



OPEN

## Spatial and genetic clustering of *Plasmodium falciparum* and *Plasmodium vivax* infections in a low-transmission area of Ethiopia

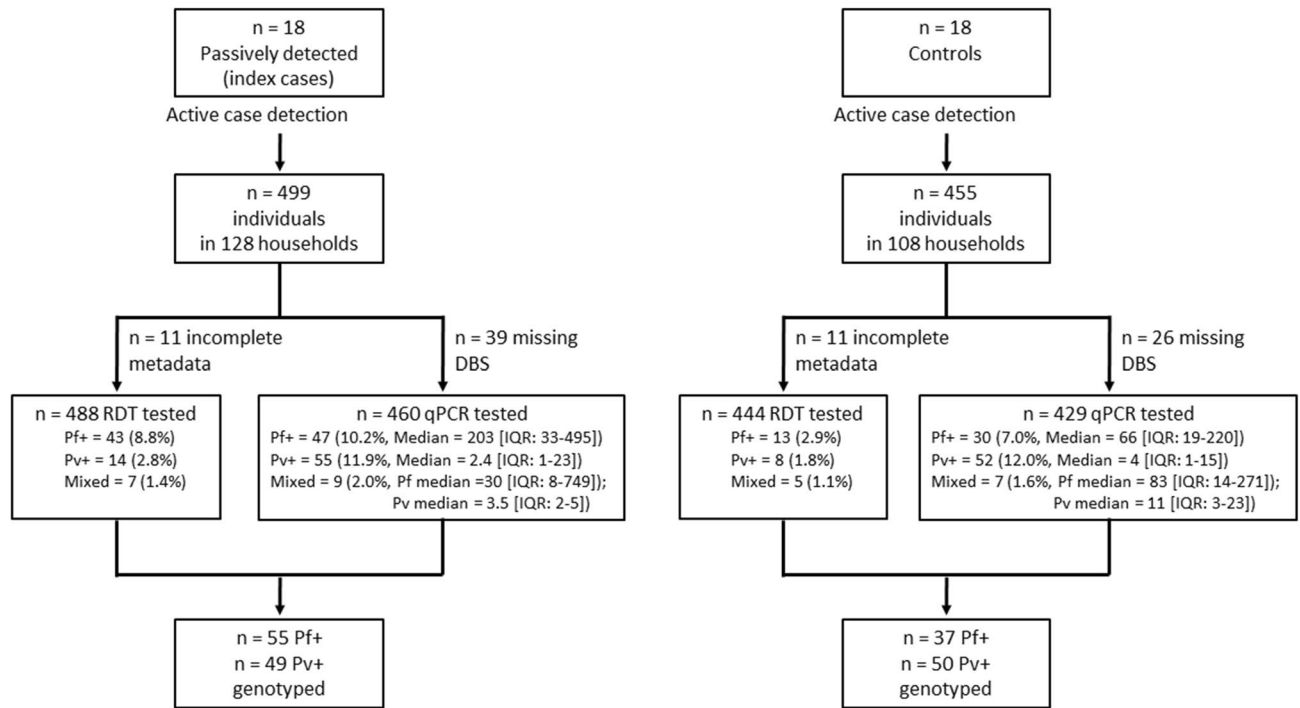
Sofonias K. Tessema<sup>1,9</sup>, Mulualem Belachew<sup>2,9</sup>, Cristian Koepfli<sup>3</sup>, Kjerstin Lanke<sup>4</sup>, Tiffany Huwe<sup>3</sup>, Wakweya Chali<sup>2</sup>, Girma Shumie<sup>2</sup>, Elias F. Mekuria<sup>5</sup>, Chris Drakeley<sup>6</sup>, Endalamaw Gadisa<sup>2</sup>, Bryan Greenhouse<sup>1,7</sup>, Teun Bousema<sup>4,6</sup> & Fitsum G. Tadesse<sup>2,4,8</sup>✉

The distribution of malaria infections is heterogeneous in space and time, especially in low transmission settings. Understanding this clustering may allow identification and targeting of pockets of transmission. In Adama district, Ethiopia, *Plasmodium falciparum* and *P. vivax* malaria patients and controls were examined, together with household members and immediate neighbors. Rapid diagnostic test and quantitative PCR (qPCR) were used for the detection of infections that were genetically characterized by a panel of microsatellite loci for *P. falciparum* (26) and *P. vivax* (11), respectively. Individuals living in households of clinical *P. falciparum* patients were more likely to have qPCR detected *P. falciparum* infections (22.0%, 9/41) compared to individuals in control households (8.7%, 37/426; odds ratio, 2.9; 95% confidence interval, 1.3–6.4;  $P = .007$ ). Genetically related *P. falciparum*, but not *P. vivax* infections showed strong clustering within households. Genotyping revealed a marked temporal cluster of *P. falciparum* infections, almost exclusively comprised of clinical cases. These findings uncover previously unappreciated transmission dynamics and support a rational approach to reactive case detection strategies for *P. falciparum* in Ethiopia.

Recent success in malaria control in endemic countries has re-ignited elimination efforts<sup>1</sup>. Transition from control to elimination requires improved understanding of transmission patterns within communities<sup>2</sup> to better target the remaining infections that sustain or seed transmission. In low endemic settings, malaria infections are often characterized by low parasite density, commonly below the detection limit of rapid diagnostic tests (RDT) and microscopy<sup>3,4</sup>. These low-density infections are typically asymptomatic and may contribute substantially to onwards transmission to mosquitoes<sup>5</sup>.

As malaria transmission intensity declines, the distribution of malaria infections becomes increasingly heterogeneous, and infections may cluster within villages or households<sup>6–8</sup>. If this clustering reflects pockets of active transmission, finding and targeting these clustered infections may accelerate targeted elimination efforts<sup>9</sup>. Based on this assumption, individuals living in close proximity to passively detected cases are screened and treated in reactive case detection (RCD) strategies in low transmission settings<sup>8</sup>. The optimal and logistically feasible population that needs to be targeted in RCD remains unclear<sup>10,11</sup>. Several studies have attempted to determine whether geographical clusters of infections show evidence of increased genetic relatedness and thus truly reflect related transmission events<sup>12–18</sup>. Few studies have coupled household-level case clustering studies with information on

<sup>1</sup>EPPIcenter Research Program, Division of HIV, ID, and Global Medicine, Department of Medicine, University of California, San Francisco, CA, USA. <sup>2</sup>Armauer Hansen Research Institute, PO Box 1005, Addis Ababa, Ethiopia. <sup>3</sup>Department of Biological Sciences, Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556, USA. <sup>4</sup>Department of Medical Microbiology, Radboud University Medical Center, 6500 HB Nijmegen, The Netherlands. <sup>5</sup>Faculty of Public Health, Jimma University, Jimma, Ethiopia. <sup>6</sup>Department of Immunology and Infection, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK. <sup>7</sup>Chan Zuckerberg Biohub, San Francisco, CA 94158, USA. <sup>8</sup>Institute of Biotechnology, Addis Ababa University, PO Box 1176, Addis Ababa, Ethiopia. <sup>9</sup>These authors contributed equally: Sofonias K. Tessema and Mulualem Belachew. ✉email: fitsum.tadesse@radboudumc.nl



