RESEARCH

CrossMark

Comparison of allele frequencies of *Plasmodium falciparum* merozoite antigens in malaria infections sampled in different years in a Kenyan population

Lynette Isabella Ochola-Oyier[1*], John Okombo[1], Njoroge Wagatua[1], Jacob Ochieng[1], Kevin K. Tetteh[2], Greg Fegan[1], Philip Bejon[1] and Kevin Marsh[1]

Abstract

Background: *Plasmodium falciparum* merozoite antigens elicit antibody responses in malaria-endemic populations, some of which are clinically protective, which is one of the reasons why merozoite antigens are the focus of malaria vaccine development efforts. Polymorphisms in several merozoite antigen-encoding genes are thought to arise as a result of selection by the human immune system.

Methods: The allele frequency distribution of 15 merozoite antigens over a two-year period, 2007 and 2008, was examined in parasites obtained from children with uncomplicated malaria. In the same population, allele frequency changes pre- and post-anti-malarial treatment were also examined. Any gene which showed a significant shift in allele frequencies was also assessed longitudinally in asymptomatic and complicated malaria infections.

Results: Fluctuating allele frequencies were identified in codons 147 and 148 of reticulocyte-binding homologue (Rh) 5, with a shift from HD to YH haplotypes over the two-year period in uncomplicated malaria infections. However, in both the asymptomatic and complicated malaria infections YH was the dominant and stable haplotype over the two-year and ten-year periods, respectively. A logistic regression analysis of all three malaria infection populations between 2007 and 2009 revealed, that the chance of being infected with the HD haplotype decreased with time from 2007 to 2009 and increased in the uncomplicated and asymptomatic infections.

Conclusion: Rh5 codons 147 and 148 showed heterogeneity at both an individual and population level and may be under some degree of immune selection.

Keywords: Rh5, Merozoite, *Plasmodium*, Malaria

Background

Genes involved in immune interactions tend to contain adaptive variation that is maintained by balancing selection [1, 2]. In *Plasmodium falciparum* balancing selection is evidenced by the presence of multiple polymorphisms in several merozoite antigens, such as merozoite surface protein-1 (MSP-1) [3], apical membrane antigen-1 (AMA-1) [4], erythrocyte binding antigen-175 (EBA-175) [5], MSP3 [6], MSP Duffy Binding Ligand-1 (DBL-1) [7], MSPDBL2 [8], reticulocyte binding homologues Rh2a and Rh2b [9, 10].

Since merozoite proteins are targets of host immunity [10–13] polymorphisms are thought to be maintained by selection from host immune responses. There is evidence to suggest polymorphisms could result in primarily allele-specific immune responses [6, 14, 15]. Therefore, an individual may resist infection to one parasite allele for which they have pre-existing immunity but not to a heterologous allele that they have not yet encountered.

*Correspondence: liochola@kemri-wellcome.org
[1] KEMRI-Wellcome Trust Collaborative Programme, P.O. Box 230, Kilifi 80108, Kenya
Full list of author information is available at the end of the article

BioMed Central

Ochola-Oyier *et al. Malar J (2016) 15:261*

Page 2 of 11

Initially over time, rare alleles attracting less 'immune attention' would be expected to progressively proliferate while dominant alleles are cleared by the immunity they induce [16]. Thus, the polymorphisms are potentially maintained by a mechanism of frequency-dependent selection. Allele-specific immunity may limit the efficacy of vaccines developed against the merozoite blood stage antigens [17–19].

Stable balancing selection predicts constant intermediate allele frequencies over long periods [20], which change more slowly over time than they would under neutral genetic drift (2). This has been demonstrated in merozoite antigens in longitudinal studies, which primarily assessed repeat regions in MSP1 and MSP2 alleles [16, 21–24]. Similar observations were made with other merozoite antigens SURFIN4.1 [25], SURFIN4.2 [26] and GLURP [27]. In contrast, under fluctuating selection due to spatial and temporal variation, allele frequencies dynamically change over time becoming common when advantageous and rare when disadvantageous [28, 29], and allele frequencies should change more quickly over time than they would under neutral genetic drift [2]. There are a few cases where some merozoite antigen polymorphisms fluctuate in frequency over time [16, 21, 30, 31].

The basis of the frequency-dependent selection operating on merozoite antigens is not entirely understood. It is likely to reflect both the function of merozoite antigens as invasion receptors (erythrocyte-merozoite interactions) and the impact of the immune system on the merozoite. Understanding the distribution of merozoite antigen polymorphisms in natural parasite populations can inform the vaccine design process. The main aim of this study was to analyse the allele frequency distribution of 15 merozoite antigens in uncomplicated malaria infections from children under 5 years of age in an endemic area in Kilifi, Kenya over a two-year period, in order to identify single nucleotide polymorphisms (SNPs) under selection. Any significantly fluctuating SNPs were then investigated in asymptomatic and complicated malaria infections. Two SNPs (codons 147 and 148) in Rh5 were identified, which significantly fluctuated over the two-year period in uncomplicated malaria infections and remained stable in the asymptomatic and complicated malaria infections. Within the same population (the children under 5 years of age with uncomplicated malaria) the changes in allele frequencies pre- and post-anti-malarial treatment were also studied. Rh5 codons 147 and 148 were more likely to change during the course of an infection pre- and post-drug treatment. Additionally, the extent of genetic diversity was examined within and between the genes usling linkage disequilibrium (LD).

Methods

Longitudinal *Plasmodium falciparum* population samples

Merozoite antigenic diversity was examined by genotyping 15 genes (see Additional file 1) in 100 parasites obtained from children aged up to five years with uncomplicated malaria, between 2007 (n = 62) and 2008 (n = 38), recruited into a randomized, controlled trial of dihydroartemisinin-piperaquine versus artemether–lumefantrine in Pingilikani sub-location, Kilifi County (Trial registration: Controlled-Trials.com ISRCTN88705995) [32]. Any gene that showed a significant temporal change in allele frequency was also sequenced in three other parasite populations (to determine if the change was observed in other malaria infection populations): (1) asymptomatic malaria infections from an annual community survey of parasite-positive children aged up to 12 years; 266 cross-sectional venous bleed samples were obtained before the malaria season in 2007 (n = 108), 2008 (n = 104) and 2009 (n = 54) in Junju sub-location, Kilifi County, Kenya. Junju and Pingilikani are generally high to moderate transmission areas, although transmission is falling in Kilifi County as a whole [33]. Pingilikani and Junju neighbour each other and the dispensaries are at least 10 km apart; (2) 300 isolates taken from parasite-positive children aged up to 12 years admitted to the Kilifi District Hospital (KDH) between 2007 and 2009 (100 isolates each year). This was later extended to 100 samples each over a ten-year period from 2000 to 2010. These samples were defined as complicated malaria infections and were also representative of the parasite population in Kilifi County since KDH serves a population within an 891 km^2 area of 15 administrative locations within the Kilifi Health and Demographic Surveillance System (KHDSS) [34]; and, (3) 11 genotypically distinct laboratory isolates obtained from geographically diverse *P. falciparum* sources (3D7, cloned from an airport malaria case in The Netherlands; RO33 and Palo Alto, from Africa; K1, Dd2, FCC2 and D10, V1/S from Southeast Asia; Wellcome, nominally from Africa but suspected to have been previously cross-contaminated by parasites of unknown source during culture; HB3 from Honduras and IT from Brazil). Parasite DNA was extracted from packed frozen erythrocytes using the QIAamp DNA Blood Mini Kit (QIAGEN, UK).

