

1 **Emergence and evolution of *Plasmodium falciparum* histidine-rich protein 2** 2 **and 3 deletion mutant parasites in Ethiopia**

3

4 Sindew M. Feleke^{1a*}, Emily N. Reichert^{2a}, Hussein Mohammed¹, Bokretsion G.
5 Brhane¹, Kalkidan Mekete¹, Hassen Mamo³, Beyene Petros³, Hiwot Solomon⁴, Ebba
6 Abate¹, Chris Hennelly², Madeline Denton², Corinna Keeler², Nicholas J. Hathaway⁵,
7 Jonathan J. Juliano², Jeffrey A. Bailey⁶, Eric Rogier⁷, Jane Cunningham^{8b*}, Ozkan
8 Aydemir^{6b}, Jonathan B. Parr^{2b*}

9 (a Co-first authors; b Co-senior authors; * Corresponding authors)

10

11 **Affiliations**

- 12 1. Ethiopian Public Health Institute (EPHI), Addis Ababa, Ethiopia
- 13 2. Institute for Global Health and Infectious Diseases and Department of
14 Medicine, Division of Infectious Diseases, University of North Carolina at
15 Chapel Hill (UNC), USA
- 16 3. Department of Microbial, Cellular and Molecular Biology, College of Natural
17 and Computational Sciences, Addis Ababa University, Addis Ababa, Ethiopia
- 18 4. Ministry of Health (MoH), Ethiopia
- 19 5. Department of Medicine, University of Massachusetts Medical School,
20 Worcester, Massachusetts, USA
- 21 6. Department of Pathology and Laboratory Medicine, Warren Alpert Medical
22 School, Brown University, Providence, RI, USA
- 23 7. Division of Parasitic Diseases and Malaria, Center for Disease Control and
24 Prevention (CDC), Atlanta, USA
- 25 8. Global Malaria Programme, World Health Organization (WHO), Geneva,
26 Switzerland

27

28

29 **Key words:** hrp2, hrp3, malaria, molecular inversion probe, genomics, deletion,
30 evolution, Ethiopia

31 **Abstract**

32 Malaria diagnostic testing in Africa is threatened by *Plasmodium falciparum* parasites
33 lacking histidine-rich protein 2 (*pfhrp2*) and 3 (*pfhrp3*) genes. Among 12,572 subjects
34 enrolled along Ethiopia's borders with Eritrea, Sudan, and South Sudan and using
35 multiple assays, we estimate HRP2-based rapid diagnostic tests would miss 9.7%
36 (95% CI 8.5-11.1) of falciparum malaria cases due to *pfhrp2* deletion. Established
37 and novel genomic tools reveal distinct subtelomeric deletion patterns,
38 well-established *pfhrp3* deletions, and recent expansion of *pfhrp2* deletion. Current
39 diagnostic strategies need to be urgently reconsidered in Ethiopia, and expanded
40 surveillance is needed throughout the Horn of Africa.

41 *Plasmodium falciparum* strains that evade diagnosis by rapid diagnostic tests
42 (RDTs) represent a major threat to malaria control and elimination efforts^{1,2}. Malaria
43 RDTs detect antigens produced by *Plasmodium* parasites, including *P. falciparum*
44 histidine-rich protein 2 (HRP2), parasite lactate dehydrogenase (LDH), and aldolase.
45 HRP2 has advantages over other biomarkers due to its abundance in the
46 bloodstream, repetitive binding epitopes, and falciparum-specificity^{3,4}. Most
47 HRP2-based RDTs also exhibit some cross-reactivity to a closely related protein
48 (HRP3). HRP2-based RDTs are currently the predominant malaria diagnostic test
49 employed throughout sub-Saharan Africa⁵⁻⁷.

50 Deletion mutations involving the histidine-rich protein 2 and/or 3 (*pfhrp2/3*)
51 genes allow parasite strains to escape HRP2-based RDT detection^{8,9}. First described
52 in clinical samples from Peru in 2010, these subtelomeric deletions on chromosomes
53 8 (*pfhrp2*) and 13 (*pfhrp3*) are frequently large (≥ 20 kb), encompass multiple genes,
54 and are difficult to study using existing methods⁹⁻¹¹. Improved PCR and serological
55 approaches can be used to increase confidence in deletion prevalence
56 estimates¹²⁻¹⁴, but our understanding of the evolutionary history of *pfhrp2/3*-deleted
57 *P. falciparum* is limited and largely informed by analysis of a small number of
58 microsatellite markers¹⁵⁻¹⁷. Recent genomic analyses have begun to expand our
59 understanding of *pfhrp2/3*-deleted *P. falciparum*¹⁸⁻²⁰ but continue to be hindered by
60 the challenges of assembling the highly repetitive and paralogous sequences of *P.*
61 *falciparum*'s subtelomeres²¹. New tools are needed to support surveillance of
62 *pfhrp2/3* deletions and to determine their true prevalence and the forces impacting
63 their evolution and spread.

64 Increasing reports of these "diagnostic resistant" *pfhrp2/3*-deleted parasites in

65 Africa in 2017-2018 prompted calls for urgent surveillance in affected regions,
66 including countries in the Horn of Africa like Ethiopia^{15,22–25}. Ethiopia is Africa's
67 second most populous country, and 68% of its population is at risk of malaria
68 exposure²⁶. *P. falciparum* infection accounts for the majority of malaria deaths²⁷ and
69 approximately 70% of all cases²⁶. RDTs were first introduced in Ethiopia in 2004, and
70 the country's current test-treat-track strategy requires parasitological confirmation
71 either by quality microscopy or RDT prior to antimalarial treatment²⁸. *P.*
72 *falciparum*-*Plasmodium vivax* (HRP2/Pv-specific-LDH) combination RDTs are the
73 sole diagnostic test used in most settings. Over the last decade, Ethiopia has
74 achieved remarkable progress in the fight against malaria through strong
75 preventative and case management interventions, including engagement of
76 volunteers to provide diagnostic services at a local level²⁸. Reports of highly
77 prevalent *pfhrp2/3*-deleted parasites in neighboring Eritrea suggest that these gains
78 could be threatened^{15,24}. Rapid assessment of the epidemiology of *pfhrp2/3* deletions
79 in Ethiopia and surrounding regions is required to determine whether a change in
80 malaria diagnostic testing policy is warranted.

81 Here, we describe the first prospective, multi-site study of *pfhrp2/3*-deleted *P.*
82 *falciparum* in sub-Saharan Africa based on WHO's *pfhrp2/3* deletion surveillance
83 protocol²⁹, released in 2018 to encourage a harmonized and representative
84 approach to *pfhrp2/3* deletion surveillance and accurate reporting. Including sites
85 spanning Ethiopia's borders with Eritrea, Sudan, and South Sudan, we apply both
86 established and novel genomic tools to determine the genetic epidemiology of
87 *pfhrp2/3*-deleted *P. falciparum*, confirming deletions using multiple PCR assays¹⁴, an
88 ultrasensitive bead-based immunoassay for antigen detection¹³, whole-genome

89 sequencing (WGS)³⁰, and/or molecular inversion probe (MIP) deep sequencing³¹.
90 Using a MIP panel designed for high-throughput *pfhrp2/3* genotyping, we map and
91 categorize deletion breakpoints and evaluate their flanking regions for evidence of
92 recent evolutionary pressure favoring *pfhrp2/3*-deleted parasites in Ethiopia.

93

94 **RESULTS**

95 ***Study Population and RDT Results***

96 A total of 12,572 study participants (56% male, 44% female) between the
97 ages of 0 to 99 years who presented with one or more of symptoms consistent with
98 malaria at 108 health facilities in the Amhara, Tigray, and Gambella regions between
99 November 2017 and April 2018 were enrolled (**Table 1**). Median participant age was
100 19 years (interquartile range [IQR]: 8-30). From the same finger prick, participants
101 were tested with two RDTs, including the routine HRP2/Pv-specific-LDH RDT
102 combination test [CareStart *Pf/Pv* RDT (Access Bio, Somerset, NJ; product code RM
103 VM-02571)] and the survey HRP2/Pf-specific-LDH RDT [SD Bioline Malaria Ag P.f.
104 RDT (Alere, Waltham, MA; product code 05FK90)].

105 Overall, 2,714 (22%) study participants were *P. falciparum* positive by at least
106 one RDT (any HRP2 or Pf-LDH positive band); among these, 361 (13.3%; 95%
107 confidence interval [CI] 12.1-14.7%) had a discordant RDT profile suggestive of
108 *pfhrp2/3*-deleted *P. falciparum* infection, which was defined as HRP2-negative by
109 both RDTs but Pf-LDH positive. Among the 2,714 samples that were *P. falciparum*
110 positive by RDT, the northern region of Tigray had the highest proportion of
111 infections with discordant RDT profiles at 140/689 (20.4%; 95% CI 17.5-23.7%),
112 followed by Amhara with 211/1342 (15.8%; 13.9-17.8) and Gambella with 10/683

113 (1.5%; 0.7-2.8), as shown in **Figure 1**.

114

115 ***Pfhrp2/3* deletion PCR genotyping**

116 820 samples with complete demographic and clinical data from Amhara (n =
117 524), Tigray (n = 225), and Gambella (n = 71) underwent molecular analysis. These
118 samples were collected from subjects with the discordant RDT profile and a subset
119 of subjects with other RDT results (**Supplementary Figure 1**), including a randomly
120 selected 248/361 (68.7%) of those with the discordant RDT profile of interest
121 (HRP2-, Pf-LDH+), as well as 465/2115 (22.0%) randomly selected *P. falciparum*
122 RDT HRP2 positives (HRP2+, Pf-LDH+) as controls. The remaining 107 samples
123 included 90 with inconclusive HRP2 results (HRP2+ by only 1 RDT, of which 67 were
124 Pf-LDH- and 23 were Pf-LDH+) and 17 negative controls (HRP2-, Pf-LDH-).
125 Quantitative real-time PCR (qPCR) targeting the *P. falciparum* lactate
126 dehydrogenase (*pfldh*) gene confirmed parasitemia in 731/820 (89%) samples, with
127 a geometric mean (GM) of 1,390.7 parasites/ μ L (geometric standard deviation
128 [geoSD]: 9.8). Further analysis was restricted to the 610 samples with >100
129 parasites/ μ L to avoid misclassification of *pfhrp2/3* deletions due to low parasitemia
130 (**Supplementary Figure 2**); 176 (28.9%) had the discordant RDT profile.

131 Infection by *pfhrp2/3*-negative parasites was common among these 610
132 subjects when assessed by PCR, with 355 (58%; 95% CI 54-62) lacking detectable
133 *pfhrp2* and/or *pfhrp3*, and 136 (22%; 19-26) lacking both *pfhrp2* and *pfhrp3*. For
134 those lacking only one gene, *pfhrp3*-negative infections [192 (31%; 28-35)
135 *pfhrp2*+/*pfhrp3*-] were more prevalent than *pfhrp2*-negative infections [27 (4.4%; 3-6)
136 *pfhrp2*-/*pfhrp3*+]. Concordance between *pfhrp2*-negative PCR results and the

137 discordant RDT profile was good (Cohen's kappa 0.66). Overall, among samples
138 with the discordant RDT profile, 64.8% (95% CI 57-72) were *pfhrp2*-/*3*- and 8.0%
139 (5-13) *pfhrp2*-/*3*+, with an additional 15.9% (11-22) *pfhrp2*+/*3*- (**Supplementary**
140 **Figure 3**). Interestingly, of samples HRP2+ by both RDTs, *pfhrp3* could not be
141 amplified in 42.6% (38-48). We observed expected agreement between the results of
142 *pfhrp2/3* PCR assays, RDTs, and a bead-based HRP2 immunoassay applied to a
143 randomly selected subset of samples (see **Supplementary Results, Table 2**). No
144 associations between *pfhrp2/3* PCR result and age, sex, or parasitemia were
145 identified (**Supplementary Table 1**).

146

147 ***Pfhrp2/3* deletion prevalence estimates**

148 Incorporating RDT and PCR results, we estimated that 9.7% (95% CI
149 8.5-11.1) of all *P. falciparum* infections across all study sites would have
150 false-negative HRP2-based RDT results due to *pfhrp2* deletions. Regional
151 prevalence of false-negative RDTs due to *pfhrp2*-deleted parasites varied, with the
152 highest estimates in Tigray (14.9%; 12.5-17.7), followed by Amhara (11.5%;
153 9.8-13.4) and Gambella (1.1%; 0.6-2.0). Our prevalence estimates only include
154 samples with both the discordant RDT profile and a *pfhrp2*-negative call by PCR.
155 Parasites with a deletion of *pfhrp2* but intact *pfhrp3* and sufficient cross-reactive
156 HRP3 to trigger a positive HRP2 band on either RDT are not included in these
157 estimates. Thus, the estimated prevalence of false-negative RDT results caused by
158 *pfhrp2* deletions likely underestimates the true prevalence of *pfhrp2*-deleted
159 parasites in this study.

160

161 ***Pfhrp2/3* deletion characterization using MIP sequencing**

162 To enable mapping of *pfhrp2/3* deletion regions and population genetic
163 analyses in large-scale epidemiological studies, we developed a targeted panel of
164 241 MIPs for highly multiplexed deep sequencing of *pfhrp2*, *pfhrp3*, and flanking
165 genes on chromosomes 8 and 13. A tiled design strategy was employed that
166 involved multiple, overlapping probes spanning each gene target. We detected 244
167 of 273 targets with sufficient mapping quality and depth across multiple segments of
168 both *pfhrp2* and *pfhrp3* and their flanking regions, spanning positions 1,344,451 to
169 1,397,773 and 2,780,863 to 2,853,533 of chromosomes 8 and 13, respectively.
170 Fourteen total probes were used to target different segments of *pfhrp2* (n = 8
171 probes) and *pfhrp3* (n = 6). Among *P. falciparum* PCR-positive samples collected
172 from 926 subjects and subjected to MIP capture and sequencing, 375 (40.5%) had
173 sufficient depth of coverage to make high-confidence calls. In total, 43,541,045 reads
174 were devoted to this sample set, or roughly half of a single NextSeq 550 mid-output
175 flow cell. The median parasite density for samples successfully called using MIP
176 sequencing was 5,077 p/μL (SD: 1.6 x 10⁴), compared to 264 p/μL (SD: 5.9 x 10³) for
177 samples with failed MIP calls. Analysis of variant-called MIP sequences confirmed
178 mixed infections with complexity of infection ≥2 in only 45 (12%) subjects; the
179 majority (n=330, 88%) were infected by a single *P. falciparum* strain.

