

**Evolution of *pfhrp2* and *pfhrp3* deletions in Equatorial Guinea between the pre – and post – RDT introduction.**

Irene Molina - de la Fuente<sup>1,2,3</sup>, M. Andreína Pacheco<sup>4</sup>, Luz García<sup>2,3</sup>, Vicenta González<sup>2,3</sup>, Matilde Riloha<sup>5</sup>, Consuelo Oki<sup>5</sup>, Agustín Benito<sup>2,3</sup>, Ananias A. Escalante<sup>4</sup>, Pedro Berzosa<sup>2,3</sup>

<sup>1</sup>Biomedicine and biotechnology Department, University of Alcalá, Ctra.Madrid-Barcelona Km.33,600, 28871, Alcalá de Henares, Spain

<sup>2</sup>National Centre of Tropical Medicine, Carlos III Institute of Health, C/ Sinesio Delgado 10, 28029, Madrid, Spain

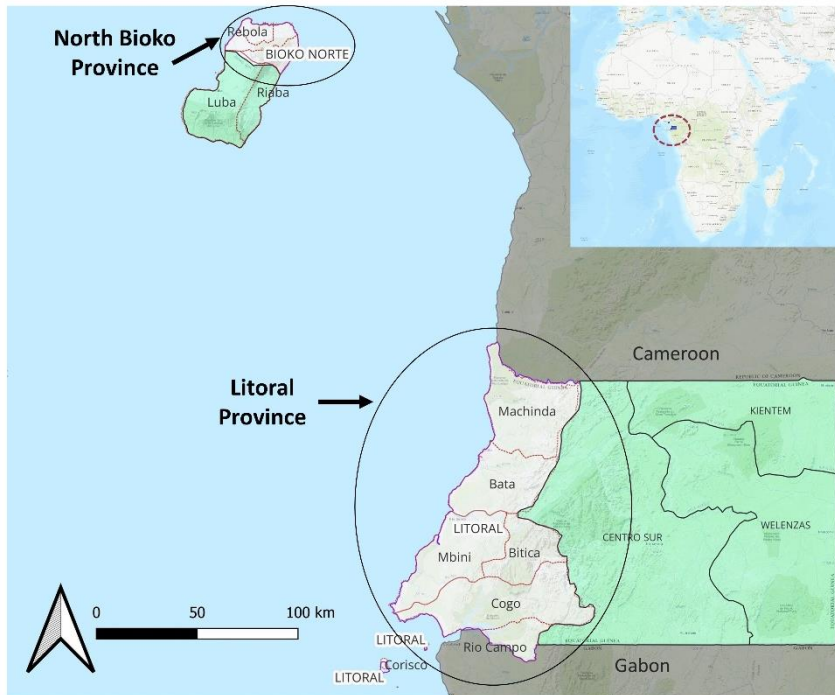
<sup>3</sup>Consortio Centro de Investigación Biomédica en Red – CIBERINFEC ISCIII, C/ Sinesio Delgado 10, 28029, Madrid, Spain

<sup>4</sup> Biology Department/Institute of Genomics and Evolutionary Medicine (iGEM), Temple University, (SERC - 645), 1925 N. 12 St, Philadelphia, PA, 19122-1801, USA.

<sup>5</sup>National Programme for Malaria Control, Ministry of Health and Social Welfare (MINSABS), Malabo, Equatorial Guinea

## Supplementary material

**Figure S1.** The Equatorial Guinea map shows the two provinces where samples were collected, Litoral and Bioko Norte.



**Table S1.** Primers for each microsatellite combined with GeneScan350 ROX dye size standard. bp: base pairs of the amplified fragment. Ch: chromosome

