Speciation in Brazilian Atlantic Forest Mosquitoes: A Mini-Review of the Anopheles cruzii Species Complex

Luísa D.P. Rona¹, Carlos J. Carvalho-Pinto² and Alexandre A. Peixoto³

¹Universidade Federal do Rio de Janeiro / Polo de Xerém, Duque de Caxias - RJ,
²Departamento de Microbiologia e Parasitologia, CCB, Universidade Federal de Santa Catarina, Florianópolis - SC,
³Laboratório de Biologia Molecular de Insetos, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil

1. Introduction

Anopheles (Kerteszia) cruzii s.l. (Diptera: Culicidae) has long been known as the primary vector of human and simian malaria parasites in southern and southeastern Brazil (Deane et al., 1970; 1971; Rachou, 1958). Between 1930 and 1960, An. cruzii together with Anopheles (Kerteszia) bellator and Anopheles (Kerteszia) homunculus were considered the main vectors of malaria once endemic in southern Brazil. Vector control has reduced or even interrupted malaria transmission in some areas, but An. cruzii is still responsible for several oligosymptomatic malaria cases in southern and southeastern Brazil. This mosquito is also a vector of simian malaria in Rio de Janeiro and São Paulo States (Deane et al., 1970). Studies on seasonal and vertical distribution of An. cruzii demonstrated high vertical mobility from ground level to tree tops and this behavior could be responsible for human infection by simian Plasmodium species (Deane et al., 1984; Marrelli et al., 2007; Ueno et al., 2007).

The distribution of this mosquito follows the coast of the Brazilian Atlantic forest (Consoli & Lourenço-de-Oliveira, 1994; Zavortink, 1973), which provides an excellent environment for An. cruzii, since it is an ecosystem abundant in bromeliads, the larval habitat for this anopheline (Pittendrigh, 1949; Rachou, 1958; Veloso et al., 1956). The adults are found in a variety of habitats, from sea level in coastal areas to the mountains. Females are strongly anthropophilic and blood-feed preferably during the evening (Aragão, 1964; Corrêa et al., 1961; Veloso et al., 1956), perhaps biting more than one host to complete egg maturation, which is epidemiologically relevant for malaria transmission (Bona & Navarro-Silva, 2006; Wilkerson & Peyton, 1991). However, notwithstanding its importance as a malaria vector, there are not many population genetic studies of An. cruzii (e.g. Calado et al., 2006; Carvalho-Pinto & Lourenço-de-Oliveira, 2004; Malafronte et al., 2007; Ramirez & Dessen, 2000a,b; see also below).

The possibility that An. cruzii could represent more than one species was first suggested by morphological differences observed among populations from the states of Santa Catarina
and Rio de Janeiro (Zavortink, 1973). Later it was revealed that southern and southeastern Brazilian populations of *An. cruzii* are polymorphic for chromosomal inversions (Ramírez *et al*., 1994; Ramírez & Dessen, 1994). The authors found evidence for the occurrence of genetically distinct *An. cruzii* populations with three different sets of inversions on the X chromosome, defined as forms A, B and C. In populations where two forms are sympatric no heterozygotes were detected, suggesting the absence or limited gene flow between the two groups (Ramírez *et al*., 1994; Ramírez & Dessen, 1994, 2000a,b).

The possibility that *An. cruzii* may represent a complex of cryptic species was also supported by isoenzymatic profiles from 10 distinct loci of several *An. cruzii* populations. This analysis indicated two genetically isolated groups, one from northeastern Brazil (Itaparica Island - Bahia State) and the other from southeastern and southern Brazil (Nova Iguaçu - Rio de Janeiro State, Cananéia - São Paulo State and Florianópolis - Santa Catarina State) (Carvalho-Pinto & Lourenço-de-Oliveira, 2004).

These papers, which proposed that *An. cruzii* is a species complex, led to further studies using molecular markers to investigate the genetic differentiation among populations of this malaria vector. For example, Malafrante *et al* (2007) found some differences between ITS2 sequences comparing a number of southern and southeastern *An. cruzii* populations from Brazil. Similar results were observed by Calado *et al* (2006), using PCR-RAPD and PCR-RFLP of the ITS2 region.

We used a number of single-copy genes to investigate the molecular differentiation and gene flow among the putative sibling species of this complex (Rona *et al*., 2009, 2010a,b). The results and the main conclusions of these analyses are discussed in more detail below.

