Supplementary Text 1

As described in the Methods, the degenerate primers we use amplify the DBL*a* domain of *var* genes. We performed an additional validation of our PCR (using these degenerate DBL*a* primers) and sequencing methodology to understand the margin of error of detection of all DBL*a* types in an isolate. DBL*a* types were translated into amino acid sequences and classified as upsA or non-upsA, using the classifyDBLalpha pipeline (Ruybal-Pesántez et al. 2017) to examine whether the expected genomic proportions of upsA/non-upsA were obtained in each isolate.

Given our study was conducted in South America, we used the Honduran laboratory reference strain HB3 as a benchmark for the expected number of total DBL*a* types (i.e. repertoire size), as well as the proportion of upsA/non-upsA. Whole genome sequencing of HB3 has identified 44 *var* genes, with 8 upsA DBL*a* types (defined by DBL*a* domain 1), 34 non-upsA DBL*a* types (DBL*a* domains 0 and 2) and 2 upsE (*var2csa* genes, defined by DBLpam domains and do not have DBL*a* domains) (Rask et al. 2010). We independently verified this using 37 technical replicates of *var* DBL*a* PCR amplification and illumina sequencing of our HB3 laboratory isolate. It is worth noting that since HB3 has two *var2csa* genes that do not have DBL*a* domains, these will not be amplified by our degenerate primers, so we expect that only 42 of the *var* genes of HB3 could be amplified.

HB3 technical replicates

From the data obtained from 37 HB3 technical replicates, we identified the expected repertoire sizes with a median of 39 DBL*a* types (range: 36-41). A median of 7 upsA DBL*a* types (range: 6-8) and a median of 33 non-upsA DBL*a* types (range: 30-34) were identified (Figure 1). The median genomic proportion of upsA DBL*a* types was 17.5% (range: 15.4-19.5%) and 82.5% (range: 80.5-84.6%) for non-upsA DBL*a* types (Figure 2). We found that of the 46 identified in the 37 technical replicates, 40 of them were consistently identified in the majority of replicate isolates (range 21 to 37 replicates). Of these 40 types, 7 were ups-A and 33 were non-upsA. All of these findings are in line with what is expected from whole genome sequencing data (Rask et al. 2010).

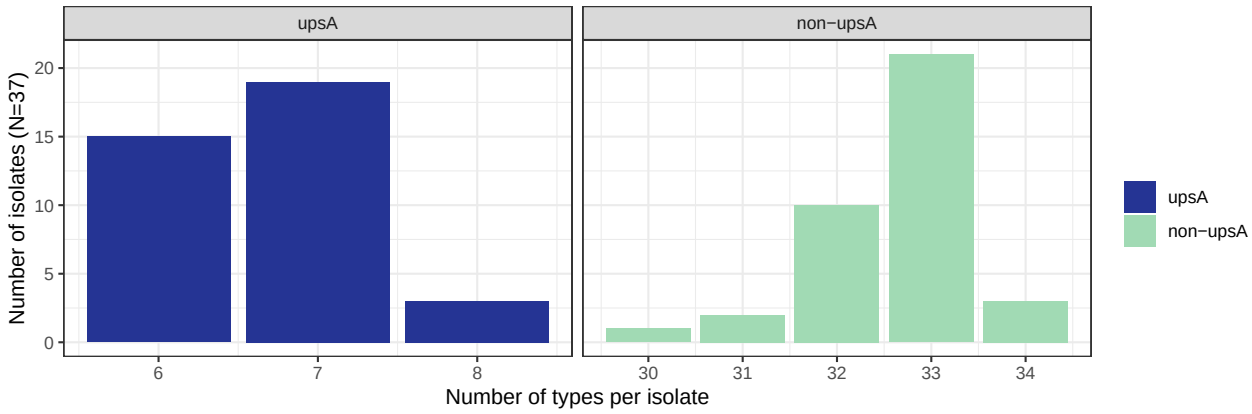

Figure 1: The distribution of the number of upsA and non-upsA types identified in each HB3 isolate repertoire.

