



Queensland University of Technology
Brisbane Australia

This may be the author's version of a work that was submitted/accepted for publication in the following source:

Baker, Joanne, Ho, Mei-Fong, Pelecanos, Anita, [Gatton, Michelle](#), Chen, Nanhua, Abdullah, Salim, Albertini, Audrey, Arie, Frederic, Barnwell, John, Bell, David, Cunningham, Jane, Djalle, Djibrine, Echeverry, Diego, Gamboa, Dionicia, Hii, Jeffery, Kyaw, Myat Phone, Luchavez, Jennifer, Membi, Christopher, Menard, Didier, Murillo, Claribel, Nhem, Sina, Ogutu, Bernhards, Onyor, Pamela, Oyibo, Wellington, Wang, Shan Qing, McCarthy, James, & Cheng, Qin
(2010)

Global sequence variation in the histidine-rich proteins 2 and 3 of *Plasmodium falciparum*: implications for the performance of malaria rapid diagnostic tests.

Malaria Journal, 9, Article number: 1291-12.

This file was downloaded from: <https://eprints.qut.edu.au/217899/>

© Consult author(s) regarding copyright matters

This work is covered by copyright. Unless the document is being made available under a Creative Commons Licence, you must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a Creative Commons License (or other specified license) then refer to the Licence for details of permitted re-use. It is a condition of access that users recognise and abide by the legal requirements associated with these rights. If you believe that this work infringes copyright please provide details by email to qut.copyright@qut.edu.au

Notice: *Please note that this document may not be the Version of Record (i.e. published version) of the work. Author manuscript versions (as Submitted for peer review or as Accepted for publication after peer review) can be identified by an absence of publisher branding and/or typeset appearance. If there is any doubt, please refer to the published source.*

<https://doi.org/10.1186/1475-2875-9-129>

RESEARCH

Open Access

Global sequence variation in the histidine-rich proteins 2 and 3 of *Plasmodium falciparum*: implications for the performance of malaria rapid diagnostic tests

Joanne Baker^{1,21}, Mei-Fong Ho^{1,2}, Anita Pelecanos^{2,3}, Michelle Gatton^{2,3}, Nanhua Chen¹, Salim Abdullah⁴, Audrey Albertini⁵, Frederic Ariey⁶, John Barnwell⁷, David Bell^{5,8}, Jane Cunningham⁹, Djibrine Djalle¹⁰, Diego F Echeverry¹¹, Dionicia Gamboa^{12,13}, Jeffery Hii¹⁴, Myat Phone Kyaw¹⁵, Jennifer Luchavez¹⁶, Christopher Membi⁴, Didier Menard¹⁷, Claribel Murillo¹¹, Sina Nhem⁶, Bernhards Ogutu¹⁸, Pamela Onyior¹⁸, Wellington Oyibo¹⁹, Shan Qing Wang²⁰, James McCarthy² and Qin Cheng^{*1,3}

Abstract

Background: Accurate diagnosis is essential for prompt and appropriate treatment of malaria. While rapid diagnostic tests (RDTs) offer great potential to improve malaria diagnosis, the sensitivity of RDTs has been reported to be highly variable. One possible factor contributing to variable test performance is the diversity of parasite antigens. This is of particular concern for *Plasmodium falciparum* histidine-rich protein 2 (PfHRP2)-detecting RDTs since PfHRP2 has been reported to be highly variable in isolates of the Asia-Pacific region.

Methods: The *pfhrp2* exon 2 fragment from 458 isolates of *P. falciparum* collected from 38 countries was amplified and sequenced. For a subset of 80 isolates, the exon 2 fragment of histidine-rich protein 3 (*pfhrp3*) was also amplified and sequenced. DNA sequence and statistical analysis of the variation observed in these genes was conducted. The potential impact of the *pfhrp2* variation on RDT detection rates was examined by analysing the relationship between sequence characteristics of this gene and the results of the WHO product testing of malaria RDTs: Round 1 (2008), for 34 PfHRP2-detecting RDTs.

Results: Sequence analysis revealed extensive variations in the number and arrangement of various repeats encoded by the genes in parasite populations world-wide. However, no statistically robust correlation between gene structure and RDT detection rate for *P. falciparum* parasites at 200 parasites per microlitre was identified.

Conclusions: The results suggest that despite extreme sequence variation, diversity of PfHRP2 does not appear to be a major cause of RDT sensitivity variation.

Background

Malaria is one of the most important infectious diseases of humanity, and continues to cause significant mortality and morbidity worldwide. Early diagnosis is important for case management and treatment of the disease, and in guiding treatment for non-malaria fevers. Symptom-based clinical diagnosis is inaccurate, and contributes to

poor management of febrile illness, over-treatment of malaria, and may promote drug resistance to current anti-malarials [1].

Rapid diagnostic tests (RDTs) for malaria have the potential to improve case management and thereby reduce morbidity and mortality, especially in remote areas, facilitating the timely delivery of appropriate treatment. Indeed, many RDTs today can achieve excellent sensitivity and specificity for *Plasmodium falciparum* at a parasitaemia greater than 500 parasites per microlitre

* Correspondence: qin.cheng@defence.gov.au

¹ Department of Drug Resistance and Diagnostics, Australian Army Malaria Institute, Brisbane, Australia

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

(parasites/ μ L) [2]. At lower parasitaemia, however, variability in sensitivity is more common [3-7].

Today, over 150 malaria RDTs are commercially available, with most using a *P. falciparum* detecting component targeting *P. falciparum* histidine-rich protein 2 (PfHRP2). The gene encoding the protein, *pfhrp2*, is a single copy subtelomeric gene located on chromosome 7 encoding an amino acid sequence containing 34% histidine, 37% alanine and 10% aspartic acid [8-11]. PfHRP2 is characterized by multiple contiguous repeats of the sequences AHH and AHHAAD [8,9]. PfHRP2 is a 60-105 kD water-soluble protein specific to *P. falciparum*, synthesized and present throughout the asexual life cycle, identified as a surface-exposed protein in infected erythrocytes [8-10,12-18]. The protein is also found circulating in the peripheral blood of infected individuals [19]. These features make PfHRP2 a good target for diagnosis of *P. falciparum* infection.

Pfhrp3 encodes *P. falciparum* histidine-rich protein 3 (PfHRP3), also known as the small histidine-rich protein (SHARP), located near one end of chromosome 13 [17,20]. *Pfhrp3* shares many structural similarities with *pfhrp2*. Both genes have an interrupted structure and contain a signal peptide sequence in exon 1 followed by an intron. The intron is followed by the main coding region, exon 2. Exon 2 in both *pfhrp2* and *pfhrp3* encodes histidine-rich amino acid repeats beginning 75-90 nucleotides downstream from its start [18]. Although the histidine composition of PfHRP3 is slightly less than that of PfHRP2 (28% compared to 34%), both genes share many histidine and alanine rich repeats [10]. It has been suggested that, due to their similarity, both genes are related, derived from an ancestral duplication and interchromosomal divergence from a common ancestral gene, and may complement each other in function [8,10,17,18]. Antibodies against PfHRP2 cross-react with PfHRP3 [8,18]. Thus, PfHRP3 also contributes to the detection of *P. falciparum* infections in PfHRP2-detecting malaria RDTs.