Pre- and post-drug treatment *Plasmodium falciparum* samples

The uncomplicated malaria samples were obtained from the randomized, controlled drug trial described above. The pre-treatment sample (day 0) was obtained on recruitment of a child into the study. The children were followed up weekly until day 63 and finally on day 84, resulting in a post-treatment sample (range 19–84 days,

Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 3 of 11

mean number of days = 50) if a child was parasite slide-positive during a follow-up visit. The samples were previously genotyped using MSP1, MSP2 and glutamate-rich protein (GLURP) [32] and a large majority of the post-treatment samples were re-infections (about 64 %) and 8 % were recrudescent infections the remaining were indeterminate. Parasite genomic DNA was extracted from dried filter-paper blood spots using the boiling method described elsewhere [35]. Briefly, this involved snipping a small piece of the dried filter paper, fixing it in methanol and boiling at 100 °C in TB Elution buffer to extract the DNA. All the above studies were reviewed and approved by the Scientific Steering Committee and the Ethics Committee of the Kenya Medical Research Institute. Consent for study participation was obtained from the parents or legal guardian of the study participants.

PCR and sequencing of polymorphic *Plasmodium falciparum* genes

The 15 merozoite antigens were amplified using High Fidelity Taq polymerase (Roche) (details of regions amplified and primers are shown in Additional file 1). PCR-generated products were visualized on 1 % agarose gels, purified using ethanol precipitation and directly sequenced using the PCR primers, additional sequencing primers, BIG DYE terminator chemistry v3.1 (Applied Biosystems, UK) and an ABI 3130xl capillary sequencer (Applied Biosystems, UK). Nucleotide positions which displayed a peak within a peak in the electropherogram were defined as a mixed base and excluded from further analysis. Sequences were assembled, edited and aligned using SeqMan and MegAlign (DNASTAR, Madison, WI, USA). SNPs were identified and using their corresponding amino acids, haplotypes were defined.

Statistical analysis of longitudinal *Plasmodium falciparum* polymorphic genes

To analyse SNP frequency dynamics in the 15 merozoite genes, comparisons were conducted within the uncomplicated malaria parasite population for SNPs that showed a significant temporal shift (between 2007 and 2008) using the Fisher's exact test. SNPs with a >5 % minor allele frequency (MAF) were translated into their corresponding amino acids, in order to define haplotypes, which were also analysed. Any gene which maintained significance after a Bonferroni correction for multiple comparisons [36] was analysed further for temporal allele frequency changes in a primarily asymptomatic malaria parasite population between 2007 and 2009, and in a complicated malaria parasite population obtained from KDH between 2007 and 2009. A logistic regression analysis was performed on all the three parasite populations, controlling for year, age and population. Each gene

was then examined for allele frequency changes over a ten-year period in the KDH sample set. Within the ten-year complicated malaria infection samples, the homogeneity of the infected population was determined and investigated if there were any linear trends of location, age and severe malaria syndrome, coma, deep breathing, and anaemia [37] using the 'tabodds' command in Stata. All analyses were conducted using Stata v11 (Stata-Corp, TX, USA) and R. Using the available latitude and longitude co-ordinates, the Rh5 haplotypes from the uncomplicated (Pingilikani), asymptomatic (Junju) and complicated (KDH) malaria parasite populations were plotted on to the KHDSS map [34]. The spatial scan statistic (SaTScan) software [38] was then employed as in Bejon et al. [39], to identify significant haplotype clusters, using a Bernoulli (case–control) model to test for significant clusters between the two dominant haplotypes and a multinomial model to test between all the haplotypes in each of the parasite populations between 2007 and 2009.

Statistical analysis of pre- and post-drug treatment *Plasmodium falciparum* polymorphic genes

The overall population proportions of merozoite gene allelic pairs which were the same and different pre- and post-treatment from the same individual were determined. The parasite pairs were defined using haplotypes consisting of the amino acid codes for genes (*rh1, rh5, eba140, ebl-1, msp1*$_{42}$) with a maximum of six SNPs and in genes (*rh2a, rh2b, rh4, eba175, eba181, ama1, msp3, msp6, mspdbl1, mspdbl2*), containing greater than six SNPs we assessed individual SNPs. The number of haplotype changes per parasite pair, pre- and post-treatment, were determined for each gene. A non-parametric test (McNemar's test) for matched paired data was then done to analyse the proportion of heterologous compared to homologous haplotypes at each locus. Mixed infections were excluded from the analysis. To determine whether SNPs were in LD, all possible pair-wise comparisons between SNPs across all the genes were performed using a pair-wise correlation analysis with a Bonferroni correction for multiple comparisons [36].

Results

Merozoite gene haplotype frequencies in the uncomplicated malaria infections between 2007 and 2008

Most haplotypes were present as stable frequencies over the two-year period. There were two dominant haplotypes for EBL1-QFFFF.VN and QLFLSKS, EBA181-NKSFN and NQSFN and EBA140-VSTK and INTK; the 3D7 reference EBA140 haplotype, INKK, was not observed in this population (see Additional file 2). The 5′ end of the MSP1 42 kDa sequence contained a 3-bp deletion which has not been described before (SNP4414-6),

Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 4 of 11

resulting in three stable and high frequency haplotypes, FKFFDD (3D7 reference), FKFFDN, FFFDD. MSP3 K1 and MSP6 3D7 were the prevalent and stable haplotypes. The MSPDBL1 and MSPDBL2 haplotypes were defined as previously described [40]. Shifts in haplotype frequencies were observed in MSPDBL1 from AHQIARY to ALTAIKY, and for MSPDBL2, AHQAIRY was the dominant haplotype in both years. The variation in the Rh family was characterized by assessing previously defined polymorphic regions (see Additional file 1) [41]. The C-terminal end of Rh1 included two dipeptide (HNQN) repeats. The common dipeptide repeats represented the 3D7 reference sequence, 4HNQN, which shifted to 3HN2QN in 2008 and 4HN2QN remained stable over the two years. Rh2a and Rh4 both contained as the stable and widespread genotype the 3D7 reference haplotypes, KAKQQR and (IHTNENNINN) (EHTNENNINN)(EHTNEKNINN)(EHANEKNIYN) (EHTNENNINY), respectively. In Rh2b, two previously described deletions, a 192-bp deletion (codon 2717) [42] and a 585-bp deletion (codon 2954) [42, 43] were identified. The 156-bp insertion, Q (codon 2795) and a 585-bp deletion (156insQ585del) was dominant and stable, while 156del585del and 192ins585del were only present at high frequency in 2007 and 2008, respectively. The 3D7 reference genotype, 156ins585ins, was not observed in this population. A shift was also observed in Rh5 from N(DYKNVDYKNV)HDY to N(DYKNVDYKNV) YHY. The previously described EBA175 6-bp deletion [5] was identified and distinct fluctuations in the haplotypes from EEKKSISENKKI to EKEKPISEN-KKK between 2007 and 2008 were observed. AMA1 was defined using the amino acids in cluster 1 loop of domain I (codons 196, 197, 199, 200, 201, 204, 206, 207) [44], DQRHFDKY (16 %) and NGRDLNEY (14 %) were the two high frequency haplotypes. Sequences for all 15 genes were submitted to GenBank and are available under the accession codes EBL1:KU525776–KU525801, MSP6:KU525802–KU525809, KU525811, KU525813–KU525817, KU525819, KU525822–KU525844, Rh2b:KU525845–KU525879, Rh5:KU525880–KU525986, Rh2a:KU525988–KU526029, MSPDBL1:KU526030–KU526063, EBA140:KU526064–KU526129, Rh1:KU526130–KU526176, MSPDBL2:KU526177–KU526235, EBA175: KU526236–KU526265, EBA181:KU526266–KU526296, Rh4:KU526297–KU526361, MSP1:KU526362–KU526442, Rh5_10yr_samples:KU526443–KU526857 and KU526859–KU526963, AMA1:KU526964–KU527018, MSP3: KU527019–KU527058.

