Shifts in *P. falciparum* genetic structure and gametocyte markers in the transition to elimination

Himanshu Gupta

ISGlobal, Barcelona Center for International Health Research (CRESIB), Hospital Clínic - Universitat de Barcelona / Centro de Investigação em Saúde da Manhiça <https://orcid.org/0000-0003-1391-8358>

Beatriz Galatas

ISGlobal, Barcelona Center for International Health Research (CRESIB), Hospital Clínic - Universitat de Barcelona / Centro de Investigação em Saúde da Manhiça

Arlindo Chidimatembe

Manhiça Health research Center

Wilson Simone

Manhiça Health research Center

Gloria Matambisso

Manhiça Health research Center

Arnau Pujol

ISGlobal, Barcelona Center for International Health Research (CRESIB), Hospital Clínic - Universitat de Barcelona / Centro de Investigação em Saúde da Manhiça <https://orcid.org/0000-0001-7288-6435>

Carlos Ruiz-Arenas

ISGlobal, Barcelona Global Health Institute <https://orcid.org/0000-0002-6014-3498>

Juan Gonzalez

Barcelona Institute for Global Health <https://orcid.org/0000-0003-3267-2146>

Lidia Nhamussa

Manhiça Health research Center

Pau Cisteró

ISGlobal

Regina Rabinovich

ISGlobal, Barcelona Center for International Health Research (CRESIB), Hospital Clínic - Universitat de Barcelona / Centro de Investigação em Saúde da Manhiça

Pedro Alonso

Universitat de Barcelona

Francisco Saúte

Manhiça Health research Center

Pedro Aide

Manhiça Health Research Centre <https://orcid.org/0000-0002-2379-6529>

Alfredo Mayor (✉ alfredo.mayor@isglobal.org)

Article

Keywords: malaria, elimination, gametocytes, genetic diversity

Posted Date: November 9th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-92978/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Large-scale programs targeting *Plasmodium falciparum* (*Pf*) elimination can exert strong selection pressures on the parasite population. To better understand the impact that elimination initiatives can have on *Pf* genetic structure and gametocyte carriage, we applied amplicon-based sequencing of two polymorphic *Pf* genes and quantitative reverse-transcription PCR targeting gametocyte-specific genes to *Pf* isolates collected in Magude District (Southern Mozambique) before and after an elimination initiative. The 71% reduction of *Pf* prevalence achieved in 2 years was followed by reductions in *Pf* genetic diversity and increases in between-infection similarity. These genetic shifts were accompanied by increases in the relative transcript number of the female mature gametocyte marker *pfs25*, the *pfap2g* transcription factor that drives gametocytogenesis and the sexual ring marker *pfgexp02*, suggesting the parasite ability of adapting its sexual investment during elimination initiatives. Reactive interventions that target *Pf* sexual stages may be required to achieve complete interruption of transmission.

Introduction

Large-scale programs such as mass drug administration (MDA) which aim to quickly reduce the malaria parasite biomass in a community and to prevent new infections for a certain period can exert intensive selective pressures on these parasites.¹ While the impact of MDA on *Plasmodium falciparum* (*Pf*) genetic markers of antimalarial resistance has been assessed in few studies,² there is limited evidence about the changes in parasite population structure and transmission dynamics during the critical transition to elimination. Understanding these molecular processes can guide the design new strategies to enhance the long-term success of future programs.

There is great genetic diversity within *Pf* populations, with people living in endemic areas being frequently and simultaneously infected by several parasite strains.³ Bottlenecks in parasite population driven by elimination efforts are expected to reduce the genetic diversity of *Pf* infections and increase their similarity due to inbreeding and recent common ancestry.⁴ As genetic diversity is vital for malaria parasite' ability to evade the host immunity and adapt to changing environmental conditions,⁵ these shifts in the *Pf* population structure can affect the dynamics of infections. Theoretical^{6,7} and mouse models^{8,9} suggest that competitive interactions between genetically distinct malaria parasites within an infection impose greater demands on replication at the expenses of reproduction. Moreover, expression of *Pf* genes that regulate the parasite's replicative and reproductive programmes, such as AP2-G (the transcription factor that initiates reproduction),¹⁰ has been shown to vary depending on the intensity of malaria transmission.¹¹ Therefore, changes in the in-host ecology of *Pf* infections after drastic changes in transmission may impact the production of gametocytes and determine the transmission success of parasites escaping elimination efforts.

Here we employed parasites collected before and after the deployment of an intensive combination of interventions in Magude district (southern Mozambique)¹² to interrogate for shifts in population structure

and gametocyte carriage. We hypothesized that reductions in genetic variation, as a result of drastic reductions in transmission intensity, would lead to increases in gametocyte production due to lower within-host competition. For this genetic investigation, we utilized amplicon next-generation sequencing to evaluate genetic diversity and similarity based on the presence and relative abundance of haplotypes of two highly polymorphic *Pf* genes. Subsequently, we used a quantitative reverse transcriptase PCR (RT-qPCR) assay to detect transcripts of female (*pfs25*) and male (*pfs230*) mature gametocytes,^{13,14} as well as of *pfap2g*¹⁰ and one of its earliest responders, *pfgexp02*,¹⁵ which marks recently formed sexual forms.¹⁶

