OXFORD

## Morphology, Systematics, Evolution

# Mitochondrial Genomes of *Anopheles* (*Kerteszia*) (Diptera: Culicidae) From the Atlantic Forest, Brazil

T.M.P. Oliveira,[1] P. G. Foster,[2] E. S. Bergo,[3] S. S. Nagaki,[1] S. S. Sanabani,[4] O. Marinotti,[5] P. N. Marinotti,[5] and M.A.M. Sallum[1,6]

[1]Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, Ave. Doutor Arnaldo 715, 01246-904 São Paulo, SP, Brazil (porangaba@usp.br; sayuri20@uol.com.br; masallum@usp.br), [2]Department of Life Sciences, Natural History Museum, Cromwell Rd., London, UK (p.foster2@gmail.com), [3]Superintendência de Controle de Endemias, Secretaria de Estado da Saúde de São Paulo, R. Rui Barbosa, 1672, 14810-095 Araraquara, SP, Brazil (edusteber@uol.com.br), [4] Departamento de Patologia, LIM 03, Hospital das Clinicas (HC), Escola de Medicina, Universidade of São Paulo, Ave. Dr. Enéas de Carvalho Aguiar, 05403-000 São Paulo, Brazil (sabyem_63@yahoo.com), [5] Department of Molecular Biology and Biochemistry, University of California Irvine, 2315 McGaugh Hall, Irvine, CA 92697 (omarinotti@gmail.com; pedromarinotti@gmail.com), and [6]Corresponding author, e-mail: masallum@usp.br

## Abstract

Mitochondrial genome sequences are widely used as molecular markers for phylogenetic studies of mosquito species complexes, such as the *Anopheles albitarsis* complex. Except for a few studies that employed a limited number of nuclear or mitochondrial loci to address the genetic structure and species status of *Anopheles cruzii*, *Anopheles bellator*, and *Anopheles homunculus*, little is known about genetic markers that can be employed in studies focusing on *Kerteszia* species. The complete mitochondrial genomes of seven specimens of *An. bellator*, *An. cruzii*, *An. homunculus*, and *Anopheles laneanus* were sequenced using long-range polymerase chain reaction and Illumina sequencing. The mitochondrial genomes varied from 15,446 to 15,738 bp in length and contained 37 genes (13 protein-encoding genes, 2 rRNA genes [*12S rRNA* and *16S rRNA*] and 22 tRNA genes), and the AT-rich control region, as all do other *Anopheles* mitochondrial genomes sequenced to date. Specimens from four populations of *An. cruzii* showed differences in codon composition.

**Key words**: Anopheles bellator, Anopheles cruzii, Anopheles homunculus, Anopheles laneanus, mitochondrial genome

Malaria is one of the most important vector-borne diseases of people living in tropical and subtropical regions. Six species of the Apicomplexa parasitic protozoans of the genus *Plasmodium* (Sutherland et al. 2010) cause malaria in humans. These *Plasmodium* species can be transmitted by approximately 41 dominant *Anopheles* Meigen species or species complexes worldwide (Sinka et al. 2012). In the neotropical region, nine *Anopheles* species have been associated with the dynamics of human malaria and with transmission of malaria parasites among nonhuman primates (Duarte et al. 2006, 2008; Yamasaki et al. 2011; Sinka et al. 2012).

*Anopheles* (*Kerteszia*) mosquitoes are widely distributed from Mexico to southern Brazil (Zavortink 1973, Marrelli et al. 2007). There are 12 species recognized (Collucci and Sallum 2003); some of them are involved in the dynamics of malaria transmission (Forattini 2002). The following six species of the subgenus *Kerteszia* have been reported in Brazil: *Anopheles bellator* Dyar & Knab, *Anopheles cruzii* Dyar & Knab, *Anopheles homunculus* Komp, occurring along coastal and mountain areas of Serra do Mar in the Atlantic Forest domain (Zavortink 1973, Sallum et al. 2009); *Anopheles laneanus* Corrêa & Cerqueira, found in areas of Serra da Mantiqueira (Sallum et al. 2000); *Anopheles bambusicolus* Komp,

which was reported from inland areas of the Atlantic Forest (Luz et al. 1987); and *Anopheles neivai* Howard, Dyar & Knab, related to the northern Amazon River basin (Hutchings et al. 2005).

Among *Kerteszia* species that occur in Brazil, *An. cruzii*, *An. bellator*, and *An. homunculus* are associated with the transmission of *Plasmodium vivax* (Zavortink 1973, Forattini 2002) in the Southeastern region of Brazil. Recently, specimens of *An. cruzii* from southeastern Atlantic Forest were found naturally infected with *Plasmodium falciparum* (Laporta et al. 2015).

Chromosomal banding pattern of the ovarian polytene chromosomes suggest that there are three putative species under the name *An. cruzii*, *An. cruzii* A, *An. cruzii* B, and *An. cruzii* C (Ramirez and Dessen 1996, 2000). A series of four studies employing various nuclear loci (Rona et al. 2009, 2010a,b, 2013) corroborated this hypothesis. Unfortunately, it is not possible to correlate the differences reported by Ramirez and Dessen (1996, 2000) with those found by Rona et al. (2009, 2010a,b, 2013). This indicates that further studies will be necessary to test and corroborate the hypotheses of *An. cruzii* may be a species complex. *An. homunculus* may be a species complex in South America. Morphological differences in the male genitalia, fourth-instar larva and pupa, and differences in the

ITS2 region of the ribosomal gene suggest that *An. homunculus* may be a species complex (Sallum et al. 2009). Additionally, that *Anopheles anoplus* Komp is currently in synonymy with *An. homunculus* shows the need to address the taxonomic status of these two species.