Micrt.	Ch	PCR	Primer				bp
			Forward		Reverse		
<b>Poly A</b>	6	1°	PolyA-F	AAA ATA TAG ACG AAC AGA	PolyA-R	ATC AGA TAA TTG TTG GTA	159-204
		2°	PolyA-2-F	AAA ATA TAG ACG AAC AGA	PolyA-3IR	<b>*6-FAM-</b> GAA ATT ATA ACT CTA CCA	
<b>TA1</b>	4	1°	TA1-3F	CTA CAT GCC TAA TGA GCA	TA1-R	TTT TAT CTT CAT CCC CAC	114-241
		2°	TA1-2-F	<b>*6-FAM-</b> CCG TCA TAA GTG CAG AGC	TA1-2-F	TTT TAT CTT CAT CCC CAC	
<b>PfPK2</b>	12	1°	PfPK2-F	CTT TCA TCG ATA CTA CGA	PfPK2-R	CCT CAG ACT GAA ATG CAT	159-192
		2°	PK2-2-F	TAG TAA CGA TGT TGA CAA	PK2-2-R	<b>*HEX-</b> AAA AAG GAG GAT AAA TAC AT	
<b>TA109</b>	6	1°	TA109_3-F	TAG GGA ACA TCA TAA GGA T	TA109_3-R	CCT ATA CCA AAC ATG CTA AA	154-208
		2°	TA109-2-F	<b>*HEX -</b> GGTTAAATCAGGACAACAT	TA109-2-R	CCT ATA CCA AAC ATG CTA AA	
<b>2490</b>	10	1°	2490-F	TTC TAA ATA GAT CCA AAG	2490-R	ATG ATG TGC AGA TGA CGA	75 - 84
		2°	2490-2-F	TTC TAA ATA GAT CCA AAG	2490-2-R	<b>*6-FAM -</b> AGA ATT ATT GAA TGC AC	
<b>313</b>	2		313-F :	<b>*6-FAM -</b> TCCCTTTTAAAATAGAAGAAA	313-R	GATTATATGAAAGGATACATG	225 – 246
<b>383</b>	3		383-F	<b>*HEX -</b> AATAGGAACAAATCATATTG	383-R	AGATATCCAGGTAATAAAAAG	122 - 149

**Table S2.** Frequency of deletion in exon 1-2 and exon 2 of *pfhrp2* and *pfhrp3* genes by province per year. Significance of differences were assessed using  $\chi^2$ . F= frequency of deletion in percentage.

Province	Year	N	<i>Pfhrp2</i>			<i>Pfhrp3</i>		
			Exon1-2	Exon 2	Any deletion	Exon 1-2	Exon 2	Any deletion
			F % (n) 95% CI	F % (n) 95% CI	F % (n) 95% CI	F % (n) 95% CI	F % (n) 95% CI	F % (n) 95% CI
North Bioko	1999 - 2001	127	0.79 (1) 0.14 – 4.33	4.72 (6) 2.18 – 9.92	5.51 (7) 2.70 – 10.94	4.72 (6) 2.18 – 9.92	3.15 (4) 1.21 – 7.82	7.87 (10) 3.84 – 14.00
	2018	131	3.88 (4) 1.19 – 7.59	8.40 (11) 4.75 – 14.41	11.45 (15) 7.06 – 7.52	12.21 (16) 7.66 – 18.92	5.34 (7) 2.61 – 10.62	16.79 (22) 11.35 – 24.11
	<i>p – value</i>		0.385	0.348	0.216	0.054	0.573	0.047
Litoral	2001	113	0.0 (0) 0.00 – 3.29	2.65 (3) 0.91 – 7.52	2.65 (3) 0.91 – 7.52	6.19 (7) 3.03 – 12.24	13.27 (15) 8.21 – 20.75	17.70 (20) 11.76 – 25.76
	2019	195	36.41 (71) 29.65 – 43.59	19.49 (38) 14.17 – 25.75	43.08 (84) 36.02 – 50.34	37.95 (74) 31.11 – 45.16	18.97 (37) 13.73 – 25.19	50.77 (99) 43.53 – 57.98
	<i>P - value</i>		< 0.01	< 0.01	< 0.01	< 0.01	0.222	< 0.01

**Table S3.** Frequency of each deletion haplotype by province and year. The significance of differences was assessed using an X square. N = number of samples

Province	Year	N	Frequency of combination in percentage (number of samples) 95% Confidence Interval			
			<i>pfhrp2 - / pfhrp3 -</i>	<i>pfhrp2 - / pfhrp3 +</i>	<i>pfhrp2 + / pfhrp3 -</i>	<i>pfhrp2 + / pfhrp3 +</i>
North Bioko	1999 - 2001	127	0.0	4.72 (6) 1.75 – 10.0	3.15 (4) 0.86 – 7.87	92.13 (117) 86.0 – 96.16
	2018	131	1.53 (2) 0.16 – 5.41	6.87 (9) 3.19 – 12.64	3.82 (5) 1.25 – 8.68	87.79 (115) 80.92 – 92.85
	<i>p - value</i>		0.4459			
Litoral	2001	113	0.88 (1) 0.02 – 4.83	1.77 (2) 0.22 - 6.25	12.39 (14) 6.94 – 19.91	84.96 (96) 77.01 – 90.99
	2019	195	7.18 (14) 3.98 – 11.75	12.31 (24) 8.05 – 17.76	11.79 (23) 7.63 – 17.17	71.28 (139) 64.38 – 77.52
	<i>p - value</i>		< 0.005			