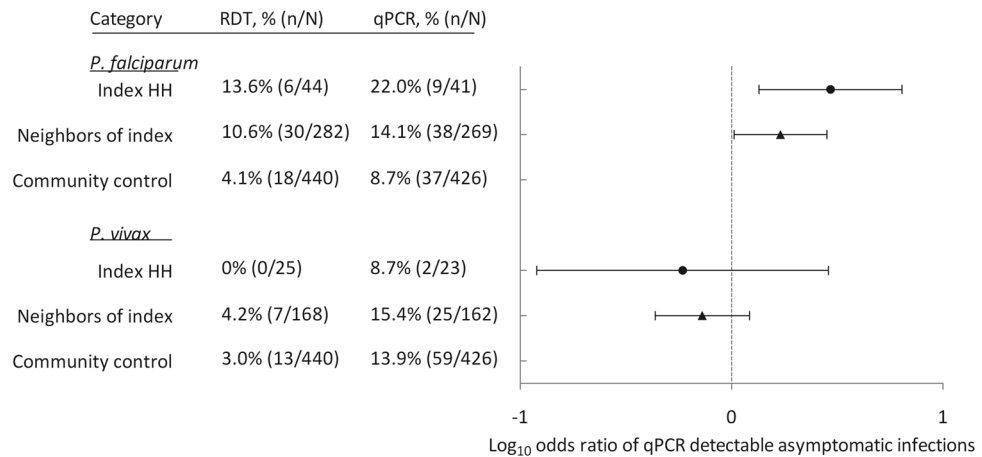
**Figure 1.** Flow chart of samples collected and genotyped. Indicated in the figure is the number of index cases and controls and individuals screened around them. Parasite identification results by RDT and qPCR are also indicated. Of all samples with complete metadata that were tested using both RDT and qPCR ( $n = 889$ ) only 5 *P. falciparum* and 1 mixed species infection were qPCR negative whilst being positive. All RDT positive samples were genotyped except these six samples. Parasite densities (18S copies/ $\mu\text{L}$ ) are indicated for each species in (median, interquartile range [IQR]). Indicated at the bottom are number of samples that were genotyped for each species. DBS, dried blood spot; RDT, rapid diagnostic test; *Pf*, *P. falciparum*; *Pv*, *P. vivax*; qPCR, quantitative polymerase chain reaction; IQR, interquartile range.

the genetic relatedness of infections. Such studies can provide a better understanding of fine-scale transmission patterns and estimate the contribution to transmission of infections that are not detected by routine interventions.

In this study, the prevalence and distribution of asymptomatic *P. falciparum* and *P. vivax* infections surrounding passively detected RDT-confirmed clinical malaria patients was assessed using RDT and quantitative PCR (qPCR) in Batu Degaga, Ethiopia. Spatial and temporal relatedness and clustering of infections within and between households of passively detected malaria infections and controls were investigated by genotyping 26 and 11 microsatellite loci for *P. falciparum* and *P. vivax*, respectively, which had been shown previously to be useful to describe population structure<sup>12,19–25</sup>.

## Results

**Prevalence and clustering of malaria infections.** Family members and neighbors of 18 passively detected index cases with RDT confirmed *P. falciparum* ( $n = 10$ ), *P. vivax* ( $n = 6$ ) or mixed species infections ( $n = 2$ ) and 18 controls without malaria infections participated in this study. In total, 499 individuals from 128 households around index cases and 454 individuals from 108 households around controls were successfully sampled (Fig. 1). The median number of individuals examined was 28 around index cases (interquartile range [IQR], 24–31) and 26 around controls (IQR, 22–30), with a median of 4 individuals (IQR, 3–5) examined per household. Reported history of malaria in the last 3 weeks was substantially higher among individuals in index HHs (38.8%, 31/88;  $P < 0.001$ ) compared with their neighbors (22.2%, 92/414) and those from and around control HHs (4.3%, 19/447). In the screened population surrounding index cases and controls (excluding the index cases), prevalence of infection was 7.9% (72/913) by RDT and 21.6% by qPCR (189/877). At least one infected individual was found in 28% (66/236) and 48% (114/236;  $P < 0.001$ ) of households by RDT and qPCR, respectively. *P. falciparum* infections detected in households of neighbors of index cases had higher parasite density (median parasite density by qPCR, 203 copies/ $\mu\text{L}$ ; IQR, 33–576) than infections detected in households of neighbors of controls (median, 75 copies/ $\mu\text{L}$ ; IQR, 17–209;  $P = 0.037$ ); whilst such difference was not detected in *P. vivax* infections ( $P = 0.27$ ). Fewer *P. falciparum* infections were detected by qPCR in households of controls (8.6%, 37/432) compared to index case households (20.3%, 12/59; OR, 2.7; 95% CI 1.3–5.6;  $P = 0.01$ ) after excluding index cases. There was no evidence for differences in parasite prevalence by PCR between households of control and neighbors of index cases (10.4%, 41/395; OR 1.2; 95% CI 0.8–2.0;  $P = 0.37$ ; Fig. 2). For *P. vivax* infections, no such differences were detected in qPCR parasite prevalence in relation to proximity to the household of the index case ( $P = 0.53$ ; Fig. 2).



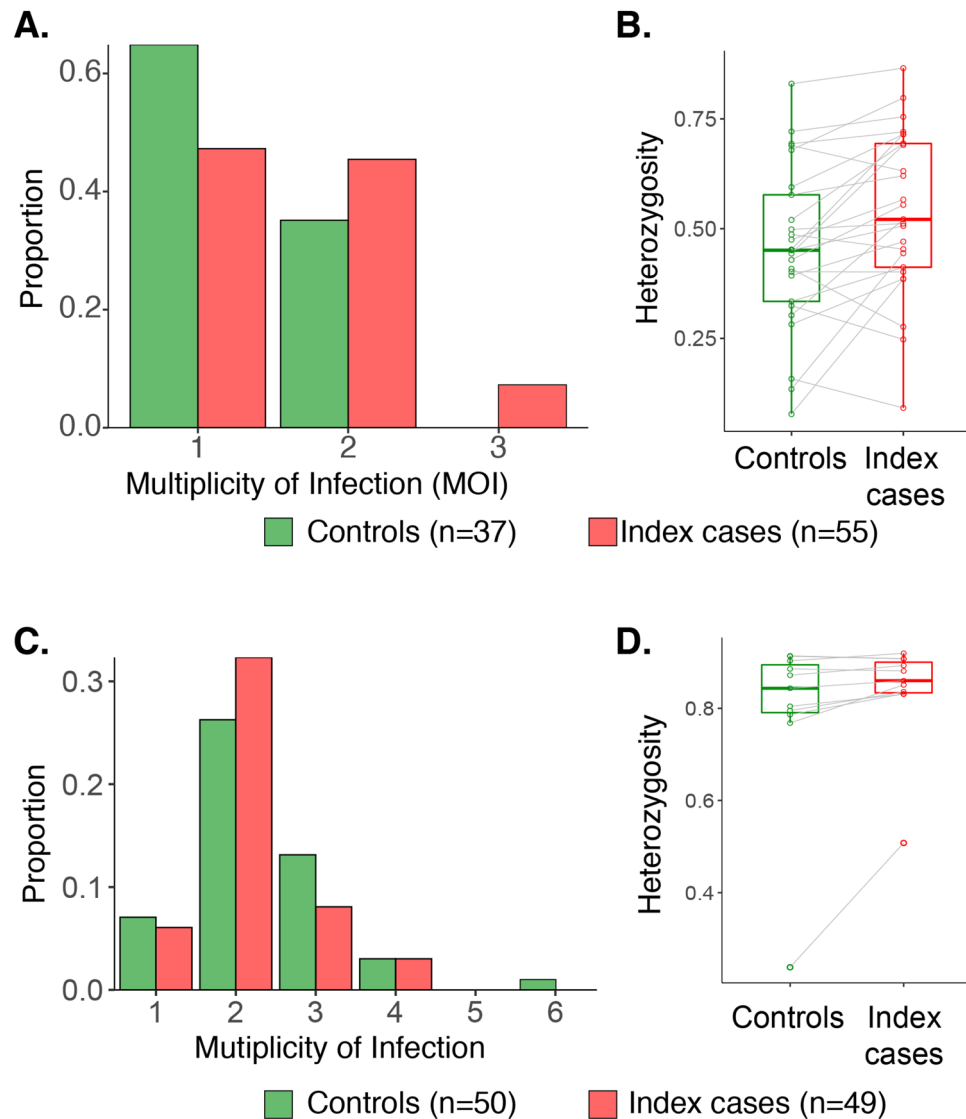
**Figure 2.** Comparison of the odds of detecting infections around households. The forest plot shows the odds ratio and 95% confidence interval for detecting asymptomatic malaria infections by 18S based qPCR in households of index cases (Index HH) and neighbors of index cases compared to control households (community control) for *P. falciparum* and *P. vivax* infections separately after excluding the index cases. Indicated on the X-axis the  $\text{Log}_{10}$  transformed odds ratio and on Y-axis are the different comparison groups. The prevalence of RDT and qPCR detectable asymptomatic infections for each household category and species are also indicated for each group to the left of the figure.

