

Novel molecular evidence of population structure in *Anopheles (Kerteszia) bellator* from Brazilian Atlantic Forest

Kamila Voges¹, Marcela Possato Correa da Rosa¹, Betina Westphal-Ferreira²,
Mario Antonio Navarro-Silva², Carime Lessa Mansur Pontes³, André Nóbrega Pitaluga^{4,5,6},
Carlos José de Carvalho-Pinto^{3,6}, Luísa DP Rona^{1,5,6/+}

¹Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Departamento de Biologia Celular, Embriologia e Genética, Florianópolis, SC, Brasil

²Universidade Federal do Paraná, Setor de Ciências Biológicas, Departamento de Zoologia, Curitiba, PR, Brasil

³Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Departamento de Microbiologia, Imunologia e Parasitologia, Florianópolis, SC, Brasil

⁴Fundação Oswaldo Cruz-Fiocruz, Instituto Oswaldo Cruz, Laboratório de Biologia Molecular de Parasitas e Vetores, Rio de Janeiro, RJ, Brasil

⁵Imperial College London, Department of Life Sciences, London, United Kingdom

⁶Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular, Rio de Janeiro, RJ, Brasil

Anopheles bellator is a primary malaria vector in the Atlantic Forest. Partial sequences of *timeless* and *Clock* genes were used to assess the genetic differentiation of five Brazilian populations, which showed strong population structure (e.g. high F_{ST} values and fixed differences) in all pairwise comparisons between Bahia sample and the others from Paraná, São Paulo and Rio de Janeiro states. Also, the resulting phylogenetic trees clearly grouped the sequences from Bahia in a different cluster with high bootstrap values. Among southern and southeastern populations low levels of genetic differentiation were found suggesting a general stability of the genetic structure.

Key words: *Anopheles (Kerteszia) bellator* - bromeliad-malaria - Brazilian Atlantic Forest

Anopheles (Kerteszia) bellator, a mosquito species whose immature stages develop in water accumulated in bromeliads, occurs in Trinidad and Tobago and coastal areas spanning eastern Venezuela to southern Brazil.^(1,2) Interestingly, it is only considered a main human malaria vector in the two extremes of its distribution area,^(3,4) and despite its epidemiological importance, there are few molecular studies concerning this species. Carvalho-Pinto and Lourenço-de-Oliveira⁽⁵⁾ evaluated *An. bellator* isoenzymes in populations from three locations in Brazil [Santa Catarina (SC), São Paulo (SP) and Bahia (BA)], and one from Trinidad Island (Trinidad and Tobago). They concluded that the Trinidad Island population was genetically distant from the Brazilian populations, mostly from BA. Also, regarding the three Brazilian samples, SC and SP were genetically close to each other, and quite distant from BA. Recently, a handful of studies have tried to understand the evolutionary relationship among different species from *Kerteszia* subgenus (including *An. bellator*, *Anopheles cruzii* and *Anopheles homunculus*),^(6,7) rather than investigating the genetic diversity between *An. bellator* populations. Therefore, to date, there have

been no further investigations regarding the relationship between *An. bellator* populations from Brazil.

From the south of Brazil to the north-east, *An. bellator*, *An. homunculus* and *An. cruzii* are sympatric in the Atlantic Forest.^(1,8) However, while *An. homunculus* populations from these regions showed no clear genetic separation — they were in fact considered a single species — *An. cruzii* was shown to be a complex of at least two siblings: one major group with broad range dispersion from south and south-east Brazil and a distinct sibling species in BA.^(8,9,10,11) This knowledge raises an interesting question: are *An. bellator* populations highly structured like those from *An. cruzii*; or, like *An. homunculus*, is it a species without significant genetic differentiation in its wide range distribution? To answer this question and to better understand the genetic variability and divergence among *An. bellator* populations, partial fragments of *timeless* and *Clock* genes were evaluated in five Brazilian locations: Ilha do Mel (PR), Cananéia (SP), Abraão and Sítio Forte (RJ) and Camacan (BA). *Clock* and *timeless* loci are involved in the circadian rhythms of insects,⁽¹²⁾ and have previously been shown to be good molecular markers to define different species in the *An. cruzii* complex.^(10,11,13,14)

