Multiple genetic origins of histidine-rich protein 2 gene deletion in *Plasmodium falciparum* parasites from Peru

Sheila Akinyi1, Tonya Hayden1, Dionicia Gamboa2,3, Katherine Torres2, Jorge Bendezu2, Joseph F. Abdallah1, Sean M. Griffith1, Wilmer Marquino Quezada4, Nancy Arrospide4, Alexandre Macedo De Oliveira1, Carmen Lucas5, Alan J. Magill5, David J. Bacon5, John W. Barnwell1 & Venkatachalam Udhayakumar1

**Malaria Branch, Division of Parasitic Diseases and Malaria, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA 30333, 2Instituto de Medicina Tropical “Alexander von Humboldt” (IMTAvH), Universidad Peruana Cayetano Heredia, AP 4314, Lima 100, Peru, 3Departamento de Ciencias Celulares y Moleculares, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Av. Honorio Delgado 430 SMP Lima 31, Lima, Peru, 4Instituto Nacional de Salud, Lima, Peru, 5Department of Parasitology, U.S. Naval Medical Research Unit No. 6 (NAMRU-6), Av. Venezuela Cuadra 36 s/n, Callao 2, Peru.

**OPEN**

The majority of malaria rapid diagnostic tests (RDTs) detect *Plasmodium falciparum* histidine-rich protein 2 (PFHRP2), encoded by the *pfhrp2* gene. Recently, *P. falciparum* isolates from Peru were found to lack *pfhrp2* leading to false-negative RDT results. We hypothesized that *pfhrp2*-deleted parasites in Peru derived from a single genetic event. We evaluated the parasite population structure and *pfhrp2* haplotype of samples collected between 1998 and 2005 using seven neutral and seven chromosome 8 microsatellite markers, respectively. Five distinct *pfhrp2* haplotypes, corresponding to five neutral microsatellite-based clonal lineages, were detected in 1998-2001; *pfhrp2* deletions occurred within four haplotypes. In 2003-2005, outcrossing among the parasite lineages resulted in eight population clusters that inherited the five *pfhrp2* haplotypes seen previously and a new haplotype; *pfhrp2* deletions occurred within four of these haplotypes. These findings indicate that the genetic origin of *pfhrp2* deletion in Peru was not a single event, but likely occurred multiple times.

Malaria rapid diagnostic tests (RDTs) are widely utilized by malaria control programs as part of case management because of their ease of use. Malaria RDTs are lateral-flow immunochromatographic devices that detect specific *Plasmodium* parasite antigens in blood samples1. At least three target antigens are captured by various commercially available malaria RDTs: histidine-rich protein 2 (HRP2), lactate dehydrogenase (LDH) and aldolase. Moreover, while most of the LDH and aldolase-based tests are pan-specific, the HRP2-based RDTs are specific for the unique *Plasmodium falciparum* protein (PFHRP2).

The *pfhrp2* gene (PlasmoDB gene ID: PF3D7_0831800) is 1063 bp and located subtelomerically on chromosome 82–4. The gene is immediately flanked upstream by a *Plasmodium* exported protein of unknown function (pseudogene), PF3D7_0831900, and downstream by a putative heat shock protein 70 gene, PF3D7_0831700 (Figure 1). Although the deletion of *pfhrp2* in laboratory-adapted *P. falciparum* strains has previously been documented, this phenomenon had not been observed in natural *P. falciparum* populations until 2010 when *P. falciparum* isolates collected from the Peruvian Amazon were shown to be *pfhrp2*-negative5–7. Even more recently, *pfhrp2*-negative parasite isolates have been detected in Brazil, Mali, Senegal and India8–11. The prevalence of this phenotype suggested that the *pfhrp2*-negative parasites were viable in vivo and capable of causing malaria. Extensive characterization of four *pfhrp2*-negative *P. falciparum* clinical isolates from Peru by whole genome microarray analysis revealed an approximately 20 kb deletion of the genome showing that the deletion is not restricted to *pfhrp2* itself but extends to a number of neighboring genes12.

The *P. falciparum* parasite population structure in the Peruvian Amazon has previously been characterized as clonal13,14. Neutral microsatellite profiles of clinical samples collected between 1998 and 2001 revealed that the *P. falciparum* parasites in Peru consisted of at least five distinct clonal lineages. These clonal lineages, designated as A, B, C, D and E, each displayed a distinct drug resistance and neutral microsatellite-based genetic profile14.
The objective of our study was to characterize the origin and possible reasons for the observed temporal and geographical expansion of \( pfhrp2 \)-negative \( P. falciparum \) parasites in the Peruvian Amazon, where the highest levels of deletions have been documented. We hypothesized that evidence of hitchhiking would be revealed once we analyzed the molecular signatures around Peruvian \( pfhrp2 \) deletions if it evolved from a rare genetic event(s) and spread rapidly under selective pressure. To test our hypothesis, we utilized PCR-based amplification methods and microsatellite analysis to detect deletion of \( pfhrp2 \) and its flanking genes in historical samples (malaria re-emergence peaked in the late 1990s in Peru), then ascertained whether the deletion genotype was restricted to certain clonal lineages. We then examined the pattern of deletion around the \( pfhrp2 \) gene.

