

Molecular phylogeny of the Myzorhynchella Section of *Anopheles* (*Nyssorhynchus*) (Diptera: Culicidae): genetic support for recently described and resurrected species

Brian Patrick Bourke¹, Sandra Sayuri Nagaki¹, Eduardo Sterlino Bergo², Jáder da Cruz Cardoso^{1,3}, Maria Anice Mureb Sallum^{1/+}

¹Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, Avenida Dr. Arnaldo 715, 01246-904, São Paulo, SP, Brasil ²Superintendência de Controle de Endemias, Secretaria de Estado da Saúde de São Paulo, Araraquara, SP, Brasil ³Divisão de Vigilância Ambiental em Saúde, Centro Estadual de Vigilância em Saúde, Secretaria da Saúde do Estado do Rio Grande do Sul, Porto Alegre, RS, Brasil

Phylogenetic relationships among species of the Myzorhynchella Section of Anopheles (Nyssorhynchus) were investigated using the nuclear ribosomal DNA second internal transcribed spacer (ITS2), the nuclear white gene and mitochondrial cytochrome oxidase subunit I (COI) regions. The recently described Anopheles pristinus and resurrected Anopheles guarani were also included in the study. Bayesian phylogenetic analyses found Anopheles parvus to be the most distantly related species within the Section, a finding that is consistent with morphology. An. pristinus and An. guarani were clearly resolved from Anopheles antunesi and Anopheles lutzii, respectively. An. lutzii collected in the same mountain range as the type locality were found within a strongly supported clade, whereas individuals from the southern state of Rio Grande do Sul, tentatively identified as An. lutzii based on adult female external morphology, were distinct from An. lutzii, An. antunesi and from each other, and may therefore represent two new sympatric species. A more detailed examination of An. lutzii sensu lato along its known geographic range is recommended to resolve these anomalous relationships.

Key words: phylogeny - *Anopheles* - Myzorhynchella - white - ITS2 - COI

Mosquitoes (Culicidae) are highly diverse geographically widespread taxa containing approximately 3,529 species (Harbach 2011a). They are also of major medical importance as a consequence of being vectors of pathogens that cause diseases in humans such as dengue, filariasis and malaria (WHO 1989). All known vectors of malarial parasites belong to the genus *Anopheles* Meigen and are responsible for an estimated 243 million malaria cases annually (WHO 2009). Within the Americas there are an estimated one million cases of malaria annually, the majority of which are caused by vectors of the subgenus *Nyssorhynchus* Blanchard (Zimmerman 1992, WHO 2009). This subgenus contains 39 species (Harbach 2011b, Nagaki et al. 2011), at least seven of which are known to harbour malarial parasites (Rosa-Freitas et al. 1998). Whereas the subgenus is a well supported monophyletic group (Sallum et al. 2000, 2002, Harbach & Kitching 2005), relationships among species within this group are less well defined.

The subgenus has traditionally been divided into three sections based on morphological characters (Peyton et al. 1992): the *Argyritarsis*, *Albimanus* and *Myzorhynchella* Sections. Until recently the *Myzorhynchella* Section was comprised of four nominal species, *Anopheles*

lutzii Cruz, *Anopheles parvus* (Chagas), *Anopheles nigratarsis* (Chagas) and *Anopheles antunesi* Galvão & Amaral (Harbach 2004). The *Myzorhynchella* Section is currently the only section in *Nyssorhynchus* without species implicated in malaria transmission and the only section with some genetic support as a natural grouping. Genetic data also indicate that species diversity within the *Myzorhynchella* Section is perhaps underestimated due to the existence of species complexes (Bourke et al. 2010). Nagaki et al. (2011) recently resurrected *Anopheles guarani* from *An. lutzii* and described a new species, *Anopheles pristinus* Nagaki & Sallum, distinguished from *An. antunesi*. Members of species complexes frequently vary in their ability to transmit malaria and an effective assessment of mosquito susceptibility to the malaria parasite relies on the ability to accurately resolve and identify species. Recent studies have worked towards identifying malaria refractory genes in *Anopheles* (Marshall & Taylor 2009, Corby-Harris et al. 2010). If the *Myzorhynchella* Section is a natural grouping within *Nyssorhynchus* without vectors of human malarial protozoa, it may prove to be a useful group for the study and identification of such genes.

The current study therefore seeks to describe the phylogenetic relationships among species within the *Myzorhynchella* Section, assess the validity of recently described *An. pristinus* and *An. guarani* and determine the species status of anomalous individuals using the nuclear ribosomal DNA second internal transcribed spacer (ITS2), the mitochondrial cytochrome oxidase subunit I (COI) and the single copy nuclear *white* gene.

Financial support: FAPESP (2005/53973-0 to MAMS), CNPq (BPP 300351/2008-9 to MAMS)

+ Corresponding author: masallum@usp.br

Received 28 February 2011

Accepted 13 July 2011

MATERIALS AND METHODS

Mosquito collection - A description of the specimens used in this study can be found in Table I. These specimens included the offspring of females caught in the field using a Shannon trap (Shannon 1939) and larvae and pupae collected from immature habitats, which were then raised to adulthood. Species identification of all but two specimens was based on either adult male genitalia or fourth-instar larval characteristics. Specimens *An. lutzii* A325 and *An. lutzii* B369 were identified only by female morphology.

DNA extraction - DNA was extracted from each specimen according to the animal tissue DNA extraction protocol provided by the QIAgen DNeasy® Blood and Tissue Kit (QIAGEN Ltd, Crawley, UK). All extractions were diluted to 200 µL with the buffer provided and extraction solutions were retained for storage at -80°C in the entomological frozen collection of the School of Public Health, University of São Paulo, Brazil.

ITS2 region - This region was amplified using the 5.8SF (5'-ATCACTCGGCTCGTGGATCG-3') and 28SR (5'-ATGCTTAAATTTAGGGGGTAGTC-3') primers using the same protocol adopted by Sallum et al. (2008). The polymerase chain reaction (PCR) was conducted in a total volume of 25 µL containing 1 µL of DNA extraction solution, 1 × PCR buffer (Invitrogen), 1.5 mM MgCl₂ (Invitrogen), 1.25 µL dimethyl sulfoxide (Sigma), 0.1 µM of each primer, 200 mM each dNTPs (Amresco) and 1.25 U Taq Platinum polymerase (Invitrogen). The reaction proceeded under the following temperature regime: 94°C for 2 min, 34 cycles of 94°C for 30 s, 57°C for 30 s and 72°C for 30 s and a final extension at 72°C for 10 min.

