Measuring changes in *Plasmodium falciparum* census population size in response to sequential malaria control interventions

Kathryn E. Tiedje[1,2§], Qi Zhan[3,4§], Shazia Ruybal-Pésantez[2†], Gerry Tonkin-Hill[2,5], Qixin He[4‡], Mun Hua Tan[1], Dionne C. Argyropoulos[1], Samantha L. Deed[1,2], Anita Ghansah[6], Oscar Bangre[7], Abraham R. Oduro[7], Kwadwo A. Koram[8], Mercedes Pascual[9,10], and Karen P. Day[1,2]

[1] Department of Microbiology and Immunology, Bio21 Institute and Peter Doherty Institute, The University of Melbourne; Melbourne, Australia
[2] School of BioSciences, Bio21 Institute, The University of Melbourne; Melbourne, Australia
[3] Committee on Genetics, Genomics and Systems Biology, The University of Chicago; Chicago, Illinois, USA
[4] Department of Ecology and Evolution, The University of Chicago; Chicago, Illinois, USA
[5] Bioinformatics Division, Walter and Eliza Hall Institute; Melbourne, Australia
[6] Department of Parasitology, Noguchi Memorial Institute for Medical Research, University of Ghana; Legon, Ghana
[7] Navrongo Health Research Centre, Ghana Health Service; Navrongo, Ghana
[8] Epidemiology Department, Noguchi Memorial Institute for Medical Research, University of Ghana; Legon, Ghana
[9] Department of Biology and Department of Environmental Sciences, New York University, New York, New York, USA
[10] Santa Fe Institute, Santa Fe, New Mexico, USA

† *Current affiliations:* Department of Infectious Disease Epidemiology and MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, United Kingdom

‡ *Current affiliation:* Department of Biological Sciences, Purdue University; West Lafayette, United States

§ These authors share first authorship

* Corresponding Author: Karen P. Day, Department of Microbiology and Immunology, Bio21 Institute and Peter Doherty Institute, The University of Melbourne, 30 Flemington Road, Parkville, VIC, 3011, Australia. E-mail: Karen.Day@unimelb.edu.au, Phone: +613 8344 2377

RUNNING TITLE: Impact of IRS and SMC on census population size

ABSTRACT

Here we introduce a new endpoint "census population size" to evaluate the epidemiology and control of *Plasmodium falciparum* infections, where the parasite, rather than the infected human host, is the unit of measurement. To calculate census population size, we rely on a definition of parasite variation known as multiplicity of infection (MOI_{var}), based on the hyper-diversity of the *var* multigene family. We present a Bayesian approach to estimate MOI_{var} from sequencing and counting the number of unique DBLα tags (or DBLα types) of *var* genes, and derive from it census population size by summation of MOI_{var} in the human population. We track changes in this parasite population size and structure through sequential malaria interventions by indoor residual spraying (IRS) and seasonal malaria chemoprevention (SMC) from 2012 to 2017 in an area of high-seasonal malaria transmission in northern Ghana. Following IRS, which reduced transmission intensity by > 90% and decreased parasite prevalence by ~40-50%, significant reductions in *var* diversity, MOI_{var}, and population size were observed in ~2,000 humans across all ages. These changes, consistent with the loss of diverse parasite genomes, were short lived and 32-months after IRS was discontinued and SMC was introduced, *var* diversity and population size rebounded in all age groups except for the younger children (1-5 years) targeted by SMC. Despite major perturbations from IRS and SMC interventions, the parasite population remained very large and retained the *var* population genetic characteristics of a high-transmission system (high *var* diversity; low *var* repertoire similarity) demonstrating the resilience of *P. falciparum* to short-term interventions in high-burden countries of sub-Saharan Africa.

INTRODUCTION

Malaria in high-transmission endemic areas of sub-Saharan Africa (SSA) is characterised by vast diversity of the *Plasmodium falciparum* parasites from the perspective of antigenic variation (Chen et al., 2011; Day et al., 2017; Otto et al., 2019; Ruybal-Pesántez et al., 2022, 2017). As with other hosts of hyper-variable pathogens (Futse et al., 2008), children experiencing clinical episodes of malaria, eventually become immune to disease but not to infection. This results in a large reservoir of chronic asymptomatic infections, in hosts of all ages, sustaining transmission to mosquitos. Given the goal of malaria eradication by 2050, it is therefore of interest to examine how the parasite population changes following perturbation by major intervention efforts, both in terms of its size and underlying population genetics.

So, what do we mean by the parasite population size in the case of *P. falciparum* and how do we measure it? Parasite prevalence, detected by microscopy or more sensitive molecular diagnostics (e.g., PCR), describes the proportion of infected human hosts. Studies of *P. falciparum* genetic diversity have shown that the majority of people in high-transmission endemic areas harbour diverse multiclonal infections measured as the complexity or multiplicity of infection (MOI) (e.g., (Anderson et al., 2000; Paul et al., 1995; Smith et al., 1999; Sumner et al., 2021)) with complex population dynamics (Bruce et al., 2000; Farnert et al., 1997). These genetic data indicate much larger parasite population sizes than observed by prevalence of infection alone. Thus, from an ecological perspective, we can consider a human host as a patch carrying a number of "antigenically-distinct infections" of *P. falciparum*. The sum of these antigenically-distinct infections over all hosts provides us with a census of the parasite count of relevance to monitoring and evaluating malaria interventions. We refer to this census population size hereafter

78 simply as population size but make clear that this measure is distinct from effective population size (N_e)
79 as measured by neutral variation.
80
81 Diversity of *P. falciparum* single copy surface antigen genes such as circumsporozoite protein (*csp*),
82 merozoite surface protein 1 (*msp1*) or 2 (*msp2*), and apical membrane antigen 1 (*ama1*) have each been
83 widely used to measure MOI (e.g., (Falk et al., 2006; Lerch et al., 2017; Nelson et al., 2019)). They have
84 become part of newer genetic panels (e.g., Paragon v1 (Tessema et al., 2020) and AMPLseq v1 (LaVerriere
85 et al., 2022)) specifically for MOI determination. Typically, MOI is reported as the maximum number of
86 alleles or single locus haplotypes present at the most diverse of these antigen-encoding loci rather than
87 the number of unique multilocus haplotypes of these genes combined, as it is challenging to accurately
88 reconstruct or phase these haplotypes in hosts with an MOI > 3 (Lerch et al., 2019). Each of these genes
89 is under balancing selection with a few geographically-common haplotypes and many very rare haplotypes
90 in moderate- to high-transmission settings (Markwalter et al., 2022; Sumner et al., 2021). Where there is
91 a high probability of co-occurrence of two or more common single locus haplotypes in a host, genotyping
92 each of these single copy antigen genes alone will underestimate MOI. Single nucleotide polymorphism
93 (SNP) panels have been used to define the presence of multiclonal infections with limited reliability to
94 estimate MOI for highly complex infections, typical in high transmission, even with the use of
95 computational methods (Labbé et al., 2023).
96
97 As an alternative to genotyping single copy antigen genes and biallelic SNP panels to estimate MOI, we
98 have proposed the use of a fingerprinting methodology known as *var*coding to genotype the hyper-diverse
99 *var* multigene family (~50-60 *var* genes per haploid genome). This method employs a ~450bp region of a

100 *var* gene, known as a DBLα tag encoding the immunogenic Duffy-binding-like alpha (DBLα) domain of
101 PfEMP1 (*Plasmodium falciparum* erythrocyte membrane protein 1), the major surface antigen of the
102 blood stages (Zhang and Deitsch, 2022). Bioinformatic analyses of a large database of exon 1 sequences
103 of *var* genes showed a predominantly 1-to-1 DBLα-*var* relationship, such that each DBLα tag typically
104 represents a unique *var* gene, especially in high transmission (Tan et al., 2023). The extensive diversity of
105 DBLα tags together with the very low percentage of *var* genes shared between parasites (Chen et al.,
106 2011; Day et al., 2017; Ruybal-Pesántez et al., 2022, 2017) facilitate measuring MOI by amplifying, pooling,
107 sequencing, and counting the number of unique DBLα tags (or DBLα types) in a host (Ruybal-Pesántez et
108 al., 2022; Tiedje et al., 2022). From a single PCR with degenerate primers and amplicon sequencing, the
109 method specifically counts the most diverse DBLα types, designated non-upsA, per infection to arrive at
110 a metric we call MOI_{var}. It is not based on assigning haplotypes but exploits the fact that *var* repertoires
111 are non-overlapping, especially in high transmission. Instead, it assumes a set number of non-upsA types
112 per genome based on repeated sampling of 3D7 control isolates accounting for PCR sampling errors to
113 calculate MOI_{var} (Ghansah et al., 2023; Ruybal-Pesántez et al., 2022; Tiedje et al., 2022). Consequently,
114 rather than looking at the diversity of a single copy antigen-encoding gene like *csp, msp2,* or *ama1* to
115 calculate MOI, by *var*coding we are looking at sets of up to 45 non-upsA DBLα types per genome. Prior
116 work has shown that *var*coding is more sensitive to measure MOI in high transmission where there is an
117 extremely high prevalence of multiclonal infections that cannot be accurately phased with either biallelic
118 SNP panels (Ghansah et al., 2023; Labbé et al., 2023; Tessema et al., 2020; Watson et al., 2021) or
119 combinations of single copy antigen genes (Sumner et al., 2021).
120

121 Here we report an investigation of changes in parasite census population size and structure through two
122 sequential malaria control interventions between 2012 and 2017 in Bongo District located in the Upper
123 East Region of Ghana, one of the 12 highest burden countries in Africa (World Health Organization, 2022).
124 We present a novel Bayesian modification to the published *var*coding approach (Ghansah et al., 2023;
125 Ruybal-Pesántez et al., 2022; Tiedje et al., 2022) that takes into account under-sampling of non-upsA
126 DBLα types in an isolate to estimate MOI_{var} (Ruybal-Pesántez et al., 2022; Tiedje et al., 2022) and therefore
127 population size. We document *P. falciparum* prevalence as well as *var* diversity and population structure
128 from baseline in 2012 through a major perturbation by a short-term indoor residual spraying (IRS)
129 campaign managed under operational conditions, which reduced transmission intensity by > 90% as
130 measured by the entomological inoculation rate (EIR) and decreased parasite prevalence by ~40-50%
131 (Tiedje et al., 2022). Next, we followed what happened to parasite population size more than two years
132 after the IRS intervention was discontinued and seasonal malaria chemoprevention (SMC) was introduced
133 for children between the ages of 3-59 months (i.e., < 5 years) (Wagman et al., 2018). Detectable changes
134 in parasite population size were seen as a consequence of the IRS intervention but this quantity rapidly
135 rebounded 32-months after the intervention ceased. Overall, throughout the IRS, SMC, and subsequent
136 rebound, the parasite population in humans remained large in size and retained the *var* population
137 genetic characteristics of high transmission (i.e., high *var* diversity, low *var* repertoire overlap),
138 demonstrating the overall resilience of the species to survive significant short-term perturbations.

MATERIALS AND METHODS

Human subject ethical approval

The study was reviewed/approved by the ethics committees at the Navrongo Health Research Centre (Ghana), Noguchi Memorial Institute for Medical Research (Ghana), The University of Melbourne (Australia), and The University of Chicago (United States). Details on the study area, study population, inclusion/exclusion criteria, and data collection procedures have been previously described (Tiedje et al., 2022, 2017).