**Results**

**Comparison of population structure in parasites collected in 1998–2001 and 2003–2005.** A total of 188 samples collected in Peru between 1998 and 2005 were used in this study (Table 1). Previous neutral microsatellite marker analysis of a set of \( P. falciparum \) samples collected in Peru during peak malaria expansion in the post-malaria eradication era (1998–2000) indicated a clonal population structure consisting of five clonal lineages (A, B, C, D and E). In the current study, we utilized the same seven neutral microsatellite markers (Table 2) to analyze samples collected between 1998 and 2001 and found that our results mirrored the previous finding of five clonal lineages as illustrated in Figure 2a.

![PF3D7 Chromosome 8](image)

**Figure 1 | Location of \( pfhrp2 \) and its flanking genes.** \( Pfhrp2 \) is located subtelomerically (position 1,374,236 to 1,375,299) on chromosome 8. The seven microsatellites flanking \( pfhrp2 \) are shown in boxes; each value indicates the distance (in kilobases) of a microsatellite locus either upstream (negative values) or downstream (positive values) of the start and stop codons of \( pfhrp2 \), respectively. The two genes flanking \( pfhrp2 \) are shown and their old gene IDs (Mal7p1.xxx) are included below the current PlasmoDB designations. Gene location and information were obtained from PlasmoDB v.9.1 (http://plasmodb.org/plasmo/).

Similar neutral microsatellite analysis performed on samples collected in 2003–2005 predicted the presence of at least eight different population clusters (Figure 2b). Only the D lineage persisted after 1998–2001 to form one of the eight clusters; all other lineages (A, B, C and E) were undetectable in the samples collected in 2003–2005. In addition to lineage D, seven new clusters emerged: A/B; A/C; B/C; C/D; C/D/E; X and Z (Figure 2b). Analysis of the data from neutral microsatellite analysis, combined with the Structure program output predicting the membership of certain samples into more than one cluster (Figure 2b), revealed that of the seven newly emerged clusters (A/B, A/C, B/C, C/D and C/D/E) were hybrids resulting from outcrossing of the five lineages A, B, C, D and E that were present in 1998–2001. The samples in each new hybrid cluster shared neutral microsatellite allele sizes found in the original five lineages A to E (Data not shown). Clusters X and Z consisted of isolates that had some neutral microsatellite locus sizes that had been seen in the 1998–2001 samples while other locus sizes were new.

The genetic relationship among the parasite isolates was further analyzed using the median joining network method in which neutral microsatellite markers were employed (Figure 3). Network analysis allowed for closer analysis of the genetic relationships among the clusters predicted by the Structure program. As observed in a previous study, samples collected in 1998–2001 had limited diversity within each of the five clonal lineages (Figure 3a). In contrast, greater genetic diversity was evident among the 2003–2005 isolates from the Network analysis; individual parasites were no longer grouped strictly by their Structure-assigned clusters (Figure 3b). Newly emerged A/C and B/C clusters formed independent single major

**Table 1 | Origin of \( P. falciparum \) samples collected from Peru**

| Study area (Department) | Collection years | Number of samples | Data set       |
|-------------------------|-----------------|------------------|----------------|
| Padre cocha (Loreto), Caballococha (Loreto) | 1998–2000 | 82 | ‘1998–2001’ |
| Bellavista (Piura) | 2001 | 10 | ‘1998–2001’ |
| Iquitos (Loreto) | 2003–2005 | 96 | ‘2003–2005’ |
groups (except two isolates in B/C that showed minor variations) indicating high levels of genetic similarity within each of these clusters. Although the C/D/E cluster formed a major group within the network, the cluster also showed some genetic diversity, with seven isolates forming groups outside of the major group (Figure 3b). The other four clusters (C/D, A/B, X and Z) were more randomly distributed, indicating high levels of genetic variation among each of these clusters, most likely due to outcrossing. It was evident that three members of the Z lineage clustered together with one set of A/B lineage isolates due to similarities in their neutral microsatellite allele sizes (Figure 3b).

