

RESEARCH

Open Access

# Analysis of *pfhrp2* genetic diversity in Senegal and implications for use of rapid diagnostic tests

Awa B Deme<sup>1</sup>, Daniel J Park<sup>2</sup>, Amy K Bei<sup>1,3</sup>, Ousmane Sarr<sup>1</sup>, Aida Sadikh Badiane<sup>1</sup>, Papa El Hadji Omar Gueye<sup>1</sup>, Ambroise Ahouidi<sup>1</sup>, Omar Ndir<sup>4</sup>, Souleymane Mboup<sup>1</sup>, Dyann F Wirth<sup>2,3</sup>, Daouda Ndiaye<sup>4</sup> and Sarah K Volkman<sup>2,3,5\*</sup>

## Abstract

**Background:** The Senegalese National Malaria Control Programme has recommended use of rapid diagnostic tests (RDTs) that target the histidine-rich protein 2 (HRP2), specific to *Plasmodium falciparum*, to diagnose malaria cases. The target antigen has been shown to be polymorphic, which may explain the variability in HRP2-based RDT results reported in field studies. The genetic diversity of the *pfhrp2* gene has not been investigated in depth in many African countries. The goal of this study is to determine the extent of polymorphism in *pfhrp2* among Senegal, Mali and Uganda parasite populations, and discuss the implications of these findings on the utility of RDTs that are based on HRP2 detection.

**Methods:** Sequencing data from the *pfhrp2* locus were used to analyze the genetic diversity of this gene among three populations, with different transmission dynamics and malaria parasite ecologies. Nucleotide diversity ( $\pi$ ) and non-synonymous nucleotide diversity ( $\pi_{NS}$ ) were studied in the *pfhrp2* gene from isolates obtained in Senegal. Amino acid repeat length polymorphisms in the PfHRP2 antigen were characterized and parameters of genetic diversity, such as frequency and correlation between repeats in these populations, were assessed.

**Results:** The diversity survey of the *pfhrp2* gene identified 29 SNPs as well as insertion and deletion polymorphisms within a 918 bp region. The Senegal *pfhrp2* exhibited a substantial level of diversity [ $\pi = 0.00559$  and  $\pi_{NS} = 0.014111$  ( $\pi_S = 0.0291627$ )], similar to several polymorphic genes, such as *msp1*, involved in immune responses, and the gene encoding the SURFIN polymorphic antigen, which are surface exposed parasite proteins. Extensive repeat length polymorphisms in PfHRP2, as well as similar patterns in the number, organization and the type of predicted amino acid repeats were observed among the three populations, characterized by an occurrence of Type 2, Type 4 and Type 7 repeats.

**Conclusions:** These results warrant deeper monitoring of the RDT target antigen diversity and emphasize that development of other essential genes as a target for diagnostic tools is critical.

**Keywords:** *Plasmodium falciparum*, HRP2, Diversity, RDT, Senegal

## Background

For over 50 years various strategies have been developed to reduce the burden of *Plasmodium falciparum*, the most virulent malaria species and the primary cause of malaria-related mortality across the globe [1]. International funding programmes have first tasked the eradication

of this disease and, as a follow-up for regional elimination, have provided tools such as insecticide-treated bed nets (ITNs), indoor residual spraying (IRS), artemisinin-based combination therapy (ACT), and rapid diagnostic tests (RDTs) to many malaria-endemic regions. This report focuses on the implications of genetic diversity in the *pfhrp2* gene that may compromise the use of RDTs for the diagnosis of malaria.

In 2006, the Senegalese National Malaria Control Programme (NMCP) scaled up the use of malaria RDTs at all health facilities in Senegal for case management [2]

\* Correspondence: [svolkman@hsph.harvard.edu](mailto:svolkman@hsph.harvard.edu)

<sup>2</sup>Broad Institute: The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

<sup>3</sup>Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA

Full list of author information is available at the end of the article

with the ultimate goal being to treat all (and only) true malaria infections, as misdiagnosis of malaria can contribute to drug resistance [3]. The number of RDTs currently used in Senegal is between 900,000 and 1,200,000 annually [4]. RDTs are ideal for rural conditions in developing countries where well-equipped health facilities and expert microscopists are lacking, as they give rapid results, require no electrical equipment or temperature-sensitive reagents, and can be performed without extensive training. The technical performance of these tests has been described previously [5], and although they are sensitive and specific to detect most cases of malaria at parasitaemia greater than 200 parasites/ $\mu$ l, product performance can vary widely at low parasitaemia [6], and these tools may miss individuals harbouring a very low parasitaemia.

RDTs differ in the antigens they detect. Some are based on detection of the *P. falciparum* specific histidine-rich protein-2 (PfHRP2), or *P. falciparum* specific lactate dehydrogenase (pLDH), while others recognize antigens common to *P. falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, and *Plasmodium malariae* (pan-species pLDH and aldolase) [7]. *Plasmodium* aldolase and lactate dehydrogenase (pLDH) are highly conserved RDT targets, [8,9] in contrast to the *P. falciparum* PfHRP2 which is considered to be more variable [10,11]. In Senegal, HRP2-based RDTs are most commonly used in health facilities because of their previously demonstrated sensitivity and specificity and because *P. falciparum* represents the major *Plasmodium* species (more than 98%) in Senegal [12], with *P. malariae* and *P. ovale* comprising less than 2% of malaria cases, and no *P. vivax* (0%) observed in 2013 [4].

The *pfhrp2* gene of the *P. falciparum* 3D7 is localized on chromosome 8 (PF3D7\_0831800, PlasmoDB Version 9.3) and contains two exons separated by an intron. The first exon encodes a signal peptide and the second exon contains a protein-export motif followed by specific histidine-rich repeats, which have been described previously [10,11]. PfHRP2 is synthesized throughout the asexual life cycle and in early sexual stages of *P. falciparum* [13]. The protein is characterized by the presence of a number of variable tandem repeats, each one named with a Type based on the motif being repeated [10,11].

Specific amino acid repeats AHHAAD (Type 7), AHH AHHAD (Type 2) have been described as possible epitopes targeted by the monoclonal antibodies used to detect HRP2 in some RDTs [14]. Many studies of the genetic diversity of this antigen show extensive variation within isolates of the same country and between isolates of different countries [10,11,15]. In Senegal, the genetic diversity of the gene target by the RDT has not yet been investigated in depth. The goal of this study is to determine the extent of genetic polymorphism in the *pfhrp2* gene in a Senegalese population in comparison with two

other diverse African populations: Mali and Uganda. Indeed, previous reports from Mali have indicated the loss of the HRP2 locus from some parasites, and this study investigates whether changes in this locus among Senegalese parasites may compromise the performance of the RDTs so widely used for malaria diagnosis.

To accomplish this goal, the *pfhrp2* gene was sequenced using patient samples from Senegal, and along with available sequencing data from Mali and Uganda (three endemic African countries with different malaria ecologies and epidemiology), was used to assess the nucleotide polymorphism as well as the distribution of specific PfHRP2 amino acid repeats.