Study design and sample collection

Using an interrupted time-series study design, four age-stratified cross-sectional surveys of ~2,000 participants per survey were undertaken to investigate the impacts of IRS and SMC in combination with long-lasting insecticidal nets (LLINs) impregnated with pyrethroids under operational conditions on the asymptomatic *P. falciparum* reservoir from two proximal catchment areas (i.e., Vea/Gowrie and Soe, with a sampling area of ~60 km^2) in Bongo District, Ghana (Tiedje et al., 2022, 2017) (Table supplement 1). Bongo District, located in the Upper East Region, is categorized as high-seasonal malaria transmission based on the World Health Organization's (WHO) "A Framework for Malaria Elimination" (WHO/GMP, 2017) where *P. falciparum* prevalence was ≥ 35% at baseline in 2012 (Tiedje et al., 2022, 2017). These ~2,000 participants of all ages (1-97 years) represent ~15% of the total population that reside in these two catchment areas in Bongo (Tiedje et al., 2017). The four cross-sectional surveys were completed at the end of the wet season (i.e., high-transmission season) and the study can be separated into four distinct study time points: (1) October 2012 (Survey 1) prior to the IRS and SMC (i.e., baseline), (2) October 2014 (Survey 2) two months after the second round of IRS (Actellic 50EC), (3) October 2015 (Survey 3) seven

161 months after the third round of IRS using a long-acting non-pyrethroid insecticide (Actellic 300CS) (Gogue
162 et al., 2020; US President's Malaria Initiative Africa IRS (AIRS) Project, 2016), and finally (4) October 2017
163 (Survey 4) 32-months after the discontinuation of IRS, but during the deployment of SMC to all children
164 3-59 months (i.e., < 5 years) (Figure 1). LLINs (i.e., PermaNet 2.0, Olyset, or DawaPlus 2.0) were mass
165 distributed in Bongo District by the National Malaria Elimination Programme (NMEP)/Ghana Health
166 Service (GHS) between 2010-2012 and again in 2016 following the discontinuation of IRS (Gogue et al.,
167 2020; Tiedje et al., 2022; US Agency for International Development (USAID) Global Health Supply Chain
168 Program, 2020). In addition, to maintain high coverage of LLINs between these campaigns continuous
169 distribution was undertaken using routine services (i.e., antenatal clinics, school distributions,
170 immunization visits, etc.) (Tiedje et al., 2022). Over the course of this study self-reported LLIN usage the
171 previous night remained high across all age groups, from 89.1% in 2012 (pre-IRS), 83.5% in 2014 (during
172 IRS), 90.6% in 2015 (post-IRS), to 96.8% in 2017 (SMC). Details on the study area, study population, and
173 data collection procedures have been previously described (Tiedje et al., 2022, 2017).
174

175 Details of the IRS and SMC interventions

176 Starting in 2013, the AngloGold Ashanti Malaria Control Programme (AGAMal) in a public-private
177 partnership with the Global Fund, scaled up IRS across all of the Upper East Region of northern Ghana
178 (Gogue et al., 2020). As part of this initiative, three rounds of IRS with organophosphate formulations (i.e.,
179 non-pyrethroid) were rolled out prior to the start of the wet season between 2013 and 2015 (Figure 1A)
180 in Bongo District (Tiedje et al., 2022). Based on AGAMal's operational reports, IRS coverage in Bongo
181 District was 91.8% in Round 1, 95.6% in Round 2, and finally 96.6% in Round 3 (AGAMal, personal
182 communication). To monitor the impact of the IRS on the local vector population, monthly entomological

183 collections were undertaken between February 2013 and September 2015 (Tiedje et al., 2022). Using
184 these surveys, we observed that the monthly entomological inoculation rate (EIR) (infective
185 bites/person/month (ib/p/m)), a measure of local transmission intensity, declined by > 90% at the peak of
186 the wet season between August 2013 (pre-IRS) (EIR = 5.3 ib/p/m) and August 2015 (post-IRS) (EIR = 0.4
187 ib/p/m) (Tiedje et al., 2022). Following the IRS, SMC was rolled out in the Upper East Region by the
188 NMEP)/(GHS) starting in 2016 (Gogue et al., 2020) (Figure 1A). SMC is the intermittent administration of
189 full treatment courses of an antimalarial to children between the ages of 3-59 months (i.e., < 5 years)
190 (WHO, 2012). Like other countries, the SMC drug combination of choice in Ghana is amodiaquine plus
191 sulfadoxine-pyrimethamine, which is administered at monthly intervals during the high-transmission
192 season (i.e., wet season) (WHO, 2012). The goal of this age-targeted intervention is to both clear current
193 malaria infections and prevent malarial illness by maintaining a therapeutic concentration of an
194 antimalarial in blood over the period of greatest risk (i.e., high-transmission season). Reported SMC
195 coverage in Bongo District was 92.6% in 2016 (two rounds between August – September 2016) and 94.6%
196 in 2017 (four rounds between September – December 2017) (Gogue et al., 2020).
197

198 ***Var* genotyping and sequence analysis**

199 Genomic DNA was extracted from the dried blood spots for all participants with a confirmed microscopic
200 asymptomatic *P. falciparum* infection (i.e., isolate) (2,572 isolates, Table supplement 2) using the QIAamp
201 DNA mini kit (QIAGEN, USA) as previously described (Tiedje et al., 2017). For *var* genotyping or *var*coding,
202 the sequence region within the *var* genes encoding the DBLα domains of PfEMP1 were amplified using a
203 single-step PCR, pooled, and sequenced on an Illumina platform using the MiSeq 2×300bp paired-end
204 protocol (Supplemental Methods, Figure supplement 1). The raw sequence data was then processed using

205 our previously published customized bioinformatic pipelines (Supplemental Methods, Figure supplement
206 2). For additional information on the use of these bioinformatic pipelines a detailed tutorial is available
207 (https://github.com/UniMelb-Day-Lab/tutorialDBLalpha).
208
209 DBLα sequencing data was obtained from 2,397 *P. falciparum* isolates (93.2%) (Table supplement 2). This
210 genotyping success was acceptable given that we were working with low-density asymptomatic infections
211 (Table supplement 1). Using a cut-off of ≥ 20 DBLα types, DBLα sequencing data was obtained from 2,099
212 *P. falciparum* isolates (81.6%) with a total of 289,049 DBLα sequences and 53,238 unique DBLα types
213 being identified in the study population (Table supplement 2). The median *P. falciparum* density was ~4
214 times higher for isolates with ≥ 20 DBLα types compared to those that gave no or < 20 DBLα types (520
215 [IQR: 200-1880] parasites/µL vs. 120 [IQR: 40-200] parasites/µL, respectively).
216
217 DBLα type diversity
218 We monitored the impacts of the sequential interventions (i.e., IRS and SMC) on diversity by measuring
219 changes in the population genetics of DBLα types at the population level (i.e., *P. falciparum* reservoir).
220 Diversity was monitored using two measures, DBLα type richness and DBLα type frequency. Richness was
221 defined as the number of unique DBLα types observed (i.e., DBLα type pool size) in each survey or study
222 time point (i.e., 2012, 2014, 2015, and 2017). DBLα type richness, however, does not provide any
223 information about the relative frequencies of the different DBLα types in the population, as they are all
224 weighted equally whether they are observed once or more frequently (e.g., observed in > 20 isolates per
225 survey). To further examine the impacts of the interventions on DBLα type diversity we also assessed the
226 frequency of each unique DBLα type in 2012, 2014, 2015, and 2017. Here we defined DBLα type frequency

227 as the number of times (i.e., number of isolates) a DBLα type was observed in each survey. Both upsA and
228 non-upsA DBLα type diversities were measured due to their different biological features, chromosomal
229 positions (i.e., subtelomeric regions vs. internal or central regions), as well as population genetics (Falk et
230 al., 2009; Gardner et al., 2002; Jensen et al., 2004; Kaestli et al., 2006; Kalmbach et al., 2010; Kraemer et
231 al., 2007; Kraemer and Smith, 2006; Kyriacou et al., 2006; Lavstsen et al., 2003; Normark et al., 2007;
232 Rottmann et al., 2006; Warimwe et al., 2012, 2009; Zhang and Deitsch, 2022). The proportions of upsA
233 and non-upsA *var* genes in a repertoire or single genome have been defined as ~15-20% and ~80-85%,
234 respectively, based on whole genome sequencing (Rask et al., 2010). The upsA and non-upsA DBLα type
235 proportions were partitioned as expected in our analyses, with the median proportions at the repertoire
236 level being comparable in 2012 (19% upsA and 81% non-upsA), 2014 (22% upsA and 78% non-upsA), 2015
237 (21% upsA and 79% non-upsA), and 2017 (20% upsA and 80% non-upsA).
238
239 Repertoire similarity as defined by pairwise type sharing
240 To estimate genetic similarity between the DBLα repertoires (i.e., unique DBLα types identified in each
241 isolate) identified from two isolates, pairwise type sharing (PTS) was calculated between all pairs of
242 isolates in each survey as previously described (Barry et al., 2007). PTS, analogous to the Sørensen Index,
243 is a similarity statistic to evaluate the proportion of DBLα types shared between two isolate repertoires
244 (i.e., DBLα repertoire overlap) and ranges from 0 (i.e., no DBLα repertoire overlap) to 1 (i.e., identical DBLα
245 isolate repertoires), where < 0.50 = unrelated, 0.5 = recent recombinants/siblings, > 0.5 = related, and 1
246 = clones. PTS is a measure of identity-by-state (IBS) used to assess repertoire similarity between isolates
247 and is not used to infer inheritance from a recent common ancestor (i.e., identity-by-decent (IBD)) (Speed
248 and Balding, 2015).

249

250 DBLα isolate repertoire size

251 For this study we have exploited the unique population structure of non-overlapping DBLα isolate
252 repertoires to estimate isolate MOI_{var}. To calculate MOI_{var}, the non-upsA DBLα types were chosen since
253 not only are they more diverse and less conserved between isolate repertoires (i.e., low median PTS_{non-}
254 $_{upsA} \leq 0.020$) compared to the upsA DBLα types, but they have also been shown to have a more specific 1-
255 to-1 relationship with a single *var* gene than upsA (Tan et al., 2023). The low to non-existent overlap of
256 repertoires enables an estimation of MOI that relies on the number of non-upsA DBLα types sequenced
257 from an individual's isolate (Ruybal-Pesántez et al., 2022; Tiedje et al., 2022). A constant repertoire size
258 or number of DBLα types in a parasite genome can be used to convert the number of types sequenced in
259 an isolate to estimate MOI (Ruybal-Pesántez et al., 2022; Tiedje et al., 2022). This approach, however,
260 neglects the measurement error in this size introduced by targeted PCR and amplicon sequencing of *var*
261 genes in an isolate.

262

263 Bayesian estimation of MOI_{var} and associated census population size

264 Here, we extend the method to a Bayesian formulation and estimate the posterior distribution for each
265 sampled individual for the probability of different MOI values. From individual posterior distributions, we
266 can then obtain the estimated MOI frequency distribution for the population as a whole. The two pieces
267 of information required for our approach are the measurement error and the prior distribution of MOI.
268 The measurement error is simply the repertoire size distribution, that is, the distribution of the number
269 of non-upsA DBLα types sequenced given MOI = 1, which is empirically available (Figure supplement 3
270 (Labbé et al., 2023)). We refer to it as $P(s \mid MOI = 1)$ where *s* here denotes repertoire size. More generally

when MOI ≥ 1, *s* denotes the number of non-upsA DBLα types sequenced or isolate size. The prior distribution of MOI refers to the belief we have for what the actual MOI distribution might look like at the population level before empirical evidence is taken into consideration. For example, the prior distribution of MOI is likely to centre around a higher value in high-transmission endemic areas than in low-transmission ones.

We can obtain P(s | MOI = m) from the serial convolution of the repertoire size distribution P(s | MOI = 1) and P(s | MOI = m - 1). Starting with the repertoire size distribution given a single infection, we can derive P(s | MOI = m) for *m* equal to 2,3,..., up until a maximum value of 20 (empirically determined), as follows:

$$P(s \mid MOI = m) = \sum_{x=L}^{U} P(x \mid MOI = 1) \times P(s - x \mid MOI = m - 1) \qquad 1.1$$

where *L* and *U* are the lower and upper limit for the repertoire size, 10 and 45 respectively from the empirical repertoire size distribution (Figure supplement 3 (Labbé et al., 2023)).

For simplicity, we begin with a uniform prior. We use Bayes' rule to derive a posterior distribution of MOI given a certain number of non-upsA DBLα types sequenced from an individual:

$$P(MOI = j \mid s) = \frac{P(s \mid MOI = j) \times P(MOI = j)}{\sum_{i=1}^{k} P(s \mid MOI = i) \times P(MOI = i)} \qquad 1.2$$

where *k* is the maximum value of MOI, here 20, as empirically determined.

290 To obtain the MOI distribution at the population level, we could either simply pool the maximum *a*
291 *posteriori* MOI estimate for each sampled individual, or use a technique called mixture distribution. For
292 the latter, we weighed each posterior MOI distribution for each sampled individual equally and sum over
293 all posterior distributions at the individual level to derive the MOI distribution at the population level:

294 $$f(MOI = m) = \sum_{i=1}^{n} \frac{1}{n} P(MOI = m \mid s_i) \qquad 1.3$$

295 where *n* is the number of sampled individuals. These two approaches gave similar results for our empirical
296 survey data as determined by the Kolmogorov-Smirnov Test. The obtained distance statistic is close to 0
297 and the corresponding *p-value* is non-significant across all surveys, indicating that the two estimates were
298 drawn from the same distribution (Table supplements 3 and 4). Additionally, the difference between the
299 mean MOI values at the population level obtained from the two approaches is small (Table supplements
300 3 and 4). Given this similarity, we present the results based on pooling the maximum *a posteriori* MOI
301 estimates for each sampled individual in the main text and include the results based on mixture
302 distribution in the Supplemental Information. Note that we focused on individuals who had confirmed
303 microscopic asymptomatic P. *falciparum* infections for our MOI estimation.