**Genetic complexity and diversity of *P. falciparum* and *P. vivax* infections.** A total of 92 *P. falciparum* infections and 99 *P. vivax* infections were successfully genotyped. The number of genetically distinct parasite clones (i.e. multiplicity of infections, MOI) and population level genetic diversity were compared for infections around index and control cases for each species. In *P. falciparum* infections, the mean MOI was 1.5 (range: 1–3) and infections detected in and around households of index cases had higher MOI compared to those detected in and around control households (mean MOI: 1.6 vs. 1.4;  $P=0.04$ ; Fig. 3A). When stratified by index and controls, infections detected around index cases showed moderately higher genetic diversity compared to those detected around controls in *P. falciparum* (Mean  $H_E$ : 0.56 vs. 0.49;  $P=0.03$ ; Fig. 3B).

In contrast to the sympatric *P. falciparum* populations, within-host diversity was consistently higher in *P. vivax*, the mean MOI was higher in *P. vivax* (mean MOI = 2.2; range: 1–6), with no difference between index and control cases (mean MOI 2.2 vs. 2.3;  $P=0.35$ ; Fig. 3C). The overall genetic diversity of *P. vivax* infections was high (Mean  $H_E \pm SE$ :  $0.82 \pm 0.03$ ). Similar to *P. falciparum*, infections detected around index cases showed moderately higher genetic diversity compared to those detected around controls (Mean  $H_E$ : 0.84 vs. 0.79;  $P=0.08$ ; Fig. 3D).

**Spatial and temporal relatedness of *P. falciparum* and *P. vivax* infections.** Pairwise genetic relatedness in 70 *P. falciparum* samples from 57 households (with allele calls from at least 15 loci) and 91 *P. vivax* samples from 79 households (with allele calls from at least 6 loci) were compared using the IBS metric. These data were used to determine how genetic relatedness of infections varied in space and time. *P. falciparum* infections sampled from the same household were more strongly related than infections from different households (median genetic relatedness: 0.81 vs. 0.37,  $P<0.001$ ; Fig. 4A). The degree of relatedness did not differ depending on household distance. *P. falciparum* infections that were genetically related (IBS > 0.5) and sampled from different households (Fig. 4B), were sampled within a shorter time frame compared to unrelated infections (median number of days: 17 vs. 25,  $P<0.001$ ; Fig. 4C), suggesting temporal clustering of genetically related *P. falciparum* infections. In contrast, overall pairwise genetic relatedness between *P. vivax* infections was low, with no spatial (Fig. 4D,E) or temporal clustering (Fig. 4F).

**Fine-scale population structure of *P. falciparum* but not *P. vivax* infections.** Analysis of dominant alleles at each locus using MavericK identified three sub-populations ( $K=3$ ), with the highest model evidence for both *P. falciparum* and *P. vivax*. Most parasites had their origin clearly assigned to a single cluster; few of them were admixed as shown by bars partitioned into  $K$  colored segments (Fig. 5). *P. falciparum* infections detected around index cases vs. controls were comprised of different populations (Fig. 5A), with infections belonging to group-3 almost exclusively detected in and around index cases households (23% vs. 4%,  $P=0.010$ ) and group-1 overrepresented in controls (72% vs. 42%,  $P=0.003$ ). Infections assigned to group-3 were identified in five independent RCD investigations and included one individual in the control group whose infection was genetically related to one of the index cases (IBS = 0.7). Analyses of the elapsed time between samples in each group showed that infections assigned to group-3 were clustered temporally (median number of days = 16, IQR: 0–30) compared to group-1 (23, IQR: 12–42;  $P=0.007$ ) and group-2 (25, IQR: 7–41;  $P=0.004$ ). These observations support the hypothesis that infections identified through investigation around index cases are linked possibly due to recent, local spread of group-3 haplotypes during the study period. No *P. vivax* genetic groups were overrepresented either to the control or index cases (Fig. 5B).

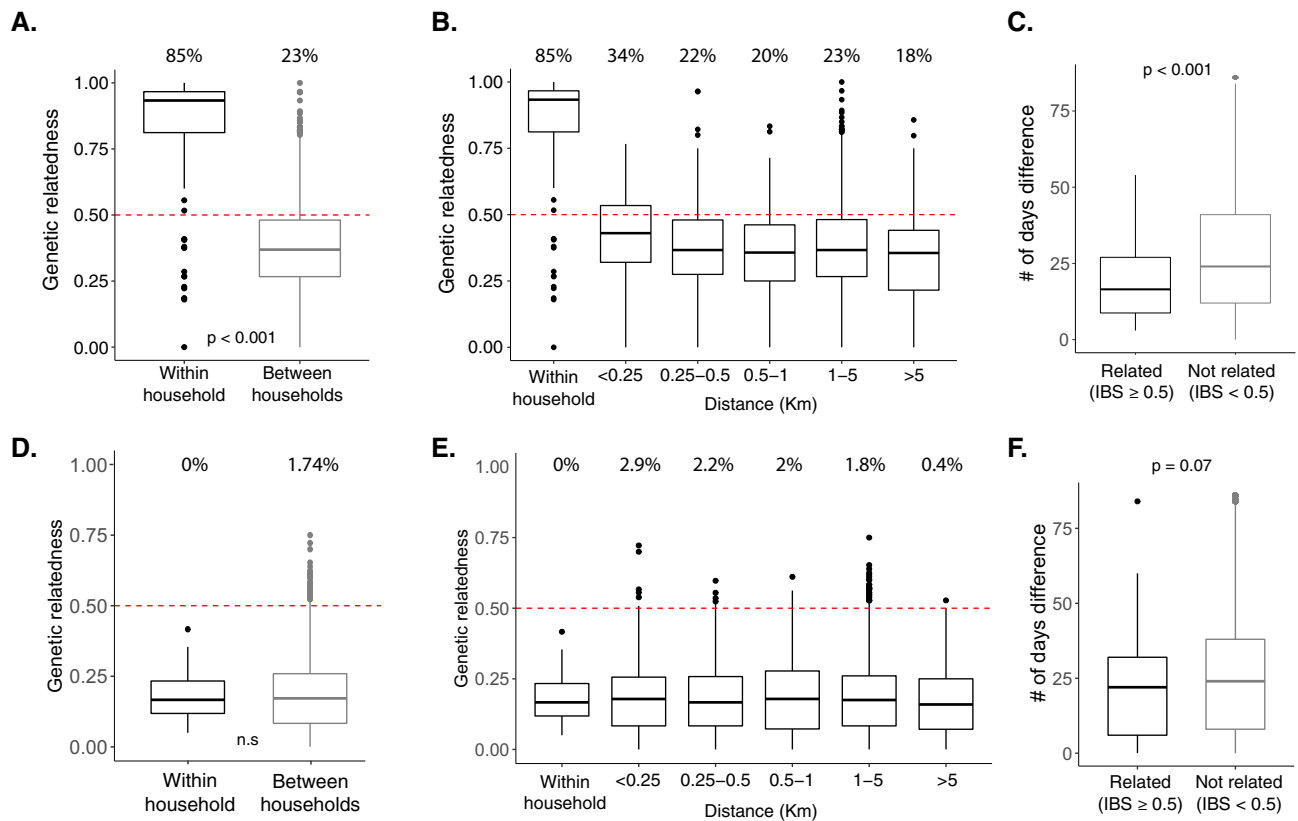


**Figure 3.** Genetic complexity and diversity of *P. falciparum* and *P. vivax* infections around index cases and controls. Distributions of multiplicity of infection (A and C) and the expected heterozygosity (B and D) of markers from samples collected in and neighboring households of index cases and controls are indicated for *P. falciparum* (A and B) and *P. vivax* (C and D). In (B and D), each dot represents one microsatellite marker. Boxplots indicate the medians and interquartile ranges of the heterozygosity of 25 loci for *P. falciparum* and 11 loci for *P. vivax*. The *P. falciparum* AS1 locus was excluded from the analyses due to the lack of diversity in this population. The whiskers indicate the highest and lowest values.