Genetic diversity pre- and post-treatment

Apart from *ama1, msp1, eba175, mspdbl2, mspdbl1* and *rh2b* SNPs, which were 100 % different parasite pairs,

though the latter two genes included an analysis of three pairs, in the rest of the genes the number of parasite pairs that were the same were <50 %. However, Rh4 (68 %) and Rh5 (70 %) (see Additional file 3) contained a larger number of parasite pairs that were the same. Only Rh5 showed a significant (McNemar's $\chi^2 = 4$, p = 0.046) change in haplotype frequencies with an increased odds (OR 1.3, 95 % CI 1–1.7) of different haplotype parasite pairs compared to the same pairs (Table 1) pre- and post-treatment.

Linkage disequilibrium analysis

A LD analysis was conducted for all genes except *eba181*, which had very few SNPs to analyse. Within gene analysis, all genes contained SNPs in LD except *rh1*, the indels in *msp1, eba175, ebl-1* and *rh2b*, the repeat region of *rh4*, SNPs 439 and 442 in *rh5*, SNPs 553 and 716 in *eba140* and multiple SNPs across the *ama1* gene (see Additional file 4). There are potential recombination hot spots between the two LD blocks in *rh2a* at the 5′ and 3′ ends of the sequenced region, the major LD block towards the 5′ end and a smaller LD block towards the 3′ end of *mspdbl2* and the small LD block at the 5′ end followed by a large central LD block and a minor LD block towards the 3′ end of *mspdbl1*. SNPs in LD between genes were identified (see Additional file 4); *mspdbl2*, the gene with the most SNPs (>200) showed the largest number of between gene LD, with SNPs in LD with six other genes, *eba140, eba175, rh4, ama1, msp3* and *mspdbl1*. SNPs in *mspdbl1, msp3* and *ama1* were in LD with five other genes, *rh2a* and *eba175* with four genes, *msp6* and *eba140* with two genes and *ebl-1, rh2b, msp1, rh4* and *rh1* with only one gene. Of interest, *rh5* SNP 608 appears to be in LD with *eba175* SNP 1434 and *rh5* SNP 442 with *rh2a* SNPs 8459, 8460, 8463, 8464, and 8465. A summary of the between-gene LD SNP analysis is depicted in Fig. 1.

Frequency distribution of the merozoite gene polymorphisms in the uncomplicated malaria infections between 2007 and 2008

A large majority of the SNPs were stably maintained over the two-year period. Eight genes contained some SNPs which showed significant (p < 0.05) temporal fluctuations (heterogeneity), Rh1–QN repeat, MSP3–SNPs 219 and 1029, MSPDBL2–SNPs 580 and 587, EBA175-SNPs 1168, 1750 and 1990, AMA1–SNPs 724, 727, 799 and 922 and MSPDBL1 contained a group of about 85 SNPs which were in LD between SNPs 467 and 855, of which SNPs downstream from SNP524 had a p value of 0.008. There were very few SNPs with a p value <0.01, Rh2b–SNP 8149 (192 bp indel), EBA175–SNP1442 and AMA1–SNP1478. Only the Rh5 SNPs 439 and 442 were highly significant, p = 0.00011 (see Additional file 5) and remained significant after a Bonferroni correction.

Ochola-Oyier *et al. Malar J (2016) 15:261*

Page 5 of 11

Table 1 McNemar's χ² test of haplotype proportions from pre- and post-treatment parasite pairs

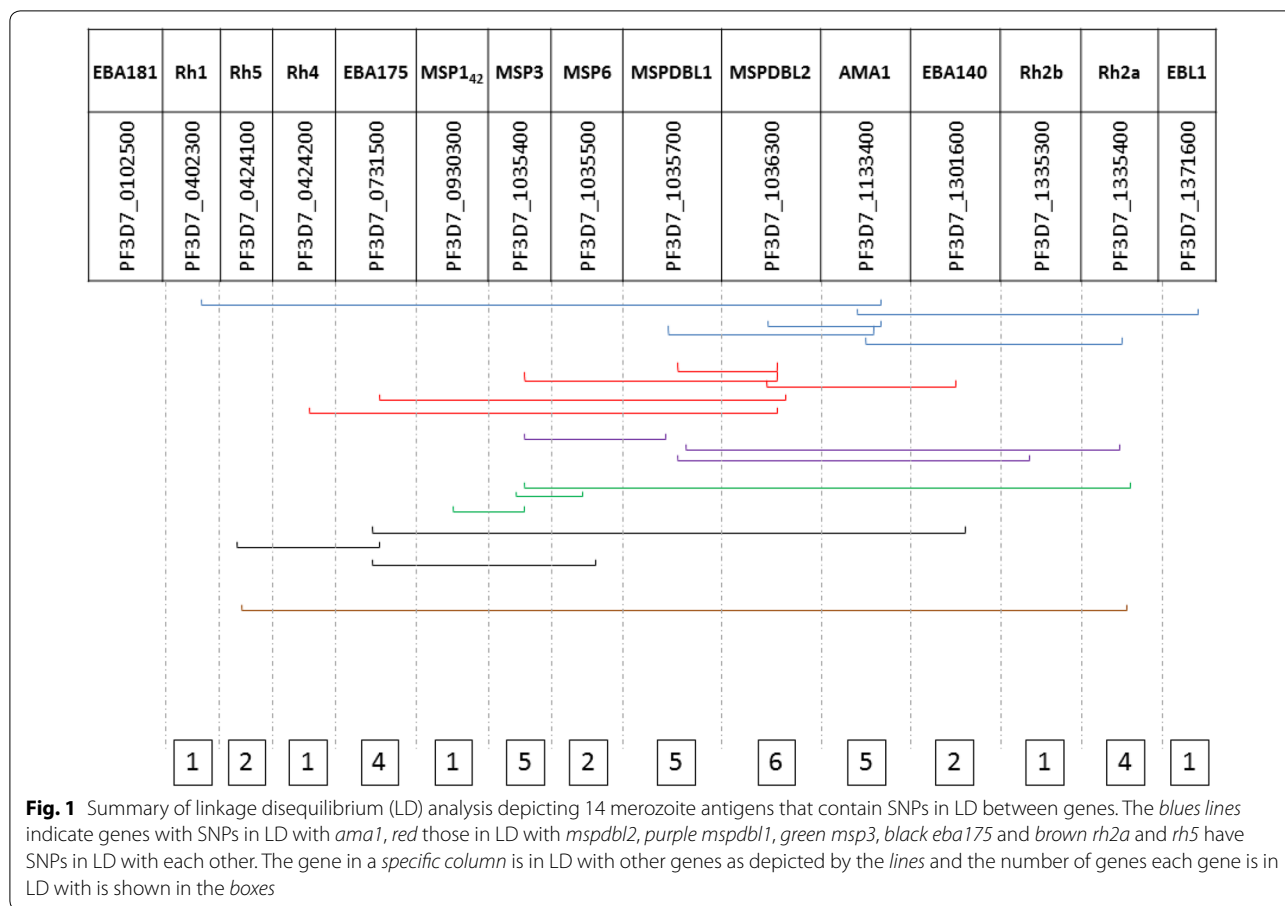| | Pre- and post-treatment | | | |
	Pre	Post	McNemar's χ² (p value)	OR (95 % CI)
AMA1		DQRHFDKY	Other	
	DQRHFDKY	0	0	
	Other	2	8	2 (0.16)	ND
EBA140		VSTK	Other	
	VSTK	5	2	
	Other	3	5	0.2 (0.65)	0.9 (0.5–1.6)
EBA181		Q	K	
	Q	5	4	
	K	2	2	0.67 (0.41)	1.3 (0.7–2.4)
EBA175		EKEKPISENKKK	Other	
	EKEKPISENKKK			
	Other	No data		
EBL-1		QFFFF.VN	Other	
	QFFFF.VN			
	Other	No data		
MSP1		KD	Other	
	KD	7	2	
	Other	5	3	1.29 (0.26)	0.8 (0.5–1.2)
MSP3		K1	3D7	
	K1	4	2	
	3D7	3	0	0.2 (0.65)	0.9 (0.4–1.7)
MSP6		K1	3D7	
	K1	1	2	
	3D7	3	9	0.2 (0.65)	0.8 (0.2–2.7)
MSPDBL1		ALTAIKY	Other	
	ALTAIKY			
	Other	No data		
MSPDBL2		AHQAIRY	Other	
	AHQAIRY	2	6	
	Other	7	3	0.08 (0.78)	0.9 (0.4–2.0)
Rh1		4HN2QN	Other	
	4HN2QN	1	2	
	Other	1	3	0.33 (0.56)	1.5 (0.4–6.0)
Rh2b		156insQ585del	Other	
	156insQ585del			
	Other	No data		
Rh5		HD	YH	
	HD	13	4	
	YH	0	3	*4 (0.046)*	1.3 (1.0–1.7)