## Methods

### Parasite isolates

Both the Ethics Committee of the Ministry of Health in Senegal and the Institutional Review Board of the Harvard School of Public Health approved this study. Parasite isolates from Senegal used in this study were collected from malaria febrile patients, after informed consent, during the malaria transmission season (August-December), from 2001 to 2010, at three health facilities: Pikine, Thies and Velingara. Pikine, a suburban area (approximately 15 km from Dakar) is classified as low transmission area (<5 cases per 1,000 individuals). Thies (70 km from Dakar), is a hypo-endemic urban site with an average of five malaria cases per 1,000 people. For Pikine and Thies, the entomological inoculation rate (EIR) is low less than five infectious bites per year. Velingara (in the south of Senegal, 565 km from Dakar) is a highly endemic area, with the level of transmission approximately 21 malaria cases per 1,000 people and the EIR around 100 to 300 infectious bites per year [16]. Overall, 54 Senegal isolates were used in this study (ten isolates from Pikine, 40 isolates from Thies and four isolates from Velingara).

Mali is a border country of Senegal. The samples used in this study are from Bandiagara in central eastern Mali, where malaria is endemic and the *P. falciparum* transmission is intense with seasonal peaks in July-October [1]. Uganda is located in East Africa and is highly endemic for malaria, which is the leading cause of morbidity and mortality [17]. However, samples were collected in Kampala, an urban centre in Uganda where malaria is meso-endemic, arising perennially with peaks during the two rainy seasons from August to December and from February to June [17].

### DNA extraction

*Plasmodium falciparum* isolates from Senegal were culture-adapted and deposited at the MR4 reagent repository. Genomic DNA was isolated, for sequencing, using Qiagen reagents according to the manufacturer instructions.

### DNA genotyping, sequencing and DNA translation

Samples were determined to be monogenomic and genetically distinct by a 24 single nucleotide polymorphism (SNP) molecular barcode [18]. Genomic DNA was sequenced using Illumina Hi-Seq machines, aligned to the 3D7 reference assembly, and variants were identified using previously described methods [19]. Using variant and invariant sites, the *pfhrp2* gene coding sequence (spliced CDS sequence) was reconstructed for each sample, inserting an 'N' to represent an undetermined base for sites that could not be resolved due to insufficient sequence coverage. The SNP calls for the Uganda and Mali isolates are downloadable from the Broad website [20].

Two isolates from Senegal SenT113.09 and SenV042.05 were PCR re-sequenced across the *pfhrp2* region to confirm a frameshift mutation identified with GATK software analysis of the sequence data (Additional file 1). Primers specific to *pfhrp2* amplifying the frameshift position (nucleotides position 571 and 574 of SenT113.09 and SenV042.05 respectively) were designed (Forward: **GCA CGCCGTTTTGCTCCGTACT** and Reverse: **GCACG CAATGTGTGGCGGCTTC**). PCR amplification was carried out in a final reaction volume of 30  $\mu$ l containing DNA template (100 ng), primers (50  $\mu$ M for each primer), dNTPs (10 mM for each dNTP), and 0.6  $\mu$ l Polymerase (Pfu Ultra II Fusion polymerase, Agilent Technologies, Inc). PCR amplification conditions were: denaturation at 95°C for 2 min, followed by 35 cycles of 95°C for 25 sec, 55°C for 25 sec and 60°C for 90 sec, with a final extension phase at 72°C for 10 min. Quality control of PCR reactions was performed by gel electrophoresis. PCR products were purified with Qiagen columns and sequenced in both the forward and reverse directions using the same PCR primers as were used for the amplification. DNA sequencing was performed by GENEWIZ [21] and amino-acid sequences were deduced from the nucleotide sequences obtained using TRANSEQ [22]. Nucleotide and amino acid sequences were aligned with the 3D7 reference sequence (PlasmoDB MAL7P1.231) using Bio Edit Sequence Alignment Editor Software (version 7.1.3.0). Polymorphic repeats were characterized as described [15].

### Nucleotide diversity

The genetic diversity of the *pfhrp2* gene was measured for the Senegal population ( $n = 54$ ) using the nucleotide diversity ( $\pi$ ), which quantifies the average number of nucleotide differences per site between DNA sequences among the sample population; and the  $\pi_{NS}$  and  $\pi_S$  (non-synonymous and synonymous polymorphism, respectively) values were calculated by dividing the average number of non-synonymous or synonymous pairwise differences by the number of non-synonymous or synonymous sites computed for all genes in the *P. falciparum* genome (Additional file 2) [23]. For Mali ( $n = 11$ ) and

Uganda ( $n = 9$ ) isolates, DnaSp [24] was used to predict nucleotide diversity ( $\pi$ ) due to a smaller than ideal sample size for estimating SNP diversity.

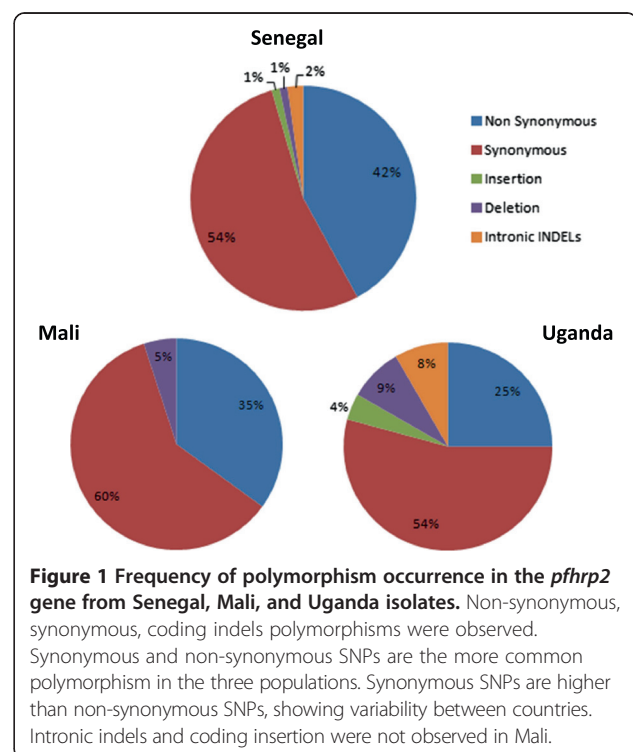
### Statistical analysis

Differences in the median number of each type of amino acid repeat were tested by Kruskal-Wallis for Senegal, Mali and Uganda *pfhrp2* sequences for each individual repeat ( $\alpha = 0.05$ ). The Spearman rank test was used to evaluate the correlation between pairs of repeats in a given population. For all statistical tests a significance threshold of  $\alpha = 0.05$  was used.