Methods

Study design and participants

The study analyzed *Pf* isolates collected in the rural district of Magude, located in the north-west of Maputo province, southern Mozambique,¹² before (2015) and after (2017) the implementation of the first phase of the Magude project. This was a malaria elimination feasibility project that consisted of two MDA rounds per year (at the beginning of the rainy season [November-December] and 4–6 weeks later) over two transmission seasons (2015-16 and 2016-17), in combination with an improved surveillance system, annual blanket indoor residual spraying, programmatically distributed long-lasting insecticide treated nets and standard case management. As a result, all-age *Pf* prevalence by rapid diagnostic test decreased from 9.1% in May 2015 to 2.6% in May 2017.¹² 50 µL dried blood spots (DBS) were collected onto Whatman 3 mm filter paper from a random selection of individuals included in an age-stratified community-based cross-sectional survey conducted at the end of the transmission season at baseline (May 2015; n = 1035) and three months after the last MDA round (May 2017; n = 3752),¹² as well as among consenting participants prior to first MDA round (November 2015, n = 1322). DBS collected in May 2015 and 2017 were also preserved in 250 µl RNAprotect (Qiagen, Hilden, Germany) and stored in -80°C. A minimum sample size of 63 and 45 per group was targeted to detect a 2-fold reduction in the proportion of monogenomic infections (assumed to be 0.5 at baseline) and mean *pfs25* relative transcript levels (1.33, standard deviation [SD] 1.12¹⁷ at baseline), respectively, with a 80% power and a two-sided significance test of 0.05.

Procedures

DNA extracted from DBS was used for *Pf* detection in duplicate by real-time quantitative PCR (qPCR) targeting 18S rRNA (**Appendix methods**). After confirming enough parasite DNA material using a nested PCR targeting *pfcr*,² *Pf* positive samples were amplified using gene-specific and indexing primers (**Appendix methods and table s1**) targeting 2 polymorphic regions of *Pf* genes coding for the circumsporozoite protein (*pfscsp*), and apical membrane antigen 1 (*pfama1*).¹⁸ Pooled libraries were sequenced in an Illumina MiSeq instrument using paired end 2×300bp reads and a MiSeq v3 (600 cycles) flow cell. The sequencing quality and performance were assessed using gDNA extracted from *Pf* strains (3D7, 7G8, Dd2 and HB3) mixed at different concentrations (**Appendix table s2**) as well as 40 *Pf* isolates

with the complexity of infection (COI) determined on the basis of the length of *pfmsp1* and *pfmsp2* amplicon sizes (**Appendix methods**). After sample demultiplexing and removal of primer sequences, *HaploTypR* pipeline version 0.3 was used to define *pfama1* and *pfmsp* haplotypes. Samples with less than 25 reads for any of the markers were removed. Reads from both amplicons were trimmed to remove bad quality bases and merged reads containing forward and reverse sequences were generated. A SNP was defined as a nucleotide position with a > 50% mismatch rate in the sequence reads using *Pf* strain 3D7 as reference from at least two independent samples. A haplotype was defined as a combination of alleles in an entire amplicon. A minimum of 3 reads and of 0.1% within-host haplotype frequency was required to call a haplotype. To discard false positive haplotypes, we removed chimeric or singleton reads as well as haplotypes only present at a frequency lower than 1%.

RNA extracted from *Pf*-positive filter papers stored in RNAprotect was subjected to reverse transcription using PrimeScript™ RT Master Mix reagents (Takara). cDNA levels of *pfs25*, *pfs230*, *pfap2g* and *pfgep02* genes, as well as the *PF08_0085* (ubiquitin-conjugating enzyme) housekeeping (HK) gene, were assessed using a RT-qPCR^{13,14,16,17} (**Appendix methods**). Samples with a Ct>40 were considered as not amplified. The transcript copy number of each gene was quantified by extrapolation against a standard curve of five serial dilutions of gDNA obtained from the 3D7 strain containing known numbers of ring-infected erythrocytes. Zeros were eliminated from the copy number data by adding to all values half of the limit of detection (LOD) of the qPCR, as calculated by the lowest concentration of target gene that can be detected in triplicate serial dilutions of 3D7 gDNA (**Appendix table s3**). The relative copy number (RCN) was calculated as the ratio of transcripts detected from target and HK genes.

Definitions and data analysis

Age was categorized as younger or equal to 15 years or older, and parasite densities quantified by qPCR as lower than 200 parasites/μL or higher. Private haplotypes were designed to those only present in pre- or post-MDA populations, and rare haplotypes to those occurring just once in the complete set of samples analyzed. Haplotypes detected were used to compute the COI (i.e. number of haplotypes present in an infection) and the presence of monogenomic infections (i.e. if a sample has one haplotype per locus). The relative abundance of a given haplotype, defined as the number of reads over all remaining reads within a sample, was used to calculate the Shannon-index¹⁹ as

$$H = - \sum_{n=i}^R p_i \log (p_i),$$

where R is the number of haplotypes in a sample and p_i is the frequency of haplotype i .

Genetic similarity of *Pf* populations was quantified through several metrics which take into account the presence of haplotypes and also their relative abundance. Binary sharing indicates whether a common haplotype exists between two samples; Jaccard distance measures the fraction of haplotypes that two

samples have in common with respect to all the haplotypes common in any of the samples; and Pearson correlation coefficient measures the linear correlation of the haplotype frequencies between two samples as

$$\rho_{X,Y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_X \sigma_Y},$$

where x and y are the haplotype frequencies of two samples, i represents the i th haplotype, and σ_X, σ_Y are the standard deviations of the two samples (**Appendix methods**).