304

305 To examine alternative priors, we considered empirical MOI distributions described in the literature
306 including the Poisson, hyper-Poisson, and negative binomial distributions (Dietz, 1988; Henry, 2020). The
307 hyper-Poisson and negative binomial distributions can capture the overdispersion seen in the empirical
308 distribution of MOI for certain areas and caused by factors such as heterogeneous biting. We therefore
309 focused on a negative binomial distribution and investigated changing its parameters to generate priors
310 with different means spanning a wide range of MOI values (mean MOI within [~1.5, ~6.7]), including those

311 seen in high-transmission endemic areas. A uniform prior and a zero-truncated negative binomial
312 distribution with parameters within the range typical of high-transmission endemic regions (higher mean
313 MOI, for example, ~4.3 versus ~6.7, with tails for higher MOI values in the range of 10-20) produce similar
314 MOI estimates at the population level (Table supplements 5 and 6). However, when setting the parameter
315 range of the zero-truncated negative binomial to be representative of those in low-transmission endemic
316 regions where the empirical MOI distribution centres around monoclonal infections with the majority of
317 MOIs = 1 or 2 (mean MOI ≈ 1.5, no tail at higher MOI values), the final population-level MOI distribution
318 does deviate more from that based on the aforementioned prior and parameter choices (Table
319 supplement 6). The final individual- and population-level MOI estimates are not sensitive to the specifics
320 of the prior MOI distribution as long as this distribution captures the tail for higher MOI values with above-
321 zero probability. The obtained Kolmogorov-Smirnov Test distance statistics and their corresponding *p-*
322 *values*, the Pearson correlation tests and their corresponding *p-values*, as well as the difference in mean
323 MOI values, for the comparison of the MOI estimates obtained with the different priors are included in
324 Table supplements 5 and 6. Given these comparisons, we provide in our analyses the estimated
325 population MOI distribution using a uniform prior.
326

327 Statistical analysis

328 We used the R v4.3.1 for all data analyses with the collection of R packages in *tidyverse* being used for
329 data curation along with *base*, *stats, gtsummary,* and *epiR* for the statistical analyses (R Core Team, 2018;
330 Sjoberg, D et al., 2021; Stevenson, 2020; Wickham et al., 2019). Continuous variables are presented as
331 medians with interquartile ranges (IQRs) and discrete variables are presented using the observed or
332 calculated values with the 95% confidence interval (95% CIs) or ± 2 standard deviations (± 2SD). Kaplan-

333 Meier survival curves were generated for the time (i.e., number of surveys) to first event (i.e., when the
334 DBLα type was no longer observed/detected) comparing the upsA and non-upsA DBLα types; *p-values*
335 were determined using the log-rank test using the R packages *survival* and *survminer* (Kassambara et al.,
336 2021; Therneau, 2023). The time interval to first event considered for all survival curves was the number
337 of surveys or year (i.e., 2012, 2014, 2015, and 2017) that each DBLα type was observed and only includes
338 those upsA (N = 2,218) and non-upsA (N = 33,159) DBLα types that were seen at baseline in 2012 (i.e.,
339 those DBLα types observed prior to the IRS intervention) (Table supplement 2).
340

341 # RESULTS

342 Between 2013 and 2015, three-rounds of IRS with non-pyrethroid insecticides were implemented across
343 all of Bongo District. Coincident with the > 90% decrease in transmission following IRS (Tiedje et al., 2022),
344 the prevalence of microscopic *P. falciparum* infections compared to the 2012 baseline survey (pre-IRS)
345 declined by 45.2% and 35.7% following the second (2012 to 2014) and the third (2012 to 2015) round of
346 IRS, respectively (Figure 1B, Table supplement 1). These declines in parasite prevalence were observed
347 across all ages, with the greatest impacts being observed on the younger children (1-5 years) who were
348 ~3 times less likely to have an infection in 2015 (post-IRS) compared to 2012 (pre-IRS) (Figure 1C, Table
349 supplement 1). These reductions were however short-lived and in 2017, 32-months after the
350 discontinuation of IRS, but during SMC, overall *P. falciparum* prevalence rebounded to 41.2%, or near pre-
351 IRS levels (Figure 1B, Table supplement 1). Importantly, this increase in the prevalence of infection in 2017
352 was only observed among the older age groups (i.e., ≥ 6 years) (Figure 1C, Table supplement 1). This
353 difference by age group in 2017 can be attributed to SMC, which only targets children between 3-59
354 months (i.e., < 5 years). A notable increase in parasite prevalence for adolescents (11-20 years) and adults
355 (> 20 years) was found in 2017 relative to the 2012 (pre-IRS) (Figure 1C, Table supplement 1).

356

357 Next, we wanted to explore changes in population size measured by MOI_{var}. As this metric is based on
358 non-overlap of *var* repertoire diversity of individual isolates, specifically non-upsA DBLα types, we
359 investigated whether DBLα isolate repertoire similarity (or overlap), as measured by PTS, increased
360 following the sequential interventions (i.e., IRS and SMC). Figure 2 shows that median PTS values for both
361 upsA and non-upsA DBLα types remained low in all surveys, although the PTS distributions for both groups
362 changed significantly at each of the study time points relative to the 2012 baseline survey (pre-IRS) (*p*-

363 *values* < 0.001, Kruskal-Wallis test) (Figure 2). Somewhat unexpectedly, the change was in the direction
364 of reduced similarity (i.e., less overlap) with lower median PTS scores and a larger number of isolates
365 sharing no DBLα types (i.e., PTS = 0) in 2014, 2015, and 2017 compared to 2012. Relevant to measurement
366 of MOI_{var}, the median PTS scores for non-upsA DBLα types were lower following the IRS intervention
367 ($PTS_{non\text{-}upsA}$: 2014 = 0.013 and 2015 = 0.013 vs. $PTS_{non\text{-}upsA}$: 2012 = 0.020). In 2017, the non-upsA PTS
368 distributions shifted back towards higher median PTS scores ($PTS_{non\text{-}upsA}$ = 0.016) and fewer isolates shared
369 no DBLα types relative to 2014 and 2015 (Figure 2). To verify this pattern was not influenced by multiclonal
370 infections (MOI_{var} > 1) we also examined isolates with monoclonal infections (MOI_{var} = 1) and found that
371 this non-overlapping structure persisted regardless of infection complexity, particularly for the non-upsA
372 DBLα types (Figure supplement 4). These PTS data make clear that we were dealing with a large, highly
373 diverse parasite population where 99.9% of the isolate comparisons in all surveys had $PTS_{non\text{-}upsA}$ scores ≤
374 0.1 (i.e., shared ≤ 10% of their non-upsA DBLα types), indicating that DBLα isolate repertoires were highly
375 unrelated (Figure 2). In fact, throughout the IRS, SMC, and subsequent rebound, very few DBLα isolate
376 repertoires were observed to be related, with < 0.003% isolate comparisons in each survey having a $PTS_{non\text{-}upsA}$ ≥ 0.5 (i.e., siblings or recent recombinants) (Figure 2).
377
378
379 The raw data of non-upsA DBLα isolate repertoire sizes (Figure supplement 5) were used to estimate
380 MOI_{var} as adjusted using the Bayesian approach based on pooling the maximum *a posteriori* MOI estimates
381 (Figure 3) and the mixture distribution (Figure supplement 6). We observed that at baseline in 2012, the
382 majority (89.2%) of the population across all ages carried multiclonal infections (median MOI_{var} = 4 [IQR:
383 2 – 6]) (Figure 3A). Following the IRS intervention, the estimated MOI_{var} distributions became more
384 positively skewed, indicating that a lower proportion of participants harboured multiclonal infections with

385 a lower median MOI_{var} in 2014 (64.5%; median MOI_{var} = 2 [IQR: 1 – 3]) and 2015 (71.4%; median MOI_{var} =
386 2 [IQR: 1 – 3]) compared to 2012 (Figure 3A). These reductions in median MOI_{var} and the proportion of
387 multiclonal infections, which were observed across all age groups (Figure 3B), are consistent with the >
388 90% decrease in transmission intensity following the IRS in turn reducing exposure to new parasite
389 genomes. However, in 2017, both median MOI_{var} (3 [IQR: 2 – 4]) and the proportion of multiclonal
390 infections (78.9%) rebounded in all age groups, even among the younger children (1-5 years)
391 predominantly targeted by SMC (Figure 3). While the prevalence of infection in 2017 remained low for
392 the younger children (1-5 years), those infected still carried multiclonal infections (84.1% of those
393 infected) (Figure 3B). Although the MOI_{var} distributions across all age groups started to rebound in 2017
394 (i.e., less positively skewed compared to 2014 and 2015) they had not fully recovered to the 2012 baseline
395 patterns (Figure 3). This was most apparent among the younger children (1-5 years), as a larger proportion
396 of isolates in 2017, compared to 2012, had MOI_{var} values equal to one or two, while a smaller proportion
397 had MOI_{var} values ≥ 5 (Figure 3B).
398
399 Census population size, measured as the number of *P. falciparum var* repertoires circulating in the
400 population during each survey, was estimated by summation of isolate MOI_{var} (Figure 4, Table supplement
401 7). In 2014 during IRS, this number decreased by 72.5% relative to the 2012 baseline survey (pre-IRS)
402 (Figure 4CE), whereas prevalence decreased by 54.5% (Figure 4CE). Although census population size
403 increased slightly in 2015 relative to 2014 (Figure 4A), there were still 64.6% fewer *var* repertoires in the
404 population compared to 2012 (Figure 4CE) in comparison to a 42.6% decrease in prevalence (Figure 4CE).
405 Importantly this loss of *var* repertoires in 2014 and 2015 following the IRS intervention was seen for all
406 age groups (Figure 4B), with the greatest overall reductions (> 85%) being observed for the younger

407 children (1-5 years) (Figure 4DF). However, in 2017, the number of diverse *var* repertoires in the
408 population rebounded, more than doubling between 2015 and 2017 (Figure 4CE). This increase in the
409 number of *var* repertoires was seen for all age groups in 2017, except for the younger children (1-5 years)
410 where those up to 59 months were targeted by SMC (Figure 4BD). In fact, the greatest overall increase
411 was observed for the adolescents (11-20 years), where the number of *var* repertoires in 2017 was ~1.4
412 times higher compared to 2012 (Figure 4F). Similar trends in the number *var* repertoires were also
413 observed for the older children (6-10 years) and adults (>20 years) in 2017, although the rebound was not
414 as striking as that detected for the adolescents.
415
416 As census population size changed considerably during the sequential IRS and SMC interventions, we
417 investigated how the removal or loss of *P. falciparum var* repertoires and subsequent rebound in 2017
418 altered DBLα type richness, measured as the number of unique upsA and non-upsA DBLα types in the
419 parasite population in each survey. Richness at baseline in 2012 (pre-IRS) was high with a large number of
420 unique DBLα types (upsA = 2,218; non-upsA = 33,159) (Figure 5, Table supplement 2) and limited overlap
421 of *var* repertoires (i.e., median $PTS_{non\text{-}upsA} \leq 0.020$) seen in a relatively small study population of 685
422 microscopically positive individuals (Figure 2). In 2014, as *P. falciparum* prevalence and census population
423 size declined (Figure 4) so too did the number of DBLα types, resulting in a 32.2% and 55.3% reduction in
424 richness for the upsA and non-upsA DBLα types, respectively, compared to 2012 (Figure 5, Table
425 supplement 2). Again in 2015, as *P. falciparum* prevalence and population size remained low (Figure 4)
426 DBLα type richness was still less than that observed in 2012 (24.6% and 46.0% reduction for upsA and
427 non-upsA DBLα types, respectively) (Figure 5, Table supplement 2). Finally, in 2017, we found that upsA