italic value indicates a significant increase in the odds of the Rh5 haplotypes changing compared to staying the same between parasite pairs before and after treatment

Other represents genes with more than 2 high frequency haplotypes, the dominant allele was compared to all other alleles in the population, *ND* not determined, *No data* insufficient pairs for the analysis <4

Distribution of PfRh5 polymorphisms in all three malaria infection populations

Rh5 codons 48, 197, 203, and 204 were identified in the 11 cultured isolates (see Additional file 6), identifying the previously described codons by Hayton and colleagues [45]. Only codons 197 and 203 were observed in the Kilifi population, however, codon 197 was a singleton SNP in the asymptomatic infections. In the uncomplicated malaria infections, two other rare variants, three DYKNV repeats (codon 132) in two isolates instead of

Ochola-Oyier *et al. Malar J (2016) 15:261*

Page 6 of 11

Fig. 1 Summary of linkage disequilibrium (LD) analysis depicting 14 merozoite antigens that contain SNPs in LD between genes. The *blues lines* indicate genes with SNPs in LD with *ama1*, *red* those in LD with *mspdbl2*, *purple mspdbl1*, *green msp3*, *black eba175* and *brown rh2a* and *rh5* have SNPs in LD with each other. The gene in a *specific column* is in LD with other genes as depicted by the *lines* and the number of genes each gene is in LD with is shown in the *boxes*

the common 3D7 haplotype of DYKNVDYKNV (two repeats) and a singleton at SNP262 (codon 88) were observed. Furthermore, two additional high frequency SNPs in codons 147 and 148 were identified that were not observed in the laboratory isolates. Codons 147 and 148 were previously described in the MalariaGen data [46, 47]. These codons resulted in two haplotypes, YH and the non-reference 3D7 haplotype HD. A rare combination of YD was also observed in the complicated malaria infections. There were two high frequency, stable haplotypes in the asymptomatic infections YHSY and HDSY with YHSY being dominant over the two-year period from 2008 to 2009. After multiple attempts, we were unable to obtain PCR amplicons from the 2007 samples. The YHY and HDY haplotypes were prevalent and YHY was the dominant and stable haplotype in the complicated malaria infections from 2007 to 2009.

PfRh5 analysis in all three malaria infection populations
The uncomplicated, asymptomatic and complicated malaria infections were compared between the years 2007–2009. Though we excluded mixed infections over this time period, we noticed that the number of mixed infections in the uncomplicated and complicated malaria infections were 6 %

(n = 44) and 5 % (n = 150), respectively, about fourfold lower than asymptomatic malaria infections (24 %, n = 54). The analysis focused on the two dominant codons 147 and 148, which showed a highly significant shift in haplotype frequencies between 2007 and 2008 (p = 0.00011), from HD to YH in the uncomplicated malaria infections (Fig. 2). In contrast, there was a dominance of the YH haplotype over the two-year period in the asymptomatic infections (p = 0.19) and over the three-year period (2007–2009) in the complicated malaria infections (p = 0.16). Interestingly, in the uncomplicated malaria infections there was a significant difference between the mean ages 2.5 and 3.5 in 2007 and 2008, respectively (t test unequal, p = 0.0007). However, there was no significant difference in the asymptomatic infections between 2008 and 2009 (mean age 6.4 and 6.6, respectively, t test unequal, p = 0.51) and in the complicated malaria infections (mean age three in all years, t test unequal, p = 0.96). The significant shift in haplotypes in the uncomplicated malaria infections may be due to the significant mean age difference between the years. However, a logistic regression analysis suggested age had no effect on the infecting Rh5 haplotypes, but the chance of being infected with HD parasites decreased with time from 2007 to 2009 and the probability of HD infections increased moving from uncomplicated to
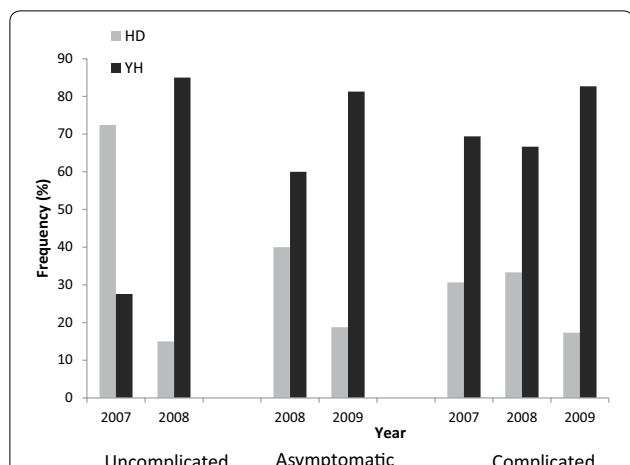
Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 7 of 11

Fig. 2 The distribution of Rh5 high frequency haplotypes, YH and HD, in all three malaria infections. The uncomplicated malaria infections from 2007 to 2008 showed a significant shift in haplotype frequency from HD to YH in 2008 (p = 0.00011). However, in the asymptomatic malaria infections between 2008 and 2009, YH was the dominant haplotype in both years and YH was also dominant in the complicated malaria infections from 2007 to 2009

asymptomatic malaria infections (Table 2). Also, there were no significant haplotype frequency differences between 2008 and 2009 in uncomplicated and asymptomatic infections. The haplotypes from 2007 to 2009 were plotted on to the KHDSS map (Fig. 3) to examine geographical clustering of haplotypes. Altogether SaTScan did not pick out any significant haplotype (YH or HD) clusters. However, a within-year stratification picked out a significant cluster in the north of the district with more YH infecting parasites when compared to all other possible clusters in 2007 (relative risk = 2, p = 0.0024) (Fig. 3b).

