BMC Infectious Diseases

RESEARCH ARTICLE

Open Access

CrossMark

Merozoite surface protein-1 genetic diversity in *Plasmodium malariae* and *Plasmodium brasilianum* from Brazil

Lilian O. Guimarães[1], Gerhard Wunderlich[2], João M. P. Alves[2], Marina G. Bueno[3], Fabio Röhe[4], José L. Catão-Dias[3], Amanda Neves[5], Rosely S. Malafronte[5,6], Izilda Curado[7], Wilson Domingues[8] and Karin Kirchgatter[1*]

Abstract

Background: The merozoite surface protein 1 (MSP1) gene encodes the major surface antigen of invasive forms of the *Plasmodium* erythrocytic stages and is considered a candidate vaccine antigen against malaria. Due to its polymorphisms, MSP1 is also useful for strain discrimination and consists of a good genetic marker. Sequence diversity in MSP1 has been analyzed in field isolates of three human parasites: *P. falciparum*, *P. vivax*, and *P. ovale*. However, the extent of variation in another human parasite, *P. malariae*, remains unknown. This parasite shows widespread, uneven distribution in tropical and subtropical regions throughout South America, Asia, and Africa. Interestingly, it is genetically indistinguishable from *P. brasilianum*, a parasite known to infect New World monkeys in Central and South America.

Methods: Specific fragments (1 to 5) covering 60 % of the MSP1 gene (mainly the putatively polymorphic regions), were amplified by PCR in isolates of *P. malariae* and *P. brasilianum* from different geographic origin and hosts. Sequencing of the PCR-amplified products or cloned PCR fragments was performed and the sequences were used to construct a phylogenetic tree by the maximum likelihood method. Data were computed to give insights into the evolutionary and phylogenetic relationships of these parasites.

Results: Except for fragment 4, sequences from all other fragments consisted of unpublished sequences. The most polymorphic gene region was fragment 2, and in samples where this region lacks polymorphism, all other regions are also identical. The low variability of the *P. malariae msp1* sequences of these isolates and the identification of the same haplotype in those collected many years apart at different locations is compatible with a low transmission rate. We also found greater diversity among *P. brasilianum* isolates compared with *P. malariae* ones. Lastly, the sequences were segregated according to their geographic origins and hosts, showing a strong genetic and geographic structure.

Conclusions: Our data show that there is a low level of sequence diversity and a possible absence of allelic dimorphism of MSP1 in these parasites as opposed to other *Plasmodium* species. *P. brasilianum* strains apparently show greater divergence in comparison to *P. malariae*, thus *P. malariae* could derive from *P. brasilianum*, as it has been proposed.

Keywords: Malaria, *Plasmodium malariae*, Genetic diversity, Merozoite surface protein 1, Brazil, Atlantic forest

* Correspondence: karink@usp.br
[1]Núcleo de Estudos em Malária, Superintendência de Controle de Endemias/Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, SP 05403-000, Brazil
Full list of author information is available at the end of the article

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 2 of 11

Background

Malaria is a mosquito-borne disease that, in humans, is caused by five different species of *Plasmodium.* Most cases are caused by either *P. falciparum* or *P. vivax*, but human infections can also be caused by *P. ovale*, *P. malariae*, and, in parts of Southeast Asia, by *P. knowlesi*, a monkey malaria [1]. *P. malariae* is also very closely related to a malarial parasite that infects 12 genera of New World primates, *P. brasilianum* [2]. *P. malariae* appears to be genetically indistinguishable from *P. brasilianum*, although complete genome data supporting this conclusion is still unavailable. This points to a recent host transfer between humans and monkeys, but the direction of transfer is still controversial [3–5].

P. malariae is a cosmopolitan parasite which develops where the summer isotherm does not fall below 15 °C (59 °F). Its distribution is variable and spotty [2]. It has been identified throughout tropical regions in Africa, Asia, and South America. *P. brasilianum* has been reported in South and Central America. There are reports of natural infections of monkeys in Panama, Venezuela, Colombia, Peru, Brazil, and French Guiana [2, 6]. In Brazil, *P. malariae* and *P. brasilianum* are sparsely distributed in the Amazon Region and also in extra-Amazonian Atlantic forest regions, where malaria transmission is considered hypoendemic and controlled [7–9]. This tropical rainforest has an ample diversity of bromeliads that are an ideal environment for *Anopheles* mosquitoes of the subgenus *Kertesia*, mainly *An. (K.) cruzii*, which use the axils of these plants as larval habitat [10]. In this region, although only a small number of clinical cases of malaria are registered, a large portion of the population shows serological evidence of recent exposure to variants of *P. vivax* or *P. malariae* [11], suggesting a high prevalence of asymptomatic infection. These asymptomatic cases are often missed or underestimated because control programs are focused on passive and active investigation of symptomatic cases.

The merozoite surface protein 1 (MSP1) is the major surface antigen of invasive forms of the erythrocytic stages of *Plasmodium* and has been proposed as a vaccine antigen against malaria [12]. However, an extensive polymorphism in this protein was found in both *P. falciparum* and *P. vivax* from different geographical areas, representing a major obstacle to the development of an effective vaccine [13, 14]. Although the sequences from *P. falciparum* and *P. vivax* MSP1 genes have been described a long time ago [15, 16], *msp1* genes were more recently identified in *P. malariae* and *P. ovale* [17]. Recently, a study showed low levels of sequence diversity in MSP1 of *P. ovale* among Thai isolates [18]. However, until now, only two studies of *P. malariae* and/or *P. brasilianum msp1* sequences were performed. The first one, using only ~200 bp in the N-terminal region of the gene, showed that *P. brasilianum* has limited polymorphism in French Guiana [6]. A low variability (97 to 100 % identity) was also found using a longer fragment (~600 bp), also in the N-terminal region, between seven *P. brasilianum* sequences from monkeys of the state of Rondônia (Western Brazilian Amazon) [19].

MSP1 genes from *Plasmodium* species infecting mammals exhibit sizes of ~5 kb. The single copy *msp1* gene of *P. falciparum* and *P. vivax* is divided into blocks, based on analysis of interspecies sequence diversity: seven highly variable blocks are interspersed with ten conserved regions [20]. Although there is no evidence that the *msp1* sequences from all species could be divided in blocks, regions of interspecies conservation and variability across the entire MSP1 protein can be clearly identified [17]. In this study, the genetic polymorphism of five *msp1* gene regions, adding up to ~3000 bp, was analyzed among field isolates collected from different hosts in Brazil. These data were used to examine genetic diversity in relation to geographic origin and hosts.