MATERIALS AND METHODS

Sampling sites - The mosquitoes used in this study were captured along the Brazilian Atlantic Forest at the following localities: Ilha do Mel (PR), Cananéia (SP), Abraão and Sítio Forte (RJ), and Camacan (BA) (Fig. 1). The maps found in Fig. 1 were made using the packages *maptools*, *maps* and *GISTools*^(15,16,17) from R Software version 3.3.2.⁽¹⁸⁾

doi: 10.1590/0074-02760180598

Financial support: FIOCRUZ (ANP), CNPq (grant no. 442300/2014-0) and Newton International Fellowship / Royal Society (grant no. NF161472) (LDPR).

+ Corresponding author: luisa.rona@imperial.ac.uk; luisa.rona@ufsc.br

https://orcid.org/0000-0002-3400-3950

Received 18 December 2018

Accepted 3 April 2019



An. bellator adults and immatures were obtained, treated and preserved as described in de Rezende Dias et al.⁽¹⁴⁾ The immatures were collected in bromeliads located on rocks, in Restinga areas. Species identification was carried out according to Consoli and Lourenço-de-Oliveira.⁽²⁾ The detailed field collection information, such as sex and the life stage in which the samples were collected (adult or larvae), of each specimen is shown in Supplementary data (Table I).

Molecular data - The *timeless* and *Clock* partial gene sequences from the five *An. bellator* populations were obtained by polymerase chain reaction (PCR), cloning and sequencing as described in de Rezende Dias et al.⁽¹⁴⁾ The primers used in the PCR reactions were those described in Rona et al.^(10,11) The sequencing was carried out in an ABI3130 DNA sequencer at Myleus Biotechnology. For each mosquito, at least eight clones were sequenced to mitigate PCR errors, and allow the identification of the two alleles. Consensus sequences representing the two alleles were generated. The individuals were classified as homozygotes when only one haplotype was observed among the eight sequences. Sequences were submitted to GenBank under the Accession Numbers: MG755641 - MG755734. An alignment of all *An. bellator* sequences for each gene is presented in Supplementary data (Figs. 1,2).

Sequences from the *timeless* gene fragment from *Anopheles (Kerteszia) laneanus* and *Anopheles (Kerteszia) cruzii* s.s. were used in order to verify the relationship of *An. bellator* samples analysed in this study

with other closely related *Kerteszia* species. The sequences from *An. laneanus* were obtained by PCR, cloning and sequencing as described above and submitted to GenBank under the Accession Numbers: KT803975 - KT803978. The sequences from *An. cruzii* s.s. were those previously published by Rona et al. (10) (GenBank accession numbers FJ408732, FJ408735, FJ408746, FJ408786, FJ408788, FJ408813, FJ408816, FJ408832, FJ408736, FJ408738 for SC population and FJ408838, FJ408840, FJ408848, FJ408854, FJ408855, FJ408856, FJ408862, FJ408864 for BA population).

DNA sequence analysis - *Anopheles bellator* DNA sequences were aligned with MAFFT⁽¹⁹⁾ and phylogenetic trees were constructed for each gene under the Maximum Likelihood method using the *ape* and *phangorn* packages^(20,21) in R version 3.3.2.⁽¹⁸⁾ The best-fit substitution models HKY+G (*Clock*) and TIM3e+G+I (*timeless*) were selected following the AIC criterion using the modelTest function from *phangorn*.⁽²¹⁾ An additional *timeless* phylogenetic tree using the coding regions from *An. bellator*, *An. laneanus* and *An. cruzii* s.s. were also constructed (K80+G model) as described above.