## Results

### Analysis of the nucleotide diversity in the *pfhrp2* gene

A total of 67 polymorphic sites among the 74 isolates were identified, with 29 of these polymorphisms described as “singletons”, appearing only in one isolate, ten as “doubletons” found in two samples, and 28 as “multiples” that appear in at least in three or more samples (Additional file 1). The full-length CDS of the *pfhrp2* gene ranges from 900 to 921 bp, and sequences were aligned using Clustal W (Additional file 3). Polymorphisms in the *pfhrp2* sequence from the three countries include: 35 synonymous and 20 non-synonymous SNP positions, as well as 11 insertion-deletions (indels) (Figure 1). For all defined populations, synonymous SNPs are more common than non-synonymous SNPs (Figure 1), and

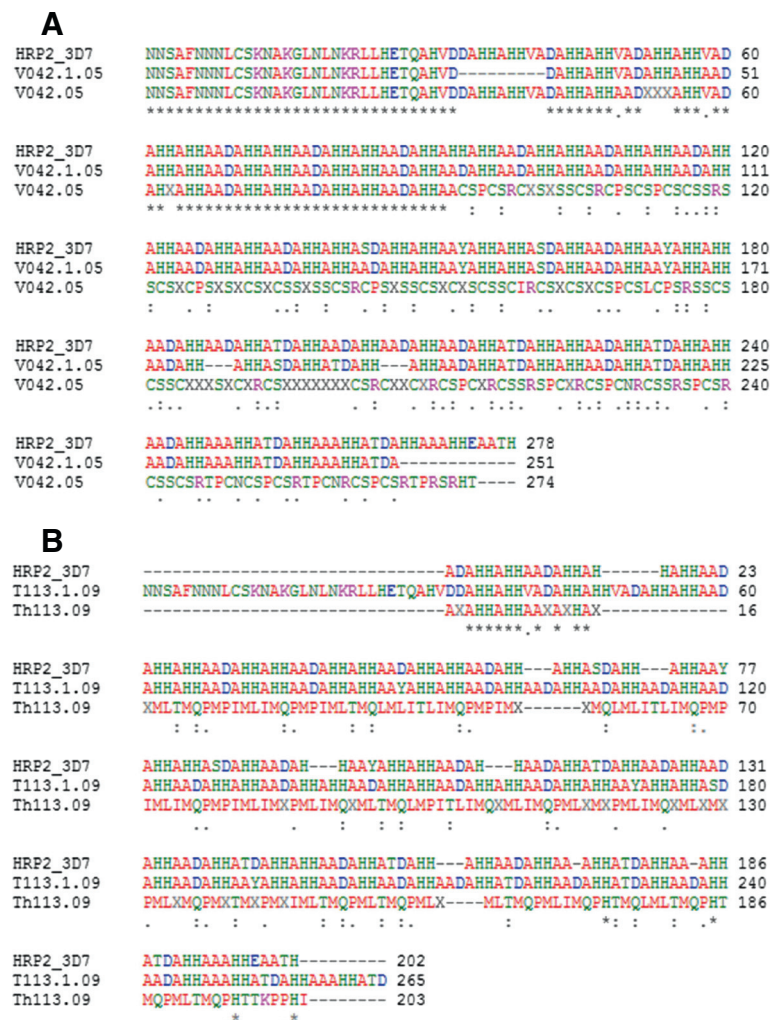


polymorphism varies within population, showing an intra- and inter-variability between isolates.

Of the 11 indels called by sequencing data, seven are predicted to result in insertions or deletions of entire codons (i.e. a multiple of three bp), two are intronic and two are suggestive of frameshifts within the coding sequence. The two frameshifting indels are both singletons and existed only in isolates SenT113.09 (insertion of two nucleotides) and SenV042.05 (deletion of two nucleotides). To confirm these frameshifts, the two isolates were PCR re-sequenced (with the new sequences being named SenT113.1.09 and SenV042.1.05). The amplification of genomic DNA by primers surrounding polymorphic amino acid (aa) positions 571 and 574, gave an amplicon of approximately 900 bp in length. The alignment of

the translated PfHRP2 protein (Figure 2) revealed an identity between the new sequences SenT113.1.09 and SenV042.1.05 (79.7 and 91.1%, respectively) with the 3D7 reference, and the lack of any frameshift, suggesting that errors in Illumina sequencing or GATK's Unified Genotyper produced a false call in these two cases. Consequently, the effects of the two frameshifting indels were ignored in further analysis.

The nucleotide diversity ( $\pi$ ) and  $\pi_{NS}$  diversity of the *pfhrp2* gene were then studied in depth among Senegalese isolates to investigate whether *pfhrp2* was remarkable in this population, since at a protein level a high  $\pi_{NS}$  may affect the PfHRP2 protein detection. Thus, *pfhrp2*  $\pi$  and  $\pi_{NS}$  were compared to other genes in the *P. falciparum* genome. As a point of reference, housekeeping



**Figure 2 PCR re-sequencing.** To confirm the frameshifting polymorphisms in isolates SenTh113.09 and SenV042.05, PCR amplification and re-sequencing of the region surrounding the polymorphic positions corresponding to aa571 and aa574 in the coding sequence were performed. **A:** Alignment of the re-sequenced SenV042.1.05 PfHRP2 sequences with 3D7 isolate. The two sequences share 91.1% identity. **B:** Alignment of the re-sequenced SenTh113.09 PfHRP2 sequence with 3D7 isolate showed 79.7% identity. These results suggest that errors in Illumina sequencing or GATK's Unified Genotyper produced a false call in these two cases.



genes such as 18S RNA (PF3D7\_1371000) have zero nucleotide diversity  $\pi = 0.000$  and  $\pi_{NS} = 0.000$  (Figure 3). The nucleotide diversity for *pfhrp2* is  $\pi = 0.00559$ , falling near highly polymorphic genes such as merozoite surface proteins (*msp6*, *msp1*,  $\pi > 0.007$ ) (Figure 3A). The distribution of  $\pi_{NS}$  varies widely in frequency among *P. falciparum* genes (Figure 3B, Additional file 2) and 89% of *P. falciparum* genes have a lower  $\pi_{NS}$  than *pfhrp2*. The  $\pi_{NS} = 0.0141$  ( $\pi_S = 0.0291627$ ) for *pfhrp2* is similar to several genes encoding SURFIN antigens (PF3D7\_0800700) involved in cyto-adherence ( $\pi_{NS} = 0.014065$ ) and a *glycophorin binding protein* family (PF3D7\_1401000) ( $\pi_{NS} = 0.01574$ ).