Methods

Samples

Seventeen samples of *P. malariae/P. brasilianum* genomic DNA were used in this study (Table 1). Regarding the human samples, three were obtained in the Amazon Region (samples 23PA, 50PA, and 66PA) and eight were collected in the Atlantic forest (A, I11, 58, 72, 23a, 190, 157a, and 222a). Among these last samples, two were obtained from patients involved in a case of transfusion malaria: one sample was obtained from a splenectomized patient (58) who was infected by blood transfusion and another from her blood donor, an asymptomatic carrier (72), who probably acquired the infection in the location of Palestina, Juquitiba [21]. The remaining four samples (23a, 190, 157a, and 222a) were obtained during malaria control activities of this malaria focus. The non-human samples analyzed were: one *P. malariae/P. brasilianum* sample obtained from a pool of three *Anopheles Kerteszia cruzii* mosquitoes collected in Itanhaém (Atlantic forest region) and four *P. brasilianum* samples obtained from simian hosts captured in Igapó-Açú, roughly 300 km from Manaus (AM) (*Callicebus caligatus*) and Abunã, roughly 200 km from Porto Velho (RO) (*Callicebus dubius* and *Pithecia sp.*), both located in the Brazilian Amazon. The Peruvian III strain of *P. brasilianum* isolated from *Saimiri sciureus* monkeys of Iquitos in 1987 was used as a control [22]. For all the samples collected, the presence of *P. malariae/P. brasilianum* was verified using two PCR protocols and sequencing: (i) a nested PCR amplifying a small subunit ribosomal RNA (ssrRNA) based on the technique originally described by Snounou et al. [23, 24], in which the first amplification reaction was modified in order to increase its sensitivity [25]; and (ii) a nested PCR amplifying the mitochondrial cytochrome b gene (*cytb*) [26].

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 3 of 11

Table 1 Human and non-human samples used in this study

Human sample	Year of Collection	Origin	
23PA	1996	Peixoto de Azevedo, MT, Amazon Region, BR	
50PA	1996	Peixoto de Azevedo, MT, Amazon Region, BR	
66PA	1996	Peixoto de Azevedo, MT, Amazon Region, BR	
A	2000	Iguape, SP, Atlantic Forest, BR	
I11	2002	Iporanga, SP, Atlantic Forest, BR	
58	2004	São Paulo, SP (Blood transfusion from Patient 72)	
72	2004	Juquitiba, SP, Atlantic Forest, BR	
23a	2005	Juquitiba, SP, Atlantic Forest, BR	
190	2006	Juquitiba, SP, Atlantic Forest, BR	
157a	2007	Juquitiba, SP, Atlantic Forest, BR	
222a	2007	Juquitiba, SP, Atlantic Forest, BR	
Sample	Year of Collection	Source	Origin
PIII	1987	*Saimiri sciureus*	Iquitos, Peru
M95	2010	*Anopheles (K.) cruzii*	Itanhaém, SP, Atlantic Forest, BR
P169	2010	*Callicebus dubius*	Porto Velho, RO, Amazon Region, BR
P171	2010	*Pithecia sp.*	Porto Velho, RO, Amazon Region, BR
P177	2011	*Callicebus caligatus*	Manaus, AM, Amazon Region, BR
P182	2011	*Callicebus caligatus*	Manaus, AM, Amazon Region, BR

SP São Paulo state, *AC* Acre state, *RO* Rondônia state, *AM* Amazonas state, *BR* Brazil

Design of PmMSP1-specific primers

Ideally, primer design would have been guided by an alignment of *P. malariae/P. brasilianum msp1* sequences; however, only one complete *msp1* sequence has been published. Therefore, we have used an alignment of *msp1* from different *Plasmodium* species. We assume that regions conserved among different species are possibly also conserved within the same species. Thus, to focus our efforts on parts most likely to be genetically diverse, we performed an alignment for determination of percent amino acid sequence identity among nineteen MSP1 sequences, representing 16 *Plasmodium* species, collected from GenBank (Table 2). Calculation of amino acid sequence similarity levels was done using the plotcon program from the EMBOSS package [27], with plotting performed in R. Three regions with the lowest similarities among *Plasmodium* species were chosen. Additionally, two regions with high similarities interspecies were also chosen including the C-terminal region of the MSP1 (MSP1$_{19}$), the leading vaccine candidate. Therefore, five regions of the *msp1* gene were chosen for our analysis (Fig. 1).

To obtain *P. malariae/P. brasilianum*-specific primers from the MSP1 gene, nucleotide sequences flanking the five regions were designed based on the only sequence available for *P. malariae* (MM1A isolate from Cameroon, GenBank #FJ824669). Thus, these primers were located within conserved sequence stretches flanking the regions of interest but include nucleotide sequences characteristic of *P. malariae/P. brasilianum*. Primers nested within the first primer pair were also designed and were used in a nested protocol in order to increase the sensitivity of detection, since *P. malariae* or *P. brasilianum* infections usually present very low parasitemias. All oligonucleotides were checked for specificity by using the Primer-BLAST tool provided by the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). In experimental PCR, these oligonucleotides were also specific for *P. malariae/P. brasilianum* DNA, showing no amplification with DNAs from other *Plasmodium* species or hosts as templates (data not shown). Table 3 contains the sequences of all oligonucleotides used herein.

Amplification and sequencing of the PmMSP1

Primer sequences and amplification conditions used to obtain each fragment and the resulting sizes of the amplicons are shown in Table 3. Each 25 µl reaction mixture for nested1 amplifications contained 5 µl of DNA template, 250 nM of each primer, 2 mM MgCl$_2$, PCR buffer (50 mM KCl, 20 mM Tris–HCl, pH 8.8), 0.125 mM each of dATP, dCTP, dGTP, dTTP, and 0.5 units of Platinum® *Taq* DNA Polymerase (Invitrogen). One microliter of the nested1 amplification product served as the DNA template for each of the 25 µl nested2 amplifications. The concentration of the nested2 primers and other