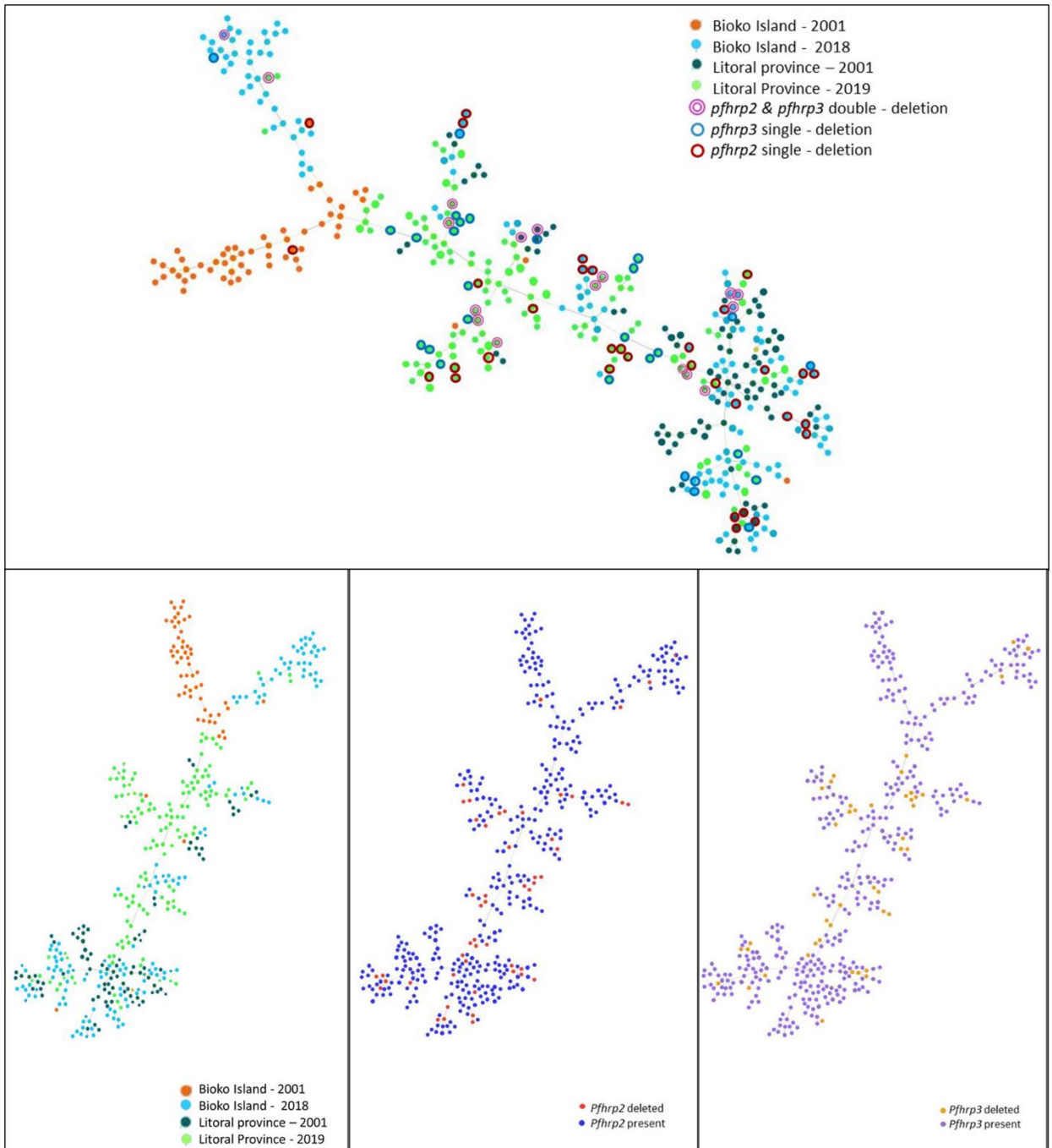
**Table S4.** Frequency of deletions in flanking regions of *pfhrp2* and *pfhrp3* by province and year. The significance of differences was assessed using an X square.