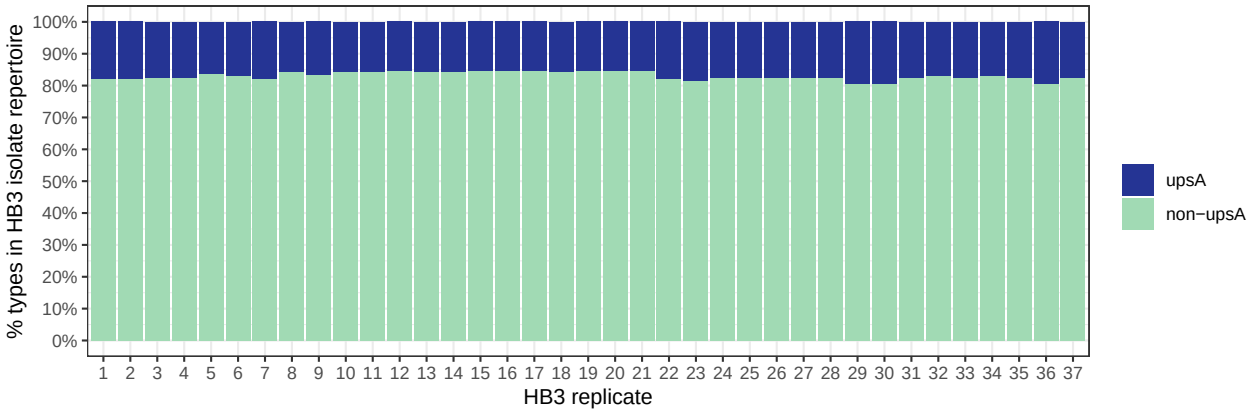

Figure 2: The proportion of upsA and non-upsA types identified in each HB3 isolate repertoire.

Field isolate genomic proportions of upsA and non-upsA DBL*a* types

The 543 unique DBL*a* types identified in the 186 South American *P. falciparum* isolates (N = 58 Ecuadorian P. falciparum isolates from this study, N = 128 previously published *P. falciparum* isolates from Colombia, French Guiana, Peru and Venezuela) were translated into amino acid sequences and classified as upsA or non-upsA, using the classifyDBLalpha pipeline (Ruybal-Pesántez et al. 2017). There were 79 upsA and 464 non-upsA types.

Looking first at the 195 types (26 upsA and 169 non-upsA) identified in Ecuadorian isolates, we obtained a median genomic proportion of upsA of 10.8% (range: 5.1-18.2%) and 89.2% (range: 81.8-94.9%) of non-upsA types for all isolate *var*codes (Figure 3).

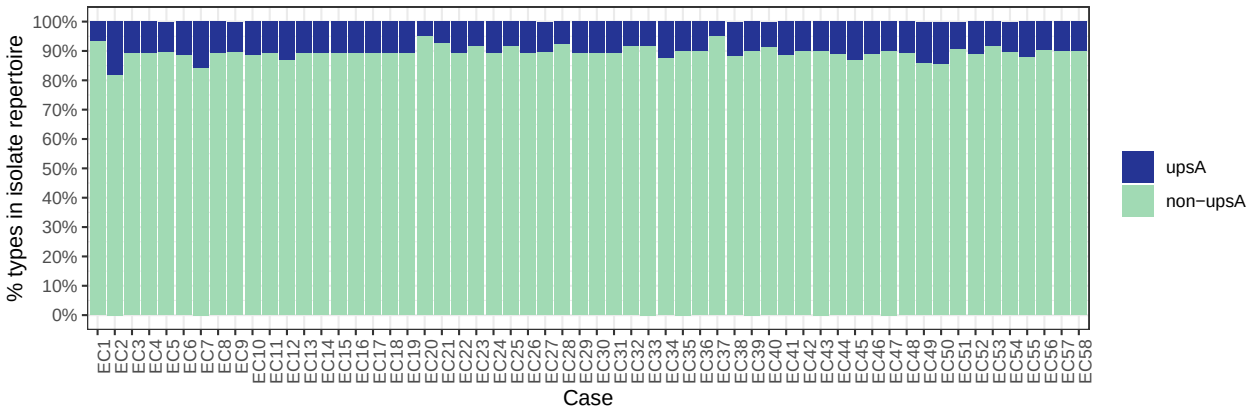

Figure 3: The proportion of upsA and non-upsA types in each isolate repertoire. The case numbers in the x-axis can be used to identify the clinical information of the participant in Table 1.