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 4 of 11

Table 2 Merozoite surface protein 1 (MSP1) sequences of *Plasmodium* species

Plasmodium species (strain)	Host species	GenBank accession no.	
		Aminoacid	Nucleotide
berghei	*Grammomys surdaster*	AAF13063.1	AF187232.1
chabaudi	*Thamnonys rutilans*	AAA29499.1	L22982.1
coatneyi	*Macaca fascicularis*	BAF74048.1	AB266180.1
cynomolgi	*Macaca sinica*	BAF74063.1	AB266195.1
falciparum (MAD20)	*Homo sapiens*	A26868	X05624.2
falciparum (K1)	*Homo sapiens*	CAA27070.1	X03371.1
fragile	*Macaca sinica*	BAF74049.1	AB266181.1
gallinaceum	*Gallus gallus*	CAH10838.1	AJ809338.1
hylobati	*Hylobati moloch*	BAF74050.1	AB266182.1
inui	*Cynopithecus niger*	BAF74051.1	AB266183.1
knowlesi	*Macaca fascicularis*	BAF74052.1	AB266184.1
malariae (MM1A)	*Homo sapiens*	ACZ51237.1	FJ824669
ovale (OM1A)	*Homo sapiens*	ACZ51238.1	FJ824670
ovale (OM1B)	*Homo sapiens*	ACZ51239.1	FJ824671
reichenowi	*Pan troglodytes*	CAH10285.1	AJ786604.1
simiovale	*Macaca sinica*	BAF74053.1	AB266185.1
vivax (Belem)	*Homo sapiens*	AAN86208.1	AF435594.1
vivax (Sal1)	*Homo sapiens*	EDL45115.1	AAKM01000007.1
yoelii	*Thamnonys rutilans*	EAA17822.1	AABL01001865.1

constituents were identical to nested1 amplifications. The PCR products of nested2 amplifications were submitted to electrophoresis in 1 % agarose gel stained with 5×10^5-diluted GelRed nucleic acid stain (Biotium, Hayward, CA, USA). PCR fragments were purified and directly sequenced. For fragment 2, due to the presence of the microsatellite, PCR-fragments were cloned in pGEM-T Easy Vector (Promega) and positive clones were sequenced. Sequencing was performed using the Big Dye Terminator v3.0 Cycle Sequencing Kit in an ABI Genetic Analyzer (ABI, USA). The sequences obtained were aligned with the published *P. malariae msp1* sequence (FJ824669) using Clustal X version1.81.

Phylogenetic analysis

Phylogenetic analysis of *msp1* sequences was performed by maximum likelihood using RAxML v. 8.1.6, with data partitioned in three sets, based on codon position. The GTR model with gamma-distributed substitution heterogeneity rates was used, with base frequencies estimated
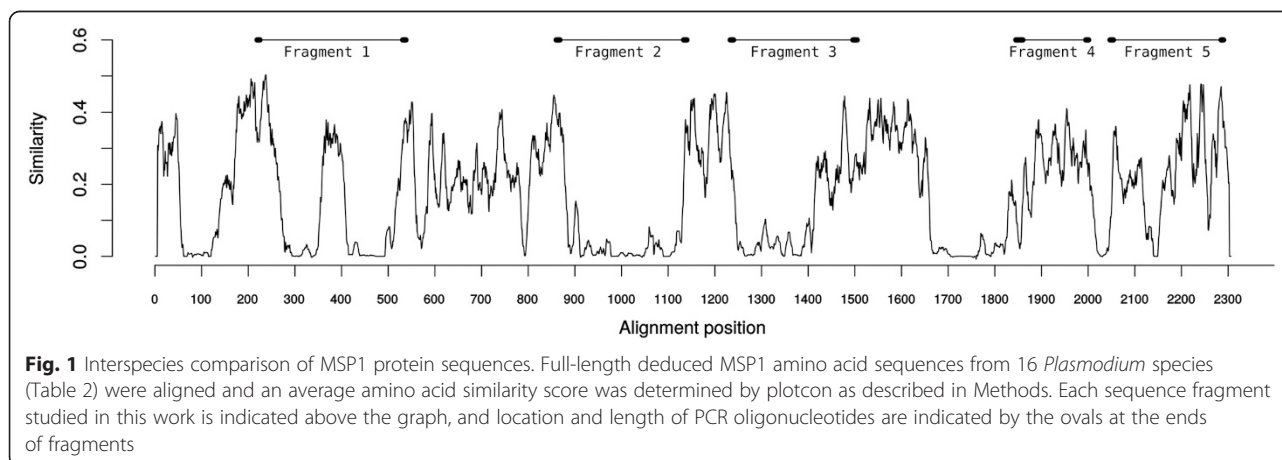

Fig. 1 Interspecies comparison of MSP1 protein sequences. Full-length deduced MSP1 amino acid sequences from 16 *Plasmodium* species (Table 2) were aligned and an average amino acid similarity score was determined by plotcon as described in Methods. Each sequence fragment studied in this work is indicated above the graph, and location and length of PCR oligonucleotides are indicated by the ovals at the ends of fragments

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 5 of 11

Table 3 PCR primers and amplification conditions

Region	PCR primers	Amplification conditions	Amplicon size[a]
Fragment 1			
First Reaction	F:5′-GCA TCT AAA TGG CTA CTG TGA TAT AC-3′	94 °C 5 min (1×); 94 °C 1 min, 53 °C 2 min, 72 °C 2 min (25×); 72 °C 5 min (1×)	595 bp
	R:5′-GGC ATC TGT AAA TAG ACC ATC C-3′		
Nested Reaction	F:5′-TAA TGA AAA GGA ATT AGA AAT G-3′	94 °C 5 min (1×); 94 °C 1 min, 43 °C 2 min, 72 °C 2 min (30×); 72 °C 5 min (1×)	533 bp
	R:5′-CCA TAT TGA ATA CTA TAT TTT TC-3′		
Fragment 2			
First Reaction	F: 5′-CCA TAC TAT TTA ATT GCA CTA AAG-3′	94 °C 5 min (1×); 94 °C 1 min, 58 °C 2 min, 72 °C 2 min (25×); 72 °C 5 min (1×)	723 bp
	R:5′-ACA CAC ATA AGC AGT TTT CAA AAA G-3′		
Nested Reaction	F:5′-AGG GAA ATT GAT AAA TTA AAT ATT TC-3′	94 °C 5 min (1×); 94 °C 1 min, 58 °C 2 min, 72 °C 2 min (30×); 72 °C 5 min (1×)	666 bp
	R:5′-CTT TTC AAG ATA TTG CAA TTT GGA-3′		
Fragment 3			
First Reaction	F:5′-GCA ATG TTC TAC AAA ATC AGT ACA AAG-3′	94 °C 5 min (1×); 94 °C 1 min, 53 °C 2 min, 72 °C 2 min (25×); 72 °C 5 min (1×)	765 bp
	R:5′-CTA CAA AAG CTG CTA GTA CAT GTC T-3′		
Nested Reaction	F:5′-GAA GAA GAA TGA CAA ACT TAA AAA C-3′	94 °C 5 min (1×); 94 °C 1 min, 43 °C 2 min, 72 °C 2 min (30×); 72 °C 5 min (1×)	632 bp
	R:5′-CTT TTA AAA TAC TAT ATT CTT TAA TAT G-3′		
Fragment 4			
First Reaction	F:5′-GAG GTG ATC GTG TTT CCC ATT G-3′	94 °C 5 min (1×); 94 °C 1 min, 48 °C 2 min, 72 °C 2 min (25×); 72 °C 5 min (1×)	532 bp
	R:5′-CAT TTT CAG TTT TTT TAT CAC CC-3′		
Nested Reaction	F:5′-GTA AGA AAG AAA AGG AAA ATC CAT TAG-3′	94 °C 5 min (1×); 94 °C 1 min, 49 °C 2 min, 72 °C 2 min (30×); 72 °C 5 min (1×)	443 bp
	R:5′-CTG AGT CTT GTA CAA CTT GGT C-3′		
Fragment 5			
First Reaction	F:5′-GAC CAA GTT GTA CAA GAC TCA G-3′	94 °C 5 min (1×); 94 °C 1 min, 48 °C 2 min, 72 °C 2 min (25×); 72 °C 5 min (1×)	737 bp
	R:5′-GTA AGT TAA ACA TAA TTA ATA AAG CTG-3′		
Nested Reaction	F: 5′-GGG TGA TAA AAA AAC TGA AAA TG-3′	94 °C 5 min (1×); 94 °C 1 min, 49 °C 2 min, 72 °C 2 min (30×); 72 °C 5 min (1×)	631 bp
	R:5′-GCC GAG GAA ACT TGA AGA AC-3′		