As part of the World Health Organization (WHO) and Foundation for Innovative and New Diagnostics (FIND) Malaria RDT Quality Assurance Programme, the levels of diversity for antigens targeted by malaria RDTs have been systematically investigated. While parasite aldolase and pLDH appear to be highly conserved [21-23], *pfhrp2* was found to be highly variable. In the preliminary analysis of 74 isolates from mostly Southwest Pacific and Asian countries, a significant sequence variation in *pfhrp2* and *pfhrp3* was observed in isolates within the same country and between different countries [24]. This raised a serious concern that the sequence variation could result in significant variation in the presence and frequency of epitopes recognized by monoclonal antibodies (MABs) and hence impact on the RDT detection sensitivities for

different parasites. This concern was strengthened by a regression analysis based on 16 cultured parasite lines tested where the number of type 2 (AHHAHHAAD) and type 7 (AHHAAD) repeats in PfHRP2 were identified to be a contributing factor to the variable sensitivity reported at low level parasitaemia (below 250 parasites/ μ L) [24,25]. While this established extensive variation and a preliminary link between sequence diversity and RDT sensitivity within this sample set, it did not comprehensively cover diversity from all areas of global malaria transmission or testing sensitivity on a large set of samples with a wide range of RDT products.

The aims of the current study were to extend the diversity investigation to include isolates from African and South American countries, and to better understand the implication of global diversity in *pfhrp2* and *pfhrp3* on the performance of PfHRP2-detecting RDTs. In this paper, the global diversity of *pfhrp2* and *pfhrp3* was examined and the distribution of variants mapped. Furthermore, the results of the recently completed WHO product testing of malaria RDTs: Round 1 (2008) [26] were used to examine the effect of PfHRP2 sequence structure on RDT sensitivity.

Methods

Parasite lines and isolates

Field isolates of *P. falciparum* were obtained from patients and preserved on filter paper, and laboratory lines were cultured at the Australian Army Malaria Institute (AMI). Patient blood sample collection was coordinated by WHO, TDR and FIND and conducted by investigators and institutions within the WHO-FIND Malaria RDT Quality Assurance Programme. Protocols were approved by each collection country's Ethics Review Board and the WHO Research Ethics Review Committee. Fingerprick blood samples were obtained from consenting individuals, and three drops of patient blood were collected onto Whatman Filter paper (Grade 1, 9.0 cm, Whatman International Ltd, Maidstone, England) and air-dried for storage. All filter papers were sent to AMI for processing which was approved by the Australian Defence Human Research Ethics Committee (ADHREC 377/05). Confirmation of parasitaemia was obtained by microscopy for each patient and was also conducted for cultured lines. The 458 samples used to examine *pfhrp2* (including 74 reported in [24]) and their country of origins are detailed in Table 1. The subset of 80 samples that was used to examine *pfhrp3* is also listed in Table 1.

DNA isolation, PCR amplification and sequencing of *pfhrp2* and *pfhrp3*

Parasite DNA was extracted, from both patient blood preserved on filter paper and from packed red cells of cultured isolates, using the QIAgen QiaBlood kit (QIA-

Table 1: Country Origins (ISO code) and number *Plasmodium falciparum* isolates sequenced for *pfhrp2* and *pfhrp3* and the number of different *pfhrp2* sequence types.

Region/Country	<i>pfhrp2</i>		<i>pfhrp3</i>		Region/Country	<i>pfhrp2</i>		<i>pfhrp3</i>	
	<i>n</i>	No. types	<i>n</i>	No. types		<i>N</i>	No. types	<i>n</i>	No. types
<u>AFRICA</u>					<u>CENTRAL & SOUTH AMERICA</u>				
Benin (BJ)	1	1			Brazil (BR)	9	6	2	1
Burkina Faso (BF)	1	1			Colombia (CO)	12	8	1	1
Cameroon (CM)	2	2			Ecuador (EC)	1	1		
Central African Republic (CF)	13	13	1	1	Haiti (HT)	10	6	4	4
Gambia (GM)	1	1	1	1	Honduras (HN)	3	2		
Ghana (GH)	6	6			Peru (PE)	18	4	6	2
Guinea (GN)	1	1			Santa Lucia (LC)	1	1		
Kenya (KE)	30	30	6	6	Suriname (SR)	3	1	2	1
Liberia (LR)	3	3			<i>Sub total</i>	57		15	
Madagascar (MG)	17	17			<u>ASIA</u>				
Malawi (MW)	2	2			Cambodia (KH)	32	24	1	1
Nigeria (NG)	80	74	16	11	China (CN)	10	7	4	3
Niger (NE)	1	1			Indonesia (ID)	5	4		
Sierra Leone (SL)	2	2	1	1	Malaysia (MY)	2	2	1	1
Sudan (SD)	3	3			Myanmar (MM)	5	5		
Tanzania (TZ)	39	34	2	2	Philippines (PH)	45	27	7	4
Uganda (UG)	2	2			Thailand (TH)	7	6	4	4
Zambia (ZM)	2	2			Vietnam (VN)	5	4		
<i>Sub total</i>	206		27		<i>Sub total</i>	111		17	
<u>SOUTHWEST PACIFIC</u>									
Papua New Guinea (PG)	17	12	7	6					
Solomon Is. (SB)	35	17	13	9					
East Timor (TP)	24	12							
Vanuatu (VU)	8	4	1	0					
<i>Sub total</i>	84		21						
Global total	458		80						

GEN, Germany) following the manufacturer's instructions. Using the same primers and PCR conditions described previously [24], close to full length exon 2 of the *pfhrp2* (+65bp to stop codon) and *pfhrp3* (+56bp to 1bp before stop codon) genes were amplified and sequenced (ABI), and the nucleotide sequence translated to corresponding amino acids (aa).

Translation of DNA sequences and analysis of repeats

The same bar code system as previously reported was used to ascertain the repeat types present in *pfhrp2* and *pfhrp3* [24]. Six additional codes were described, and were designated types 19-24. These newly identified

repeats are detailed in Table 2. The sequences reported in this article have been deposited in the GenBank database (GenBank accession numbers for *pfhrp2* FJ871160 to FJ871401; GenBank accession numbers for *pfhrp3* GU194966 to GU195043). The sequences were aligned using Microsoft Excel software to examine similarity between isolates. Identical sequences were identified and assigned to the same grouping, while different sequences were listed separately.

Statistical analyses

Variability of repeat frequencies for *pfhrp2* and *pfhrp3* within and between isolates of different geographic origin

Table 2: The presence (+) and absence (-) of amino acid repeats in PfHRP2 and PfHRP3.

Code	Repeat	PfHRP2	PfHRP3
1	AHHAHHVAD	+	+
2	AHHAHHAAD	+	+
3	AHHAHHAAY	+	-
4	AHH	+	+
5	AHHAHHASD	+	-
6	AHHATD	+	-
7	AHHAAD	+	+
8	AHHAAY	+	-
9	AAY	+	-
10	AHHAAHHATD	+	-
11	AHN	+	-
12	AHHAAHHEAATH	+	-
13	AHHASD	+	-
14	AHHAHHATD	+	-
15	AHHAHHAAN	-	+
16	AHHAAN	-	+
17	AHHDG	-	+
18	AHHDD	-	+
<u>19</u>	<u>AHHAA</u>	+	-
<u>20</u>	<u>SHHDD</u>	+	+
<u>21</u>	<u>AHHAHHATY</u>	+	-
<u>22</u>	<u>AHHAHHAGD</u>	+	-
<u>23</u>	<u>ARHAAD</u>	+	-
<u>24</u>	<u>AHHTHHAAD</u>	+	-

Repeat types not previously reported are underlined.

was examined statistically, and observations made on differences in exon length, motif use, and frequency distribution. GraphPad PRISM[®] (GraphPad Software, La Jolla, USA) and Statistical Package for the Social Sciences[®] (SPSS Inc, Chicago, USA) were employed to analyse the sequence data for the following purposes:

1. Geographical differences in the sequence characteristics. Differences in the total number of amino acids and the number of repeats of each type, between countries (38 countries) and between regions (four different regions, grouped to Africa, Southwest Pacific, Asia and South America) were tested using the Kruskal-Wallis test. Tukey's test was used for post hoc multiple comparisons. To investigate whether a correlation existed between amino acid length of PfHRP2 and PfHRP3, we examined the relationship of total amino acid length for both genes. Further analysis of *pfhrp2* repeat types 2 and 7 and *pfhrp3* repeat types 15, 16, 17 and 18, was conducted

to look for individual country differences in the number of these repeats. The one sample t-test was used to determine whether there were significant differences between a given country's mean repeat number, compared to mean number for all countries combined.