180 Among 367 (97.9%) MIP-called samples with matching PCR data, 85 (23.2%)
181 were *pfhrp2-/3-* by PCR. MIP sequencing results indicated that 126/367 (34.3%; 95%
182 CI 30-39) were *pfhrp2-*, 264/367 (71.9%; 67-76) *pfhrp3-*, and 116/367 (31.6%; 27-37)
183 *pfhrp2-/3-* by MIP sequencing (**Figure 2, Table 2**). Receiver-operator curve analysis
184 indicated the optimal parasite density threshold above which samples had sufficient

185 coverage for MIP calling was approximately 925 p/μL, although this threshold is
186 project-specific and is expected to improve with additional sequencing effort
187 **(Supplementary Figure 4)**.

188

189 ***Comparison of pfhrp2/3 MIP sequencing to PCR, HRP2 immunoassay, and***
190 ***whole-genome sequencing results***

191 Of samples called *pfhrp2*- by MIP sequencing, 82.0% were *pfhrp2*- by PCR
192 and 73.3% had the discordant RDT profile. Similarly, of samples called *pfhrp3*- by
193 MIP sequencing, 76.7% were *pfhrp3*- by PCR. While differences between
194 genotyping results were apparent and expected due to differences in targets and
195 methodologies (PCR is better suited for samples with low parasite density than MIP
196 sequencing), there was overall strong concordance between RDT, PCR, and MIP
197 results **(Table 2)**. Comparison of results from MIP and PCR *pfhrp2/3* deletion
198 genotyping revealed excellent agreement between the two approaches for *pfhrp2*
199 (Cohen's kappa: 0.82) and good agreement for *pfhrp3* (Cohen's kappa: 0.63).

200 Comparison to the bead-based HRP2 immunoassay results provided
201 additional confidence in the validity of our *pfhrp2/3* deletion calls using MIP
202 sequencing. 175 MIP-called samples also had bead-based antigen detection results
203 available. Despite fundamental differences in the targets of these two approaches
204 (*pfhrp2* gene versus HRP2 antigen, which can linger after clearance of infection),³²
205 observed agreement between the two methods was consistent with expectation. Of
206 those samples *pfhrp2*+ by MIP sequencing, 94.1% were HRP2+ by bead-based
207 antigen immunoassay, whereas 79.5% of those *pfhrp2*- by MIPs were also HRP2- by
208 the immunoassay.

209 We used whole-genome sequencing (WGS) to evaluate *pfhrp2/3* MIP
210 sequencing results and breakpoint regions. Among 14 samples subjected to both
211 WGS and MIP sequencing, median WGS depth of coverage was 20 reads/locus
212 (range 4-38). While the distribution of aligned reads was uneven in the regions
213 flanking *pfhrp2* and *pfhrp3*, visual inspection of WGS coverage supported 13 (93%)
214 *pfhrp2* and 14 (100%) *pfhrp3* deletion calls made using MIP sequencing data
215 (**Figure 3**). For the single discordant *pfhrp2* deletion call (lab ID: 1314), *pfhrp2* PCR
216 results were consistent with the MIP sequencing result. Precise mapping of
217 breakpoint regions using WGS was not possible due to regions of very high
218 coverage (“jackpotting”) resulting from selective whole-genome amplification and low
219 coverage due to ambiguous read mapping to repetitive and paralogous loci.
220 However, breakpoint regions identified using MIPs were consistent with WGS
221 coverage centromeric to *pfhrp2* and *pfhrp3* on chromosomes 8 and 13, respectively,
222 with the exception of calls in chromosome 13’s multi-copy 28S rRNA gene.
223 Discordance in these calls was expected due to ambiguous mapping of short-read
224 sequences to a multi-copy gene. MIP results from well-characterized lab strains 3D7
225 (*pfhrp2+/3+*), DD2 (*pfhrp2-/3+*), and HB3 (*pfhrp2+/3-*) were consistent with
226 whole-genome alignments of published short-read data. Telomeric deletion
227 breakpoint assessment was limited by a small number of successful MIP targets
228 telomeric to both genes. However, the concordance in *pfhrp2/3* deletion calls and
229 centromeric deletion breakpoint regions by MIP and WGS techniques confirmed the
230 utility of MIPs for identifying *pfhrp2/3* deletions and determining their extent and
231 breakpoint regions.
232

233 ***Pfhrp2/3 deletion breakpoint profiling***

234 Compared to PCR, bead-based immunoassay, or RDT diagnosis, MIP
235 sequencing was unique in its ability to reveal distinct subtelomeric structural profiles
236 along chromosomes 8 and 13 into which samples could be categorized: three for
237 *pfhrp2*⁺ samples (chr8-P1, chr8-P2, chr8-P3), one for *pfhrp2*⁻ (chr8-P4), one for
238 *pfhrp3*⁺ (chr13-P1), and three for *pfhrp3*⁻ (chr13-P2, chr13-P3, chr13-P4) (**Figure 2**).
239 All *pfhrp2*⁻ samples had the same subtelomeric structural profile (chr8-P4), although
240 two other subtelomeric deletions were identified on chromosome 8 that did not
241 involve *pfhrp2* (chr8-P2, chr8-P3). These deletions involve members of the *rifin* and
242 *stevor* gene families, as well as genes of unknown function.

243 The structural profile of most samples identified as *pfhrp3*⁻ (chr13-P3 and
244 chr13-P4) differed in the presence or absence of a segment of chromosome 13
245 directly telomeric to *pfhrp3* (position 2,852,540 - 2,853,533) encoding a member of
246 the acyl-coA synthetase family (PF3D7_1372400). All chr13-P3 and chr13-P4
247 deletions resulted in loss of genes with roles in red blood cell invasion
248 (PF3D7_1371700, serine/threonine kinase and member of the FIKK family;
249 PF3D7_1371600, erythrocyte binding-like protein 1 [*EBL-1*]),^{33,34} while they were
250 present in all chr13-P1 (*pfhrp3*-intact) parasites. The chr13-P2 deletion profile was
251 observed in only one sample from Amhara's Metema district. We did not observe an
252 association between subtelomeric structural profile and the number of symptoms
253 experienced by subjects (**Supplementary Figure 5, Supplementary Results**) or
254 geographic region (**Supplementary Table 4**).

255 Analysis of 25 genomes from *P. falciparum* samples collected in Ethiopia in
256 2013 and 2015 and available in the MalariaGEN database (**Supplementary Figures**

257 **6-7)** uncovered chromosome 13 subtelomeric structural profiles similar to those
258 identified by MIP sequencing: 9 samples with coverage consistent with chr13-P3
259 (*pfhrp3*-deleted), 2 samples with chr13-P4 (*pfhrp3*-deleted), and 14 samples with
260 chr13-P1 (*pfhrp3*-intact).¹⁹

261

262 **Genetic signatures of evolutionary selection**

263 Extended haplotype homozygosity (EHH) statistics revealed signatures of
264 recent positive selection in the flanking region centromeric to *pfhrp2* deletions on
265 chromosome 8 but not in flanking regions around *pfhrp3* deletions on chromosome
266 13³⁵. 91 and 17 biallelic SNPs within the 28kb and 27kb regions centromeric to
267 *pfhrp2* and *pfhrp3* deletions, respectively, were used to calculate EHH statistics. 327
268 samples with *pfhrp2* deletion calls using MIP sequencing and sufficient variant data
269 were included in the EHH analysis, including 212 *pfhrp2*-intact and 115
270 *pfhrp2*-deleted haplotypes. EHH remained very high for parasites with the *pfhrp2*
271 deletion (0.968) along the entire 28kb analyzed, whereas homozygosity around the
272 *pfhrp2*-intact (wild-type) allele quickly broke down (**Figure 4A**). A similar pattern was
273 observed when deletion profiles were analyzed separately; chr8-P4 EHH remained
274 high and chr8-P1-P3 EHH quickly broke down (**Supplementary Figure 8**).

275 We further confirmed high EHH around the *pfhrp2* deletion allele using WGS
276 data. Comparing 23 whole-genome sequenced samples from this study and the 25
277 published MalariaGEN samples described above, we were able to extend our
278 analysis and confirm an EHH length of >143kb centromeric to the deletion
279 (**Supplementary Figure 9A**). These findings suggest a recent selective sweep,

280 indicative of strong evolutionary pressure favoring *pfhrp2*-deleted *P. falciparum*
281 parasites.

282 A different pattern was observed in the regions flanking *pfhrp3* (**Figure 4B**).
283 162 samples with *pfhrp3* deletion calls using MIP sequencing and sufficient variant
284 data were included in the EHH analysis, including 37 *pfhrp3*-intact and 125
285 *pfhrp3*-deleted haplotypes with three distinct subtelomeric structural profiles. No
286 variant calls were made in the 15.5 kb region immediately centromeric to *pfhrp3* to
287 avoid ambiguity in read mapping to the duplicated DNA segment containing
288 multicopy genes including 5.8S, 28S rRNA. EHH quickly decreased below 0.5 for
289 *pfhrp3* deletion alleles as well as the *pfhrp3*-intact allele within 1 kb of available
290 SNPs. When deletion profiles were analysed as separate alleles, the EHH pattern
291 was similarly low for chr13-P1, P3 and P4 (**Supplementary Figure 10**). Comparison
292 of EHH around the *pfhrp3*-intact and P3-like *pfhrp3* deletions using WGS data from
293 the 25 MalariaGEN samples confirmed our finding that the EHH quickly decreased
294 for both *pfhrp3*-intact and *pfhrp3*-deletion alleles (**Supplementary Figure 9B**). Taken
295 together, these findings suggest that each *pfhrp3* deletion profile arose multiple times
296 independently, and/or they have been present in the parasite population for sufficient
297 time for homozygosity due to genetic hitchhiking to be degraded by recombination
298 with different haplotypes.

299

300 **DISCUSSION**

301 Using the largest prospective study of *pfhrp2/3*-deleted *P. falciparum*
302 performed to-date and complementary molecular, immunological, and novel
303 sequencing assays, we provide clear evidence that *pfhrp2/3*-deleted parasites are

304 circulating in multiple sites along Ethiopia's borders with Sudan and Eritrea. Analysis
305 of flanking haplotypes suggests that the *pfhrp2* deletion mutation emerged and
306 recently expanded from a single origin, while *pfhrp3* deletion mutations have existed
307 for a longer time span and likely have multiple origins. As expected, we did not
308 observe perfect concordance between RDT results, PCR, a bead-based
309 immunoassay, WGS, and MIP sequencing results. However, the preponderance of
310 evidence from these diverse platforms provides robust confirmation of deletions and
311 supports the use of World Health Organization (WHO) protocols for rapid *pfhrp2/3*
312 deletion surveillance.²⁹ The prevalence of false-negative HRP2-based RDT results
313 due to *pfhrp2* deletions is estimated at 9.7% overall and up to 11.5% and 14.9% in
314 the Amhara and Tigray regions, respectively. These estimates exceed WHO
315 minimum criteria (>5%) for a change in national diagnostic testing strategy.
316 *Pfhrp2/3*-deleted parasites threaten recent progress made by Ethiopia's malaria
317 control and elimination program, and raise concerns about ongoing use of and
318 exclusive reliance on HRP2-based RDTs in the region for diagnosis of falciparum
319 malaria.

320 Eritrea's alarming reports of false-negative RDTs due to *pfhrp2/3*-deleted
321 parasites prompted an immediate change in national diagnostic testing policy in
322 2016^{15,36}. Recent evidence from Sudan, Djibouti, and Somalia suggests that the Horn
323 of Africa may already be heavily affected by *pfhrp2/3*-deleted parasites^{37,38}, though
324 results from ongoing surveillance efforts are not yet publicly available. Within
325 affected regions in Ethiopia, we observed spatial heterogeneity in *P. falciparum* RDT
326 profiles by district, with prevalence of the discordant HRP2-, PfLDH+ RDT profile
327 ranging from 0.9 to 30% (**Supplementary Table 2**). While finer scale spatial

328 analyses were not possible due to our health facility sampling approach, this finding
329 is consistent with prior studies showing variation within countries and by region²⁵.
330 Differences in transmission intensity, treatment-seeking behavior, diagnostic testing
331 capacity, and seasonality may account for some of the spatial variation in *pfhrp2/3*
332 deletion prevalence estimates^{17,39-41}. Although the factors driving emergence of these
333 parasites in some regions but not others remain poorly understood, our study
334 suggests that *pfhrp2*-deleted parasites may have spread widely within Ethiopia from
335 a single origin. This finding is consistent with early microsatellite analysis of
336 *pfhrp2/3*-deleted strains in Eritrea, in which 30 of 31 (96.8%) *pfhrp2*-deleted strains
337 fell into a single genetically related cluster,¹⁵ and raises concern about clonal
338 expansion of *pfhrp2*-deleted strains in the Horn of Africa.

339 Using a multi-faceted approach, we validate the use of MIP sequencing for
340 high-throughput *pfhrp2/3* deletion genotyping, deletion profiling, and population
341 genetic analysis. Comparison of MIP sequencing to other approaches demonstrated
342 that it can be used for cost-effective (approximately \$10-15 per sample) and scalable
343 deletion genotyping in samples with parasite densities of approximately 1,000
344 parasites/ μ L. While this threshold can likely be improved by additional sequencing of
345 samples with inadequate sequencing depth-of-coverage, in this case, the equivalent
346 of half a NextSeq 550 flow cell enabled visualization of deletion breakpoint regions
347 and variant calling in *P. falciparum*'s subtelomeres in a large portion of samples,
348 without the need for costly enrichment and WGS.

