Natural history of malaria infections during early childhood in twins

Bronner P. Gonçalves[1‡], Raúl Pérez-Caballero[1‡], Amadou Barry[2], Santara Gaoussou[2], Alexandra Lewin[3], Djibrilla Issiaka[2], Sekouba Keita[2], Bacary S. Diarra[2], Almahamoudou Mahamar[2], Oumar Attaher[2], David L. Narum[1], Jonathan D. Kurtis[4], Alassane Dicko[2], Patrick E. Duffy[1], Michal Fried[1*]

[1] Laboratory of Malaria Immunology and Vaccinology, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), Bethesda, US

[2] Malaria Research and Training Center, University of Sciences, Techniques and Technologies of Bamako, Bamako, Mali

[3]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, United Kingdom

[4]Center for International Health Research, Rhode Island Hospital, and Department of Pathology and Laboratory Medicine, Brown University Medical School, Providence, Rhode Island, USA

[‡]Authors BPG and RP contributed equally to this work

Running Title: Malaria infection and disease in twins

***Corresponding Author:** michal.fried@nih.gov; 240.747.7880

Abstract

Background: The frequency and clinical presentation of malaria infections show marked heterogeneity in epidemiological studies. However, deeper understanding of this variability is hampered by the difficulty in quantifying all relevant factors. Here, we report the history of malaria infections in twins, who are exposed to the same *in utero* milieu, share genetic factors and are similarly exposed to vectors.

Methods: Data were obtained from a Malian longitudinal birth cohort. Samples from 25 twin pairs were examined for malaria infection and antibody responses. Bayesian models were developed on the number of infections during follow-up.

Results: In 16/25 pairs, both children were infected and often developed symptoms. In 8/25 pairs, only one twin was infected, but usually only once or twice. Statistical models suggest this pattern is not inconsistent with twin siblings having the same underlying infection rate. In a pair with discordant hemoglobin genotype, parasite densities were consistently lower in the child with hemoglobin AS, but antibody levels were similar.

Conclusions: By using a novel design, we describe residual variation in malaria phenotypes in naturally-matched children and confirm the important role of environmental factors, as suggested by the between twin pair heterogeneity in malaria history.

Key Words: malaria, infection, twins, pathogenesis, hemoglobin S

Introduction

Children living in endemic areas differ in at least three important aspects with regard to malaria infection history [1, 2]: the number of falciparum infections experienced; the levels of parasitemia when blood-stage infections are established; and the clinical outcomes of these infections. Several factors have been suggested to contribute to these differences. For example, host genetic factors [3-5] are thought to influence clinical expression and differences in parasitemia. Factors that modify *in utero* environment have also been linked to between-child differences in malaria outcomes [6, 7]. Although many epidemiological studies aim to quantify these different types of variations and identify their causes, the potentially high number of important host factors complicates the understanding of the different and common processes that lead to variability in these phenotypes.

Here, we propose that comparing the history of malaria infections in twins throughout early childhood would generate new insights into malaria pathogenesis. Twins are exposed to the same *in utero* milieu, and share genetic factors, either partially or entirely, when monozygotic. Furthermore, in some areas, twins often sleep together throughout childhood, implying similar exposure to malaria vectors. By studying the frequency of infections, parasite levels and clinical malaria presentations in twins, we learn about residual variability of malaria phenotypes in children naturally matched for disease-modifying factors and exposure.

Methods

Study population and clinical procedures

The longitudinal cohort study of mother-infant pairs was conducted from February 2011 to July 2016 in the health district of Ouelessebougou located 80 km south of Bamako, Mali, an area of intense seasonal malaria transmission. Study protocol was approved by the institutional review boards at the National Institute of Allergy and Infectious Diseases, National Institutes of Health, and by the Ethics Committee of the Faculty of Medicine, Pharmacy and Dentistry, University of Bamako. Written informed consent was obtained from the study participants or the parents/guardians of pregnant adolescent after receiving a study explanation form and oral explanation from a study clinician in their native language. The protocol is registered at Clinicaltrials.gov under identifier NCT01168271. Children were followed from birth up to 5 years. Scheduled and unscheduled visits included clinical examination and blood smear microscopy for the detection of malaria parasites. Children were seen monthly during the malaria transmission season, every two months during the dry season; note (i) that during the study period before August 2013 follow-up was more frequent for children aged less than 24 months, who were seen every two weeks during the transmission season, and (ii) the same visit scheduled applied to children within pairs. Throughout the study, any time a child in the study was sick, she or he would also be seen (unscheduled visit). Severe malaria was defined as parasitemia detected by blood smear microscopy and one of the following World Health Organization definitions of severe malaria: coma (Blantyre score below 3), two or more convulsions in the past 24 hours prostration (inability to sit unaided or in younger infants inability to move/feed); hemoglobin below 6 g/dL; respiratory

distress (hyperventilation with deep breathing, intercostal recessions and/or irregular breathing).