Prevalence of pfhrp2-deleted parasites in samples collected between 1998 and 2005. Our analysis revealed that pfhrp2 was deleted in parasite isolates belonging to clonal lineages A, B, C and D in 1998–2001, with the highest prevalence of deletion found among samples from the D lineage (60%) followed by B (22.2%) and C (20.8%) lineages. Pfhrp2-negative parasites were less common in the A lineage (4.3%) (Figure 4). Furthermore, the prevalence of pfhrp2 deletion was around 20% in 1998–2001 and increased significantly to 40% by 2003–2005 ($p = 0.0015$; Table 3). Genetic deletions were not restricted to the pfhrp2 gene alone, but were found to extend to neighboring genes. A greater proportion of parasites were found to have deleted PF3D7_0831900 (the gene located 5’ of pfhrp2) than those that had deleted the 3’ flanking gene, PF3D7_0831700 (Table 3).

Parasite haplotype structure in relation to the pfhrp2 deletion phenotype. We had shown previously that each clonal lineage present in 1998–2001 (A to E) possessed a unique drug resistance haplotype\textsuperscript{14}. In the current study, the regions flanking pfhrp2 were genotyped using seven microsatellite markers located on chromosome 8 in order to understand the haplotype structure. Five of the seven pfhrp2 microsatellite loci (−41, 1.4, 2.5, 5.2 and 15, which were amplified in both pfhrp2-positive and pfhrp2-negative isolates. The loci −9.5, and −3.8 were deleted in many pfhrp2-negative samples.

Table 2 | Baseline (1998–2001) Neutral and Pfhrp2 flanking microsatellite marker profiles for Peru samples. The table shows the microsatellite loci sizes obtained for each of the five clonal lineages from samples collected in 1998–2001. TA1, Polyα, PPK2, TA109, 2490, C2M34 and C3M69 are neutral microsatellite loci\textsuperscript{18–23}. −41, −9.5, −3.8 are microsatellite loci located upstream of pfhrp2 while 1.4, 2.5, 5.2 and 15 are loci located downstream of pfhrp2 on chromosome 8. Pfhrp2 haplotypes were determined based on the five pfhrp2 flanking loci −41, 1.4, 2.5, 5.2 and 15, which were amplified in both pfhrp2-positive and pfhrp2-negative isolates. The loci −9.5, and −3.8 were deleted in many pfhrp2-negative samples.

| Locus | Chromosome | Clonal lineage | TA1 | Polyα | PPK2 | TA109 | 2490 | C2M34 | C3M69 | −41 | −9.5 | −3.8 | 1.4 | 2.5 | 5.2 | 15 |
|-------|------------|----------------|-----|-------|------|-------|-----|-------|-------|-----|------|------|----|----|----|----|
| A     | 6          | A              | 169 | 183   | 166  | 164   | 84  | 239   | 134   | 120 | 182  | 141  | 231| 142| 171| 168|
| B     | 172        | B              | 172 | 183   | 172  | 164   | 84  | 226   | 149   | 120 | 173  | 139  | 246| 155| 180| 172|
| C     | 178        | C              | 178 | 164   | 163  | 160   | 80  | 246   | 136   | 120 | 173  | 139  | 256| 161| 196| 177|
| D     | 172        | D              | 172 | 161   | 175  | 160   | 80  | 233   | 122   | 124 | 180  | 143  | 246| 151| 170| 176|
| E     | 174        | E              | 174 | 148   | 175  | 161   | 75  | 225   | 138   | 119 | 172  | 153  | 240| 161| 217| 162|

Figure 2 | Bayesian cluster analysis of Iquitos samples. The predicted number of likely clusters ($K$) for (a) samples collected in 1998–2001 ($N = 92$) was $K = 5$ while for (b) samples collected in 2003–2005 ($N = 96$) was $K = 8$. Each color corresponds to a population classified by Structure v 2.3.3 and each individual isolate is represented by a vertical bar. The Y axis represents the estimated proportion of membership of an individual to each predicted population cluster.
15) were used to identify five unique pfhrp2 haplotypes because these markers were detectable in most of the isolates genotyped in this study (Table 2). Two pfhrp2 microsatellite loci (9.5 and 3.8) were not used to define the pfhrp2 haplotypes because they had been deleted in a number of pfhrp2-negative parasite isolates. These five pfhrp2 haplotypes, a, b, c, d and e, corresponded with the five clonal lineages of parasites (A to E) previously described (Figure 5a; Table 2).

To investigate the evolving pattern of pfhrp2 deletion in the Peruvian Amazon over time, the deletion patterns were determined for parasite isolates belonging to the five clonal lineages observed in 1998–2001 and the eight lineages seen in 2003–2005. There was clear evidence for pfhrp2 deletion in all of the observed clonal lineages except E (Figure 4). The distribution of pfhrp2 deletions among the pfhrp2 haplotypes is illustrated in the Network diagrams of Figure 5.

In 1998–2001, the highest proportion of pfhrp2 deletion was found in the d haplotype (9/15; 60%) followed by b (4/18; 22.2%), c (5/25 20%) and a (1/23; 4.3%) (Figure 5a). Deletion was not observed in the e haplotype (Figure 5a).