Categorical and continuous data were compared between pre- and post-intervention using Fisher's exact test and Student t test, respectively. qPCR parasite densities and RCNs were log transformed. Logistic, negative binomial and linear regressions were used for the analysis of binary, counts and continuous data, respectively. Univariate and multivariate regression models adjusted by age and parasite density were conducted to assess the association between genetic and transcriptional variables with study period. A Jack-Knife resampling approach was used to calculate the variance of each similarity metric (binary sharing, Jaccard distance and Pearson CC) as well as of the fraction of sample pairs above different similarity thresholds. From this, error propagation was used to obtain the error of the differences between groups. Statistical tests were performed using Stata software version 15 (Stata Corp, College Station, Tx) and Python 3 programming language. The statistical significance was defined as a p value < 0.05.

Ethics

Study ethical approval was obtained from the National Mozambican Ethical Review Committee (Mozambique) and Hospital Clínic (Barcelona, Spain) ethics review committees, and signed written informed consent was obtained from all participants or from guardians/parents in the case of minors.

Results

Study population and samples analyzed

Pf infection was confirmed by qPCR in 408 (6.7%) of the 6109 residents in Magude District from whom DBS were collected (101 out of 1035 [9.8%] in May 2015; 168 out of 1322 [12.7%] in November 2015, and 139 out of 3752 [3.7%] in May 2017). Ninety-nine (58.9%) and 107 (76.9%) of the *Pf*-positive samples from November 2015 and May 2017, respectively, were available for *pfama1* and *pfmsp* sequencing after confirming enough parasite DNA material through a nested PCR targeting *pfprt*. Among these samples, 85 (85.9%) and 93 (86.9%) amplified both genes and were included in the library for deep sequencing. Seventeen samples were excluded due to having less than 25 reads either in *pfama* or *pfmsp*, and 2 were further discarded due to incomplete demographic data (Fig. 1). Ninety (89.1%) and 80 (57.6%) of the *Pf*-positive samples from May 2015 and May 2017, respectively, were available in RNA protect and therefore used for the analysis of gametocyte-specific markers. Among these, 51 (56.7%) and 57 (71.2%) amplified

the HK gene (Fig. 1), and one pre-intervention sample was excluded due to incomplete demographic data. Study participants before and after the intervention from whom genetic and transcriptional data was available were similar in area of residence, age, sex and qPCR-quantified *Pf* densities (Table 1). Also, the selected participants were representative of the overall population of *Pf*-infected individuals detected by qPCR, except for parasite densities which were higher among those with successfully analyzed *Pf* infections (**Appendix table s4**).

Table 1
Main characteristics of study participants.

	Genetic structure			Gametocyte markers		
	Nov 2015	May 2017		May 2015	May 2017	
	(n = 78)	(n = 81)	p	(n = 50)	(n = 57)	p
Period	Nov 2015	May 2017		May 2015	May 2017	
Study Area, n (%)						
Magude Sede	47 (60)	45 (55)	0.654	34 (68)	31 (54)	0.099
Motaze	14 (18)	15 (19)		11 (22)	12 (21)	
Panjane	10 (13)	8 (10)		5 (10)	6 (11)	
Mahele	2 (3)	6 (7)		0 (0)	3 (5)	
Mapulanguene	5 (6)	7 (9)		0 (0)	5 (9)	
Age, Mean (SD)	14.2 (1,2)	11.5 (1.7)	0.198	8.5 (10.9)	10.5 (12.8)	0.400
Gender, n (%)						
Female	39 (50)	40 (49)	1.000	22 (44)	29 (51)	0.562
Male	39 (50)	41 (51)		28 (56)	28 (49)	
Pf density, GM (SD)	106.3 (254.6)	165.4 (535.1)	0.330	169.2 (466.3)	508.2 (1537.8)	0.053
<i>Pf</i> density as assessed by qPCR; GM, Geometric mean; SD, Standard deviation						

Pf haplotype distribution

Across all successfully genotyped infections with complete demographic data (159 out of 178 [89%]), read coverage was similar in samples collected in 2015 and 2017 (**Appendix table s5**), with a mean of 6221 (SD 5416) and 7965 (SD 5908) reads for *pfama1* and *pfmsp*, respectively. Parasite densities of samples that were successfully genotyped tended to have higher parasite densities (110 parasites/μl, IQR: 17–825), compared to those that failed the genotyping (41 parasites/μl, IQR: 5–798, $p = 0.089$; **Appendix figure s1**). After haplotype assignment and quality filtering of reads, we identified 30 SNPs for

pfama1 and 24 for *pfensp*. In total, 45 distinct haplotypes were detected at the *pfama1* locus and 57 at the *pfensp* locus (Fig. 2A and B), with 9 *pfama1* haplotypes and 14 *pfensp* haplotypes being detected only once (rare haplotypes). The number of haplotypes detected per locus was higher in samples collected from 2015 (43 for *pfama1* and 49 for *pfensp*) than in samples from 2017 (23 for *pfama1* and 40 for *pfensp*). Out of the 45 *pfama1* haplotypes, 22 were shared in both periods and 23 were private of study period (21 in 2015 and 2 in 2017). Out of the 57 *pfensp* haplotypes, 32 were shared and 25 were private (17 in 2015 and 8 in 2017). The proportion of *Pf* isolates with rare haplotypes was 10% (8 out of the 78) in 2015 and 0.1% (1 out of 81) in 2017 ($p = 0.039$) for *pfama1* and 18% (13 out of the 78) in 2015 and 5% (4 out of 81) in 2017 ($p = 0.024$; Fig. 2C). Pre-intervention private haplotypes were detected in less samples than shared haplotypes (Fig. 2D).