[a]This size is estimated based on that of the sequence from isolate MM1 deposited in GenBank (#FJ824669)

by maximum likelihood, and 1000 bootstrap pseudore-plicates to assess clade support. The substitution model used was chosen based on a preliminary jModelTest analysis [28], which selected model TPM3uf + I + G; since this model is not available in RAxML, we have used the closest model available (GTR + G), without a proportion of invariants in order to avoid modeling the same aspect of molecular evolution twice. The tree was drawn in Dendroscope [29], with cosmetic adjustments performed in Inkscape (http://www.inkscape.org). The resulting protein sequence alignment and phylogenetic tree are deposited in TreeBase under the accession number S18381.

Ethical considerations

The study protocol was approved by the Research Ethics Committee of the Institute of Biomedical Sciences of the University of São Paulo, Brazil (CEPSH 749/2006 and 065/2008), which included getting patient consent for research use of their blood samples. All the procedures adopted in this work fully complied with specific federal permits issued by the Brazilian Ministry of the Environment (SISBIO, process numbers 18861–3 and 24319–3).

Results

Amplification and sequencing of MSP1 fragments from *P. malariae/P. brasilianum* isolates

Five regions of the *msp1* gene, totalizing ~3000 bp, were amplified and sequenced for the samples presented in Table 1 (GenBank accession numbers JX045639-JX045645 and KR072215-KR072284). In addition to these samples, other *P. malariae* specimens were also tested, but proved negative for PCR amplification of the *msp1* gene. The quantity of target DNA was likely too small in these samples, even though *P. malariae* was positive by nested PCR of the ssrRNA genes. This probably occurred because ssrRNA genes are multiple copy and *msp1* is possibly a single copy gene in *P. malariae*, similar to other *Plasmodium* species. There are also some isolates with sequences of some fragments missing, since genomic DNA from these samples was not sufficient to perform all amplification reactions (Fig. 2).
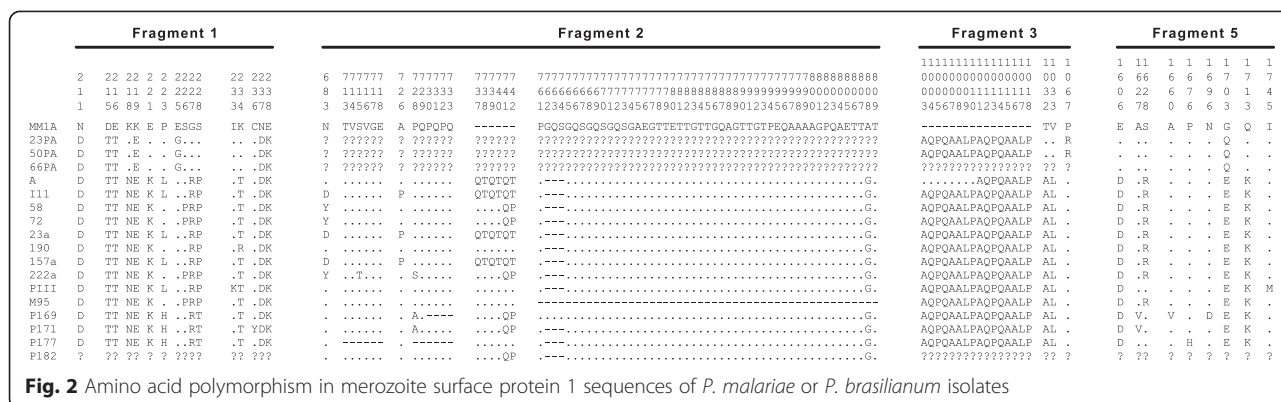
Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 6 of 11

Fig. 2 Amino acid polymorphism in merozoite surface protein 1 sequences of *P. malariae* or *P. brasilianum* isolates

Sequence diversity seems limited to specific regions inside analyzed fragments

Fragment 1 encodes a 177 residue peptide in our samples, corresponding to amino acids 160 to 336 in the unaligned MM1 sequence (GenBank #FJ824669). Only the region between amino acids 211 and 238 was polymorphic in our alignment (Fig. 2). Seven different amino acid sequences were found for this fragment. The three samples from the Amazon Region were identical in this *msp1* region and different from all the rest. The samples from *P. brasilianum* were similar to each other and different from human samples but closer to those from Atlantic forest (Fig. 2). Fragment 2 corresponds to amino acids 637 to 858 in the MM1 sequence, but the polymorphic amino acids were found only between positions 683 and 809 of the alignment (Fig. 2). Ten different amino acid sequences were found in this fragment with sizes ranging from 173 to 225 residues. Unfortunately, it was not possible to determine the sequences of the human isolates from the Amazon Region due to the limited amount of genomic DNA available. Many indels were found in relation to the MM1A sequence. There is also a region of a six amino acid indel in four of our sequences from human isolates collected in the Atlantic forest, while six others have only two extra amino acids in comparison to the MM1A sequence. The sequence from mosquito-derived material presented a great deletion in this region (49 amino acids). From the eight *P. malariae* samples, five different sequences were recorded, while each of the five *P. brasilianum* samples presented a unique sequence.