349 Based on analysis of MIP sequencing and available WGS data, we posit one
350 potential model by which *pfhrp2/3*-deleted parasite populations may have evolved in
351 the Horn of Africa. Findings from this study suggest that parasite populations with

352 *pfhrp3* deletions expanded in the more distant past and potentially arose multiple
353 times independently, based on low EHH surrounding *pfhrp3*, multiple deletion profile
354 patterns, the high overall frequency of *pfhrp3*-deleted parasites, and their presence
355 in older samples from 2013 in the MalariaGEN study. In this milieu, recent strong
356 selection favoring parasites with deletions of *pfhrp2* likely occurred due to
357 “test-track-treat” policies that rely upon HRP2-based RDTs and allow parasites with
358 deletions of both genes, or in some cases one of the two genes, to escape
359 treatment. Implicit in this model is the assumption that forces apart from RDT-derived
360 pressure are also driving the evolution of *pfhrp2/3* deletions. First, test-track-treat
361 policies do not explain the prevalence and genetic evidence of well-established
362 *pfhrp3* deletions in Ethiopia, as malaria due to *pfhrp2+/3-* parasites should be
363 detectable by HRP2-based RDTs⁴². Second, *pfhrp2/3*-deleted parasites are highly
364 prevalent in South America,²⁵ where RDT-based treatment decisions have never
365 been common. Third, the prevalence of *pfhrp2/3*-deleted parasites appears to have
366 remained stable in Eritrea despite removal of HRP2-based RDTs two years ago.³⁸

367 What other advantages might *pfhrp2/3*-deleted parasites have over those with
368 intact genes? Our limited understanding of the biology of these deletions makes this
369 question hard to answer. Several lines of inquiry may be relevant: 1) They may be
370 better adapted to low transmission intensity settings than other strains.

371 *Pfhrp2/3*-deleted parasites appear to be more common in regions with lower
372 transmission and, presumably, lower complexities of infection.⁴⁰ This trend might
373 simply be an artifact of the assays used to detect them - i.e., neither PCR, antigen
374 immunoassays, nor common sequencing methodologies are well suited to detect a
375 *pfhrp2/3*-deleted strain when *pfhrp2/3*-intact strains have co-infected a human host.

376 The high frequency of monoclonal samples in our MIP sequence analysis provides
377 support for this hypothesis. However, it is also possible that the *pfhrp2/3* genes are
378 an asset when within-host competition is common but a liability when it is rare. This
379 makes sense if parasites in low-transmission settings gain an advantage through
380 improved gametocytogenesis and transmission, for example. 2) Loss of *pfhrp2/3* or
381 flanking genes may alter parasite virulence. Evidence is accumulating that HRP2
382 plays a role in cerebral malaria and endothelial inflammation during severe
383 malaria^{43,44}. People infected by *pfhrp2/3*-deleted parasites may have less severe
384 disease and therefore be less likely to seek treatment, increasing the likelihood of
385 onward transmission. However, we cannot exclude the possibility that *pfhrp2/3* are
386 lost as a consequence of selection on other genes. For example, the flanking gene
387 *EBL-1* is almost uniformly lost in *pfhrp3*-deleted parasites in this cohort and appears
388 to play a role during invasion of red blood cells^{34,45}. Similarly, members of the *rifin*
389 and *stevor* gene families with potential roles in parasite virulence were lost in the
390 subtelomeric deletions observed in this study^{46,47}. We did not observe evidence of an
391 association between virulence and subtelomeric deletions in our cohort, but limited
392 clinical data prevents us from assessing the hypothesis rigorously. 3) Loss of
393 *pfhrp2/3* or flanking genes may improve transmissibility to or from mosquitoes. To
394 our knowledge, this phenomenon has not been studied. These and other hypotheses
395 require experimental and improved epidemiological analyses. Regardless of the
396 evolutionary forces at play, our findings strongly suggest that the evolution of
397 *pfhrp2/3*-deleted parasites in Ethiopia was a multi-step process that involved earlier
398 expansion of *pfhrp3*- than *pfhrp2*-deleted parasite populations.

399 This study has several limitations. First, the study design prioritized evaluation

400 of samples with discordant RDT results (HRP2- but Pf-LDH+) for rapid assessment
401 of false-negative RDTs due to *pfhrp2/3* deletions in the context of clinical treatment.
402 This feature of the WHO protocol is intentional as it captures clinically significant
403 *pfhrp2/3* deletions and enables real-time, efficient signaling to malaria control
404 programs of a potential problem, but it also introduces selection bias that requires
405 careful consideration when estimating the true prevalence of *pfhrp2/3*-deleted
406 parasites. We overcame this limitation by using a conservative approach that
407 incorporated both RDT and PCR data to estimate false-negative RDT results due to
408 *pfhrp2* deletions. This metric is relevant to control programs, but does not capture
409 asymptomatic or low-parasite-density infections by *pfhrp2/3*-deleted parasites.
410 Second, only a subset of samples underwent advanced analysis, and clinical data
411 was not available for all subjects. This was not unexpected for a pragmatic field
412 study of this size. We do not believe that it introduced sufficient bias into our
413 prevalence estimates or population genetic analysis to change our conclusions.
414 Third, we cannot comment on changes in selection pressure over time because the
415 study was cross-sectional. Fourth, we only sampled three regions of Ethiopia, which
416 is a diverse and populous country. In response, the Federal Ministry of Health is now
417 conducting a country-wide survey that will enable comparison of *pfhrp2/3* deletions
418 over time in select sites.

419

420 **CONCLUSION**

421 Leveraging a large prospective study, established molecular and antigen
422 detection methods, and a novel targeted sequencing approach, we demonstrate that
423 *pfhrp2/3*-deleted *P. falciparum* is a common cause of false-negative RDT results

424 among subjects presenting with symptomatic malaria in three regions of Ethiopia.
425 The genomic tools employed in this study reveal complex origins of these parasites.
426 Recent, strong selective pressures favoring *pfhrp2*-deleted parasites appear to have
427 occurred on a background of pre-existing *pfhrp3* deletions. Existing malaria control
428 programs in the region are threatened by expansion of these parasite strains, and
429 surveillance is urgently needed to inform decisions about when alternative malaria
430 diagnostics should be deployed.

431

432 **METHODS**

433 ***Study Design and Data Collection***

434 We performed a cross-sectional, multi-site study in eleven districts along
435 Ethiopia's borders with Eritrea, Sudan, and South Sudan, located within three of its
436 nine administrative regions. On average, ten health facilities were selected from
437 each district, including four districts of Amhara Region (northwest Ethiopia), six
438 districts of Tigray Region (north Ethiopia), and one district of Gambella Region
439 (southwest Ethiopia) during the 2017-2018 peak malaria transmission season
440 (September-December) (see **Figure 1**). Per WHO protocol²⁹, each facility passively
441 enrolled participants presenting with symptoms of malaria (fever, headache, joint
442 pain, feeling cold, nausea, and/or poor appetite), with sample size proportionally
443 allocated to each facility based on the previous year's malaria case load. All
444 participants provided informed consent, participated in interview questionnaires, and
445 underwent blood collection for RDT testing using two types of RDTs. Ethical approval
446 was obtained from the Ethiopia Public Health Institute (EPHI) Institutional Review
447 Board (IRB; protocol EPHI-IRB-033-2017) and WHO Research Ethics Review

448 Committee (protocol: ERC.0003174 001). Processing of de-identified samples and
449 data at UNC was determined to constitute non-human subjects research by the UNC
450 IRB (study 17-0155). The study was determined to be non-research by the Centers
451 for Disease Control and Prevention Human Subjects office (0900f3eb81bb60b9).
452 Experiments were performed in accordance with relevant guidelines and regulations.
453

454 **Field Sample Evaluation**

455 Study participants were evaluated using both a CareStart *Pf/Pv*
456 (HRP2/Pv-pLDH) RDT (Access Bio, Somerset, NJ; product code RM VM-02571) and
457 an SD Bioline Malaria Ag P.f. (HRP2/Pf-LDH) RDT (Alere, Waltham, MA; product
458 Ocode 05FK90). For the CareStart RDT, 5 μ L of capillary whole blood was collected
459 by finger prick and transferred to the RDT sample well, along with 60 μ L of buffer
460 solution. Results were read at 20 minutes. The SD Bioline RDT followed the same
461 protocol, but with 4 drops of buffer added and results read in a 15-30 minute window.
462 Participants testing positive by either RDT were first prescribed treatment, according
463 to Ethiopian national guidelines.⁴⁸

464 Cases with any positive HRP2 or Pf-LDH RDT band were considered positive
465 for *P. falciparum* malaria. Cases Pf-LDH-positive but HRP2-negative on both RDTs
466 were considered potential candidates for *pfhrp2/3* gene deletion and defined as
467 'discordant.' These participants, along with a subset of HRP2-positive and -negative
468 controls, provided further informed consent for additional blood collection for dried
469 blood spot (DBS) preparation. At least two DBS samples (50 μ L/spot) were collected
470 on Whatmann 903 protein saver cards (GE Healthcare, Chicago, IL) from consenting
471 participants. DBS were stored in plastic bags with desiccant. A randomly selected

472 subset of DBS were sent for molecular analysis to the University of North Carolina at
473 Chapel Hill and for serological analysis to the Centers for Disease Control and
474 Prevention (CDC, Atlanta, GA).

475

476 **DNA Extraction and PCR assays**

477 DNA was extracted from three 6mm punches per DBS sample using
478 Chelex-100 and saponin as previously described⁴⁹. Quantitative PCR (qPCR) assays
479 were first performed in duplicate for *pfl dh*⁵⁰. To avoid the risk of misclassification due
480 to DNA concentrations below the limit of detection for *pfhrp2/3* PCR assays, further
481 analysis was restricted to samples with >100 parasites/ μ L by qPCR
482 (**Supplementary Figure 2**).¹⁴ PCR assays targeting exon 2 of *pfhrp2* and *pfhrp3*
483 were then performed in duplicate as previously described,⁷ except that PCRs were
484 performed as single-step, 45-cycle assays, using 10 μ L template and AmpliTaq Gold
485 360 Master Mix (Thermo Fisher Scientific, Waltham, MA) in 25 μ L reaction volume.
486 In addition to no-template and *P. falciparum* 3D7 strain (*pfhrp2*+/*3*+) positive controls,
487 *pfhrp2* assays included an additional DD2 strain (*pfhrp2*-/*3*+) control and *pfhrp3*
488 assays included an additional HB3 strain (*pfhrp2*+/*3*-) control. Finally, an additional
489 single-copy gene, real-time PCR assay targeting *P. falciparum* beta-tubulin was
490 performed to confirm that sufficient parasite DNA remained in samples with a
491 negative *pfhrp2/3* PCR result.¹⁴ *Pfhrp2/3* genotyping calls were made in samples
492 with *pfl dh* qPCR parasitemia >100 parasites/ μ L to avoid misclassification in the
493 setting of amplification failure due to low target DNA concentration. A *pfhrp2* or
494 *pfhrp3* positive call required \geq 1 replicate with distinct band(s) with expected fragment
495 length. A negative call required both *pfhrp2* or *pfhrp3* replicates to be negative.

496 Detailed reaction conditions for all PCR assays are described in the **Supplementary**
497 **File**.

498

499 **Serological assays**

500 The presence of HRP2, pan-LDH, and aldolase antigenemia was assessed in
501 a subset of DBS samples (single 6mm punch) using a multiplex bead-based
502 immunoassay as previously described¹³. Within this multiplex assay, capture and
503 detection antibodies against the HRP2 antigen would also recognize similar epitopes
504 on the HRP3 antigen, so unique signals for these two antigens cannot be obtained.

505

506 **Prevalence estimates**

507 We estimated the prevalence of *P. falciparum* infections expected to have
508 false-negative HRP2-based RDT results due to *pfhrp2* deletions as follows. First, we
509 calculated the proportion of all RDT-positive *P. falciparum* cases (HRP2+ or PfLDH+
510 on any RDT) with the discordant RDT profile (HRP2- on both RDTs, but PfLDH+),
511 overall and by region. Second, we calculated the observed concordance between
512 the discordant RDT profile and a *pfhrp2*-negative PCR call, overall. Prevalence
513 estimates and 95% CIs were then back-transformed overall and by region using the
514 `ci.impt` function within the *asbio* R package, which generates CIs for the product of
515 two proportions using delta derivation. This allowed us to estimate with confidence
516 the proportion of *P. falciparum* infections with both *pfhrp2* deletions and
517 false-negative HRP2-based RDT results, overall and by region. As a sensitivity
518 analysis, we also estimated the proportion of those with a discordant RDT and a
519 *pfhrp2*-negative PCR call (directly multiplying the true proportion of Pf-positive

520 individuals with a discordant RDT profile, overall and by region, by 0.727, or the
521 overall proportion of discordant RDT samples that had a *pfhrp2*- PCR result). 95%
522 CIs were then generated using bootstrapping (1000 iterations). The prevalence
523 estimates and CIs generated by the two approaches were similar (**Supplementary**
524 **Table 3**).

525

526 ***Pfhrp2/3 molecular inversion probe (MIP) development***

527 *Pfhrp2*, *pfhrp3*, and the flanking regions within a 100kb window surrounding
528 each gene were targeted for MIP designs using *MIPTools*⁵¹. A tiled design strategy
529 was employed that involved multiple, overlapping probes spanning each gene target.
530 22 genes flanking *pfhrp2* and 31 genes flanking *pfhrp3* were used in the design, of
531 which 11 and 19 were successful on the first design try, respectively. A second
532 attempt was not made for designs for the flanking genes. A total of 241 probes; 9 for
533 *pfhrp2*, 9 for *pfhrp3* and 223 probes for the flanking genes were designed. MIPs
534 were designed using the 3D7 (v3) reference genome avoiding hybridization arms in
535 variant regions when possible. 80 alternative probes accommodating potential
536 variants in the highly variable *pfhrp2* and *pfhrp3* genes were also created. A 15.5 kb
537 segment centromeric to *pfhrp3* on chromosome 13 between positions
538 2792000-2807500 is duplicated on chromosome 11 between positions
539 1918007-1933488, with 99.4% sequence identity. Therefore, the target genes falling
540 into this region were multicopy genes and their probes were designed to bind to both
541 loci on the genome (see **Supplementary Table 5** for the design overview including
542 all genes targeted, MIPs designed and genomic coordinates). Probes were ordered
543 from Integrated DNA Technologies (Coralville, Indiana, USA) as 200 pmol ultramer

544 oligos. Probe sequences are provided in the **Supplementary Table 6**.

545

546 ***MIP capture and deep sequencing of clinical samples***

547 All DNA samples extracted by UNC underwent MIP capture using the capture
548 and amplification methods exactly as described by Verity *et al.*⁵², with the exception
549 of oligonucleotides (the *pfhrp2/3* MIP oligonucleotide panel described above was
550 used) and controls (we selected a different set of controls that are informative for
551 *pfhrp2/3* deletion characterization). All MIP captures included multiple controls: 3D7
552 (*pfhrp2+/3+*), DD2 (*pfhrp2-/3+*), HB3 (*pfhrp2+/3-*) laboratory strains; as well as LC
553 and HC mixes (1% HB3, 10% DD2, 89% 3D7) at 250 and 1000 parasites/ μ l
554 densities, respectively. Samples were sequenced on the Illumina NextSeq 550
555 instrument using 150bp paired-end sequencing and dual indexing.