Within-host *Pf* diversity

Most study participants had polygenomic infections, with single haplotypes detected in 48% (76/159; *pfama1*) and 33% (52/159; *pfensp*) of genotyped infections. Of 83 participants with polygenomic infection at the *pfama1* locus, 67 (81%) also had more than one *pfama1* haplotype (and vice versa—63%). Additionally, within individual participants, the number of *pfama1* and *pfensp* haplotypes was highly correlated (**Appendix figure s2**; $\rho = 0.83$, $p < 0.001$, Spearman Rank test). The mean number of *pfama1* haplotypes detected per sample was lower in 2017 (1.9, SD: 1.4) than in 2015 (2.6, SD: 2.2; $p = 0.012$; Fig. 3A). Such a difference was not observed for *pfensp* haplotypes (3.0, SD: 0.23 in 2015 and 3.1, SD: 0.32 in 2017; $p = 0.734$). Mean Shannon index was lower in 2017 than in 2015 for both loci (*pfama1*: 0.36, SD: 0.05 in 2015 and 0.14, SD: 0.04 in 2017, $p = 0.002$; *pfensp*: 0.43, SD: 0.06 in 2015 and 0.26, SD: 0.05 in 2017, $p = 0.018$; Fig. 3A and B). The proportion of *pfama1* monogenomic infections was higher in 2017 (56%; 45/81) than in 2015 (40%, 31/78; $p = 0.047$). A similar trend, although not statistically significant, was observed for *pfensp* (29% [23/78] in 2015 and 36% [29/81] in 2017, $p = 0.397$; Fig. 3B). No statistically significant association was observed with area of sample collection and age of study participants (**Appendix table s6**), except for the proportion of *pfensp* monogenomic infections which were higher among the older individuals (18/37; 49%) than among younger ones 34/122, 28%, $p = 0.027$; **Appendix table s7-8**). The differences in within-host diversity metrics were maintained after adjusting the regression analysis by parasite density and age of study participants (**Appendix table s8**).

Between-infection *Pf* similarity indices

We employed 3 metrics that describe genetic similarity between pairs of *Pf* infections: binary sharing, Jaccard distance and Pearson CC. There was no evidence of statistically significant differences in the proportion of sample pairs that shared at least one haplotype (binary sharing; **Appendix figure s3** and **Appendix table s9**). Pearson CC and Jaccard distance (Fig. 4A, B, C and D) of *pfama1* increased from 2015 to 2017. Similarly, the fraction of sample pairs with high Pearson CC (> 0.9999) increased after the intervention (Fig. 4E) for both *pfama1* (from 1.03–4.97%, $p < 0.001$) and *pfensp* (from 1.03–3.4%, $p = 0.019$). Similar trends but not statistically significant were observed for highly similar parasites defined

on the basis of Jaccard distance (Fig. 4F). No correlation was observed between genetic relatedness and gender, area, use of antimalarial, age and parasite density (**Appendix figure s4**).

Pf gametocyte-specific transcripts

Pf isolates with RT-qPCR and complete demographic data ($n = 107$) had higher parasite densities (546 parasites/ μ L, IQR: 23-2243) than those that failed (5, IQR: 3-1180, $p < 0.001$; **Appendix figure s1**). The proportion of *Pf* isolates with measurable gametocyte-specific transcripts ranged between 94% (101/107) for *pfs25* and 61% (46/107) for *pfs230*, with intermediate values for *pfap2g* (75%; 80/107) and *pfexp02* (69%; 74/107) (Fig. 5A and **Appendix table s10**). Relative levels of female-specific gametocyte transcripts (*pfs25*) increased from 2015 to 2017, as did the proportion of parasites showing detectable *pfap2g* and *pfexp02* transcripts and their levels (Fig. 5A and **B** and **Appendix table s10**). The statistical significance was maintained in multivariate regression models (**Appendix table s11**). Both the univariate and the multivariate regression models (**Appendix table s11**) showed that *pfap2g* and *pfexp02* transcript levels increased with increasing parasite densities, while the opposite trend was observed for *pfs25*. No difference was observed by the area where the *Pf* isolates were collected (**Appendix table s6**).

Discussion

The modulation of *Pf* reproductive investment in response to changes in transmission is expected from theory^{6,7} and from the effect of environmental stressors on gametocyte conversion rates in experimental models²⁰. Here we show that a 71% drop in *Pf* prevalence achieved in 2 years through the scale up of intervention in southern Mozambique is associated with reductions in *Pf* genetic diversity and increases in parasite similarity, overall indicating a severe bottleneck in the parasite population. To the best of our knowledge, this investigation presents for the first time evidences of an increased expression of gametocyte-specific genes accompanying this shift in parasite genetic structure. These changes in the *Pf* transmission potential during the transition to elimination may affect the success of subsequent interventions aiming to achieve complete interruption of transmission.