COI mitochondrial gene - This region was amplified using LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGT-GACCAAAAATCA-3') primers (Folmer et al. 1994). DNA working solution was first made by diluting the DNA extraction solution to 1:20 using ultra-pure autoclaved water. The PCR reaction was conducted as above for ITS2. The reaction proceeded under the following temperature profile: 95°C for 2 min, 35 cycles of 94°C for 1 min, 57°C for 1 min and 72°C for 1 min and a final extension at 72°C for 7 min.

White nuclear gene - This gene was amplified using WZ2E and WZ11 primers (Besansky & Fahey 1997). This amplification product then served as a template in a sequencing reaction using internal primers W1F (5'-GATCAARAAGATCTGYGACTCGTT-3') and W2R (5'-GCCATCGAGATGGAGGAGCTG-3'). The initial PCR reaction contained approximately 3 µL of DNA extraction solution in a total volume of 25 µL containing 1 × PCR buffer (Invitrogen), 1.5 mM of MgCl₂ (Invitrogen), 200 mM of each dNTP (Amresco), 2 µM of each primer and 0.625 U Taq DNA Polymerase (Invitrogen). The reaction proceeded under the following temperature regime: 94°C for 5 min, 35 cycles at 94°C for 30 s, an annealing temperature of 50°C for 1 min and then 72°C for

2 min followed by a final extension at 72°C for 10 min. The next step then involved taking this PCR product for sequencing using the W1F and W2R internal primers.

Cloning - ITS2 PCR amplicons obtained from *An. parvus* were purified using PEG precipitation (20% polyethylene glycol 8000/2.5 M NaCl) and cloned into pGem-T Easy Vector (Promega, Madison, WI, USA). Three positive clones were sequenced. For the nuclear *white* gene, females were cloned when the direct sequence showed either unreadable peaks or double peaks. Individuals that had either ITS2 or *white* gene cloned are shown in Table I.

Sequencing and sequence alignment - Sequencing reactions were carried out in both directions using a Big Dye Terminator cycle sequencing kit v3.1 (Applied Biosystems) and Applied Biosystems 3130 DNA Analyzer (Applied Biosystems). The COI and *white* gene sequences were aligned first by nucleotides using ClustalX (Thompson et al. 1997) and then by amino acid using TranslatorX (Abascal et al. 2010). The intron in the *white* gene could not be reliably aligned and so was excluded from further analyses. In addition, only a single randomly selected sequence was used to represent cloned individuals in the combined gene analysis.

In the case of ITS2, sequences were first annotated for the 5.8S and 28S ends. Template ITS2 secondary structure was then predicted for *An. antunesi* RJ036 with Predict ITS2 Structure in the ITS2 Database (Koetschan et al. 2010). The remaining secondary structures were modelled at this source with Custom Modelling using the predicted *An. antunesi* RJ036 ITS2 secondary structure (above) as a template. Sequences with secondary structures were then aligned and edited in 4Sale (Seibel et al. 2006, 2008). It was not possible to align the complete ITS2 sequences and so those regions that could not be aligned were excluded from further analysis. Additionally, we were unable to align ITS2 sequences with an appropriate outgroup and so trees constructed using ITS2 sequence data (including the combined gene trees) are unrooted.

Phylogenetic analyses - Bayesian analyses were applied to ITS2, COI and *white* sequence data using a partitioning strategy to allow different partitions to have their own model characteristics (composition, rate matrix and among-site variation) and to allow for among-partition rate variation. Datasets could be left unpartitioned, partitioned by gene, partitioned by codon position or partitioned by both gene and codon position. Optimal evolutionary models were determined for isolated partitions using the Akaike Information Criterion (AIC) in Modeltest (Posada & Crandall 1998).

Optimal models for partitioned data were calculated using Bayes factors (BF) according to the formula $BF_{21} = 2[\ln(HM_2) - \ln(HM_1)]$, where HM_2 and HM_1 are the harmonic means of the posterior sample of likelihoods from the pair of partition strategies being compared (Brown & Lemmon 2007). Positive values of BF_{21} are indicative of support for the partition denoted two over the partition denoted one. A BF_{21} value greater than 10 indicates

TABLE I

Specimen information, including specimen and voucher numbers, localities, geographical coordinates, species included in the study, and GenBank accessions of the second internal transcribed spacer (ITS2) ribosomal DNA, cytochrome oxidase subunit I (COI) mitochondrial and *white* nuclear genes

Sample	Specimen.	Locality (state)	Coordinates	Species	ITS2	COI	<i>white</i>		
RJ0311	RJ03(11)	Itatiaia (RJ)	-44.622139,	<i>Anopheles antunesi</i>	GU989325	GU989343	JN392474		
RJ0312	RJ03(12)		-22.416306		GU989326	GU989344	JN392475		
RJ0313	RJ03(13)		GU989327		GU989345	JN392476			
RJ036	RJ03(6)		GU989324		GU989342	JN392473			
VP0917	VP09-17	Pindamonhangaba (SP)	-45.515500,	<i>An. antunesi</i>	GU989328	GU989346	JN392477		
VP11b	VP11b		-22.758806		GU989329	GU989347	FJ147290		
PR29	PR29	Foz do Iguaçu (PR)	-54.586667, -25.480556	<i>Anopheles guarani</i>	JN023041	JF923659	JN392478		
SP0210	SP02(10)-5	Pariquera-Açu (SP)	-47.949056,	<i>Anopheles lutzii</i>	JN023045	JF923665	FJ147284		
SP0211	SP02(11)-9		-24.749583		JN023046	JF923666	JN392483		
SP0212	SP02(12)-1		JN023047		JF923667	JN392484			
SP1213	SP02(13)-3		JN023048		JF923668	JN392485			
SP0214	SP02(14)-6		JN023049		JF923669	JN392486			
SP0215	SP02(15)-5		JN023050		JF923670	JN392487			
SP029	SP02(9)-2		JN023044		JF923664	FJ147283			
A325	A325		Maquiné (RS)		-50.213611,	<i>An. lutzii</i>	JN023042	JF923662	JN392479
B369	B369	-29.661111		JN023043	JF923663		-		
Clone 3					-	-	JN392480		
Clone 4					-	-	JN392481		
Clone 5					-	-	JN392482		
PR2818	PR28(18)-1	Guaira (PR)	-54.290556,	<i>Anopheles parvus</i>	JN023064	JF923678	FJ147288		
PR2851	PR28(5)-1		-24.271500		-	JF923677	JN392498		
Clone 1					JN023062	-	-		
Clone 2					JN023063	-	-		
PR2865	PR28(65)-6				JN023065	JF923679	JN392499		
AS51	AS5-1	Pilar de Goiás (GO)	-49.539834,	<i>An. parvus</i>	JN023051	JF923671	JN392488		
AS52	AS5-2		-14.809194		-	JF923672	-		
Clone 1						JN023052	-	JN392489	
Clone 2						JN023053	-	JN392490	
Clone 3						JN023054	-	JN392491	
AS53	AS5-3					JN023055	JF923673	-	
Clone 1						-	-	JN392492	
Clone 3						-	-	JN392493	
AS54	AS5-4					-	JF923674	-	
Clone 1						JN023056	-	JN392494	
Clone 2				JN023057	-	JN392495			
Clone 3				JN023058	-	-			
MG562	MG56-2	Campo Belo (MG)	-45.121778, -20.795111	<i>An. parvus</i>	JN02306	JF923676	-		
Clone 2					-	-	JN392496		
Clone 3					-	-	JN392497		
MG07920	MG07(9)-20	Frutal (MG)	-49.076500, -20.025278	<i>An. parvus</i>	-	JF923675	FJ147287		
Clone 1					JN023059	-	-		
Clone 2					JN023060	-	-		
SP50a	SP50a	Pindamonhangaba (SP)	-45.515278,	<i>Anopheles pristinus</i>	GU989333	GU989351	JN392500		
SP50b	SP50b		-22.758472		GU989334	GU989352	JN392501		
SP51100	SP51-100		GU989335		GU989353	JN392502			
SP53100	SP53-100		GU989338		GU989356	JN392505			
SP53101	SP53-101		GU989339		GU989357	JN392506			
SP534	SP53-4		GU989336		GU989354	JN392503			
SP535	SP53-5		GU989337		GU989355	JN392504			
SP552	SP55(2)		GU989340		GU989358	JN392507			
SP554	SP55(4)		GU989341		GU989359	JN392508			
VP11a	VP11a		Pindamonhangaba (SP)		-45.515500, -22.758806	<i>An. pristinus</i>	GU989331	GU989348	FJ147289