For fragment 3, corresponding to amino acids 950 to 1159 (210 aa) in the MM1 sequence, only the region between the amino acids 1003 and 1067 was polymorphic in our samples (Fig. 2). An insertion of 8 or 16 residues (either one or two copies of AQPQAALP) in relation to the MM1 isolate was found in all of our sequences. However, only three different sequences were found for this fragment. The three samples from the Amazon Region were identical in this *msp1* region and different from all the rest due to just one amino acid (R in position 1067). *P. brasilianum*

samples were identical to each other and also to human isolates from the Atlantic forest (excluding the sequence from isolate A, which presented the only 8 amino acid insertion). Fragment 4, with 147 residues corresponding to amino acids 1366 to 1512 in the MM1 sequence, was found to be monomorphic in all the isolates analysed in this study, including *P. malariae* and *P. brasilianum*, from mosquito, simian or human hosts (Fig. 2). Even the nucleotide sequence was the same as that found for isolate MM1, from Cameroon.

In the C-terminal region, fragment 5 was amplified, corresponding to amino acids 1527 to 1736 (210 aa) in the MM1 sequence. Its variability was restricted to 9 amino acids, between residues 1606 and 1745 (Fig. 2). The $MSP1_{19}$ region corresponds to 89 amino acids, between positions 1641 and 1729. Sequences from Amazonian isolates differed from the sequence from Cameroon only by a glutamine residue (position 1703) that was not found in other sequences from Brazil. Sequences from Atlantic forest hosts presented three conserved amino acids: D, E and K, respectively in positions 1606, 1703, and 1713 that were also found in the Amazonian simian isolates from Brazil and Peru. All *P. brasilianum* sequences were unique, and the one from Peru presented a methionine in position 1745 that was the only polymorphism found in this position.

A more quantitative presentation of the amounts of differences between *msp1* fragments analyzed in this work is presented both for each separate fragment as well as in aggregated form representing the concatenation of all fragments (Additional file 1).

Similar haplotypes in related infections from different patients

Regarding the malaria case acquired by blood transfusion previously reported [21], the *P. malariae* sample obtained from the donor (72) showed the same sequence as the recipient's (58). This provides additional evidence that the donor and recipient were infected with the same strain, as expected.

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 7 of 11

Phylogeny analysis shows genetic and geographic structuring of isolates

Figure 3 shows the result from the partitioned maximum likelihood phylogenetic analysis, with branch lengths omitted due to the very high level of sequence similarity between samples (the expected substitution per site bar is 0.001 units long). We found a strong genetic and geographic structure in the phylogeny. The *msp1* sequences were segregated according to their geographic origin and hosts. Regarding sequences obtained from human isolates from the Amazon, we found a well-supported clade that separates them from all other Brazilian sequences. The *msp1* sequences obtained from human isolates collected in the Atlantic forest were also clustered in a clade that was positioned near sequences from simian isolates obtained in the Amazon Region, forming a well-supported clade with a high support value. In general, clades presented low (<50) support values, due to the very high similarity among sequences from the same geographic origin or hosts and the consequent lack in phylogenetic signal (sequences

that are too similar do not present enough differences to allow their separation in reliable clades). A Bayesian inference was also performed and showed similar results.

Discussion

This work presents the first molecular epidemiological data based on the *msp1* gene for *P. malariae/P. brasilianum* samples from Brazil that were obtained from different hosts and from different gene regions. Except for fragment 4, which was monomorphic in our samples, all sequences were different from the MM1A sequence, which was found in a human isolate from Cameroon [17]. As expected, the most polymorphic region was found in fragment 2, already known for including imperfect repeats [17]. Analysis of the percent similarity among the fragment 2 amino acid sequences from all isolates showed identities ranging from 89 to 100 %, and the lowest percentage of identity (89 %) was found between sequences obtained from simian and human samples. Similar results were found in another region of the *msp1* gene (upstream from
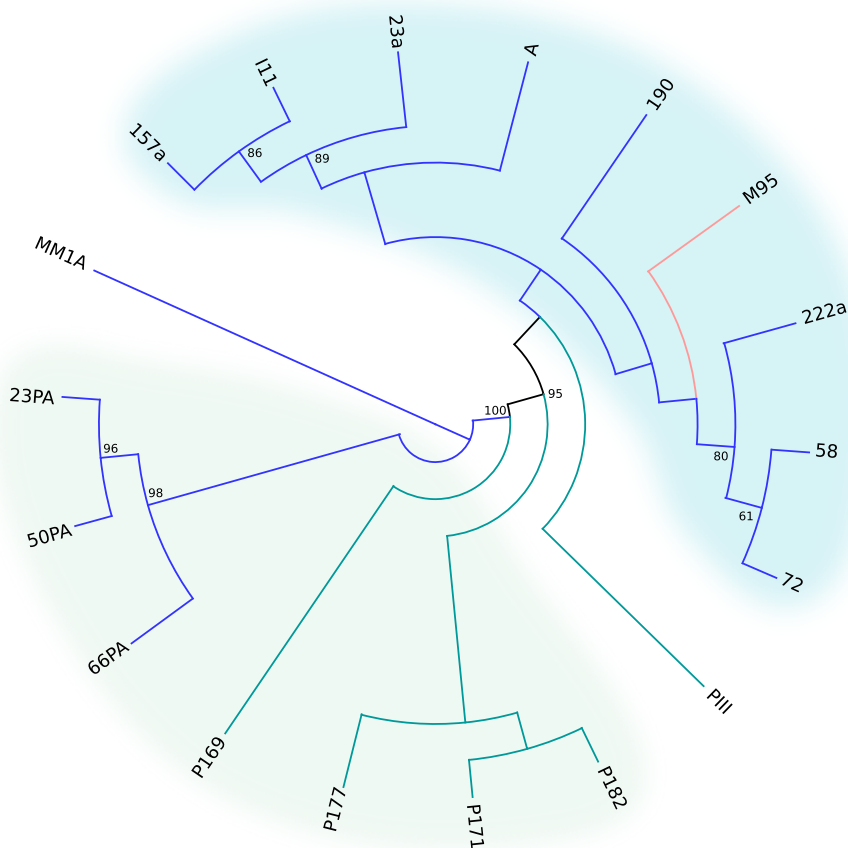

Fig. 3 Maximum likelihood phylogeny pattern (branch lengths not displayed) of *msp1* sequences of *P. malariae* (human samples, dark blue branches), *P. brasilianum* (simian samples, light blue branches) and *P. malariae/P. brasilianum* (mosquito sample, red branch) isolates from Cameroon (MM1), Brazilian Amazon Region (green background) or Atlantic forest (blue background). Names of isolates are given according to Table 1. Numbers on nodes indicate bootstrap support values in percentage. Only values above 50 are shown

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 8 of 11

this work's fragment 1, corresponding to amino acids 12 to 140 in the MM1 sequence), which showed a polymorphism limited with identities ranging from 93 to 100 % between *P. malariae* and *P. brasilianum* sequences from French Guiana [6].