## Discussion

In the current study, we examined spatial clustering of *Plasmodium* infections in a low endemic area in Ethiopia with a focus on asymptomatic infections in the vicinity of passively detected clinical index cases. Asymptomatic *P. falciparum* infections were clustered in index case households whilst there was no such evidence for clustering of *P. vivax* infections. *P. vivax* infections were more genetically complex and diverse than *P. falciparum* infections, with no detectable spatial or temporal clustering. We identified fine-scale focal transmission and a genetic cluster around index cases that may represent a clonal *P. falciparum* expansion, highlighting the added value of genotyping to understand the contrasting transmission epidemiology of the two *Plasmodium* species.

Spatial and temporal heterogeneity in the occurrence of infections is widely acknowledged in malaria<sup>26,27</sup> and other infectious diseases<sup>28–30</sup>. Several countries have adopted RCD approaches that take advantage of infection clustering and assume that transmission chains, or at least transmission sources, can be interrupted by targeting infected household members. In some settings where symptomatic and asymptomatic malaria cases do not overlap spatially, clusters of infections can be missed by RCD<sup>31</sup>. This highlights the need for detailed studies on *Plasmodium* transmission patterns in settings aiming for elimination<sup>32</sup>. In the current study, we moved beyond spatial patterns in infection prevalence and examined small-scale transmission patterns by genotyping microsatellite markers. The overall genetic diversity (i.e. the mean expected heterozygosity) in *P. falciparum* infections



**Figure 4.** Spatial and temporal genetic relatedness of *P. falciparum* and *P. vivax* infections. Pairwise genetic relatedness is shown for sample pairs collected within and between households for *P. falciparum* (A) and *P. vivax* (D). The proportion of comparisons in each category that are related (above the 0.5 genetic relatedness threshold demarcated by the red dashed line) is shown at the top. Pairwise genetic relatedness is shown for sample pairs by physical distance bin for *P. falciparum* (B) and *P. vivax* (E). Box plots indicating the distributions of elapsed days for genetically related (IBS > 0.5) and unrelated infections are shown for *P. falciparum* (C) and *P. vivax* (F). Boxplots indicate the median and interquartile ranges, and the whiskers indicate the highest and lowest values.

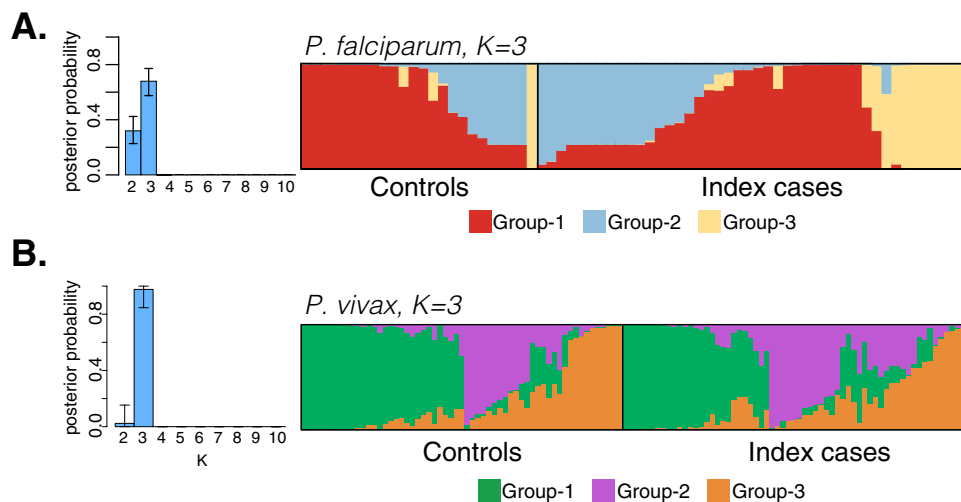
was lower (Mean  $H_E \pm SE$ :  $0.47 \pm 0.03$ ) than previously observed in other low transmission areas of Eswatini, Zambia, and Namibia (Mean  $H_E$ : 0.71–0.79)<sup>12,22,33</sup> suggesting infrequent recombination and smaller parasite population size in the study site.

We observed a clear overlap of symptomatic and asymptomatic *P. falciparum* infections; the odds of detecting an infection within the household of clinical patients was significantly higher than in neighboring households<sup>8,18,34,35</sup>. MOI and genetic diversity of infections were higher around passively detected *P. falciparum* compared to controls, indicative of higher transmission. We observed that the very strong household-level clustering for *P. falciparum* was matched with strong genetic relatedness of infections within a household. Density of asymptomatic *P. falciparum* infections surrounding clinical cases was higher than in infections detected in control households. This might indicate that the former infections were acquired more recently compared to older, very low density infections detected in controls. All of these findings corroborate that RCD is useful to detect clusters of recently acquired *P. falciparum* infections. Results from control households, however, show that a substantial number of infections is not linked to index cases. As found in other studies<sup>36,37</sup>, RCD alone is thus unlikely to achieve elimination.

We detected a genetic cluster of *P. falciparum* infections, present in multiple households, that was overrepresented in and around index cases. These findings suggest a possible population expansion of a single parasite genotype. We can hypothesize that possible expansion, occurring against the background of wide-spread asymptomatic infections, could indicate introduction of a clone from another site. Its association with clinical cases suggests that some infections disproportionately give rise to symptoms<sup>18</sup>.

As in other sites where *P. falciparum* and *P. vivax* are co-endemic, *P. vivax* was more diverse, showed high MOI, and exhibited lower levels of population structure<sup>38–41</sup>. *P. vivax* diversity was similar to previous studies from Madagascar (Mean  $H_E$ : 0.80), PNG highlands (Mean  $H_E$ : 0.82) and lowlands (Mean  $H_E$ : 0.80) and higher than South Korea (Mean  $H_E$ : 0.56)<sup>42</sup>. There was no significant difference in the odds to detect an asymptomatic *P. vivax* infection in the household or neighbors of an index case compared to a control household. There was no difference in the MOI and heterogeneity of infections around passively detected *P. vivax* infections and controls. No spatial clustering was observed for *P. vivax*. The loss of relatedness and household level clustering may reflect differences in the biology of these two *Plasmodium* species where the majority of clinical *P. vivax* cases are attributable to relapsing episodes and do not necessarily reflect an active circulating infection that involves





**Figure 5.** Fine-scale population structure and focal transmission of genetically related infections. Population cluster analysis of *P. falciparum* (A) and *P. vivax* (B) microsatellite haplotypes from dominant alleles for cases detected around index cases and controls. Individual ancestry coefficients and optimum K value ( $K = 3$ ) are shown as inferred by Maverick<sup>70</sup>. Each vertical bar represents an individual haplotype, and its membership to the three population groups is defined by the different colors. Black borders separate individuals detected around controls and index cases.