Table 2 Logistic regression analysis of Rh5 YH and HD alleles in all three malaria infections between 2007 and 2009

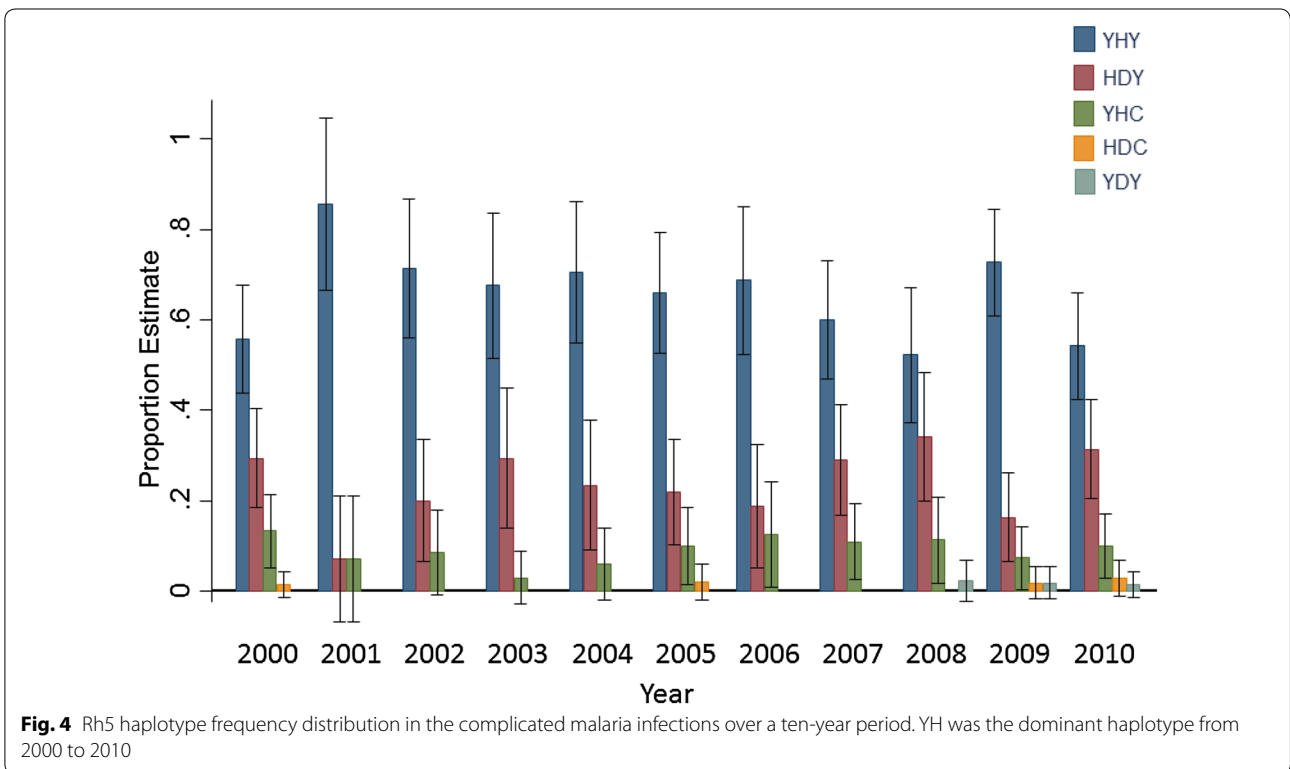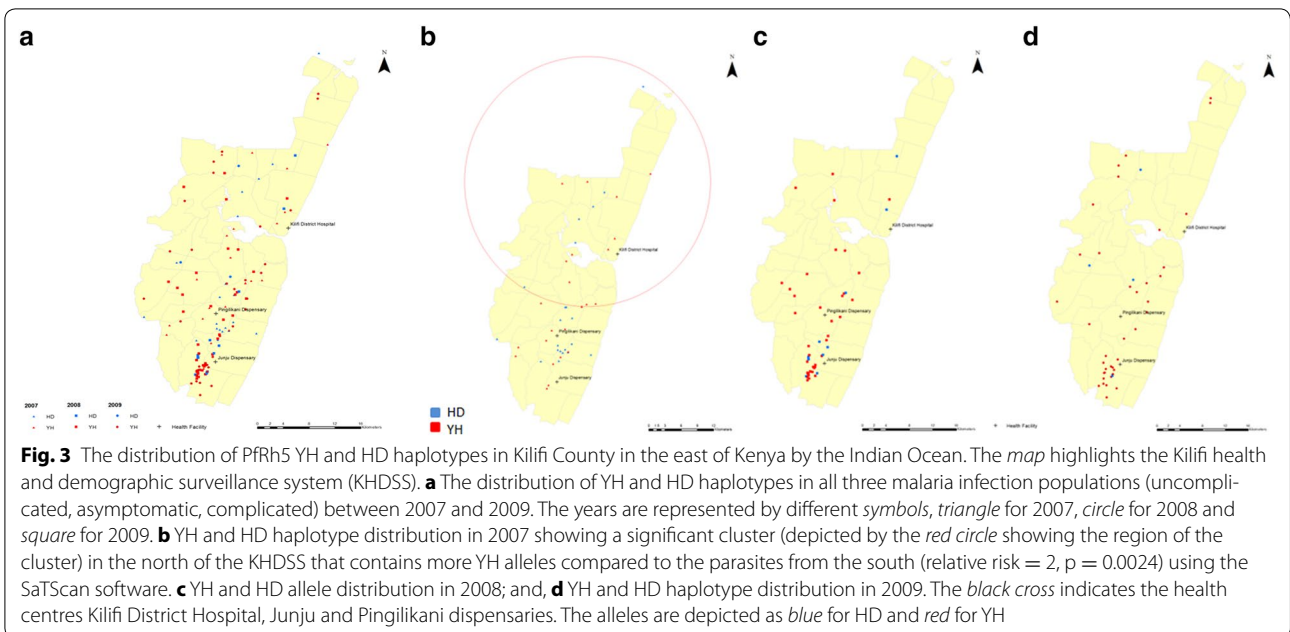|	Haplotype		Odds ratio (95 % CI)	p value
	YH	**HD**		
Age			1.1 (0.89–1.39)	0.349
Year				
2007	42 (54)	36 (46)	1	
2008	60 (68)	27 (31)	0.11 (0.03–0.37)	<0.001
2009	51 (80)	12 (19)	0.05 (0.01–0.29)	0.001
Population				
Severe	105 (73)	38 (27)	1	
Uncomplicated	25 (51)	24 (49)	4.66 (1.75–12.43)	0.002
Asymptomatic	28 (68)	13 (32)	7.43 (1.57–35.16)	0.011

Stable Rh5 polymorphisms over a ten-year period of declining malaria transmission

The stable haplotype frequencies of a dominant YHY haplotype, followed by HDY and YHC haplotypes observed between 2007 and 2009, extended over a ten-year period from 2000 to 2010. On occasion rare variants, HDC, were observed in 2000, 2005, 2009 and 2010 and also YDY in 2008, 2009 and 2010 (Fig. 4). There is evidence of a lack of homogeneity (p = 0.0001) and a linear trend (p < 0.001) of an increase in the odds of individuals admitted to KDH coming from the south of the district and from an older age group, >5 years (p = 0.02 and 0.0016, respectively) over the ten-year period. Consequently, the increase in the number of cases from the south is associated with increase in the age of the children admitted to KDH. This population of individuals with complicated malaria had at least one of the severe malaria syndromes of deep breathing, anaemia (haemoglobin <5 g/dl) and cerebral malaria (Blantyre coma score <3) [37]. Over the ten-year period there is evidence of a linear trend of a decrease in the odds of cerebral malaria (p = 0.0004), there is a lack of homogeneity in the odds of population exhibiting deep breathing (p = 0.002), and there were no differences in the individuals with anaemia. However, when location and age were adjusted for in the analysis there were no trends observed in any of the severe malaria syndromes.