Mitochondrial and nuclear DNA sequences provided useful markers for species identification (Bourke et al. 2013), inferring phylogenetic relationships within the subfamily Anophelinae (Sallum et al. 2002, Foster et al. 2013), investigating *Anopheles* species complex (Krzywinski et al. 2006, 2011; Moreno et al. 2010; Logue et al. 2013), and *Plasmodium* parasites (Liu et al. 2010, Taylor et al. 2013). However, questions regarding phylogenetic relationships, definitions of species complex, and species identification remain poorly known or even unknown for the majority of the groups of the Anophelinae. Consequently, it is important to have other sources of DNA markers available to show genetic variation and define the limits of species complexes and thus to incriminate species as potential vectors of *Plasmodium*. Therefore, the mitochondrial genome sequences of *An. bellator*, *An. cruzii*, *An. homunculus*, and *An. laneanus* will provide a rich source of markers to strengthen investigations on a group of mosquitoes that includes species that are vectors of malarial parasites.

The objective of this study was to use next-generation sequencing for characterization of the mitochondrial genomes of *An. bellator*, *An. cruzii*, *An. homunculus*, and *An. laneanus*, which will represent a foundation for future studies of *Kerteszia* as well as other Anophelinae. The mitochondrial genomes of seven newly determined sequences of four species of the subgenus *Kerteszia* were compared with those of the *Anopheles albitarsis* complex and *Anopheles darlingi* Root reported to date and publicly available in GenBank.

## Materials and Methods

### Mosquito Sampling

In total, seven mosquitoes were employed in this study: four individuals of *An. cruzii* and one of each species, *An. bellator*, *An. homunculus*, and *An. laneanus*. Individuals of *An. cruzii* are from distinct coastal regions of the Atlantic Forest (Fig. 1; Table 1). The specimens were adults, either offspring of field captured gravid females or from larvae collected from the tanks of bromeliad phytotelmata. Morphological characteristics of the male genitalia, fourth-instar larvae, and pupae were utilized for species identification using the key proposed by Forattini (2002). Adults were stored in 95% ethanol at −80°C until DNA extraction.

### Genomic DNA Isolation

DNA was extracted from each specimen using the animal tissue DNA extraction protocol provided by the QIAgen DNeasy Blood and Tissue Kit (QIAgen Ltd, Crawley, UK). DNA was eluted with 200 μl Buffer AE (10 mM Tris·Cl; 0.5 mM ethylenediaminetetraacetic acid; pH 9.0) and stored at −80°C in the entomological frozen collection of the Faculdade de Saúde Pública, Universidade de São Paulo, Brazil.

### Polymerase Chain Reaction Amplification of the Mitochondrial Genome

The mitochondrial genome of each individual was amplified using two sets of primers (Table 2; Fig. 2). Long-range polymerase chain reaction (PCR) reactions performed using GoTaq Long PCR Master Mix (Promega, WI) and the primers HPK16Saa and HPK16Sbb both proposed by Hwang et al. (2011) amplified a DNA fragment of

approximately 15,300 bp. PCR amplification protocol consisted of an initial denaturation step at 94°C for 2 min, 39 cycles of 94°C for 15 s, 65°C for 20 s, and 65°C for 15 min and a final extension at 72°C for 10 min. PCR amplicons were purified using DNA Clean & Concentrador (Zymo Research, CA) and quantified using the Qubit 2.0 fluorometer (LifeTechnologies, OR). The remaining portion of the mitochondrial genome was amplified by PCR using Platinum Taq DNA polymerase (Invitrogen, CA), with the primers 16Sa and 16Sb, to generate DNA fragment of approximately 650 bp. PCR amplification proceeded under the following temperature regime: 95°C for 3 min, 10 cycles of 95°C for 30 s, 40°C for 1 min and 72°C for 1 min, following 25 cycles of 95°C for 30 s, 50°C for 1 min, and 72°C for 1 min and a final extension at 72°C for 7 min. The PCR amplicons were purified by PEG precipitation (20% polyethylene glycol 8,000/2.5 M NaCl).

### Nextera DNA Sample Preparation and Sequencing

Next-generation sequencing and Sanger technologies were employed to obtain mitochondrial DNA sequences of seven specimens of four species of the subgenus *Kerteszia*. Barcoded libraries were constructed for long-range PCR products using Nextera XT DNA Sample Preparation Kit (Illumina, IL) and sequenced on the Illumina MiSeq platform with paired-end 250 bp chemistry. Sanger technology (Sanger et al. 1997) was adopted to sequence the short PCR amplicons of approximately 650 bp.

### Mitochondrial Genome Assembly and Annotation

Sequences of the mitochondrial genomes were assembled and annotated using the Mira v4 software, and visualized in Tablet (Chevreux et al. 1999, Milne et al. 2013). Mitochondrial genes were annotated using the MITOS web server (Bernt et al. 2013), and protein coding genes verified in the Wise2 v2.2 package using a HMMER2 model (Birney and Durbin 2000, Finn et al. 2011). The boundaries of the rRNAs (16S and 12S) and the tRNAs were determined by sequence similarity to rRNAs and tRNAs of other *Anopheles* species. Secondary structures of tRNAs were predicted within tRNAscan-SE (Lowe and Eddy 1997). The circular map was drawn using the MacVector v.14.0.3 software (MacVector, Inc., Cary, NC).

### GC- and AT-Skews, Nucleotide and Amino Acid Composition, and Relative Synonymous Codon Usage

DNA sequence alignments were performed with the MEGA 6.06 software program (Tamura et al. 2013) using the ClustalW algorithm. The GC- and AT-skews (Perna and Kocher 1995) were employed to address compositional asymmetries in DNA sequences using the following formulas, GC-skew (G-C)/(G + C) and AT-skew (A-T)/(A + T), which give a symmetrical metric from −1.0 to 1.0. As such, a positive GC-skew indicates richness of G over C, whereas a negative GC-skew shows richness of C over G.