556

557 ***Subtelomeric profiling and variant calling with MIP data***

558 Read mapping and variant calling were carried out using *MIPTools*
559 (v0.19.12.13)⁵¹. *MIPTools* uses the *MIPWrangler* algorithm (v1.2.0)⁵³ to create high
560 quality consensus sequences from sequence read data utilizing unique molecular
561 indexes (UMIs) of MIPs, maps those sequences to the reference genome using *bwa*
562 (v0.7.17) and remove off target sequences as described previously^{31,52}. Deletion calls
563 were limited to the samples that had high coverage to avoid false positives.
564 Considering the high frequency of large deletions present in the sample set, the
565 coverage threshold was based on a subset of probes that were present on > 60% of
566 the samples, none of which overlapped with the chromosome 8 or 13 deletions.
567 Samples with a median coverage of < 5 UMIs for this subset of probes were

568 excluded from analysis.

569 Structural profiling was performed using the UMI count table (**Supplementary**
570 **Table 7**). The count table was converted to a presence/absence table such that if a
571 probe had > 1 UMI for a given sample, it was accepted as present (i.e., not deleted).
572 Samples were clustered into subtelomeric structural profile groups based on this
573 table using the hierarchical clustering algorithm AgglomerativeClustering of the
574 Python module Scikit-learn (v0.20)⁵⁴ using only the regions involved in the deletion
575 events of the corresponding chromosome (position > 1372615 for chromosome 8
576 and position > 2806319 for chromosome 13). Samples were grouped into their final
577 subtelomeric structural profile based on visual inspection of the resulting clusters.

578 Initial variant calls were made using *freebayes* (v1.3.1) via *MIPTools* with the
579 following options: --pooled-continuous --min-base-quality 1 --min-alternate-fraction
580 0.01 --min-alternate-count 2 --haplotype-length -1 --min-alternate-total 10
581 --use-best-n-alleles 70 --genotype-qualities. Variants were processed using *MIPTools*
582 to filter for: variant quality > 1, genotype quality > 1, average alternate allele quality >
583 15, minimum depth > 2 UMIs; and make final genotype calls based on the major
584 allele (within sample allele frequency > 0.5). In addition, the following variants were
585 removed from the final call set: those that were observed as a major allele in less
586 than two samples (singletons), not supported by more than two UMIs in at least three
587 samples, present on multicopy genes, and indels. Variant calls were further filtered
588 for missingness to avoid imputation in EHH calculations: samples missing calls for
589 >50% of the variants were removed, variants missing calls in >50% of the samples
590 were removed. Variants calls were converted to .map and .hap files (**Supplementary**
591 **Table 8**) for use with the *rehh* package in R.

592

593 ***Assessment of MIP calls using whole-genome sequencing***

594 We performed WGS on a subset of samples to assess the accuracy of MIP
595 *pfhrp2/3* deletion calls. DNA extracted from samples with discordant RDT results
596 were selected for *P. falciparum* selective whole-genome amplification (sWGA) and
597 whole-genome sequencing as described previously³⁰. In brief, DNA was first
598 subjected to two separate sWGA reactions using the Probe_10 primer set described
599 by Oyola et al.⁵⁵ and the JP9 primer set.³⁰ sWGA products were then pooled in equal
600 volumes and acoustically sheared using a Covaris E220 instrument prior to
601 sequencing library preparation using Kappa Hyper library preps. Indexed libraries
602 were then pooled and sequenced on an Illumina HiSeq 4000 instrument using
603 150bp, paired-end sequencing. Sequencing reads were deposited into NCBI's
604 Sequence Read Archive (accession numbers pending).

605

606 ***Published whole genome sequencing data retrieval***

607 Fastq files from 25 Ethiopian samples included in the MalariaGEN genome
608 variation project¹⁹ and 3 laboratory strains (3D7, HB3 and DD2) from MalariaGEN
609 genetic crosses project⁵⁶ were downloaded from the European Nucleotide Archive
610 using fasterq-dump (v2.10.8) and sample accession numbers (**Supplementary**
611 **Table 9**).

612

613 ***WGS data analysis***

614 All fastq files were processed as follows. Adapter and quality trimming was
615 performed using *Trimmomatic* (v.0.39) with the recommended options (seed

616 mismatches:2, palindrome clip threshold:30, simple clip threshold:10,
617 minAdapterLength:2, keepBothReads LEADING:3 TRAILING:3 MINLEN:36).
618 Trimmed fastq files were mapped to 3D7 reference genome (v3.0) concatenated to
619 human genome (hg38) to avoid incorrect mapping of reads originating from host
620 DNA using *bowtie2* (v2.3.0) with the '--very-sensitive' option. Reads mapping to the
621 parasite chromosomes were selected and optical duplicates were removed using the
622 *sambamba* (v0.7.1) view and markdup commands, respectively. Read coverage was
623 calculated using *samtools* (v1.9) depth command with options '-a -Q1 -d0', filtering
624 reads with mapping quality of zero. Variants were called only for the regions of
625 interest using *freebayes* (v1.3.1) with the following options: '--use-best-n-alleles 70
626 --pooled-continuous --min-alternate-fraction 0.01 --min-alternate-count 2
627 --min-alternate-total 10 --genotype-qualities --haplotype-length -1
628 --min-mapping-quality 15 -r region'. Regions of interest were from 300 kb
629 centromeric to the deletions to chromosome ends (positions 1074000-1472805 and
630 2505000-2925236 for chromosomes 8 and 13, respectively).

631 Variants were filtered for: variant quality > 20, genotype quality > 15, average
632 alternate allele quality > 15, minimum depth > 4 reads. In addition, the following
633 variants were removed from the final call set: those that were never observed as a
634 major allele in any sample, not supported by more than 10 reads in at least one
635 sample, and indels. Final genotype calls were based on the major allele (within
636 sample allele frequency > 0.5). Variant calls were further filtered for missingness to
637 avoid imputation in EHH calculations: samples missing calls for >95% of the variants
638 were removed, variants missing calls in >10% of the samples were removed.
639 Telomeric profiling of the published genomes was carried out by visual inspection of

640 depth-of-coverage plots (**Supplementary Figures 8 and 9**). Summary statistics were
641 generated (**Supplementary Table 10**) using the python pandas module (v0.23).

642

643 ***Statistical and population genetic analysis***

644 Data collected during the participant's study visit (clinical data and RDT
645 results) was linked to laboratory results via the barcode number transcribed on DBS
646 sent to the UNC and CDC laboratories. Samples in the dataset with missing or
647 duplicate barcodes were arbitrated using original paper questionnaires by the EPHI
648 data center. Ultimately, an analysis dataset that included both PCR and field data
649 was created including all samples we could confidently merge by both barcode
650 number and region label.

651 Statistical analysis was performed using R (version 3.6.0, R Core Team,
652 Vienna, Austria, 2019; www.R-project.org). ArcGIS (Desktop Version 10.5, ESRI,
653 Redlands, CA, 2016) was utilized for mapping, with additional annotation performed
654 using PowerPoint (version 16.31, Microsoft, Redmond, WA, 2019).

655 Extended haplotype homozygosity (EHH) statistics were calculated to
656 evaluate the regions flanking the *pfhrp2* and *pfhrp3* genes for signatures of recent
657 positive selection³⁵ using the *rehh* package (version 3.1.2)⁵⁷. EHH statistics were
658 calculated using the *data2haplohh* and *calc_ehh* functions, haplotype furcations
659 were calculated using *calc_furcation*, and plots were generated using the package's
660 plot function and annotated using Inkscape (version 0.92).

661 Complexity of infection (COI) for each sample was calculated using *McCOILR*
662 (v1.3.0, <https://github.com/OJWatson/McCOILR>), an *Rcpp* wrapper for *THE REAL*
663 *McCOIL*⁵⁸ with the options *maxCOI=25*, *totalrun=2000*, *burnin=500*, *M0=15*,

664 err_method=3. The same variant set used in the EHH analysis was used for the COI
665 calculations, except that variants whose within-sample allele frequency were
666 between 0.05 and 0.95 were called heterozygote for COI analysis.

667 **ACKNOWLEDGMENTS**

668 The authors thank the research teams for conducting field work and the subjects for
669 participating in the study. They would also like to thank Steven R. Meshnick
670 posthumously for contributions to the laboratory analyses and data analysis. The
671 following reagents were obtained through BEI Resources, NIAID, NIH: Genomic
672 DNA from *P. falciparum* strain 3D7, MRA-102G, contributed by Daniel J. Carucci; *P.*
673 *falciparum* strain HB3, MRA-155G, contributed by Thomas E. Wellems; *P. falciparum*
674 strain Dd2, MRA-150G, contributed by David Walliker. The authors also would like to
675 acknowledge MSF Holland for supporting the field study in Gambella region.

676

677 *Author contributions*

678 SMF, JAC, and JBP conceived the study. SMF, HM, BGB, HM, HS, BP, EA
679 supervised and/or conducted field work. OA, CH, MD, ER performed laboratory
680 assays and experiments. ENR, OA, CK, JJJ, JAB, ER, JBP analyzed laboratory
681 data. ENR and OA produced the tables and figures. SMF and ENR wrote the first
682 draft with assistance from OA and JBP. All authors critically reviewed and approved
683 the final manuscript.

684

685 *Disclaimer*

686 The findings and conclusions in this report are those of the authors and do not
687 necessarily represent the official position of the CDC.

688

689 *Data availability*

690 Genomic sequencing data will be available through the Sequence Read Archive
691 (BioSample accession numbers pending). De-identified datasets generated during
692 the current study will be available as supplementary files. Code used during data
693 analysis will be made available on GitHub.

694

695 *Funding*

696 This work was funded by the Global Fund to Fight AIDS, Tuberculosis, and Malaria
697 through the Ministry of Health-Ethiopia (EPHI5405 to SMF) and the World Health
698 Organization (JAC, JBP). It was also partially supported by MSF Holland, which
699 supported field work in Gambella Region, the Doris Duke Charitable Foundation
700 (JBP), and the US National Institutes of Health (R01AI132547 to JJJ, JAB, OA, and
701 JBP; K24AI134990 to JJJ).

702

703 *Competing interests*

704 JBP reports research support from Gilead Sciences, honoraria from Virology
705 Education for medical education teaching, and non-financial support from Abbott
706 Diagnostics, all outside the scope of the current work. SMF reports research support
707 from AccessBio, outside the scope of the current work.

708 **TABLES**

709

710 **Table 1. Characteristics of study subjects and RDT results.** Abbreviations: SD,
711 standard deviation; RDT, rapid diagnostic test; HRP2, histidine-rich protein 2;
712 Pv-LDH, *P. vivax* parasite lactate dehydrogenase; Pf-LDH, *P. falciparum* parasite
713 lactate dehydrogenase.

714

	Amhara	Gambella	Tigray	Overall
Subjects, n	3,879	2,335	6,357	12,572
Age, median years (IQR)	20 (10-28)	12 (5-19)	21 (9-37)	19 (8-30)
Female sex, n (%)	1,492 (38.5)	1,055 (45.2)	3,008 (47.3)	5,555 (44.2)
Location, n (%)				
Rural	2,350 (60.6)	923 (39.5)	4,445 (69.9)	7,718 (61.4)
Urban	1,282 (33.0)	82 (3.5)	1,609 (25.3)	2,973 (23.6)
Missing	247 (6.4)	1,330 (57.0)	303 (4.8)	1,881 (15.0)
Fever, n (%)	3,607 (93.0)	2,255 (96.6)	5,593 (88.0)	11,455 (91.1)
CareStart RDT, n (%)				
HRP2+, Pv-LDH+	59 (1.5)	509 (21.8)	25 (0.4)	593 (4.7)
HRP2+ only	1,053 (27.1)	165 (7.1)	507 (8.0)	1,725 (13.7)
Pv-LDH+ only	241 (6.2)	11 (0.5)	338 (5.3)	590 (4.7)
Negative	2,518 (64.9)	1,650 (70.7)	5,486 (86.3)	9,654 (76.8)
Invalid	8 (0.2)	0 (0.0)	1 (0.0)	9 (0.1)
SD-Bioline RDT, n (%)				
HRP2+, Pf-LDH+	719 (18.5)	552 (23.6)	276 (4.3)	1,547 (12.3)
HRP2+ only	297 (7.7)	106 (4.5)	201 (3.2)	604 (4.8)
Pf-LDH+ only	239 (6.2)	11 (0.5)	168 (2.6)	418 (3.3)
Negative	2,609 (67.3)	1,665 (71.3)	5,705 (89.7)	9,979 (79.4)
Invalid	15 (0.4)	0 (0.0)	2 (0.0)	17 (0.1)

715

716 **Table 2. Assay results across platforms.** PCR, bead-based antigen immunoassay,
717 molecular inversion probe (MIP) deep sequencing, and whole-genome sequencing
718 (WGS) for samples *P. falciparum*-positive by RDT are shown.
719

	RDT results		
	2 HRP2+, 1 PflLDH+	2 HRP2-, 1 PflLDH+	1 HRP2+, 1 HRP2-
PCR result, n	379	176	47
<i>pfhrp2+/3+</i>	210 (55%)	20 (11%)	19 (41%)
<i>pfhrp2-/3-</i>	9 (2%)	114 (65%)	10 (22%)
<i>pfhrp2-/3+</i>	7 (2%)	14 (8%)	5 (11%)
<i>pfhrp2+/3-</i>	152 (40%)	28 (16%)	12 (26%)
HRP2 immunoassay result, n	243	167	42
HRP2+	224 (92%)	40 (24%)	30 (71%)
HRP2-	19 (8%)	127 (76%)	12 (29%)
MIP result, n	198	104	30
<i>pfhrp2+/3+</i>	77 (39%)	5 (5%)	5 (17%)
<i>pfhrp2-/3-</i>	10 (5%)	84 (81%)	10 (33%)
<i>pfhrp2-/3+</i>	3 (2%)	1 (1%)	4 (13%)
<i>pfhrp2+/3-</i>	108 (55%)	14 (13%)	11 (37%)
WGS result, n*	0	22	0
<i>pfhrp2+/3+</i>	NA	0 (0%)	NA
<i>pfhrp2-/3-</i>	NA	22 (100%)	NA
<i>pfhrp2-/3+</i>	NA	0 (0%)	NA
<i>pfhrp2+/3-</i>	NA	0 (0%)	NA