428 and non-upsA DBLα type richness rebounded relative to 2014 and 2015, coincident with the increase in
429 *P. falciparum* prevalence and census population size (Figure 4, Figure 5).
430
431 Given this reduction in DBLα type richness following the IRS intervention and subsequent rebound in 2017,
432 we wanted to explore whether the loss of richness was influenced by the frequency of individual DBLα
433 types in the parasite population within and among surveys. To answer this, we defined the relative
434 frequency of individual DBLα types in all isolates in each survey (Figure 6, Figure supplement 7). We
435 discovered that individual upsA and non-upsA DBLα types were not all at equal frequencies within a survey
436 and among surveys. They could be classified as frequent (i.e., observed in 11-20 or > 20 isolates), less
437 frequent (i.e., observed in 2-10 isolates) or only seen once, at baseline in 2012 (Figure 6AB). In 2014 and
438 2015, following IRS, there was a significant increase in the proportion of upsA and non-upsA DBLα types
439 in the lower frequency categories (*p-value* < 0.001, Mann-Whitney U test), with all DBLα types becoming
440 rarer in the population (Figure 6CD). This change can be attributed to the removal of *P. falciparum var*
441 repertoires (Figure 4) with associated loss of upsA and non-upsA DBLα type richness (Figure 5), which
442 disproportionally affected those DBLα types seen once. This shift to all DBLα types becoming rarer
443 following IRS changed in 2017, where the proportion of DBLα types in the more frequent categories (i.e.,
444 2-10, 11-20, or > 20 isolates) significantly increased while the proportion seen once decreased (*p-values*
445 < 0.001, Mann-Whitney U tests) (Figure 6CD). Data in Figure 6 (A-D) pointed to a differential effect of the
446 IRS intervention and subsequent rebound on the less frequent upsA and non-upsA DBLα types vs. those
447 that were classified as frequent, where those DBLα types that were most frequent persisted longitudinally.
448

449 To explore this observation further we restricted the longitudinal analysis to those DBLα types from the
450 baseline survey in 2012 (pre-IRS). We compared the probability of survival for the DBLα types identified
451 at baseline in 2012 and found that the upsA DBLα types persisted significantly longer in the population
452 relative to the non-upsA DBLα types ($p < 0.001$, log-rank test), despite the IRS intervention. The simple
453 explanation being that although the upsA DBLα types had lower richness (Figure 5), a larger proportion
454 was classified as frequent indicating that multiple copies existed in the population compared to the non-
455 upsA DBLα types (Figure 6CD). Furthermore, when we examined survival using the frequency categories,
456 the upsA and non-upsA DBLα types that were observed at multiple study time points (i.e., 2012, 2014,
457 2015, and 2017), albeit in different isolate repertoires, were those that were most frequent (i.e., observed
458 in 11-20 and >20 isolates) in the population at baseline in 2012 (Figure 6EF). As expected, the DBLα types
459 that were only observed once in 2012, were significantly less likely to be seen longitudinally (*p-value* <
460 0.001, log-rank test) (Figure 6EF). These differential changes in DBLα type richness with respect to rare vs.
461 frequent DBLα types are a consequence of changes in census population size with interventions (Figure
462 4) where each isolate repertoire is composed of many rare DBLα types as defined by PTS (Figure 2).
463

464 # DISCUSSION

465 *P. falciparum* populations in high-transmission endemic areas in SSA, are characterised by extensive
466 diversity, high rates of recombination, as well as frequent multiclonal infections. Here we defined census
467 population size of *P. falciparum* to understand total parasite diversity in a human population and explore
468 the age-specific efficacy of malaria interventions to reduce this metric in such areas as typified by Bongo,
469 Ghana. Census population size proved more informative than parasite prevalence alone because it
470 captures "within" host parasite population size as MOI, rather than using the infected host per se as a unit
471 of population size. Whilst the concept of census population size is agnostic of how you measure MOI, the
472 extensive DBLα isolate repertoire diversity presented makes a strong case for fingerprinting parasite
473 isolates by *var*coding in high transmission. This is opposed to looking at allelic diversity of a single copy
474 antigen gene, such as *csp*.

475

476 Census population size is a total enumeration or count of infections in a given population sample and over
477 a given time period in an ecological sense, distinct from the formal effective population size (N_e) used in
478 population genetics (Charlesworth, 2009). Given the low overlap between *var* repertoires of parasites
479 observed in monoclonal infections (MOI = 1), the census population size calculated in Bongo, Ghana
480 translates to a diversity of strains or repertoires. The distinction of census population size in terms of
481 infection counts and effective population size from population genetics has been made before for
482 pathogens, including the seasonal influenza virus and the measles virus (Bedford et al., 2011); it is also a
483 distinction made in the ecological literature for non-pathogen populations (Palstra and Fraser, 2012). The
484 census population size of a given population sample depends of course on sample size and was used here
485 for comparisons across time of samples of the same depth (i.e., ~2,000 individuals). There is, however, a

486 simple map between census population size and mean MOI, as one can simply divide or multiply by the
487 sample size, respectively, to convert between the two quantities (Table supplement 7). Therefore, one
488 can extrapolate from the census population size of a given population sample to that of the whole
489 population of local hosts to compare across studies that differ in sampling depth. What is needed for this
490 extrapolation is a stable mean MOI relative to the sample size or sampling depth, which is indeed the case
491 in this study (Figure supplement 8) and can be easily checked in other studies. Given the typical duration
492 of infection, we expect our population size to be representative of a per-generation measure.
493
494 By *var*coding we identified a very large census parasite population size in a relatively small human
495 population of ~2,000 individuals at baseline and captured age-specific changes in this metric in response
496 to sequential malaria control interventions. IRS reduced the MOI_{var} parasite population size substantially
497 with the greatest reductions (85%) seen in the younger children (1-5 years). More than two years after
498 cessation of IRS, rebound in 2017 was rapid in all age groups, except for the younger children (1-5 years)
499 where those up to 59 months were targeted by SMC. Population sizes in adolescents (11-20 years) and
500 adults (> 20 years) showed they carried more infections in 2017 than at baseline in 2012. This is indicative
501 of a loss of immunity during IRS which may relate to the observed loss of *var* richness, especially the many
502 rare types. This warrants further investigation of changes in variant-specific immunity. During and
503 following the IRS and SMC interventions, *var* diversity remained high and *var* repertoire overlap remained
504 low, reflecting characteristic properties of high transmission and demonstrating the overall resilience of
505 the species to survive significant short-term perturbations. Combining interventions and targeting older
506 age groups or the whole community with chemoprevention would no doubt have a much greater impact
507 on reducing the diversity of the reservoir of infection.

508

509 What was striking about the Bongo study was the speed with which rebound in MOI_{var} per person and
510 census population size occurred, once the short-term IRS was discontinued. We looked for a potential
511 explanation in our genetic data. PTS and population frequency data showed that many of the DBLα types
512 occurred in multiple repertoires or genomes. This enabled the survival of these more frequent DBLα types
513 through the interventions, facilitating rebound by maintenance of this diversity. The other notable
514 population genetic result of our study was the failure to increase similarity (or relatedness by state) of *var*
515 repertoires by reducing transmission by > 90% via IRS. From a baseline of a very large population size with
516 very low overlap in repertoires you need outcrossing to create relatedness. However, this was less likely
517 to happen due to reduced transmission as a result of IRS. Rebound, with associated increases in
518 transmission, led to a small increase in *var* repertoire similarity. This is the opposite to what has been
519 observed in areas of lower transmission under intense malaria control where the intensity of interventions
520 led to increased genome similarity, as assessed by IBD, from a starting point of much lower genome
521 diversity and greater relatedness (Daniels et al., 2015).

522

523 Our molecular approach to measure population size has been to sum MOI_{var} in individual hosts with
524 microscopically-detectable infections. Like any diagnostic method there are limits to sensitivity and
525 specificity, which can be more or less tolerated dependent upon the purpose of the study. Here we have
526 looked at relative changes in population size with sequential interventions using an interrupted time series
527 study design and observed changes by measuring MOI_{var}. We have accounted for missing DBLα type data
528 where complete *var* repertoires may not have been sequenced using a Bayesian method based on
529 empirical knowledge of the measurement error. This approach has conceptual relation to the Bayesian

530 approach by Johnson and Larremore (Johnson and Larremore, 2022) to estimate complete repertoire size
531 of, and overlap between, monoclonal infections from incomplete sampling of DBLα types. Our
532 measurement of population size based on MOI_{var} will be subject to other sampling errors which may in
533 the end be more significant (discussed in detail in Labbé et al. (Labbé et al., 2023)). For example, low
534 parasitemia typical of asymptomatic infections, small blood volumes, clinical status, and/or within host
535 dynamics, including synchronicity, will all create sampling problems, but these are common to all
536 measures of MOI.
537
538 Previously, we have drawn attention to the potential underestimation of the number of DBLα types of
539 related parasites generated by a cross, when using *var* genotyping (Labbé et al., 2023). Such related
540 parasites must be created frequently in high transmission due to extensive outcrossing (Babiker et al.,
541 1994; Paul et al., 1995). Single clone genomics experiments using biallelic SNPs from whole genome
542 sequencing data have also detected related parasites using IBD in clinical infections from humans in a
543 high-transmission area of Malawi (Nkhoma et al., 2020). Here we have analysed low- to moderate-density,
544 chronic, asymptomatic infections (see Table supplement 1) under strong immune selection in semi-
545 immune hosts whom we have shown select against parasites with high PTS scores consistent with
546 relatedness by decent (Day et al., 2017; He et al., 2018; Ruybal-Pesántez et al., 2022). When considering
547 the importance of possible exclusion of parasites related by descent, the only sure way to detect such
548 parasites in high transmission is by single cell genomics, a methodology of limited application to malaria
549 surveillance due to practicality and cost of scale up. Again, the error from failure to sample infections
550 related by descent must be weighed up against the issues of under-sampling as described above.
551

552 The Bayesian approach of the *var*coding method relies on the low or limiting similarity of *var* repertoires
553 infecting individual human hosts. As such, it would appear to break down as the *var* repertoire overlap
554 moves away from extremely low, and therefore, for locations with lower transmission intensity.
555 Interestingly, this is not the case in the numerical simulations of Labbé et al. (Labbé et al., 2023), for a
556 gradient of three transmission intensities, from high to low, with the original *var*coding method
557 performing well across the gradient. This robustness of the method may arise from a nonlinear and fast
558 transition from low to high overlap that is accompanied by MOI changing rapidly from primarily
559 multiclonal (MOI > 1) to monoclonal (MOI = 1) infections. This matter needs to be investigated further in
560 the future, including ways to extend the Bayesian approach to explicitly include the distribution of *var*
561 repertoire overlap.
562
563 In summary, our findings provide parasite population insights into why rebound is the inevitable
564 consequence of such short-term IRS interventions unless you simultaneously target the highly diverse,
565 long-lived parasite population in humans, not just children < 5 years by SMC. Of potential translational
566 significance for malaria molecular surveillance, we identify new metrics, especially MOI_{var} and census
567 population size as well as *var* frequency category, informative to monitor and evaluate interventions in
568 high-transmission areas with multiclonal infections and high rates of outcrossing. Such metrics could be
569 used longitudinally to detect incremental gains of transmission-reducing interventions, including IRS,
570 LLINs, and vaccines to perturb the high-transmission characteristics of the parasite population in humans
571 in high-burden countries in SSA.
572

ACKNOWLEDGMENTS. This research was supported by Fogarty International Center at the National Institutes of Health through the joint NIH-NSF-NIFA Ecology and Evolution of Infectious Diseases award R01-TW009670 to K.A.K, M.P, and K.P.D; and the National Institute of Allergy and Infectious Diseases, National Institutes of Health through the joint NIH-NSF-NIFA Ecology and Evolution of Infectious Diseases award R01-AI149779 to A.R.O, K.A.K, M.P, and K.P.D. We wish to thank the participants, communities, and the Ghana Health Service in Bongo District, Ghana for their willingness to participate in this study. We would like to thank the field teams in Bongo for their technical assistance in the field, as well as the laboratory personnel at the Navrongo Health Research Centre for their expertise and for undertaking the sample collections and parasitological assessments.