Even though four of the five clonal lineages (A, B, C, and E) could not be detected by 2003–2005, all five pfhrp2 haplotypes (a, b, c, d and e) that were present in 1998–2001 persisted in the 2003–2005 set of samples (Figure 5). We also observed four isolates from the 2003–2005 sample set with pfhrp2 microsatellite alleles that had not been detected in previous years. The allele sizes differed among these samples. We assigned them to a sixth haplotype, ‘ψ’ (Figure 5b). The highest level of deletion was found within the d haplotype (23/26; 88.5%). This was followed by c (4/16; 25%), b (7/40; 17.5%) and a (1/16; 6.25%). The one sample belonging to the e haplotype was pfhrp2-positive, while all four ψ haplotype samples were pfhrp2-negative.

Discussion

We utilized microsatellites to investigate the genetic origins of pfhrp2 deletion in Peruvian P. falciparum samples. We have shown that pfhrp2-negative parasites in the Peruvian Amazon have multiple origins. Deletion of pfhrp2 in at least four out of the five clonal
Figure 4 | Prevalence of pfhrp2 among the clonal lineages identified in Iquitos. Clinical samples were collected in 1998–2001 (N = 92) and 2003–2005 (N = 96). Dark grey boxes represent the proportion of pfhrp2-positive samples while light grey boxes represent pfhrp2-negative isolates. The clonal lineage assignments are indicated along the x-axis.

lineages found in 1998–2001 (A to E) strongly suggests that pfhrp2 deletion has evolved independently in multiple genetic backgrounds (Figures 4 and 5).

Malaria transmission was reduced to very low levels in Peru after the implementation of the malaria eradication program in the 1950s. The disease re-emerged in the 1990s with peak P. falciparum transmission occurring between 1998–2000. Our earlier work showed that the population structure of P. falciparum in Peru was clonal in nature during the peak transmission years and at least five clonal lineages were found. Given this background and our access to historical samples that were collected during this peak period and in subsequent years (until 2005), we were interested in determining the influence of increased transmission and clonal population structure on the evolution of pfhrp2 gene deletions. In this study, we have found that each of these previously defined clonal populations has a distinct pfhrp2 haplotype background based on five pfhrp2 microsatellite loci: −41, 1, 4, 2.5, 5.2 and 15 (Table 2). This distinction has allowed us to determine the genetic origins of pfhrp2 deletion in these populations. However, since the historical samples available to us only went as far back as 1998, we could not further assess when these deletion events originated and spread. It is also unclear whether the deletion events originated in Peru and expanded recently due to some selective pressure.

We speculated previously, based on the drug resistance haplotype history, that clonal lineages C, D and E may have been descendants of P. falciparum populations from the Pacific Coast/Andean region of Colombia and Ecuador. On the other hand, clonal lineages A and B were ancestrally related to parasites from the Amazon region of Brazil. The occurrence of pfhrp2 deletions in these two distinct ancestral lineages of parasites raises the question whether this deletion occurred before these parasites migrated to Peru or whether it was a more recent event. If these deletion events occurred before the introduction of the clonal populations to Peru, then one would expect to find pfhrp2-negative parasites in the Colombian coast/Andean region as well as in the Brazilian Amazon. Analysis of parasite samples from these regions will provide further insights into the evolutionary history of these deletions.

The observed increase in the proportion of pfhrp2-negative parasites from 20% in 1998–2001 to 40% in 2003–2005 (Table 3) has major implications for the implementation of PfHRP2-based RDTs in the Peruvian Amazon. In one study, a comparable increase in pfhrp2-negative parasites was observed while testing a limited number of samples collected in the Peruvian Amazon during a ten-year period from 1996 to 2006. It is not clear whether any biological selection process may have favored the expansion of pfhrp2 deleted parasites in this region and if it is still ongoing.

Pfhrp2 deletion was restricted to the four clonal lineages A to D observed in 1998–2001, but was also observed in the hybrid clusters that emerged between 2003–2005 (Figures 4 and 5b). Our data (Figure 3b) is consistent with previous studies suggesting the emergence of new clonal parasite lineages due to outcrossing among the ancestral lineages. Although seven entirely new hybrid clusters emerged between 2003 and 2005 (clonal lineages A, B, C and E were not detected), most of these parasites had each inherited one of the five pfhrp2 haplotypes (a to e) found at the earlier time point. The continued inheritance of a to e haplotypes among the newly emerged clusters suggest that these parasites are not necessarily due to new introductions but due to admixture of historical samples present in this region. There was limited recombination among the pfhrp2 haplotypes because only four of the ninety six isolates from