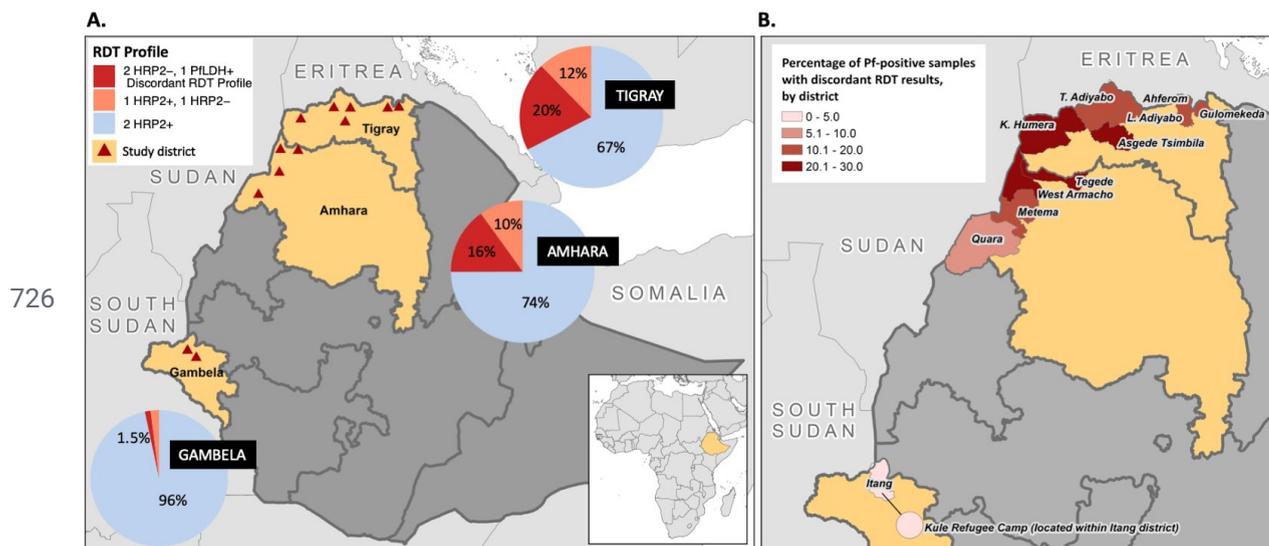
720
721
722

* Zero median WGS coverage across 1,000 base-pair windows encompassing *pfhrp2* or *pfhrp3*, in samples with clinical data.

723 **FIGURES**

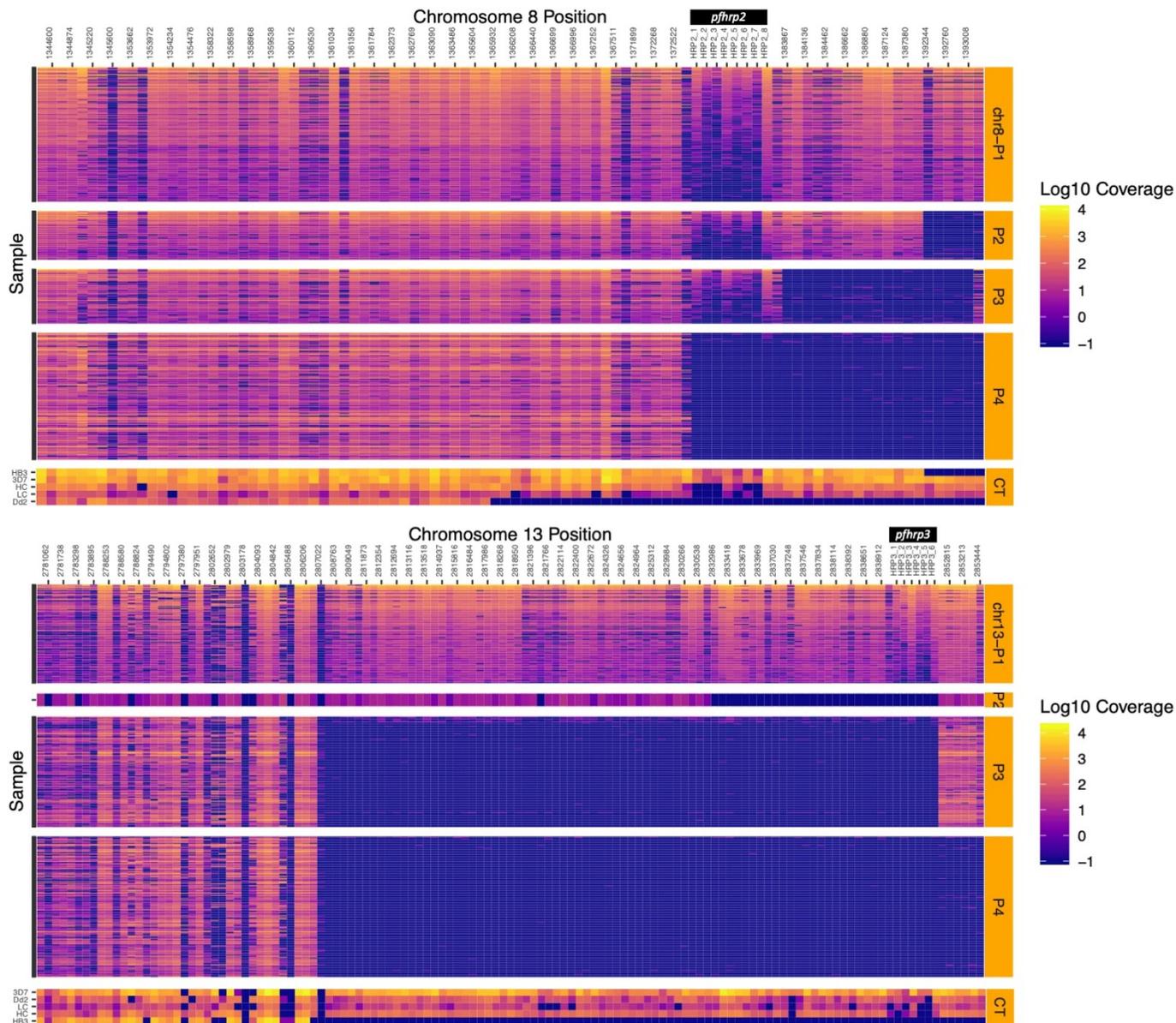
724

725

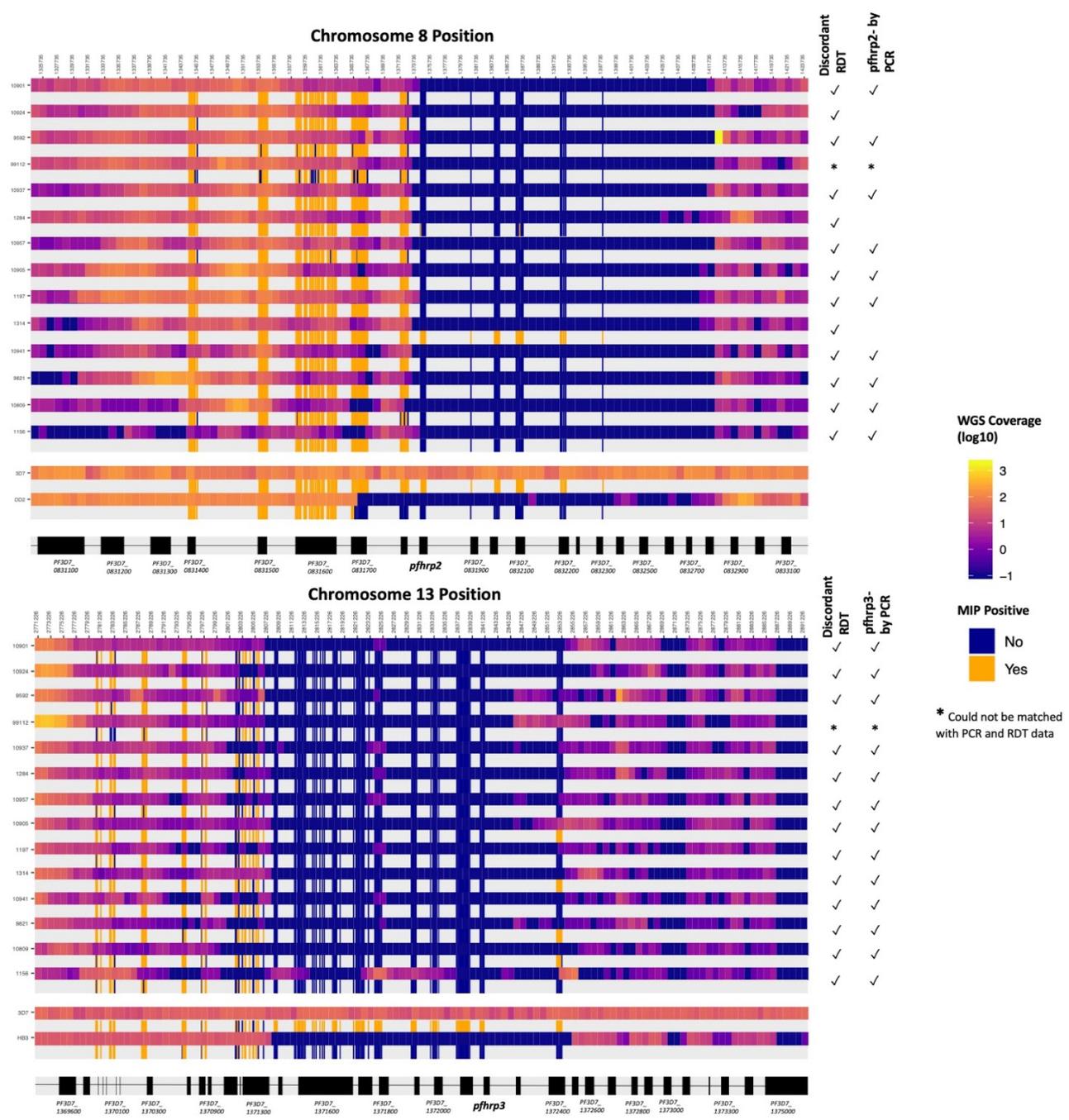


727 **Figure 1. Distribution of *P. falciparum*-positive RDT results and discordant**
728 **profiles suggestive of *pfhrp2/3* gene deletions.** A) Aggregated results from both
729 RDTs, CareStart *Pf/Pv* (HRP2/Pv-LDH) RDT and SD Bioline Malaria Ag P.f.
730 (HRP2/Pf-LDH) RDT, displayed by region for all *P. falciparum* infections. The '2
731 HRP2-, 1 Pf-LDH+' discordant RDT profile indicates potential infection by
732 *pfhrp2/3*-deleted *P. falciparum*. Triangles represent the enrollment sites, including 11
733 districts and the Kule refugee camp within the Itang district in Gambella. B) The
734 percentage of study participants identified with *P. falciparum* infection by RDT who
735 had the discordant RDT profile, by district.

736



737 **Figure 2. Deletion profiling using MIP sequencing of *pfhrp2* (chromosome 8),**
 738 ***pfhrp3* (chromosome 13), and flanking regions applied to 375 field samples.**
 739 Samples are grouped by subteleric structural profile, with control strains denoted
 740 CT, as labeled along the right y-axis. Columns represent each MIP target segment,
 741 rows represent individual samples, and the color scale represents \log_{10} unique
 742 molecular identifier (UMI) depth-of-coverage at each location. Columns are labeled
 743 by the midpoint of each probe's target region.

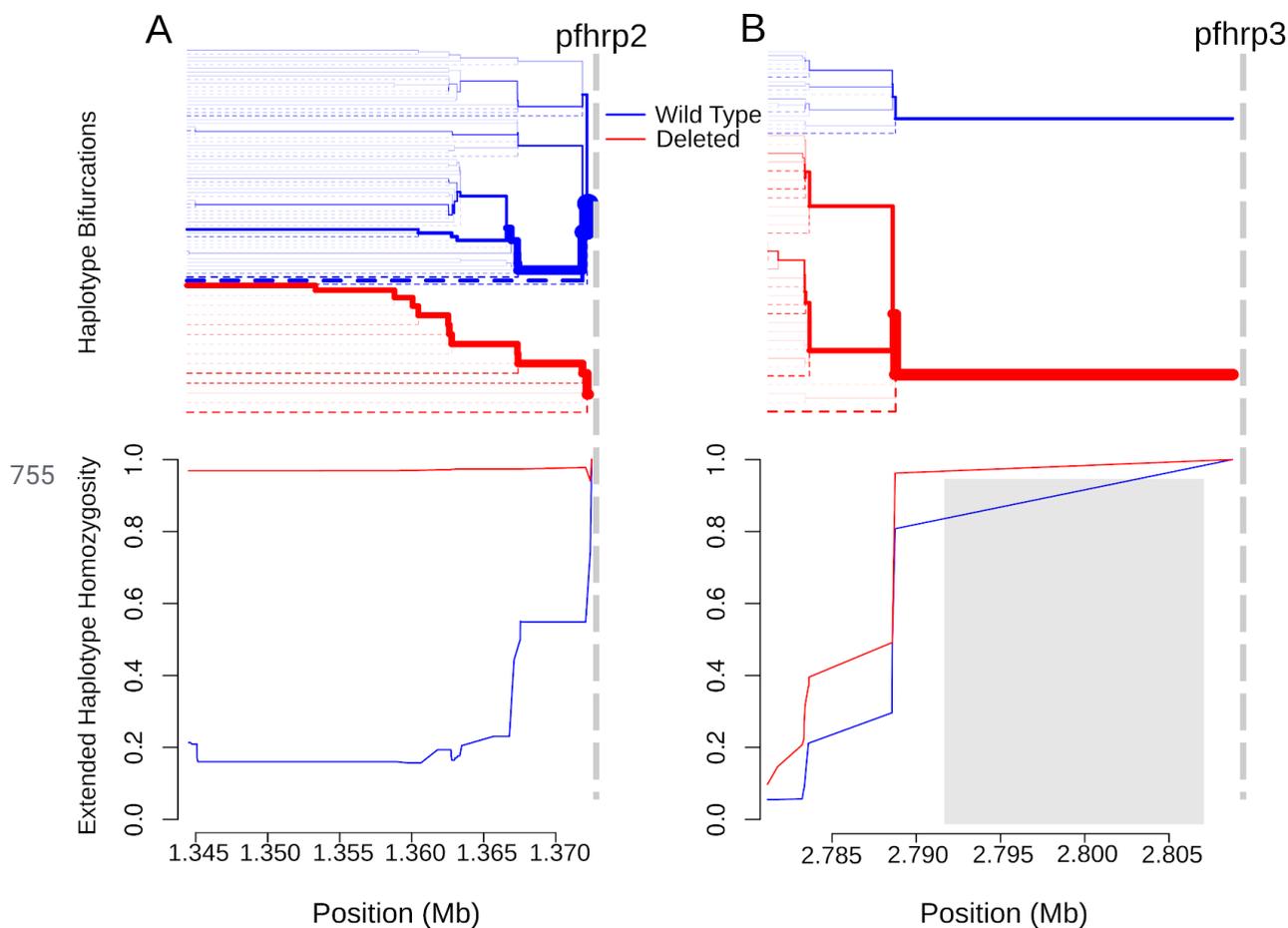


744

745 **Figure 3. Comparison of MIP and WGS *pfhrrp2/3* deletion calls and breakpoint**
 746 **regions.** Among the 14 clinical samples subjected to both methods, each sample is
 747 represented by two adjacent rows representing WGS (top) and MIP (bottom)
 748 coverage results. WGS coverage is displayed as the log₁₀ median number of aligned
 749 reads per 1kb window. MIP results are colored by whether each probe captured its
 750 target, with intervening regions not targeted in the MIP panel uncolored. Sample
 751 numbers (lab_ID) are provided at left. The locations of *pfhrrp2*, *pfhrrp3*, and flanking
 752 genes are shown in black with non-genic regions in gray.