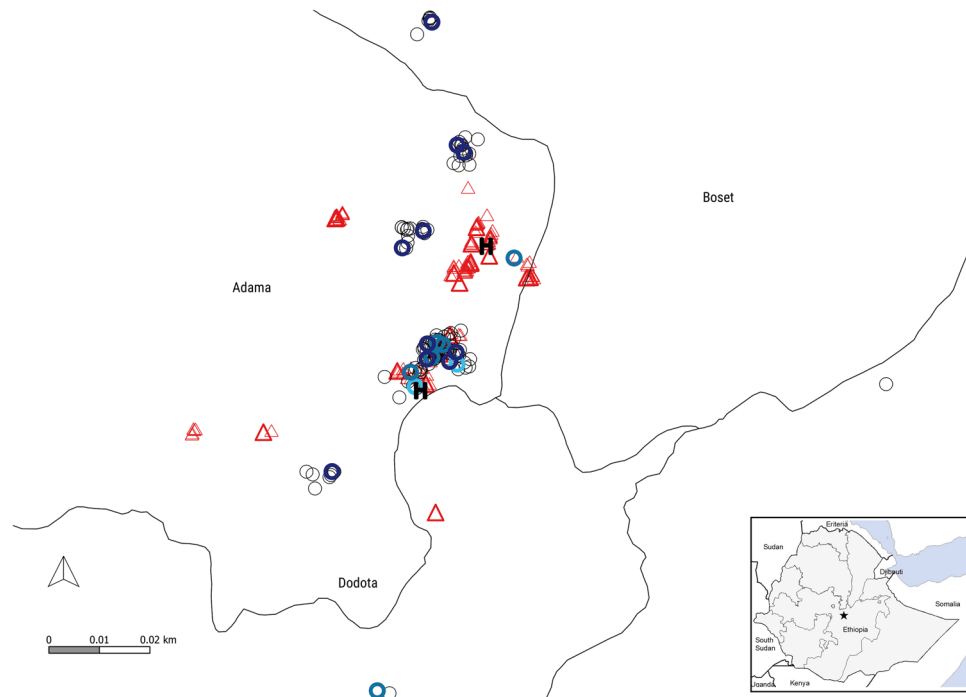
mosquito bites<sup>43,44</sup>. This would contribute to considerable variation in the time between the original inoculation by mosquito bite and parasite detection in the bloodstream, and thus diffuse any genomic signal. In summary, RCD is not more effective than random screening to identify asymptomatic *P. vivax* infections in the study site.

Our findings highlight the value of genotyping approaches to infer spatial and temporal clustering of *P. falciparum* infections. Whilst infections are prevalent outside households of index cases, and thus may be missed by RCD approaches, a disproportionate number of infections may be targeted by RCD<sup>45</sup>. Our finding of a possible population expansion of a single parasite genotype arising from a clone that was frequently detected in clinical cases further supports RCD approaches that may interrupt clinically relevant transmission chains. The lack of spatial (genetic) clustering in *P. vivax* may be attributable to the limitations of the methods used in this study, in addition to its unique biology discussed above, calling for more robust approaches. The biology of *P. vivax*; existence of hypnozoites<sup>43,46,47</sup>, generation of gametocytes at the very early stage<sup>48</sup>, lower parasite densities<sup>49,50</sup>, and higher vectoral capacity<sup>51–53</sup>, makes it a harder species to control and one that may require a different strategy to that of *P. falciparum*<sup>54,55</sup>. Future studies may benefit from more sample size and including genotyping of mosquitoes paired with the human sampling in study households<sup>56</sup>. In summary, for *P. vivax* RCD may not be more efficient than untargeted screening to identify asymptomatic infections in the study site. Control strategies against *P. vivax* need to target hypnozoites using radical cure, as the majority of infections detected are attributable to relapsing episodes rather than active infections<sup>57</sup>. Justified by the very low prevalence of glucose-6-phosphate dehydrogenase deficiency<sup>58–64</sup>. Ethiopia is currently implementing a 14-day dose primaquine for radical cure in *P. vivax*<sup>65–67</sup>.

## Methods

**Study site, population and ethics statement.** This study was conducted between October and December 2016 in Batu Degaga within Adama district, Ethiopia. Adama district is located at an altitude range of 1400–2300 m above sea level, and has an estimated population of 183,502 within 38,230 households (District Health Office Report). *P. falciparum* and *P. vivax* are co-endemic in the district<sup>5,68</sup> with 60% of the infections attributed to *P. vivax*<sup>5</sup>. The area is characterized by unstable seasonal transmission that peaks following the two rainy seasons: September to November (major) and April to May (minor). Two health posts, the lowest level governmental health facilities, serve the study area. Only one of the two health posts (the southern health post, Fig. 6) was providing service during the entire study period. The northern health post was opened only for 3 days during the study period.

The study was approved by the ethics review boards of the Department of Medical Biochemistry at Addis Ababa University (DRERC 19/16), Armauer Hansen Research Institute (PO52/14), the National Research Ethics Review Committee (310/109/2016), and the London School of Hygiene and Tropical Medicine (10628). Written informed consent was obtained from all participants and/or parents/guardians. All experiments were performed in accordance with permission obtained from participants on the consent forms, following relevant guidelines and regulations. If participant was found RDT positive while being febrile during household level sampling, he/she was referred to the nearest health post to be treated with current first-line antimalarial drugs according to the Ethiopian national malaria treatment guidelines, artemether–lumefantrine for *P. falciparum* and chloroquine for *P. vivax*.



**Figure 6.** Map of study area and distribution of study households. Shown in the right bottom corner is a map of Ethiopia with boundaries indicating administrative regions (asterisk indicating the study district, Adama, Oromia region). The location of index households (thick circles; *P. falciparum*, dark blue; *P. vivax*, blue; mixed species, light blue) and neighbors of index households (thin circles) surveyed during the study are indicated together with control households (thick triangles) and neighbor households of controls (thin triangles). The two health posts are indicated in black and bold letter ‘H’. The entire study area was 0.031 km<sup>2</sup>. The QGIS software version 3.14.16 (QGIS developer team, Open Source Geospatial Foundation Project) was used to map study households (<https://issues.qgis.org/projects/qgis/>). Coordinates of study household were geo-located with a handheld GPS receiver (GPSMAP 62 s; Garmin International).

**Sample collection and laboratory procedures.** Patients with (history of) fever that presented to the two health posts and tested positive by RDT (First Response Malaria Ag. pLDH/HRP2 Combo Card Test; Premier Medical Corporation Limited, Valsad, Gujarat, India) were classified as index cases. Patients that were suspected of malaria but tested RDT negative on the same day were included as controls. Individuals who attended the health post for vaccination, trauma, pregnancy test, antenatal care, and family planning were not considered. Index cases were independent of one another (i.e. they wouldn’t be found in the same cluster of households assessed for another index case). A total of 147 patients attended the clinic during the study period of whom 52 presented with main complaint of fever. Fever was confirmed in 44 of them of whom 24 were RDT positive for malaria. Patients who were referred to the nearby health center for better treatment (n = 4), were not residents of the study area (n = 2), or did not have completed information (n = 2) were excluded from the study. Controls were recruited among patients that presented to the same health post where the index case was recruited.

During the study period, 0–6 patients attended the clinic in a day and the health extension workers were available for 2–3 days a week from 9:00–12:00 a.m. The northern health post was opened occasionally and the same health extension workers were responsible and two index cases and two controls were recruited. Household members and members of the six nearest neighboring households of index cases and controls were tested for malaria by RDT within 2 days after the index or control cases were identified<sup>8</sup> and had their household geo-located with a handheld GPS receiver (GPSMAP 62 s; Garmin International). Neighboring households were selected based on proximity to the respective index case or control household upon obtaining permission from the head of the house. Socio-demographic and epidemiological data were collected, together with finger prick blood samples to diagnose malaria using RDT and to prepare dried blood spots (DBS) on 3MM Whatman papers (Whatman, Maidstone, UK) from all study participants including index cases, controls, neighbors and household members.