Antibody detection by ELISA

Blood samples were taken, and plasma was used to detect total IgG against AMA-1 and MSP1[8, 9]. ELISA assays were performed using Immulon® 4HBX (high binding), flat-bottom, 96-well microtiter plates (Thermo Scientific, Rochester, NY). Briefly, plates were coated with 200 ng per well (50 µl/well) of each one of the antigens, diluted in 0.05M carbonate–bicarbonate buffer pH 9.6 and incubated at 4 °C overnight. Plates were blocked with 120 µl/well of blocking buffer containing 5% skim milk diluted in tris buffered saline (TBS, Fisher Scientific®) and incubated at RT for 1 h. Wells were washed with TBS 0.05% Tween 20 and 50 µl/well of plasma diluted in antibody diluting buffer were added in duplicate and incubated at RT for 1 h. After washing, 50 µl/well of goat anti-human antibody (KPL, SeraCare®) were added at 1:300 and incubated at RT for 1 h. After the plates were washed, 50 µl/well of phosphatase substrate (Sigma Aldrich®) diluted in coating buffer were added and incubated at RT for 20 min. Optical density was measured at 405 nm using a microplate photometer (Infinite® M200 PRO, Tecan).

Anti-PfGARP and anti-PfSEA-1 antibody assays

To measure IgG anti-rPfGARP (glutamic acid rich protein) and anti-rPfSEA-1(Schizont Egress Antigen-1) antibody levels, we developed a bead-based assay according to previous published methods[10, 11]. In brief, the beads in 96-well plates were washed three times with ABE to remove particulates. Plasma samples were diluted 1:80 in ABE buffer and 50 µl

of diluted sample was added to the beads and incubated in the dark on a shaking platform at 1000 rpm for 30 seconds followed by 300 rpm for 30 minutes at RT. These conditions were used in successive incubations. Excess antibody was removed using a vacuum manifold followed by three washes in ABE. Twenty-five µl of biotinylated human IgG detection antibody (Pharmingen) diluted 1:500 in ABE was added to the beads and incubated as indicated. Fifty µl of phycoerythrin-conjugated streptavidin (BD PharmingenTM, Franklin Lakes, NJ) diluted 1:500 in ABE was added to the beads and incubated as mentioned. Excess phycoerythrin conjugated streptavidin was removed followed by three washes in ABE. The beads were then re-suspended in 125 µl of ABE and analyzed on the BioPlex100 system.

The reader was set to read a minimum of 50 beads with identical unique detection signal. Fluorescence values for BSA beads were subtracted from rPfGARP and rPfSEA-1 beads and the results were expressed as median fluorescent intensity (MFI). The cut-off for detectable anti-PfGARP/PfSEA-1 antibody levels was defined as fluorescence values greater than the mean plus 2 s.d. fluorescence level of 10 healthy children. Anti-rPfGARP and anti-rPfSEA-1 antibody levels are presented in the *Supplementary Appendix*.

qPCR

Following DNA extraction, real-time quantitative PCR (qPCR) to amplify the 18S small sub-unit rRNA gen (ssrDNA) of *P. falciparum* was carried out on a 7500 Real-Time PCR System (Applied Biosystems, USA). Briefly, a PCR mixture was prepared with 2.5 µl of DNA template, a 5 µM concentration of each primer and a 1.5 µM concentration of TaqMan®

TAMRA Probe labelled with 6-carboxyfuorescein (FAM) as a reporter. The results were analyzed using default settings on the Applied Biosystems 7500 Fast Real-Time PCR System Sequence Detection Software v1.4.1.

Statistical analysis

Descriptive analyses were performed using Pandas and matplotlib libraries in Python. Positive blood smears within 28 days of a previous positive smear, and with no negative smears during the same period, were considered to be a single infection. Bayesian statistical models were fit, using PyStan, to explore whether observed patterns were consistent with either Poisson, Negative Binomial or a Hierarchical Poisson model, to account for the grouping of children in twin pairs. In all models, the prior used for the log-mean parameter (for the Hierarchical Poisson model, presented in equations I and II, this corresponds to μ_g) was $Normal(0, 5^2)$, which covers all plausible values and beyond; for the scale parameter (σ_g) used in the Hierarchical Poisson model we used a weakly informative prior of $Normal(0, 1)$. Sensitivity analyses were performed with different prior assumptions, with similar results (see **Table S1** and **Figure S8**); models were also fit that used the log number of routine visits, rather than the log follow-up duration, as the offset term, and were consistent with findings in the *Results* section.