753

754



756

757 **Figure 4. Extended haplotype homozygosity (bottom) and the bifurcation**
758 **diagrams showing haplotype branching (top) centromeric to the *pfhrp2* (A) and**
759 ***pfhrp3* (B) deletions based on MIP data. Vertical dashed lines indicate the**
760 **centromeric end of deletions. No variant calls were made within the 15.5 kb region**
761 **on chromosome 13 which is duplicated on chromosome 11, demarcated by the gray**
762 **box (B). Abbreviations: Mb, mega-base.**

763

764

765 **SUPPLEMENTARY MATERIAL**

766

767 **Supplementary Results**

768

769 *Association between malaria symptoms, geographical location, and subtelomeric*
770 *structural variants*

771 We did not observe an association between subtelomeric deletion profile and
772 the number of symptoms experienced by subjects (**Supplementary Figure 5**).

773 Because the majority of subjects (96.5%, 2620/2714) who tested positive for *P.*
774 *falciparum* by RDT were febrile, fever alone was not sufficient to evaluate disease
775 severity. Therefore, as a crude metric of disease severity, we calculated the total
776 number of symptoms (six total were assessed: fever, headache, joint pain, feeling
777 cold, nausea, and lack of appetite). No significant differences in the total number of
778 symptoms by deletion profile were revealed for chromosome 8 (one-way ANOVA $p =$
779 0.83) or chromosome 13 ($p = 0.72$). No obvious spatial patterns in subtelomeric
780 deletion profiles were apparent at the regional level (**Supplementary Table 2**).

781

782 *Comparison of pfhrp2/3 PCR and HRP2 bead-based assays*

783 We observed expected agreement between the results of *pfhrp2/3* PCR
784 assays, RDTs, and a bead-based immunoassay applied to a subset of 456 samples.
785 93% (95% CI 86-96) of samples *pfhrp2+/3+* by PCR tested positive for HRP2
786 antigen (GM 40,284 pg/mL HRP2, geoSD 7.5). In comparison, 19% (12-29) of
787 *pfhrp2-/3-* samples were HRP2+, with a GM of 2,089 pg/mL HRP2 (geoSD 5.5).
788 HRP2+ but *pfhrp2-* PCR results are expected in a subset of subjects because HRP2
789 antigenemia can persist for weeks after clearance of parasitemia³². 92% (95% CI

790 88-95) of samples HRP2+, Pf-LDH+ by RDT were HRP2+ by the antigen assay (GM
791 34,536 pg/mL, geoSD 6.5), compared to 24% (95% CI 18-31) of those with the
792 discordant HRP2-, Pf-LDH+ RDT profile of interest (GM 2,455 pg/mL, geoSD: 7.2)
793 **(Table 2)**.

794

795 *Subtelomeric profiling and variant calling using MIPs*

796 241 MIPs mapped to 273 targeted loci on the reference genome, including 32
797 extra loci accounting for the multicopy genes on chromosome 11. Probes failing to
798 amplify in >90% of the samples were removed from the analysis, leaving 244 loci.
799 841 of 1014 samples and controls had sequence data after read mapping. 20 of 841
800 belonged to control strains (positive controls). None of the 20 negative controls had
801 any sequence mapping to the reference genome. Deletion calls were only made in
802 samples with sufficient depth of UMI coverage (see Methods), leaving 375
803 high-coverage samples from the study cohort and 6 controls in the final call set.

804 **Supplementary Files**

805 *Supplementary tables are compiled into a single file for ease of viewing.*

806

807 **Supplementary Table 1.** PCR results by age, sex, and parasite density.

808 **Supplementary Table 2.** RDT profile by district for individuals *P.*
809 *falciparum*-positive by RDT.

810 **Supplementary Table 3.** Prevalence estimate sensitivity analysis.

811 **Supplementary Table 4.** MIP subtelomeric structural profiles by region.

812 **Supplementary Table 5.** *Pfhrp2/3* MIP panel design overview, including genes
813 targeted, MIPs designed, and genomic coordinates.

814 **Supplementary Table 6.** *Pfhrp2/3* MIP panel probe sequences.

815 **Supplementary Table 7.** Absolute (A) and normalized (B) unique molecular
816 identifier (UMI) counts by sample and locus.

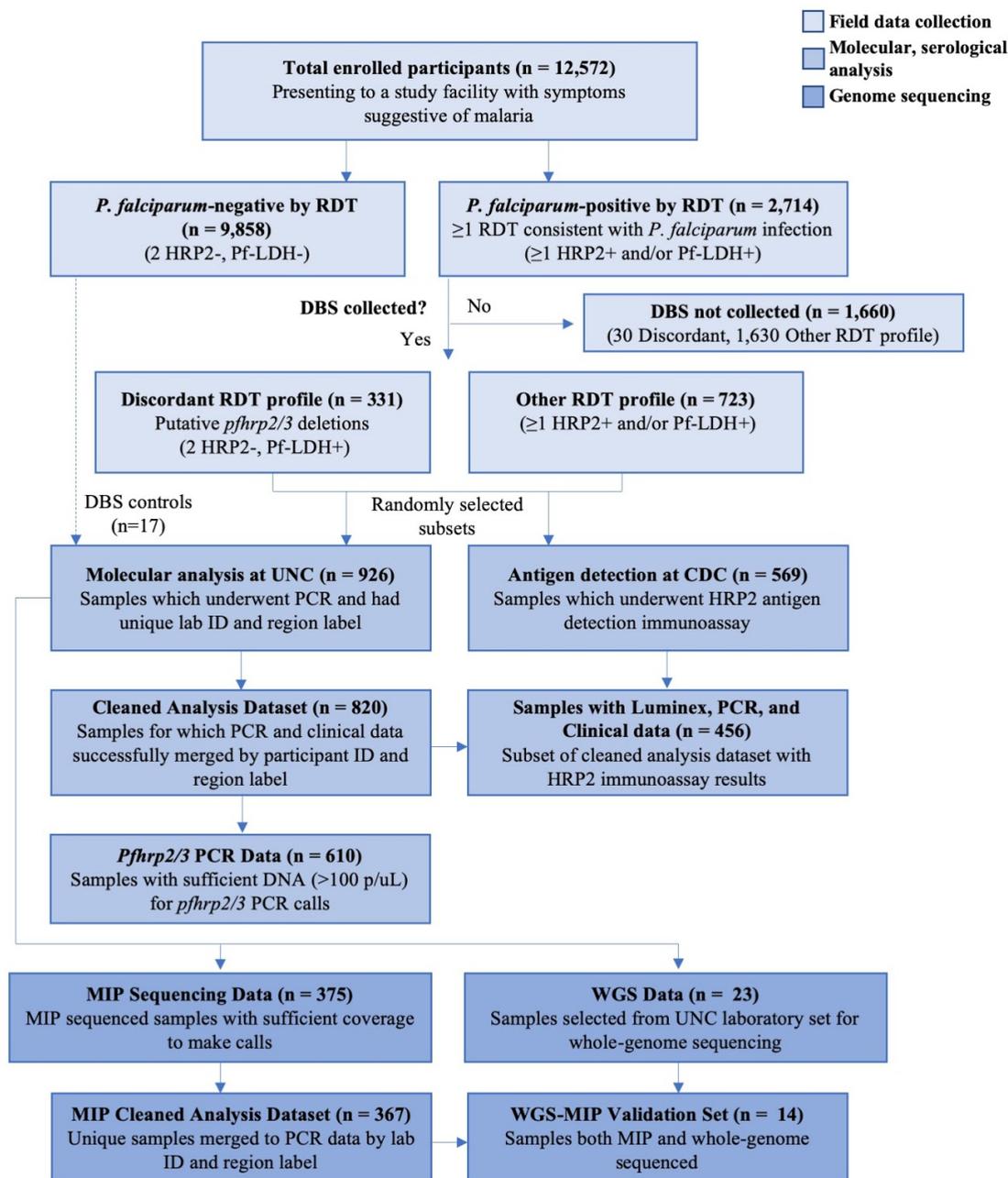
817 **Supplementary Table 8.** MIP variant calls, which were converted to .map and
818 .hap files.

819 **Supplementary Table 9.** ENA accession numbers of previously published WGS
820 data.

821 **Supplementary Table 10.** Summary statistics of WGS coverage for samples
822 sequenced in this study.

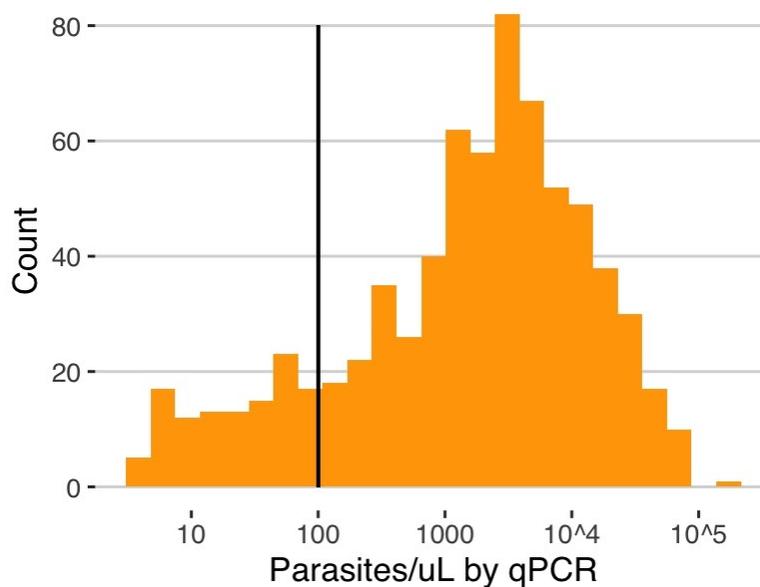
823 **Supplementary File.** PCR reaction conditions.

824 **Supplementary Figures**



825

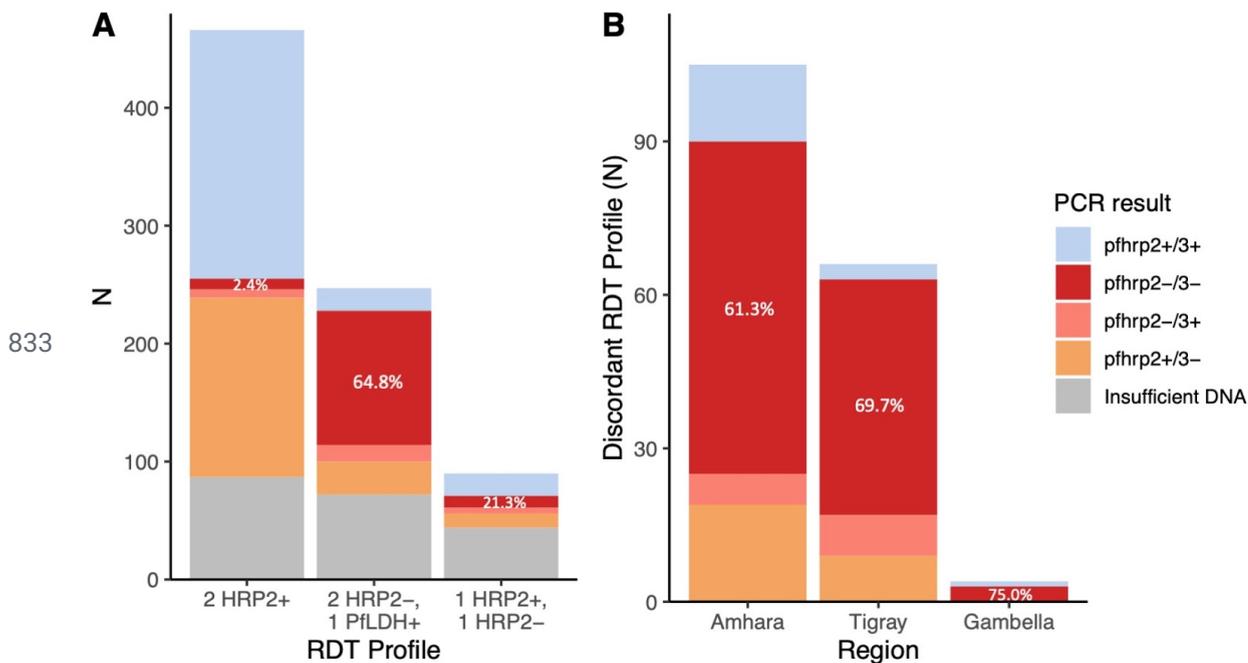
826 **Supplementary Figure 1. Study samples and assays performed.**



827

828 **Supplementary Figure 2. Parasite density distribution.** *Pf1dh* quantitative PCR
829 (qPCR) results used to assess parasite density and determine which samples were
830 eligible for *pfhrp2/3* deletion genotyping using a series of PCR assays. *Pfhrp2/3*
831 deletions were only called in samples with >100 parasites/ μ L (solid line).