DNA was extracted from a 6 mm diameter punch, using MagNa Pure LC 2.0 Instrument and Total Nucleic Acid Kit—High Performance (Roche Life Sciences) with prior treatment using Buffer ATL (QIAGEN) and Proteinase K (QIAGEN)<sup>5</sup>. Parasites were quantified using species-specific qPCR that targeted the 18S small subunit rRNA gene<sup>69</sup>. qPCR positive samples were genotyped using 26 microsatellite markers for *P. falciparum* and 11 markers for *P. vivax* using previously established laboratory and data analyses protocols<sup>12,19–23</sup>. Mixed species infections were genotyped with both *P. falciparum* and *P. vivax* markers. Genotyping was successful in 92/93 *P. falciparum* and 99/123 *P. vivax* positive samples. For haplotype reconstruction in multiclonal infections, the predominant peak was included.

**Statistical analysis.** GraphPad Prism version 5.03 for Windows (GraphPad Software, La Jolla California USA, <http://www.graphpad.com>); STATA 13 (StataCorp., TX, USA) and R (R Core Team, Vienna, Austria) were used for data analysis. The QGIS software version 3.14.16 (QGIS developer team, Open Source Geospatial Foundation Project) was used to map study households. Continuous variables were presented as median and interquartile range (IQR). Tests of association between two categorical variables were performed using the Chi squared test, and comparison of non-normally distributed continuous variables by Mann–Whitney U test. Odds ratios (OR) with 95% confidence intervals (95% CI) were calculated using Generalized Estimating Equations to determine the association between being parasite positive and living in households of index cases or surrounding houses. These estimates were adjusted for clustering of species-specific observations from the same household with robust standard errors and used community members not belonging to index cases or neighboring households as a reference.

Multiplicity of Infection (MOI) was defined as the highest number of alleles detected by at least two loci, reducing the impact of false positive allele calls from a single, outlier locus<sup>12,22</sup>. Population-level genetic diversity was determined by the expected heterozygosity ( $H_E$ ) of each locus.  $H_E$  was calculated for each locus using the formula  $H_E = \frac{n}{n-1} [1 - \sum p_i^2]$ , where  $n$  = the number of isolates analyzed and  $p_i$  = the allele frequency of the  $i$ th allele in the population. Mean  $H_E$  was calculated by taking the average of  $H_E$  across all loci. To assess the genetic relatedness and clustering of *P. falciparum* and *P. vivax* infections, pairwise genetic relatedness was determined for each pair of samples using a modified identity by state (IBS) metric<sup>22</sup>. Briefly, pairwise IBS was calculated based on the number of shared alleles between isolates, allowing estimation of genetic relatedness from monoclonal as well as polyclonal samples. Samples were considered related if the pairwise genetic relatedness was greater than or equal to 0.5 (i.e. at least 50% of microsatellite markers shared the same alleles). This threshold reflects the degree of relatedness of meiotic siblings, which are the result of sexual recombination of different clones in the mosquito. Population structure was assessed using the program Maverick<sup>70</sup>, under the admixture model with 500 burn-in iterations, 5000 sampling iterations and 20 thermodynamic rungs.

**Consent for publication.** This work does not contain any individual person's data in any form (including any individual details, images or videos).

### Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 17 June 2020; Accepted: 2 November 2020

Published online: 17 November 2020

### References

1. Feachem, R. G. *et al.* Shrinking the malaria map: Progress and prospects. *Lancet* **376**, 1566–1578. [https://doi.org/10.1016/s0140-6736\(10\)61270-6](https://doi.org/10.1016/s0140-6736(10)61270-6) (2010).
2. Rabinovich, R. N. malERA: An updated research agenda for characterising the reservoir and measuring transmission in malaria elimination and eradication. *PLoS Med.* **14**, e1002452 (2017).
3. Slater, H. C., Walker, P. G., Bousema, T., Okell, L. C. & Ghani, A. C. The potential impact of adding ivermectin to a mass treatment intervention to reduce malaria transmission: A modelling study. *J. Infect. Dis.* **210**, 1972–1980. <https://doi.org/10.1093/infdis/jiu351> (2014).
4. Slater, H. C. *et al.* Assessing the impact of next-generation rapid diagnostic tests on *Plasmodium falciparum* malaria elimination strategies. *Nature* **528**, S94–101. <https://doi.org/10.1038/nature16040> (2015).
5. Tadesse, F. G. *et al.* The relative contribution of symptomatic and asymptomatic *Plasmodium vivax* and *Plasmodium falciparum* infections to the infectious reservoir in a low-endemic setting in Ethiopia. *Clin. Infect. Dis.* **66**, 1883–1891. <https://doi.org/10.1093/cid/cix1123> (2018).
6. Bousema, T. *et al.* Hitting hotspots: Spatial targeting of malaria for control and elimination. *PLoS Med.* **9**, e1001165. <https://doi.org/10.1371/journal.pmed.1001165> (2012).
7. Bejon, P. *et al.* A micro-epidemiological analysis of febrile malaria in Coastal Kenya showing hotspots within hotspots. *Elife* **3**, e02130–e02130. <https://doi.org/10.7554/eLife.02130> (2014).
8. Sturrock, H. J. *et al.* Reactive case detection for malaria elimination: Real-life experience from an ongoing program in Swaziland. *PLoS ONE* **8**, e63830. <https://doi.org/10.1371/journal.pone.0063830> (2013).
9. Mosha, J. F. *et al.* Hot spot or not: A comparison of spatial statistical methods to predict prospective malaria infections. *Malar. J.* **13**, 53. <https://doi.org/10.1186/1475-2875-13-53> (2014).
10. Moonen, B. *et al.* Operational strategies to achieve and maintain malaria elimination. *Lancet* **376**, 1592–1603. [https://doi.org/10.1016/S0140-6736\(10\)61269-X](https://doi.org/10.1016/S0140-6736(10)61269-X) (2010).
11. Pampana, E. *A Textbook of Malaria Eradication.* (1963).
12. Pringle, J. C. *et al.* Genetic evidence of focal *Plasmodium falciparum* transmission in a pre-elimination setting in southern Province, Zambia. *J. Infect. Dis.* **219**, 1254–1263. <https://doi.org/10.1093/infdis/jiy640> (2019).
13. Taylor, A. R. *et al.* Quantifying connectivity between local *Plasmodium falciparum* malaria parasite populations using identity by descent. *PLoS Genet.* **13**, e1007065. <https://doi.org/10.1371/journal.pgen.1007065> (2017).
14. Tessema, S. K. *et al.* Applying next-generation sequencing to track falciparum malaria in sub-Saharan Africa. *Malar. J.* **18**, 268. <https://doi.org/10.1186/s12936-019-2880-1> (2019).
15. Pringle, J. C. *et al.* High *Plasmodium falciparum* genetic diversity and temporal stability despite control efforts in high transmission settings along the international border between Zambia and the Democratic Republic of the Congo. *Malar. J.* **18**, 400. <https://doi.org/10.1186/s12936-019-3023-4> (2019).
16. Liu, Y. *et al.* Confirmation of the absence of local transmission and geographic assignment of imported falciparum malaria cases to China using microsatellite panel. *Malar. J.* **19**, 244. <https://doi.org/10.1186/s12936-020-03316-3> (2020).
17. Manrique, P. *et al.* Microsatellite analysis reveals connectivity among geographically distant transmission zones of *Plasmodium vivax* in the Peruvian Amazon: A critical barrier to regional malaria elimination. *PLoS Neglect. Trop. Dis.* <https://doi.org/10.1371/journal.pntd.0007876> (2019).