2. Predictive value of sequences to RDT performance. The sequence characteristics for the 79 isolates used in the WHO Product Testing of Malaria RDTs (Round 1) were selected [26]. These geographically varying isolates had different *pfhrp2* sequences and repeat structures, classified as type A (≥ 100), B (50-99), or C (< 50), according to the frequency of their type 2 \times type 7 repeats [24]. Using only the results from the PfHRP2-detecting RDTs, the percent of the products testing positive to each parasite isolate at 200 parasites/ μ L was determined and linear regression used to investigate whether sequence structure and length influenced the detection rate.

Results

Variation in *pfhrp2*

The PCR product of *pfhrp2* exon 2 varied markedly in size between different parasite isolates. The amplified exon 2 size within the entire set of 458 isolates/laboratory lines examined ranged from 561bp to 918bp, with an average of 783 bp. The size variation was largely attributed to variation in numbers of 27- and 18- bp repeats. Three hundred and eighteen different *pfhrp2* sequences were identified consisting of combinations of 20 different amino acid repeats (Table 2). Two hundred and fifty nine isolates had a unique *pfhrp2* sequence, while the remaining 59 sequences were seen in > 1 parasite isolate. Of these 59 sequences, 36 (61%) were shared by isolates from the same country, while the remaining 23 (39%) were common to isolates from different countries (Table 3), of which 10 (44%) were shared within the region and 13 (56%), between different regions.

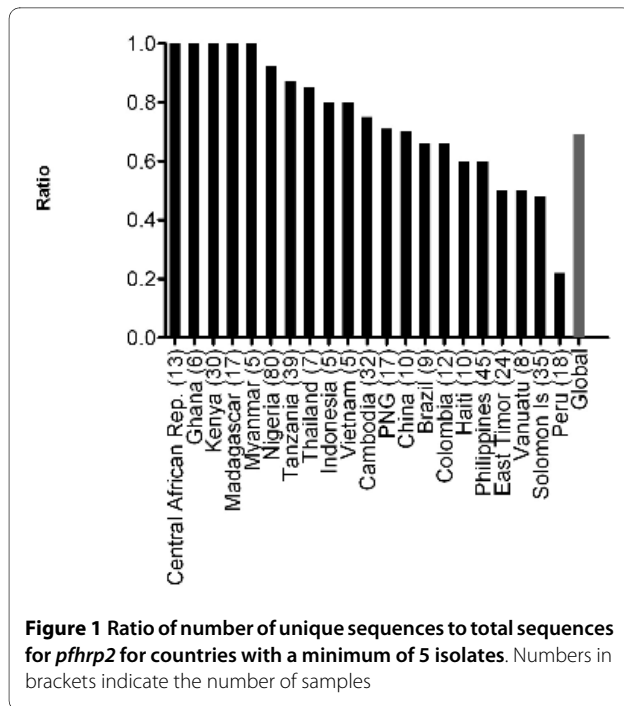
The number of different *pfhrp2* sequence types observed in each country is listed in Table 1. To demonstrate the level of sequence diversity the ratio of the number of *pfhrp2* sequence types to total number of sequences in a country was calculated for all countries. The overall ratio is 0.69 (318 unique sequence/458 isolates) with a unique *pfhrp2* type observed in every 1.44 parasite isolates examined. The ratio varied between different countries. Figure 1 shows the range of ratios for countries with more than 5 *pfhrp2* sequences analysed. Isolates from Peru ($n = 18$, ratio 0.22) showed the least *pfhrp2* sequence variability among all countries sampled with the 18 isolates sharing only 4 different sequence types. The Solomon Islands and The Philippines showed medium range sequence diversity. Of the Solomon Island isolates ($n = 35$, ratio 0.48), 28 samples combined to give

Table 3: The country (by ISO code) distribution of shared *pfhrp2* and *pfhrp3* sequences.

<i>pfhrp2</i>					
Seq. type	Shared in Country	Seq. type	Shared in Country	Seq. type	Shared in Country
a	SB, PH	u	KE, TP, TH	ao	NG
b	ID, NG	v	CN	ap	CO, HT
c	PG, VU	w	SB	aq	CO, PE
d	UG, KH, MY, PH	x	BR, PE,	ar	GH, NG
e	SB, PH	y	HN, LC	as	TH
f	SB, PH	z	NG	at	CO, KH
g	PH	aa	LR, PG	au	NG
h	SD, SB	ab	BR	av	NG, TZ, PE, KH
i	SB	ac	PH	aw	KH
j	PH	ad	HT	ax	KH
k	PH	ae	PH	ay	KH, ID, MM
l	PH	af	TP	az	KH
m	VN	ag	BR	ba	CO
n	VU	ah	CN	bb	NG
o	PH	ai	CN	bc	PH
p	TP	aj	BR, SR	bd	NG, TZ
q	SB	ak	NG	be	PG, ID
r	TP	al	NG	bf	NG, SB
s	SB	am	PG, SB	bg	SB
t	GH, PH, TH	an	PE		
<i>pfhrp3</i>					
bh	PE	bn	CN, SR	bs	NG, SB
bi	KE, TH, CO, HT	bo	NG, CF, MY	bt	PG, SB, HT
bj	NG	bp	PG, SB	bu	PG, KH
bk	NG	bq	KE, PH	bv	NG, VU
bl	PH	br	PG, TH	bw	SB
bm	NG, CN, PE, BR				

10 different sequence types, often identical to sequences found in isolates from other countries. The remaining seven Solomon Island isolates each had a distinct unique sequence. The Philippines isolates ($n = 45$, ratio 0.6) showed slightly greater diversity, with the 32 sequences combining to give 14 unique sequences, and the remaining 13 isolates having unique sequences. The greatest variability in *pfhrp2* sequence was found in the Central African Republic ($n = 13$), Ghana ($n = 6$), Kenya ($n = 30$), Madagascar ($n = 17$) and Myanmar ($n = 5$), where no sequence was identical to another either within the country (ratio = 1.0) or with any other country.

The length of the deduced PfHRP2 sequence encoded by the exon 2 varies from 187 to 306 aa (average 252.2 aa, Figure 2). When compared to the global mean, the mean of amino acid length was found to be significantly lower in Vanuatu (mean = 228.3, $P < 0.01$) and Cambodia (mean = 236.3, $P < 0.01$) and significantly higher in China (mean = 270.1, $P < 0.01$), Brazil (mean = 261.1, $P < 0.05$) and Haiti (mean = 262.2, $P < 0.05$, Figure 2). In Vietnam (mean = 274.4) the total aa length was greater than the global mean of 252.2, however due to small sample size it was not possible to gauge statistical significance. The lowest overall variation in the number of total amino



acids for *pfhrp2* for this sample set was seen in Indonesia (range 238-260 aa), the highest variation in Papua New Guinea (PNG) (194-306 aa), Nigeria (200-300 aa), and Madagascar (203-303 aa) (Additional file 1: Table S1).