DnaSP v5 software⁽²²⁾ was used to calculate intra-population statistics, including nucleotide diversity and number of segregating sites, and also to ensure that all loci met the neutrality model by calculating the Tajima's D.⁽²³⁾ The $P_{RO}S_{EQ}$ v 2.91⁽²⁴⁾ software was used to perform inter-population analyses, including the identification of fixed and shared polymorphisms and the generation of pair-wise estimates of population differentiation (e.g. F_{ST} , D_{xy} and Da).

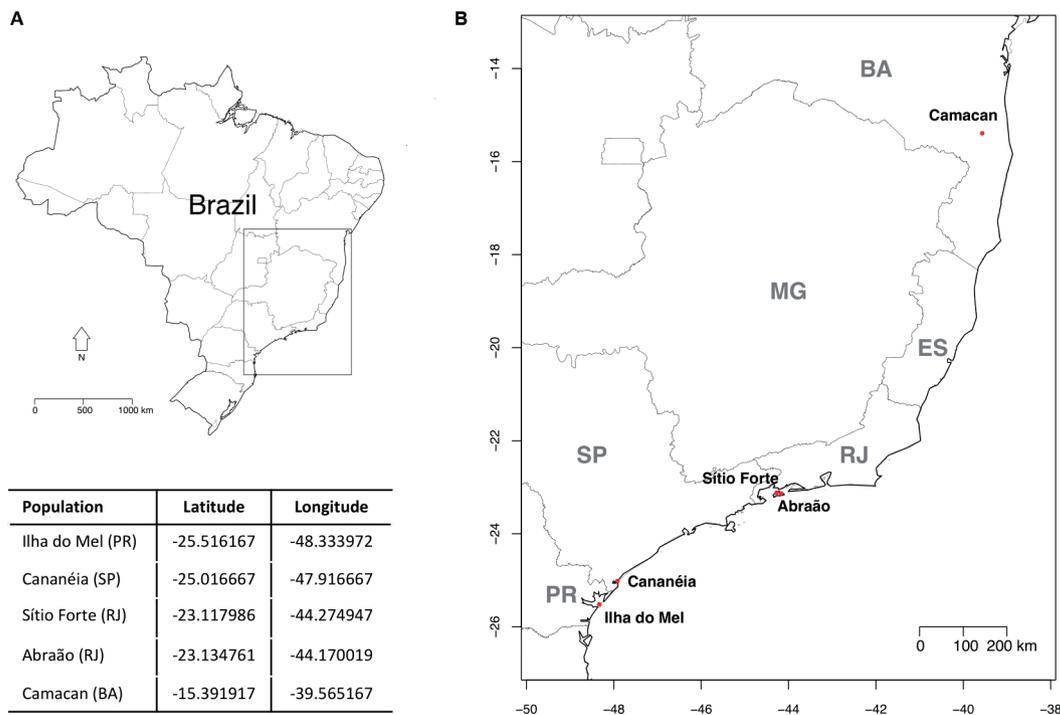


Fig. 1: *Anopheles bellator* collection sites in the Brazilian Atlantic Rainforest. (A): map of Brazil. (B): magnification of the black box from (A) displaying the sites where *An. bellator* samples were collected. The x and y axes drawn in the map (B) display the longitude and latitude, respectively.

RESULTS AND DISCUSSION

Population structure analysis - A total of 94 sequences (two alleles from each *An. bellator* individual) were analysed for the two loci [Supplementary data (Tabela II)]. *Clock* and *timeless* fragment sequences were 164 bp and 368 bp long, respectively [Supplementary data (Figs 1,2)]. Polymorphism data from a variety of measures — e.g. nucleotide diversity (θ and π values) — showed that in general the *timeless* gene has lower genetic diversity compared with *Clock* [Supplementary data (Tabela II)], although the first has a bigger fragment size. None of the Tajima's D results were significant, which means that there is no evidence of natural selection acting in any of the populations [Supplementary data (Tabela II)].