18. Amambua-Ngwa, A. *et al.* Long-distance transmission patterns modelled from SNP barcodes of *Plasmodium falciparum* infections in The Gambia. *Sci. Rep.* **9**, 13515. <https://doi.org/10.1038/s41598-019-49991-4> (2019).
19. Anderson, T., Su, X.-Z., Bockarie, M., Lagog, M. & Day, K. Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology* **119**, 113–125 (1999).
20. Koepfli, C. *et al.* A large *Plasmodium vivax* reservoir and little population structure in the south Pacific. *PLoS ONE* **8**, e66041. <https://doi.org/10.1371/journal.pone.0066041> (2013).
21. Koepfli, C. *et al.* *Plasmodium vivax* diversity and population structure across four continents. *PLoS Negl. Trop. Dis.* **9**, e0003872. <https://doi.org/10.1371/journal.pntd.0003872> (2015).
22. Tessema, S. *et al.* Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. *Elife* **8**, e43510. <https://doi.org/10.7554/eLife.43510> (2019).
23. Karunaweera, N., Ferreira, M., Hartl, D. & Wirth, D. Fourteen polymorphic microsatellite DNA markers for the human malaria parasite *Plasmodium vivax*. *Mol. Ecol. Notes* **7**, 172–175 (2007).
24. Kittichai, V., Koepfli, C., Nguitraogot, W., Sattabongkot, J. & Cui, L. Substantial population structure of *Plasmodium vivax* in Thailand facilitates identification of the sources of residual transmission. *PLoS Negl. Trop. Dis.* **11**, e0005930. <https://doi.org/10.1371/journal.pntd.0005930> (2017).
25. Waltmann, A. *et al.* Increasingly inbred and fragmented populations of *Plasmodium vivax* associated with the eastward decline in malaria transmission across the Southwest Pacific. *PLoS Negl. Trop. Dis.* **12**, e0006146. <https://doi.org/10.1371/journal.pntd.0006146> (2018).
26. Noor, A. M. *et al.* The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000–10: A spatial and temporal analysis of transmission intensity. *Lancet* **383**, 1739–1747. [https://doi.org/10.1016/S0140-6736\(13\)62566-0](https://doi.org/10.1016/S0140-6736(13)62566-0) (2014).
27. Clark, T. D. *et al.* Factors determining the heterogeneity of malaria incidence in children in Kampala, Uganda. *J. Infect. Dis.* **198**, 393–400 (2008).
28. Bhunia, G. S., Kesari, S., Chatterjee, N., Kumar, V. & Das, P. Spatial and temporal variation and hotspot detection of kala-azar disease in Vaishali district (Bihar), India. *BMC Infect. Dis.* **13**, 64 (2013).
29. Chowell, G. *et al.* Spatial and temporal dynamics of dengue fever in Peru: 1994–2006. *Epidemiol. Infect.* **136**, 1667–1677 (2008).
30. Ostfeld, R. S., Glass, G. E. & Keesing, F. Spatial epidemiology: An emerging (or re-emerging) discipline. *Trends Ecol. Evol.* **20**, 328–336 (2005).
31. Bejon, P. *et al.* Stable and unstable malaria hotspots in longitudinal cohort studies in Kenya. *PLoS Med.* <https://doi.org/10.1371/journal.pmed.1000304> (2010).
32. World Health Organization. *Disease surveillance for malaria elimination: an operational manual*, [https://apps.who.int/iris/bitstream/handle/10665/44852/9789241503334\\_eng.pdf;jsessionid=7476190A6D5CC72F80C2BF900AF55F86?sequence=1](https://apps.who.int/iris/bitstream/handle/10665/44852/9789241503334_eng.pdf;jsessionid=7476190A6D5CC72F80C2BF900AF55F86?sequence=1) (2012).
33. Roh, M. E. *et al.* High genetic diversity of *Plasmodium falciparum* in the low-transmission setting of the Kingdom of Eswatini. *J. Infect. Dis.* **220**, 1346–1354. <https://doi.org/10.1093/infdis/jiz305> (2019).
34. Stresman, G. H. *et al.* A method of active case detection to target reservoirs of asymptomatic malaria and gametocyte carriers in a rural area in Southern Province, Zambia. *Malar. J.* **9**, 265. <https://doi.org/10.1186/1475-2875-9-265> (2010).
35. Stresman, G. H. *et al.* Do hotspots fuel malaria transmission: A village-scale spatio-temporal analysis of a 2-year cohort study in The Gambia. *BMC Med.* **16**, 160. <https://doi.org/10.1186/s12916-018-1141-4> (2018).
36. Chitnis, N. *et al.* Theory of reactive interventions in the elimination and control of malaria. *Malar. J.* **18**, 266. <https://doi.org/10.1186/s12936-019-2882-z> (2019).
37. Stuck, L. *et al.* Malaria infection prevalence and sensitivity of reactive case detection in Zanzibar. *Int. J. Infect. Dis.* **97**, 337–346. <https://doi.org/10.1016/j.ijid.2020.06.017> (2020).
38. Jennison, C. *et al.* *Plasmodium vivax* populations are more genetically diverse and less structured than sympatric *Plasmodium falciparum* populations. *PLoS Negl. Trop. Dis.* **9**, e0003634. <https://doi.org/10.1371/journal.pntd.0003634> (2015).
39. Koepfli, C., Waltmann, A., Ome-Kaius, M., Robinson, L. J. & Mueller, I. Multiplicity of infection is a poor predictor of village-level *Plasmodium vivax* and *P. falciparum* population prevalence in the southwest Pacific. *Open Forum Infect. Dis.* <https://doi.org/10.1093/ofid/ofy240> (2018).
40. Noviyanti, R. *et al.* Contrasting transmission dynamics of co-endemic *Plasmodium vivax* and *P. falciparum*: Implications for malaria control and elimination. *PLoS Negl. Trop. Dis.* **9**, e0003739. <https://doi.org/10.1371/journal.pntd.0003739> (2015).
41. Chenet, S. M., Schneider, K. A., Villegas, L. & Escalante, A. A. Local population structure of *Plasmodium*: Impact on malaria control and elimination. *Malar. J.* **11**, 412. <https://doi.org/10.1186/1475-2875-11-412> (2012).
42. Iwagami, M. *et al.* Population structure and transmission dynamics of *Plasmodium vivax* in the Republic of Korea based on microsatellite DNA analysis. *PLoS Negl. Trop. Dis.* **6**, e1592–e1592. <https://doi.org/10.1371/journal.pntd.0001592> (2012).
43. Golassa, L. & White, M. T. Population-level estimates of the proportion of *Plasmodium vivax* blood-stage infections attributable to relapses among febrile patients attending Adama Malaria Diagnostic Centre, East Shoa Zone, Oromia, Ethiopia. *Malar. J.* **16**, 301. <https://doi.org/10.1186/s12936-017-1944-3> (2017).
44. Chu, C. S. & White, N. J. Management of relapsing *Plasmodium vivax* malaria. *Expert Rev. Anti Infect. Ther.* **14**, 885–900. <https://doi.org/10.1080/14787210.2016.1220304> (2016).
45. Littrell, M. *et al.* Case investigation and reactive case detection for malaria elimination in northern Senegal. *Malar. J.* **12**, 331. <https://doi.org/10.1186/1475-2875-12-331> (2013).
46. Baird, J. K., Schwartz, E. & Hoffman, S. L. Prevention and treatment of vivax malaria. *Curr. Infect. Dis. Rep.* **9**, 39–46 (2007).
47. Robinson, L. J. *et al.* Strategies for understanding and reducing the *Plasmodium vivax* and *Plasmodium ovale* hypnozoite reservoir in Papua New Guinean children: A randomised placebo-controlled trial and mathematical model. *PLoS Med.* **12**, e1001891. <https://doi.org/10.1371/journal.pmed.1001891> (2015).
48. Douglas, N. M. *et al.* Gametocyte dynamics and the role of drugs in reducing the transmission potential of *Plasmodium vivax*. *J. Infect. Dis.* **208**, 801–812. <https://doi.org/10.1093/infdis/jit261> (2013).
49. Karyana, M. *et al.* Malaria morbidity in Papua Indonesia, an area with multidrug resistant *Plasmodium vivax* and *Plasmodium falciparum*. *Malar. J.* **7**, 148 (2008).
50. Van den Eede, P. *et al.* *Plasmodium vivax* sub-patent infections after radical treatment are common in Peruvian patients: Results of a 1-year prospective cohort study. *PLoS ONE* **6**, e16257. <https://doi.org/10.1371/journal.pone.0016257> (2011).
51. Macdonald, G. *The Epidemiology and Control of Malaria*. (1957).
52. Nigatu, W., Abebe, M. & Dejene, A. *Plasmodium vivax* and *P. falciparum* epidemiology in Gambella, south-west Ethiopia. *Trop. Med. Parasitol.* **43**, 181–185 (1992).
53. Sombon, P., Suwonkerd, W. & Lines, J. D. Susceptibility of Thai zoophilic *Anophelines* and suspected malaria vectors to local strains of human malaria parasites. *Southeast Asian J. Trop. Med. Public Health* **25**, 766–770 (1994).
54. Sattabongkot, J., Tsuboi, T., Zollner, G. E., Sirichaisinthop, J. & Cui, L. *Plasmodium vivax* transmission: Chances for control?. *Trends Parasitol.* **20**, 192–198. <https://doi.org/10.1016/j.pt.2004.02.001> (2004).
55. Olliaro, P. L. *et al.* Implications of *Plasmodium vivax* biology for control, elimination, and research. *Am. J. Trop. Med. Hyg.* **95**, 4–14. <https://doi.org/10.4269/ajtmh.16-0160> (2016).
56. Gonçalves, B. P. *et al.* Examining the human infectious reservoir for *Plasmodium falciparum* malaria in areas of differing transmission intensity. *Nat. Commun.* **8**, 1133 (2017).