Table 3 | Proportion of pfhrp2 and flanking gene deletions in samples from different sites of the Peruvian Amazon. Data show the number (and percentage) of parasites with gene deletions out of total samples tested.

| Collection sites | PF3D7_0831900 | pfhrp2 | PF3D7_0831700 |
|------------------|----------------|--------|----------------|
| Caballococha     | 22/36          | 13/36  | 10/36          |
| Padre cocha      | 14/46          | 6/46   | 3/46           |
| Bellavista       | 0/10           | 0/10   | 0/10           |
| **Total (1998–2001)** | **36/92 (39.1%)** | **19/92 (20.7%)** | **13/92 (14.1%)** |
| Iquitos          | 43/96          | 39/96  | 11/96          |
| **Total (2003–2005)** | **44.8%** | **40.6%*** | **11.5%** |

*The increase in the proportion of pfhrp2-negative parasites from 1998–2001 to 2003–2005 was determined to be statistically significant (p = 0.0015) using the chi-square test.
2003–2005 showed a different haplotype structure (ψ) while the rest retained their ancestral haplotype structure.

Although this study has clearly indicated multiple origins for pfhrp2 deletion, it remains to be determined if there is selective pressure favoring the spread of pfhrp2-negative parasites. This question could not be adequately answered in this study due to limited numbers of genetic loci and geographical areas studied, and the use of historical samples from a relatively short time period (1998–2005). However, we can hypothesize various reasons for the spread of pfhrp2-deleted parasites. One possibility is that changes in drug treatment policies around this period (chloroquine and sulfadoxine-pyrimethamine (SP) were replaced with artemisinin combination therapy, ACT, in 2001) could have favored the expansion of these new hybrids. However, there was no clear evidence to support this hypothesis. It should be noted that only lineages A and B carried highly SP-resistant dihydrofolate reductase (pfldhfr; located on chromosome 4) and dihydropteroate synthase (pfldhps; located on chromosome 8) genotypes. We had previously observed a decline in the highly resistant pfldhfr and pfldhps lineages from 1998–2001 to 2006. This decline may have impacted the evolution of these clonal lineages. However, pfhrp2 haplotype β from the B clonal lineage did not disappear within the same time period, suggesting that while drug policy changes may have impacted the fitness of the SP-resistant genotypes, they had no apparent effect on the pfhrp2 locus on chromosome 8.

Continuous use of PfHRP2-based RDTs may also create a selective pressure favoring the expansion of pfhrp2-negative parasite lineages. However, RDTs were not routinely used for malaria case management in Peru because microscopic analysis was available in most health facilities. Therefore, it is unlikely that any limited use of PfHRP2-based RDTs by itself contributed to the selection of a pfhrp2-negative parasite population in Peru between 1998 and 2006. It is unclear whether bottlenecks and uneven distribution of limited P. falciparum clonal lineages, in addition to their underlying population structure, could account for the selection of pfhrp2-negative parasites in Peru. Since there is a wider deletion in the genomic

---

Figure 5 | Prevalence of pfhrp2 deletion among the P. falciparum haplotypes found in Peru. The genetic relationships among parasite isolates collected in (a) 1998–2001 and (b) 2003–2005 were constructed using the seven neutral microsatellite loci shown on Table 2. Blue color represents the proportion of pfhrp2-positive samples while yellow indicates the proportion of pfhrp2-negative samples. Additionally, the pfhrp2 microsatellite-based haplotypes are labeled and highlighted. (b) Circle circumferences are proportional to the number of samples that belonged to a particular haplotype.
region flanking pfhrp2, there is also the possibility that selection may be acting on a nearby target that is totally independent of pfhrp2 and is yet to be discovered.

Pfhrp2 deletions are significant because of the impact they have on the use of PfHRP2-based RDTs for malaria detection in South America. For this reason, the ability to monitor parasites that are pfhrp2-positive but are negative for the upstream flanking gene PF3D7_0831900 and surrounding microsatellite loci may allow one to predict where subsequent generations of pfhrp2-negative parasites could appear. The biological significance of the presence of pfhrp2-negative P. falciparum parasites in Peru is unclear. Given our findings regarding the genetic origins of pfhrp2-deleted P. falciparum parasites in the Peruvian Amazon, the next step would be to determine if any selective force is influencing the increase in parasites with this deletion. The nature, gene target and effect of this selective force on the P. falciparum parasite population will be of interest to those living in regions where pfhrp2-deletion has recently been reported, including Mali and India, albeit at much lower levels than that seen in Peru11. Further studies of parasite populations from other regions will provide clarity regarding this genotype.