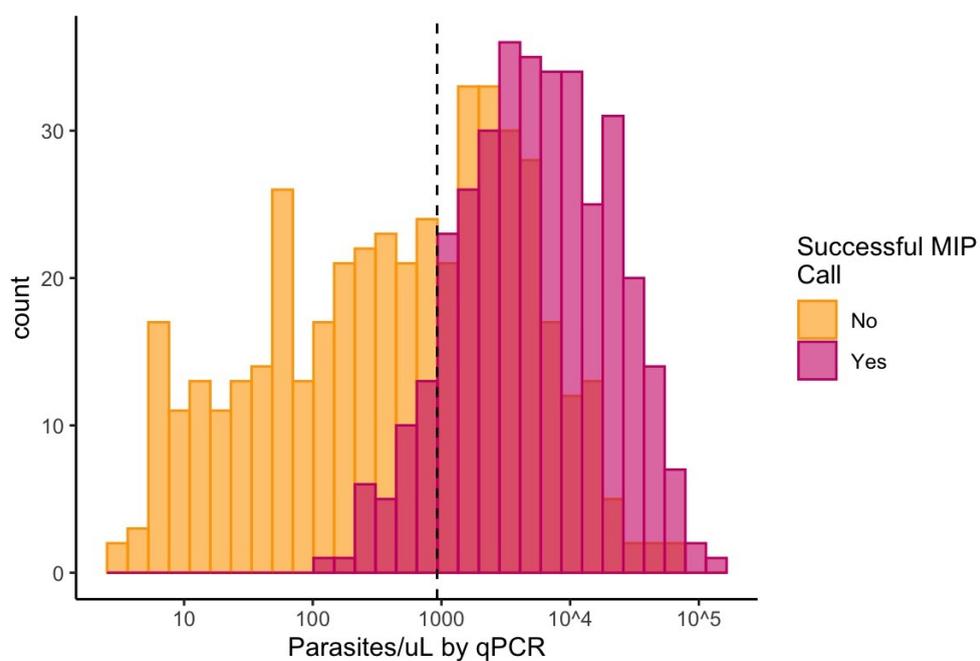
832



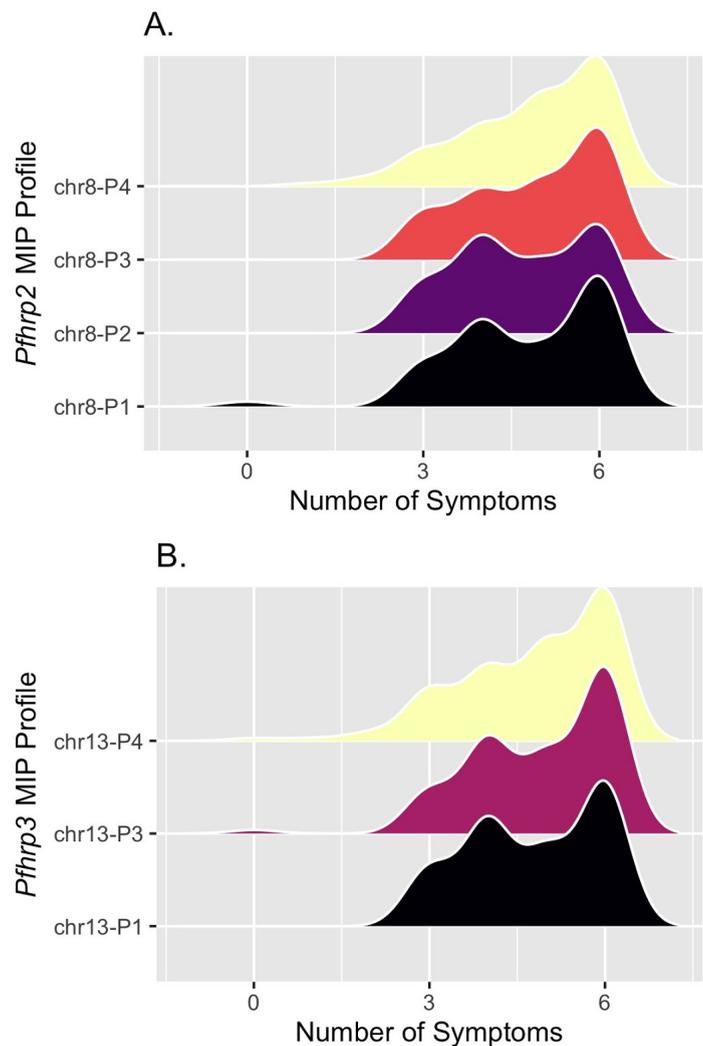
834 **Supplementary Figure 3. *Pfhrp2* and *pfhrp3* PCR results by RDT profile.** A)
 835 Concordance between RDT profile and PCR *pfhrp2/3* result for *P. falciparum*
 836 samples with >100 parasites/ μ L. B) *Pfhrp2/3* PCR results for participants with the
 837 discordant RDT profile and sufficient DNA for molecular analysis (n = 176), by study
 838 region.

839

840



841 **Supplementary Figure 4. Successful MIP deletion calls versus qPCR parasite**
842 **density.** Comparison of MIP call results and qPCR parasite densities suggests a
843 project-specific threshold for MIP calling of approximately 925 p/μL of whole blood
844 (dashed line).

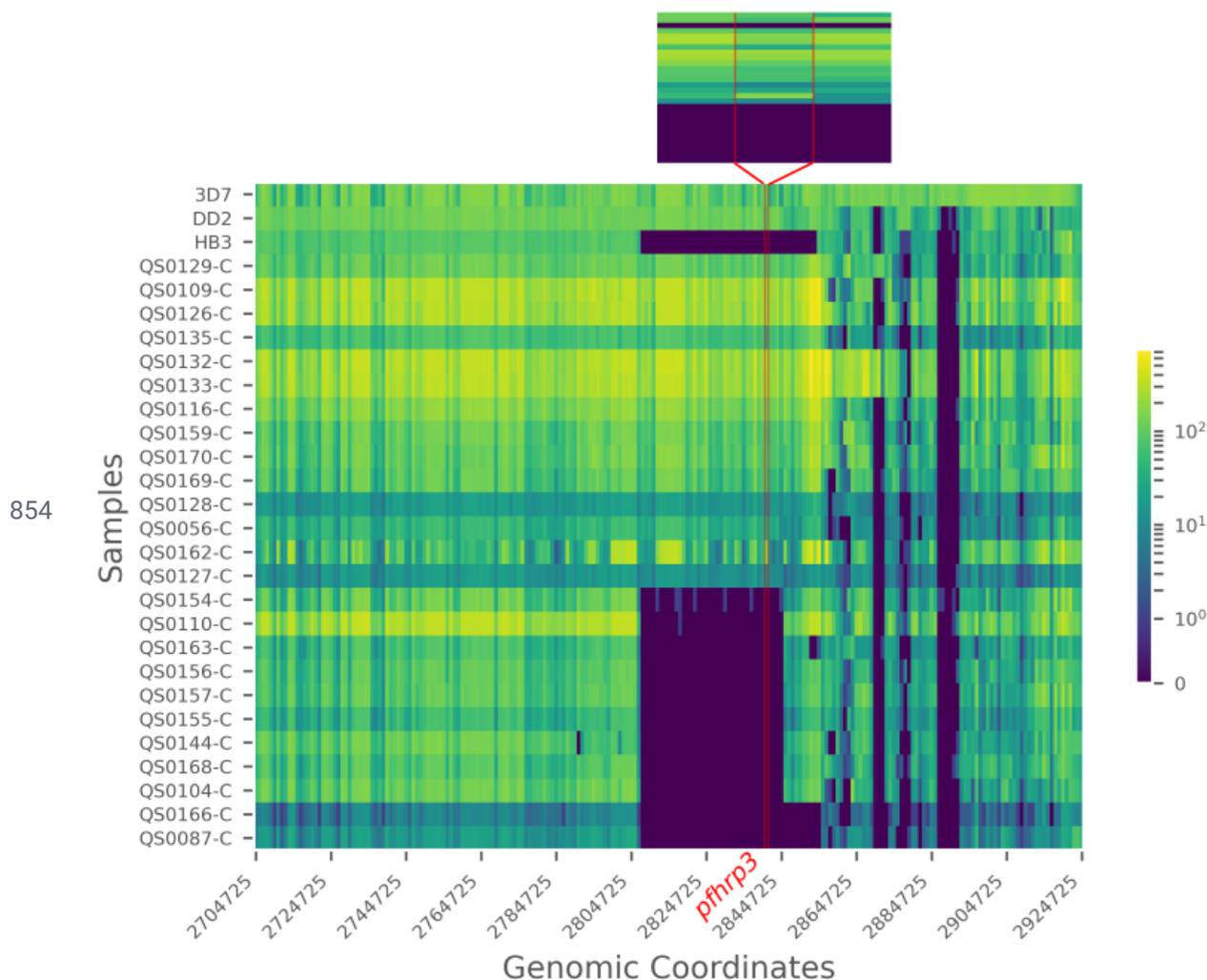


845

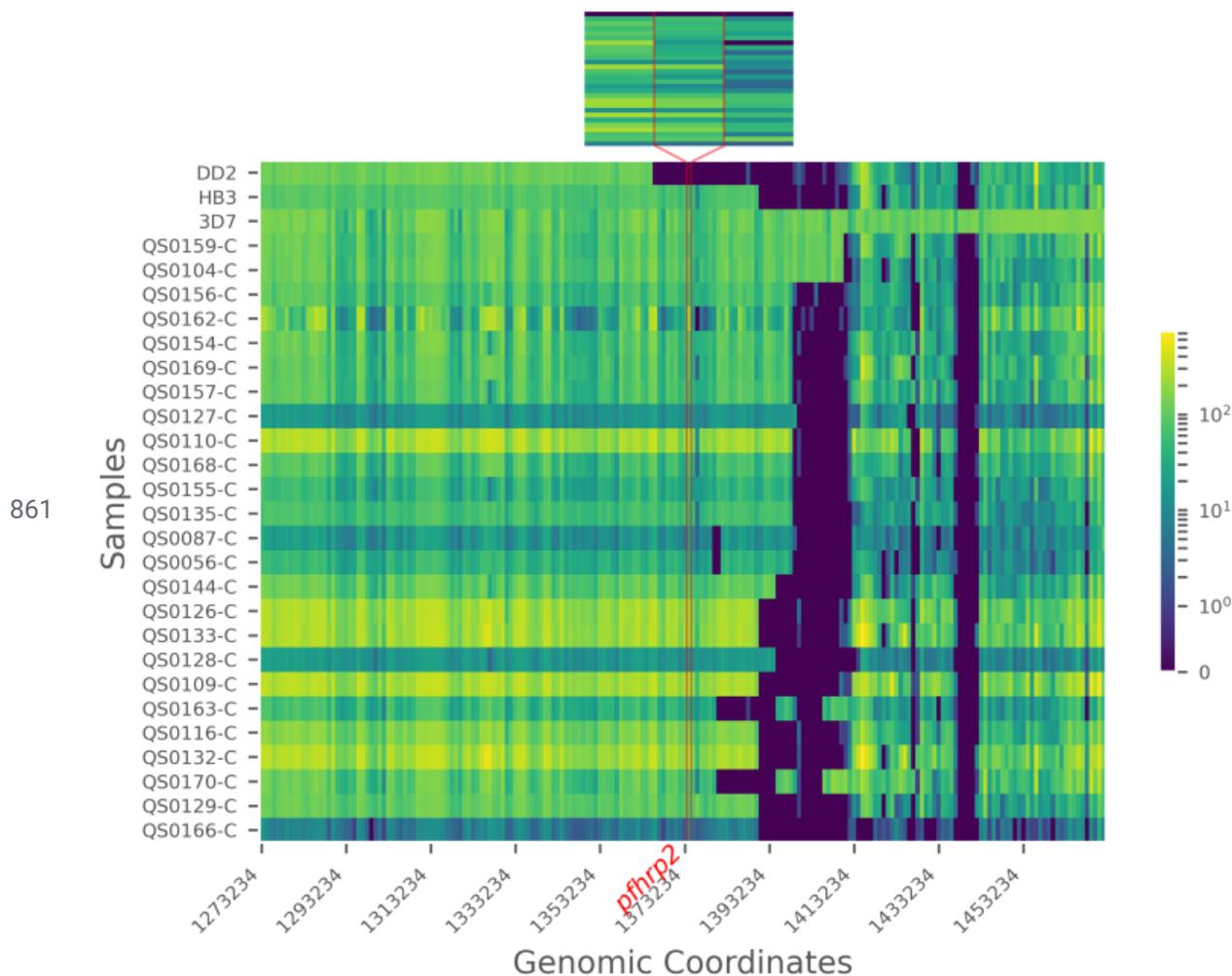
846

847 **Supplementary Figure 5. Disease severity by subtelomeric structural profile.**