Province	Year	N	Flanking <i>Pfhrp2</i>		Flanking <i>Pfhrp3</i>	
			heat shock protein 70	<i>Plasmodium</i> exported protein (PHIST) - pseudogene	<i>Plasmodium</i> exported protein (PHISTb)	acyl-CoA synthetase
			PF3D7_0831700	PF3D7_0831900	PF3D7_1372100	PF3D7_1372400
			F (n) 95% CI	F (n) 95% CI	F (n) 95% CI	F (n) 95% CI
North Bioko	1999 - 2001	127	0.79 (1) 0.14 – 4.33	6,30 (8) 3.23 – 11.94	0,0 (0) 0.00 – 2.94	8.66 (11) 4.91 – 14.85
	2018	131	0.76 (1) 0.02 – 4.18	45.04 (59) 36.78 – 53.58	2.29 (3) 0.18 – 6.52	22.90 (30) 16.54 – 30.81
	<i>p – value</i>		1	< 0.05	0.259	0.003
Litoral	2001	113	0.00 (0) 0.00 – 3.29	9.73 (11) 5.52 – 16.59	6.19 (7) 3.03 – 12.24	7.08 (8) 3.63 – 13.35
	2019	195	29.23 (57) 22.95 – 36.15	49.23 (96) 42.02 – 56.47	50.77 (99) 43.53 - 57.98	41.03 (80) 34.05 – 48.28
	<i>p – value</i>		< 0.05	< 0.05	< 0.05	< 0.05

**Table S5. Genetic polymorphism of infections included.** \* It refers to monoclonal isolates or isolates with more than one allele only for one locus. That analysis excluded polyclonal infections with more than one allele for more than one locus because using microsatellites it is not possible to know what alleles conform each clone.

Province	Year	Number of isolates	Multiple infection	Single infections	Percentage of multiple infections	95% CI for percentage of multiple infections	Total isolates included for population structure*	Total clones included
<b>Litoral</b>	2001	81	62	19	76.5	65.8 – 85.2	44	74
	2019	105	70	35	66.7	56.8 – 75.6	68	105
	<i>p-value</i> ( $\chi^2$ )				0.1907 (1.712)			
<b>North Bioko</b>	2001	74	41	33	55.4	43.4 – 67.0	38	51
	2018	105	66	39	62.9	52.9 – 72.1	74	124
	<i>p-value</i> ( $\chi^2$ )				0.3973 (0.7165)			
<b>Overall</b>		365	239	126	65.5	60.4 – 70.4	150	354

**Figure S2.** A minimum spanning tree of microsatellite allelic data showing genetic relatedness of *Plasmodium falciparum* populations according to their origin or *pfhrp2/3* deletion status.



**Table S6.** Number of samples included for microsatellites analysis and exon 2 genetic analysed by year and province.

Province	Year	Aminoacidic analysis		Microsatellite analysis
		<i>Pfhrp2</i>	<i>Pfhrp3</i>	
North Bioko	1999 – 2001	26	23	56
	2018	26	20	147
Litoral	2001	44	19	91
	2019	48	23	136

**Table S7.** Frequency and occurrence of amino acid repeat types in exon 2 of *pfhrp2*. N: Range of number of repetitions of each amino acid repeat. F = Occurrence of each type of repeat. The repetitions that are not showed were not found in the analysed sequences.