In summary, our study provides evidence that the expansion of pfhrp2-negative P. falciparum parasites in Peru originated from multiple genetic backgrounds, rather than being caused by selection for a single, rare deletion event. Reports of low levels of pfhrp2-negative parasites in other geographically distant regions suggest that the pfhrp2 deletion may continue to occur randomly on other genetic backgrounds. Further understanding of the biological features that may favor the selection and spread of parasites with genomic deletions in the pfhrp2 region will be essential to determine the evolutionary basis for the spread of these parasites.

Methods

Peru study sites and P. falciparum clinical isolates. The details of study sample collection year and location are provided in Table 1. The 92 P. falciparum parasite samples used in this study (collected from 1998–2001) were obtained using sample collection protocols that were approved by the Ethical Review Committees on the Instituto Nacional de Salud, U.S. Naval Medical Research Center Institutional Review Board, the Universidad Peruana Cayetano Heredia and the U.S. Centers for Disease Control and Prevention approved the study. Written informed consent was provided by study participants and/or their legal guardians.

DNA isolation and genotyping of pfhrp2 and flanking genes. DNA was isolated from filter paper blood spots or whole blood using the QiAamp DNA mini kit (QiAGEN, Valencia, CA) in accordance with the manufacturer’s instructions. Confirmation of P. falciparum infection was conducted by PCR amplification of the 18S ribosomal gene using methods described by Singh et al20. Samples that were positive for P. falciparum were then tested for pfhrp2. Two sets of primers were designed to amplify a 228 bp fragment of pfhrp2 using a nested PCR protocol. The outward forward primer was 5’-GGTTTCCCTCTCAA-AAAATAAAG-3’ and outward reverse primer was 5’-CTCTATAGTGTGTTGAGGAT-3’. The secondary reaction utilized 5’-GTATATTACCCTCGTCCGTGTTTGGC-3’ (forward) and 5’-CTCACAAGGTATATATAATCGGAAA-3’ (reverse) primers. The cycling conditions were 95°C for 1 min; 30 cycles of 95°C for 15 sec, 55°C for 30 sec, 68°C for 30 sec; 68°C for 1 min (secondary reaction). 18S rRNA-positive samples were identified as pfhrp2-negative if there was a failure to amplify the latter gene after two attempts.

Flanking genes pfhrp2 - PF3D7_0831900 (5,353 kb upstream) and PF3D7_0831700 (6,177 kb downstream) (Figure 1) were amplified using the following primers: For PF3D7_0831900, the outward forward primer was 5’-GAT ATC ATT AGA AAA CAA GAG CTT AG-3’ and outward reverse primer was 5’-TAT CCA ATC CTT CCT TGG CAA CAC C-3’; The secondary reaction utilized 5’-TAT GAA CGC ATT ATAAGT GAG CCA G-3’ (forward) and 5’-TAT CCA ATC CTT CCT TGG CAA CAC C-3’ (reverse) primers. For PF3D7_0831700, the outward forward primer was 5’-AGA CAA GCA GCT ACC AAA AAT GAT GCA GAG C-3’ and outward reverse primer was 5’-TAA ATG TGT ATC TCC TGA GGT AGC -3’; The secondary reaction utilized 5’-CCA TTT GTG TTT TAA ATG TTA GGT C-3 (forward) and 5’-TAA ATG TGT ATC TCC TGA GGT AGC -3’ (reverse) primers.

For all samples, the reaction conditions were as follows: 95°C/30 sec; 30 cycles of 95°C/30 sec, X°C/30 sec, 68°C/30 sec; 68°C/5 min, where X, the annealing temperature, was between 55–62°C depending on the primer pair used.

Multilocus genotyping. Whole genome amplified DNA (REPLi-g Whole Genome Amplification Kit, Qiagen, Valencia, CA) samples were assayed for seven putatively neutral microsatellites and several microsatellites flanking pfhrp2. The neutral loci have previously been described and used in various studies in South America20–22. TA1 was found on chromosome 6; poly A, poly C, chromosome 4; PIPK2, chromosome 12; TA109, chromosome 6; and 2490, chromosome 10. The last two neutral microsatellite loci used were C2M34 (chromosome 2) and C3M69 (chromosome 3)23. Additionally, we amplified a pair of primary microsatellites flanking pfhrp2 at the following annealing temperatures: −9.5 kb, −3.8 kb, 1.4 kb, 2.5 kb, 5.2 kb and 15 kb (Figure 1). Upstream distances (negative values) were calculated from the start codon of pfhrp2 and downstream distances (positive values) calculated from the gene’s stop codon. If a sample failed to amplify any of the fourteen microsatellite loci, the PCR was repeated. If there was no amplification after two attempts, it was recorded that the particular locus had been deleted.

The amplification products were labeled with fluorescent dyes (HEX or FAM) and assayed for size on an Applied Biosystems 3130xl sequencer. The fragments were then scored with GeneMapper software v.3.7 (Applied Biosystems, Foster City, CA) using default microsatellite settings, whereby bands smaller than 500 relative fluorescence units (rfu) were defined as background. Samples for which we obtained no amplification in some loci were re-analyzed to complete the haplotypes.