Genetic diversity of the *Pf* population in Magude district decreased after deploying an intensive package of interventions, as indicated by the consistent decline in *pfama1* and *pfmsp* Shannon diversity index, as well as the reductions in *pfama1* COI and frequency of polygenomic infections. This decline in genetic diversity was not due to differences in sequencing depth, as parasite densities and sequencing coverage were similar in both study periods. The decrease in the overall number of detected haplotypes and the loss of the rarest haplotypes rather suggests that the severe bottleneck in the parasite population probably drove this reduction in genetic diversity. Drops in diversity were accompanied by enhanced between-infection genetic similarity in parasites surviving the intervention, as evidenced by the proportion of highly-similar parasite pairs (Pearson CC > 0.95) which increased from 2–8% for *pfama1* and from 3–6% for *pfmsp*. Our results suggest that genetic parameters may be useful to monitor the success interventions deployed at the community level,²¹ especially when the parameters take into account the relative abundance of haplotypes (i.e., Shannon index and Pearson CC). Importantly, less than 100

samples per group were enough to detect signals of a bottleneck in the parasite population driven by a 71% drop in *Pf* prevalence.¹² Moreover, the enhanced genetic similarity after the bottleneck in the parasite population may increase the performance of genetic tools to identify imported infections with unrelated haplotypes.²²

The reduction in *Pf* within-host genetic diversity and the enhanced between-infection similarity of infections after the drastic drop in transmission were coincident with an increase in the detectability and levels of *Pf* transcripts specifically expressed by female gametocytes (*pfs25*), sexually committed (*pfap2g*) and recently formed gametocytes (*pfgexp02*). Similar increases in gametocyte carriage have been described after reductions in transmission achieved through the use of a combination of interventions, including insecticide spraying and MDA, in Nigeria (Garki project) in the 1970s²³ and in Netherlands New Guinea in the 1950s.²⁴ Reductions in the within-host diversity might have contributed to this plastic response if less competitive interactions eventually favors gametocyte investment at the expenses of asexual replication.^{6,7,25} This could partly explain the low parasite densities in historical endemic areas that have experienced transmission reductions, but where immunity is still present, especially among adults.²⁶ Such counter-adaptation would allow the parasite to recover from severe bottlenecks by ensuring the continuation of some level of anti-disease immunity (therefore less detection and less treatment), while at the same time betting on its transmissibility to increase the population size. Alternatively, higher gametocyte conversion rates may have been selected by reductions in the Anopheline population driven by vector campaigns deployed during the elimination initiative, such as intra-door residual spraying. Further studies are required to assess if changes in sexual differentiation after sharp declines of malaria transmission are mainly driven by plastic responses or adaptive selection of *Pf* genes such as *pfgdvr1*^{27–29}. Independently of the mechanisms used by the parasite, the higher investment in gametocyte production observed in this study suggest that reactive focal MDA approaches with gametocytocidal antimalarials such as primaquine, coupled with glucose-6-phosphate dehydrogenase deficiency testing,³⁰ may increase the impact of strategies aiming to interrupt or prevent the re-establishment of malaria transmission.³¹

The results of this study are subjected to several limitations. First, *Pf* infections that were successfully analyzed correspond to those in the upper range of parasite densities, given the limit of detection of the transcriptional and sequencing approaches.^{13,18} Second, the similarity metrics used were based on the coincidence of individual haplotypes between infections rather than on identity by descent. Third, we cannot discard that other factors may have contributed to this reduction in within-host genetic diversity, such as expansion of a small number of alleles with increased success in the new transmission condition, a shorter time to acquire multiple infections after the elimination initiative or selection of *pfama1* and *pfcsp* specific variants due to changes in host's immunity.³² However, the contribution of these other factors were mitigated in our study owing to the restricted temporal scale of sampling and the abundance of unique haplotypes at each locus. Fourth, a few parasite isolates remained for genetic analysis after the marked reduction in the incidence of *Pf* malaria in Magude, although numbers were enough to detect changes in diversity and transcriptional markers. Finally, the study focused on the short-

term impact of the intervention and further studies are needed to determine the duration of these changes.

In summary, this is the first study to show that when malaria control programmes lead to a drastic reduction in transmission intensity, diversity declines while similarity and gametocyte production increases, supporting the concept that *Pf* can facultatively trade off investment in sexual transmissible stages to maximize their within-host competitive ability.^{8,9} If associated with increased infectivity, this higher carriage of sexual stages may increase the ability of infections that survive elimination efforts to expand and recover to baseline levels, unless interventions that target the *Pf* sexual stages are deployed to sustain gains and eventually achieve interruption of transmission. The identification of biological mechanisms involved in these increases in gametocyte production can allow the design of new approaches for early programmatic reactions in settings approaching elimination.

Declarations

Acknowledgements

We would like to express our gratitude to the community of Magude and the district authorities, for allowing this project to take place and for greatly contributing to its successful implementation and evaluation. We would also like to thank all the field workers, the staff of the Hospitals, clinical officers, field supervisors, data manager and those who worked in the lab for their contribution to the collection and organization of samples shipment. We also thank Mariona Bustamante, Jochen Hecht and Anita Lerch for their guidance on the sequencing of samples and analysis of data using HaplotypeR, and Alfred Cortes and Harvie Portugaliza for providing the primer sequence to amplify *pfap2g* and *pfgexp002*. Finally, we thank the Mozambican National Malaria Control Program and the Ministry of Health for supporting and providing guidance throughout the project. The following reagents were obtained through BEI Resources, NIAID, NIH: *Plasmodium falciparum*, Strain 3D7A (MRA-151), HB3 (MRA-155), Dd2 (MRA-156) and 7G8 (MRA-152), contributed by David Walliker and Thomas E. Wellems.