Results

Study participants

We report data from a birth cohort study with intensive follow-up performed in Southwest Mali. 25 twin pairs, 2 of which monozygotic, were identified and included in this analysis.

These participants were recruited at birth between April 2011 and February 2015 and were followed for 30.7 months (median, interquartile range [IQR] 18.5 – 39.9). The median number of study visits, scheduled or unscheduled, was 33 (IQR 16 - 44) and the range was 7 - 91. Half (25/50) of these children were female. Of the 50 children, 36 had AA genotype, 9 were heterozygous for hemoglobin C mutation (AC genotype) and 5 had AS genotype. There were 9 twin pairs with at least one child with hemoglobin mutations; in 4 of these 9 pairs, only one of the two twins had hemoglobin mutations. Additional information about the pairs is presented in **Table 1**.

Malaria infections during childhood

Overall, 1,717 blood smears were performed, 126 performed in monozygotic twins. Asexual falciparum parasites were identified in 190 smears - 137 during scheduled and 53 during unscheduled, walk-in visits. 40/50 children had at least one falciparum-positive visit; the median age at the first falciparum-positive smear was 11.4 months (IQR 5.9–14.4). In 24/25 pairs, at least one of the twins had a positive smear. Both children in the pair with no parasitemia were followed for more than 15 months. 36/40 children with at least one positive smear developed clinical malaria (range of number of clinical episodes per child, 1–13).

Figures 1 (the subset of twin pairs with longest follow-up) and **S1** (all twin pairs) show patterns of infection and parasitemia. Study pairs were broadly categorized: pairs where both children were infected at least once (concordant group; group 1 in **Figures 1** and **S1**); pairs where only one of the twins had infection (discordant group; group 2 in **Figures 1**

and **S1**); twins with different hemoglobin genotypes (AA versus AS [i.e. heterozygous for sickle cell mutation] hemoglobin type; group 3 in the same figure). Although numbers of infections in children in the same pair were often similar (**Figures S1** and **S2**), in 8/25 pairs, all with follow-up longer than a year, only one twin had parasitemia detected. The number of positive smears in this group was low: in 5/8 pairs the infected sibling had one or two positive smears, which might suggest lower exposure to sporozoite-infected mosquitoes in the household or area where these pairs lived. Note that 16/28 and 4/16 births in groups 1 and 2, as defined above, happened during the transmission season months (July – December); furthermore, median (IQR) numbers of scheduled visits per month of follow-up in the two groups were similar: 0.84 (0.71 – 0.97) and 0.74 (0.66 – 0.88) scheduled visits per month, respectively. To assess whether this pattern, where only one of the twins had infections detected by microscopy, represents reduced exposure or better control of parasitemia at submicroscopic levels, available samples from children with no positive blood smears were tested with a sensitive molecular assay (qPCR targeting 18S; 2 samples for 3 children, 1 sample for 4 children). Only one child without patent infection had parasites detected by molecular methods.

To further understand the distribution of infections within twin pairs, we fit a series of Bayesian models to the number of independent malaria infections, accounting for differences in follow-up duration: Poisson model (where the same incidence parameter is assumed for all children in the study), Negative Binomial model (which also assumes the same incidence parameter but allows for more between child variability, without regard to twin pairs) and Hierarchical Poisson model (where a separate incidence parameter is

allowed for each twin pair). The Hierarchical Poisson model is shown below in equations I and II:

$$\mu_j \sim Normal\ (\mu_g, \sigma_g^2\ ) \quad \text{(I)}$$

$$y_{ij} \sim Poisson\ (e^{(\mu_j + offset)}\ ) \quad \text{(II)}$$

where log-rate μ_j for pair j is assumed to follow a Normal distribution with location parameter μ_g and scale σ_g. The number of infections detected for child i in pair j is represented by y_{ij} ; the offset term corresponds to the log of the follow-up duration.

The Poisson model assuming the same incidence for all twin pairs does not fit the data well. As can be seen in **Figure 2A**, the data are consistent with the Hierarchical Poisson model, suggesting that there is between pair variability in incidence parameters, but not necessarily within twin pair variability. In particular, consistency with the latter suggests that the observed degree of discordance in malaria infection history among twins in some pairs could have arisen even if there were no differences in the underlying incidence within pairs of siblings, i.e. if both twins in each pair were equally exposed to malaria parasites and had on average the same number of events per time unit. In other words, the relatively low expected numbers of infections per time unit estimated under the Hierarchical Poisson model for pairs of group 2, defined in **Figure 1**, are consistent with the pattern where one of the twins had no infections detected during follow-up and the other twin in the same pair had only a few infections. On the other hand, the underlying infection rates estimated in the Hierarchical Poisson model suggest significant variation between pairs (**Figure 2B**).