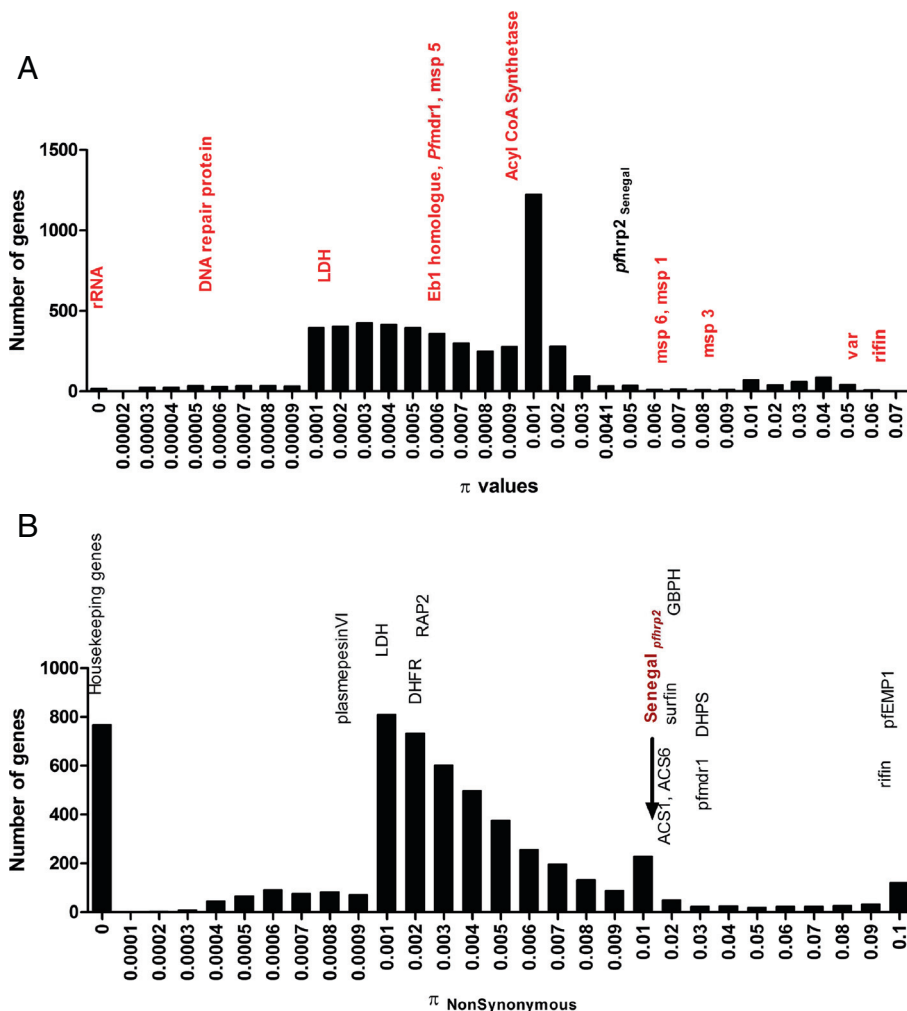
For Mali (n = 11) and Uganda (n = 9) isolates, the predicted nucleotide diversity for *pfhrp2* using DnaSp, revealed  $\pi = 5.76 \times 10^{-3}$  and  $\pi = 1.065 \times 10^{-2}$  for Mali

and Uganda, respectively, which were both greater than Senegal *pfhrp2* ( $\pi = 0.00559$ ).

#### Analysis of the PfHRP2 amino acid repeat polymorphisms

The length of the PfHRP2 for the reference strain 3D7 is 305 amino acids. Amino acid sequences were deduced from the nucleotide sequences and aligned using BioEdit (Clustal W). Variations in the length of the PfHRP2 sequences were observed. These fragments ranged in size from 299–307 aa (Additional file 4), with an average length of 303 aas.

The PfHRP2 protein is composed of a varying number of amino acids repeats [15]. The motif VLSAAVFAS VLLLDNNSAFNNL (termed Type 0), is conserved in all Senegalese, Malian and Ugandan PfHRP2 sequences (Table 1). In this study, 13 of the 20 different



**Figure 3 Nucleotide diversity ( $\pi$ ) and non-synonymous polymorphism ( $\pi_{NS}$ ).** **A:** Comparison of the nucleotide ( $\pi$ ) of *pfhrp2* in the Senegalese population, to other *P. falciparum* genes. All  $\pi$  values were rank in the *P. falciparum* genome  $\pi = 0.0000$  to  $\pi = 0.07$ . The diversity  $\pi_{pfhrp2} = 0.00559$  for Senegal falls within regions containing polymorphic genes such as *msp1* or *msp6*. **B:** The non-synonymous polymorphism ( $\pi_{NS} = 0.014111$ ) for *pfhrp2* gene showed that 89% of genes have a lower  $\pi_{NS}$  than *pfhrp2*.

**Table 1 The frequency of amino-acid repeats of PfHRP2**

Types	Repeat found in the HRP2 sequence	Repeat frequency in Senegal isolates N = 54	Repeat frequency in Mali isolates N = 11	Repeat frequency in Uganda isolates N = 9
0	VLSAAVFASVLLLDNNSAFNNNL	<b>54 (100%)</b>	<b>11 (100%)</b>	<b>9 (100%)</b>
1	AHHAHHVAD	52 (96.29%)	<b>11 (100%)</b>	<b>9 (100%)</b>
2	AHHAHHAAD	51 (94.4%)	<b>11 (100%)</b>	<b>9 (100%)</b>
3	AHHAHHAAY	29 (53.7%)	5 (45.45%)	7 (78%)
4	AHH	52 (96.29%)	<b>11 (100%)</b>	<b>9 (100%)</b>
5	AHHAHHASD	20 (37.07%)	2 (18.18%)	0%
6	AHHATD	45 (83.33%)	10 (90.9%)	<b>9 (100%)</b>
7	AHHAAD	51 (94.44%)	<b>11 (100%)</b>	<b>9 (100%)</b>
8	AHHAAY	37 (68.51%)	8 (72.72%)	<b>9 (100%)</b>
9	AAY	3 (5.55%)	0%	0%
10	AHHAACHATD	33 (61.1%)	6 (54.54%)	<b>9 (100%)</b>
11	AHN	0%	0%	0%
12	AHHAACHHEAATH	40 (74.07%)	7 (63.63%)	<b>9 (100%)</b>
13	AHHASD	1 (1.85%)	0%	1 (11%)
14	AHHAHHATD	0%	0%	0%
19	AHHAA	43 (79.62%)	<b>11 (100%)</b>	8 (88.88%)
20	SHHD	0%	0%	0%
21	AHHAHHATY	0%	0%	0%
22	AHHAHHAGD	0%	0%	0%
23	ARHAAD	0%	0%	0%
24	AHHTHHAAD	0%	0%	0%