Funding

We would like to thank La Caixa and Bill and Melinda Gates Foundations for providing the funds for this study (grant number INV-008483/OPP1115265). A.M. is supported by the Departament d'Universitats i Recerca de la Generalitat de Catalunya, Agència de Gestió d'Ajuts Universitaris i de Recerca (2017SGR664). HG was supported (January 2017-January 2019) by the Science and Engineering Research Board (SERB), Department of Science & Technology, Government of India (Overseas Postdoctoral Fellowship, SB/OS/PDF043/2015-16). We acknowledge support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. The Centro de Investigaçao em Saude de Manhica (CISM) is supported by the Government of Mozambique and the Spanish Agency for International Development Cooperation (AECID). We also acknowledge

support from the Spanish Ministry of Science and Innovation through the “Centro de Excelencia Severo Ochoa 2019-2023” Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program This research is part of ISGlobal’s Program on the Molecular Mechanisms of Malaria, which is partially supported by the Fundación Ramón Areces. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors’ contributions

HG carried out the molecular determinations, and together with CR, JRG and AP analyzed the data. PC, AC, GM contributed to the analysis of samples. FS, PA, BG, PLA, LN, NRR, WS participated in the design of the interventions, fieldwork and collected clinical and epidemiological data together with the biological samples. AM, HG and BG participated in the study design. AM conceptualized the study, interpreted the results and wrote the draft of this manuscript. All authors read and approved the final manuscript.

Data availability

Anonymized data used for this analysis is available upon request.

Competing interests

The authors declare that they have no competing interests.

References

1. Webster, J. P., Molyneux, D. H., Hotez, P. J. & Fenwick, A. The contribution of mass drug administration to global health: past, present and future. *Philos Trans R Soc Lond B Biol Sci* <background-color:#FFCC66;bu>369</background-color:#FFCC66;bu>, 20130434, doi:10.1098/rstb.2013.0434 (2014).
2. Gupta, H. *et al.* Effect of mass dihydroartemisinin–piperaquine administration in southern Mozambique on the carriage of molecular markers of antimalarial resistance *PLoS One* **In press** (2020).
3. Read, A. F. & Taylor, L. H. The ecology of genetically diverse infections. *Science* <background-color:#FFCC66;bu>292</background-color:#FFCC66;bu>, 1099–1102, doi:10.1126/science.1059410 (2001).
4. Daniels, R. F. *et al.* Evidence for Reduced Malaria Parasite Population after Application of Population-Level Antimalarial Drug Strategies in Southern Province, Zambia. *Am J Trop Med Hyg*, doi:10.4269/ajtmh.19-0666 (2020).
5. Crompton, P. D. *et al.* Malaria immunity in man and mosquito: insights into unsolved mysteries of a deadly infectious disease. *Annu Rev Immunol* <background-color:#FFCC66;bu>32</background-color:#FFCC66;bu>, 157–187, doi:10.1146/annurev-immunol-032713-120220 (2014).

6. Reece, S. E., Ramiro, R. S. & Nussey, D. H. Plastic parasites: sophisticated strategies for survival and reproduction? *Evol Appl* <background-color:#FFCC66;bu>2</background-color:#FFCC66;bu>, 11–23, doi:10.1111/j.1752-4571.2008.00060.x (2009).
7. Greischar, M. A., Mideo, N., Read, A. F. & Bjornstad, O. N. Predicting optimal transmission investment in malaria parasites. *Evolution* <background-color:#FFCC66;bu>70</background-color:#FFCC66;bu>, 1542–1558, doi:10.1111/evo.12969 (2016).
8. Pollitt, L. C. *et al.* Competition and the evolution of reproductive restraint in malaria parasites. *Am Nat* <background-color:#FFCC66;bu>177</background-color:#FFCC66;bu>, 358–367, doi:10.1086/658175 (2011).
9. Reece, S. E., Drew, D. R. & Gardner, A. Sex ratio adjustment and kin discrimination in malaria parasites. *Nature* <background-color:#FFCC66;bu>453</background-color:#FFCC66;bu>, 609–614, doi:10.1038/nature06954 (2008).
10. Kafsack, B. F. *et al.* A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* <background-color:#FFCC66;bu>507</background-color:#FFCC66;bu>, 248–252, doi:10.1038/nature12920 (2014).
11. Rono, M. K. *et al.* Adaptation of *Plasmodium falciparum* to its transmission environment. *Nat Ecol Evol* <background-color:#FFCC66;bu>2</background-color:#FFCC66;bu>, 377–387, doi:10.1038/s41559-017-0419-9 (2018).
12. Galatas, B. *et al.* A multiphase program for malaria elimination in southern Mozambique (the Magude project): A before-after study. *PLoS Med* <background-color:#FFCC66;bu>17</background-color:#FFCC66;bu>, e1003227, doi:10.1371/journal.pmed.1003227 (2020).
13. Schneider, P. *et al.* Quantification of female and male *Plasmodium falciparum* gametocytes by reverse transcriptase quantitative PCR. *Mol Biochem Parasitol* <background-color:#FFCC66;bu>199</background-color:#FFCC66;bu>, 29–33, doi:10.1016/j.molbiopara.2015.03.006 (2015).
14. Aguilar, R. *et al.* Molecular evidence for the localization of *Plasmodium falciparum* immature gametocytes in bone marrow. *Blood* <background-color:#FFCC66;bu>123</background-color:#FFCC66;bu>, 959–966, doi:10.1182/blood-2013-08-520767 (2014).
15. Silvestrini, F. *et al.* Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* <background-color:#FFCC66;bu>9</background-color:#FFCC66;bu>, 1437–1448, doi:10.1074/mcp.M900479-MCP200 (2010).
16. Portugaliza, H. P., Llorca-Batlle, O., Rosanas-Urgell, A. & Cortes, A. Reporter lines based on the gexp02 promoter enable early quantification of sexual conversion rates in the malaria parasite *Plasmodium falciparum*. *Sci Rep* <background-color:#FFCC66;bu>9</background-color:#FFCC66;bu>, 14595, doi:10.1038/s41598-019-50768-y (2019).
17. Gupta, H. *et al.* Differential expression of var subgroups and PfSir2a genes in afebrile *Plasmodium falciparum* malaria: a matched case-control study. *Malar J* <background-