GO: Goiás; MG: Minas Gerais; PR: Paraná; RJ: Rio de Janeiro; RS: Rio Grande do Sul; SP: São Paulo.

significant support for partition two, values between 10 and -10 indicate ambiguity and values less than -10 indicate significant support for partition one. All Bayesian analyses were performed using MrBayes (Ronquist & Huelsenbeck 2003) and each analysis consisted of two runs to provide confirmation of convergence of posterior probability distribution.

For COI, *white*, ITS2 and all but one of the combined analyses of gene partition strategies, each run was four million generations long and the first two million were discarded as burn-in. The Metropolis-coupled Markov chain Monte Carlo strategy was used with four heated chains; adequate mixing was achieved by setting the chain temperature to 0.1 for isolated genes and combined genes. Convergence of topology between the two runs was monitored using the average standard deviation of split frequencies - this index consistently fell to below 0.015 in the post-burn-in samples. Convergence was also monitored by noting that the potential scale reduction factor values were all approximately 1.0 in the post-burn-in samples. The one exceptional partition strategy (partitioning by gene and codon position with among partition rate variation) did not converge (runs were varied from two-40 million generations, with four-eight heated chains and chain temperatures from 0.05-0.25) and so was excluded from further analysis. Post burn-in samples were then used to construct a consensus tree containing nodes with greater than 70% posterior probability support. Trees were drawn using the R package APE (Paradis et al. 2004).

Pairwise genetic distances for COI were calculated in MEGA4 (Tamura et al. 2007) using Kimura's (1980) two-parameter (K2P) model. Uncorrected pairwise p-distances were calculated in MEGA4 from a separate *An. parvus* ITS2 alignment using ClustalX.

RESULTS

The alignment included sequences from 35 individuals. Difficulties with aligning ITS2 (at positions 143-199, 223-285, 386-635 and 714-797) and the intron in the *white* gene resulted in these positions being excluded from the phylogenetic analyses. Non-overlapping sites from the *white* gene were also excluded. In total, 1,539 sites were included in the analyses, consisting of 363 sites from ITS2, 657 sites from COI and 519 sites from the *white* gene. GenBank accessions of the nuclear *white*, mitochondrial COI and the ITS2 are in Table I. *Anopheles gambiae* Giles and *Anopheles strodei* Root were used as outgroup taxa in the analyses of COI and *white* gene. It was not possible to find an appropriate outgroup for the ITS2 sequences and so ITS2 and combined COI-*white*-ITS gene trees were unrooted.

The optimal models of evolution determined for each sequence partition are displayed in Table II. Where the selected rate matrix was unavailable in MrBayes (i.e. TrN and TIM), the most similar rate matrix available was selected (i.e. GTR). ITS2 is a non-coding region and so was left unpartitioned in the analysis. The best evolutionary model for this region was the TrNef+I model (substituted with the GTR+I model in Bayesian analyses). BF (Tables III, IV, V) found the optimal partitioning

TABLE II
Models used for gene and codon positions

Gene	Codon position	Modeltest	Model used
ITS2 + COI + <i>white</i>	-	GTR + I + G	GTR + I + G
ITS2 ^a	-	K80 + I	HKY + I
ITS2 ^b	-	TrNef + I	GTR + I
COI	-	GTR + I + G	GTR + I + G
	1	HKY	HKY
	2	TrN + I + G	TrN + I + G
<i>white</i>	3	TrN + I	GTR + I
	-	GTR + G	GTR + G
	1	TrN + I	GTR + I
COI ^c	2	HKY + G	HKY + G
	3	TrNef	GTR
	-	GTR + G	GTR + G
<i>white</i> ^d	1	HKY	HKY
	2	TrN + I	GTR + I
	3	TrN + I	GTR + I
	-	GTR + I + G	GTR + I + G
	1	GTR + I	GTR + I
	2	HKY + G	HKY + G
	3	K81 + I	GTR + I

a: including clones for isolated gene tree; b: excluding clones for combined gene tree; c: with outgroup; d: with outgroup and clones; COI: cytochrome oxidase subunit I; ITS2: second internal transcribed spacer.

TABLE III

Bayes factors calculated from the harmonic mean of the likelihoods for all partitions at cytochrome oxidase subunit I

	None	Codon (-)	Codon (+)
None	-	126.35	563.01
Codon (-)	-126.35	-	436.66
Codon (+)	-563.01	-436.66	-

(+), (-): inclusion and exclusion of among partition rate variation, respectively.