57. Mueller, I. *et al.* Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. *Lancet Infect. Dis.* [https://doi.org/10.1016/s1473-3099\(09\)70177-x](https://doi.org/10.1016/s1473-3099(09)70177-x) (2009).
58. Assefa, A. *et al.* Glucose-6-phosphate dehydrogenase (G6PD) deficiency in Ethiopia: Absence of common African and Mediterranean allelic variants in a nationwide study. *Malar. J.* **17**, 388. <https://doi.org/10.1186/s12936-018-2538-4> (2018).
59. Carter, T. E. *et al.* Glucose-6-phosphate dehydrogenase deficiency genetic variants in malaria patients in southwestern Ethiopia. *Am. J. Trop. Med. Hyg.* **98**, 83–87. <https://doi.org/10.4269/ajtmh.17-0557> (2018).
60. Kießling, N. *et al.* Glucose-6-phosphate dehydrogenase activity measured by spectrophotometry and associated genetic variants from the Oromiya zone, Ethiopia. *Malar. J.* **17**, 358. <https://doi.org/10.1186/s12936-018-2510-3> (2018).
61. Lo, E. *et al.* Prevalence and distribution of G6PD deficiency: Implication for the use of primaquine in malaria treatment in Ethiopia. *Malar. J.* **18**, 340. <https://doi.org/10.1186/s12936-019-2981-x> (2019).
62. Shitaye, G. *et al.* Low and heterogeneous prevalence of glucose-6-phosphate dehydrogenase deficiency in different settings in Ethiopia using phenotyping and genotyping approaches. *Malar. J.* **17**, 281. <https://doi.org/10.1186/s12936-018-2437-8> (2018).
63. Tadesse, F. G. *et al.* Submicroscopic carriage of *Plasmodium falciparum* and *Plasmodium vivax* in a low endemic area in Ethiopia where no parasitaemia was detected by microscopy or rapid diagnostic test. *Malar. J.* **14**, 303. <https://doi.org/10.1186/s12936-015-0821-1> (2015).
64. Tsegaye, A., Golassa, L., Mamo, H. & Erko, B. Glucose-6-phosphate dehydrogenase deficiency among malaria suspects attending Gambella hospital, southwest Ethiopia. *Malar. J.* **13**, 438. <https://doi.org/10.1186/1475-2875-13-438> (2014).
65. Abreha, T. *et al.* Comparison of artemether-lumefantrine and chloroquine with and without primaquine for the treatment of *Plasmodium vivax* infection in Ethiopia: A randomized controlled trial. *PLoS Med.* **14**, e1002299. <https://doi.org/10.1371/journal.pmed.1002299> (2017).
66. Commons, R. J. *et al.* The effect of chloroquine dose and primaquine on *Plasmodium vivax* recurrence: A WorldWide Antimalarial Resistance Network systematic review and individual patient pooled meta-analysis. *Lancet Infect. Dis.* **18**, 1025–1034. [https://doi.org/10.1016/s1473-3099\(18\)30348-7](https://doi.org/10.1016/s1473-3099(18)30348-7) (2018).
67. Taylor, W. R. J. *et al.* Short-course primaquine for the radical cure of *Plasmodium vivax* malaria: A multicentre, randomised, placebo-controlled non-inferiority trial. *Lancet* **394**, 929–938. [https://doi.org/10.1016/s0140-6736\(19\)31285-1](https://doi.org/10.1016/s0140-6736(19)31285-1) (2019).
68. Peterson, I., Borrell, L. N., El-Sadr, W. & Teklehaimanot, A. Individual and household level factors associated with malaria incidence in a highland region of Ethiopia: A multilevel analysis. *Am. J. Trop. Med. Hyg.* **80**, 103–111. <https://doi.org/10.4269/ajtmh.2009.80.103> (2009).
69. Tadesse, F. G. *et al.* The shape of the iceberg: Quantification of submicroscopic *Plasmodium falciparum* and *Plasmodium vivax* parasitaemia and gametocytaemia in five low endemic settings in Ethiopia. *Malar. J.* **16**, 99 (2017).
70. Verity, R. & Nichols, R. A. Estimating the number of subpopulations (K) in structured populations. *Genetics* **203**, 1827–1839. <https://doi.org/10.1534/genetics.115.180992> (2016).

## Acknowledgements

We would like to acknowledge the study participants for their willingness and Adama district malaria focal person, Teshome Bacha, for assisting in organizing the field work.

## Author contributions

Conceived and designed the study protocol: F.G.T., T.B., E.G., M.B.; Carried out the sample collection: M.B., F.G.T., E.G., W.C., G.S.; Carried out geospatial analysis and prepared the map of the study households: E.F.M.; Carried out the molecular experiments (DNA extraction and qPCR): K.L., M.B., F.G.T., W.C., G.S.; Carried out the genotyping experiments: S.K.T., C.K., T.H., B.G.; Carried out data analysis and interpretation of results: F.G.T., S.K.T., B.G., C.K., C.D., T.B.; Drafted the manuscript: F.G.T., M.B., S.K.T., C.K., T.B., B.G., C.D.; All authors critically revised the manuscript for intellectual content, read and approved the final manuscript.

## Funding

This work was supported by grants from the Netherlands Organization for Scientific Research (Vidi fellowship) [NWO project number 016.158.306] and fellowship from the European Research Council [ERC-2014-StG 639776] to TB; the Bill & Melinda Gates foundation [INDIE; OP P1173572] to TB, FGT and CD; the Netherlands organization for international cooperation in higher education (Nuffic) [grant number NFP-PhD.14/150] to FGT and the Armauer Hansen Research Institute (via its core funding from Norwegian Agency for Development Cooperation and Swedish International Development Cooperation) to FGT and MK. BG is a Chan Zuckerberg Biohub investigator.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to F.G.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020