared polymorphisms and considerable genetic differentiation among the three species of the *An. triannulatus* complex. The results confirm that *An. triannulatus* s.s., *An. halophylus* and *An. triannulatus* species *C* are distinct taxa, with the latter two likely representing a more recent speciation event.

Key words: *Anopheles triannulatus* - *Anopheles halophylus* - *cpr* gene - *timeless* gene - species complex

*Anopheles* (Nyssorhynchus) *triannulatus* s.l. has a wide geographic distribution ranging from Argentina-Nicaragua and Trinidad (Farin & Linthicum 1981, Silva-do-Nascimento & Lourenço-de-Oliveira 2002, Chadee & Wilkerson 2005, Silva-do-Nascimento et al. 2006). Although *An. triannulatus* s.l. is considered zoophilic in most of its territory, it may bite humans indoors and close to houses. It is believed to be a malaria vector when at present at high population densities (Charlwood & Wilkes 1981, Rubio-Palis 1994). Natural infection by *Plasmodium falciparum* and *Plasmodium vivax* s.l. has been reported for *An. triannulatus* s.l. (Oliveira-Ferreira et al. 1990, Tadei & Dutary-Thatcher 2000, Póvoa et al. 2001, Galardo et al. 2007).

Morphological and molecular studies on *An. triannulatus* s.l. from several sites in South America revealed the existence of three distinct sympatric cryptic species only in Salobra, central-western Brazil: *An. triannulatus* s.s. (Netva & Pinto 1922), *Anopheles halophylus* (Silva-do-Nascimento & Lourenço-de-Oliveira 2002) and *An. triannulatus* species *C* (Silva-do-Nascimento & Lourenço-de-Oliveira 2002, 2007, Silva-do-Nascimento et al. 2006). Genetic variability analysis among these sympatric species using allozymes and random amplification of polymorphic DNA (RAPD) detected a barrier to gene flow among them and corroborated their status as separate species and also showed that *An. halophylus* and *An. triannulatus* species *C* are more closely related to each other than to *An. triannulatus* s.s. (Silva-do-Nascimento et al. 2006). In addition, analysis of the population density dynamics in central-western Brazil revealed greater differences between *An. triannulatus* s.s. and *An. halophylus/*An. triannulatus* species *C* than between the latter two species, suggesting that the genetic divergence is correlated with differences in larval exploitation and spatial segregation (Silva-do-Nascimento & Lourenço-de-Oliveira 2007). Additionally, slight differences were observed in the peak time periods of biting activity at sunset among the three sympatric species in Salobra. Roughly, *An. triannulatus* s.s. and *An. triannulatus* *C* peak after sunset, while *An. halophylus* bites almost as frequently early as late in the evening. However, during the peak season, *An. halophylus* and *An. triannulatus* *C* appear to present higher frequencies just after sunset, while *An. triannulatus* s.s. is caught in similar numbers just after sunset and in the evening (Silva-do-Nascimento & Lourenço-de-Oliveira 2007). These interspecific biting time differences are hypothesised to be under the control of the endogenous circadian clock.