In all pairwise comparisons with the BA population, very high and significant F_{ST} values (ranging from ~ 0.62 to 0.67) were found, there were few shared polymorphisms and a large number of fixed differences (Table), some of which caused amino acid changes [Supplementary data (Figs 1,2)]. The average number of nucleotide substitutions per site (D_{xy}) and the number of net nucleotide substitutions per site between populations (D_a) was also higher in the BA pairwise comparisons, confirming the F_{ST} measurements. The southern and southeastern populations exhibited limited differentiation, characterized by shared polymorphisms and no fixed differences.

Haplotype genealogies analysis - The genealogical relationship of haplotypes was inferred by Maximum Likelihood for the two genes (Fig. 2). The resulting

trees showed no clear separation between the sequences from PR, SP and RJ (henceforth called *An. bellator* A), however they clearly grouped the sequences from BA (henceforth called *Anopheles bellator* B) in a different cluster with high bootstrap values.

In order to verify the genealogical relationship between *An. bellator* A and B with other closely related *Kerteszia* species (*An. cruzii* s.s. and *An. laneanus*), a Maximum Likelihood tree was constructed using the coding regions of the *timeless* gene fragment, since the introns showed a substantial number of indels in the alignment [Supplementary data (Fig. 3)]. The resulting phylogenetic tree showed that *An. bellator* A and B are a monophyletic group, clustering them in a separated clade with high bootstrap value. The final topology was very similar to those from Foster et al.⁽²⁵⁾ and Lorenz et al.⁽⁶⁾ using mitochondrial and genomic sequences: *An. cruzii* s.s. and *An. laneanus* are genetically closer when compared with *An. bellator*. An attempt was also made to compare *An. bellator* *Clock* sequences with those from different *Kerteszia* species available in the GenBank. Although, the same problem found in the *timeless* non-coding regions (indels) was also disclosed in the *Clock* alignment, but in a more challenging way, since more than 70% of the fragment is composed by an intron.

***Anopheles bellator* A and B** - Two highly *An. bellator* structured groups (*An. bellator* A and B) were disclosed in this study. The population divergence between them is well supported by (i) fixed differences (Table), (ii) high F_{ST} values (average of 0.65) (Table), and (iii) it is the earliest divergent node in both phylogenetic trees with strong