REFERENCES

Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP, Estrada-Franco JG, Richardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP. 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite Plasmodium falciparum. *Mol Biol Evol* **17**:1467–1482. doi:10.1093/oxfordjournals.molbev.a026247

Babiker HA, Ranford-Cartwright LC, Currie D, Charlwood JD, Billingsley P, Teuscher T, Walliker D. 1994. Random mating in a natural population of the malaria parasite Plasmodium falciparum. *Parasitology* **109**:413–421. doi:10.1017/S0031182000080665

Barry AE, Leliwa-Sytek A, Tavul L, Imrie H, Migot-Nabias F, Brown SM, McVean GA V, Day KP. 2007. Population genomics of the immune evasion (var) genes of Plasmodium falciparum. *PLoS Pathog* **3**:e34. doi:10.1371/journal.ppat.0030034

Bedford T, Cobey S, Pascual M. 2011. Strength and tempo of selection revealed in viral gene genealogies. *BMC Evol Biol* **11**. doi:10.1186/1471-2148-11-220

Bruce MC, Galinski MR, Barnwell JW, Donnelly CA, Walmsley M, Alpers MP, Walliker D, Day KP. 2000. Genetic diversity and dynamics of plasmodium falciparum and P. vivax populations in multiply infected children with asymptomatic malaria infections in Papua New Guinea. *Parasitology* **121 (Pt 3**:257–272. doi:10.1017/S0031182099006356

Charlesworth B. 2009. Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet*. doi:10.1038/nrg2526

Chen DS, Barry AE, Leliwa-Sytek A, Smith T-AA, Peterson I, Brown SM, Migot-Nabias F, Deloron P, Kortok MM, Marsh K, Daily JP, Ndiaye D, Sarr O, Mboup S, Day KP. 2011. A molecular epidemiological study of var gene diversity to characterize the reservoir of Plasmodium falciparum in humans in Africa. *PLoS One* **6**:e16629. doi:10.1371/journal.pone.0016629

Daniels RF, Schaffner SF, Wenger E a., Proctor JL, Chang H-H, Wong W, Baro N, Ndiaye D, Fall FB, Ndiop M, Ba M, Milner D a., Taylor TE, Neafsey DE, Volkman SK, Eckhoff P a., Hartl DL, Wirth DF. 2015. Modeling malaria genomics reveals transmission decline and rebound in Senegal. *PNAS* **112**:7067–7072. doi:10.1073/pnas.1505691112

Day KP, Artzy-Randrup Y, Tiedje KE, Rougeron V, Chen DS, Rask TS, Rorick MM, Migot-Nabias F, Deloron P, Luty AJF, Pascual M. 2017. Evidence of Strain Structure in Plasmodium falciparum Var Gene Repertoires in Children from Gabon, West Africa. *PNAS* **114**:E4103–E4111. doi:10.1073/pnas.1613018114

Dietz K. 1988. Mathematical models for transmission and control of malariaMalaria: Principles and Practice of Malariology.

Falk N, Kaestli M, Qi W, Ott M, Baea K, Cortés A, Beck H-P. 2009. Analysis of Plasmodium falciparum var genes expressed in children from Papua New Guinea. *Journal of infectious diseases* **200**:347–356. doi:10.1086/600071

Falk N, Maire N, Sama W, Owusu-Agyei S, Smith T, Beck HP, Felger I. 2006. Comparison of PCR-RFLP and genescan-based genotyping for analyzing infection dynamics of Plasmodium falciparum. *American Journal of Tropical Medicine and Hygiene* **74**:944–950.

624 Farnert A, Snounou G, Rooth I, Bjorkman A. 1997. Daily dynamics of Plasmodium falciparum
625 subpopulations in asymptomatic children in a holoendemic area. *American Journal of Tropical*
626 *Medicine and Hygiene* **56**:538–547. doi:10.4269/ajtmh.1997.56.538
627 Futse JE, Brayton KA, Dark MJ, Knowles DP, Palmer GH. 2008. Superinfection as a driver of genomic
628 diversification in antigenically variant pathogens. *Proc Natl Acad Sci U S A* **105**:2123–2127.
629 doi:10.1073/pnas.0710333105
630 Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S,
631 Paulsen IT, James K, Eisen J a, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M-S, Nene V, Shallom
632 SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin
633 DMA, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph S a, McFadden GI, Cummings LM, Subramanian
634 GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B. 2002.
635 Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**:498–511.
636 doi:10.1038/nature01097
637 Ghansah A, Tiedje KE, Argyropoulos DC, Onwona CO, Deed SL, Labbé F, Oduro AR, Koram KA, Pascual M,
638 Day KP. 2023. Comparison of molecular surveillance methods to assess changes in the population
639 genetics of Plasmodium falciparum in high transmission. *Frontiers in Parasitology* **2**:1067966.
640 Gogue C, Wagman J, Tynuv K, Saibu A, Yihdego Y, Malm K, Mohamed W, Akplu W, Tagoe T, Ofosu A,
641 Williams I, Asiedu S, Richardson J, Fornadel C, Slutsker L, Robertson M. 2020. An observational
642 analysis of the impact of indoor residual spraying in Northern, Upper East, and Upper West Regions
643 of Ghana: 2014 through 2017. *Malar J* **19**:1–13. doi:10.1186/s12936-020-03318-1
644 He Q, Pilosof S, Tiedje KE, Ruybal-Pesántez S, Artzy-Randrup Y, Baskerville EB, Day KP, Pascual M. 2018.
645 Networks of genetic similarity reveal non-neutral processes shape strain structure in Plasmodium
646 falciparum. *Nat Commun* **9**:1817. doi:10.1038/s41467-018-04219-3
647 Henry JM. 2020. A hybrid model for the effects of treatment and demography on malaria superinfection.
648 *J Theor Biol* **491**:110194.
649 Jensen ATR, Magistrado P, Sharp S, Joergensen L, Lavstsen T, Chiucchiuini A, Salanti A, Vestergaard LS,
650 Lusingu JP, Hermsen R, Sauerwein R, Christensen J, Nielsen MA, Hviid L, Sutherland C, Staalsoe T,
651 Theander TG. 2004. Plasmodium falciparum associated with severe childhood malaria preferentially
652 expresses PfEMP1 encoded by group A var genes. *J Exp Med* **199**:1179–1190.
653 doi:10.1084/jem.20040274
654 Johnson EK, Larremore DB. 2022. Bayesian estimation of community size and overlap from random
655 subsamples. *PLoS Comput Biol* **18**:1–16. doi:10.1371/journal.pcbi.1010451
656 Kaestli M, Cockburn I a, Cortés A, Baea K, Rowe JA, Beck H-P. 2006. Virulence of malaria is associated with
657 differential expression of Plasmodium falciparum var gene subgroups in a case-control study. *J Infect*
658 *Dis* **193**:1567–1574. doi:JID35650 [pii]\n10.1086/503776
659 Kalmbach Y, Rottmann M, Kombila M, Kremsner PG, Beck H-P, Kun JFJ. 2010. Differential var gene
660 expression in children with malaria and antidromic effects on host gene expression. *Journal of*
661 *Infectious Diseases* **202**:313–317. doi:10.1086/653586
662 Kassambara A, Kosinski M, Biecek P, Scheipl F. 2021. survminer: Drawing Survival Curves using ggplot2.
663 Kraemer SM, Kyes SA, Aggarwal G, Springer AL, Nelson SO, Christodoulou Z, Smith LM, Wang W, Levin E,
664 Newbold CI, Myler PJ, Smith JD. 2007. Patterns of gene recombination shape var gene repertoires in
665 Plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics* **8**:45.
666 doi:10.1186/1471-2164-8-45

667 Kraemer SM, Smith JD. 2006. A family affair: var genes, PfEMP1 binding, and malaria disease. *Curr Opin*
668 *Microbiol*. doi:10.1016/j.mib.2006.06.006
669 Kyriacou HM, Stone GN, Challis RJ, Raza A, Lyke KE, Thera MA, Koné AK, Doumbo OK, Plowe C V., Rowe JA.
670 2006. Differential var gene transcription in Plasmodium falciparum isolates from patients with
671 cerebral malaria compared to hyperparasitaemia. *Mol Biochem Parasitol* **150**:211–218.
672 doi:10.1016/j.molbiopara.2006.08.005
673 Labbé F, He Q, Zhan Q, Tiedje KE, Argyropoulos DC, Tan MH, Ghansah A, Day KP, Pascual M. 2023. Neutral
674 vs . non-neutral genetic footprints of Plasmodium falciparum multiclonal infections. *PLoS Comput*
675 *Biol* **19**:e1010816. doi:doi.org/10.1101/2022.06.27.497801
676 LaVerriere E, Schwabl P, Carrasquilla M, Taylor AR, Johnson ZM, Shieh M, Panchal R, Straub TJ, Kuzma R,
677 Watson S, Buckee CO, Andrade CM, Portugal S, Crompton PD, Traore B, Rayner JC, Corredor V, James
678 K, Cox H, Early AM, MacInnis BL, Neafsey DE. 2022. Design and implementation of multiplexed
679 amplicon sequencing panels to serve genomic epidemiology of infectious disease: A malaria case
680 study. *Mol Ecol Resour* 2285–2303. doi:10.1111/1755-0998.13622
681 Lavstsen T, Salanti A, Jensen ATR, Arnot DE, Theander TG. 2003. Sub-grouping of Plasmodium falciparum
682 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar J* **2**:27.
683 doi:10.1186/1475-2875-2-27
684 Lerch A, Koepfli C, Hofmann NE, Kattenberg JH, Rosanas-Urgell A, Betuela I, Mueller I, Felger I. 2019.
685 Longitudinal tracking and quantification of individual Plasmodium falciparum clones in complex
686 infections. *Sci Rep* **9**:1–8. doi:10.1038/s41598-019-39656-7
687 Lerch A, Koepfli C, Hofmann NE, Messerli C, Wilcox S, Kattenberg JH, Betuela I, O'Connor L, Mueller I,
688 Felger I. 2017. Development of amplicon deep sequencing markers and data analysis pipeline for
689 genotyping multi-clonal malaria infections. *BMC Genomics* **18**:1–13. doi:10.1186/s12864-017-4260-
690 y
691 Markwalter CF, Menya D, Wesolowski A, Esimit D, Lokoel G, Kipkoech J, Freedman E, Sumner KM, Abel L,
692 Ambani G, Meredith HR, Taylor SM, Obala AA, O'Meara WP. 2022. Plasmodium falciparum
693 importation does not sustain malaria transmission in a semi-arid region of Kenya. *PLOS Global Public*
694 *Health* **2**:e0000807. doi:10.1371/journal.pgph.0000807
695 Nelson CS, Sumner KM, Freedman E, Saelens JW, Obala AA, Mangeni JN, Taylor SM, O'Meara WP. 2019.
696 High-resolution micro-epidemiology of parasite spatial and temporal dynamics in a high malaria
697 transmission setting in Kenya. *Nat Commun* **10**. doi:10.1038/s41467-019-13578-4
698 Nkhoma SC, Trevino SG, Gorena KM, Nair S, Khoswe S, Jett C, Garcia R, Daniel B, Dia A, Terlouw DJ, Ward
699 SA, Anderson TJC, Cheeseman IH. 2020. Co-transmission of Related Malaria Parasite Lineages Shapes
700 Within-Host Parasite Diversity. *Cell Host Microbe* **27**:93-103.e4. doi:10.1016/j.chom.2019.12.001
701 Normark J, Nilsson D, Ribacke U, Winter G, Moll K, Wheelock CE, Bayarugaba J, Kironde F, Egwang TG,
702 Chen Q, Andersson B, Wahlgren M. 2007. PfEMP1-DBL1alpha amino acid motifs in severe disease
703 states of Plasmodium falciparum malaria. *Proc Natl Acad Sci U S A* **104**:15835–40.
704 doi:10.1073/pnas.0610485104
705 Otto TD, Assefa SA, Böhme U, Sanders MJ, Kwiatkowski DP, Berriman M, Newbold C. 2019. Evolutionary
706 analysis of the most polymorphic gene family in falciparum malaria. *Wellcome Open Res* **4**:1–29.
707 doi:10.12688/wellcomeopenres.15590.1
708 Palstra FP, Fraser DJ. 2012. Effective/census population size ratio estimation: A compendium and
709 appraisal. *Ecol Evol* **2**:2357–2365. doi:10.1002/ece3.329