Cluster analysis. In order to examine the population structure of P. falciparum isolates collected in Peru between 1998–2001 and 2003–2005, we used a Bayesian approach to infer the number of genetically related clusters (K) from the individual microsatellite haplotype profiles generated using seven neutral microsatellites. Neutral microsatellite analysis was implemented with Structure v2.3.34. Any locus that we failed to amplify was assigned a null value (−9) as recommended by Structure. Although we had prior information on the presence of five clonal lineages in this population between 1998 and 200113,14, we predicted that there would be a change in the number of resulting clusters due to recombination among the clonal lineages. Therefore, we chose to test the likelihood of finding between one and ten clusters in this population (K = 1 to K = 10). We performed twenty replicates of the clustering algorithm for each value of K with a burn-in period of 10,000 iterations and 100,000 Markov Chain Monte Carlo replications. We used the admixture model with correlated allele frequencies24. The most likely number of clusters was defined by calculating the AK value as described by Evanno et al25. The Evanno method was implemented by inputting our data from Structure into the Structure Harvester program (http://taylor0.biology.ucla.edu/struct_harvest/).

Network analysis. To determine the relatedness of the samples collected in 1998–2001 and 2003–2005, we created median-joining network diagrams using Network v. 4.6.1.0 (fluxus-engineering.com)27 and neutral microsatellite data from the seven neutral microsatellites. Any locus that we failed to amplify was assigned a null value (99) for the purposes of analysis.

Statistical analysis. The proportion of pfhrp2-negative parasites collected in 1998–2001 and in 2003–2005 were compared using the chi-square test. An α of 0.05 was our threshold of statistical significance.

Ethics statement. The Ethical Review Committees of the Instituto Nacional de Salud, Universidad Peruana Cayetano Heredia, US Naval Medical Research Center Institutional Review Board and Institutional Review Boards of the U.S. Army and U.S. Navy as well as the U.S. Centers for Disease Control and Prevention approved the original studies. Permission to use samples retrospectively for this study was obtained from the CDC. Written informed consent was provided by study participants and/or their legal guardians.

1. Moody, A. Rapid diagnostic tests for malaria parasites. Clin Microbiol Rev 15, 66–78 (2002).
2. Rock, E. P. et al. Comparative analysis of the Plasmodium falciparum histidine-rich proteins HRP-1, HRP-2 and HRP-III in malaria parasites of diverse origin. Parasitology 95 (Pt 2), 209–227 (1987).
3. Scherf, A. & Mattei, D. Cloning and characterization of chromosome breakpoints of Plasmodium falciparum: breakage and new telomere formation occurs frequently and randomly in subtelomeric genes. Nucleic Acids Res 20, 1491–1496 (1992).
4. Weilens, T. E. & Howard, R. J. Homologous genes encode two distinct histidine-rich protein-3 proteins in a cloned library of Plasmodium falciparum. Proc Natl Acad Sci U S A 83, 6065–6068 (1986).
5. Kemp, D. J., Thompson, J. K., Walliker, D. & Corcoran, L. M. Molecular karyotype of Plasmodium falciparum: conserved linkage groups and expandable histidine-rich protein genes. Proc Natl Acad Sci U S A 84, 7672–7676 (1987).
6. Chess, A. & Ravelich, J. The large deletions result from breakage and healing of P. falciparum chromosomes. Cell 55, 869–874 (1988).
7. Gamboa, D. et al. 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 5, e8091 (2010).
8. Houze, S., Hubert, V., Le Pessec, G., Le Bras, J. & Clain, J. Combined deletions of pfrp2 and pfrp3 genes result in Plasmodium falciparum malaria false-negative rapid diagnostic test. *J Clin Microbiol* **49**, 2694–2696 (2011).

9. Koita, O. A. *et al.* False-Negative Rapid Diagnostic Tests for Malaria and Deletion of the Histidine-Rich Repeat Region of the hrp2 Gene[dagger]. *Am J Trop Med Hyg* **86**, 194–198 (2012).

10. Wurtz, N. *et al.* Pfrp2 and pfrp3 polymorphisms in Plasmodium falciparum isolates from Dakar, Senegal: impact on rapid malaria diagnostic tests. *Malar J* **12**, 34 (2013).

11. Kumar, N. *et al.* Genetic deletion of HRP2 and HRP3 in Indian Plasmodium falciparum population and false negative malaria rapid diagnostic test. *Acta tropica* (2012).

12. Dharia, N. V. *et al.* Genome scanning of Amazonian Plasmodium falciparum shows subtelomeric instability and clindamycin-resistant parasites. *Genome Res* **20**, 1534–1544 (2010).