Almost all (455/458, 99.3%) deduced amino acid sequences of PfHRP2 began with the type 1 repeat and universally (100%) concluded with the type 12 repeat. While most sequences had conserved start and end amino acid repeats, the sequence varied considerably in the organisation of the repeats and the number of each repeat in the central region of the sequence. A conserved motif of repeat types 2, 3, 5, 7, 8, 2, 7 was found in the central region of the sequence in 201 of the 458 isolates sequenced (44%). In a further 38% (174/458) of the sequences part of this motif was present. A total of 83 sequences (18%) did not contain the motif.

A total of 20 different types of amino acid repeats were identified in 458 PfHRP2 sequences (Table 2). Repeat type 2 and type 12 were observed in 100% of the isolates sequenced (Additional file 2: Table S2). Types 1, 6 and 7 were found in over 97% of isolates (Additional file 2: Table S2). The prevalence of repeat types 3, 5, 8 and 10 was found to differ between the geographic areas considered: 100% in parasites from some areas and between 70% and 100% in other areas. The type 4 repeat was present in less than 50% of isolates in all regions. The remaining 10 types were limited to a few isolates, and some were limited to a cluster of countries (Additional file 2: Table S2).

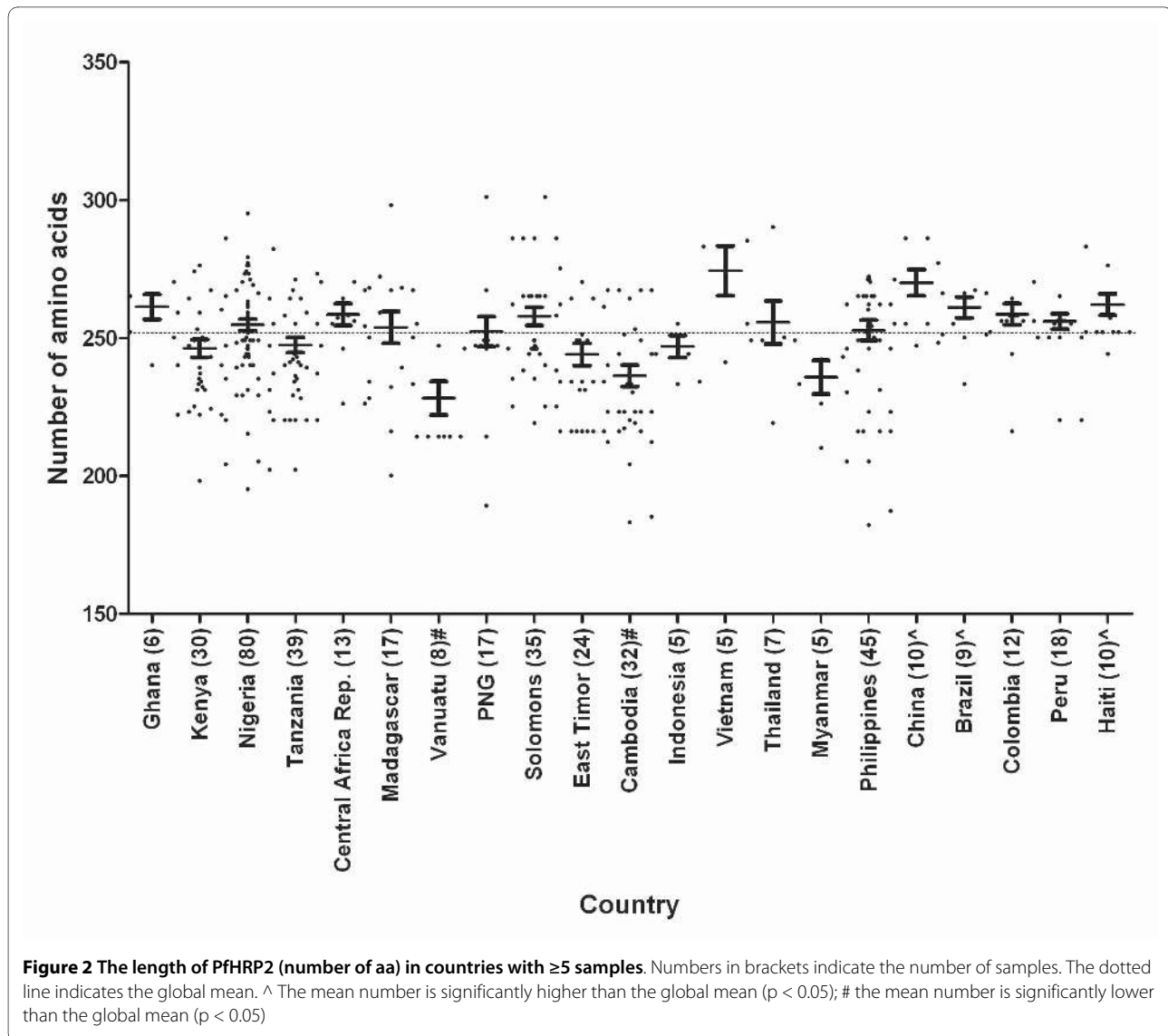
With the exception of type 12 the number of each repeat type was observed to vary between different parasite isolates from the same country and between coun-

tries. When grouped by region, significant differences were found between regions for the number of repeat types 1, 2, 6 and 7 ($P < 0.0001$), as well as 3, 4, 5, 10, 11, 13 and 14 ($P < 0.05$). Further analysis was performed to examine differences in the number of types 2 and 7 repeats between countries (countries with less than 5 samples were excluded). Figures 3 and 4 illustrate the variations across different countries in the number of repeat types 2 and 7, respectively. Countries with significantly higher or lower numbers of type 2 and 7 repeats than their global means are identified and indicated in Additional file 1: Table S1.

Variation in *pfhrp3*

The size of the *pfhrp3* gene fragment amplified by PCR ranged from 294 bp to 552 bp, largely due to variation in numbers of 18- and 15- bp repeats. 42 different sequence types were identified from 80 isolates originating from 20 countries. Twenty-six isolates had a unique *pfhrp3* sequence, while the remaining 54 isolates combined to give a total 16 sequence types. Of these 16 sequence types, 4 types were shared by isolates from the same country and the remaining 12 sequence types were common to isolates from different countries (Table 3). The ratio of unique over total number of sequences calculated for *pfhrp3* was 0.525 (42/80) indicating that there is a different sequence every 1.90 parasite isolates examined for *pfhrp3*. Overall the level of diversity in *pfhrp3*, as measured by the proportion of unique sequences was significantly lower than that observed for *pfhrp2* ($P < 0.01$). The lowest ratio was observed in Peru ($n = 6$, ratio 0.33), followed by Philippines ($n = 7$, ratio 0.57) and Solomon Islands ($n = 13$, ratio 0.69), a similar rank observed for *pfhrp2* for these countries. A higher ratio was observed for PNG ($n = 7$, ratio 0.85) and Kenya ($n = 6$, ratio 1.0).

A total of nine different amino acid repeats were identified from the PfHRP3 sequences (Table 2). All deduced amino acid sequences of the PfHRP3 began with the type 1 repeat, and ended with a type 4 repeat. All sequences contained each of the nine different amino acid repeat types with varying frequency (Additional file 3: Table S3). Two sequences also contained one copy of the type 2 repeat. Most PfHRP3 sequences had one non-repetitive region while some isolates from three countries showed two non-repetitive regions (Additional file 3: Table S3). The non-repetitive region located in the centre of the sequence was identical for all sequence types, except for those sequences with two non-repetitive regions which consistently ended in a type 20 repeat. Upstream of the non-repetitive region, the sequences showed variability in the number of type 16 repeats (AHHAAN). Downstream of the non-repetitive region varied in the number of type 17 and 18 repeats in isolates from different countries (Additional file 3: Table S3).