Discussion

The stable and intermediate allele frequencies of all genes except Rh2a and Rh4 (which had only one dominant allele from the region sequenced) are suggestive of these polymorphisms being maintained by stable balancing selection. In particular, the antigens which have previously shown evidence of being under balancing selection in cross-sectional population studies such as AMA1 [11, 48], MSPDBL1 [7], MSPDBL2 [8], and MSP3 [6]. The stable alleles are being maintained by a constant selection pressure on these antigens obtained from parasites, resulting in uncomplicated malaria in children under the age of five over a two-year period.

Though it does not seem to matter if an individual is infected with the YH or HD alleles, what was striking was the shift in allele frequencies observed only in Rh5. This suggests that Rh5 may be under different selection pressures over the two-year period examined, resulting in fluctuating allele frequencies. At an individual level, what appeared to be frequency-dependent selection was observed in the pre- and post-treatment infections in only Rh5. All the changes were from pre-treatment HD infections to YH post-treatment infections. The Rh5 codons 147 and 148 showed a greater likelihood of changing compared to staying the same. In vaccine-induced selection, subsequent infections post-vaccine

Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 8 of 11

Fig. 3 The distribution of PfRh5 YH and HD haplotypes in Kilifi County in the east of Kenya by the Indian Ocean. The *map* highlights the Kilifi health and demographic surveillance system (KHDSS). **a** The distribution of YH and HD haplotypes in all three malaria infection populations (uncomplicated, asymptomatic, complicated) between 2007 and 2009. The years are represented by different *symbols, triangle* for 2007, *circle* for 2008 and *square* for 2009. **b** YH and HD haplotype distribution in 2007 showing a significant cluster (depicted by the *red circle* showing the region of the cluster) in the north of the KHDSS that contains more YH alleles compared to the parasites from the south (relative risk = 2, p = 0.0024) using the SaTScan software. **c** YH and HD allele distribution in 2008; and, **d** YH and HD haplotype distribution in 2009. The *black cross* indicates the health centres Kilifi District Hospital, Junju and Pingilikani dispensaries. The alleles are depicted as *blue* for HD and *red* for YH

Fig. 4 Rh5 haplotype frequency distribution in the complicated malaria infections over a ten-year period. YH was the dominant haplotype from 2000 to 2010

administration were by parasites of the non-vaccine allele [17, 19]. Thus, similar to the scenario in this study, Rh5 appears to show the same effect of post-treatment infections with predominantly an alternative allele after drug clearance. This variation may be reflected in the Rh5 allele frequencies at a population level, providing the first possible explanation for the observed heterogeneity in Rh5 allele frequencies in the uncomplicated malaria

Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 9 of 11

infections. The temporal variation in selection of Rh5 variants may be due to frequency-dependent selection, such that in this population of children with low levels of acquired immunity, the rare YH variants in 2007 have an advantage over the common HD variant and thus YH rises to high frequency in 2008. Furthermore, in the asymptomatic and complicated malaria infections the Rh5 alleles were temporally maintained at stable frequencies, also suggesting frequency-dependent selection. Therefore, at an individual level, allele frequencies may change between infections but at a population level they even out into stable frequencies over time. Since all individuals will eventually encounter all Rh5 variants, the initially rare variant will rise in frequency until equilibrium is attained between all circulating variants.

The second, plausible reason is provided by evidence of differences in clinical outcome, such that children who presented with uncomplicated or asymptomatic malaria were more likely to be infected with HD allele parasites than those with complicated malaria infections. The heterogeneity in selection pressures may therefore be due to differences in the levels of acquired immunity in the individuals with mild and complicated malaria. The variation in the spectrum of disease from asymptomatic to complicated or severe malaria has previously been described and it was primarily attributed to prior exposure and acquired immunity [49].

The third explanation may be as a result of the evidence of spatial patterns of selection in 2007 of a difference in allele frequencies between the north and south of Kilifi County, with more HD infecting parasites in the south compared to the north. These allele frequency differences may be a result of the differences in malaria transmission, which is low in the north and moderate to high in the south [50], and this would alter the levels of immunity of individuals in both regions. A previous study in Senegal, showed higher immune responses to recombinant Rh5 protein in a high malaria transmission area compared to a lower transmission region [51]. Therefore, it is possible that there is selection for more variants in the higher transmission areas and we observed a higher frequency of the HD variant in the south of Kilifi. However, perhaps the missing data from the asymptomatic malaria infections in 2007, also from the south, which was expected from the observed stable frequencies in 2008 and 2009 to be predominantly YH, may have resulted in a sampling bias, resulting in the observed skew in allele frequencies. Nevertheless, the data do provide evidence that fluctuating selection may maintain the variation in Rh5 polymorphisms in the uncomplicated malaria population.

The fourth and more extreme hypothesis is that the fluctuation in Rh5 allele frequencies may be a result of a parasite population bottleneck effect, which greatly reduced the parasite population harbouring the HD allele, perhaps by some variation in transmission intensity that altered the allele frequency over time. However, it may simply be a case of genetic drift in a small localized population that results in the random fluctuations in allele frequencies.

The stable Rh5 allele frequencies in the complicated malaria infections were maintained over a ten-year period on a backdrop of changes in the host population demography. An increase over time of more children coming from the south who are older (>5 years) and a decrease in cerebral malaria cases over time. Suggesting therefore, that even a change over time in the infected host population dynamics does not alter the frequency of Rh5 polymorphisms, perhaps due to the fact that the constant selection pressure is the immune pressure from these individuals. Thus, even though there is a shift to an older age group at a population level, all the individuals still encounter all Rh5 variants, maintaining a stable allele frequency over time.

Overall, the stable and fluctuating Rh5 polymorphism frequencies appear to be the result of the host's ability to mount an immune response to prevent infection, and it may be the major driving force maintaining the Rh5 polymorphisms in codons 147 and 148. Furthermore, the Rh5 laboratory strain sequences were similar to each other except for the Palo Alto strain, however they differ greatly from the field isolates which contain SNPs in codons 147 and 148. These SNPs were not observed in the laboratory strains.

Recently, the partial Rh5 protein structure has been described [52], which suggests that codons 147 and 148 fall outside the basigin-binding region. They may simply be involved in evading host immune responses and redirecting the immune system away from the important functional binding regions of the protein. Previous studies have demonstrated that Rh5 may be a poorly immunogenic antigen or the responses may be short lived [53, 54] and immune responses are likely to be allele-transcending [47, 54]. Yet, this study potentially suggests that immunity may maintain the high frequency of the polymorphisms in codons 147 and 148. Although the observed allele changes in Rh5 are at a genetic level, the functional impact of the polymorphisms remains to be determined, since cross-strain neutralizing antibodies have previously been demonstrated in vitro [53].

The LD analysis examined the impact of the potential interactions between the 15 merozoite antigens, identifying polymorphic combinations between different genes that may be favoured and maintained by natural selection, potentially due to epistatic selection [55, 56]. Rh5 appears to contain SNPs in LD with SNPs in EBA175 and Rh2a. This interaction between the polymorphisms of different merozoite loci may be due to a functional

Ochola-Oyier *et al. Malar J (2016) 15:261*

Page 10 of 11

relationship, providing a fitness advantage such as evading the immune system or invading erythrocytes and emphasizing the overall complex nature of merozoite invasion [57] or immune evasion.