### 2. Molecular markers and the genetic differentiation among Brazilian populations of *An. cruzii* s.l.

The *timeless* gene is a locus involved in the control of circadian activity rhythms in *Drosophila* (reviewed in Hardin 2005). It also controls mating rhythms (Sakai & Ishida, 2001) and its orthologues in mosquitoes are potentially involved in maintaining temporal reproductive isolation between closely related species. Rona *et al* (2009) isolated a fragment of the *timeless* gene in *An. cruzii* and used it to assess the genetic differentiation among six populations of this malaria vector within its geographic distribution range in Brazil: Florianópolis - Santa Catarina State, Cananéia and Juquitiba - São Paulo State, Itatiaia - Rio de Janeiro State, Santa Teresa - Espírito Santo State and Itaparica Island - Bahia State (Figure 1).

Very strong evidence was obtained for the existence of a different species in Itaparica, a finding that supports the isoenzyme study mentioned above (Carvalho-Pinto & Lourenço-de-Oliveira, 2004). Extremely high $F_{ST}$ values and an elevated number of fixed differences (Table 1) were observed between this northeastern population and the other five studied localities. In addition, the data also suggest that some populations from southern and southeastern regions might also constitute different incipient species. Moderately high $F_{ST}$ values were found when comparing Itatiaia with Florianópolis, Cananéia, Juquitiba and Santa Teresa, suggesting perhaps that this population is in a process of differentiation and incipient speciation (Table 1).
Fig. 1. Localities of the six Brazilian An. cruzii populations studied in Rona et al. (2009). Values in table are approximated distances between localities in km. (Source: IBGE and Google Maps). All mosquitoes used in this study were females captured at the following localities along the Brazilian Atlantic forest: Florianópolis, Santa Catarina State (SC) (27°31'S / 48°30'W), Cananéia and Juquitiba, São Paulo State (SP) (25°01'S / 47°55'W and 23°57'S / 47°03'W), Itatiaia, Rio de Janeiro State (RJ) (22°27'S / 44°36'W), Santa Teresa, Espírito Santo State (ES) (19°56'S / 40°35'W) and Itaparica Island (Jaguaribe), Bahia State (BA) (13°05'S / 38°48'W) (Modified from Rona et al., 2009).
Table 1. Genetic differentiation between An. cruzii populations using the timeless gene. The pair-wise estimates of population differentiation ($F_{ST}$) are shown in the upper right matrix and the numbers of fixed differences between each pair of populations are shown in the lower left matrix of the table. In all cases the $F_{ST}$ values were significant (significance evaluated by 1000 random permutations). The sequences were aligned using ClustalX software (Thompson et al., 1997) and population genetics analysis was carried out using DNASP 4.0 (Rozas et al., 2003) and PROSEQ v 2.91 (Filatov & Charlesworth, 1999) (Modified from Rona et al., 2009).

These results were supported by a Neighbor-joining tree (Figure 2). The An. cruzii sequences from Itaparica (Bahia) were clearly separated in an isolated branch indicating that this northeastern population has diverged significantly from the other populations, in agreement with the isoenzyme analysis (Carvalho-Pinto & Lourenço-de-Oliveira, 2004). In addition, although no clear separation between the timeless sequences from Florianópolis, Cananéia, Juquitiba and Santa Teresa was observed, the sequences from Itatiaia do not appear at a random, showing some clustering. Therefore, a process of incipient speciation seems to be occurring between Itatiaia and the other studied southern and southeastern populations.

To investigate in more detail the genetic differentiation between the southern/southeastern and northeastern siblings of An. cruzii, a multilocus analysis was carried out comparing Itaparica to Florianópolis (Rona et al., 2010a). The aim of this study was to determine if there is still gene flow between the two sibling species and to estimate their divergence time. This analysis was implemented using six loci, three circadian clock genes (timeless, Clock and cycle) and three encoding ribosomal proteins (Rp49, RpS29 and RpS2). As mentioned above, circadian clock genes (Hardin, 2005), such as timeless, Clock and cycle, are putatively involved in the control of mating rhythms and therefore are potentially important in maintaining temporal reproductive isolation between closely related species (Sakai & Ishida, 2001, Tauber et al., 2003). The analysis revealed very high $F_{ST}$ values (ranging from 0.58 to 0.89) and fixed differences between these two cryptic species in all six loci, irrespective of their
function. The divergence time and the migration rate parameters were estimated for all combined loci. Figure 3 shows the posterior probability distributions for each of the three parameters estimated using the IM program. The results suggested that the two species have not exchanged migrants since their separation and that they possibly diverged between 1.1 and 3.6 million years ago (Rona et al., 2010a). In fact, the divergence time between the southern and northeastern species fall within the Pleistocene, a period of intense climatic changes (Cantolla, 2003; Ravelo et al., 2004).