The total number of clinical malaria episodes was 141. Of the four children who became infected during follow-up and did not develop clinical malaria, two were only followed for a few months, another child only developed infection at the age of ~6 months and the other child had only one infection at the age of ~13 months, despite her twin developing two clinical malaria episodes in the same period. As can be seen in **Figures 1** and **S1**, children continued to develop clinical malaria throughout their follow-up, sometimes with more than one episode during the same transmission season. The numbers of clinical episodes were generally similar for twins, except for those pairs where one twin had AS genotype and the other AA genotype (**Figure 3**).

Non-falciparum malaria infections were also observed: 5 positive smears for *Plasmodium malariae*, all in different pairs, and 7 positive smears for *Plasmodium ovale* (3 different children in 2 different pairs). In three clinical malaria episodes, only non-falciparum malaria species were observed. Additionally, malaria transmission stages, gametocytes, were detected in 18 visits (11 different children in 9 different pairs).

Twins with different hemoglobin types

There were three pairs with discordant hemoglobin S mutation status. In one of these pairs, while parasitemia in 8/10 positive-smears of the AA child was > 1000 parasites/300 WBCs, none of the 17 positive-smears in the AS child had >1000 parasites/300 WBCs. The second twin pair with discordant hemoglobin S status had different follow-up duration: while the AA child was followed for nearly 4 years, the AS child died at the age of 2.5 years. If we consider only the first ~2.5 years of follow-up of the AA child, there were 10 positive blood

smears, while the AS child only had one falciparum positive visit. We tested 2 samples from this child with AS genotype using qPCR and did not observe submicroscopic malaria infection. In the third pair, no infections were detected in both AA and AS children.

Longitudinal seroreactivity

In addition to quantifying parasite carriage throughout childhood, we also tested samples, including cord-blood samples, for antibodies against the malaria blood-stage antigens. Cord-blood levels of antibodies against four different blood-stage antigens were similar between twins (**Figure S5**). In **Figures 4**, **S6** and **S7**, we show longitudinal changes in antibodies as children age. Samples collected during infancy had lower antibody levels compared to birth, as maternal antibodies are cleared from infants' blood. For a subset of children, in samples collected after infancy, antibody levels that were as high as those observed at birth were, at least temporarily, detected. In 5/8 pairs in the discordant group (i.e. pairs where only one child had evidence of infection), the child with at least one positive smear had higher antibody levels following an infection than the twin with no infections. In one of the pairs of twins with dissimilar hemoglobin genotype, although the AS child had consistently lower parasitemia, antibody levels were similar to the levels of the AA sibling.

Discussion

Falciparum malaria burden is not equally distributed in endemic areas [12]. Here, we provided a detailed description of the history of malaria infections in children who have

both similar long-term exposure to vectors and share individual-level factors thought to influence infection outcome.

Although there is substantial evidence for the role of host genetics in modifying malaria disease, with recent genome-wide association studies identifying several new protective variants [13], there are only a handful of malaria epidemiological studies that focused on twins, to our knowledge none of which included intensive longitudinal follow-up from birth. These studies primarily aimed to quantify the host genetic contribution to specific malaria phenotypes by comparing monozygotic twin pairs and dizygotic twins. For example, in The Gambia, twin pairs, recruited on average at the age of 5 years, were followed during rainy seasons and a stronger correlation in the development of fever was observed for monozygotic twins compared to dizygotic twins [14]. In another malaria twin study, it was shown that correlations of the IgG isotype levels were generally higher in the monozygotic versus dizygotic twins, which was interpreted as reflecting a degree of genetic regulation [15]. In addition to studies on twins, other studies have assessed familial aggregation of malaria phenotypes to understand joint contribution of shared genetic and environmental factors [16]. Whilst our sample size of 25 twin pairs is relatively small, the dataset included more than 1,500 parasitological observations during the age period when most clinical burden occurs. More importantly, by having information from birth, we were able to describe patterns of infections in these pairs of siblings as these children acquire partial immunity from repeated exposure.

One advantage of the design adopted is that study participants in each twin pair were expected to have similar exposure to *Anopheles* vectors. Indeed, although recent studies suggest that even within households there is variation in the number of mosquito bites different individuals receive [17], twins do not differ in two important determinants of frequency of malaria vector exposure: age, which is associated with attractiveness to mosquitoes possibly through body size, and bed net use, which is likely to be shared by twins in the same room.