13. Branch, O. H. *et al.* Plasmodium falciparum genetic diversity maintained and amplified over 5 years of a low transmission endemic in the Peruvian Amazon. *Mol Biol Evol* **28**, 1973–1986 (2011).

14. Griffing, S. M. *et al.* South American Plasmodium falciparum after the Malaria Eradication Era: Clonal Population Expansion and Survival of the Fittest Hybrids. *PLoS One* **6**, e23486 (2011).

15. Bacon, D. J. *et al.* Dynamics of malaria drug resistance patterns in the Amazon basin region following changes in Peruvian national treatment policy for uncomplicated malaria. *Antimicrob Agents Chemother* **53**, 2042–2051 (2009).

16. Zhou, Z. *et al.* Decline in sulfadoxine-pyrimethamine-resistant alleles after change in drug policy in the Amazon region of Peru. *Antimicrob Agents Chemother* **52**, 739–741 (2008).

17. Singh, B. *et al.* A genus- and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies. *Am J Trop Med Hyg* **60**, 687–692 (1999).

18. Anderson, T. J. C. *et al.* Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. *Molecular Biology and Evolution* **17**, 1467–1482 (2000).

19. Anderson, T. J. C., Su, X. Z., Bockarie, M., Lagog, M. & Day, K. P. Twelve microsatellite markers for characterization of Plasmodium falciparum from finger-prick blood samples. *Parasitology* **119**, 113–125 (1999).

20. dalla Martha, R. C., Tada, M. S., Ferreira, R. G., da Silva, L. H. & Wunderlich, G. Microsatellite characterization of Plasmodium falciparum from symptomatic and non-symptomatic infections from the Western Amazon reveals the existence of non-symptomatic infection-associated genotypes. *Mem Inst Oswaldo Cruz* **102**, 293–298 (2007).

21. Hoffmann, E. H., Ribolla, P. E. & Ferreira, M. U. Genetic relatedness of Plasmodium falciparum isolates and the origin of allelic diversity at the merozoite surface protein-1 (MSP-1) locus in Brazil and Vietnam. *Malar J* **2**, 24 (2003).

22. Machado, R. L. *et al.* Genetic structure of Plasmodium falciparum populations in the Brazilian Amazon region. *J Infect Dis* **190**, 1547–1555 (2004).

23. McCollum, A. M., Mueller, K., Villegas, L., Udhayakumar, V. & Escalan, A. A. Common origin and fixation of Plasmodium falciparum dfhr and dhps mutations associated with sulfadoxine-pyrimethamine resistance in a low-transmission area in South America. *Antimicrob Agents Chemother* **51**, 2085–2091 (2007).

24. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).

25. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).

26. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**, 2611–2620 (2005).

27. Bandelt, H. J., Forster, P. & Rohl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**, 37–48 (1999).

**Acknowledgements**

We wish to thank all members of the study team who originally contributed to the collection of samples and study participants for their permission to use these samples. We acknowledge funding support from the Amazon Malaria Initiative which is supported by the U.S. Agency for International Development. S.A. and T.H. were supported by the American Society of Microbiology/CDC Postdoctoral Fellowship Program. D.G. is supported by DCG (Framework Agreement 03, 2008–2013, Project 95502); D.G. and J.B. are supported by the Peru/Brazil International Center of Excellence in Malaria Research U19AI089681 (NIH/NIAID, United States Public Health Service, USA). K.T. is supported by the NIH/Fogarty International Center Global Infectious Diseases Training Grant D43TW007120 (IMTAvH-UCSD). A.M.O., C.L., J.W.B. and V.U. are employees of the U.S. Government and D.B. is a military service member. This work was prepared as part of their official duties. We thank Dr. Fredrik Vannberg of the Georgia Institute of Technology and Mr. Ira F. Goldman of the Centers for Disease Control and Prevention for critical review of the manuscript.

**Author contributions**

S.A., T.H., J.W.B. and V.U. conceived and designed the study. A.M.O., J.W.B. and V.U. supervised the project. D.G., K.T., J.B., W.M.Q., N.A., C.L., A.J.M. and D.J.B. provided the field samples. S.A., J.F.A. and S.M.G. performed the experiments. S.A., T.H., J.F.A. and S.M.G. analyzed the data. S.A., A.M.O., J.W.B. and V.U. wrote the paper.

**Additional information**

Disclaimer: The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defence, nor the U.S. Government.

Competing financial interests: The authors declare no competing financial interests.

**How to cite this article**: Akinyi, S. *et al.* Multiple genetic origins of histidine-rich protein 2 gene deletion in *Plasmodium falciparum* parasites from Peru. *Sci. Rep.* **3**, 2797; DOI:10.1038/srep02797 (2013).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported license. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0