Fig. 2. Neighbor-joining tree using timeless nucleotide sequences of the Anopheles cruzii populations carried out using MEGA 4.0 (Tamura et al., 2007) with Kimura 2-parameters distance. Numbers on the nodes represent the percentage bootstrap values based on 1000 replications. Flo: Florianópolis population; Can: Cananéia; Juq: Juquitiba; Ita: Itatiaia; San: Santa Teresa; Bahia: Itaparica Island population. (Source: Rona et al., 2009).
Fig. 3. Posterior probability distributions for each of the three demographic parameters estimated using IM: divergence time between Florianópolis and Itaparica, and migration rates in both directions. The estimated mutation rate, based on Drosophila, was used to convert the divergence time parameter $t$ to the number of years since population splitting. Four IM simulations using different seed numbers were plotted for each parameter estimate. All curves are shown including the range of the priors. The IM program is an implementation of the Isolation with Migration model and is based on the MCMC (Markov Chain Monte Carlo) simulations of genealogies (Hey & Nielsen, 2004). Initial IM runs were performed in order to establish appropriate upper limits for the priors of each demographic parameter mentioned above. These preliminary simulations generated marginal
distributions that facilitated the choice of parameter values used in the final IM analyses. The convergence was assessed through multiple long runs (four independent MCMC runs with different seed numbers were carried out with at least 30,000,000 recorded steps after a burn-in of 100,000 steps) and by monitoring the ESS values, the update acceptance rates and the trend lines. The Infinite Sites model (Kimura, 1969) was chosen as the mutation model in the IM simulations because the two species are closely related and all genes are nuclear. The optimal recombination-filtered block was extracted from each gene alignment using the IMGC program, which also removes haplotypes that represent likely recombinant sequences (Woerner et al., 2007). See Rona et al. (2010a) for more details. (Modified from Rona et al., 2010a).

Fig. 4. Haplotype network of cpr sequences. Each color represents one population of An. cruzii. Each circle represents a different haplotype with size proportional to its relative frequency. The number of sequences of each haplotype is given in brackets. The small white circles represent missing intermediates and the lines connecting the haplotypes represent one mutational step between two observed haplotypes. Each individual of Itatiaia population is discriminated next to the respective haplotype. The haplotype network was estimated using TCS1.21 (Clement et al., 2000). (Modified from Rona et al., 2010b).
As mentioned above, the analysis of the molecular polymorphism and genetic differentiation of the *timeless* gene among *An. cruzii* populations from southern and southeastern Brazil suggested that the population from Itatiaia (Rio de Janeiro State) is in a process of differentiation and incipient speciation (Rona *et al.*, 2009). To analyze the divergence between these populations, a fragment of the *cpr* gene, a *locus* involved in metabolic insecticide resistance and odorant clearance in insects, was used. High $F_{ST}$ values, some fixed differences and few shared polymorphisms were found between Itatiaia and the other populations (Florianópolis, Cananéia, Juquitiba and Santa Teresa). Moreover, an haplotype network constructed using the *cpr* sequences shows that Itatiaia is clearly separated in an isolated group (Figure 4) suggesting that this population represents a different species in the *An. cruzii* complex (Rona *et al.*, 2010b).

In addition, a more detailed analysis of the Itatiaia *cpr* sequences revealed that this sample might enclose two different sets of individuals. Based on the number of uninterrupted AG repeats found in the intron included in the studied fragment, the Itatiaia population can be divided in two groups: one called Itatiaia A (04 to 06 AG repeats) and the second, called Itatiaia B (03 AG repeats) (Figure 5). In fact, the separation between the two groups is also

<table>
<thead>
<tr>
<th>Sample</th>
<th>AG Repeat Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ita02a</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita02b</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita03a</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita03b</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita04a</td>
<td>AGAGAGAG</td>
</tr>
<tr>
<td>Ita04b</td>
<td>AGAGAGAG</td>
</tr>
<tr>
<td>Ita05a/b</td>
<td>AGAGAG</td>
</tr>
<tr>
<td>Ita06a/b</td>
<td>AGAGAG</td>
</tr>
<tr>
<td>Ita07a/b</td>
<td>AGAGAG</td>
</tr>
<tr>
<td>Ita08a</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita08b</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita09a/b</td>
<td>AGAGAG</td>
</tr>
<tr>
<td>Ita10a/b</td>
<td>AGAGAGAGAG</td>
</tr>
<tr>
<td>Ita11a</td>
<td>AGAGAGAG</td>
</tr>
<tr>
<td>Ita11b</td>
<td>AGAGAGAG</td>
</tr>
<tr>
<td>Ita12a</td>
<td>AGAGAG</td>
</tr>
<tr>
<td>Ita12b</td>
<td>AGAGAGAGAG</td>
</tr>
</tbody>
</table>