- color:#FFCC66;bu>18</background-color:#FFCC66;bu>, 326, doi:10.1186/s12936-019-2963-z (2019).
18. Lerch, A. *et al.* Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections. *BMC Genomics* <background-color:#FFCC66;bu>18</background-color:#FFCC66;bu>, 864, doi:10.1186/s12864-017-4260-y (2017).
 19. Shannon, C. E. The mathematical theory of communication. 1963. *MD Comput* **14**, 306–317 (1997).
 20. Bousema, T. & Drakeley, C. Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. *Clin Microbiol Rev* <background-color:#FFCC66;bu>24</background-color:#FFCC66;bu>, 377–410, doi:10.1128/CMR.00051-10 (2011).
 21. Watson, O. J. *et al.* Evaluating the performance of malaria genetics for inferring changes in transmission intensity using transmission modelling. *Mol Biol Evol*, doi:10.1093/molbev/msaa225 (2020).
 22. Wesolowski, A. *et al.* Mapping malaria by combining parasite genomic and epidemiologic data. *BMC Med* <background-color:#FFCC66;bu>16</background-color:#FFCC66;bu>, 190, doi:10.1186/s12916-018-1181-9 (2018).
 23. Molineaux, L., Storey, J., Cohen, J. E. & Thomas, A. A longitudinal study of human malaria in the West African Savanna in the absence of control measures: relationships between different *Plasmodium* species, in particular *P. falciparum* and *P. malariae*. *Am J Trop Med Hyg* <background-color:#FFCC66;bu>29</background-color:#FFCC66;bu>, 725–737, doi:10.4269/ajtmh.1980.29.725 (1980).
 24. Metselaar, D. Relative increase in the prevalence of *Plasmodium falciparum* some years after the beginning of a house-spraying campaign in Netherlands New Guinea. *Trans R Soc Trop Med Hyg* <background-color:#FFCC66;bu>54</background-color:#FFCC66;bu>, 523–528, doi:10.1016/0035-9203(60)90026-2 (1960).
 25. Mackinnon, M. J. & Read, A. F. Virulence in malaria: an evolutionary viewpoint. *Philos Trans R Soc Lond B Biol Sci* <background-color:#FFCC66;bu>359</background-color:#FFCC66;bu>, 965–986, doi:10.1098/rstb.2003.1414 (2004).
 26. Slater, H. C. *et al.* The temporal dynamics and infectiousness of subpatent *Plasmodium falciparum* infections in relation to parasite density. *Nat Commun* <background-color:#FFCC66;bu>10</background-color:#FFCC66;bu>, 1433, doi:10.1038/s41467-019-09441-1 (2019).
 27. Duffy, C. W. *et al.* Multi-population genomic analysis of malaria parasites indicates local selection and differentiation at the *gdv1* locus regulating sexual development. *Sci Rep* <background-color:#FFCC66;bu>8</background-color:#FFCC66;bu>, 15763, doi:10.1038/s41598-018-34078-3 (2018).

28. Mobegi, V. A. *et al.* Genome-wide analysis of selection on the malaria parasite *Plasmodium falciparum* in West African populations of differing infection endemicity. *Mol Biol Evol* <background-color:#FFCC66;bu>31</background-color:#FFCC66;bu>, 1490–1499, doi:10.1093/molbev/msu106 (2014).
29. Usui, M. *et al.* *Plasmodium falciparum* sexual differentiation in malaria patients is associated with host factors and GDV1-dependent genes. *Nat Commun* <background-color:#FFCC66;bu>10</background-color:#FFCC66;bu>, 2140, doi:10.1038/s41467-019-10172-6 (2019).
30. Galatas, B. *et al.* Heterogeneity of G6PD deficiency prevalence in Mozambique: a school-based cross-sectional survey in three different regions. *Malar J* <background-color:#FFCC66;bu>16</background-color:#FFCC66;bu>, 36, doi:10.1186/s12936-016-1674-y (2017).
31. Marasinghe, M. M. *et al.* Mass radical treatment of a group of foreign workers to mitigate the risk of re-establishment of malaria in Sri Lanka. *Malar J* <background-color:#FFCC66;bu>19</background-color:#FFCC66;bu>, 346, doi:10.1186/s12936-020-03419-x (2020).
32. Escalante, A. A., Cornejo, O. E., Rojas, A., Udhayakumar, V. & Lal, A. A. Assessing the effect of natural selection in malaria parasites. *Trends Parasitol* <background-color:#FFCC66;bu>20</background-color:#FFCC66;bu>, 388–395, doi:10.1016/j.pt.2004.06.002 (2004).

Figures

Figure 1

Study profile, sequencing and transcriptional analysis. Real time quantitative PCR positive subjects were selected for transcriptional and sequence analysis after the detection of Pf infections in 6109 dried blood spots. 22% samples failed haplotype assignment at both loci due to lack of amplification product or low number of sequencing reads. Sequencing data were excluded for 2 study participants due to incomplete demographic data. 36% samples failed transcriptional analyses due to lack of amplification of the housekeeping gene.