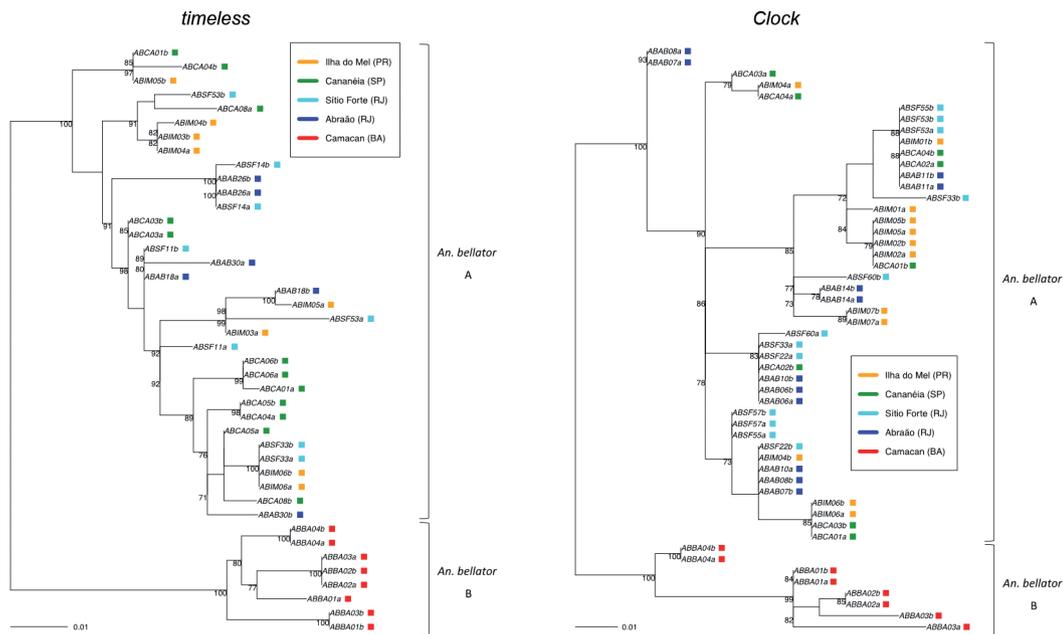


Fig. 2: maximum likelihood trees of *Anopheles bellator* *timeless* and *Clock* sequences (TIM3e+G+I and HKY+G models, respectively). These phylogenies show that two structured groups were identified in the *An. bellator* populations from Brazil: *An. bellator* B occurs in the North-East Atlantic Rainforest (Camacan) and *An. bellator* A is found in the other more Southern Brazilian regions. Numbers on the nodes represent the percentage bootstrap values based on 1000 replications. Lowercase letters in the haplotype names specify the two alleles (A or B) obtained from each individual.

TABLE
Genetic differentiation between all *Anopheles bellator* populations

Populations	Gene	Km	F_{ST}	D_{xy}	D_a	S_s	S_f
Abraão (RJ) x Sítio Forte (RJ)	<i>timeless</i>	10.89	-0.1039	0.0177	-0.0018	11	00
	Clock		-0.0157	0.0322	-0.0005	08	00
Ilha do Mel (PR) x Abraão (RJ)	<i>timeless</i>	496.5	0.0411	0.0215	0.0009	09	00
	Clock		0.1134	0.0404	0.0046	07	00
Ilha do Mel (PR) x Sítio Forte (RJ)	<i>timeless</i>	488.6	0.0695	0.0224	0.0016	11	00
	Clock		0.0904	0.0377	0.0034	08	00
Ilha do Mel (PR) x Cananéia (SP)	<i>timeless</i>	70.1	0.1147	0.0233	0.0027	10	00
	Clock		-0.0508	0.0376	-0.0019	10	00
Cananéia (SP) x Sítio Forte (RJ)	<i>timeless</i>	425.7	0.1158*	0.0221	0.0026	08	00
	Clock		0.0001	0.0364	0.0000	09	00
Cananéia (SP) x Abraão (RJ)	<i>timeless</i>	434.1	0.1311*	0.0222	0.0029	06	00
	Clock		0.0086	0.0383	0.0003	08	00
Camacan (BA) x Ilha do Mel (PR)	<i>timeless</i>	1449	0.6471*	0.0621	0.0402	00	07
	Clock		0.6482*	0.1168	0.0757	02	05
Camacan (BA) x Cananéia (SP)	<i>timeless</i>	1379	0.6694*	0.0622	0.0417	00	09
	Clock		0.6287*	0.1163	0.0731	02	05
Camacan (BA) x Abraão (RJ)	<i>timeless</i>	987.1	0.6722*	0.0628	0.0422	01	10
	Clock		0.6236*	0.1050	0.0655	02	03
Camacan (BA) x Sítio Forte (RJ)	<i>timeless</i>	990.1	0.6758*	0.0641	0.0434	00	09
	Clock		0.6562*	0.1105	0.0725	02	05

F_{ST} : pair-wise estimates of population differentiation. The significance of F_{ST} values was evaluated by 1,000 random permutations (* $p < 0.05$). D_{xy} : average number of nucleotide substitutions per site between populations; D_a : number of net nucleotide substitutions per site between populations.⁽²⁹⁾ S_s : number of shared polymorphisms between the two populations. S_f : number of fixed differences between the two populations. Km: the approximated geographic distances between localities in km.