The circadian clock controls several important biological activities, including mating rhythms, which may be involved in temporal reproductive isolation between closely related species (Sakai & Ishida 2001, Tauber et al. 2003). The molecular control of circadian rhythms involves several clock genes (Hardin 2005), among them *timeless* (Myers et al. 1995). Rona et al. (2009) found genetic variation in the *timeless* gene that was useful for distinguishing two cryptic species among Brazilian populations in the *Anopheles cruzii* complex of the subgenus *Kerteszia*. Additionally, their results suggested that populations of *An. cruzii* from South and Southeast Brazil also constitute distinct incipient species of this complex.
The cpr gene has been successfully used as a molecular marker in the study of the An. cruzii complex (Rona et al. 2010). cpr encodes the enzyme NADH-cytochrome P450 reductase (CPR) and has been sequenced in several insect species (Koener et al. 1993, Hovemann et al. 1997, Nikou et al. 2003). CPR is encoded by a single-copy gene that is located on chromosome X in Anopheles gambiae (Nikou et al. 2003). The cytochrome P450 gene family is involved in metabolic insecticide resistance and requires CPR to function. Apart from its previously reported functions, cytochrome P450 enzymes were found to play a role in the regulation of sex-specific gene expression in Drosophila melanogaster (Kasai & Tomita 2003). P450 exclusive expression in only one sex can have important roles in hormone metabolism and sexual behaviour (Wen & Scott 2001). Studies conducted on the expression of CPR showed a high mRNA expression in the antenna, suggesting a probable function in olfactory clearance in Drosophila (Hovemann et al. 1997) and in the moth Mamestra brassicae (Maibeche-Coise et al. 2005). Olfactory cues are important environmental stimuli affecting mosquito behaviour, playing significant roles not only in the location of oviposition sites and food sources, but also in mating. Therefore, genes involved in the regulation of antennal responses to pheromones are potentially important in sexual isolation between closely related species.

In the present work, we analysed partial fragments of both timeless and cpr to assess the genetic variability and divergence between three sympatric cryptic species of the An. triannulatus complex in Brazil.

**MATERIALS AND METHODS**

**Mosquito collection** - The mosquitoes used in this study were females from Salobra (20°12'40"S, 56°29'30"W), municipality of Miranda, Mato Grosso do Sul. Blood-fed females were collected from horse bait between 5:00-8:00 pm in 1996 and 1997. Mosquitoes were identified on the basis of morphological characteristics of male genitalia, larva and egg, and allozyme patterns (Consoli & Lourenço-de-Oliveira 1994, Silva-do-Nascimento & Lourenço-de-Oliveira 2002, Silva-do-Nascimento et al. 2006). Six specimens of each species [An. triannulatus s.s. (ATS), An. halophyta (AHSB), An. triannulatus species C (SNSB)] were analysed. Analyses of gene sequences were performed by using a single female from each progeny of each field-collected female (where morphological and molecular analysis of the progeny allowed specific identification).

**DNA extraction, polymerase chain reaction (PCR) and sequencing of the timeless and cpr fragments in An. triannulatus s.l.** - Genomic DNA was extracted individually from females following Jowett (1998). The initial timeless gene fragment was isolated from An. triannulatus s.s. using a pair of specific primers based on a fragment of the Anopheles darlingi timeless gene (forward-3' darltim02a and reverse-3' darltim03) (C Gentile & AA Peixoto, unpublished observations) (Table I). PCR was performed in an Eppendorf Mastercycler thermocycler using the following cycling conditions: 15 cycles at 94°C for 60 s, 55°C (decreasing 1°C/cycle) for 90 s and 72°C for 60 s, followed by 20 cycles of 94°C for 60 s, 50°C for 90 s and 72°C for 60 s. The obtained products were then purified and cloned with the Zero Blunt TOPO PCR cloning it (Invitrogen). Sequencing of positive clones was performed in an ABI Prism 3730 DNA sequencer at the Oswaldo Cruz Institute using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems). The identity of the cloned fragments was confirmed by BlastX analysis using the GenBank database. After checking the sequence identity, two new specific primers named 5'timtri and 3'timtri (Table I) were designed and used to amplify a ~310 bp fragment of the An. triannulatus timeless gene. To obtain fragments of the An. triannulatus cpr gene, a pair of degenerate primers (5'Cpr01deg and 3'Cpr01deg) (Rona et al. 2010) were used (Table I) to amplify a ~210 bp fragment of the An. triannulatus cpr gene. The fragments of both genes include intron sequences.