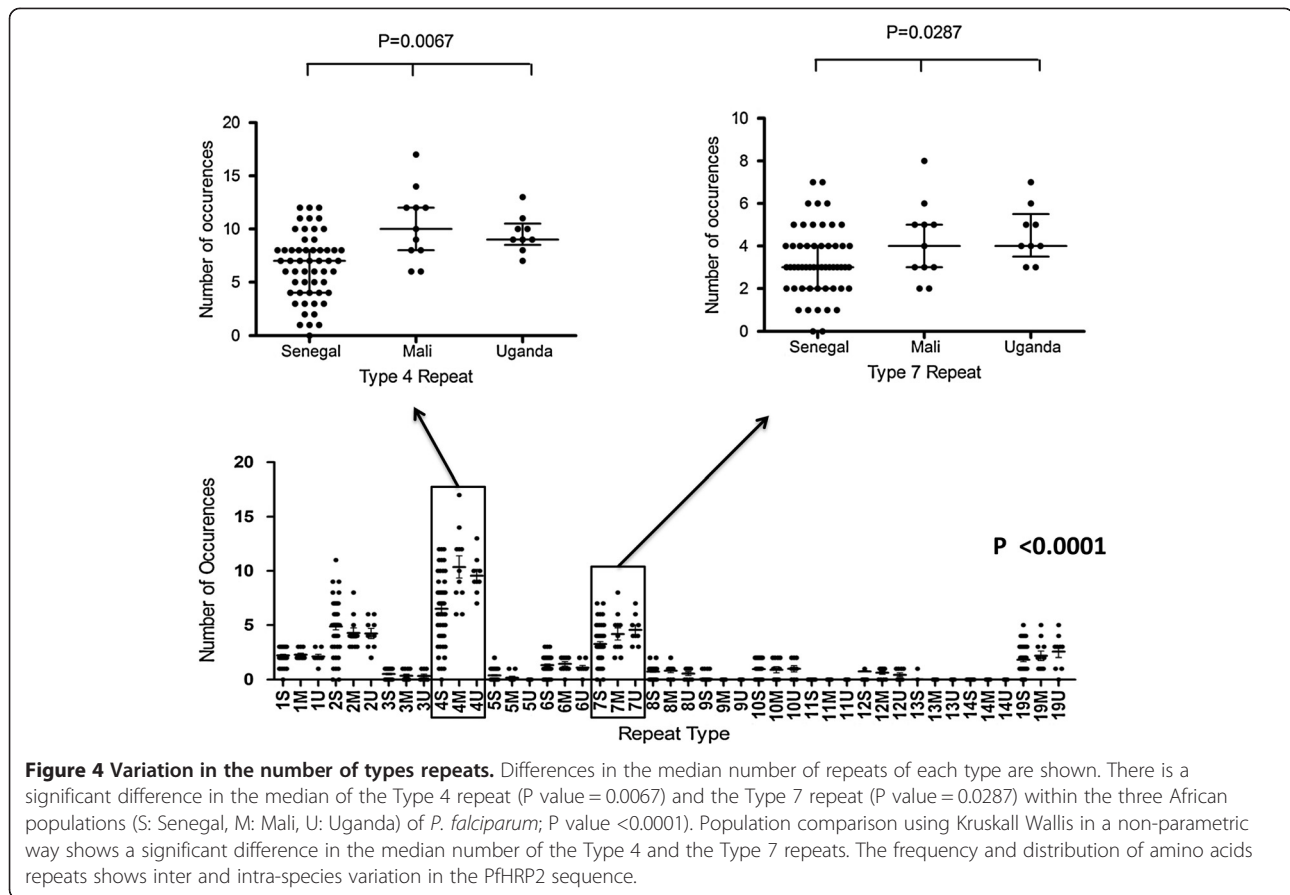
Repeat Types found in Senegal, Mali, and Uganda isolates. Repeat Types found in 100% of isolates are in bold, 0% and other percentages are in plain text.

amino acid repeats previously reported were detected (Table 1) [10,15]. PfHRP2 sequences begin with the Type 1 (AHHAHHVAD) motif and end with the Type 12 (AHHAACHHEAATH) motif (Additional file 4). The frequency distributions of PfHRP2 type repeats within the three African populations show an overall significant difference in the median number of occurrences of the different repeat types for each of the three populations (Kruskall-Wallis,  $p < 0.0001$ ) (Figure 4). The frequencies of Type 4 and Type 7 repeats, repeats thought to be the targets of RDTs, were significantly different between the three populations (Figure 4) (Kruskall-Wallis,  $p = 0.0067$ ,  $p = 0.0287$ , respectively) suggesting that the frequency and distribution of amino acid repeats could possibly reflect population level variation in the PfHRP2 protein.

The overall frequency of the Type 2, Type 4 and Type 7 repeats were compared to other regions of PfHRP2. The Type 2 (range 0–11), Type 4 (range 0–17) and Type 7 (range 0–8) were among the more variable repeats that also had the highest frequencies in this study for Senegal, Mali and Uganda (Figure 3 and Additional file 5). The other repeat types showed lower frequencies (range 0–3). Similar results were reported in other studies [11,25] where the Type 2 (range 2–14) and Type 7 (range 0–4) were the most frequent types observed compared to other repeat types [25].

Senegal isolates lack Type 11 and Type 14 repeats, as do Mali and Uganda isolates. However, 5.5% of Senegal isolates have the Type 9 repeat, which is absent from Mali and Uganda samples. The Type 13 was not observed in Mali isolates and one isolate from both Senegal and Uganda contained this motif (Table 1). Similarly, while the Type 5 repeat is absent in Uganda isolates, it is present in 37.07% of Senegal samples and in 18.18% of Mali samples (Table 1, Figure 3).

In addition to the frequency of the repeats, a difference in the organization of the repeats in the PfHRP2 sequences was observed. Some Types are always found in the same position in the protein: Type 3 (AHHAHHAAY) at aa168-176; Type 5 (AHHAHHASD) at aa177-185; Type 8 (AHHAAY) at aa192-197; Type 9 (AAY) at aa192-197; and, Type 12 (AHHAACHHEAATH) at aa289-301. Others motifs such as Type 1 (AHHAHHVAD), Type 2 (AHHAHHAAD), Type 4 (AHH) and Type 10 (AHHAACHHATD) can be present in clusters of two, three or five copies and are distributed throughout the protein (Additional file 4). When present, Type 1 and Type 10 motifs are found with at least two copies that are always repeated in tandem, at aa57-83 and aa267-aa288, respectively. Interestingly, the extent of Type 2 tandem repeats varies between isolates, and can be observed as a cluster of tandem repeats, or smaller clusters separated by other



repeats. Type 4 is mostly present in two tandem repeats usually at position aa75-80, but is also found at other positions in the protein. The Type 6 repeat (AHHATD) most often occurs in three positions (aa213-218; aa237-242; aa252-257) while Type 7 (AHHAAD) and Type 19 (AHHAA) are widely dispersed in PfHRP2.

In addition to variation between populations in repeat structure and frequency, there was a significant correlation between repeat Type 1-Type 2 and Type 3-Type 5 ( $r_s = 0.6158$   $P < 0.0001$ ,  $r_s = 0.5151$   $P < 0.0001$ , respectively; Bonferroni correction  $P < 1.84 \times 10^{-4}$ ) in the Senegalese PfHRP2 sequences. For Mali PfHRP2 Types repeats, a significant correlation was observed between repeat Type 10-Type 19 ( $r_s = -0.8945$ ,  $P < 0.0001$ ). No significant correlation was observed in repeat types from among Ugandan sequences, possibly due to the small sample size in Uganda ( $n = 9$ ) which makes statistical interpretation more challenging (Table 2).

## Discussion

*Plasmodium falciparum* tests targeting the PfHRP2 antigen have demonstrated high detection rates [26], however test performance showed variability between lots and between similar products. Recently *P. falciparum*

genetic factors have been investigated as a possible cause of variability in RDT detection rates [15,25,27]. As genetic diversity in *P. falciparum* is associated with transmission intensity, such findings could significantly interfere with the ability to accurately detect all malaria cases, and therefore have great potential to impact control strategies and efforts to eliminate or eradicate the disease [28]. For this reason, it has become increasingly important to characterize the diversity of target antigens used in RDTs to predict the impact genetic variation has on diagnosis, since it has the potential to affect the sensitivity of these diagnostic tests. In this study, the genetic diversity ( $\pi$  and  $\pi_{NS}$ ) in the *pfhrp2* gene was described

**Table 2 Correlations found between repeat types**

Country	Type repeats correlated	rs	P value
Senegal	Type 1 vs Type 2	0.6158	< 0.0001
	Type 3 vs Type 5	0.5151	< 0.0001
Mali	Type 10 vs Type 19	- 0.8945	< 0.0001
Uganda	No statistical correlation found		

Spearman rank analysis was performed on all pairs of repeat motifs in a given population. The Bonferroni correction was applied. The total number of tests is  $n = 273$  for the three populations and for all statistical tests  $\alpha = 0.05$ , given a threshold  $P < 1.83 \times 10^{-4}$ . Significant correlations are listed. (vs: versus, rs: Spearman rank).

and polymorphisms in the PfHRP2 protein among parasites obtained from human infections in three African countries with distinct patterns of malaria transmission were analyzed, to provide additional information about the diversity of this locus to that described previously [10,11,25].