statistical support (Fig. 2). The results shown here are in agreement with a former study that used allozymes as genetic markers and showed that among all pairwise comparisons between Brazilian *An. bellator* samples, the highest genetic distances were found between Itaparica Island (BA) and the other more southern populations from Florianópolis (SC) and Cananéia (SP).⁽⁵⁾ However, the BA population (Camacan) used in this study is 300 Km away (to the south direction) from Itaparica Island (BA), which raises new questions like: Does *An. bellator* from Camacan belong to the same group as *An. bellator* from Itaparica Island? Also, are there additional structured *An. bellator* groups in the Atlantic Forest? A more extensive sampling across the geographic distribution of *An. bellator*, mainly focusing in Espírito Santo (ES) and BA, and the use of a larger number of molecular markers are desirable in order to better understand the evolutionary forces that establish and maintain the population structure of *An. bellator*. Accordingly, a number of other studies have shown the same separation pattern between northern and southern Atlantic Forest segments in many fauna (including *An. cruzii* s.s.) and flora groups, which converge around northern ES and southern BA, suggesting a common vicariant event.^(10,11,26)

The high genetic differentiation observed between BA and the more southern populations might also be explained by the geographic distance between them (at least 980 Km away) (Table). However, we would expect higher F_{ST} values in the pairwise comparisons between PR and SP, with the two populations from RJ (average F_{ST} of only 0.07), which do not have a negligible distance apart (~ 500 km). Even so, among the southern and southeastern populations, we found a widespread lack of significant differentiation.

This study uncovers two structured *An. bellator* groups in the Brazilian Atlantic Forest: *An. bellator* A is widespread in Southern and Southeastern regions, and *An. bellator* B is found in the north-east Brazil. Interestingly, in spite of the fact that malaria cases are accounted for yearly in BA, *An. bellator* has never been viewed as a malaria vector in this state^(27,28) as it is for the Southern Brazilian regions.⁽²⁾ So, future studies would be valuable in order to differentiate the two *An. bellator* groups regarding malaria susceptibility.

ACKNOWLEDGEMENTS

To Dr Edmundo Carlos Grisard and Dr Patricia Hermes Stoco for their logistic support during this study; to Bibiana Sgorla, Clisten Fátima Staffen and Mari Dalva Staffen for

their technical assistance; to David A Ellis for kindly reviewing the English and to Dr Rosely dos Santos Malafrente and Dr Maria Anice M Sallum for providing *Anopheles bellator* from Cananéia and *Anopheles laneanus*, respectively.

AUTHORS' CONTRIBUTION

LDPR, KV, MPCR and CLMP participated in data generation and analysis; LDPR, BWF, MANS, CJCP and ANP carried mosquito collections and morphological identification; LDPR drafted the manuscript; ANP and CJCP helped in the paper drafting by critically reading the original manuscript; LDPR was the principal investigator, participated in its design and coordination. All authors read and approved the final manuscript.