Fig. 5. Schematic representation of the AG repeat variable region in the DNA sequences of the *cpr* gene fragment from the Itatiaia population. The sequences of homozygote individuals were grouped and are represented as a/b. The haplotypes with exactly three AG repeats are in red. According to this classification the individuals Ita2, Ita3, Ita4, Ita8, Ita10 and Ita11 belong to Itatiaia A (genotype “4-6/4-6”), the mosquitoes Ita5, Ita6, Ita7 and Ita9 belong to Itatiaia B (genotype “3/3”) and individual Ita12 is the only “hybrid” between the two groups (genotype “3/4-6”). Inspection of the data shows that the Itatiaia sample is not in Hardy-Weinberg equilibrium suggesting the possibility that two sympatric sibling species might exist in this locality. (Modified from Rona *et al.*, 2010b).
evident in the haplotype network of *cpr* sequences shown in Figure 4. Besides, the *F*<sub>ST</sub> value (considering gaps as single mutations) between Itatiaia A and B is quite large (0.67) and highly significant (*P*<0.001) despite the small sample sizes. To confirm, with another locus, the hypothesis that the Itatiaia population might include two incipient sympatric sibling species, the *timeless* data (Rona et al., 2009) from the same sample were reanalyzed. As for the *cpr* data, the *timeless* sequences were divided into Itatiaia A and Itatiaia B. The *timeless* gene also suggests that the sequences might belong to two different sibling species with a highly significant *F*<sub>ST</sub> value (0.34; *P*<0.001) (Rona et al., 2010b).

Further work is clearly needed in this locality and an analysis of a number of other molecular markers might allow a more precise estimate of the differentiation and gene flow between the two putative Itatiaia siblings and between this and other localities in southern Brazil. It will be also important to extend our analyses to a number of other populations along the distribution area of *An. cruzii* as this might provide a more complete representation of the evolutionary history of this species complex. These studies are currently under way.

### 3. Conclusion

In this chapter we reviewed some of our results on *An. cruzii* with emphasis on how the molecular data is providing insights on the evolution of this complex of cryptic species, an example of speciation in Brazilian Atlantic Forest Mosquitoes. Our results and previously published data from other groups suggest that this complex is formed by a number of siblings or incipient species with different levels of genetic divergence and gene flow.

Population genetic studies using molecular markers often revealed complexes of cryptic sibling species in Anopheles mosquitoes with wide geographical distributions (Krzywinska & Besansky, 2003). This is the case of the *An. cruzii* complex, an excellent model for studying ecological vicariance and endemic regions in the Brazilian Atlantic Forest due to its broad geographic range, from southern to northeastern Brazil, and its dependence on forested areas as larval habitat. The appearance of ecological barriers caused by climatic changes as in glaciation periods is a possible explanation for the genetic structure found in this species complex. *An. cruzii* is a forest obligate mosquito and these cooling periods are known to cause forest fragmentation (Cantolla, 2003; Ravelo et al., 2004), which probably affected the distribution of intraspecific lineages and might have split a single ancestral species into isolated groups.

The genetic pattern exhibited by the *An. cruzii* complex is compatible with a historical scenario of populations isolated during the Pleistocene ecological changes (Carnaval et al., 2009). The subdivision of the Brazilian Atlantic Forest has been recognized as a cause of endemcity, for example, in bats (Martins et al., 2009) and pit vipers (Grazziotin et al., 2006) and climatic changes have been proposed to explain the differentiation among many forest-obligate species (Carnaval et al., 2009; Marroig et al., 2004; Pedro et al., 2008).

Understanding the forces that shaped the Brazilian Atlantic Forest diversity is essential to explain the biodiversity of this important and endangered ecosystem and might help the conservation programs selecting the endemic areas that should be considered conservation priorities.
4. Acknowledgment

The authors would like to thank Dr Rosely Malafronte (Instituto de Medicina Tropical de São Paulo), Dr Monique Motta (FIOCRUZ – Rio de Janeiro) and Claudiney dos Santos (Unidade de Medicina Tropical - Universidade Federal do Espirito Santo) for providing most of the An. cruzii samples used in our work. The authors are also indebted to Dr André Pitaluga for helping prepare Figure 1. Our work is supported by grants from the Howard Hughes Medical Institute, FIOCRUZ, Faperj and CNPq.

5. References