To confirm the patterns we observed in Ecuadorian isolates, we also compared them to the other South American isolates. In the other South American isolate *var*codes the genomic proportions of upsA/non-upsA were similar, with a median proportion of upsA of 9-14% and 86-91% for non-upsA types (Table 1).

With regards to the number of upsA identified in all the South American isolate *var*codes, we identified a median of 4-5 upsA types and 31-42.5 non-upsA types (Table 2).

Table 1: The proportion of upsA types in South American isolates

country	min	median	mean	max
Colombia	0.00	0.09	0.09	0.14
Ecuador	0.05	0.11	0.11	0.18
French Guiana	0.04	0.11	0.11	0.18
Peru	0.08	0.14	0.14	0.20
Venezuela	0.05	0.11	0.12	0.21

Table 2: The number of upsA types identified in South American isolates

country	min	median	mean	max
Colombia	0	4	4	6
Ecuador	2	4	4	6
French Guiana	2	5	6	14
Peru	2	5	5	7
Venezuela	2	4	4	9

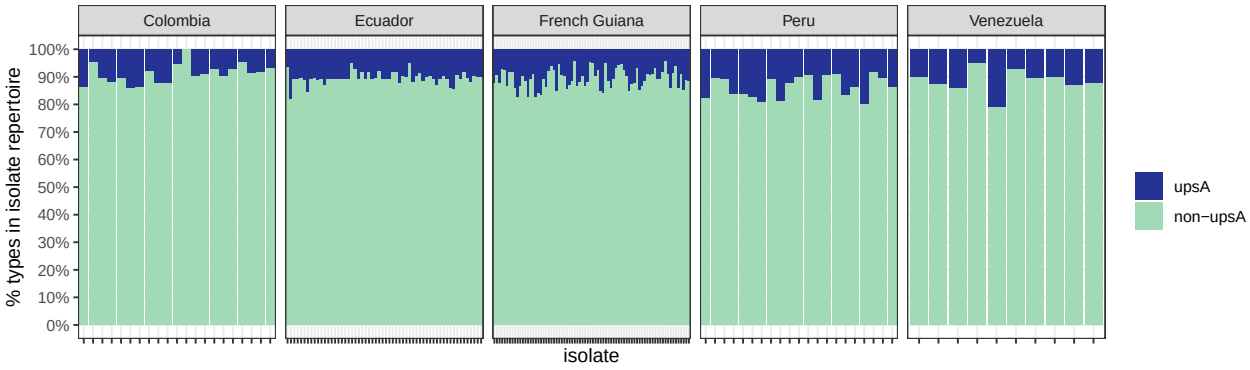

Figure 4: The proportion of upsA and non-upsA types in each isolate repertoire stratified by country.

Inheritance of DBL*a* types in the recombinant *var*codes

For the outbreak *var*code1, its 47 DBL*a* types were classified as 5 upsA and 42 non-upsA. In Figure 5 we look specifically at the "inheritance" of upsA (blues) vs non-upsA (greens) types from the parental outbreak *var*code1 in the case of the parasites with recombinant *var*codes (*var*codes3,4,6,7). This provides a proxy to examine inheritance of types with regards to their chromosomal location. The proportion of the outbreak types that were inherited in the recombinant parasite *var*codes is indicated in the darker shades of blue or green, showing that the proportion of inherited types varied both by *var*code and upsA/non-upsA. Overall, the DBL*a* type sharing patterns in parasites with recombinant varcodes are consistent with inheritance of 50% types, with a higher proportion of non-upsA inherited types (~40-70%, i.e. 17 to 30 of the 42 types) vs upsA (~20-50%, i.e. 1 to 3 of the 5 upsA types). The exception was *var*code7 where 60-80% of upsA were inherited, i.e. 3-4 of the 5 upsA types. The lighter shades correspond to those types that were not inherited from the outbreak clone but from the other parent.