REFERENCES

- Aragão MB. Geographic distribution and abundance of *Anopheles* species (*Kerteszia*) (Diptera, Culicidae). Rev Bras Malariol Doenças Trop. 1964; 16: 73-109.
- Consoli RAGB, Lourenço-de-Oliveira R. Principais mosquitos de importância sanitária no Brasil. Rio de Janeiro: Fiocruz; 1994. 225 pp.
- Smith LB. Bromeliad malaria. Annual Rep Smithson Inst. 1952; 385-98.
- Rachou RG, Ferreira MO, Lobo AG, Pires WM. Epidemiology of malaria in Southern Brazil. Rev Bras Malariol Doenças Trop. 1955; 6: 177-88.
- de Carvalho-Pinto CJ, Lourenço-de-Oliveira R. Isoenzymatic analysis of four *Anopheles (Kerteszia) bellator* Dyar & Knab (Diptera: Culicidae) populations. Mem Inst Oswaldo Cruz. 2003; 98(8): 1045-8.
- Lorenz C, Patané JS, Suesdek L. Morphogenetic characterisation, date of divergence, and evolutionary relationships of malaria vectors *Anopheles cruzii* and *Anopheles homunculus*. Infect Genet Evol. 2015; 35: 144-52.
- Oliveira TM, Foster PG, Bergo ES, Nagaki SS, Sanabani SS, Marinotti O, et al. Mitochondrial genomes of *Anopheles (Kerteszia)* (Diptera: Culicidae) from the Atlantic Forest, Brazil. J Med Entomol. 2016; 53(4): 790-7.
- Cardoso JC, Bergo ES, Oliveira TMP, Sant'ana DC, Motoki MT, Sallum MAM. New records of *Anopheles homunculus* in Central and Serra do Mar biodiversity corridors of the Atlantic Forest, Brazil. J Am Mosq Control Assoc. 2012; 28(1): 1-5.
- de Carvalho-Pinto CJ, Lourenço-de-Oliveira R. Isoenzymatic analysis of four *Anopheles (Kerteszia) cruzii* (Diptera: Culicidae) populations of Brazil. Mem Inst Oswaldo Cruz. 2004; 99(5):471-5.
- Rona LD, Carvalho-Pinto CJ, Gentile C, Grisard EC, Peixoto AA. Assessing the molecular divergence between *Anopheles (Kerteszia) cruzii* populations from Brazil using the *timeless* gene: further evidence of a species complex. Malar J. 2009; 8: 60.
- Rona LD, Carvalho-Pinto CJ, Mazzoni CJ, Peixoto AA. Estimation of divergence time between two sibling species of the *Anopheles (Kerteszia) cruzii* complex using a multilocus approach. BMC Evol Biol. 2010; 10: 91.
- Sakai T, Ishida N. Circadian rhythms of female mating activity governed by clock genes in *Drosophila*. Proc Natl Acad Sci. 2001; 98(16): 9221-5.
- Rona LDP, Carvalho-Pinto CJ, Peixoto AA. Evidence for the occurrence of two sympatric sibling species within the *Anopheles (Kerteszia) cruzii* complex in southeast Brazil and the detection of asymmetric introgression between them using a multilocus analysis. BMC Evol Biol. 2013; 13: 207.
- de Rezende Dias G, Fujii TTS, Fogel BF, Lourenço-de-Oliveira R, Silva-do-Nascimento TF, Pitaluga AN, et al. Cryptic diversity in an Atlantic Forest malaria vector from the mountains of South-East Brazil. Parasit Vectors. 2018; 11(1): 36.
- Brunsdon C, Chen H. GISTools: some further GIS capabilities for R. R package version 0.7-4. 2014. Available from: <https://CRAN.R-project.org/package=GISTools>.
- Bivand R, Lewin-Koh N. mapproj: tools for reading and handling spatial objects. R package version 0.9-2. 2017. Available from: <https://CRAN.R-project.org/package=mapproj>.
- Becker RA, Wilks AR, Brownrigg R, Minka TP, Deckmyn A. maps: Draw Geographical Maps. R package version 3.2.0. 2017. Available from: <https://CRAN.R-project.org/package=maps>.
- R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2014.
- Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4): 772-80.
- Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004; 20(2): 289-90.
- Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics. 2011; 27(4): 592-3.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 2003; 19(18): 2496-7.
- Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 1989; 123(3): 585-95.
- Filatov D. ProSeq: a software for preparation and evolutionary analysis of DNA sequence data sets. Mol Ecol Notes. 2002; 2: 621-4.
- Foster PG, de Oliveira TMP, Bergo ES, Conn JE, Sant'Ana DC, Nagaki SS, et al. Phylogeny of *Anophelinae* using mitochondrial protein coding genes. R Soc Open Sci. 2017; 4(11): 170758.
- Patterson BD, Costa LP. Bones, clones, and biomes: the history and geography of recent neotropical mammals. Chicago: University of Chicago Press; 2012. 432 pp.
- Rachou RG. Anofelinos do Brasil: Comportamento das espécies vetoras de malaria. Rev Bras Malar Doenças Trop. 1958; 10: 145-81.
- Deane LM. Malaria vectors in Brazil. Mem Inst Oswaldo Cruz. 1986; 81(Suppl. II): 5-14.
- Nei M, Kumar S. Molecular evolution and phylogenetics. New York: Oxford University Press; 2000. 352 pp.