To evaluate nucleotide and amino acid composition and potential codon usage (CU) bias (relative synonymous CU [RSCU]) in the set of protein-encoding genes, we employed the MEGA 6.06 software (Tamura et al. 2013). An RSCU value equal or close to 1.0 indicates no CU bias, whereas a value higher than 1.0 indicates that a codon was used more frequently than expected. The opposite is proposed for an RSCU value lower than 1.0. The invertebrate mitochondrial genetic code implemented in MEGA 6.06 (Tamura et al. 2013) was employed to translate the DNA sequences of the protein coding genes into amino acid sequences.
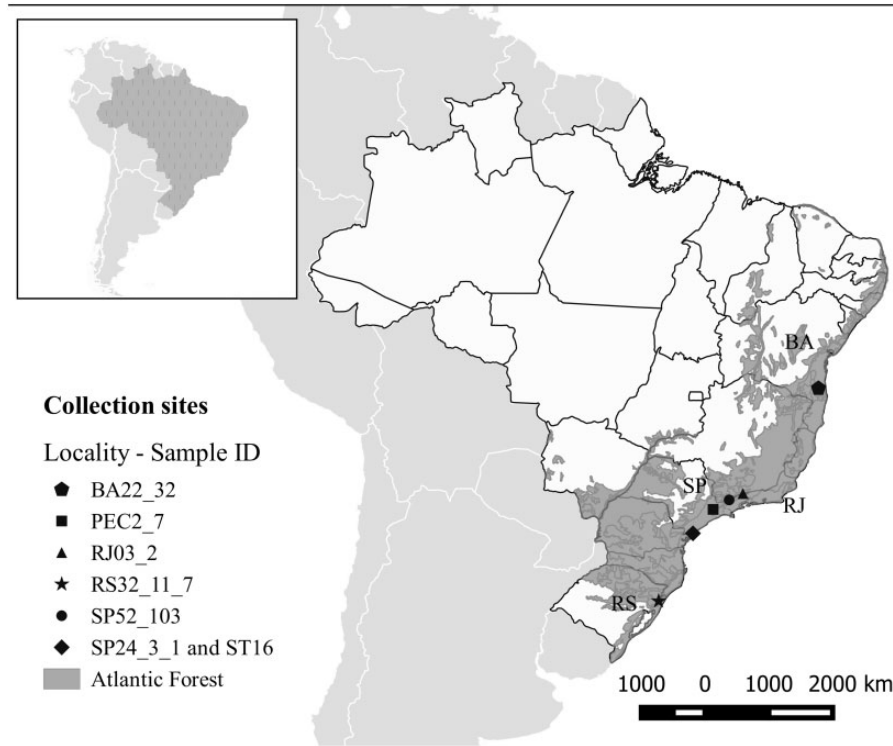
**Fig. 1.** Collection sites of species of *Anopheles* (*Kerteszia*) along coastal regions of the Atlantic Forest, Brazil. The specimens SP24_3_1 (*An. bellator*), ST16 (*An. cruzii*), PEC2_7 (*An. cruzii*) and SP52_103 (*An. laneanus*) were collected in São Paulo state; BA22_32 (*An. homunculus*) and RS32_11_7 (*An. cruzii*) were collected in Bahia and Rio Grande do Sul states, respectively.

**Table 1.** Collection sites, information, and their respective geographical coordinates by state/city in Brazil and GenBank accession numbers of male adults sequenced for this study

| Species | Sample ID | State/city | Latitude | Longitude | GenBank |
|---|---|---|---|---|---|
| *An. bellator* | SP24_3_1 | São Paulo/Cananéia | 25° 1′0″ S | 47° 57′0″ W | KU551287 |
| *An. cruzii* | PEC_2_7 | São Paulo/São Paulo | 23° 24′5″ S | 46° 35′24″ W | KU551284 |
| *An. cruzii* | RJ03_2 | Rio de Janeiro/Itatiaia | 22° 20′0″ S | 44° 35′0″ W | KU551285 |
| *An. cruzii* | RS32_11_7 | Rio Grande do Sul/Maquiné | 29° 35′22.4″ S | 50° 15′45.5″ W | KU551286 |
| *An. cruzii* | ST16 | São Paulo/Cananéia | 25° 1′0″ S | 47° 57′0″ W | KU551289 |
| *An. homunculus* | BA22_32 | Bahia/Camacan | 15° 14′5″ S | 39° 29′45″ W | KU551283 |
| *An. laneanus* | SP52-103 | São Paulo/Pindamonhangaba | 22° 45′50″ S | 45° 30′87″ W | KU551288 |

**Table 2.** Sequences of the primers employed for amplification of mitochondrial genomes using both long-range and short-range PCR approaches

| Primer | Primer sequences (5′–3′) | References | PCR length (bp) |
|---|---|---|---|
| HPK16Saa | ATGCTACCTTTGCACRGTCAAGATACYGCGGC | Hwang et al. (2001) | ~15,300 |
| HPK16Sbb | CTTATCGAYAAAAAAGWTTGCGACCTCGATGTTG | Hwang et al. (2001) | ~15,300 |
| 16Sa | CGCCTGTTTATCAAAAACAT | Simon et al. (1994) | ~650 |
| 16Sb | CCGGTTGAACTCAGATCA | Kambhampati and Smith (1995) | ~650 |