710 Paul RE, Packer MJ, Walmsley M, Lagog M, Ranford-Cartwright LC, Paru R, Day KP. 1995. Mating patterns
711 in malaria parasite populations of Papua New Guinea. *Science (1979)* **269**:1709–1711.
712 R Core Team. 2018. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical*
713 *Computing*.
714 Rask TS, Hansen D, Theander TG, Pedersen AG, Lavstsen T, Gorm Pedersen A, Lavstsen T, Pedersen AG,
715 Lavstsen T, Gorm Pedersen A, Lavstsen T. 2010. Plasmodium falciparum erythrocyte membrane
716 protein 1 diversity in seven genomes - divide and conquer. *PLoS Comput Biol* **6**:e1000933.
717 doi:10.1371/journal.pcbi.1000933
718 Rottmann M, Lavstsen T, Mugasa JP, Kaestli M, Jensen ATR, Müller D, Theander T, Beck H-PP. 2006.
719 Differential expression of var gene groups is associated with morbidity caused by Plasmodium
720 falciparum infection in Tanzanian children. *Infect Immun* **74**:3904–3911. doi:10.1128/IAI.02073-05
721 Ruybal-Pesántez S, Tiedje KE, Pilosof S, Tonkin-Hill G, He Q, Rask TS, Amenga-Etego L, Oduro AR, Koram
722 KA, Pascual M, Day KP. 2022. Age-specific patterns of DBLa var diversity can explain why residents of
723 high malaria transmission areas remain susceptible to Plasmodium falciparum blood stage infection
724 throughout life. *Int J Parasitol* **20**:721–731.
725 Ruybal-Pesántez S, Tiedje KE, Tonkin-Hill G, Rask TS, Kamya MR, Greenhouse B, Dorsey G, Duffy MF, Day
726 KP. 2017. Population genomics of virulence genes of Plasmodium falciparum in clinical isolates from
727 Uganda. *Sci Rep* **7**:11810. doi:10.1038/s41598-017-11814-9
728 Sjoberg, D, Whiting K, Curry M, Lavery J, Larmarange J. 2021. Reproducible Summary Tables with the
729 gtsummary Package. *R J* **13**:570–580.
730 Smith T, Felger I, Fraser-Hurt N, Beck H-PP. 1999. Effect of insecticide-treated falciparum infections bed
731 nets on the dynamics of multiple Plasmodium falciparum infections. *Trans R Soc Trop Med Hyg* **93**
732 **(Supp)**:S1/53-S1/57. doi:10.1016/S0035-9203(99)90328-0
733 Speed D, Balding DJ. 2015. Relatedness in the post-genomic era: Is it still useful? *Nat Rev Genet* **16**:33–44.
734 doi:10.1038/nrg3821
735 Stevenson M. 2020. epiR: Tools for the analysis of epidemiological data.
736 Sumner KM, Freedman E, Abel L, Obala A, Pence BW, Wesolowski A, Meshnick SR, Prudhomme-O'Meara
737 W, Taylor SM. 2021. Genotyping cognate Plasmodium falciparum in humans and mosquitoes to
738 estimate onward transmission of asymptomatic infections. *Nat Commun* **12**:1–12.
739 doi:10.1038/s41467-021-21269-2
740 Tan MH, Shim H, Chan Y, Day KP. 2023. Unravelling chaos for malaria surveillance : Relationship between
741 DBLα types and var genes in Plasmodium falciparum. *Frontiers in Parasitology* **1**.
742 doi:https://doi.org/10.3389/fpara.2022.1006341
743 Tessema SK, Hathaway NJ, Teyssier NB, Murphy M, Chen A, Aydemir O, Duarte EM, Simone W, Colborn J,
744 Saute F, Crawford E, Aide P, Bailey JA, Greenhouse B. 2020. Sensitive, highly multiplexed sequencing
745 of microhaplotypes from the Plasmodium falciparum heterozygome. *Journal of Infectious Diseases*
746 **225**:1227–1237.
747 Therneau T. 2023. A Package for Survival Analysis in R.
748 Tiedje KE, Oduro AR, Agongo G, Anyorigiya T, Azongo D, Awine T, Ghansah A, Pascual M, Koram KA, Day
749 KP. 2017. Seasonal Variation in the Epidemiology of Asymptomatic Plasmodium falciparum Infections
750 Across Two Catchment Areas in Bongo District, Ghana. *Am J Trop Med Hyg* **97**:199–212.
751 doi:10.4269/ajtmh.16-0959
752 Tiedje KE, Oduro AR, Bangre O, Amenga-Etego L, Dadzie SK, Appawu MA, Frempong K, Asoala V, Charles
753 A Narh, Samantha L Deed, Ruybal-Pesántez S, Argyropoulos DC, Anita Ghansah, Agyei SA, Segbaya S,

754 Desewu K, Williams I, Simpson JA, Malm K, Pascual M, Koram KA, Day KP. 2022. Indoor residual
755 spraying with a non-pyrethroid insecticide reduces the reservoir of Plasmodium falciparum in a high-
756 transmission area in northern Ghana. *PLOS Global Public Health* **2**:e0000285.
757 doi:10.1371/journal.pgph.0000285
758 US Agency for International Development (USAID) Global Health Supply Chain Program. 2020. Technical
759 Breif: Data visability makes all the difference in Ghana's 2018 LLIN mass distribution campaign.
760 US President's Malaria Initiative Africa IRS (AIRS) Project. 2016. Entomological monitoring of the PMI AIRS
761 program in northern Ghana: 2016 annual report. Bethesda, Marland, USA.
762 Wagman J, Gogue C, Tynuv K, Mihigo J, Bankineza E, Bah M, Diallo D, Saibu A, Richardson JH, Kone D,
763 Fomba S, Bernson J, Steketee R, Slutsker L, Robertson M. 2018. An observational analysis of the
764 impact of indoor residual spraying with non-pyrethroid insecticides on the incidence of malaria in
765 Ségou Region, Mali: 2012–2015. *Malar J* **17**:19. doi:10.1186/s12936-017-2168-2
766 Warimwe GM, Fegan G, Musyoki JN, Newton CRJC, Opiyo M, Githinji G, Andisi C, Menza F, Kitsao B, Marsh
767 K, Bull PC. 2012. Prognostic indicators of life-threatening malaria are associated with distinct parasite
768 variant antigen profiles. *Sci Transl Med* **4**:129ra45. doi:10.1126/scitranslmed.3003247
769 Warimwe GM, Keane TM, Fegan G, Musyoki JN, Newton CRJC, Pain A, Berriman M, Marsh K, Bull PC. 2009.
770 Plasmodium falciparum var gene expression is modified by host immunity. *Proc Natl Acad Sci U S A*
771 **106**:21801–21806.
772 Watson OJ, Okell LC, Hellewell J, Slater HC, Unwin HJT, Omedo I, Bejon P, Snow RW, Noor AM, Rockett K,
773 Hubbart C, Nankabirwa JI, Greenhouse B, Chang HH, Ghani AC, Verity R. 2021. Evaluating the
774 Performance of Malaria Genetics for Inferring Changes in Transmission Intensity Using Transmission
775 Modeling. *Mol Biol Evol* **38**:274–289. doi:10.1093/molbev/msaa225
776 WHO. 2012. WHO Policy Recommendation: Seasonal Malaria Chemoprevention (SMC) for Plasmodium
777 falciparum malaria control in highly seasonal transmission areas of the Sahel sub-region in Africa.
778 WHO/GMP. 2017. A Framework for Malaria Elimination, Geneva World Health Organization.
779 Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester
780 J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi
781 K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the Tidyverse. *J Open Source Softw*
782 **4**:1686. doi:10.21105/joss.01686
783 World Health Organization. 2022. World Malaria Report 2022, World Health Organization. doi:ISBN 978
784 92 4 1564403
785 Zhang X, Deitsch KW. 2022. The mystery of persistent, asymptomatic Plasmodium falciparum infections.
786 *Curr Opin Microbiol* **70**:102231. doi:10.1016/j.mib.2022.102231
787
788

789 **DATA AVAILABILITY STATEMENT**. The sequences utilized in this study are publicly available in GenBank
790 under BioProject Number: PRJNA 396962. All data associated with this study are available in the main text,
791 the Supporting Information, or upon reasonable request for research purposes to the corresponding
792 author, Prof. Karen Day (karen.day@unimelb.edu.au). All custom code is available in an open source
793 repository: (1) DBLαCleaner pipeline is available athttps://github.com/UniMelb-Day-Lab/DBLaCleaner, (2)
794 clusterDBLalpha pipeline is available at https://github.com/Unimelb-Day-Lab/clusterDBLalpha, and the (3)
795 classifyDBLalpha pipeline is available at https://github.com/Unimelb-Day-Lab/classifyDBLalpha. A dataset
796 to demo this custom code is available at https://github.com/UniMelb-Day-Lab/tutorialDBLalpha. For
797 additional information on the use of the Bayesian approach to estimate MOI_{var} please see
798 https://github.com/qzhan321/Bayesian-formulation-varcoding-MOI-estimation.
799
800 **BENEFIT-SHARING STATEMENT**. A research collaboration was developed with scientists from Ghana
801 based at the Navrongo Health Research Centre and the Noguchi Memorial Institute for Medical Research.
802 All collaborators are included as co-authors and the relevant results from the research has been shared
803 with key stakeholders and the local community (i.e., Paramount Chief of Bongo, divisional Chiefs, Queen
804 Mothers, and community members), the Bongo District and Upper East Regional Health Directorates, as
805 well as Ghana National Malaria Elimination Programme. Before this research was undertaken, informed
806 consent was sought and obtained from the key stakeholders, traditional leadership, and the local
807 community in Bongo District. In addition, members of the local community were trained as field workers
808 and were directly involved in liaising with the community and in the collection of the study data. The
809 contribution of these individuals to this research is described in the Acknowledgements. This research
810 addresses a priority concern regarding malaria control and the impact of interventions. These concerns

811 are relevant for both the local community in Bongo District, as well as for the National Malaria Elimination
812 Programme in Ghana.
813
814 **AUTHOR CONTRIBUTIONS.** K.P.D and M.P conceptualized the study of population size using MOI_{var}. K.A.K
815 and K.P.D conceived and designed the field study with input from A.R.O and K.E.T, who together with O.B
816 coordinated the field studies. K.E.T, S.R-P, and S.L.D processed the samples and performed the genotyping
817 experiments. G.T-H developed and validated the bioinformatic pipelines. K.E.T, S.R-P, and D.C.A
818 processed, cleaned, and curated the datasets for analysis. Q.Z developed and implemented the Bayesian
819 estimations. K.E.T, Q.Z, M.P, and K.P.D analysed the data with contributions from S.R-P, Q.H, and M-H.T.
820 K.E.T and Q.Z visualized the data. K.P.D, M.P, and K.E.T wrote the original draft of the manuscript. Q.Z.,
821 S.R-P, G.T-H, Q.H, M-H.T, D.C.A, S.L.D, O. B, A.R.O, and K.AK reviewed and edited manuscript. M.P and K.P.D
822 supervised the research. A.R.O, K.A.K, M.P, and K.P.D acquired the funding.
823
824 **COMPETING INTERESTS.** The authors declare no competing interests.