Conclusions

The shift in allele frequencies at a population level and in individual infections, pre- and post-treatment, seemed to only have an impact on the pattern of Rh5 genetic diversity and codons 147 and 148 may be genetic markers of immune pressure—not to level that results in allele-specific immune responses. However, perhaps an increase in sample size, and an extension of the uncomplicated and asymptomatic malaria parasite population analyses back to 2005 would identify more antigens operating under the same principle of both fluctuating and stable allele frequencies over time. Additionally, the authors note the lack of microsatellite analyses as a limitation, since it would help to confirm unique genotypes and ascertain that the allele frequency changes were limited to the Rh5 locus. Nonetheless, this study provides crucial data describing potential selection forces maintaining Rh5 polymorphisms in the population on a temporal context. It would also be important to determine whether the Rh5 variants elicit an immune response in this genotyped population, to determine the exact mechanism maintaining the polymorphisms in this population.

Additional files

Additional file 1. PCR conditions and primer sequences for the gene regions analysed in the 15 merozoite genes.

Additional file 2. Merozoite gene allele frequencies in the uncomplicated malaria infections between 2007 and 2008.

Additional file 3. The proportion of same and different genotype parasite pairs *in vivo* and *in vitro*.

Additional file 4. Linkage disequilibrium (LD) analysis of all high frequency (>5%) SNPs in all 15 merozoite genes, except *eba181* due to limited SNP data. The analysis was performed by conducting all possible pair-wise comparisons between all the SNPs across 14 merozoite genes using a pair-wise correlation analysis and a Bonferroni correction for multiple comparisons. The red boxes represent significant correlations between SNPs and thus SNPs in LD, orange represents correlations between SNPs that are not significant, white indicates the SNPs that are not correlated and therefore they are not in LD and grey represents cases where the SNP data is missing in one or both of the paired SNPs.

Additional file 5. SNPs and repeat region frequencies within the 15 merozoite genes from the uncomplicated malaria infections in 2007 and 2008 to identify significant SNPs.

Additional file 6. Rh5 SNPs in 11 geographically distinct laboratory isolates.

Authors' contributions
LIOO, KKT and KM conceived and designed the study; LIOO, JoO and NW conducted the research; LIOO, JoO, NW, JaO, GF, and PB assisted with the

data analysis; LIOO wrote and revised the manuscript. All authors read and approved the final manuscript.

Author details
[1] KEMRI-Wellcome Trust Collaborative Programme, P.O. Box 230, Kilifi 80108, Kenya. [2] Department of Immunology and Infection, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK.

References

1. Garrigan D, Hedrick PW. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution. 2003;57:1707–22.
2. Tennessen JA, Blouin MS. Balancing selection at a frog antimicrobial peptide locus: fluctuating immune effector alleles? Mol Biol Evol. 2008;25:2669–80.
3. Hughes AL. Positive selection and interallelic recombination at the merozoite surface antigen-1 (MSA-1) locus of *Plasmodium falciparum*. Mol Biol Evol. 1992;9:381–93.
4. Polley SD, Conway DJ. Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene. Genetics. 2001;158:1505–12.
5. Verra F, Chokejindachai W, Weedall GD, Polley SD, Mwangi TW, Marsh K, et al. Contrasting signatures of selection on the *Plasmodium falciparum* erythrocyte binding antigen gene family. Mol Biochem Parasitol. 2006;149:182–90.
6. Polley SD, Tetteh KKA, Lloyd JM, Akpogheneta OJ, Greenwood BM, Bojang KA, et al. *Plasmodium falciparum* merozoite surface protein 3 is a target of allele-specific immunity and alleles are maintained by natural selection. J Infect Dis. 2007;195:279–87.
7. Tetteh KKA, Stewart LB, Ochola LI, Amambua-Ngwa A, Thomas AW, Marsh K, et al. Prospective identification of malaria parasite genes under balancing selection. PLoS ONE. 2009;4:e5568.
8. Ochola LI, Tetteh KKA, Stewart LB, Riitho V, Marsh K, Conway DJ. Allele frequency based and polymorphism-versus-divergence indices of balancing selection in a new filtered set of polymorphic genes in *Plasmodium falciparum*. Mol Biol Evol. 2010;27:2344–51.
9. Rayner JC, Tran TM, Corredor V, Huber CS, Barnwell JW, Galinski MR. Dramatic difference in diversity between *Plasmodium falciparum* and *Plasmodium vivax* reticulocyte binding-like genes. Am J Trop Med Hyg. 2005;72:666–74.
10. Reiling L, Richards JS, Fowkes FJI, Barry AE, Triglia T, Chokejindachai W, et al. Evidence that the erythrocyte invasion ligand PfRh2 is a target of protective immunity against *Plasmodium falciparum* malaria. J Immunol. 2010;185:6157–67.
11. Osier FH, Weedall GD, Verra F, Murungi L, Tetteh KK, Bull P, et al. Allelic diversity and naturally acquired allele-specific antibody responses to *Plasmodium falciparum* apical membrane antigen 1 in Kenya. Infect Immun. 2010;78:4625–33.
12. Hodder AN, Crewther PE, Anders RF. Specificity of the protective antibody response to apical membrane antigen 1. Infect Immun. 2001;69:3286–94.
13. Tham WH, Wilson DW, Reiling L, Chen L, Beeson JG, Cowman AF. Antibodies to reticulocyte binding protein-like homologue 4 inhibit invasion of *Plasmodium falciparum* into human erythrocytes. Infect Immun. 2009;77:2427–35.
14. Crewther PE, Matthew ML, Flegg RH, Anders RF. Protective immune responses to apical membrane antigen 1 of *Plasmodium chabaudi* involve recognition of strain-specific epitopes. Infect Immun. 1996;64:3310–7.