Several previous epidemiological studies have described considerable variability in numbers of malaria episodes in children living in the same endemic area. In our study area, we observed that while in some twin pairs both children developed multiple clinical episodes, in nearly a third of the pairs, only one of the twins had parasites during follow-up. Whilst previous studies[18, 19] suggest that average malaria infection duration is longer than the time interval between scheduled study visits, we cannot rule out the possibility that children with no parasites detected during the study had short-duration infections or subpatent infections. However, statistical analyses suggest that these observations are consistent with models that assume a common incidence for twins in the same pair. The presence in our study population of siblings with discordant hemoglobin types also allowed us to compare parasitemia during children's first infections by hemoglobin S mutation status. The observation of consistently low parasitemia in a child with AS hemoglobin genotype, whilst the twin with AA genotype had multiple infections with high parasitemia provide additional evidence for the effect of hemoglobin S heterozygosity on parasite growth as a mechanism of protection against malaria [20]. This observation also

parallels results of a Kenyan study that showed that children with higher-than-average numbers of clinical episodes often lived in the same homestead as children with average numbers of clinical episodes and that sickle cell mutation was an important factor which differentiated these groups [21]. Furthermore, after the consistent decrease in antibody levels during the first months of life as maternal antibodies wane, our antibody data show variable patterns of longitudinal changes during early childhood, including for children in the same pair (although changes in antibodies were often similar for different target antigens studied), reflecting between pair differences in parasite exposure.

CONCLUSIONS

By using a novel design, including longitudinal follow-up of twins from birth, with over 1,500 parasitological observations, and a series of statistical models involving different assumptions on the distribution of the underlying incidence risk in the study population, we described heterogeneity in malaria burden between children from different families. This could be related to differences in vector exposure, host genetics and preventative measures, such as bed net use. Our detailed longitudinal data from naturally matched children also suggested similarities in malaria infection history within twin pairs, although important differences were observed for twins having different hemoglobin genotypes. More detailed phenotyping (i.e. host and parasite transcriptomics) of infections of intensively followed matched children, for whom history of infection exposure would thus be documented, is the next methodological step that could provide additional insights into the susceptibility to repeated malaria episodes during early childhood and generate hypotheses to be tested in larger population-based studies.

Notes

Acknowledgments

We thank the women and children in Ouelessebogou, Mali for participation in the study. Jillian Neal managed qPCR analysis, Rathy Mohan managed the clinical data, staff at the community health center supported follow-up of study participants. We thank J. Patrick Gorres for editing and assistance in preparing the manuscript.

Financial support

This work was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Conflict of Interest

Authors declare that they have no competing interests.

References

1. Goncalves BP, Huang CY, Morrison R, et al. Parasite burden and severity of malaria in Tanzanian children. N Engl J Med **2014**; 370:1799-808.

2. Trape JF, Rogier C, Konate L, et al. The Dielmo project: a longitudinal study of natural malaria infection and the mechanisms of protective immunity in a community living in a holoendemic area of Senegal. Am J Trop Med Hyg **1994**; 51:123-37.

3. Taylor SM, Parobek CM, Fairhurst RM. Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. Lancet Infect Dis **2012**; 12:457-68.

4. Mackinnon MJ, Gunawardena DM, Rajakaruna J, Weerasingha S, Mendis KN, Carter R. Quantifying genetic and nongenetic contributions to malarial infection in a Sri Lankan population. Proc Natl Acad Sci U S A **2000**; 97:12661-6.

5. Mackinnon MJ, Mwangi TW, Snow RW, Marsh K, Williams TN. Heritability of malaria in Africa. PLoS Med **2005**; 2:e340.

6. Le Hesran JY, Cot M, Personne P, et al. Maternal placental infection with Plasmodium falciparum and malaria morbidity during the first 2 years of life. Am J Epidemiol **1997**; 146:826-31.

7. Mutabingwa TK, Bolla MC, Li JL, et al. Maternal malaria and gravidity interact to modify infant susceptibility to malaria. PLoS Med **2005**; 2:e407.

8. Miura K, Herrera R, Diouf A, et al. Overcoming allelic specificity by immunization with five allelic forms of Plasmodium falciparum apical membrane antigen 1. Infect Immun **2013**; 81:1491-501.

9. Shimp RL, Jr., Martin LB, Zhang Y, et al. Production and characterization of clinical grade Escherichia coli derived Plasmodium falciparum 42 kDa merozoite surface protein 1 (MSP1(42)) in the absence of an affinity tag. Protein Expr Purif **2006**; 50:58-67.

10. Raj DK, Das Mohapatra A, Jnawali A, et al. Anti-PfGARP activates programmed cell death of parasites and reduces severe malaria. Nature **2020**; 582:104-8.

11. Raj DK, Nixon CP, Nixon CE, et al. Antibodies to PfSEA-1 block parasite egress from RBCs and protect against malaria infection. Science **2014**; 344:871-7.

12. Mwangi TW, Fegan G, Williams TN, Kinyanjui SM, Snow RW, Marsh K. Evidence for over-dispersion in the distribution of clinical malaria episodes in children. PLoS One **2008**; 3:e2196.