The repeats on fragment 2 facilitate recombination events, making this region useful as an informative genetic marker. Our results have also shown that, most likely, an absence of polymorphism in this *msp1* region reflects stability also in other regions. The corresponding region in *P. falciparum* and *P. vivax* proved to be likewise the most polymorphic one in many studies [14, 20, 30–32]. However, as opposed to what has been found with *P. falciparum*, our study indicates a likely absence of allelic dimorphism in *P. malariae* and *P. brasilianum msp1* genes, as seen in other *Plasmodium* species [18, 33]. Nevertheless, the possibility of some sequences being so different that our oligonucleotides would be unable to amplify them cannot be ruled out. In this case, the domain architecture would be fundamentally different between *P. malariae/P. brasilianum* and other *Plasmodium* species analyzed so far. It is important to note that, more recently, low levels of sequence diversity at the *msp1* locus have also been found for *P. ovale curtisi* and *P. ovale wallikeri* from Thailand [18].

Blood stages of *Plasmodium* are haploid and the detection of multiple sequences in a human or simian sample indicates the presence of multiple genotypes. However, we also tested a sample from a mosquito, where the *Plasmodium* stages can be haploid (sporozoites) or diploid (ookinete and oocyst). Since multiple sequences were not detected, we could assume that the infections are clonal or there is just one infection in haploid stage. As the mosquitoes were used without severing of head/thorax (sporozoites) from abdomen (ookinete and oocyst), this cannot be determined. These results were obtained with sequencing of cloned fragments and confirmed also by sequencing directly from PCR. Interestingly, the fragment 2 sequence obtained from the mosquito *Plasmodium* sample showed a rather different size, presenting a region with a large deletion (about a third of the amplified sequence), but was highly conserved in its sequence when compared to sequences found in human isolates of the same region, indicating a likely recombination event in this hotspot region.

The main origin of the isolates used in this study was an area with persistently low prevalence of *P. malariae* and *P. vivax* located at the Atlantic forest region of São Paulo State [9]. The low variability of the *msp1* sequences of these isolates indicates the likely "clonal" source and low possibility of future heterologous recombination. In fact, as well as for *msp1* sequences from *P. ovale* [18], the same *P. malariae* haplotype was identified in isolates collected many years apart and in different locations. Considering that the range of flight for *An. (Ker.) cruzii* is roughly 1 km

[34], the spread of the same haplotype could be attributable to human movement. The low level of *msp1* sequence diversity from *P. malariae* could indicate, besides the low transmission rate, events of genetic bottlenecks probably caused by malaria control activities.

Relationship of *msp1* sequence with geographical features and hosts aiming to give insights into the evolutionary and phylogenetic relationships of *P. brasilianum* and *P. malariae*

In terms of geographical origin, *msp1* sequences were informative to distinguish the two main sites of sample origin (Amazon region and Atlantic forest), grouping them in two different clades. Similarly, in relation to the hosts, sequences from parasites with the same host species were clustered. According to the phylogenetic analysis, the sequence obtained from the mosquito sample was included in the group of human isolates from the Atlantic forest, indicating its probable source of infection. Interestingly, sequences from human parasites isolated from the Atlantic forest clustered with the sequences from simian parasites isolated from the Amazon region instead of with those from human parasites isolated in the Amazon region. This indicates that *P. malariae* from human residents in the Amazon Region are genetically different. Their clustering in the phylogenetic tree, with the human isolate from Africa could indicate its origin and also would explain why many of these infections are symptomatic [35]. On the other hand, for the Atlantic forest region, the clustering of sequences from human isolates with sequences from simian isolates supports the idea of zoonotic malaria in this region [7], although it is very likely that it is being currently maintained in the region due to asymptomatic individuals who are acting as main source of parasites [36].

The high genetic identity found between *P. malariae* from humans and *P. brasilianum* from New World primates suggests the occurrence of a recent transfer between these hosts [3–5], but there is no evidence to determine whether the transference occurred from monkeys to humans or vice versa. The direction of the host transfer can then be suggested by comparing the genetic diversity of the human and primate parasites [37]. Although studies focusing on these species are rare, some works indicated that *P. brasilianum* strains seem to show greater divergence than *P. malariae*, thus *P. malariae* could derive from *P. brasilianum* [37, 38]. In this work, we also identified greater diversity among *P. brasilianum* isolates that could support this hypothesis. In contrast, *msp1* sequence analysis of *Plasmodium knowlesi* from humans and macaques in Thailand revealed that the diversity of human-derived sequences exceeded that of monkey-derived sequences [39].

On the other hand, in the case of *P. malariae*, the host switch between humans and monkeys might have occurred

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 9 of 11

through a single transfer associated with a bottleneck, after the European colonization of the Americas in the 16th century [3] and thus, *P. brasilianum* would be derived from *P. malariae*. An analysis of the sequence sizes in fragment 2 could support this hypothesis. It has been thought that repeats present in longer stretches are likely to be older than short ones, as the former needs more time for accumulation of base substitutions [40]. From this point of view, our data suggests that the simian sequences (with shorter stretches in this work) could have derived from the human sequences (longer stretches in this work). Anyway, microsatellite evolution is a dynamic process during which repeats may shrink as well as expand over an evolutionary time scale [41].

Recently, *P. malariae* has been found from chimpanzees in West [42] and Central Africa [43, 44] and bonobos in Central Africa [45]. However, the mitochondrial genomes of the parasites related to *P. malariae* found in bonobos carry a six nucleotide insert that has not been observed for *P. malariae* or the South American parasite *P. brasilianum* [45]. Analysis of *msp1* genetic polymorphism in such isolates could potentially shed light on the question of origin and host switch for these parasites.

The assumption that infection diversity and multiplicity is positively correlated with the frequency of transmission is reasonable [46]. Low allelic diversity was detected in Brazilian isolates, which seems to reflect a very low transmission rate. It has been established that *P. malariae* shows a reduced growth rate [47] and one hypothesis proposed that the low parasitemia obtained *in vivo* may be caused by its preference for mature erythrocytes [48]. *P. malariae* is also known to promote infections that are frequently long-lasting (up to 40 years absence from reinfection) and non-symptomatic [49]. Exclusive infections with *P. malariae*, like those found in our study area, would also explain the very low transmission rates seen in this region. In other locations, where *P. malariae* occurs in coinfections, mainly with *P. falciparum*, it has developed increased transmission efficiency in the presence of a coinfecting parasite of another species [50]. In areas of high transmission of all species, *P. falciparum* is dominant for a while and is then being succeeded by *P. vivax*, which holds center stage for some time, then being replaced by *P. malariae*, causing a "residual infestation" [51]. At the end of the XIX century, malaria (caused by *P. falciparum*, *P. vivax* and *P. malariae*) was present throughout the entire Brazilian territory, but after a successful World Health Organization (WHO) malaria eradication campaign, initiated in 1956, the number of cases decreased drastically outside the Brazilian Amazon. Currently, malaria transmission is primarily concentrated in the Amazon. A residual, low malaria transmission (0.05 % of all Brazilian malaria cases), occurs in the Atlantic Forest, caused by *P. vivax* or *P. malariae* [9]. Thus, *P. malariae* uses its long association with man 'to learn' and to cope with difficulty and succeed [51].