Here, the global nucleotide diversity  $\pi$  values calculated for *pfhrp2* in the Senegal, Mali and Uganda isolates are similar to highly polymorphic genes families exposed to strong immune pressure, such as *msp1* and *msp6*, since these genes are highly polymorphic [29]. In addition, the  $\pi_{NS}$  for the *pfhrp2* is high, similar to  $\pi_{NS}$  for genes encoding the polymorphic SURFIN antigens expressed on *P. falciparum* merozoite and infected erythrocytes [30]. The *pfhrp2* gene is located in the subtelomeric region of chromosome 8 within a dynamic area where recombination events frequently contribute to the high genetic variation observed among genes from these regions [31]. In addition, the *pfhrp2* gene reveals differential polymorphism by geography, with distinct population profiles of synonymous and non-synonymous SNP variants, with over half these variants being synonymous SNPs. This is particularly important because non-synonymous SNP may destabilize protein structure or interfere with either ligand binding or the formation of domain-domain interfaces [32] and these changes may affect the ability of monoclonal antibodies to detect PfHRP2, thus compromising the sensitivity of the RDT.

Insertions and deletions of nucleotides observed in this study provided similar level of polymorphisms as SNPs in *pfhrp2*. Previous studies have reported deletion of the *pfhrp2* gene associated with false-negative HRP2-based RDTs in Mali, suggesting that spontaneous *hrp2* deletions could be one mechanism of genetic variation responsible for RDT failures [27]. Such deletions have been reported before in South America, where 15/275 parasites (5.5%) were found with deletions of the *pfhrp2* gene, generating concerns over the reliability of PfHRP2-based RDTs in these regions [33]. Furthermore, genetic deletions were not restricted to the *pfhrp2* alone, but were found to extend into neighbouring genes such as Hsp70 (PF3D7F\_0831700) in the 3' region of *pfhrp2* and in 5' a pseudo-gene encoding a *Plasmodium* exported protein (PF3D7F\_0831900) [34]. Such data emphasize the importance of precisely mapping deletions in *pfhrp2* in parasites from Senegal and other geographic regions.

It is important to note that the genomic DNA samples used in this study were obtained from culture-adapted *P. falciparum* isolates (and additionally from single clone infections). Further studies should focus on DNA obtained directly from patients as many infections are polygenomic, and also culturing parasites can lead to artifactual differences resulting from the process of culture adaptation,

such as breakage of *P. falciparum* chromosomes occurring frequently in the *pfhrp2*, a subtelomeric gene [35]. However, the number of isolates was quite small for Mali and Uganda to allow more stringent comparison. Nevertheless, these findings suggest that the parasite differences based upon different geography, and consequently ecology and epidemiology, are important to consider.

At the amino acid level, the PfHRP2 protein showed significant differences, in the organization and the number of repeats, likely due to chromosome breakage and healing occurring randomly in a site-unspecific manner during culture *in vitro* [35]. The composition of the PfHRP2 sequence repeats in Senegalese isolates is similar to Mali and Uganda isolates where the Type 2 (AHHAHHAAD), Type 4 (AHH) and the Type 7 (AHH AAD) are the mostly common repeats (Table 1). The antigen PfHRP2 differs in the composition and the number of repeat types compared to parasites from global endemic countries previously reported [10,11,15]. Specifically, Senegalese PfHRP2 isolates differ from Cameroon isolates (presence of Type 11, and Type 14 repeats), Thailand (presence of Type 14 repeat), Philippines (presence of Type 11 repeat), and Papua New Guinea (presence of Type 14 repeat). In this analysis Type 11 and Type 14 repeats were not observed. Additionally, Type 20 (SHHDD), Type 21 (AHHAHHATY), Type 22 (AHHAH HAGD), Type 23 (ARHAAD), and Type 24 (AHHTH HAAD) repeats reported previously [10] were not found.

These results agree with previously published studies [10,25], which found a high frequency of Type 2 and Type 7 repeats. Ideally, for a universal and sensitive RDT, the target epitope should be present in all parasites, regardless of specific population, and should be present in multiple copies [14]. For RDT sensitivity, a major factor is the binding specificity and affinity of the monoclonal antibodies for the target epitopes, thus the expression level and number of these epitopes could potentially influence the binding affinity. From these data and that of other studies, the high frequency of Type 2, Type 4 and Type 7 repeats observed in *P. falciparum* isolates could contribute to the sensitivity of the RDTs in the African populations studied.

The study of the amino acid composition of the PfHRP2 protein showed that different factors can contribute to polymorphism at the protein level: both the organization of the repeats within sequences and the position of repeats in the antigen. In fact, the frequency or distribution of the target epitopes present in a particular parasite population may have an impact on the efficiency of antigen detection in this population and the sensitivity of RDTs [10,15]. These parameters should be taken into account in the choice and the design of monoclonal antibodies that could be used for PfHRP2 based RDTs.



## Conclusion

To counter increasing malaria drug resistance and to achieve the goal of malaria eradication, the need for reliable diagnostic tools is critical. This study provides information about the genetic diversity of the *pfhrp2* gene and the predicted polymorphism in the PfHRP2 antigen from sequences obtained from three African parasite populations Senegal, Mali and Uganda. For all three populations of parasites, the higher rate of synonymous SNPs in *pfhrp2* suggests evolutionary selection of this SNP. However, for Senegal isolates, the  $\pi$  and  $\pi_{NS}$  warrant deeper follow-up investigations as they may represent functional variation, which would potentially interfere with RDT detection. *Pfhrp2* is polymorphic at the nucleotide level, shown by both synonymous and non-synonymous SNPs; nucleotide ( $\pi$ ) and non-synonymous ( $\pi_{NS}$ ) diversity similar to highly polymorphic genes such as *msp1* and gene encoding SURFIN antigen. PfHRP2 showed extensive repeat length polymorphisms with an occurrence of the Type 2, Type 4 and Type 7 repeats. Similar patterns were observed in the number, organization and the type of amino acid repeats in the protein among the three African populations. However, the PfHRP2 showed variation within isolates of the same country and between isolates of different countries, addressing concern in the efficacy of the HRP2 based RDT test. Since the polymorphism described in the target antigen is a recognized concern, comprehensive studies of possible factors influencing the sensitivity of these diagnostic tools, including genetic deletions and protein expression level, are critical in predicting the efficacy of RDTs in global populations.