FIGURES

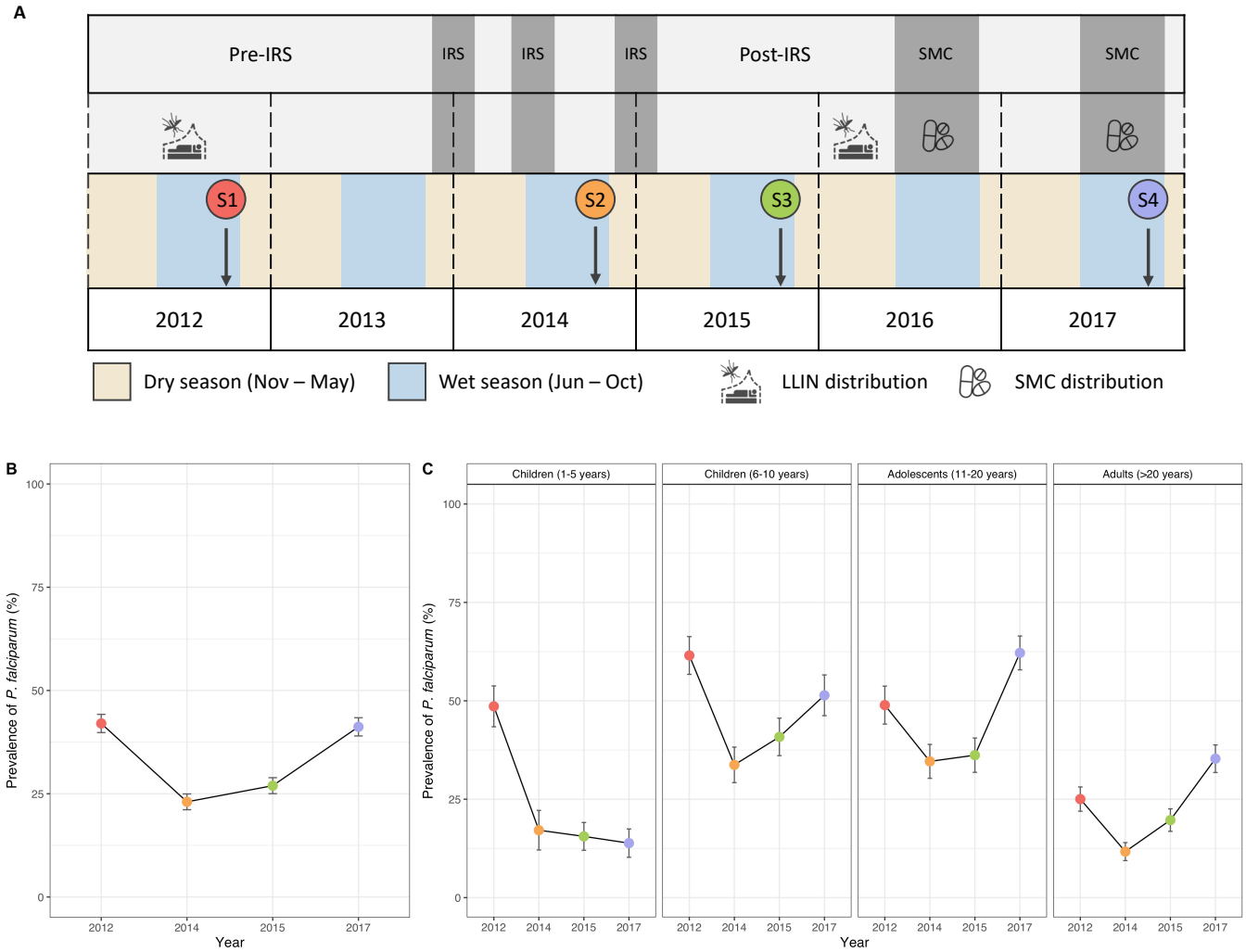

Figure 1. Study design and changes in the prevalence of microscopic *P. falciparum* infection following the IRS and SMC interventions in Bongo, Ghana. (**A**) Four age-stratified cross-sectional surveys of ~2,000 participants per survey were conducted in Bongo, Ghana at the end of the wet seasons in October 2012 (Survey 1, baseline pre-IRS, red), October 2014 (Survey 2, during IRS, orange), October 2015 (Survey 3, post-IRS, green), and October 2017 (Survey 4, SMC, purple) (Table supplement 1) (see Materials and Methods). The three rounds of IRS (grey areas) were implemented between 2013 and 2015 (Tiedje et al., 2022). SMC was distributed to all children < 5 years of age during the wet seasons in 2016 (two rounds between August – September 2016) and 2017 (four rounds between September – December 2017) (Gogue et al., 2020). Both IRS and SMC were implemented against a background of widespread LLIN usage (Tiedje et al., 2022). Prevalence of microscopic *P. falciparum* infections (%) in the (**B**) study population and (**C**) for all age groups (years) in each survey. Error bars represent the upper and lower limits of the 95% confidence interval (CI) calculated using the Wald Interval.

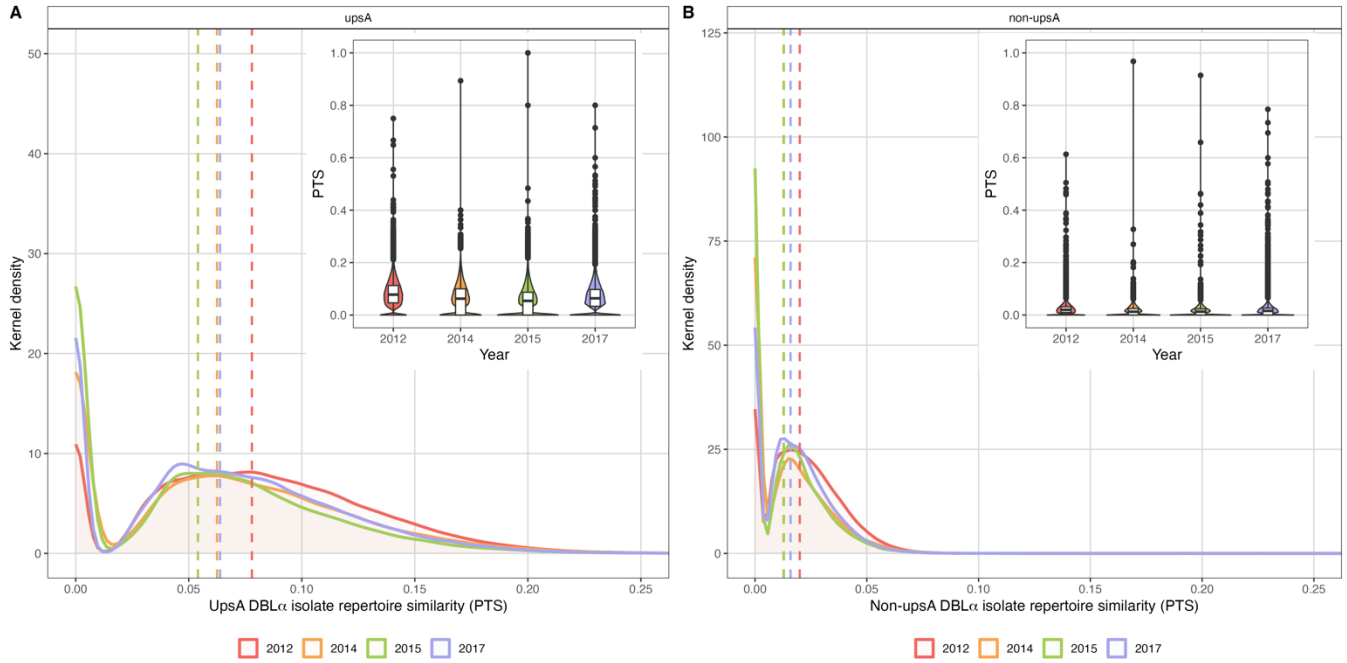

Figure 2. Sharing of upsA and non-upsA DBLα types among the DBLα isolate repertoires in 2012 (pre-IRS, red), 2014 (during IRS, orange), 2015 (post-IRS, green), and 2017 (SMC, purple). The overlapping density and violin plots (upper right-hand corners) show the distribution of PTS scores (i.e., DBLα isolate repertoire similarity) between the (**A**) upsA and (**B**) non-upsA DBLα isolate repertoires for those isolates with DBLα sequencing data (Table supplement 2) in each survey. The PTS scales in the density plots have been zoomed-in to provide better visualization of the upsA and non-upsA DBLα types PTS distributions. The colored dashed lines in the density plots indicate the median PTS scores in each survey for the upsA (2012 (red) = 0.078, 2014 (orange) = 0.063, 2015 (green) = 0.054, and 2017 (purple) = 0.064) and non-upsA (2012 (red) = 0.020, 2014 (orange) = 0.013, 2015 (green) = 0.013, and 2017 (purple) = 0.016) DBLα types. *Note:* The non-upsA median PTS values in 2014 (orange) and 2015 (green) were both 0.013 and overlap in the figure. In the PTS violin plots the central box plots indicate the medians (centre line), interquartile ranges (IQR, upper and lower quartiles), whiskers (1.5x IQR), and outliers (points).

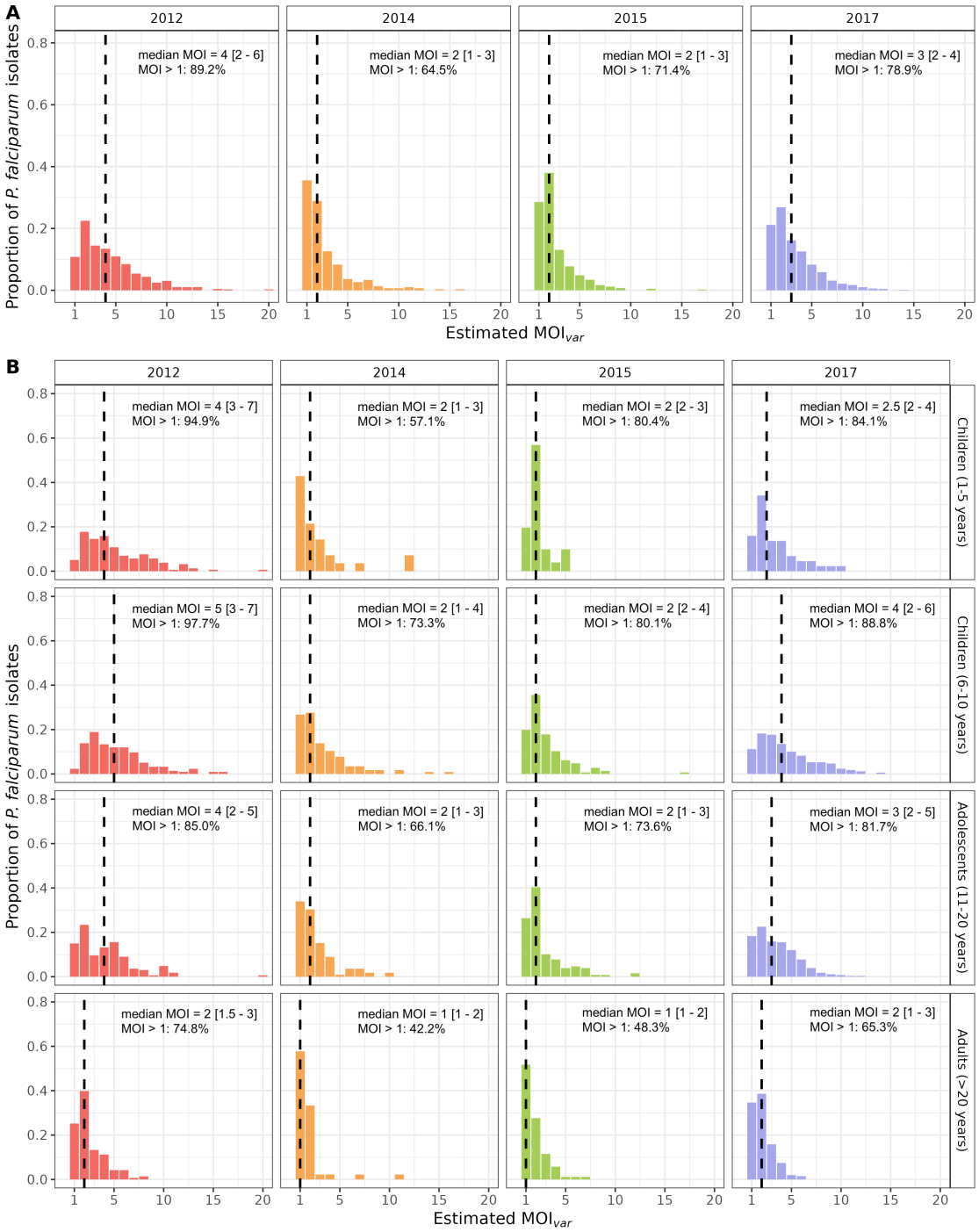

Figure 3. MOI_{var} distributions in 2012 (pre-IRS, red), 2014 (during IRS, orange), 2015 (post-IRS, green), and 2017 (SMC, purple) based on pooling the maximum *a posteriori* MOI estimates. Estimated MOI_{var} distributions for the (**A**) study population and (**B**) for all age groups (years) in each survey for those isolates with DBLα sequencing data (Table supplements 2 and 7). The median MOI_{var} values are indicated with the black dashed lines and have been provided in the top right corner (median MOI_{var} value [interquartile range, upper and lower quartiles]) along with the percentage of *P. falciparum* infections that were multiclonal (MOI_{var} > 1) in each survey and age group (years).