**Measurement of Selection by dN/dS Ratio**

To address the degree to which two homologous coding sequences differ with respect to nonsynonymous sites (d$N$), we employed the average number of nucleotide differences between the sequences per nonsynonymous site. The d$S$ value indicates the degree to which two homologous coding sequences differ at synonymous sites. The KaKs_Calculator 2.0 (Wang et al. 2010) was utilized to address the substitution ratio among homologous protein coding genes applying the MYN model (modified Yang–Nielsen algorithm) proposed by Zhang et al. (2006) for the evolutionary analysis of protein coding sequences. Results were plotted for all pairwise comparisons, and the *dN/dS* rate values were ranked into three groups. The first group (*dN/dS* < 1) included nonsynonymous mutations that were deleterious, i.e., mutations that were under purifying selection. The second group contained nonsynonymous mutations that were largely neutral, thus *dN/dS* = 1. *dN/dS* > 1, indicating that positive selections were driving amino acid substitutions.
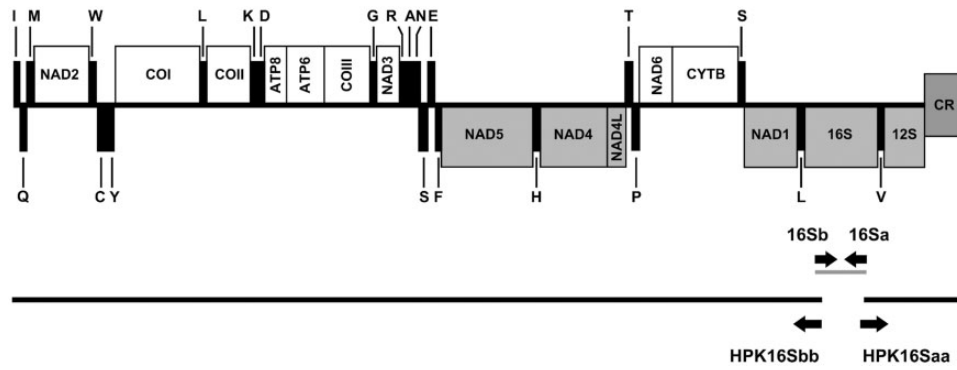
**Fig. 2.** Details of the long (black line) and short (gray line) regions amplified by PCR for either Illumina or Sanger sequencing, respectively. The arrows indicate the primers used to amplify both long and short regions.

### Nucleotide Diversity (π) Using Sliding Window Analyses

To estimate the nucleotide diversity (π) we employed sliding window approach in the DnaSP v.5.10.1 software (Librado and Rozas 2009). A 300-bp window with step size of 10 was utilized to calculate the (π) value. The ClustalW algorithm implemented in MEGA 6.06 (Tamura et al. 2013) was employed to align the mitochondrial DNA sequences; the latter were further contrasted using pairwise comparisons.

## Results

### Features, Organization, and Nucleotide Content of the Mitochondrial Genome

Using two PCR amplification strategies, we obtained the complete mitochondrial genome of seven individuals of four species of the subgenus *Kerteszia* of *Anopheles* (*An. cruzii*, *An. bellator*, *An. homunculus,* and *An. laneanus*). The primer pair HPK16Saa and HPK16Sbb (Table 2) amplified one region of ∼15 kb of the complete mitochondrial genome by long-range PCR. A second primer pair (16Sa and 16Sb) (Table 2) amplified a ∼650 bp fragment. The annotated mitochondrial genomes were deposited in GenBank with accession numbers in Table 2.

The mitochondrial genomes of *Kerteszia* mosquitoes consisted of circular DNA molecules and contained 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes (*12S rRNA* and *16S rRNA*), and an AT-rich control region (Supp Tables S1–S7 [online only]; Fig. 3). The complete mitochondrial genome of *An. cruzii* RS32_11_7 and *An. laneanus* were 15,446 bp, whereas those of *An. cruzii* PEC_2_7, *An. cruzii* ST16, *An. cruzii* RJ03_2, *An. bellator*, and *An. homunculus* varied from 15,451 to 15,738 bp. Except for variation in the length of the control region, the mitochondrial genes showed no length variation. Genes were located on either the H or the L strand (Fig. 3) and some open reading frames overlapped 1–9 nucleotides to adjacent genes (Supp Tables S1–S7 [online only]). The AT-rich control region for *Kerteszia* species was located between the *12S rRNA* and *tRNAIle* genes (Fig. 3).

The mean nucleotide content of coding strand of the mitochondrial genome sequence of *Kerteszia* species consisted of 39.9% A, 12.6% C, 9.0% G, and 38.5% T (Supp Table S8 [online only]). The length of the AT-rich region varied from 591 bp in *An. laneanus* SP52_103 (Supp Table S7 [online only]) to 885 bp in *An. homunculus*_BA22_32 (Supp Table S6 [online only]). The overall AT-content of the mitochondrial genome sequence of *Kerteszia* species varied from 78.18 to 78.56% (Supp Table S8 [online only]), and the mean overall GC-skew of the H-strand was −0.166, whereas the AT-skew was 0.017.

### Protein-Coding Genes

The mitochondrial genome of *Kerteszia* species consisted of 13 intron-less protein-coding genes. The H-strand encoded the *NAD2, COI, COII, ATP8, ATP6, COIII, NAD3, NAD6,* and *CYTB* genes, whereas the L-strand encoded the *NAD5, NAD4, NAD4L,* and *NAD1* genes. The start codons followed the ATN rule, except for the *COI* gene (Supp Tables S1–S7 [online only]), which starts with TCG. The start codons were identical for all sequences studied, except for *ATP8* and *NAD6* genes (Supp Tables S1–S7 [online only]). A complete stop codon (TAA) was observed in the *ATP6, ATP8, NAD1, NAD2, NAD3, NAD4, NAD4L, NAD5, NAD6,* and *CYTB* genes, whereas *COI, COII, COIII* genes showed a partial stop codon (T_). Taken together, the protein-coding genes summed to 11,238 bp. The amino acids *Leu, Ile,* and *Phe* were the most abundant (Supp Fig. S1 [online only]).