13. Kariuki SN, Williams TN. Human genetics and malaria resistance. Hum Genet **2020**; 139:801-11.

14. Jepson AP, Banya WA, Sisay-Joof F, Hassan-King M, Bennett S, Whittle HC. Genetic regulation of fever in Plasmodium falciparum malaria in Gambian twin children. J Infect Dis **1995**; 172:316-9.

15. Duah NO, Weiss HA, Jepson A, Tetteh KK, Whittle HC, Conway DJ. Heritability of antibody isotype and subclass responses to Plasmodium falciparum antigens. PLoS One **2009**; 4:e7381.

16. Ranque S, Safeukui I, Poudiougou B, et al. Familial aggregation of cerebral malaria and severe malarial anemia. J Infect Dis **2005**; 191:799-804.

17. Guelbeogo WM, Goncalves BP, Grignard L, et al. Variation in natural exposure to anopheles mosquitoes and its effects on malaria transmission. Elife **2018**; 7.

18. Bretscher MT, Maire N, Chitnis N, Felger I, Owusu-Agyei S, Smith T. The distribution of Plasmodium falciparum infection durations. Epidemics **2011**; 3:109-18.

19. Felger I, Maire M, Bretscher MT, et al. The dynamics of natural Plasmodium falciparum infections. PLoS One **2012**; 7:e45542.

20. Mahamar A, Attaher O, Swihart B, et al. Host factors that modify Plasmodium falciparum adhesion to endothelial receptors. Sci Rep **2017**; 7:13872.

21. Ndungu FM, Marsh K, Fegan G, et al. Identifying children with excess malaria episodes after adjusting for variation in exposure: identification from a longitudinal study using statistical count models. BMC Med **2015**; 13:183.

Figure Legends

Fig. 1. Falciparum infections during early childhood. In each panel, smear results (y-axes; in parasite counts per 300 WBCs) for both twins in each pair are represented by blue and red lines. Circles represent time points (x-axes) when smears were performed; clinical episodes are represented by black squares. Orange shaded areas represent the transmission season (July – December). To facilitate visualization of low-density infections, infections with less than 100 parasites/300 WBC are represented by '*'. Groups described in the text are shown on the top right corner in each panel. Pairs in group 1 were subcategorized in those with similar parasite levels during infection (1a) and those with discordant levels (1b), where the twin with highest parasitemia had at least one smear with counts 5-fold or higher than the maximum parasitemia of her or his sibling. Group 2 corresponds to pairs where only one of the twins had infection. Group 3 includes pairs with discordant hemoglobin S mutation status. Note that (i) **here only the two pairs in each group with the longest follow-up are presented (the same information is shown for all pairs in Figure S1)**; (ii) in group 3, only individuals with AA genotype had hyperparasitemia (here defined as > 3,750 parasites/300 WBC; see also **Figure S3**); (iii) clinical episodes where only non-falciparum parasites were detected are represented by black squares that are not linked to the plot lines and whose y-axis coordinate is defined by the non-falciparum parasitemia.

Fig. 2. Results of statistical models on the number of malaria infections. (A) We fit three different count data models: Poisson model, that assumes the same incidence parameter for all individuals; Negative Binomial model, that allows for additional variability in the incidence, but that does not explicitly allow for the clustering of children within pairs; and a Hierarchical Poisson model that allows the incidence parameter to vary between different pairs of twins but not within pairs. In this figure, the y-axis represents the number of infections (black circles) during the follow-up of children, grouped in pairs (x-axis); children with hemoglobin AS genotype are presented by red stars. For each child, posterior predictive distributions are presented for the 3 models discussed above; the different degrees of transparency of the colors represent different posterior intervals: 2.5 – 97.5, 25 – 75, 40 – 60. Whilst the Poisson model (orange color) does not fit the data well, data are consistent with both the Negative Binomial (green color) and the Hierarchical Poisson (purple) models. Of note, we also fit a Poisson model that included as covariate birth before versus after August 2013, when follow-up frequency changed (see **Figure S4**). **(B)** In this panel, the estimated rate of infection (per month) (x-axis; posterior median and 95% interval) based on the Hierarchical Poisson model is presented for each of the 25 twin pairs (y-axis). Note that the rate in Pair 5 is estimated to be relatively high, due to infection detection during the relatively short follow-up of this pair. The ordering of this panel, from higher to lower y-axis coordinates, corresponds to the same ordering, from left to right, in panel A.

Figure 3. Numbers of clinical malaria episodes for twins in the same pair (x- and y-axes). Red circles represent pairs with discordant hemoglobin S status; in these pairs, the Twin 2 (y-axis) is AS child. Noise was added to x- and y-coordinates so that circles for pairs with similar coordinates do not fully overlap.