## Additional files

### Additional file 1: Polymorphic position in the *pfhrp2* gene.

The positions of the SNPs identified are showed. The Table contains information about the SNP types, the description of the gene where the SNP is located, the position of the SNPs, the possible alternative alleles, the minor allele frequency, and the call rate. A total of 67 polymorphisms were identified, one per row, among Senegal, Mali and Uganda isolates. These polymorphisms include SNPs and indels. The indels position is reported as the position before the insertion/deletion point, and the base before it is repeated. The first letter in the isolate names indicates the site of collection (P for Pikine; V for Velingara; and, T for Thies), and the last two digits in the isolate names indicate the year of collection (e.g., P05.02 was isolated in 2002).

### Additional file 2: Nucleotide diversity synonymous and non-synonymous polymorphism in Senegalese isolates.

The Table contains 5,495 *P. falciparum* genes from the set of Senegalese isolates. Each line corresponds to a description of the gene with data on localizations in the genome, the nucleotide diversity  $\pi$ , synonymous and non-synonymous polymorphism ( $\pi_S$  and  $\pi_{NS}$ ) and additionally the ratio ( $\pi_{NS}/\pi_S$ ).

**Additional file 3: Alignment of *pfhrp2* coding sequences from, Senegal, Mali and Uganda isolates.** All CDS were aligned Using BioEdit software, with Clustal W program.

**Additional file 4: Alignment of PfHRP2 sequence.** All protein sequences were aligned against the 3D7 reference, using BioEdit software and Clustal W program.

**Additional file 5: PfHRP2 Repeat Types.** This table contains the repeat Types found in the PfHRP2 sequence from Senegal, Mali and Uganda isolates. Each line represents an isolate, with the number of the different repeat Types observed within the sequences indicated.

## Abbreviations

NMCP: National Malaria Control Programme; RDTs: Rapid diagnostic tests; HRP-2: Histidine-rich proteins-2; PCR: Polymerase chain reaction; WHO: World Health Organization; aa: Amino acid; EIR: Entomological inoculation rate; rs: Spearman rank; ACT: Artemisinin-based combination therapy; ITNs: Insecticide-treated bed nets; IRS: Indoor residual spraying; SNP: Single nucleotide polymorphism; CDS: Coding sequence; MSP: Merozoite surface protein.

## Competing interests

The authors declare that they have no any competing interests.

## Authors' contributions

Data analysis, manuscript writing, and PCR re-sequencing were performed by ABD. DP provided the sequencing data for analysis and reviewed the manuscript. AKB gave constructive advice on analysis, assisted with writing and revision of the manuscript. OS, ASB, POG, and AA reviewed the manuscript. SKV performed culture-adaptation of the Senegal and Mali parasites and prepared DNA for sequencing from these parasite populations, and assisted with the writing and revision of the manuscript. DN led the Senegal field collection sites for samples collected. DP and SKV guided the analysis performed by ABD. All authors have read and approved the final version of manuscript.

## Acknowledgements

The authors thank Daniel E Neafsey for critical reading of the manuscript, Rachel Daniels for providing data from the molecular barcode assay. We thank Eli Moss for providing  $\pi$  and  $\pi_{NS}$  computation, and all the collection team in Senegal: Ngayo Sy, Lamine Ndiaye, Younouss Diedhiou, Amadou Moctar Mbaye. The authors thank the patients who participated in these studies and all health facilities staff. We thank the NIAID Genome Sequencing Center for Infectious Disease at the Broad Institute, for generating sequencing data. All parasites used in these studies have been deposited in MR4; and, all SNPs from these sequence data can be found on <http://www.plasmodb.org>

## Author details

<sup>1</sup>Laboratory of Bacteriology-Virologie, Hopital Aristide Le Dantec, Dakar, BP 7325, Senegal. <sup>2</sup>Broad Institute: The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. <sup>3</sup>Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA 02115, USA. <sup>4</sup>Laboratoire de Parasitologie and Mycologie, Faculty of Medicine and Pharmacy, Cheikh Anta Diop University, BP 5005 Dakar, Senegal. <sup>5</sup>School of Nursing and Health Sciences, Simmons College, Boston, MA 02115, USA.

Received: 28 November 2013 Accepted: 27 January 2014

Published: 29 January 2014

## References

1. Doumbia SO, Ndiaye D, Koita OA, Diakite M, Nwakanma D, Coulibaly M, Traore SF, Keating J, Milner DA Jr, Ndiaye JL, Sene PD, Ahouidi A, Dieye TN, Gaye O, Okebe J, Ceesay SJ, Ngwa A, Oriero EC, Konaté L, Sy N, Jawara M, Faye O, Kéita M, Cissé M, Sogoba N, Poudiougou B, Diawara S, Sangaré L, Coulibaly T, Seck I, et al: **Improving malaria control in West Africa: interruption of transmission as a paradigm shift.** *Acta Trop* 2012, **121**:175–183.
2. Ly AB, Tall A, Perry R, Baril L, Badiane A, Faye J, Rogier C, Toure A, Sokhna C, Trape JF, Michel R: **Use of HRP-2-based rapid diagnostic test for *Plasmodium falciparum* malaria: assessing accuracy and cost-effectiveness in the villages of Dielmo and Ndiop, Senegal.** *Malar J* 2010, **9**:153.