### CU and RSCU

The values of CU are in Supp Table S9 (online only). Sequences of the protein-coding genes of the mitochondrial genome of *Kerteszia* species are composed of 3,736 codons. The majority of codons ended with either A or T showed values of RSCU higher than 1.0 (Supp Fig. S2 [online only]), indicating that these codons are used with a higher than expected frequency. Among the codons, AGG was absent in all sequences of *Kerteszia* species, except in *An. cruzii* RJ03_2. Specimens of *An. cruzii* showed no RSCU pattern.

### Ribosomal and Transfer RNA Genes

All 22 tRNA genes were located on the H- and L-strand, and the small (*rrn*S) and large (*rrn*L) subunits of rRNA were located on the L-strand. Predicted secondary structures of tRNAs are in the Supp Fig. S3 (online only). All except for the two *tRNASer1* and *tRNAArg* genes had their secondary structure predicted herein. Similarly, we could not predict a secondary structure for the *tRNAPhe*-coding gene of *An. cruzii* RJ03_2, RS32_11_7, ST16 and for *An. laneanus*, and for the *tRNATyr*-coding gene of *An. homunculus*. The predicted secondary structure contained an acceptor stem, a D-loop arm, an anticodon loop arm, a variable loop, and an arm with a TψC loop. The predicted secondary structure of TψC arm of *tRNATrp* showed two loops.

**Fig. 3.** Graphic representation of the gene arrangement and gene order of the mitochondrial genome of four species of *Anopheles* (*Kerteszia*). All 13 protein-coding genes, 22 tRNA genes, 2 rRNA genes (*12S rRNA* and *16S rRNA*), and the AT-rich control region are indicated in italics. Each tRNA gene is identified by its three letter abbreviation. The direction of transcription is indicated by arrows.

**Noncoding Regions**

The mitochondrial genomes of the *Kerteszia* species possessed either small or no intergenic spacers. The longest intergenic spacer was the AT-rich control region with a length that varied from 591 to 886 bp.

**Measurement of Selection by dN/dS Ratio**

Measurement of selection showed *dN/dS* values that ranged from 0.0 to 0.117 for the *Kerteszia* species under study. These values showed that natural selection is suppressing protein changes in *Kerteszia* species and that all mitochondrial protein-coding genes are under negative selection pressure. The values of the *dN/dS* ratio obtained for the *COX1, COX2, COX3, CYTB*, and *NAD1* showed

that the negative selection pressure acting on these genes were weaker than the negative selection pressure on *NAD2, NAD6*, and *ATP8* genes (Supp Fig. S4 [online only]).

**Nucleotide Diversity (π)**

Nucleotide diversity (π) along the mitochondrial DNA sequences showed the degree of polymorphism among individuals of distinct species of *Kerteszia*. The π values ranged from 0.0025 to 0.1104 (Fig. 4), with the highest values observed for the AT-control region and the lowest values for the rRNA-coding genes. Results of pairwise comparison among specimens of *An. cruzii* showed that RS32_11_7 and ST16 had the lowest nucleotide diversity, with the
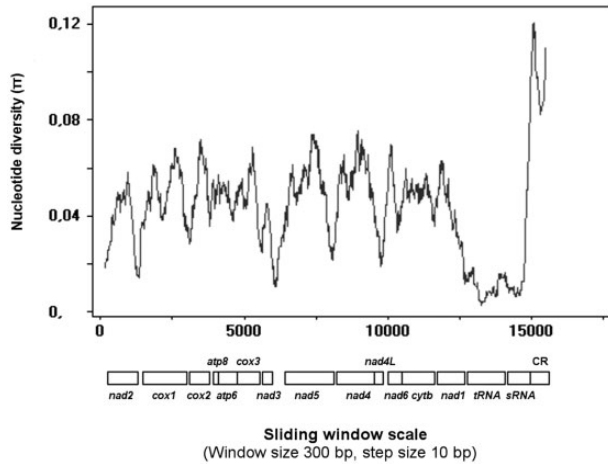
**Fig. 4.** Nucleotide diversity (π) estimated for the complete mitochondrial genome of four species of *Anopheles* (*Kerteszia*). The (π) values were estimated using a sliding window analysis employing the alignment of DNA sequences of seven individuals of four *Kerteszia* species. The size of the sliding window was 300 bp. The bars below the graphic represent the protein coding genes, the ribosomal genes and control region (CR).

*NAD5* gene as the most polymorphic region (Supp Fig. S5 [online only]). In addition, PEC_2_7 showed the highest nucleotide diversity when compared with individuals of other populations of *An. cruzii* (Supp Fig. S5 [online only]).

## Discussion

As in the majority of metazoans (Bernt et al. 2013, Osigus et al. 2013), the mitochondrial genomes of *Kertezsia* species contain 37 genes and an AT-rich control region. The gene order, boundaries, and arrangements (Figs. 2 and 3) in *Kerteszia* species were mostly consistent with the mitochondrial genomes of other *Anopheles* species characterized to date (Beard et al. 1993; Krzywinski et al. 2006, 2011; Moreno et al. 2010). Presence of small intergenic regions, genes encoded on both strands, and overlapped adjacent genes produced a compact mitochondrial genome in *Kerteszia* species as in other *Anopheles* species (Krzywinski et al. 2006, 2011; Moreno et al. 2010; Logue et al. 2013).

The mitochondrial genome of *Kerteszia* species has a clear strand bias in the nucleotide composition as do other species of *Anopheles* characterized to date. This is likely explained by an asymmetric pattern of mutation that occurred during transcription and replication processes, causing the strand bias (Francino and Ochman 1997). The values of the AT-content found in *Kerteszia* species are more similar to that reported for *An. darlingi* (Moreno et al. 2010) than for species of the *An. albitarsis* complex (Krzywinski et al. 2011). Consequently, the CU reflects the high AT-content, with a deficiency of codons containing either C or G (Supp Table S9 [online only]). Thus, the CU in protein-coding genes favored those ending with A or T residues (Supp Fig. S2 [online only]) over those with C or G.