TABLE IV

Bayes factors calculated from the harmonic mean of the likelihoods for all partitions at the *white* gene

	None	Codon (-)	Codon (+)
None	-	163.59	190.35
Codon (-)	-163.59	-	26.76
Codon (+)	-190.35	-26.76	-

(+), (-): inclusion and exclusion of among partition rate variation, respectively.

strategies for COI, *white* and combined genes. The best model for the COI gene (with an outgroup taxa) was one that partitioned the data by codon position and included among-partition rate variation (APRV) (Table III). The model chosen for the *white* gene (with an outgroup taxa) was one that partitioned the data by codon position and included APRV (Table IV). The model chosen for combined genes was one that partitioned the data by gene and codon position (Table V).

The results of Bayesian analyses show a high degree of congruence among the ITS2, COI, *white* and combined gene trees (Figs 1-4, respectively). *An. parvus*, *An. pristinus* and *An. guarani* are strongly supported as species across all trees. In addition, the majority of the *An. lutzii* individuals (7 of 9) form a strongly supported clade across all trees with support of 100% Bayesian posterior probability (BPP). Some incongruence also exists between gene trees. The remaining two *An. lutzii* individuals (A325 and

TABLE V
Bayes factors calculated from the harmonic mean of the likelihoods for all partitions at combined genes

	None	Gene (-)	Gene (+)	Codon (-)	Codon (+)	Gene/Codon (-)	Gene/Codon (+)
None	-	203.64	174.7	355.22	451.35	469.2	Na
Gene (-)	-203.64	-	-28.94	151.58	247.71	265.56	Na
Gene (+)	-174.7	28.94	-	180.52	276.65	294.5	Na
Codon (-)	-355.22	-151.58	-180.52	-	96.13	113.98	Na
Codon (+)	-451.35	-247.71	-276.65	-96.13	-	17.85	Na
Gene/Codon (-)	-469.2	-265.56	-294.5	-113.98	-17.85	-	Na
Gene/Codon (+)	Na	Na	Na	Na	Na	Na	-

Na: not applicable; (+), (-): inclusion and exclusion of among partition rate variation, respectively.

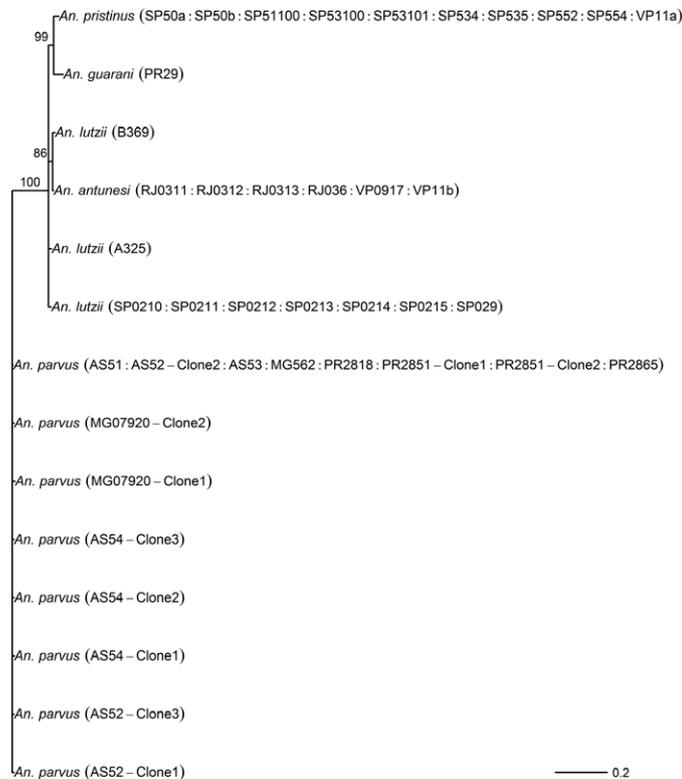


Fig. 1: Bayesian tree of the second internal transcribed spacer of species of the Myzorhynchella Section. Numbers at branches indicate Bayesian posterior probability ($\geq 70\%$).

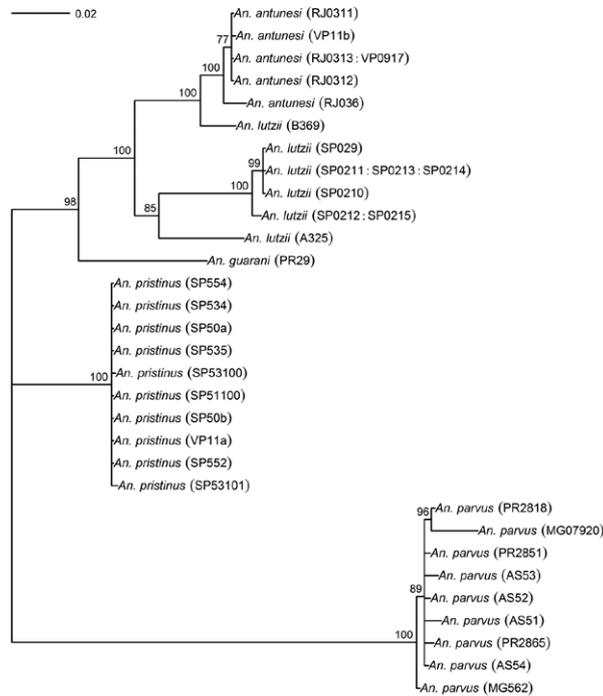


Fig. 4: Bayesian tree of the combined second internal transcribed spacer, single-copy nuclear *white* gene and mitochondrial cytochrome oxidase subunit I gene. The data were partitioned by gene and codon position. Numbers at branches indicate Bayesian posterior probability ($\geq 70\%$).

B369) cluster differently across trees. In the *white* and combined gene trees (Figs 2, 4), individual A325 forms a clade with the *An. lutzii sensu stricto* group described above ($> 85\%$ BPP), whereas the other (represented by the 3 clones from B369) forms a clade with *An. antunesi* (100% BPP). Both individuals form a clade with *An. antunesi* at the COI gene (78% BPP) (Fig. 2), whereas A325 was unclustered with *An. lutzii* at ITS2 (Fig. 1).

The COI and *white* gene trees (Figs 2, 3) were rooted with an outgroup. In the *white* gene tree (Fig. 3), *An. strodei* and *An. parvus* were sisters to a clade containing the remaining species in the Myzorhynchella Section (96% BPP). In the COI tree (Fig. 2), *An. strodei* and *An. parvus* were sisters to *An. pristinus*, *An. guarani* and the *An. lutzii-An. antunesi* clade. In the combined gene tree (Fig. 4), *An. pristinus* was sister to a clade containing *An. guarani*, *An. antunesi* and *An. lutzii* (98% BPP) and within this clade *An. guarani* was sister to a clade containing all individuals of *An. antunesi* and *An. lutzii* (100% BPP).