3. Amexo M, Tolhurst R, Barnish G, Bates I: **Malaria misdiagnosis: effects on the poor and vulnerable.** *Lancet* 2004, **364**:1896–1898.
4. Programme SNMC: **National Malaria Control Programme Report.** 2013.
5. Nkrumah B, Acquah SE, Ibrahim L, May J, Brattig N, Tannich E, Nguah SB, Adu-Sarkodie Y, Huenger F: **Comparative evaluation of two rapid field tests for malaria diagnosis: partec rapid malaria test(R) and binax Now(R) malaria rapid diagnostic test.** *BMC Infect Dis* 2011, **11**:143.
6. World Health Organization (WHO): **Malaria Rapid Diagnostic Test Performance: result of WHO product testing of malaria RDTs Round 3 (2010–2011).** 2011. <http://www.who.int/tdr/publications/documents/rdt3.pdf>.
7. Murray CK, Bell D, Gasser RA, Wongsrichanalai C: **Rapid diagnostic testing for malaria.** *Trop Med Int Health* 2003, **8**:876–883.
8. Shin HI, Kim JY, Lee WJ, Sohn Y, Lee SW, Kang YJ, Lee HW: **Polymorphism of the parasite lactate dehydrogenase gene from *Plasmodium vivax* Korean isolates.** *Malar J* 2013, **12**:166.
9. Kim JY, Kim HH, Shin HL, Sohn Y, Kim H, Lee SW, Lee WJ, Lee HW: **Genetic variation of aldolase from Korean isolates of *Plasmodium vivax* and its usefulness in serodiagnosis.** *Malar J* 2012, **11**:159.
10. Baker J, Ho MF, Pelecanos A, Gatton M, Chen N, Abdullah S, Albertini A, Ariei F, Barnwell J, Bell D, Cunningham J, Djalle D, Echeverry DF, Gamboa D, Hii J, Kyaw MP, Luchavez J, Membri C, Menard D, Murillo C, Nhem S, Ogutu B, Onyop P, Oyibo W, Wang SQ, McCarthy J, Cheng Q: **Global sequence variation in the histidine-rich proteins 2 and 3 of *Plasmodium falciparum*: implications for the performance of malaria rapid diagnostic tests.** *Malar J* 2010, **9**:129.
11. Mariette N, Barnadas C, Bouchier C, Tichit M, Menard D: **Country-wide assessment of the genetic polymorphism in *Plasmodium falciparum* and *Plasmodium vivax* antigens detected with rapid diagnostic tests for malaria.** *Malar J* 2008, **7**:219.
12. WHO: **Global Malaria Report Country Profile: Senegal.** Geneva: World Health Organization; 2013.
13. Desakorn V, Dondorp AM, Silamut K, Pongtavornpinyo W, Sahassananda D, Chotivanich K, Pitisuttithum P, Smithyman AM, Day NP, White NJ: **Stage-dependent production and release of histidine-rich protein 2 by *Plasmodium falciparum*.** *Trans R Soc Trop Med Hyg* 2005, **99**:517–524.
14. Lee N, Gatton ML, Pelecanos A, Bubbs M, Gonzalez I, Bell D, Cheng Q, McCarthy JS: **Identification of optimal epitopes for *Plasmodium falciparum* rapid diagnostic tests that target histidine-rich proteins 2 and 3.** *J Clin Microbiol* 2012, **50**:1397–1405.
15. Baker J, McCarthy J, Gatton M, Kyle DE, Belizario V, Luchavez J, Bell D, Cheng Q: **Genetic diversity of *Plasmodium falciparum* histidine-rich protein 2 (PfHRP2) and its effect on the performance of PfHRP2-based rapid diagnostic tests.** *J Infect Dis* 2005, **192**:870–877.
16. PNLP: **Plan Strategique National 2011–2015.** Dakar, Senegal: Programme National de Lutte contre le Paludisme; 2010.
17. Talisuna A, Adibaku S, Dorsey G, Kamya MR, Rosenthal PJ: **Malaria in Uganda: challenges to control on the long road to elimination. II. The path forward.** *Acta Trop* 2012, **121**:196–201.
18. Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey DE, Park DJ, Rosen D, Angelino E, Sabeti PC, Wirth DF, Wiegand RC: **A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking.** *Malar J* 2008, **7**:223.
19. Park DJ, Lukens AK, Neafsey DE, Schaffner SF, Chang HH, Valim C, Ribacke U, Van Tynne D, Galinsky K, Galligan M, Becker JS, Ndiaye D, Mboup S, Wiegand RC, Hartl DL, Sabeti PC, Wirth DF, Volkman SK: **Sequence-based association and selection scans identify drug resistance loci in the *Plasmodium falciparum* malaria parasite.** *Proc Natl Acad Sci USA* 2012, **109**:13052–13057.
20. **Genome project.** [website:https://olive.broadinstitute.org/projects/plasmodium\\_100\\_genomes/downloads/plasmodium\\_snp.tar.gz](https://olive.broadinstitute.org/projects/plasmodium_100_genomes/downloads/plasmodium_snp.tar.gz) Bl.
21. **Genewiz.** <http://www.genewiz.com>.
22. Etzold T, Argos P: **SRS—an indexing and retrieval tool for flat file data libraries.** *Comput Appl Biosci* 1993, **9**:49–57.
23. Chang HH, Park DJ, Galinsky KJ, Schaffner SF, Ndiaye D, Ndir O, Mboup S, Wiegand RC, Volkman SK, Sabeti PC, Wirth DF, Neafsey DE, Hartl DL: **Genomic sequencing of *Plasmodium falciparum* malaria parasites from Senegal reveals the demographic history of the population.** *Mol Biol Evol* 2012, **29**:3427–3439.
24. Librado P, Rozas J: **DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.** *Bioinformatics* 2009, **25**:1451–1452.
25. Kumar N, Singh JP, Pande V, Mishra N, Srivastava B, Kapoor R, Valecha N, Anvikar AR: **Genetic variation in histidine rich proteins among Indian *Plasmodium falciparum* population: possible cause of variable sensitivity of malaria rapid diagnostic tests.** *Malar J* 2012, **11**:298.
26. WHO: **Malaria rapid diagnostic test performance, round 4.** Geneva: World Health Organization; 2012.
27. Koita OA, Doumbo OK, Ouattara A, Tall LK, Konare A, Diakite M, Diallo M, Sagara I, Masinde GL, Doumbo SN, Dolo A, Tounkara A, Traoré I, Krogstad DJ: **False-negative rapid diagnostic tests for malaria and deletion of the histidine-rich repeat region of the *hrp2* gene.** *Am J Trop Med Hyg* 2012, **86**:194–198.
28. Babiker HA, Walliker D: **Current views on the population structure of *Plasmodium falciparum*: Implications for control.** *Parasitol Today* 1997, **13**:262–267.
29. Pearce JA, Triglia T, Hodder AN, Jackson DC, Cowman AF, Anders RF: ***Plasmodium falciparum* merozoite surface protein 6 is a dimorphic antigen.** *Infect Immun* 2004, **72**:2321–2328.
30. Winter G, Kawai S, Haeggstrom M, Kaneko O, von Euler A, Kawazu S, Palm D, Fernandez V, Wahlgren M: **SURFIN is a polymorphic antigen expressed on *Plasmodium falciparum* merozoites and infected erythrocytes.** *J Exp Med* 2005, **201**:1853–1863.
31. Figueiredo LM, Freitas-Junior LH, Bottius E, Olivo-Marin JC, Scherf A: **A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation.** *Embo J* 2002, **21**:815–824.
32. Burke DF, Worth CL, Priego EM, Cheng T, Smink LJ, Todd JA, Blundell TL: **Genome bioinformatic analysis of nonsynonymous SNPs.** *BMC Bioinforma* 2007, **8**:301.
33. Gamboa D, Ho MF, Bendezu J, Torres K, Chiodini PL, Barnwell JW, Incardona S, Perkins M, Bell D, McCarthy J, Cheng Q: **A large proportion of *P. falciparum* isolates in the Amazon region of Peru lack *pfhpr2* and *pfhpr3*: implications for malaria rapid diagnostic tests.** *PLoS One* 2010, **5**:0008091.
34. Akinyi S, Hayden T, Gamboa D, Torres K, Bendezu J, Abdallah JF, Griffing SM, Quezada WM, Arrospe N, De Oliveira AM, Lucas C, Magill AJ, Bacon DJ, Barnwell JW, Udhayakumar V: **Multiple genetic origins of histidine-rich protein 2 gene deletion in *Plasmodium falciparum* parasites from Peru.** *Sci Rep* 2013, **3**:2797.
35. Scherf A, Mattei D: **Cloning and characterization of chromosome breakpoints of *Plasmodium falciparum*: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes.** *Nucleic Acids Res* 1992, **20**:1491–1496.

doi:10.1186/1475-2875-13-34

**Cite this article as:** Deme et al.: Analysis of *pfhpr2* genetic diversity in Senegal and implications for use of rapid diagnostic tests. *Malaria Journal* 2014 **13**:34.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