The GC- and AT-skew has been widely invoked to account for compositional replication strand bias in protein-coding genes (Perna and Kocher 1995). A negative value of the GC-skew shows an excess of C over G on the leading strand, whereas a positive value indicates richness of G over C. The excess of C over G in the leading strand produces asymmetries (Wei et al. 2010). Results of the GC- and AT-skew of the mitochondrial genome of species of the subgenus *Kerteszia* showed that the value of the GC-skew was negative, whereas the AT-skew was positive, showing an excess of A over T.

Similarly, negative GC-skews and positive AT-skews were reported in the mitochondrial genomes of other metazoan species (Saccone et al. 1999) reported to date.

The *COI* gene starts with the TCG codon that has been reported for other *Anopheles* species (Beard et al. 1993; Mitchell et al. 1993; Krzywinski et al. 2006, 2011; Moreno et al. 2010). Three protein-encoding genes (*COI, COII,* and *COIII*) show a partial stop codon (T_) and these are possibly completed via polyadenylation (Beckenbach and Stewart 2009). CU values for *An. darlingi, An. funestus, An. gambiae,* and *An. quadrimaculatus* (Moreno et al. 2010) are higher than those found for *Kerteszia* species The absence of a pattern for RSCU among mitochondrial sequences of *An. cruzii* can be explained because it is a species complex as hypothesized by Ramirez and Dessen (1996, 2000) and Rona et al. (2009, 2010a,b, 2013) and so specimens of various populations can be subjected to different evolutionary histories. However, this hypothesis needs to be considered with caution, because the sample size is small and does not represent all populations of *An. cruzii*.

Mitochondrial genomes of *Kerteszia* species do not show any substantial differences from the genomes reported for other *Anopheles* species relative to the number of rRNA and tRNA genes (2 and 22, respectively) genes. Additionally, the gene arrangement is identical to those of other species (Fig. 3) (Beard et al. 1993; Krzywinski et al. 2006, 2011; Moreno et al. 2010), all protein-encoding genes possess the same number of nonstop codons and the AT-rich control region is highly polymorphic and responsible for controlling replication of the mitochondrial genome (Saito et al. 2005). Furthermore, the mitochondrial genome of *Kerteszia* is subject to rapid evolution; however, the rate of changes is not homogeneous throughout the genome. This is shown by differences in the values of *dN/dS* ratio among genes. Differing from other *Anopheles*, in newly sequenced *Kerteszia* species, the value of the *dN/dS* ratio is lower than 1, indicating that there is a purifying selection in progress. This value is higher than those obtained for species of the *An. albitarsis* complex, whose *dN/dS* ratio values range from 0.0 to 0.057 (Krzywinski et al. 2011).

The length, gene composition, gene arrangement, and AT-content of the mitochondrial genomes of *Kerteszia* species studied are similar to other species of the genus *Anopheles* reported to date.

Few differences were observed among individuals of *An. cruzii* collected in four regions of the Atlantic Forest (Fig. 1). For instance, specimen *An. cruzii* PEC_2_7 differed from the others by possessing both ACG and UGG codons. Additionally, CUG codon was absent in both RJ03_2 and PEC_2_7 specimens; however, it was present in ST16 and RS_32_11_7. The AGG codon was found in all specimens, except in *An. cruzii* RJ03_2. Furthermore, the *NAD6* gene protein-encoding genes of all four specimens of *An. cruzii* have identical ATT start codon, except PEC2_7, in which the start codon is ATC. Whether these differences corroborate the presence of distinct species all under the name *An. cruzii* is a question that needs singular investigation. Ideally future studies should include samples from the same populations analyzed by Ramirez and Dessen (1996, 2000) and Rona et al. (2009, 2010a,b, 2013).

Studies addressing the complete mitochondrial genome of *Anopheles* species representing all subgenera will provide new insights and likely new tools to investigate aspects of evolution within this group of public health importance (Moreno et al. 2010). Additionally, mitochondrial DNA seems to represent a useful source of informative markers to verify hypotheses by Rona et al. (2009, 2010a,b, 2013) that *An. cruzii* is a cryptic species complex with sympatric lineages in areas of the Atlantic Forest. Results presented here fill a gap in representation of mitochondrial genomes of

*Kerteszia* species, a subgenus that encompasses species, which are vectors of *Plasmodium* in forest areas where bromeliad vegetation is an important constituent of the landscape.

## Supplementary Data

Supplementary data are available at *Journal of Medical Entomology* online.

## Acknowledgments

## References Cited

Beard, C. B., D. M. Hamm, and F. H. Collins. 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Mol. Biol. 2: 103–124.

Beckenbach, A. T., and J. B. Stewart. 2009. Insect mitochondrial genomics 3: the complete mitochondrial genome sequences of representatives from two neuropteroid orders: a dobsonfly (Order Megaloptera) and a giant lacewing and an owlfly (Order Neuroptera). Genome 52: 31–38.

Bernt, M., A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, M. Middendorf, and P. F. Stadler. 2013. MITOS: improved de novo Metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69: 313–319.

Birney, E., and R. Durbin. 2000. Using GeneWise in the *Drosophila* annotation experiment. Genome. Res. 10: 547–548.

Bourke, B. P., T.M.P. Oliveira, L. Suesdek, E. S. Bergo, and M.A.M. Sallum. 2003. A multi-locus approach to barcoding in the *Anopheles strodei* subgroup (Diptera: Culicidae). Parasit. Vectors 6: 111.