Conclusions

Evidence of geographical clustering of haplotypes was verified even though using a small sample available. However, further analysis with other Brazilian samples of different geographical origins is important to characterize the epidemiology of Brazilian *P. malariae* and *P. brasilianum* isolates in a higher resolution. Comparisons between human and simian *Plasmodium* isolates are also of interest, particularly in areas of the Atlantic forest where humans and monkeys are at risk of malaria transmission by the same anopheline vectors. Additional studies concerning the characterization of antigenic diversity in vaccine candidate antigens are valuable for future vaccine trials as well as for understanding the population dynamics of *P. malariae* and *P. brasilianum* parasites.

Additional file

Additional file 1: Pairwise similarity matrix among *P. malariae/P. brasilianum* isolates based on the MSP1 protein sequence fragments. The matrices on the left show, for each fragment separately, the raw numbers of identical amino acids and total amino acids compared between each pair of isolates. The matrices on the right show the aggregated similarity or difference values for the concatenated regions.(XLS 40 kb)

Author details
[1]Núcleo de Estudos em Malária, Superintendência de Controle de Endemias/ Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, SP 05403-000, Brazil. [2]Departamento de Parasitologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, SP 05508-900, Brazil. [3]Departamento de Patologia, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, SP 05508-270, Brazil. [4]Wildlife Conservation Society, Rio de Janeiro, RJ 22461-000, Brazil. [5]Laboratório de Protozoologia, Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, SP 05403-000, Brazil. [6]Departamento de Moléstias Infecciosas e Parasitárias, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP 01246-903, Brazil. [7]Laboratório de Imunoepidemiologia, Superintendência

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 10 of 11

de Controle de Endemias, São Paulo, SP 01027-000, Brazil. [8]Laboratório de Soroepidemiologia e Imunobiologia, Instituto de Medicina Tropical de São Paulo, Universidade de São Paulo, São Paulo, SP 05403-000, Brazil.