The deduced amino acid sequence of PfHRP3 observed ranging in length from 98 to 178 aa. The lowest overall variation in the total number of amino acids for *pfhrp3* (Additional file: 3 Table S3) for this sample set was seen in China (124-150 amino acids), the highest variation in Nigeria (114-170 amino acids), the Solomon Islands (98-184 amino acids) and PNG (104-170 amino acids). When compared with a mean of 144.5 amino acids for total global isolates, the mean amino acid length was found to be significantly higher among the samples from Peru (mean = 168.7, $P < 0.01$).

PfHRP2 and PfHRP3 amino acid length correlation

To examine whether a correlation existed between amino acid length of PfHRP2 and PfHRP3, we examined the relationship of total amino acid length for both genes. No

significant correlation existed between the length of *pfhrp2* and *pfhrp3* ($P > 0.05$).

pfhrp2 and *pfhrp3* deletions

Absence of the *pfhrp2* gene was only observed in two laboratory adapted lines: Dd2 and D10 [20,27], but not in any field isolates. The lack of *pfhrp3* gene was also only observed in a laboratory-adapted line of HB3 [17].

Relationship between PfHRP2 sequence diversity and the RDT detection rate

The proportion of the 34 PfHRP2-detecting RDTs which "detected" each of the 79 parasite isolates at a density of 200 parasites/ μ l was calculated. In the WHO product testing, "detected" was defined as returning 4/4 positive tests against an isolate [26]. Linear regression analysis

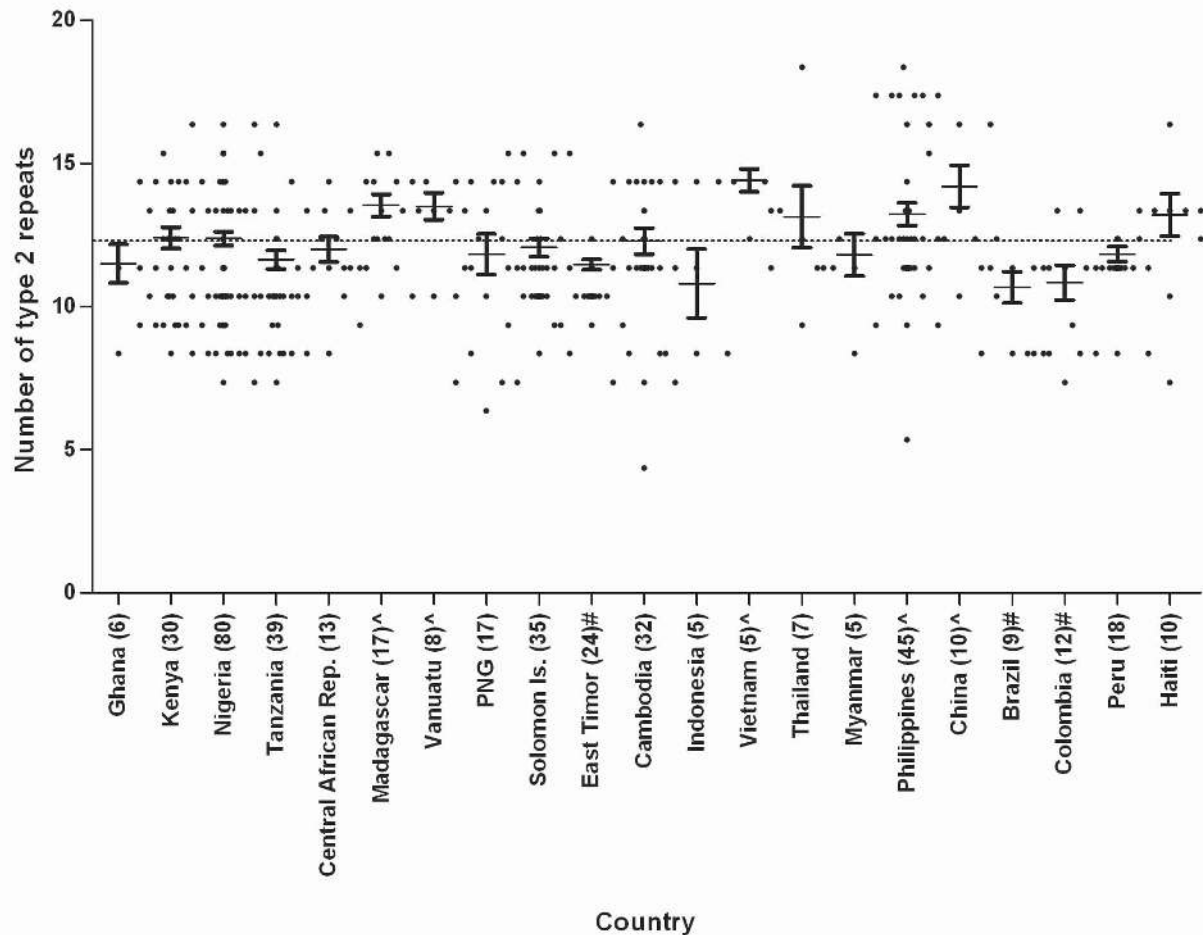


Figure 3 The number of type 2 repeat present in PfHRP2 in countries with ≥ 5 samples. Numbers in brackets indicate the number of samples. The dotted line indicates the global mean. ^ The mean number is significantly higher than the global mean ($p < 0.05$); # the mean number is significantly lower than the global mean ($p < 0.05$)

was used to investigate if there was a relationship between sequence structure of the PfHRP2 and percent detected. No significant relationship could be established between PfHRP2 structure types and percent detected ($P > 0.05$).

Discussion

PfHRP2 is the target antigen for many RDTs detecting *P. falciparum*. In this study, *pfhrp2* from 458 *P. falciparum* isolates collected from 38 countries, all major malaria endemic areas, were sequenced and evaluated. It was shown that this gene exhibits extensive diversity both within and between countries and regions. A similar protein, PfHRP3, also exhibited extensive diversity between parasite isolates examined. This comprehensive examination and analysis of the extent of the diversity and geographic variation in these genes provides important information for laboratory and field evaluation of malaria

RDTs. It also provides an indication of whether the genetic diversity in the antigen contributes to the variability in RDT sensitivity.