Chevreux, B., T. Wetter, and S. Suhai. 1999. *Proceedings of the German Conference on Bioinformatics (GCB)*; Genome sequence assembly using trace signals and additional sequence information. Comput. Sci. Biol. 99: 45–56.

Collucci, E., and M. A. M Sallum. 2003. Phylogenetic analysis of the subgenus *Kerteszia* of *Anopheles* (Diptera: Culicidae: Anophelinae) based on morphological characters. Insect Syst. Evol. 34: 361–372.

Duarte, A. M., M. A. Porto, I. Curado, R. S. Malafronte, E. H. Hoffmann, S. G. de Oliveira, A. M. da Silva, J. K. Kloetzel, and A. C. Gomes. 2006. Widespread occurrence of antibodies against circumsporozoite protein and against blood forms of *Plasmodium vivax*, *P.* falciparum and P. malariae in Brazilian wild monkeys. J. Med. Primatol. 35: 87–96.

Duarte, A.M.R.C., R. S. Malafronte, C.Cerutti Jr., I. Curado, B. R. Paiva, A. Y. Maeda, T. Yamasaki, M.E.L. Summaf, D.V.D.A. Neves, S. G. Oliveira, and A. C. Gomes. 2008. Natural *Plasmodium* infections in Brazilian wild monkeys: reservoirs for human infections? Acta Trop. 107: 179–185.

Finn, R. D., J. Clements, and S. R. Eddy. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 39 Suppl 2: W29–W37.

Forattini, O. P. 2002. Identificação, biologia, epidemiologia. *In* Culicidologia médica, vol. II. Editora da Universidade de São Paulo, São Paulo.

Foster, P. G., E. S. Bergo, B. P. Bourke, T.M.P. Oliveira, S. S. Nagaki, D. C. Sant'Ana, and M.A.M. Sallum. 2013. Phylogenetic analysis and DNA-based species confirmation in *Anopheles* (*Nyssorhynchus*). PLoS One 8: e54063.

Francino, M. P., and H. Ochman. 1997. Strand asymmetries in DNA evolution. Trends Genet. 13: 240–245.

Hutchings, R.S.G., M.A.M. Sallum, R.L.M. Ferreira, and R. W. Hutchings. 2005. Mosquitoes of the Jaú National Park and their potential importance in Brazilian Amazonia. Med. Vet. Entomol. 19: 428–441.

Hwang, U. W., C. J. Park, T. S. Young, and W. Kim. 2001. One-step PCR amplification of complete arthropod mitochondrial genomes. Mol. Plylogenet. Evol. 19: 345–352.

Kambhampati, S., and P. T. Smith. 1995. PCR primers for amplification of four insect mitochondrial gene fragments. Insect Mol. Biol. 4: 233–236.

Krzywinski, J., O. G. Grushko, and N. J. Besansky. 2006. Analysis of the complete mitochondrial DNA from *Anopheles funestus*: an improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. Mol. Phylogenet. Evol. 39: 417–423.

Krzywinski, J., C. Li, M. Morris, J. E. Conn, J. B. Lima, M. M. Povoa, and R. C. Wilkerson. 2011. Analysis of the evolutionary forces shaping mitochondrial genomes of a Neotropical malaria vector complex. Mol. Phylogenet. Evol. 58: 469–477.

Laporta, G. Z., M. N. Burattini, D. Levy, L. A. Fukuya, T.M.P. Oliveira, L.M.F. Maselli, J. E. Conn, E. Massad, S. P. Bydlowski, and M.A.M. Sallum. 2015. *Plasmodium falciparum* in the southeastern Atlantic forest: a challenge to the bromeliad-malaria paradigm? Malar. J. 14: 181.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.

Liu, W., Y. Li, G. H. Learn, R. S. Rudicell, J. D. Robertson, B. F. Keele, J.B.N. Ndjango, C. M. Sanz, D. B. Morgan, S. Locatelli, M. K. Gonder, P. J. Kranzusch, P. D. Walsh, E. Delaporte, E. Mpoudi-Ngole, A. V. Georgiev, M. N. Muller, G. M. Shaw, M. Peeters, P. M. Sharp, J. C. Rayner, and B. H. Hahn. 2010. Origin of the human malaria parasite *Plasmodium*. Nature 467: 420–425.

Logue, K., E. R. Chan, T. Phipps, S. T. Small, L. Reimer, C. Henry-Halldin, J. Sattabongkot, P. M. Siba, P. A. Zimmerman, and D. Serre. 2013. Mitochondrial genome sequences reveal deep divergences among *Anopheles punctulatus* sibling species in Papua New Guinea. Malar. J. 12: 64.

Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25: 955–964.

Luz, E., J. Consolim, O. C. Barbosa, and P. B. Torres. 1987. Larvas de *Anopheles* (Subgênero *Kerteszia*) Theobald, 1905 encontradas em criadouros artificiais, no Estado do Paraná, Brasil. Rev. Saúde Pública 21: 466–468.

Marrelli, M. T., R. S. Malafronte, M.A.M. Sallum, and D. Natal. 2007. *Kerteszia* subgenus associated with the Brazilian Atlantic forest: current knowledge and future challenges. Malar. J. 6: 127.

Milne, I., G. Stephen, M. Bayer, P.J.A. Cock, L. Pritchard, L. Cardle, P. D. Shaw, and D. Marshall. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform. 14: 193–202.

Mitchell, S. E., A. F. Cockburn, and J. A. Seawright. 1993. The mitochondrial genome of *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene organization. Genome 36: 1058–1073.

Moreno, M., O. Marinotti, J. Krzywinski, W. P. Tadei, A. A. James, N. L. Achee, and J. E. Conn. 2010. Complete mtDNA genomes of *Anopheles darlingi* and an approach to anopheline divergence time. Malar. J. 9: 127.