The primer pairs 5'timtri x 3'timtri and 5'Cpr01deg x 3'Cpr01deg were used to amplify the An. triannulatus populations (Table I). The PCR reactions for all populations were carried out using an Eppendorf Mastercycler thermocycler according to the following cycling conditions: hold at 94°C for 5 min, 35 cycles at 94°C for 30 s, 55°C (timeless) or 60°C (cpr) for 60 s and 72°C for 90 s, using the proofreading Pfu DNA polymerase (Biobots). A negative control template (water) was included in all PCR assays. PCR products were purified using Wizard SV Gel and PCR Clean-Up system (Promega) or QIAquick PCR Purification Kit protocol (Qiagen) and cloned with either the Zero Blunt TOPO PCR (Invitrogen) or pMOSBlue Blunt-Ended Cloning Kit (GE Healthcare). Eight clones of each individual were sequenced as described above to identify possible PCR induced errors and to identify the two alleles. In some cases only one haplotype (SNSB) was observed among the eight sequences and the individuals were considered homozygotes. The probability of classifying a heterozygote as a homozygote is less than 1% using this method.

**Sequence analysis** - DNA sequences were aligned using the Clustal program (Thompson et al. 1997). Analyses of polymorphisms and differentiation among populations were performed using DnaSP (Rozas et al. 2003) and PROSEQ (Filatov & Charlesworth 1999) software. MEGA 3.0 was used to perform the genealogical analysis of the sequences obtained (Kumar et al. 2004).

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences of primers at 5' → 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'darltim02a</td>
<td>GAACGCAGGTCGAGAAGGAATC (forward)</td>
</tr>
<tr>
<td>3'darltim03</td>
<td>TCGTACCGCTTAGCATTT (reverse)</td>
</tr>
<tr>
<td>5'timtri</td>
<td>CATTTCGACCTCGAGCAA (forward)</td>
</tr>
<tr>
<td>3'timtri</td>
<td>CTCTTAAAGGGCGACCTGGA (forward)</td>
</tr>
<tr>
<td>5'Cpr01deg</td>
<td>ATGAARGNATGGTNGCNGA (forward)</td>
</tr>
<tr>
<td>3'Cpr01deg</td>
<td>ATCCARTARAAATTC (reverse)</td>
</tr>
</tbody>
</table>

**TABLE I**

Sequence of primers used to amplify the timeless and cpr gene fragments
RESULTS

A total of 36 nucleotide sequences of the *timeless* gene fragment from the three species of the *An. triannulatus* complex were analysed: 12 from each species (ATSB, AHSB and SNSB). The same number was obtained from each species for the *cpr* locus. The sequences were submitted to GenBank (accessions JN099312 and JN099383). Alignment of the DNA sequences for both genes was performed (Supplementary data) and used in all sequence analysis. All base substitutions were either silent or occurred in the intron.

AHSB and SNSB showed the lowest degree of polymorphism in both *timeless* and *cpr*, with the lower number of polymorphic sites (2 and 1 for *timeless* and 1 and 0 for *cpr*, respectively) (Table II). In addition, no shared polymorphisms and some fixed differences (7 for ATSB x AHSB, 5 for ATSB x SNSB, 2 for AHSB x SNSB) were found in *cpr* (Table II).

Fig. 1 shows a neighbour-joining tree using the *timeless* sequences. The results were not conclusive for the separation between AHSB and SNSB, clustering these two samples. Nonetheless, a partial separation was observed between the sequences of the AHSB/SNSB cluster and ATSB (Fig. 1). In contrast, the neighbour-joining tree obtained for the *cpr* sequences showed a clear separation among the three sympatric populations from Salobra, consistent with the fixation index ($F_{st}$) values (Fig. 2).