COI pairwise genetic distances among *An. lutzii* and *An. antunesi* were calculated under the K2P model (Table VI). Pairwise distances among individuals within *An. antunesi* and *An. lutzii* s.s. were found to be less than 2%, whereas differences between individuals from each group were in the range 6.90-7.70%. Individuals *An. lutzii* A325 and *An. lutzii* B369 were most distant to those from the *An. lutzii* s.s. group (8.00-8.60 and 6.40-7.20%, respectively). The distances between A325 and *An. antunesi* individuals ranged from 5.80-6.40, whereas the

distances between B369 and *An. antunesi* ranged from 2.30-2.80%. *An. lutzii* A325 and *An. lutzii* B369 differed from each other by 6.4%.

Difficulty in aligning ITS2 sequences resulted in large numbers of polymorphic sites being excluded from the phylogenetic analyses. It is likely the exclusion of such sites contributed to the existence of the large polytomy observed among *An. parvus* individuals and the lack of resolution for *An. lutzii* s.s. and *An. lutzii* A325. A separate alignment of *An. parvus* individuals was constructed to include all sites and pairwise p-distances were calculated to describe variation among these individuals (p-distances) (Table VII). *An. parvus* was the only species in the study to have intragenomic variation at ITS2. *An. parvus* showed higher intragenomic (0.17-2.09%) than intraspecific (0-1.57%) variation. Two clones from a single individual from the state of Minas Gerais (MG) differed from each other by 2.09% and from other conspecifics by 0.87-1.57%. This same individual was also clearly resolved from the remaining *An. parvus* individuals at the COI gene. A separate alignment of *An. lutzii* and *An. antunesi* could not be constructed because of the frequency of ambiguous sites and so p-distances could not be calculated as in the case of *An. parvus*. Instead, we calculated the proportion of unambiguously aligned sites for these ITS2 sequences on a pairwise basis using secondary structure (Table VIII). *An. lutzii* and *An. antunesi* consisted of four unique ITS2 sequences, one each for *An. lutzii* A325, *An. lutzii* B369, *An. antunesi*

TABLE VI
Cytochrome oxidase subunit I (COI) pairwise distance Kimura's two-parameter (K2P) values (%) among individuals of *Anopheles antunesi* and *Anopheles lutzii*

	An. antunesi RJ0311	An. antunesi RJ0312	An. antunesi RJ0313	An. antunesi RJ036	An. antunesi VP0917	An. antunesi VP11b	An. lutzii B369	An. lutzii A325	An. lutzii SP0210	An. lutzii SP0211	An. lutzii SP0212	An. lutzii SP0213	An. lutzii SP0214	An. lutzii SP0215	An. lutzii SP029
<i>An. antunesi</i> RJ0311	-														
<i>An. antunesi</i> RJ0312	0.00	-													
<i>An. antunesi</i> RJ0313	0.00	0.00	-												
<i>An. antunesi</i> RJ036	1.70	1.70	1.70	-											
<i>An. antunesi</i> VP0917	0.00	0.00	0.00	1.70	-										
<i>An. antunesi</i> VP11b	0.00	0.00	0.00	1.70	0.00	-									
<i>An. lutzii</i> B369	2.30	2.30	2.30	2.80	2.30	2.30	-								
<i>An. lutzii</i> A325	5.80	5.80	5.80	6.40	5.80	5.80	6.40	-							
<i>An. lutzii</i> SP0210 ^a	7.00	7.00	7.00	7.70	7.00	7.00	7.70	8.60	-						
<i>An. lutzii</i> SP0211 ^a	7.00	7.00	7.00	7.70	7.00	7.00	7.70	8.60	0.00	-					
<i>An. lutzii</i> SP0212 ^a	6.90	6.90	6.90	7.20	6.90	6.90	7.20	8.60	1.10	1.10	-				
<i>An. lutzii</i> SP0213 ^a	7.00	7.00	7.00	7.70	7.00	7.00	7.70	8.60	0.00	0.00	0.00	-			
<i>An. lutzii</i> SP0214 ^a	7.00	7.00	7.00	7.70	7.00	7.00	7.70	8.60	0.00	0.00	1.10	0.00	-		
<i>An. lutzii</i> SP0215 ^a	6.90	6.90	6.90	7.20	6.90	6.90	7.20	8.00	1.10	1.10	0.00	1.10	1.10	-	
<i>An. lutzii</i> SP029 ^a	7.00	7.00	7.00	7.70	7.00	7.00	7.70	8.60	0.00	0.00	1.10	0.00	0.00	1.10	-

a: *An. lutzii* s.s. Mean group distances were 0.6% and 0.5% for *An. antunesi* and *An. lutzii* s.s., respectively.

and *An. lutzii* s.s., giving a total of six pairwise combinations. The results indicated that the *An. lutzii* A325 and *An. lutzii* B369 sequence pair was the most difficult to align with only 77% of sites aligned, compared to 0.91% for *An. lutzii* s.s. and *An. antunesi*.

DISCUSSION

There is scant molecular data available for species of the Myzorhynchella Section. However, in Bourke et al. (2010) phylogenetic analysis of the subgenus *Nyssorhynchus*, the Myzorhynchella Section was found to be monophyletic (at *white* and combined *white*-ND6 genes) and *An. parvus* was consistently found to be the sister to the remaining species of the Myzorhynchella Section. Our study found *An. parvus* to be sister to the Myzorhynchella clade and *An. strodei* in the *white* gene tree, whereas at COI it was found to be sister to *An. pristinus*, *An. guarani*, an *An. antunesi*-*An. lutzii sensu lato* clade and *An. strodei* from the outgroup. *An. parvus* was also the most distantly related species within the Section based on branch lengths. This distinct relationship is also supported by morphological characters that distinguish species within the Myzorhynchella Section. *An. parvus* can be separated from the remaining species in the Section by a distinct hook-like apex at the aedeagus in the male genitalia (Root 1927, Galvão 1941). Additionally, the eggs of *An. parvus* have an anterior crown-like structure and float vertically, characteristics which distinguish the species not only from species of the Myzorhynchella Section but also from *Anopheles* species in general (Forattini et al. 1998). Intraspecific variation did not support *An. parvus* as a species complex. However, a single individual from MG did show some evidence of distinction from the remaining individuals of *An. parvus*, as well as high intragenomic variation at ITS2. Concerted evolution normally maintains fixed interspecific differences and intraspecific homogeneity within ribosomal multigene families (Arnheim 1983). However, intragenomic variation occurs when the rate of mutation exceeds the rate of homogenization. Although fixed interspecific differences and intraspecific homogeneity at ITS2 has permitted unambiguous species identification in a range of closely related *Anopheles* species (Collins & Paskewitz 1996, Beebe et al. 2001, Wilkerson et al. 2004, Li & Wilkerson 2005), the intraspecific variation in ITS2 that sometimes occurs within *Anopheles* can pose a major problem for population and phylogenetic studies when it exceeds variation between populations. Such markers are likely to be of little use for the resolution of populations and may even lead to the misidentification of species (Li & Wilkerson 2007). Here we find higher sequence variation within an individual from MG than that found between all other individuals of *An. parvus*. In addition, both sequences in this individual were very different from all other sequences in *An. parvus*. Although this finding may be suggestive of population or taxonomic differences in *An. parvus*, given the extent of intragenomic variation, we consider the utility of ITS2 to be limited in more detailed future phylogenetic analysis of this species.