Both *pfhrp2* and *pfhrp3* are located in the subtelomeric regions of chromosomes. In general, genes located in telomeric and subtelomeric regions of *Plasmodium* have vast genetic diversity, and are highly susceptible to changes during recombination events [28-33]. Subtelomeric regions from different malaria species appear to have undergone rapid evolution, with significant sequence variation generated in the complex repeats in these regions [32,33]. Molecular mechanisms contributing to the generation of tandemly repeated regions, changes in the length of repeat blocks and other variation events include slipped strand mispairing followed by DNA replication or repair, unequal reciprocal combination and gene conversion [34-39]. The different organization and varying number of repeats observed in both

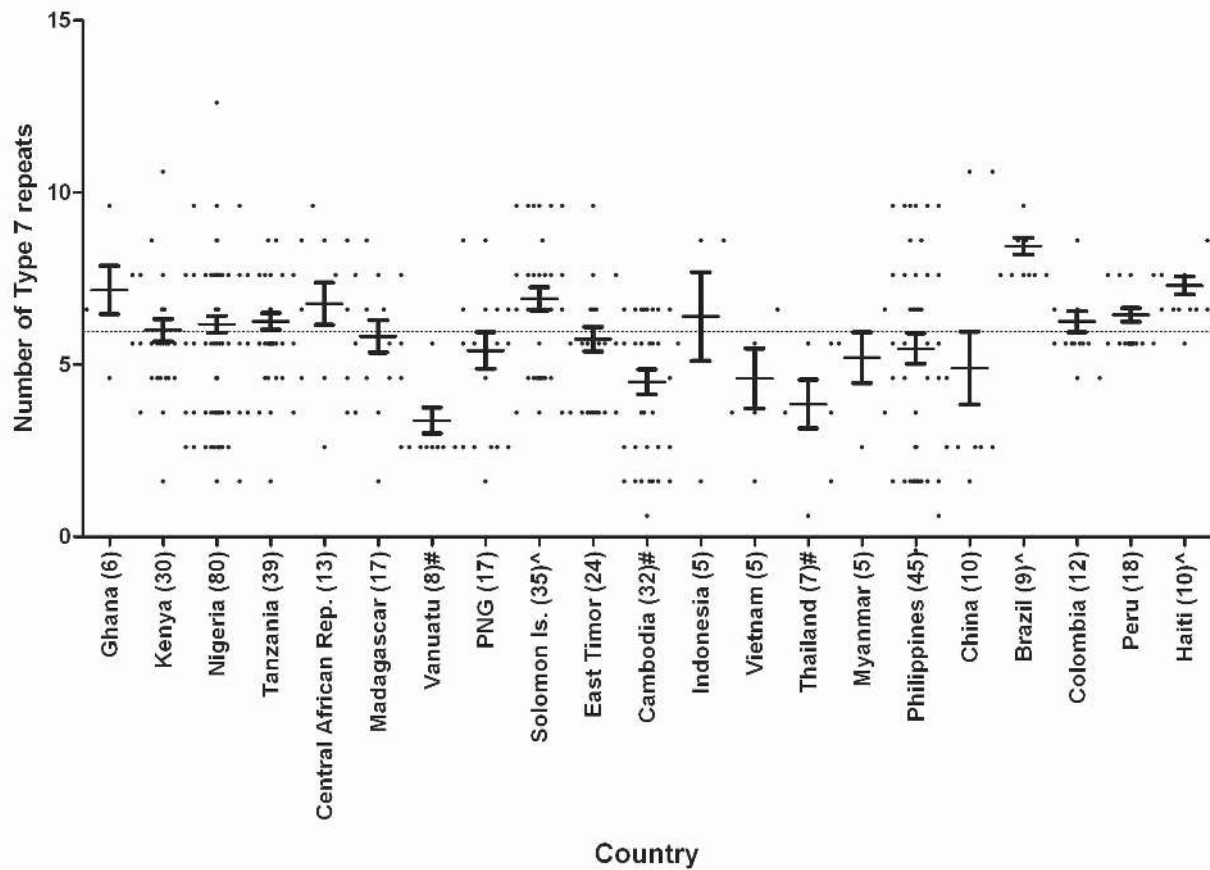


Figure 4 The number of Type 7 repeat present in PfHRP2 in countries with ≥ 5 samples. Numbers in brackets indicate the number of samples. The dotted line indicates the global mean. ^ The mean number is significantly higher than the global mean ($p < 0.05$); # the mean number is significantly lower than the global mean ($p < 0.05$)

pfhrp2 and *pfhrp3* are likely the result of frequent recombination of the chromosomes. Therefore, a correlation between malaria transmission intensity and *pfhrp2* diversity should be expected, as malaria infections in high transmission settings often involve co-infection of multiple strains which increase the probability of recombination during sexual reproduction in the mosquito vector. Indeed, this general trend was observed in this study with the ratio of different *pfhrp2* sequence types to total sequences being higher in countries with high transmission intensity such as in Africa, and lower in South American and Asian countries. However, this may also reflect differences in the geographical spread of collections within countries, which varied between sites.

The function of PfHRP2 still remains to be determined. Early theories suggested that PfHRP2 may be involved in detoxification of free haem by converting it to inactive haemozoin [40,41]. Other theories suggest that PfHRP2 may be involved in remodeling the infected erythrocyte

cytoskeleton [42] and in modulating host immune responses [43]. The extensive diversity in the *pfhrp2* sequence provides evidence that the function(s) of the molecule is not affected by the sequence diversity in exon 2 and that parasites with a particular *pfhrp2* sequence do not appear to have a significant fitness or survival advantage. The host immune responses may also contribute to the maintenance of diversity by selecting for immunologically different types.

A mechanism of chromosome breakage and healing is believed to contribute to generation of deletions in various parasite chromosomes [27]. Chromosome deletions at the *pfhrp2* locus of chromosome 7 and *pfhrp3* locus of chromosome 13 have been observed in laboratory-adapted lines [17,20,27], and in field isolates from the Amazon region of Peru [44]. This observation of gene deletions in field isolates suggests that PfHRP2-detecting RDTs may not be reliable for detecting *P. falciparum* infections in this region of South America, and also

raised a serious question as to whether parasites lacking PfHRP2 and PfHRP3 may exist in other parts of the world and what are possible fitness or survival advantages. In the 485 isolates, collected in over 30 malaria endemic countries, that we examined, we did not observe any deletions of *pfhrp2* or *pfhrp3* in any field isolates, suggesting that the parasites with gene deletions may not be widespread outside of South America. However, it should be noted that the sampling did not include some endemic regions (e.g. South and Western Asia) and the number of samples examined for some countries was quite small, not coming from different areas within the country. To ensure the performance of RDTs, efforts should be taken to monitor the existence and spread of parasites with gene deletions, especially when false negative results associated with a high parasitaemia are reported.

Pfhrp2 and *pfhrp3* variability appears to occur independent of the other, as no correlation between the lengths of the two genes was evident. However, both proteins share several repeats, such as type 1 and 2 repeats. These shared repeats are probably the basis for observed cross reactivity between PfHRP2 and PfHRP3 by monoclonal antibodies against PfHRP2 [8,18], which may contribute to the detection sensitivity of PfHRP2-detecting RDTs and reduce the effect of PfHRP2 variability on RDT sensitivity, particularly at high parasite densities.

The organization and the number of repeats in PfHRP2 vary extensively between parasite isolates. Theoretically, the presence and absence, as well as the number of repeats could affect the binding affinity of the antibodies used in RDTs to the parasites' antigen. Indeed, this preliminary analysis based on 16 cultured lines and tested on two earlier malaria RDTs, Paracheck Pf (Orchid Biomedical Systems, India) and ICT Malaria (ICT Diagnostics, South Africa), yielded a binary logistic regression model that was able to predict detection sensitivity of these two RDTs based on sequence structure [24]. However, the number of isolates used was limited and cultured parasites were used. In this study, the analysis was repeated using the recently completed WHO product testing of malaria RDTs (Round 1) results that included the testing of 34 PfHRP2-detecting RDTs tested against 79 isolates at 200 parasites/ μ L. This much larger sample size enabled more stringent analysis. The regression analysis this time did not show a correlation between PfHRP2 structure and the overall RDT detection rates. This result suggests that the performance of this group of RDTs is not greatly affected by the diversity of PfHRP2 at parasite densities of 200 parasites/ μ L or above and, therefore, should increase the confidence in the performance of the devices in various geographic locations. However, this does not exclude an effect of sequence diversity on RDT detection rates at

lower parasite densities, or by RDTs employing other monoclonal antibodies.