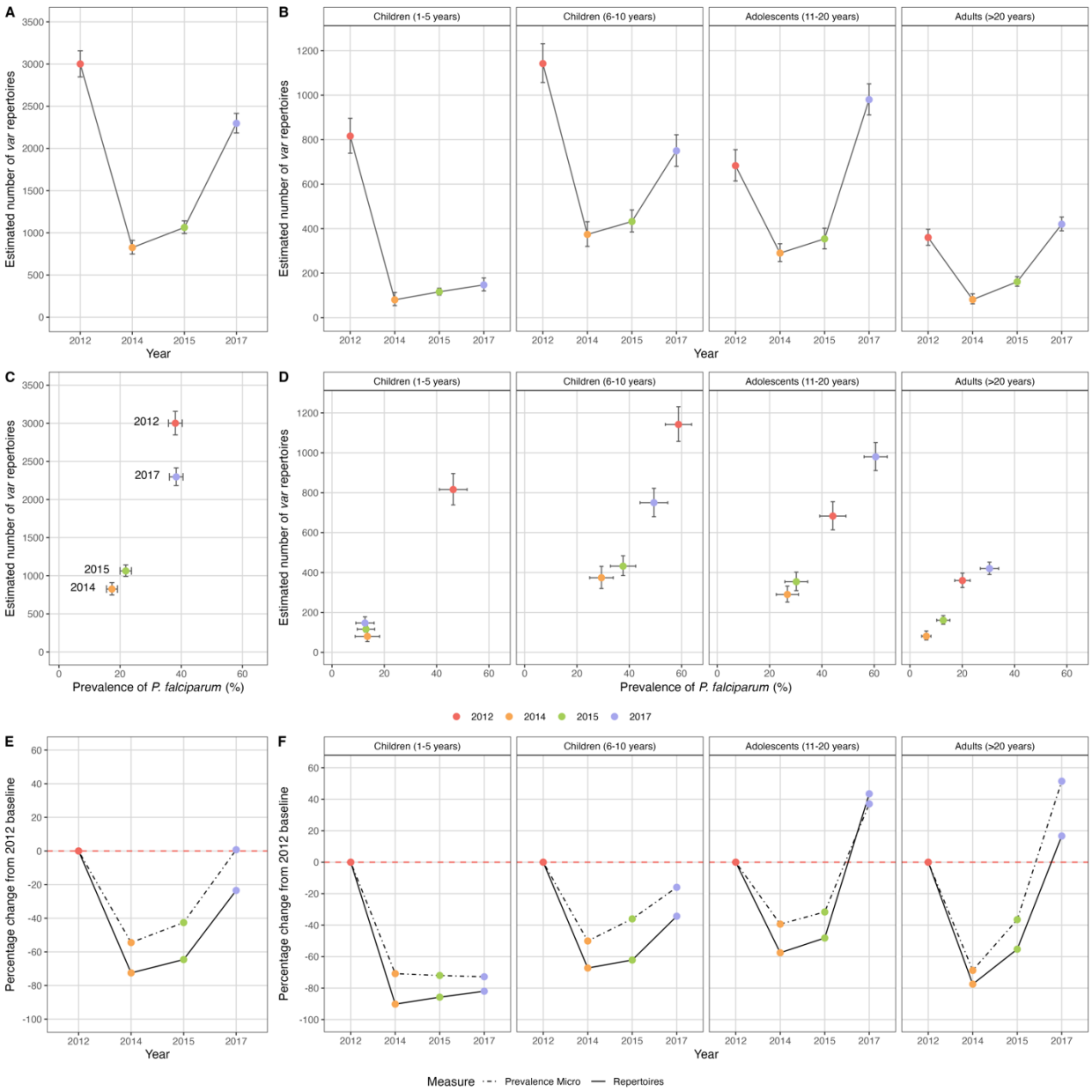

Figure 4. Estimated number and relative change in the number of *P. falciparum var* repertoires in 2012 (pre-IRS, red), 2014 (during IRS, orange), 2015 (post-IRS, green), and 2017 (SMC, purple). The estimated number of *var* repertoires (i.e., census population size) for those isolates with DBLα sequencing data (Table supplements 2 and 7) in the (**A**) study population and (**B**) for all age groups (years). The estimated number of *var* repertoires vs. *P. falciparum* prevalence for (**C**) study population and (**D**) for all age groups (years) (Table supplement 7). The percentage change in *P. falciparum* prevalence (black dotted line) and the estimated number of *var* repertoires (black solid line) in 2014, 2015, and 2017 compared to the 2012 baseline survey (red dashed horizontal line at 0% change) for the (**E**) study population and (**F**) for all age groups (years). Error bars in (**A-D**) represent the upper and lower limits of the 95% confidence intervals (95% CIs). The 95% CIs for the number of *var* repertoires (i.e., census population size) were

871 calculated based on a bootstrap approach. We resampled 10,000 replicates from the original population-
872 level distribution with replacement. Each resampled replicate has the same size as the original sample.
873 We then derive the 95% CI based on the distribution of the resampled replicates. The 95% CIs for *P.*
874 *falciparum* prevalence (%) were calculated using the Wald Interval.

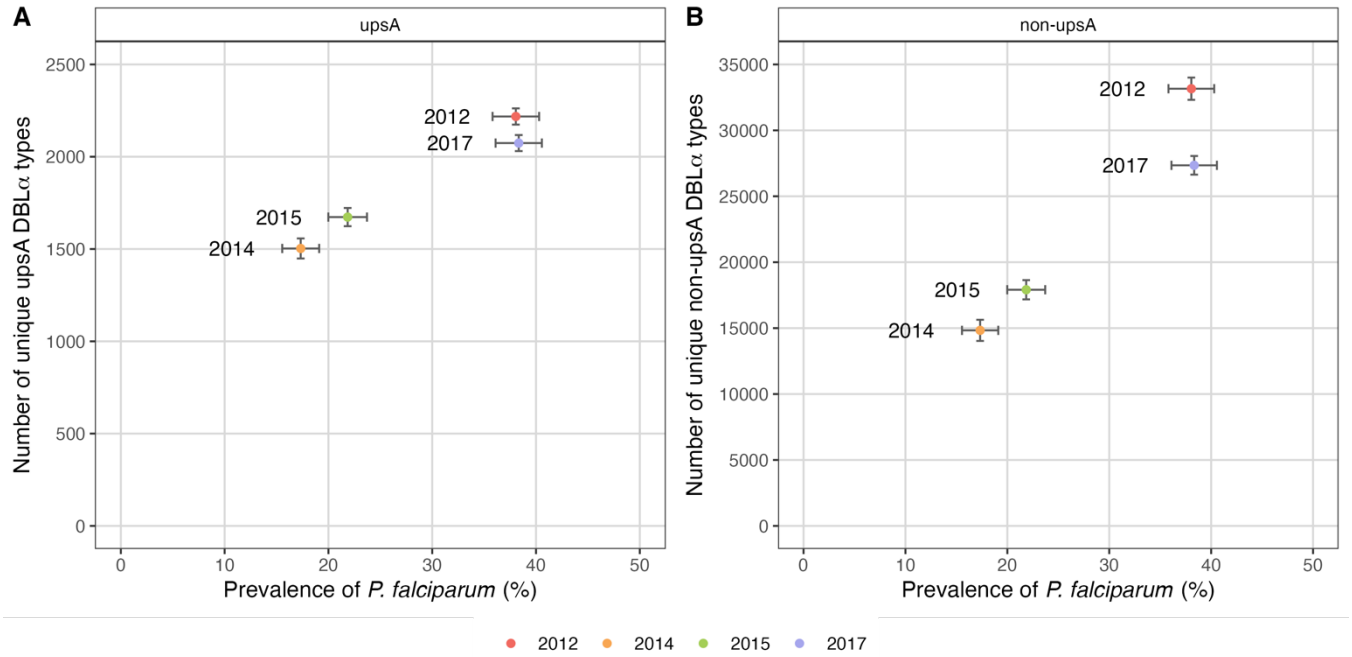

Figure 5. UpsA and non-upsA DBLα type richness in 2012 (pre-IRS, red), 2014 (during IRS, orange), 2015 (post-IRS, green), and 2017 (SMC, purple). Number of unique (**A**) upsA and (**B**) non-upsA DBLα types (i.e., richness) observed in each survey vs. *P. falciparum* prevalence based on those isolates with DBLα sequencing data (Table supplement 2). Error bars represent the upper and lower limits of the 95% confidence intervals (95% CIs) for the *P. falciparum* prevalence (%; x-axis) and ± 2 standard deviations (± 2SD) for the number of unique upsA and non-upsA DBLα types (y-axis). The 95% CIs for *P. falciparum* prevalence (%) were calculated using the Wald Interval. The ± 2SD for the number of unique upsA and non-upsA DBLα types were calculated based on a bootstrap approach. We resampled 10,000 replicates from the original population-level distribution with replacement. Each resampled replicate has the same size as the original sample. We then derive the standard deviation (SD) based on the distribution of the resampled replicates.

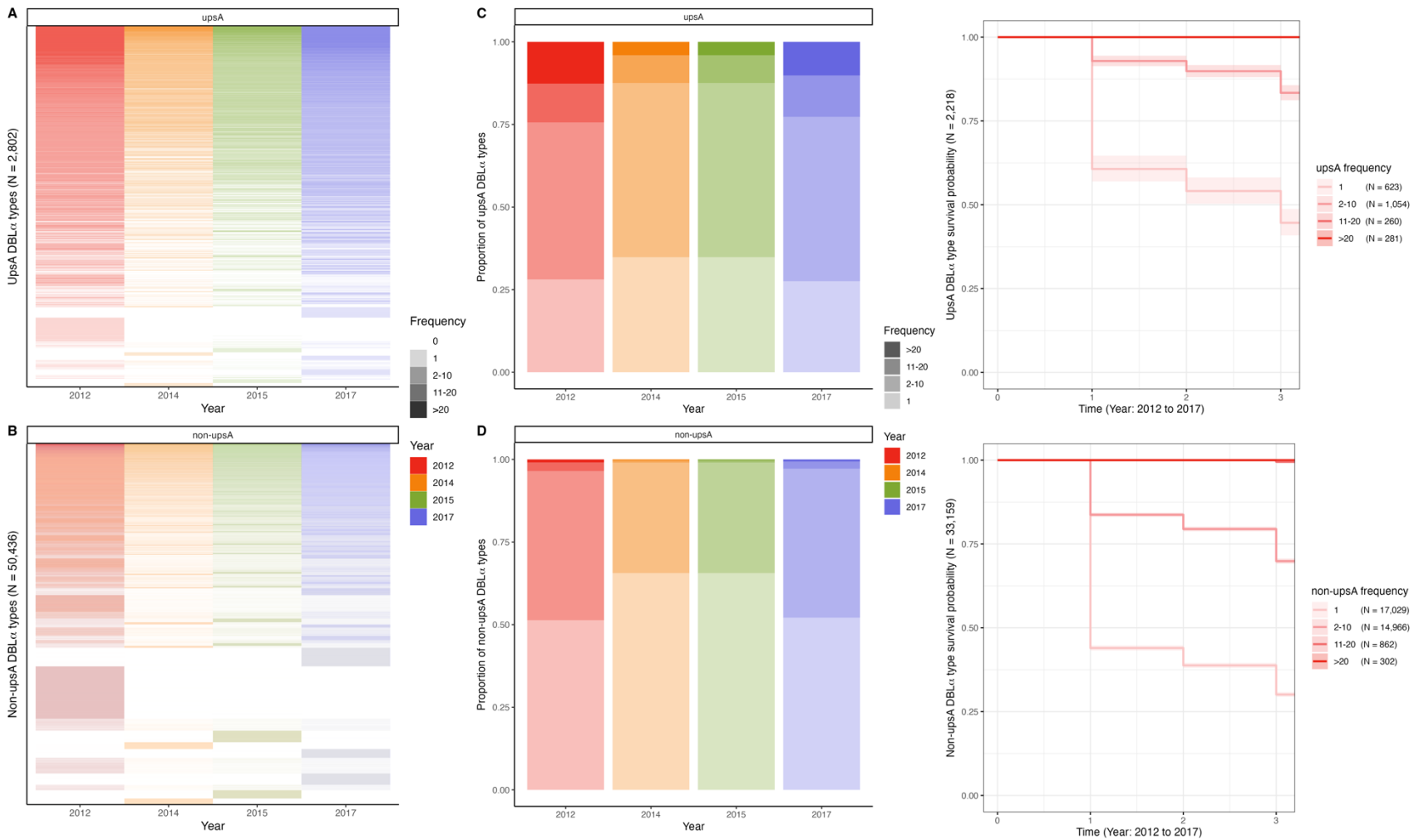

Figure 6. UpsA and non-upsA DBLα type frequencies and survival in 2012 (pre-IRS, red), 2014 (during IRS, orange), 2015 (post-IRS, green), and 2017 (SMC, purple). Heatmaps showing the patterns of diversity for the (**A**) upsA and (**B**) non-upsA DBLα types. The columns represent all the upsA and non-upsA DBLα types observed in the four surveys, and the rows represent each of the 2,802 upsA DBLα types and the 50,436 non-upsA DBLα types (Table supplement 2). White rows are used to denote the absence of a specific DBLα type, while the presence of a DBLα type is indicated in colour and further categorised (colour gradations) based on the frequency or the number of times (i.e., number of isolates) a DBLα type was observed in each survey (Frequency categories: 1, 2-10, 11-20, > 20 isolates; Note the frequency category cut-offs were chosen based on the frequency distributions in Figure supplement 7). The proportions of (**C**)

895 upsA and (**D**) non-upsA DBLα types in each survey based on the number of times (i.e., number of isolates) they were observed in each
896 survey. Kaplan-Meier survival curves for the (**E**) upsA and (**F**) non-upsA DBLα types across time (2012 to 2017) categorised based on
897 their frequency at baseline in 2012 (pre-IRS, red). The colored shaded areas represent the upper and lower limits of the 95% confidence
898 intervals (95% CIs) with the number (N) of upsA and non-upsA DBLα types in each frequency category are provided in parenthesis. These
899 survival curves include only those upsA (N = 2,218) and non-upsA (N = 33,159) DBLα types that were seen at baseline in 2012 (pre-IRS)
900 as indicated in red (Table supplement 2). The x-axis indicates time where time "0" denotes 2012 (pre-IRS), "1" denotes 2014 (during
901 IRS), "2" denotes 2015 (post-IRS), and finally "3" denotes 2017 (SMC). *Note:* In the survival curves the 11-20 and > 20 frequency
902 categories for both the (**E**) upsA and (**F**) non-upsA DBLα types overlap in the figure.