Figure 1

Study profile, sequencing and transcriptional analysis. Real time quantitative PCR positive subjects were selected for transcriptional and sequence analysis after the detection of Pf infections in 6109 dried blood spots. 22% samples failed haplotype assignment at both loci due to lack of amplification product or low number of sequencing reads. Sequencing data were excluded for 2 study participants due to incomplete demographic data. 36% samples failed transcriptional analyses due to lack of amplification of the housekeeping gene.

Figure 2

Distribution of *pfama1* and *pfmsp* haplotypes in study samples. A) Distribution of haplotype frequencies for *pfama1* and *pfmsp* in samples from 2015 and 2017. The y-axes represent the haplotype number, the x-axes the sample number, and the colour represents the haplotype frequency, ranging from 0 to 1. Each column represents the haplotype distribution for a given sample. Black points represent samples with one single haplotype, and coloured points are present in samples with several haplotypes; B) Number of *pfama1* and *pfmsp* haplotypes detected in samples; C) Number of pre-intervention samples which contain haplotypes that are not detected in 2017 or shared at both timepoints; D) Proportion of haplotypes only

detected once in the samples, by study period. Red lines indicate mean, T bars the 95%CI and p value of the t Student's test.

Figure 2

Distribution of *pfama1* and *pfmsp* haplotypes in study samples. A) Distribution of haplotype frequencies for *pfama1* and *pfmsp* in samples from 2015 and 2017. The y-axes represent the haplotype number, the x-axes the sample number, and the colour represents the haplotype frequency, ranging from 0 to 1. Each column represents the haplotype distribution for a given sample. Black points represent samples with one single haplotype, and coloured points are present in samples with several haplotypes; B) Number of

pfama1 and pfmsp haplotypes detected in samples; C) Number of pre-intervention samples which contain haplotypes that are not detected in 2017 or shared at both timepoints; D) Proportion of haplotypes only detected once in the samples, by study period. Red lines indicate mean, T bars the 95%CI and p value of the t Student's test.

Figure 3

P. falciparum genetic diversity in 2015 and 2017. Complexity of the infection (number of different haplotypes in a *P. falciparum* isolate), Shannon index (diversity index that considers the number of haplotypes in a sample and the frequency of the haplotype) and monoclonal infections (isolates with only one haplotype) detected before and after the intervention assessed for A) pfama1 and B) pfmsp. Red lines indicate geometric mean, T bars the 95%CI and asterisk if p value was below 0.05 in the multivariate analysis adjusted by parasite densities and age.

Figure 3

P. falciparum genetic diversity in 2015 and 2017. Complexity of the infection (number of different haplotypes in a *P. falciparum* isolate), Shannon index (diversity index that considers the number of haplotypes in a sample and the frequency of the haplotype) and monoclonal infections (isolates with only one haplotype) detected before and after the intervention assessed for A) *pfama1* and B) *pfcsp*. Red lines indicate geometric mean, T bars the 95%CI and asterisk if *p* value was below 0.05 in the multivariate analysis adjusted by parasite densities and age.

Figure 4

P. falciparum genetic similarity in 2015 and 2017. Heat maps of Pearson correlation coefficient (PCC) (A) and Jaccard distance (JD) (B) between sample pairs for *pfama1* and *pfmsp* in 2015 and 2017. Jaccard distance takes into account all the haplotypes present in the sample pair to measure the proportion of haplotypes that they share, while the Pearson CC also takes into account the frequency of the haplotypes present in the samples to quantify the linear correlation between the sample pairs. Mean PCC (C) and JD (D) by study period was calculated for all sample pairs and standard deviation (T bar) through a Jack-Knife resampling scheme. The fraction of sample pairs with relatedness above different thresholds (E for PCC and F for JD) were compared in samples from 2015 and 2017.

Figure 4

P. falciparum genetic similarity in 2015 and 2017. Heat maps of Pearson correlation coefficient (PCC) (A) and Jaccard distance (JD) (B) between sample pairs for *pfama1* and *pfmsp* in 2015 and 2017. Jaccard distance takes into account all the haplotypes present in the sample pair to measure the proportion of haplotypes that they share, while the Pearson CC also takes into account the frequency of the haplotypes present in the samples to quantify the linear correlation between the sample pairs. Mean PCC (C) and JD (D) by study period was calculated for all sample pairs and standard deviation (T bar) through a Jack-Knife resampling scheme. The fraction of sample pairs with relatedness above different thresholds (E for PCC and F for JD) were compared in samples from 2015 and 2017.

Figure 5

P. falciparum gametocyte-specific transcripts in 2015 and 2017. A) Rate of *P. falciparum* isolates with transcripts detectable by RT-qPCR and B) relative copy number (RCN) of transcripts in Pf isolates collected detected before (2015) and after (2017) the interventions. Gene transcripts were considered to be absent if fluorescence was not detected at cycle 40 ($C_t > 40$). Red lines indicate geometric mean, T bars the 95%CI and asterisk if p value was below 0.05 in the multivariate analysis adjusted by parasite densities and age.

Figure 5

P. falciparum gametocyte-specific transcripts in 2015 and 2017. A) Rate of *P. falciparum* isolates with transcripts detectable by RT-qPCR and B) relative copy number (RCN) of transcripts in Pf isolates collected detected before (2015) and after (2017) the interventions. Gene transcripts were considered to be absent if fluorescence was not detected at cycle 40 ($C_t > 40$). Red lines indicate geometric mean, T bars the 95%CI and asterisk if p value was below 0.05 in the multivariate analysis adjusted by parasite densities and age.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [MDAmolecularSI14thOctober2020.docx](#)
- [MDAmolecularSI14thOctober2020.docx](#)