Ochola-Oyier *et al. Malar J* (2016) 15:261

Page 11 of 11

15. Lyon JA, Angov E, Fay MP, Sullivan JS, Girourd AS, Robinson SJ, et al. Protection induced by *Plasmodium falciparum* MSP1 (42) is strain-specific, antigen and adjuvant dependent, and correlates with antibody responses. PLoS ONE. 2008;3:e2830.
16. Conway DJ, Greenwood BM, McBride JS. Longitudinal study of *Plasmodium falciparum* polymorphic antigens in a malaria-endemic population. Infect Immun. 1992;60:1122–7.
17. Genton B, Betuela I, Felger I, Al-Yaman F, Anders RF, Saul A, et al. A recombinant blood-stage malaria vaccine reduces *Plasmodium falciparum* density and exerts selective pressure on parasite populations in a phase 1-2b trial in Papua New Guinea. J Infect Dis. 2002;185:820–7.
18. Healer J, Murphy V, Hodder AN, Masciantonio R, Gemmill AW, Anders RF, et al. Allelic polymorphisms in apical membrane antigen-1 are responsible for evasion of antibody-mediated inhibition in *Plasmodium falciparum*. Mol Microbiol. 2004;52:159–68.
19. Ouattara A, Takala-Harrison S, Thera MA, Coulibaly D, Niangaly A, Saye R, et al. Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: vaccine development implications. J Infect Dis. 2013;207:511–9.
20. Takahata N, Nei M. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics. 1990;24:967–78.
21. Babiker HA, Satti G, Walliker D. Genetic changes in the population of *Plasmodium falciparum* in a Sudanese village over a three-year period. Am J Trop Med Hyg. 1995;53:7–15.
22. Ferreira MU, Liu Q, Zhou M, Kimura M, Kaneko O, Van Thien H, et al. Stable patterns of allelic diversity at the merozoite surface protein-1 locus of *Plasmodium falciparum* in clinical isolates from southern Vietnam. J Eukaryot Microbiol. 1998;45:131–6.
23. Sallenave-Sales S, Ferreira-da-Cruz MF, Faria CP, Cerruti C Jr, Daniel-Ribeiro CT, Zalis MG. *Plasmodium falciparum*: limited genetic diversity of MSP-2 in isolates circulating in Brazilian endemic areas. Exp Parasitol. 2003;103:127–35.
24. Orjuela-Sánchez P, Da Silva-Nunes M, Da Silva NS, Scopel KKG, Gonçalves RM, Malafronte RS, et al. Population dynamics of genetically diverse *Plasmodium falciparum* lineages: community-based prospective study in rural Amazonia. Parasitology. 2009;136:1097–105.
25. Xangsayarath P, Kaewthamasorn M, Yahata K, Nakazawa S, Sattabongkot J, Udomsangpetch R, et al. Positive diversifying selection on the *Plasmodium falciparum* surf4.1 gene in Thailand. Trop. Med Health. 2012;40:79–89.
26. Kaewthamasorn M, Yahata K, Alexandre JSF, Xangsayarath P, Nakazawa S, Torii M, et al. Stable allele frequency distribution of the polymorphic region of SURFIN(4.2) in *Plasmodium falciparum* isolates from Thailand. Parasitol Int. 2012;61:317–23.
27. Pratt-Riccio LR, Perce-da-Silva Dde S, Lima-Junior JC, Theisen M, Santos F, Daniel-Ribeiro CT, de Oliveira-Ferreira J, Banic DM. Genetic polymorphisms in the glutamate-rich protein of *Plasmodium falciparum* field isolates from a malaria-endemic area of Brazil. Mem Inst Oswaldo Cruz. 2013;108:523–8.
28. Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. Nature. 1999;400:667–71.
29. Tiffin P, Hacker R, Gaut BS. Population genetic evidence for rapid changes in intraspecific diversity and allelic cycling of a specialist defense gene in Zea. Genetics. 2004;168:425–34.
30. Forsyth KP, Anders RF, Cattani JA, Alpers MP. Small area variation in prevalence of an S-antigen serotype of *Plasmodium falciparum* in villages of Madang, Papua New Guinea. Am J Trop Med Hyg. 1989;40:344–50.
31. Eisen DP, Wang L, Jouin H, Murhandarwati EEH, Black CG, Mercereau-Puijalon O, et al. Antibodies elicited in adults by a primary *Plasmodium falciparum* blood-stage infection recognize different epitopes compared with immune individuals. Malar J. 2007;6:86.
32. Borrmann S, Sasi P, Mwai L, Bashraheil M, Abdallah A, Muriithi S, et al. Declining responsiveness of *Plasmodium falciparum* infections to artemisinin-based combination treatments on the Kenyan coast. PLoS ONE. 2011;6:e26005.
33. O'Meara WP, Bejon P, Mwangi TW, Okiro EA, Peshu N, Snow RW, Newton CR, Marsh K. Effect of a fall in malaria transmission on morbidity and mortality in Kilifi, Kenya. Lancet. 2008;372:1555–62.
34. Scott JAG, Bauni E, Moisi JC, Ojal J, Gatakaa H, Nyundo C, et al. Profile: the Kilifi health and demographic surveillance system (KHDSS). Int J Epidemiol. 2012;41:650–7.
35. Bereczky S, Mårtensson A, Gil JP, Färnert A. Rapid DNA extraction from archive blood spots on filter paper for genotyping of *Plasmodium falciparum*. Am J Trop Med Hyg. 2005;72:249–51.
36. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995;310:170.
37. Marsh K, Forster D, Waruiru C, Mwangi I, Winstanley M, Marsh V, et al. Indicators of life-threatening malaria in African children. N Engl J Med. 1995;332:1399–404.
38. Kulldorff M. SaTScan—Software for the spatial, temporal, and space-time scan statistics. Boston: Harvard Medical School and Harvard Pilgrim Health Care; 2010.
39. Bejon P, Williams TN, Liljander A, Noor AM, Wambua J, Ogada E, et al. Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. PLoS Med. 2010;7:e1000304.
40. Ochola-Oyier LI, Okombo J, Mwai L, Kiara SM, Pole L, Tetteh KK, et al. The MSPDBL2 codon 591 polymorphism is associated with lumefantrine in vitro drug responses in *Plasmodium falciparum* isolates from Kilifi, Kenya. Antimicrob Agents Chemother. 2015;59:1770–5.
41. Lopez-Perez M, Villasis E, Machado RL, Póvoa MM, Vinetz JM, Blair S, et al. *Plasmodium falciparum* field isolates from South America use an atypical red blood cell invasion pathway associated with invasion ligand polymorphisms. PLoS ONE. 2012;7:e47913.
42. Jennings CV, Ahouidi AD, Zilversmit M, Bei AK, Rayner J, Sarr O, et al. Molecular analysis of erythrocyte invasion in *Plasmodium falciparum* isolates from Senegal. Infect Immun. 2007;75:3531–8.
43. Ahouidi AD, Bei AK, Neafsey DE, Sarr O, Volkman S, Milner D, et al. Population genetic analysis of large sequence polymorphisms in *Plasmodium falciparum* blood-stage antigens. Infect Genet Evol. 2010;10:200–6.
44. Bai T, Becker M, Gupta A, Strike P, Murphy VJ, Anders RF, et al. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. Proc Natl Acad Sci USA. 2005;102:12736–41.
45. Hayton K, Gaur D, Liu A, Takahashi J, Henschen B, Singh S, et al. Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of *Plasmodium falciparum* invasion. Cell Host Microbe. 2008;4:40–51.
46. Manske M, Miotto O, Campino S, Auburn S, Almagro-Garcia J, Maslen G, et al. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. Nature. 2012;487:375.
47. Bustamante LY, Bartholdson SJ, Crosnier C, Campos MG, Wanaguru M, Nguon C, et al. A full-length recombinant *Plasmodium falciparum* PfRH5 protein induces inhibitory antibodies that are effective across common PfRH5 genetic variants. Vaccine. 2013;31:373–9.
48. Polley SD, Chokejindachai W, Conway DJ. Allele frequency-based analyses robustly map sequence sites under balancing selection in a malaria vaccine candidate antigen. Genetics. 2003;165:555–61.
49. Mackinnon MJ, Read AF. Virulence in malaria: an evolutionary viewpoint. Philos Trans R Soc Lond B Biol Sci. 2004;359:965–86.
50. Mwangi TW, Ross A, Snow RW, Marsh K. Case definitions of clinical malaria under different transmission conditions in Kilifi District, Kenya. J Infect Dis. 2005;191:1932–9.
51. Patel SD, Ahouidi AD, Bei AK, Dieye TN, Mboup S, Harrison SC, et al. *Plasmodium falciparum* merozoite surface antigen, PfRH5, elicits detectable levels of invasion-inhibiting antibodies in humans. J Infect Dis. 2013;208:1679–81.
52. Wright KE, Hjerrild KA, Bartlett J, Douglas AD, Jin J, Brown RE, et al. Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies. Nature. 2014;515:427–30.
53. Douglas AD, Williams AR, Illingworth JJ, Kamuyu G, Biswas S, Goodman AL, et al. The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody. Nat Commun. 2011;2:601.
54. Tran TM, Ongoiba A, Coursen J, Crosnier C, Diouf A, Huang CY, et al. Naturally acquired antibodies specific for *Plasmodium falciparum* reticulocyte-binding protein homologue 5 inhibit parasite growth and predict protection from malaria. J Infect Dis. 2014;209:789–98.
55. Smit-McBride Z, Moya A, Ayala FJ. Linkage disequilibrium in natural and experimental populations of *Drosophila melanogaster*. Genetics. 1998;120:1043–51.
56. da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE. Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. Genetics. 2010;85:293–303.
57. Baum J, Maier AG, Good RT, Simpson KM, Cowman AF. Invasion by *P. falciparum* merozoites suggests a hierarchy of molecular interactions. PLoS Pathog. 2005;1:e37.