References

1. White NJ, Pukrittayakamee S, Hien TT, Faiz MA, Mokuolu OA, Dondorp AM. Malaria Lancet. 2014;383(9918):723–35.
2. Coatney GR, Collins WE, Warren M, Contacos PG. The primate malarias [original book published 1971]. Division of parasitic disease, producers. Version 1.0. Atlanta, GA: CDC; 2003.
3. Escalante AA, Barrio E, Ayala FJ. Evolutionary origin of human and primate malarias: evidence from the circumsporozoite protein gene. Mol Biol Evol. 1995;12:616–26.
4. Escalante AA, Freeland DE, Collins WE, Lal AA. The evolution of primate malaria parasites based on the gene encoding cytochrome b from the linear mitochondrial genome. Proc Natl Acad Sci U S A. 1998;95:8124–9.
5. Qari SH, Shi YP, Pieniazek NJ, Collins WE, Lal AA. Phylogenetic relationship among the malaria parasites based on small subunit rRNA gene sequences: monophyletic nature of the human malaria parasite, Plasmodium falciparum. Mol Phylogenet Evol. 1996;6:157–65.
6. Fandeur T, Volney B, Peneau C, de Thoisy B. Monkeys of the rainforest in French Guiana are natural reservoirs for *P. brasilianum/P. malariae* malaria. Parasitol. 2000;120:11–21.
7. Deane LM. Simian malaria in Brazil. Mem Inst Oswaldo Cruz. 1992;87(S3):1–20.
8. Oliveira-Ferreira J, Lacerda MV, Brasil P, Ladislau JL, Tauil PL, Daniel-Ribeiro CT. Malaria in Brazil: an overview. Malar J. 2010;9:115.
9. Pina-Costa A, Brasil P, Di Santi SM, de Araujo MP, Suárez-Mutis MC, Santelli AC, et al. Malaria in Brazil: what happens outside the Amazonian endemic region. Mem Inst Oswaldo Cruz. 2014;109:618–33.
10. Forattini OP, Kakitani I, Massad E, Marucci D. Studies on mosquitoes (Diptera: Culicidae) and anthropic environment. 11–Biting activity and blood-seeking parity of *Anopheles (Kerteszia)* in south-eastern Brazil. Rev Saude Publica. 1996;30:107–14.
11. Curado I, Duarte AM, Lal AA, Oliveira SG, Kloetzel JK. Antibodies anti bloodstream and circumsporozoite antigens (*Plasmodium vivax* and *Plasmodium malariae/P. brasilianum*) in areas of very low malaria endemicity in Brazil. Mem Inst Oswaldo Cruz. 1997;92:235–43.
12. Greenwood BM, Targett GA. Malaria vaccines and the new malaria agenda. Clin Microbiol Infect. 2011;17:1600–7.
13. Kadekoppala M, Holder AA. Merozoite surface proteins of the malaria parasite: the MSP1 complex and the MSP7 family. Int J Parasitol. 2010;40:1155–61.
14. Brito CF, Ferreira MU. Molecular markers and genetic diversity of *Plasmodium vivax*. Mem Inst Oswaldo Cruz. 2011;106(S1):12–26.
15. Holder AA, Lockyer MJ, Odink KG, Sandhu JS, Riveros-Moreno V, Nicholls SC, et al. Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. Nature. 1985;317(6034):270–3.
16. del Portillo HA, Longacre S, Khouri E, David PH. Primary structure of the merozoite surface antigen 1 of *Plasmodium vivax* reveals sequences conserved between different *Plasmodium* species. Proc Natl Acad Sci U S A. 1991;88:4030–4.
17. Birkenmeyer L, Muerhoff AS, Dawson GJ, Desai SM. Isolation and characterization of the MSP1 genes from *Plasmodium malariae* and *Plasmodium ovale*. Am J Trop Med Hyg. 2010;82:996–1003.
18. Putaporntip C, Hughes AL, Jongwutiwes S. Low level of sequence diversity at merozoite surface protein-1 locus of *Plasmodium ovale curtisi* and *P. ovale wallikeri* from Thai isolates. PLoS One. 2013;8(3):e58962.
19. Araújo MS, Messias MR, Figueiró MR, Gil LH, Probst CM, Vidal NM, et al. Natural *Plasmodium* infection in monkeys in the state of Rondonia (Brazilian Western Amazon). Malar J. 2013;12:180.
20. Tanabe K, Mackay M, Goman M, Scaife JG. Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. J Mol Biol. 1987;195:273–87.
21. Di Santi SMF, Carvalho ME, Costa MJ, Kirchgatter K, Pereira BF, Toniolo CRC, et al. Malária transfusional causada por *Plasmodium malariae* transmitido por doador assintomático infectado na Mata Atlântica do Estado de São Paulo [abstract]. Rev Soc Bras Med Trop. 2005;38:333.
22. Collins WE, Ruebush TK 2nd, Skinner JC, Filipski VK, Broderson JR, Stanfill PS, et al. The Peruvian III strain of *Plasmodium brasilianum* in *Saimiri sciureus boliviensis* monkeys. J Parasitol. 1990;76:676–80.
23. Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, Rosário VE, et al. High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. Mol Biochem Parasitol. 1993;61:315–20.
24. Singh B, Bobogare A, Cox-Singh J, Snounou G, Shukri Abdullah M, Abdul RH. A genus- and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies. Am J Trop Med Hyg. 1999;60:687–92.
25. dos Santos LC, Curotto SM, de Moraes W, Cubas ZS, Costa-Nascimento MJ, de Barros Filho IR, et al. Detection of *Plasmodium* sp. in capybara. Vet Parasitol. 2009;163:148–51.
26. Perkins SL, Schall JJ. A molecular phylogeny of malarial parasites recovered from cytochrome b gene sequences. J Parasitol. 2002;88:972–8.
27. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000;16:276–7.
28. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9(8):772.
29. Huson DH, Scornavacca C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. Syst Biol. 2012;61(6):1061–7.
30. Rich SM, Ayala FJ. Population structure and recent evolution of *Plasmodium falciparum*. Proc Natl Acad Sci U S A. 2000;97:6994–7001.
31. Rich SM, Ferreira MU, Ayala FJ. The origin of antigenic diversity in *Plasmodium falciparum*. Parasitol Today. 2000;16:390–6.
32. Roy SW, Ferreira MU, Hartl DL. Evolution of allelic dimorphism in malarial surface antigens. Heredity (Edinb). 2008;100(2):103–10.
33. Tanabe K, Escalante A, Sakihama N, Honda M, Arisue N, Horii T, et al. Recent independent evolution of *msp1* polymorphism in *Plasmodium vivax* and related simian malaria parasites. Mol Biochem Parasitol. 2007;156(1):74–9.
34. Ferreira E, Corrêa RR, Tomich A, de Sá FT. Flying range of *Anopheles (Kerteszia) cruzii* and *Anopheles (Kerteszia) bellator* in Guaratuba, coast line of the State of Paraná, Brazil. Rev Bras Malariol Doencas Trop. 1969;21:819–22.
35. Scopel KK, Fontes CJ, Nunes AC, Horta MF, Braga EM. High prevalence of *Plasmodium malariae* infections in a Brazilian Amazon endemic area (Apiacás-Mato Grosso State) as detected by polymerase chain reaction. Acta Trop. 2004;90(1):61–4.
36. Kirchgatter K, Tubaki RM, Malafronte RS, Alves IC, Lima GF, Guimarães LO, et al. *Anopheles (Kerteszia) cruzii* (Diptera: Culicidae) in peridomiciliary area during asymptomatic malaria transmission in the Atlantic Forest: molecular identification of blood-meal sources indicates humans as primary intermediate hosts. Rev Inst Med Trop Sao Paulo. 2014;56:403–9.
37. Tazi L, Ayala FJ. Unresolved direction of host transfer of *Plasmodium vivax* v. *P. simium* and *P. malariae* v. *P. brasilianum*. Infect Genet Evol. 2011;11:209–21.
38. Guimarães LO, Bajay MM, Wunderlich G, Bueno MG, Röhe F, Catão-Dias JL, et al. The genetic diversity of *Plasmodium malariae* and *Plasmodium brasilianum* from human, simian and mosquito hosts in Brazil. Acta Trop. 2012;124:27–32.
39. Putaporntip C, Thongaree S, Jongwutiwes S. Differential sequence diversity at merozoite surface protein-1 locus of *Plasmodium knowlesi* from humans and macaques in Thailand. Infect Genet Evol. 2013;18:213–9.
40. Bell GI, Jurka J. The length distribution of perfect dimer repetitive DNA is consistent with its evolution by an unbiased single-step mutation process. J Mol Evol. 1997;44:414–21.
41. Ellegren H. Microsatellites: simple sequences with complex evolution. Nat Rev Genet. 2004;5:435–45.
42. Hayakawa T, Arisue N, Udono T, Hirai H, Sattabongkot J, Toyama T, et al. Identification of *Plasmodium malariae*, a human malaria parasite, in imported chimpanzees. PLoS One. 2009;4:e7412.
43. Duval L, Fourment M, Nerrienet E, Rousset D, Sadeuh SA, Goodman SM, et al. African apes as reservoirs of *Plasmodium falciparum* and the origin and diversification of the *Laverania* subgenus. Proc Natl Acad Sci U S A. 2010;107(23):10561–6.
44. Pacheco MA, Cranfield M, Cameron K, Escalante AA. Malarial parasite diversity in chimpanzees: the value of comparative approaches to ascertain the evolution of *Plasmodium falciparum* antigens. Malar J. 2013;12:328.
45. Krief S, Escalante AA, Pacheco MA, Mugisha L, André C, Halbwax M, et al. On the diversity of malaria parasites in African apes and

Guimarães *et al. BMC Infectious Diseases* (2015) 15:529

Page 11 of 11

the origin of *Plasmodium falciparum* from Bonobos. PLoS Pathog. 2010;6(2):e1000765.

46. Babiker HA, Walliker D. Current views on the population structure of *Plasmodium falciparum*: Implications for control. Parasitol Today. 1997;13:262–7.

47. Gilles HM, Warrell DA. Bruce-Chwatt's essential malariology. London: Edward Arnold; 1993.

48. Garnham PCC. Malaria parasites and other haemosporidia. Oxford: Blackwell; 1996.

49. Vinetz JM, Li J, McCutchan TF, Kaslow DC. *Plasmodium malariae* infection in an asymptomatic 74-year-old Greek woman with splenomegaly. N Engl J Med. 1998;338:367–71.

50. Bruce MC, Macheso A, McConnachie A, Molyneux ME. Comparative population structure of *Plasmodium malariae* and *Plasmodium falciparum* under different transmission settings in Malawi. Malar J. 2011;10:38.

51. Field JW, Shute PG. The microscopic diagnosis of human malaria. II. A morphological study of the erythrocytic parasites. Inst Med Res. 1956;24:251.