Types	Amino acid repeat	1999 – 2001		2018		2001		2019	
		N	F (%)	N	F (%)	N	F (%)	N	F (%)
Type 1	AHHAHHVAD	0 - 5	50.0	0 - 7	84.6	0 - 5	65.9	0 - 4	52.1
Type 2	AHHAHHAAD	0 - 13	88.5	3 - 18	<b>100.0</b>	2 - 15	<b>100.0</b>	3 - 15	<b>100.0</b>
Type 3	AHHAHHAAY	0 - 2	61.5	0 - 2	80.8	0 - 3	84.1	0 - 3	89.6
Type 4	AHH	0 - 14	76.9	0 - 18	76.9	0 - 31	77.3	0 - 19	85.4
Type 5	AHHAHHASD	0 - 2	61.5	0 - 2	69.2	0 - 2	77.3	0 - 2	62.5
Type 6	AHHATD	0 - 6	73.1	1 - 7	<b>100.0</b>	0 - 8	90.9	0 - 6	91.7
Type 7	AHHAAD	0 - 11	88.5	3 - 10	<b>100.0</b>	1 - 10	<b>100.0</b>	1 - 11	<b>100.0</b>
Type 8	AHHAAY	0 - 3	61.5	0 - 2	84.6	0 - 4	93.2	0 - 3	70.8
Type 9	AAAY	0 - 1	3.8	0	0.0	0	0.0	0 - 1	2.1
Type 10	AHHAAHHTD	0 - 2	11.5	0 - 2	53.8	0 - 3	47.7	0 - 2	27.1
Type 13	AHHASD	0 - 2	15.4	0 - 1	77.0	0 - 2	15.9	0 - 3	22.9
Type 14	AHHAHHATD	0 - 2	19.2	0 - 1	3.8	0 - 2	11.4	0 - 1	6.3
Type 19	AHHAA	0 - 24	92.3	0 - 1	42.3	0 - 4	31.8	0 - 3	45.8

**Table S8.** Frequency and occurrence of amino acid repeat types in exon 2 of *pfhrp3*. N: Range of number of repetitions of each amino acid repeat. F = Occurrence of each type of repeat.

Types	Amino acid repeat	1999 – 2001		2018		2001		2019	
		N	F (%)	N	F (%)	N	F (%)	N	F (%)
Type 1	AHHAHHVAD	0 - 5	26.1	0 - 3	75.0	0 - 3	57.9	0 - 3	26.1
Type 4	AHH	0 - 2	78.3	0 - 2	65.0	0 - 2	94.7	0 - 2	78.3
Type 6	AHHATD	0	0	0	0	0 - 1	26.3	0	0
Type 7	AHHAAD	1 - 1	<b>100.0</b>	1 - 2	<b>100.0</b>	1 - 1	<b>100.0</b>	0 - 1	91.3
Type 15	AHHAHHAAN	0 - 1	30.4	0 - 1	80.0	0 - 13	84.2	0 - 1	60.9
Type 16	AHHAAN	9 - 18	<b>100.0</b>	8 - 15	<b>100.0</b>	0 - 15	<b>100.0</b>	0 - 15	95.7
Type 17	AHHDG	0 - 9	95.7	0 - 7	95.0	2 - 10	<b>100.0</b>	0 - 7	91.3
Type 18	AHHDD	0 - 3	91.3	0 - 3	95.0	1 - 14	<b>100.0</b>	0 - 3	82.6
Type 19	AHHAA	0	0	0	0.0	0 - 1	36.8	0	0
Type 20	SHHDD	1 - 1	<b>100.0</b>	0 - 1	95.0	0 - 1	68.4	0 - 1	91.3

**Table S9.** Occurrence of amino acid repeat types in exon 2 of *pfhrp2*. N: number of repetitions of each amino acid repeat.

		<i>P.</i> <i>praealciparum</i>	<i>P.</i> <i>reichnowi</i>	<i>P.</i> <i>billcolinsi</i>	<i>P.</i> <i>blacklocki</i>	<i>P. gaboni</i>	<i>P. adleri</i>
Types	Amino acid repeat	N	N	N	N	N	N
Type 1	AHHAHHVAD	1	0	0	0	0	0
Type 2	AHHAHHAAD	14	<b>0</b>	0	<b>0</b>	0	<b>0</b>
Type 3	AHHAHHAAY	1	0	0	0	0	0
Type 4	AHH	2	12	9	17	26	0
Type 5	AHHAHHASD	0	0	0	0	0	0
Type 6	AHHATD	0	0	0	<b>0</b>	0	0
Type 7	AHHAAD	4	<b>3</b>	0	<b>0</b>	0	<b>0</b>
Type 8	AHHAAY	1	0	0	0	0	0
Type 9	AAAY	0	0	0	0	0	0
Type 10	AHHAAHHATD	3	0	0	0	0	0
Type 12	AHHAAAHHEAATH	1	0	0	0	0	0
Type 13	AHHASD	0	0	0	0	0	0
Type 14	AHHAHHATD	1	0	0	0	0	0
Type 16	AHHAAN	0	19	0	0	0	0
Type 17	AHHDG	0	0	0	19	0	0
Type 19	AHHAA	0	5	0	0	0	0

**Table S10.** Frequency and occurrence of amino acid repeat types in exon 2 of *pfhrp3*. N: Number of repetitions of each amino acid repeat. The repetitions that are not showed were not found in the analysed sequences.