DISCUSSION

*An. triannulatus* s.l. displays a high degree of morphological variation (Galvão 1941, Forattini 1962, Silva-do-Nascimento & Lourenço-de-Oliveira 2002). In addition to morphological variation, previous studies have shown that this malaria vector presents polymorphisms in molecular markers and exhibits behavioural variation, suggesting that it is in fact a complex of cryptic species (Santos et al. 1992, 2003, 2004, Rosa-Freitas et al. 1998, Silva-do-Nascimento & Lourenço-de-Oliveira 2002, Silva-do-Nascimento et al. 2005). The taxonomic status of the members of the *An. triannulatus* complex was reevaluated with molecular markers (Silva-do-Nascimento et al. 2006) and morphological characters (Silva-do-Nascimento & Lourenço-de-Oliveira 2002). Silva-do-Nascimento et al. (2006) suggested the existence of at least three distinct taxa occurring in sympatry in Salobra: *An. triannulatus* s.s., *An. halophylus* and *An. triannulatus* species C. The first two have been characterised (Silva-do-Nascimento & Lourenço-de-Oliveira 2002, Silva-do-Nascimento et al. 2006), but the last remains to be formally described.

Some interesting particularities emerged from studies of these three sympatric cryptic species. Subtle differences regarding the peak hematophagy period and larval habitat preferences, as those verified during the collections in Salobra (Silva-do-Nascimento & Lourenço-de-Oliveira 2007), are putative behavioural differences in sympatric incipient species that can reduce gene flow and trigger the speciation process (Diabaté et al. 2005). *Timeless* and *cpr* are genes associated with behaviour and are potentially useful tools in population genetic studies and to separate closely related species. Therefore, we used these two molecular markers to study the genetic differentiation among these three sympatric species of the *An. triannulatus* complex.

Although the *timeless* gene sequences did not differentiate *An. halophylus* and *An. triannulatus* species C, the $F_{st}$ values were moderately high and significant in the other comparisons (ATSB x AHSB and ATSB x SNSB). In addition, when using the *cpr* gene, very high and significant $F_{st}$ values were detected among the sympatric *An. triannulatus* species, revealing considerable genetic differentiation in the *An. triannulatus* complex. In addition, fixed differences and no shared polymorphisms were found in the *cpr* sequences among the three populations. AHSB and SNSB were the least polymorphic populations, showing a very low number of haplotypes. Therefore, considering the small sample sizes, the

### TABLE II

<table>
<thead>
<tr>
<th>Locus</th>
<th>Populations</th>
<th>$F_{st}$</th>
<th>$S_s$</th>
<th>$S_f$</th>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$D_{xy}$</th>
<th>$D_{a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>timeless</em></td>
<td>ATSB x AHSB</td>
<td>0.50000*</td>
<td>00</td>
<td>00</td>
<td>09</td>
<td>02</td>
<td>0.01079</td>
<td>0.00539</td>
</tr>
<tr>
<td></td>
<td>ATSB x SNSB</td>
<td>0.51818*</td>
<td>00</td>
<td>00</td>
<td>09</td>
<td>01</td>
<td>0.01079</td>
<td>0.00559</td>
</tr>
<tr>
<td></td>
<td>AHSB x SNSB</td>
<td>-0.06364</td>
<td>01</td>
<td>00</td>
<td>01</td>
<td>00</td>
<td>0.00179</td>
<td>-0.00011</td>
</tr>
<tr>
<td><em>cpr</em></td>
<td>ATSB x AHSB</td>
<td>0.91103*</td>
<td>00</td>
<td>07</td>
<td>04</td>
<td>01</td>
<td>0.03821</td>
<td>0.03481</td>
</tr>
<tr>
<td></td>
<td>ATSB x SNSB</td>
<td>0.91180*</td>
<td>00</td>
<td>05</td>
<td>04</td>
<td>00</td>
<td>0.02724</td>
<td>0.02483</td>
</tr>
<tr>
<td></td>
<td>AHSB x SNSB</td>
<td>0.90909*</td>
<td>00</td>
<td>02</td>
<td>01</td>
<td>00</td>
<td>0.01098</td>
<td>0.00998</td>
</tr>
</tbody>
</table>