TABLE VII
Pairwise p-distance values (%) among complete internal transcribed spacer sequences of *Anopheles parvus*

	AS52 (clone 1)	AS52 (clone 3)	AS54 (clone 1)	AS54 (clone 2)	AS54 (clone 3)	MG07920 (clone 1)	MG07920 (clone 2)	PR2851 (clone 1)	PR2851 (clone 2)	AS52 (clone 2), PR2818, PR2865, MG562, AS53, AS51
AS52 (clone 1)	-	-	-	-	-	-	-	-	-	-
AS52 (clone 3)	0.52	-	-	-	-	-	-	-	-	-
AS54 (clone 1)	0.52	0.35	-	-	-	-	-	-	-	-
AS54 (clone 2)	0.52	0.35	0.35	-	-	-	-	-	-	-
AS54 (clone 3)	0.70	0.52	0.52	0.52	-	-	-	-	-	-
MG07920 (clone 1)	1.22	1.05	1.05	1.05	1.22	-	-	-	-	-
MG07920 (clone 2)	1.22	1.40	1.40	1.40	1.57	2.09	-	-	-	-
PR2851 (clone 1)	0.17	0.35	0.35	0.35	0.52	1.05	1.05	-	-	-
PR2851 (clone 2)	0.35	0.17	0.17	0.17	0.35	0.87	1.22	0.17	-	-
AS52 (clone 2), PR2818, PR2865, MG562, AS53, AS51	0.35	0.17	0.17	0.17	0.35	0.87	1.22	0.17	0.00	-

TABLE VIII
Proportion of unambiguously aligned sites at internal transcribed spacer
(excluding the 5.8S and 28S regions) based on pairwise secondary structure alignments

	<i>An. lutzii</i> s.s	<i>An. antunesi</i>	<i>An. lutzii</i> A325	<i>An. lutzii</i> B369
<i>Anopheles lutzii</i> s.s.	-	-	-	-
<i>Anopheles antunesi</i>	0.91	-	-	-
<i>An. lutzii</i> A325	0.79	0.90	-	-
<i>An. lutzii</i> B369	0.93	0.97	0.77	-

There is general consistency in the remaining species relationships found at all genes, i.e. the existence of *An. lutzii* and *An. antunesi* species complexes and support for the recently described *An. guarani* and resurrected *An. pristinus*. *An. lutzii* was first described from individuals collected from the state of Rio de Janeiro (Cruz 1901). A later study synonymised *An. niger* Theobald and *An. guarani* with *An. lutzii* (Lane 1953). Consequently, the geographic distribution for this species became quite extensive, with records from Argentina, Brazil, Mexico and Paraguay. Recently, Nagaki et al. (2011) found morphological support for the resurrection of *An. guarani* from *An. lutzii* and for *An. niger* to be synonymised with *An. guarani*. Indications are that, contrary to having a continental-scale distribution, *An. lutzii* may be restricted to the Atlantic Forest of southeastern Brazil (Nagaki et al. 2011). Our analysis again found support for the distinction of *An. guarani*. *An. lutzii* on the other hand is found to be paraphyletic with respect to other species in the group. Whereas most *An. lutzii* individuals [from the state of São Paulo (SP)] consisted of one strongly supported group likely to be *An. lutzii* s.s. (with the type specimen originating from the same Serra do

Mar mountain range), the lack of monophyly in this species was caused by two individuals tentatively identified as *An. lutzii* collected from the most southerly Brazilian state of Rio Grande do Sul. These two individuals were found to consist of two undescribed sympatric species, one of which appeared more closely related to *An. antunesi* and the other which was associated with both the *An. lutzii* s.s clade and *An. antunesi*. The COI genetic distances observed between *An. lutzii* A325 and *An. antunesi* (5.80-6.40%) and *An. lutzii* s.s (8-8.60%) individuals are approximately 10 times greater than mean within group distances (0.60% and 0.50% for *An. antunesi* and *An. lutzii* s.s, respectively). This finding is consistent with the sequence threshold of 10 times the mean intraspecific variation to delimit animal species proposed by Hebert et al. (2004). Although the relationship between *An. lutzii* B369 and *An. antunesi* falls outside this threshold (4 times mean intraspecific variation), genetic distances between *An. lutzii* B369 and *An. antunesi* individuals (2.30-2.80%) remained high and were close to the 3% threshold used by Hebert et al. (2003) to resolve 196 of 200 species of Lepidoptera.

Fixed interspecific differences and intraspecific homogeneity generally found at ITS2 have proved effective at resolving many closely related *Anopheles* species and our results found that *An. antunesi* and *An. lutzii* s.s. are each represented by a single ITS2 sequence. Differences observed between *An. lutzii* A325 and *An. lutzii* B369 (based on the proportion of sites successfully aligned) were greater than interspecific differences. However, the exclusion of a large numbers of potentially important sites at ITS2 from phylogenetic analysis may have accounted for the poor resolution between *An. lutzii* A325 and *An. lutzii* in the ITS2 tree. It is notable that *An. lutzii* A325 and *An. lutzii* B369 were the only individuals in this study to be identified solely by adult female morphology. *An. lutzii* can be differentiated from other members of the Myzorhynchella Section by having small patches of pale scales at the proximal and distal ends of two white spots on the vein R_{4+5} of the wing (versus other combinations of dark and white spots) (Nagaki et al. 2011). However, important differences in the egg, larval and male genitalic morphology may exist that differentiate them from *An. lutzii* and *An. antunesi*.