Figure 4. Levels (y-axes) of antibodies against the malaria antigens AMA-1 (gray lines) and MSP-1 (green lines). As with **Figure 1**, each panel presents data of a twin pair. The ordering of the panels was kept between **Figures 1** and **4**; and continuous lines in **Figure 4** represent data of the twins represented with red lines in **Figure 1**, and the dashed line, data of the twins represented by blue lines in **Figure 1**. Blue and red circles, shown in the lower half of each panel, indicate the timing (x-axes) of positive smears. Black crosses on the serological measurements represent serological samples coinciding with smear positive results. Of note, there were twins who developed a high humoral response despite low cumulative parasite burden. Note that, as for Figure 1, only data for the twin pairs in each group with the longest follow-up are shown; similar data for all twin pairs are shown in **Figure S6**.

Table 1. Twins' characteristics and follow-up.

Positive blood smears in this table refer only to falciparum infections

		Dizygotic	Monozygotic
		N	N
Total		46	4
Female		25	0
Hemoglobin genotype			
	AA	32	4
	AC	9	0
	AS	5	0
Mother's gravidity			
	Primi	6	2
	Secundi	6	2
	Multi	34	0
Placental Malaria		6	0
		Median (IQR)	Values
Maternal age (years)		27 (21 – 34.5)	20, 22
Age at first infection (months)		10.4 (5.9 – 13.5)	14.1, 19.8, 26.2
Age at last study visit (months)		30.7 (18.0 – 40.3)	28.6 (x2), 33.3 (x2)
Number of study visits		34 (15.2 – 46.2)	27, 28, 35, 36
		N	N
Total number of positive blood smears		183	7
Total number of clinical malaria		135	6
Total number of severe malaria		2	0

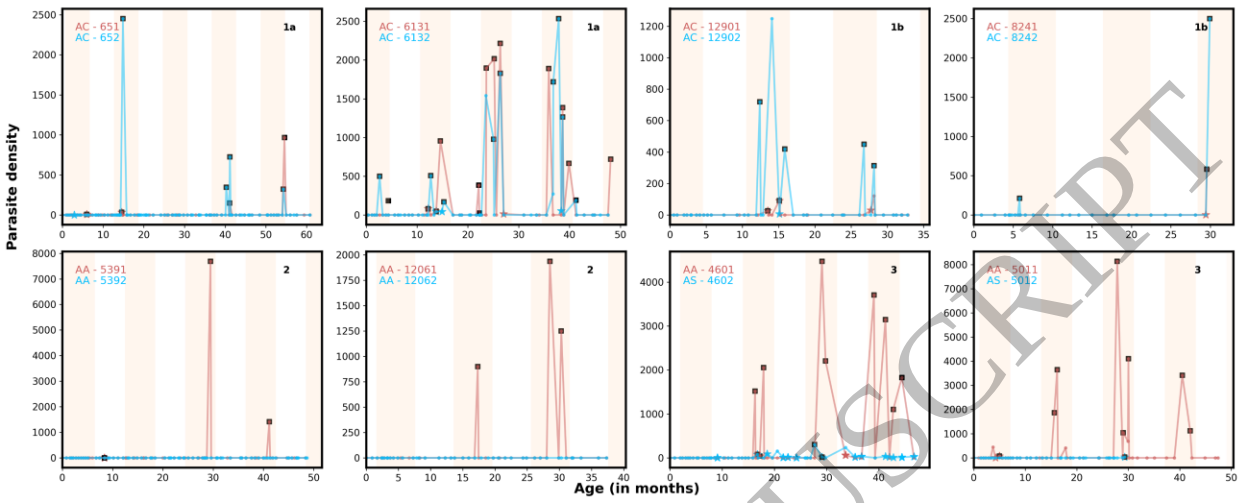

Figure 1
165x93 mm (2.5 x DPI)