$a$: p < 0.001; AHSB: *Anopheles halophylus*; ATSB: *An. triannulatus* s.s.; $D_{a}$: number of net nucleotide substitutions per site between populations; $D_{xy}$: average number of nucleotide substitutions per site between populations; fixation index ($F_{st}$): pair-wise estimates of population differentiation (significance of $F_{st}$ values evaluated by 1,000 random permutations); $S_s$: polymorphic sites; $S_f$: number of $S$ exclusive to the second population shown in the first column; $S_1$: number of fixed differences between the two populations; $S_2$: number of shared polymorphisms between the two populations.
observed high $F_{ST}$ values have to be regarded with some caution. Nevertheless, as mentioned above, the $F_{ST}$ values are significant in all comparisons, except between AHSB x SNSB using the *timeless* gene.

The lower degree of polymorphism found in *An. halophylus* and *An. triannulatus* species C populations and the higher variability of *An. triannulatus* s.s., for both *timeless* and *cpr*, corroborated previous studies using different markers (allozymes and RAPD) (Silva-do-Nascimento et al. 2006). A study using 11 allozyme loci to analyse the three sympatric populations of the *An. triannulatus* complex from Salobra showed that the enzyme mannose phosphate isomerase is a diagnostic locus (Silva-do-Nascimento et al. 2006). The analysis indicated no gene flow among these populations at this locus, as seems to be the case of *cpr*. The lower divergence found in the analysis of *timeless* might reflect the persistence of ancestral polymorphisms or introgression in some parts of the genome, indicating that a low level of gene flow might still occur or have occurred in the recent past, as observed, for example, in the *An. gambiae* complex (Wang-Sattler et al. 2007).

The present study confirms that *An. triannulatus* s.s. is more polymorphic than both *An. halophylus* and *An. triannulatus* species C that *An. halophylus* and *An. triannulatus* species C are more closely related to each other than to *An. triannulatus* s.s. and, finally, that *An. triannulatus* s.s., *An. halophylus* and *An. triannulatus* species C are distinct taxa, with the last two likely representing a more recent speciation event.

**Fig. 1:** neighbor-joining tree of the *timeless* DNA sequences. The figure shows a neighbour-joining tree based on *Anopheles triannulatus* s.l. sequences of the *timeless* gene obtained with Kimura two-parameters distance. Numbers on the nodes represent the percentage bootstrap values based on 1,000 replications. AHSB: *Anopheles halophylus* (red dot); ATSB: *An. triannulatus* s.s. (green); SNSB: *An. triannulatus* species C (blue).

**Fig. 2:** neighbor-joining tree of the *cpr* DNA sequences. The figure shows a neighbour-joining tree based on *Anopheles triannulatus* s.l. sequences of the *cpr* gene obtained with Kimura two-parameters distance. Numbers on the nodes represent the percentage bootstrap values based on 1,000 replications. AHSB: *Anopheles halophylus* (red dot); ATSB: *An. triannulatus* s.s. (green); SNSB: *An. triannulatus* species C (blue).
REFERENCES


Alignment of the *timeless* DNA sequences of *Anopheles triannulatus*. Alignment of the DNA sequences from the *timeless* gene fragment from all populations of *An. triannulatus* analysed. The translated amino acid sequence is shown above the alignment and the intron sequences are presented in the darkened regions. Dots represent identity to the first nucleotide sequence. AHSB: *Anopheles halophylus*; ATSB: *An. triannulatus* s.s.; SNSB: *An. triannulatus* species C.

Alignment of the *cpr* DNA sequences of *Anopheles triannulatus*. Alignment of the DNA sequences from the *cpr* gene fragment from all populations of *An. triannulatus* analysed. The translated amino acid sequence is shown above the alignment and the intron sequences are presented in the darkened regions. Dots represent identity to the first nucleotide sequence. AHSB: *Anopheles halophylus*; ATSB: *An. triannulatus* s.s.; SNSB: *An. triannulatus* species C.