From its original description by Galvão and Amaral (1940) from SP, *An. antunesi* has been recorded across a large geographical range from northeastern Brazil (Rebêlo et al. 2007) to Argentina (Gorham et al. 1967, Darsie 1985) and Uruguay (Rodriguez & Varela 1962, Gorham et al. 1967). However, recent examination of *An. antunesi* from the type locality has shown that individuals formerly described as *An. antunesi* can be resolved into two sympatric species, *An. antunesi* and *An. pristinus*, based on the pattern of pale and dark wing spots, male genitalia and fourth-instar larva (Nagaki et al. 2010). Our analyses support this finding by resolving *An. antunesi* and *An. pristinus* and providing strong support for the monophyly of *An. pristinus*. However, we also found *An. antunesi* clusters with *An. lutzii*, as mentioned earlier, which is a relationship that has been recovered in a previous study of *Nyssorhynchus* phylogeny (Bourke et al. 2010). As a result of these findings, we find it necessary to question the status of *An. antunesi* in much of its reported range and suggest that the reports of the species in these varied localities may be the result of misidentifications.

The main findings of the current study confirm the species status of *An. pristinus* and *An. guarani* and identify a strongly supported *An. lutzii* s.s. clade and two species complexes (*An. antunesi* and *An. lutzii* complexes). To further clarify phylogenetic relationships among species within the Myzorhynchella Section, we propose additional sampling and morphological analyses (egg, larval, pupal, male genitalic and adult female morphology) of *An. lutzii* s.l. from various localities in southern and southeastern Brazil. These individuals may then be more accurately identified, as a particular form or species, prior to additional phylogenetic analysis. In addition, the findings of Nagaki et al. (2010, 2011) and the current study underlines the need for a reevaluation of the geographic distribution of the species of the Myzorhynchella Section in general. The principal questions raised from this study, therefore, are whether published records of *An. antunesi* from outside the type locality, such as Argentina and Uruguay, refer to the nomi-

nate species or to *An. pristinus* and, similarly, whether reports of *An. lutzii* to date refer to *An. lutzii* s.s., to a distinct species in an *An. antunesi* complex, or should be classified as *An. guarani*. The reports of *An. lutzii* from Mexico are a good example of potential confusion associated with this species. The individuals of *An. niger*, originally described as *Myzorhynchella nigra* by Theobald (1907) and synonymised with *An. lutzii* (Chagas 1907, Belkin 1968), are now potentially *An. guarani*, as noted by Nagaki et al. (2011). Consequently, resolving the apparent complexes and undertaking a morphological re-examination of individuals identified in collections as *An. antunesi* and *An. lutzii* will be the basis for providing more accurate distributions of species in the Myzorhynchella Section.

ACKNOWLEDGEMENTS

To the three anonymous reviewers, for their great contribution to the improvement of the text.

REFERENCES

- Abascal F, Zardoya R, Telford MJ 2010. TranslatorX: multiple alignments of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* 38: W7-W13.
- Arnheim N 1983. Concerted evolution of multigene families. In M Nei, R Koehn, *Evolution of genes and proteins*, Sinauer Associates, New York, p. 38-61.
- Beebe NW, Maung J, van den Hurk AF, Ellis JT, Cooper RD 2001. Ribosomal DNA spacer genotypes of the *Anopheles bancroftii* group (Diptera: Culicidae) from Australia and Papua New Guinea. *Insect Mol Biol* 10: 407-413.
- Belkin JN 1968. Mosquito studies (Diptera, Culicidae) IX. The type specimens of New World mosquitoes in European museums. *Contrib Amer Entomol Inst* 3: 1-69.
- Besansky NJ, Fahey GT 1997. Utility of the *white* gene in estimating phylogenetic relationships among mosquitoes (Diptera: Culicidae). *Mol Biol Evol* 14: 442-454.
- Bourke BP, Foster PG, Bergo ES, Calado DC, Sallum MA 2010. Phylogenetic relationships among species of *Anopheles* (*Nyssorhynchus*) (Diptera, Culicidae) based on nuclear and mitochondrial gene sequences. *Acta Trop* 114: 88-96.
- Brown JM, Lemmon AR 2007. The importance of data partitioning and the utility of Bayes factors in Bayesian phylogenetics. *Syst Biol* 56: 643-655.
- Chagas C 1907. O novo gênero *Myzorhynchella* de Theobald: duas novas anophelinas brasileiras pertencentes a este gênero - *Myzorhynchella parva* (nov. sp.). *Brazil Medico* 21: 291-293.
- Collins FH, Paskewitz SM 1996. A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic *Anopheles* species. *Insect Mol Biol* 5: 1-9.
- Corby-Harris V, Drexler A, Watkins de Jong L, Antonova Y, Pakpour N, Ziegler R, Ramberg F, Lewis EE, Brown JM, Luckhart S, Riehle MA 2010. Activation of Akt signalling reduces the prevalence and intensity of malaria parasite infection and lifespan in *Anopheles stephensi* mosquitoes. *PLoS Pathog* 6: e1001003.
- Cruz OG 1901. Contribuição para o estudo dos culicídeos do Rio de Janeiro. *Brazil Medico* 15: 423-426.
- Darsie RF 1985. Mosquitoes of Argentina. Part I. Keys for identification of adult females and fourth stage larvae in English and Spanish (Diptera: Culicidae). *Mosq Syst* 17: 153-253.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase

- subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3: 294-297.
- Forattini OP, Sallum MAM, Bergo ES, Flores DC 1998. Ultrastructure of eggs of *Anopheles rondoni*, *Anopheles lutzii* and *Anopheles parvus*, three species of the subgenus *Nyssorhynchus*. *J Am Mosq Control Assoc* 14: 256-265.
- Galvão ALA 1941. Contribuição ao conhecimento das espécies de *Myzorhynchella* (Diptera: Culicidae). *Arch Zool Sao Paulo* 2: 505-576.
- Galvão ALA, Amaral ADF 1940. Estudos sobre os anofelinos do grupo *Myzorhynchella* com a descrição de uma espécie nova, *Anopheles (Nyssorhynchus) antunesi* n. sp. (Diptera: Culicidae). *Folia Clin Biol* 12: 150-160.
- Gorham JR, Stojanovich CJ, Scott HG 1967. *Clave ilustrada para los mosquitos anofelinos de Sudamerica Oriental*, Communicable Disease Centre/Public and Human Health Services, Atlanta, 62 pp.
- Harbach RE 2004. The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. *Bull Entomol Res* 95: 537-553.
- Harbach RE 2011a. [accessed 29 April 2011]. Family Culicidae Meigen, 1818. Mosquito taxonomic inventory. Available from: mosquito-taxonomic-inventory.info/family-culicidae-meigen-1818.
- Harbach RE 2011b. [accessed 29 April 2011]. Subgenus *Nyssorhynchus* Blanchard, 1902. Mosquito taxonomic inventory. Available from: mosquito-taxonomic-inventory.info/subgenus-nyssorhynchus-blanchard-1902.
- Harbach RE, Kitching LJ 2005. Reconsideration of anopheline mosquito phylogeny (Diptera: Culicidae: Anophelinae) based on morphological data. *Syst Biodiv* 3: 345-374.
- Hebert PDN, Cywinska A, Ball SL, DeWaard JR 2003. Biological identifications through DNA barcodes. *Proc R Soc Lond B* 270: 313-321.
- Hebert PDN, Stoeckle MA, Zemplak TS, Francis CM 2004. Identification of birds through DNA barcodes. *PLoS Biol* 2: 1657-1663.
- Kimura M 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 15: 111-120.
- Koetschan C, Forster F, Keller A, Schleicher T, Ruderisch B, Schwarz R, Muller T, Wolf M, Schultz J 2010. The ITS2 Database III - sequences and structures for phylogeny. *Nucleic Acids Res* 38: D275-D279.
- Lane J 1953. Neotropical Culicidae. *Dixinae, Chaoborinae and Culicinae, tribes Anophelini, Toxorhynchitini and Culicini*, vol. I, Universidade de São Paulo, São Paulo, 548 pp.
- Li C, Wilkerson RC 2005. Identification of *Anopheles (Nyssorhynchus) albitarsis* complex species (Diptera: Culicidae) using rDNA internal transcribed spacer 2-based polymerase chain reaction primers. *Mem Inst Oswaldo Cruz* 100: 495-500.
- Li C, Wilkerson RC 2007. Intragenomic rDNA ITS2 variation in the neotropical *Anopheles (Nyssorhynchus) albitarsis* complex (Diptera: Culicidae). *J Hered* 98: 51-59.
- Marshall JM, Taylor CE 2009. Malaria control with transgenic mosquitoes. *PLoS Med* 6: e1000020.
- Nagaki SS, da Silva AM, Sallum MAM 2011. Redescription of *Anopheles (Nyssorhynchus) lutzii* and resurrection of *Anopheles guarani* from synonymy with *An. lutzii* (Diptera: Culicidae). *Ann Entomol Soc Am* 104: 374-388.
- Nagaki SS, Motta MA, Sallum MAM 2010. Redescription of *Anopheles (Nyssorhynchus) antunesi* Galvão & Amaral and description of a new species of the *Myzorhynchella* Section (Diptera: Culicidae) from Serra da Mantiqueira, Brazil. *Mem Inst Oswaldo Cruz* 105: 278-285.
- Paradis E, Claude J, Strimmer K 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289-290.
- Peyton EL, Wilkerson RC, Harbach RE 1992. Comparative analysis of the subgenera *Kerteszia* and *Nyssorhynchus* of *Anopheles* (Diptera: Culicidae). *Mosq Syst* 24: 51-69.
- Posada D, Crandall KA 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Rebêlo JMM, Moraes JLP, Alves GA, Leonardo FS, da Rocha RV, Mendes WA, Costa E, Câmara LEMB, Silva MJA, Pereira YNO, Mendonça JAC 2007. Distribution of species from genus *Anopheles* (Diptera: Culicidae) in the state of Maranhão, Brazil. *Cad Saude Publica* 23: 2959-2971.
- Rodriguez MF, Varela JC 1962. *Anopheles (Myzorhynchella) antunesi*, especie nueva para el Uruguay. *An Fac Med Montev* 47: 246-249.
- Ronquist F, Huelsenbeck JP 2003. MRBAYES 3 Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
- Root FM 1927. Studies on Brazilian mosquitoes. IV. Notes on some Brazilian species of *Anopheles*. *Am J Hyg* 7: 599-605.
- Rosa-Freitas MG, Lourenço-de-Oliveira R, Carvalho-Pinto CJ, Flores-Mendoza C, Silva-do-Nascimento TF 1998. Anopheline species complexes in Brazil. Current knowledge of those related to malaria transmission. *Mem Inst Oswaldo Cruz* 93: 651-655.
- Sallum MA, Marrelli MT, Nagaki SS, Laporta GZ, Dos Santos CL 2008. Insight into *Anopheles (Nyssorhynchus)* (Diptera: Culicidae) species from Brazil. *J Med Entomol* 45: 970-981.
- Sallum MAM, Schultz TR, Foster PG, Aronstein K, Wirtz RA, Wilkerson RC 2002. Phylogeny of Anophelinae (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences. *Syst Entomol* 27: 361-382.
- Sallum MAM, Schultz TR, Wilkerson RC 2000. Phylogeny of Anophelinae based on morphological characters. *Ann Entomol Soc Am* 93: 745-775.
- Seibel PN, Muller T, Dandekar T, Schultz J, Wolf M 2006. 4SALE - A tool for synchronous RNA sequence and secondary structure alignment and editing. *BMC Bioinformatics* 7: 498.
- Seibel PN, Muller T, Dandekar T, Wolf M 2008. Synchronous visual analysis and editing of RNA sequence and secondary structure alignments using 4SALE. *BMC Res Notes* 1: 91.
- Shannon RC 1939. Methods for collecting and feeding mosquitoes in jungle yellow fever studies. *Am J Trop Med* 19: 131-138.
- Tamura K, Dudley J, Nei M, Kumar S 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.
- Theobald FV 1907. *A monograph of the Culicidae or mosquitoes*, vol. IV, British Museum of Natural History, London, 639 pp.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876-4882.
- WHO - World Health Organization 2009. World malaria report 2009. Available from: who.int/malaria/publications/atoz/9789241563901/en/index.html.
- WHO - World Health Organization 1989. *Geographical distribution of arthropod-borne diseases and their principal vectors*, document WHO/VBC/89.967, WHO, Geneva, 134 pp.
- Wilkerson RC, Reinert JF, Li C 2004. Ribosomal DNA ITS2 sequences differentiate six species in the *Anopheles crucians* complex (Diptera: Culicidae). *J Med Entomol* 41: 392-401.
- Zimmerman RH 1992. Ecology of malaria vectors in the Americas and future direction. *Mem Inst Oswaldo Cruz* 87: 371-383.