Osigus, H-J., M. Eitel, M. Bernt, A. Donath, and B. Schierwater. 2013. Mitogenomics at the base of Metazoa. Mol. Phylogenet. Evol. 69: 339–351.

Perna, N. T., and T. D. Kocher. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J. Mol. Evol. 41: 353–358.

Ramirez, C.C.L., and E.M.B. Dessen. 1996. The polytene chromosomes of the mosquito *Anopheles bellator* compared with those of *Anopheles cruzii*. Braz. J. Genet. 19: 555–558.

Ramirez, C. C., and E. M. Dessen. 2000. Chromosome differentiated populations of *Anopheles cruzii*: evidence for a third sibling species. Genetica 108: 73–80.

Rona, L. D., C. J. Carvalho-Pinto, C. Gentile, E. C. Grisard, and A. A. Peixoto. 2009. Assessing the molecular divergence between *Anopheles* (*Kerteszia*) *cruzii* populations from Brazil using the timeless gene: further evidence of a species complex. Malar. J. 8: 60.

Rona, L. D., C. J. Carvalho-Pinto, C. J. Mazzoni, and A. A. Peixoto. 2010a. Estimation of divergence time between two sibling species of the *Anopheles* (*Kerteszia*) *cruzii* complex using a multilocus approach. BMC Evol. Biol. 10: 91.

Rona, L. D., C. J. Carvalho-Pinto, and A. A. Peixoto. 2010b. Molecular evidence for the occurrence of a new sibling species within the *Anopheles* (*Kerteszia*) *cruzii* complex in south-east Brazil. Malar. J. 9: 33.

Rona, L. D., C. J. Carvalho-Pinto, and A. A. Peixoto. 2013. 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. 13: 207.

Saccone, C., G. C. De, C. Gissi, G. Pesole, and A. Reyes. 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238: 195–209.

Saito, S., K. Tamura, and T. Aotsuka. 2005. Replication origin f mitochondrial DNA in insects. Genetics 171: 1695–1705.

Sallum, M.A.M., O. P. Forattini, and R. C. Wilkerson. 2000. Redescription of the adult and larva and first description of the pupa of *Anopheles* (*Kerteszia*) *laneanus*. J. Am. Mosq. Control Assoc. 16: 86–92.

Sallum, M.A.M., T. R. Schultz, P. G. Foster, K. Aronstein, R. A. Wirtz, and R. C. Wilkerson. 2002. Phylogeny of Anophelinae (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences. Syst. Entomol. 27: 361–382.

Sallum, M.A.M., C.L.C. dos Santos, and R. C. Wilkerson. 2009. Studies on *Anopheles* (*Kerteszia*) *homunculus* Komp (Diptera: Culicidae). Zootaxa 2299: 1–8.

Sanger, F., S. Nicklen, and A. R. Coulson. 1997. DNA sequencing with chain-terminating inhibitors. Proc. Natl Acad. Sci. USA. 74: 5463–5467.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a comilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87: 651–701.

Sinka, M. E., M. J. Bangs, S. Manguin, Y. Rubio-Palis, T. Chareonviriyaphap, M. Coetzee, C. M. Mbogo, J. Hemingway, A. P. Patil, W. H. Temperley, P. W. Gething, C. W. Kabaria, T. R. Burkot, R. E. Harbach, and S. I. Hay. 2012. A global map of dominant malaria vectors. Parasit. Vectors 5: 69.

Sutherland, C. J., N. Tanomsing, D. Nolder, M. Oguike, C. Jennison, S. Pukrittayakamee, C. Dolecek, T. T. Hien, V. E. do Rosário, A. P. Arez, J. Pinto, P. Michon, A. A. Escalante, F. Nosten, M. Burke, R. Lee, M. Blaze, T. D. Otto, J. W. Barnwel, A. Pain, J. Williams, N. J. White, N.P.J. Day, G. Snounou, P. J. Lockhart, P. L. Chiodini, M. Imwong, and S. D. Polley. 2010. Two non-recombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. J. Infect. Dis. 201: 1544–1550.

Tamura, K., G. Stecher, D. Peterson, A. Filipski, and S. Kumar. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Phylogenet. Evol. 30: 2725–2729.

Taylor, J. E., M. A. Pacheco, D. J. Bacon, M. A. Beg, R. L. Machado, R. M. Fairhurst, S. Herrera, J. Y. Kim, D. Menard, M. M. Povoa, L. Villegas, G., Mulyanto L., Snounou F. Y., Cui A. A., Zeyrek and Escalante. 2013. The evolutionary history of *Plasmodium vivax* as inferred from mitochondrial genomes: parasite genetic diversity in the Americas. Mol. Biol. Evol. 30: 2050–2064.

Wang, D., Y. Zhang, Z. Zhang, J. Zhu, and J. Yu. 2010. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genom. Proteom. Bioinform. 8: 77–80.

Wei, S-J., M. Shi, X-X. Chen, M. J. Sharkey, C. van Achterberg, and G-Y. Ye. 2010. New views on strand asymmetry in insect mitochondrial genomes. PLoS One 5: e12708.

Yamasaki, T., A. M. Duarte, I. Curado, M. E. Summa, D. V. Neves, G. Wunderlich, and R. S. Malafronte. 2011. Detection of etiological agents of malaria in howler monkeys from Atlantic Forests, rescued in regions of São Paulo city. Braz. J. Med. Primatol. 40: 392–400.

Zavortink, T. J. 1973. Mosquito studies XXIX. A review of the subgenus *Kerteszia of Anopheles*. Contrib. Am. Entomol. Inst. 9: 1–54.

Zhang, Z., J. Li, X. Q. Zhao, J. Wang, G. K. Wong, and J. Yu. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genom. Proteom. Bioinform. 4: 259–263.