		<i>P. praefalciparum</i>	<i>P. reichenowi</i>	<i>P. blacklocki</i>	<i>P. gaboni</i>	<i>P. adleri</i>
Types	Amino acid repeat	N	N	N	N	N
Type 1	AHHAHVAD	1	1	0	0	0
Type 4	AHH	1	19	17	0	0
Type 5	AHHAHASD	1	0	0	0	0
Type 7	AHHAAD	1	7	0	0	0
Type 15	AHHAHHAAN	1	0	0	0	0
Type 16	AHHAAN	10	0	0	0	0
Type 17	AHHDG	6	3	19	0	0
Type 18	AHHDD	2	5	0	0	0
Type 19	AHHAA	0	2	0	0	0
Type 20	SHHDD	1	1	0	0	0

**Table S110.** Frequency of linear epitopes in exon2 of *pfhrp2* and *pfhrp3*. P = Prevalence Median (n) = Median frequency (n).

MAb	Major epitope	1999		2018		2001		2019	
		P (%)	Median (n)	P (%)	Median (n)	P (%)	Median (n)	P (%)	Median (n)
<i>Pfhrp2</i>									
3A4	AHHAHHA	92.3	12	100.0	14	100.0	14	100.0	14
2G12-1C12	DAHHAADAHH	76.9	3	100.0	5	93.2	5	97.9	4
1E1-A9	AHHAHHV	50.0	0.5	84.6	2	63.6	1	56.3	1
A6-4	HATDAHH	73.1	1.5	100.0	3	88.6	4	95.8	3
C1-13	AHHAADAHH	92.3	12	100.0	16	100.0	17	100.0	16
N7	DAHHAADAHHA	73.1	3	100.0	5	93.2	5	97.9	4
PTL-3	YAHHAHHA	76.9	2	96.2	2	95.5	3	95.8	2
S2-5	AHHASDAHHA	57.7	1	73.1	1	81.8	1	72.9	1
TC-10	TDAHHAADAHHAADA	30.8	0	80.8	1	61.4	1	64.6	1
C2-3	HAHHAHHAADAHH	15.4	0	19.2	0	27.3	0	25.0	0
Genway	AYAHHAHHAAY	0.0	0	0.0	0	2.3	0	2.1	0
<i>Pfhrp3</i>									
3A4	AHHAHHA	30.5	0	85	1	84.2	1	65.2	1
1E1-A9	AHHAHHV	26.1	0	80	1	57.9	1	26.1	0

**Table S12.** Occurrence of linear epitopes in exon2 of *hrp2* and *hrp3* in other species from subgenus *Laverania*. N = number of repetitions of epitopes.

		<i>P. praefalciparum</i>	<i>P. reichnowi</i>	<i>P. billcolinsi</i>	<i>P. blacklocki</i>	<i>P. gaboni</i>	<i>P. adleri</i>
		N	N	N	N	N	N
<b><i>hrp2</i></b>							
<b>3A4</b>	AHHAHHA	18	0	NA	7	0	0
<b>2G12-1C12</b>	DAHHAADAHH	4	2	NA	0	0	0
<b>1E1-A9</b>	AHHAHHV	1	0	NA	1	0	0
<b>A6-4</b>	HATDAHH	4	0	NA	0	0	0
<b>C1-13</b>	AHHAADAHH	18	3	NA	0	0	0
<b>N7</b>	DAHHAADAHHA	4	1	NA	0	0	0
<b>PTL-3</b>	YAHHAHHA	2	0	NA	0	0	0
<b><i>hrp3</i></b>							
<b>3A4</b>	AHHAHHA	2	0	NA	7	0	0
<b>2G12-1C12</b>	DAHHAADAHH	0	4	NA	0	0	0
<b>1E1-A9</b>	AHHAHHV	1	3	NA	1	0	0
<b>C1-13</b>	AHHAADAHH	0	7	NA	0	0	0
<b>N7</b>	DAHHAADAHHA	0	3	NA	0	0	0
<b>S2-5</b>	AHHASDAHHA	1	0	NA	0	0	0