Conclusions

PfHRP2, the most common target of malaria RDTs, is highly polymorphic throughout the malaria-endemic regions included in this study. While this may affect RDT sensitivity at very low parasite densities (< 200 parasites/ μ L), it appears unlikely to have a major impact in most endemic regions at parasite densities usually seen in clinical malaria (200 parasites/ μ L and above). The extensive diversity in the *pfhrp2* sequence also provides evidence that there has not been a strong evolutionary selection for any particular type of sequence. The better understanding of the structure of PfHRP2 and its variation contributes to the evaluation and testing of malaria RDTs, the RDT quality assurance programmes and helps improve malaria RDTs.

Additional material

Additional file 1 Table S1: Comparison of the length (aa) and number of each repeat in PfHRP2 in parasites from different geographical areas.

Additional file 2 Table S2: Frequency (in percentage) of different amino acid repeats (types 1-24) observed in PfHRP2 of isolates obtained from various geographic areas and countries (sample number > 10).

Additional file 3 Table S3: Variation of PfHRP3 sequences.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JB1 carried out molecular genetic studies, sequence alignment, statistical analyses, and drafted the manuscript. MFH carried out molecular genetic studies and submission of sequences to GenBank. AP and MG carried out statistical analyses. NC assisted with molecular genetic studies for the Nigerian isolates. SA, FA, JB2, AA, DB, JC, DD, DFE, DG, JH, MPK, JL, CM1, DM, CM2, SN, BO, PO, WO, SQW supervised, carried out and coordinated field collection of isolates from patients. DB, JM and QC conceived of the study, participated in its design and coordination and helped to draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The work is partially funded by the AusAID, the World Health Organisation (WHO)- Regional Office for the Western Pacific, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and the Bill and Melinda Gates Foundation through the Foundation for Innovative New Diagnostics (FIND). DG is supported by Directorate General for Development Cooperation (DGCD) of the Belgian Government (Framework Agreement 03, 2008-2010, project 910000) and NIH/NIAID (RO1 AI06727-03), and JM is supported by the NHMRC, Australia. The authors wish to thank investigators and health workers in many countries for their assistance in collecting samples used in this study. We also thank Dr Katherine Trenholme, QIMR, and Alyson Auliff, AMI, for their contribution of isolates and laboratory lines to this study. The opinions expressed herein are those of the authors and do not necessarily reflect those of the WHO, the Australian Defence Force or any ADF extant policy.

Author Details

¹Department of Drug Resistance and Diagnostics, Australian Army Malaria Institute, Brisbane, Australia, ²Clinical Tropical Medicine Laboratory, Queensland Institute of Medical Research, University of Queensland, Herston, Australia, ³Malaria Drug Resistance and Chemotherapy Laboratory, Queensland Institute of Medical Research, Herston, Australia, ⁴Bagamoyo/Ikara Health Research and Development Centre, Ifakara, United Republic of Tanzania, ⁵Foundation for Innovative and New Diagnostics, Geneva, Switzerland, ⁶Pasteur Institute of Cambodia, Phnom Penh, Cambodia, ⁷Centre for Disease Control and Prevention, Atlanta, USA, ⁸Global Malaria Programme, World Health Organization, Geneva, Switzerland, ⁹UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), Geneva, Switzerland, ¹⁰Institut Pasteur de Bangui, Bangui, Central African Republic, ¹¹Centro Internacional de Entrenamiento e Investigaciones Medicas (CIDEIM), Cali, Colombia, ¹²Instituto de Medicina Tropical Alexander Von Humboldt, Universidad Peruana Cayetano Heredia, Peru, ¹³Departamento de Bioquímica, Biología Molecular y Farmacología, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Peru, ¹⁴Western Pacific Regional Office of the World Health Organization, Solomon Islands, ¹⁵Department of Medical Research (Lower Myanmar), Yangon, Myanmar, ¹⁶Research Institute for Tropical Medicine, Alabang, The Philippines, ¹⁷Institut Pasteur de Madagascar, Madagascar, ¹⁸Centre for Clinical Research, Kenya Medical Research Institute, Kisumu, Kenya, ¹⁹College of Medicine, University of Lagos, Odoaraba, Lagos, Nigeria, ²⁰Hainan Provincial Centre for Disease Control and Prevention, Haikou, Hainan, China and ²¹School of Population Health, University of Queensland, Herston, Queensland, Australia