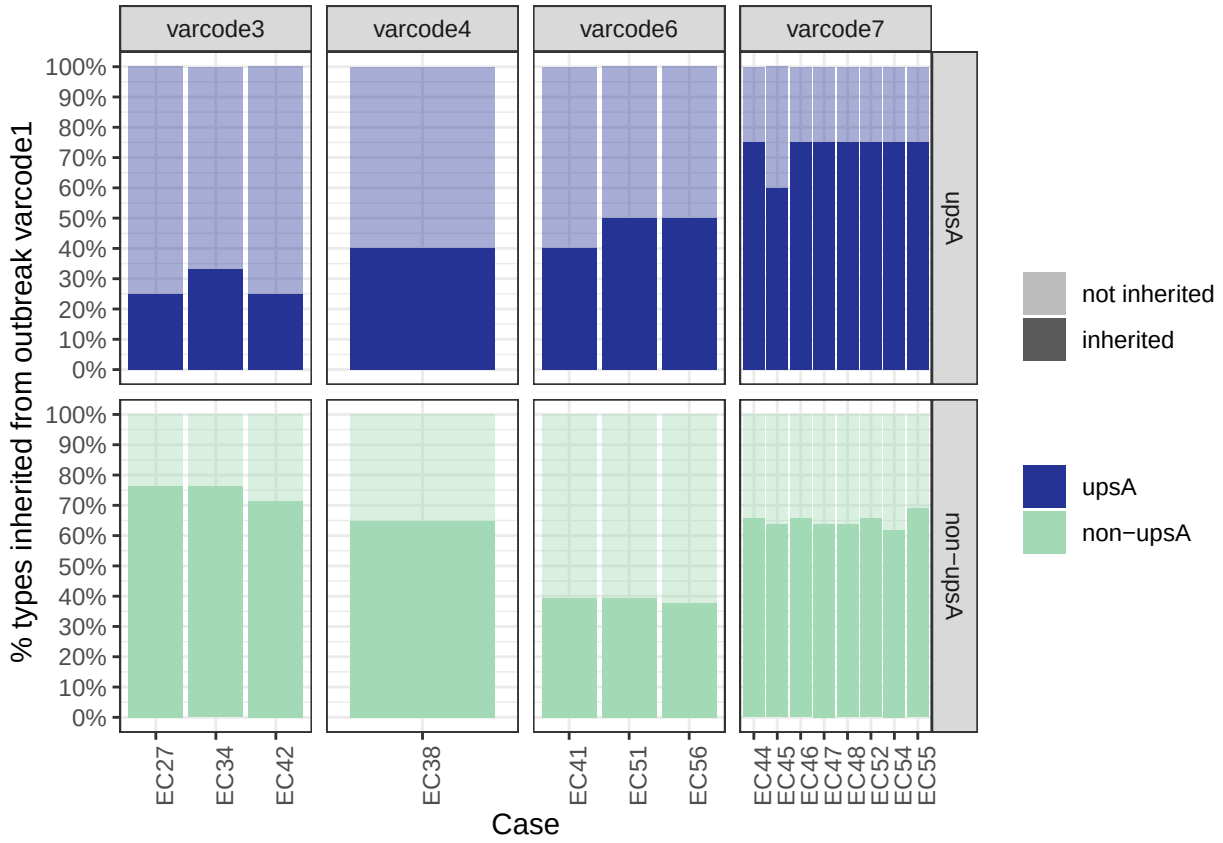

Figure 5: The proportion of the outbreak types that were inherited in the recombinant parasite varcodes is indicated in the darker shades of blue or green. The proportion of the outbreak types that were inherited in the recombinant parasite varcodes is indicated in the darker shades of blue or green and the lighter shades correspond to those types that were not inherited from the outbreak clone but from the other parent.

References

Rask, Thomas S., Daniel A. Hansen, Thor G. Theander, Anders Gorm Pedersen, and Thomas Lavstsen. 2010. "Plasmodium Falciparum Erythrocyte Membrane Protein 1 Diversity in Seven Genomes Divide and Conquer." Edited by Jonathan A. Eisen. *PLoS Computational Biology* 6 (9): e1000933. https://doi.org/10.1371/journal.pcbi.1000933.

58 Ruybal-Pesántez, Shazia, Kathryn E. Tiedje, Gerry Tonkin-Hill, Thomas S. Rask, Moses R. Kamya, Bryan
59 Greenhouse, Grant Dorsey, Michael F. Duffy, and Karen P. Day. 2017. "Population Genomics of
60 Virulence Genes of Plasmodium Falciparum in Clinical Isolates from Uganda." *Scientific Reports* 7 (1).
61 https://doi.org/10.1038/s41598-017-11814-9.