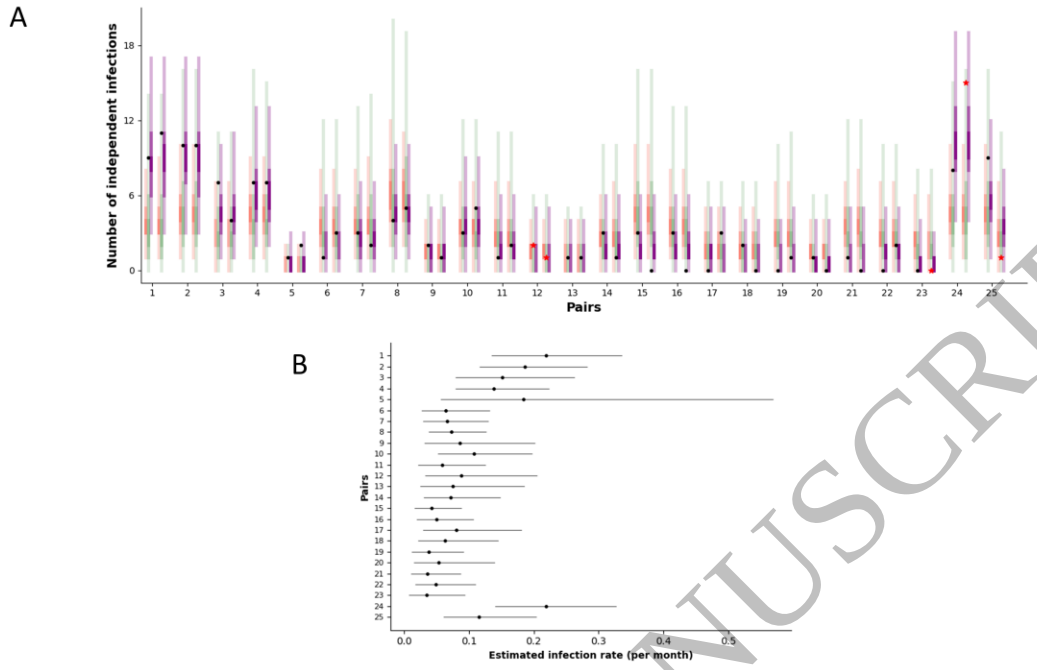

Figure 2
165x93 mm (2.5 x DPI)

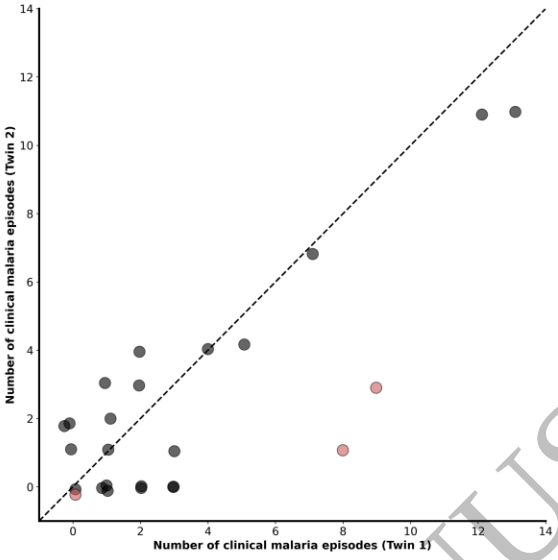

Figure 3
165x93 mm (2.5 x DPI)

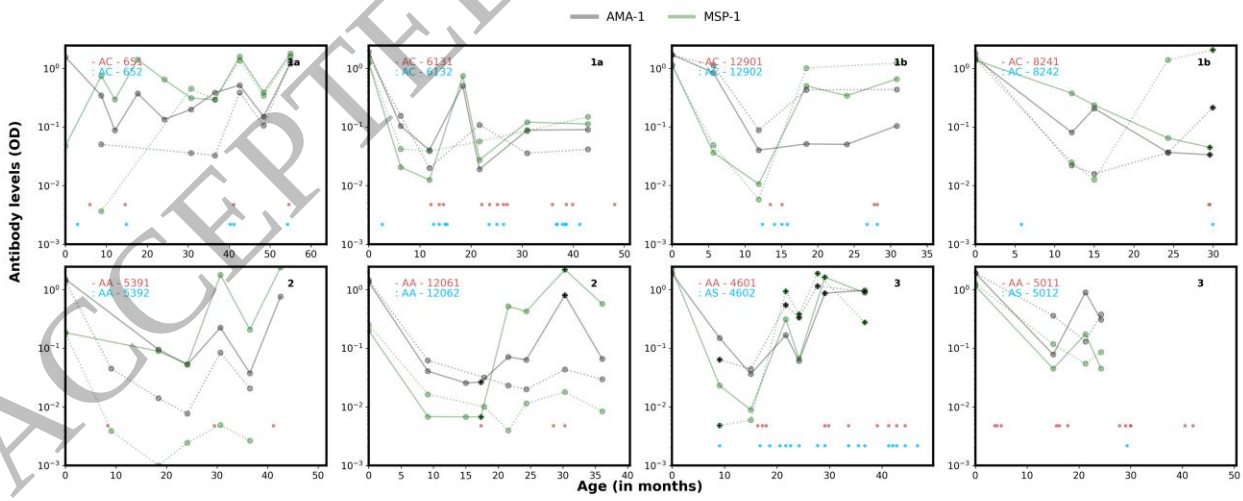

Figure 4
165x93 mm (2.5 x DPI)