Received: 8 March 2010 Accepted: 17 May 2010

Published: 17 May 2010

References

- Amexo M, Tolhurst R, Barnish G, Bates I: **Malaria misdiagnosis: effects on the poor and vulnerable.** *Lancet* 2004, **364**:1896-98.
- Wongsrichanalai C, Barcus MJ, Muth S, Sutamihardja A, Wernsdorfer WH: **A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT).** *Am J Trop Med Hyg* 2007, **77**(Suppl 6):119-127.
- Stow NW, Torrens JK, Walker J: **An assessment of the accuracy of clinical diagnosis, local microscopy and a rapid immunochromatographic card test in comparison with expert microscopy in the diagnosis of malaria in rural Kenya.** *Trans R Soc Trop Med Hyg* 1999, **93**:519-520.
- Wongsrichanalai C: **Rapid diagnostic techniques for malaria control.** *Trends Parasitol* 2001, **17**:307-309.
- Endeshaw T, Gebre T, Ngondi J, Graves PM, Shargie EB, Ejigsemahy Y, Ayele B, Yohannes G, Teferi T, Messele A, Zerihun M, Genet A, Mosher AW, Emerson PM, Richards FO: **Evaluation of light microscopy and rapid diagnostic test for the detection of malaria under operational field conditions: a household survey in Ethiopia.** *Malar J* 2008, **7**:118.
- Broek I van den, Hill O, Gordillo F, Angarita B, Hamade P, Counihan H, Guthmann JP: **Evaluation of three rapid diagnostic tests for diagnosis of *P. falciparum* and *P. vivax* in Colombia.** *Am J Trop Med Hyg* 2006, **75**:1209-15.
- Rakotonirina H, Barnadas C, Raheerijafy R, Hery H Andrianantenaina, Ratsimbaoa A, Randrianasolo L, Jahevitra M, Andriantsoanirina V, Menard D: **Accuracy and reliability of malaria diagnostic techniques for guiding febrile outpatient treatment in malaria-endemic countries.** *Am J Trop Med Hyg* 2008, **78**:217-221.
- Wellems TE, Howard RJ: **Homologous genes encode two distinct, histidine rich proteins in a cloned isolate of *Plasmodium falciparum*.** *Proc Natl Acad Sci USA* 1986, **83**:6065-6069.
- Panton LJ, McPhie P, Maloy WL, Wellems TE, Taylor DW, Howard RJ: **Purification and partial characterization of an unusual protein of *Plasmodium falciparum*: histidine rich protein II.** *Mol Biochem Parasitol* 1989, **35**:149-60.
- Rock EP, Marsh K, Saul AJ, Wellems TE, Taylor DW, Maloy WL, Howard RJ: **Comparative analysis of the *Plasmodium falciparum* histidine rich proteins HRP-I, HRP-II, and HRP-III in malaria parasites of diverse origin.** *Parasitology* 1987, **95**:209-227.
- Scherf A, Mattei D: **Cloning and characterization of chromosomes breakpoints of *Plasmodium falciparum*: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes.** *Nucleic Acids Res* 1992, **20**:1491-1496.
- Birku Y, Welday D, Ayele D, Shepherd A: **Rapid diagnosis of severe malaria based on the detection of PfHRP-2 antigen.** *Ethiop Med J* 1999, **37**:173-179.
- Howard RJ, Uni S, Aikawa M, Aley SB, Leech JH, Lew AM, Wellems TE, Rener J, Taylor DW: **Secretion of a malaria histidine rich protein (PfHRP II) from *Plasmodium falciparum*-infected erythrocytes.** *J Cell Biol* 1986, **103**:1269-1277.
- López R, Urquiza M, Curtidor H, Eduardo J Caminos, Mora H, Puentes A, Patarroyo ME: ***Plasmodium falciparum*: red blood cell binding studies of peptides derived from histidine rich KAHRP-I, HRP-II and HRP-III proteins.** *Acta Trop* 2000, **75**:349-359.
- Dondorp AM, Desakorn V, Pongtavornpinyo W, Sahassananda D, Silamut K, Chotivanich K, Newton PN, Pitisuttithum P, Smithyman AM, White NJ, Day NPJ: **Estimation of the total parasite biomass in acute *falciparum* malaria from plasma PfHRP2.** *PLoS Med* 2005, **2**:788-797.
- Hayward RE, Sullivan DJ, Day KP: ***Plasmodium falciparum*: histidine rich protein II is expressed during gametocyte development.** *Exp Parasitol* 2000, **96**:139-146.
- Wellems TE, Walliker D, Smith CL, do Rosario VE, Maloy WL, Howard RJ, Carter R, McCutchan TF: **A histidine rich protein gene marks a linkage group favoured strongly in a genetic cross of *Plasmodium falciparum*.** *Cell* 1987, **49**:633-642.
- Sharma YD: **Genomic organization, structure and possible function of histidine rich proteins of malaria parasites.** *Int J Biochem* 1988, **20**:471-477.
- Biswas S, Tomar D, Rao DN: **Investigation of the kinetics of histidine rich protein 2 and of the antibody responses to this antigen, in a group of malaria patients from India.** *Ann Trop Med Parasitol* 2005, **99**:553-562.
- Kemp DJ, Thompson JK, Walliker D, Corcoran LM: **Molecular karyotype of *Plasmodium falciparum*: conserved linkage groups and expendable histidine rich protein genes.** *Proc Natl Acad Sci USA* 1987, **84**:7672-7676.
- Turgut-Balik D, Akbulut E, Shoemark DK, Celik V, Moreton KM, Sessions RB, Holbrook JJ, Brady RL: **Cloning, sequence and expression of the lactate dehydrogenase gene from the human malaria parasite, *Plasmodium vivax*.** *Biotechnol Lett* 2004, **26**:1051-1055.
- Akbulut E, Celik V, Balik DT: **Comparative analysis at the nucleotide level of the genes encoding the lactate dehydrogenase enzyme of *Plasmodium vivax* and *Plasmodium falciparum*.** *Turkiye Parazitoloj Derg* 2005, **29**:149-153.
- Lee N, Baker J, Bell D, McCarthy J, Cheng Q: **Assessing the genetic diversity of the aldolase genes of *Plasmodium falciparum* and *Plasmodium vivax* and its potential effect on performance of aldolase-detecting rapid diagnostic tests.** *J Clin Microbiol* 2006, **44**:4547-4549.
- 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.
- Lee N, Baker J, Andrews KT, Gatton ML, Bell D, Cheng Q, McCarthy J: **Effect of sequence variation in *Plasmodium falciparum* histidine rich protein 2 on binding of specific monoclonal antibodies: implications for rapid diagnostic tests for malaria.** *J Clin Microbiol* 2006, **44**:2773-2778.
- World Health Organization: **Malaria rapid diagnostic tests performance: results of WHO product testing of malaria RDTs: round 1 (2008)** Geneva; 2009. ISBN 978 92 4 1598071
- Pologe LG, Ravetch JV: **Large deletions result from breakage and healing of *P. falciparum* chromosomes.** *Cell* 1988, **55**:869-874.
- Figueiredo LM, Freitas-Junior LH, Bottius E, Olivo-Marin J-C, Scherf A: **A central role for *Plasmodium falciparum* subtelomeric regions in spatial positioning and telomere length regulation.** *EMBO J* 2002, **21**:815-824.
- Volkman SK, Hartl DL, Wirth DF, Nielsen KM, Choi M, Batalov S, Zhou Y, Plouffe D, Le Roch KG, Abagyan R, Winzler EA: **Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*.** *Science* 2002, **298**:216-218.
- Ribacke U, Mok BW, Wirta V, Normark J, Lundeberg J, Kironde F, Egwang TG, Nilsson P, Wahlgren M: **Genome wide gene amplifications and deletions in *Plasmodium falciparum*.** *Mol Biochem Parasitol* 2007, **155**:33-44.
- Vernick KD, McCutchan TF: **Sequence and structure of a *Plasmodium falciparum* telomere.** *Mol Biochem Parasitol* 1988, **28**:85-94.
- Scherf A: ***Plasmodium* telomeres and telomere proximal gene expression.** *Semin Cell Biol* 1996, **7**:49-57.

33. Anders RF, Brown GV, Coppel RL, Kemp DJ: **Repeat structures in malaria antigens.** *PNG Med J* 1986, **29**:87-93.
34. Kemp DJ, Coppel RL, Anders RF: **Repetitive proteins and genes of malaria.** *Annu Rev Microbiol* 1987, **41**:181-208.
35. Weber JL: **Molecular biology of malaria parasites.** *Exp Parasitol* 1988, **66**:143-170.
36. Smith GP: **Evolution of repeated DNA sequences by unequal crossover.** *Science* 1976, **191**:528-535.
37. Levinson G, Gutman GA: **Slipped-strand mispairing: A major mechanism for DNA sequence evolution.** *Mol Biol Evol* 1987, **4**:203-221.
38. Dodin G, Levoir P: **Replication slippage and the dynamics of the immune response in malaria: a formal model for immunity.** *Parasitology* 2005, **131**:727-35.
39. Hughes AL: **The evolution of amino acid repeat arrays in Plasmodium and other organisms.** *J Mol Evol* 2004, **59**:528-35.
40. Sullivan DJ Jr, Gluzman IY, Goldberg DE: **Plasmodium hemozoin formation mediated by histidine-rich proteins.** *Science* 1996, **271**:219-22.
41. Choi CY, Cerda JF, Chu HA, Babcock GT, Marletta MA: **Spectroscopic characterization of the heme-binding sites in Plasmodium falciparum histidine-rich protein 2.** *Biochemistry* 1999, **38**:16916-24.
42. Benedetti CE, Kobard J, Pertinhez TA, Gatti RM, de Souza ON, Spisni A, Meneghini R: **Plasmodium falciparum histidine-rich protein II binds to actin, phosphatidylinositol 4,5-bisphosphate and erythrocyte ghosts in a pH-dependent manner and undergoes coil-to-helix transitions in anionic micelles.** *Mol Biochem Parasitol* 2003, **128**:157-66.
43. Bosshart H, Heinzelmann M: **Endotoxin-neutralizing effects of histidine-rich peptides.** *FEBS Lett* 2003, **553**:135-140.
44. 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 pfhrp2 and pfhrp3: implications for malaria rapid diagnostic tests.** *PLoS ONE* 2010, **5**:e8091.

doi: 10.1186/1475-2875-9-129

Cite this article as: Baker *et al.*, Global sequence variation in the histidine-rich proteins 2 and 3 of Plasmodium falciparum: implications for the performance of malaria rapid diagnostic tests *Malaria Journal* 2010, **9**:129

**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