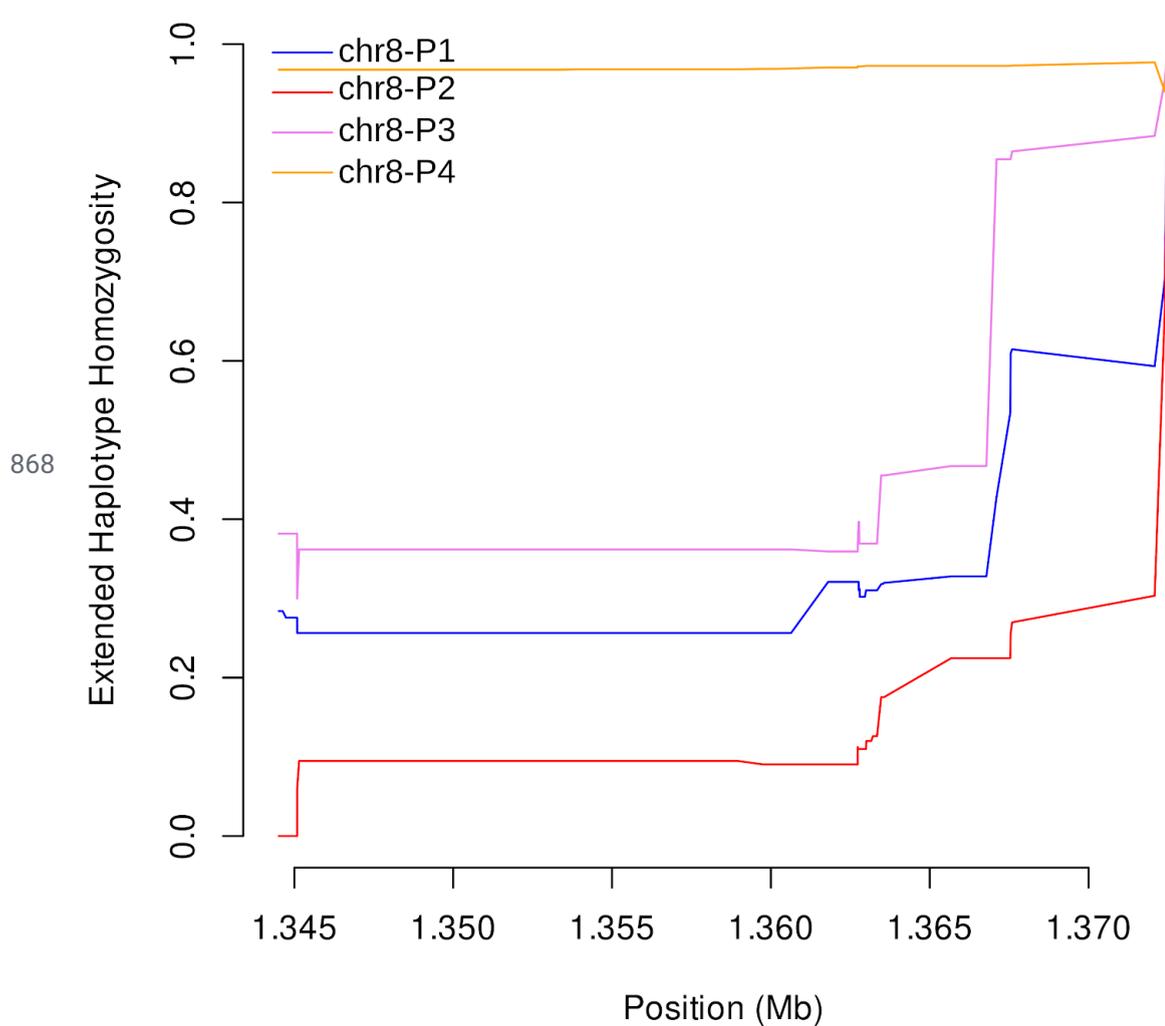
848 Smoothed distribution of disease severity for each of the broader deletion breakpoint
849 haplotypes along chromosomes 8 (A) and 13 (B) identified by MIP genomic
850 enrichment. The total number of symptoms with which participants presented was
851 used to estimate disease severity, with all participants evaluated for: fever,
852 headache, joint pain, feeling cold, nausea, and loss of appetite. Profile chr13-P2 was
853 excluded for its small sample size (n=1).



855 **Supplementary Figure 6. Chromosome 13 telomeric-end coverage (aligned**
856 **reads/locus) plot of WGS control strains and 25 published Ethiopian genomes**
857 **from 2013-2015 (MalariaGEN).** The location of *pfhrp3* is indicated by vertical red
858 lines. Inset showing *pfhrp3* and 1 kilobase flanking genomic region. Large
859 subtelomeric deletions containing *pfhrp3* are apparent in the laboratory strain HB3,
860 as well as 11 samples from Ethiopia.

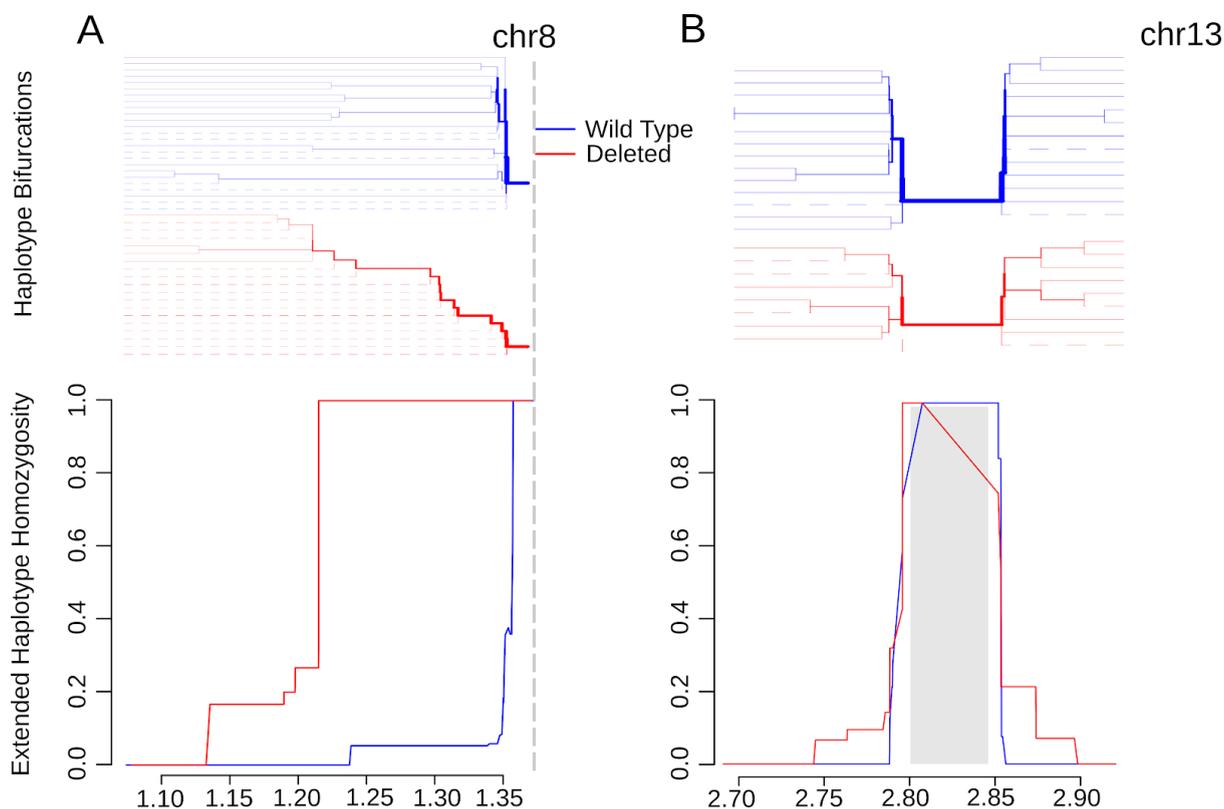


862 **Supplementary Figure 7. Chromosome 8 telomeric-end coverage (aligned**
863 **reads/locus) plot of WGS control strains and 25 published Ethiopian genomes**
864 **from 2013-2015 (MalariaGEN).** The location of *pfhrp2* is indicated by vertical red
865 lines. Inset showing *pfhrp2* and 1 kilobase flanking genomic region. Large
866 subtelomeric deletions compared to the reference strain 3D7 are apparent in most
867 samples. None of the deletions involve *pfhrp2* except the DD2 strain.



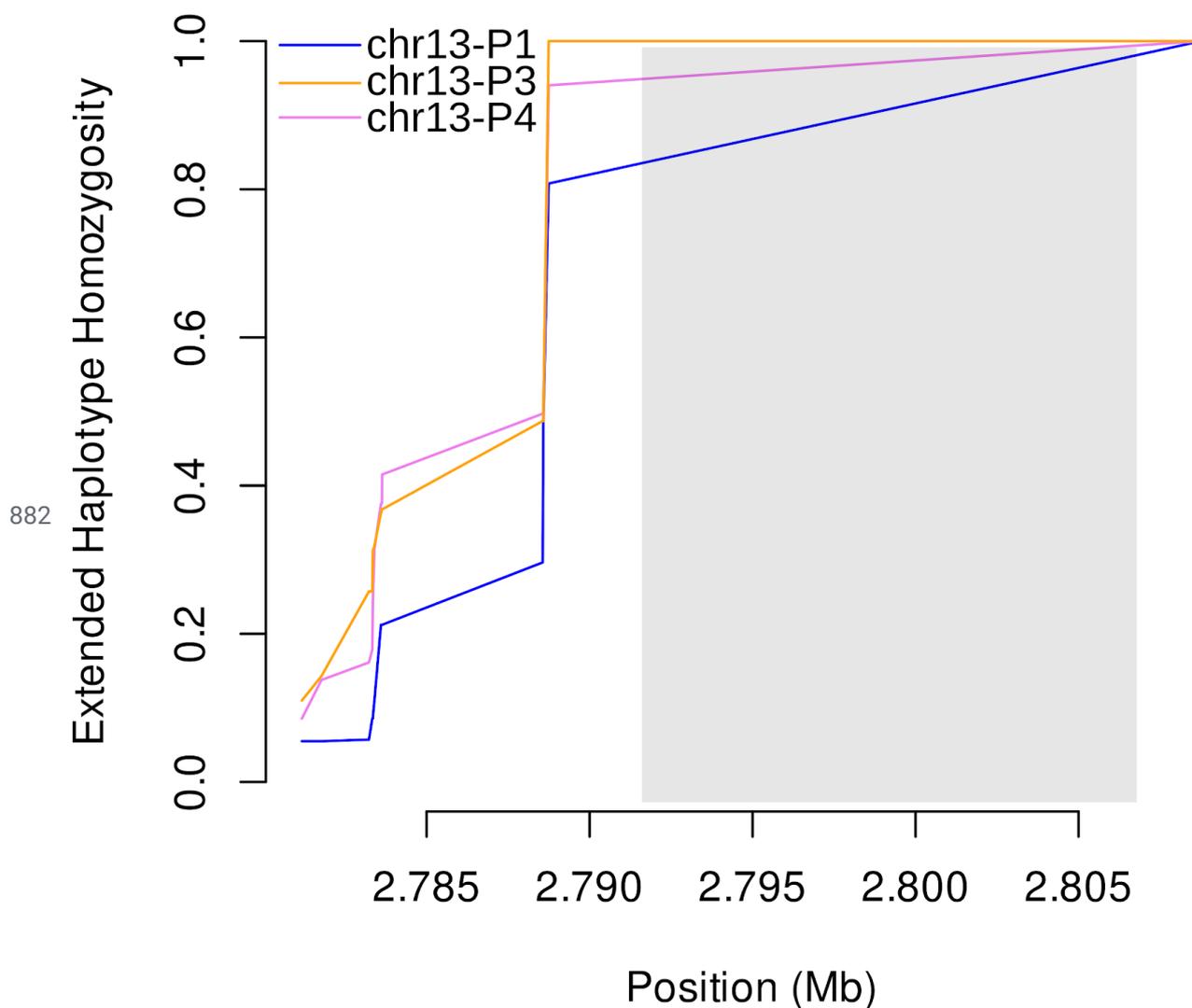
869 **Supplementary Figure 8. Extended haplotype homozygosity centromeric to**
870 **chromosome 8 subtelomeric structural profiles P1-P4 using MIP data.** The only
871 *pflhrp2*-deleted profile (chr8-P4) showed sustained EHH, whereas EHH quickly broke
872 down for the *pflhrp2*-intact profiles (P1-P3). A vertical dashed line on the right marks
873 the centromeric end of the profile 4 (*pflhrp2*) deletion.

874



876

877 **Supplementary Figure 9. Extended haplotype homozygosity (bottom) and the**
878 **bifurcation diagrams showing haplotype branching (top) centromeric to the**
879 ***pfhrp2* (A) and surrounding *pfhrp3* (B) deletions based on WGS data. Vertical**
880 **dashed line indicating the centromeric end of the chromosome 8 deletion (A). Gray**
881 **box demarcating the chromosome 13 deletion (B). Abbreviations: Mb, mega-base.**



Supplementary Figure 10. Extended haplotype homozygosity centromeric to chromosome 13 subtelomeric structural profiles P1, P3 and P4 using MIP data. Chr13-P2 profile was observed in only one sample and not included in the haplotype analysis. EHH quickly brown down for all profiles (P1: *pfhrp3*-intact, P3-P4: *pfhrp3*-deleted). A vertical dashed line on the right marks the centromeric end of the P3 and P4 deletions. No variants in the duplicated segment (gray box) were used in the EHH analysis.

891 REFERENCES

- 892 1. World Health Organization. *False-negative RDT results and implications of new*
893 *reports of P. falciparum histidine-rich protein 2/3 gene deletions*. (WHO, Geneva,
894 2016).
- 895 2. Verma, A. K., Bharti, P. K. & Das, A. HRP-2 deletion: a hole in the ship of
896 malaria elimination. *Lancet Infect. Dis.* **18**, 826–827 (2018).
- 897 3. Wellems, T. E. & Howard, R. J. Homologous genes encode two distinct
898 histidine-rich proteins in a cloned isolate of *Plasmodium falciparum*.
899 *Proceedings of the National Academy of Sciences* **83**, 6065–6069 (1986).
- 900 4. Howard, R. J. *et al.* Secretion of a malarial histidine-rich protein (Pf HRP II) from
901 *Plasmodium falciparum*-infected erythrocytes. *J. Cell Biol.* **103**, 1269 (1986).
- 902 5. Li, B. *et al.* Performance of pfHRP2 versus pLDH antigen rapid diagnostic tests
903 for the detection of *Plasmodium falciparum*: a systematic review and
904 meta-analysis. *Arch. Med. Sci.* **13**, 541–549 (2017).
- 905 6. World Health Organization. *Good practices for selecting and procuring rapid*
906 *diagnostic tests for malaria*. (WHO, Geneva, 2011).
- 907 7. Baker, J. *et al.* Genetic diversity of *Plasmodium falciparum* histidine-rich protein
908 2 (PfHRP2) and its effect on the performance of PfHRP2-based rapid diagnostic
909 tests. *J. Infect. Dis.* **192**, 870–877 (2005).
- 910 8. Cheng, Q. *et al.* *Plasmodium falciparum* parasites lacking histidine-rich protein 2
911 and 3: a review and recommendations for accurate reporting. *Malar. J.* **13**, 283
912 (2014).
- 913 9. Poti, K. E., Sullivan, D. J., Dondorp, A. M. & Woodrow, C. J. HRP2:
914 Transforming Malaria Diagnosis, but with Caveats. *Trends Parasitol.* **36**,
915 112–126 (2020).
- 916 10. Gamboa, D. *et al.* A large proportion of *P. falciparum* isolates in the Amazon
917 region of Peru lack p_{fh}rp2 and p_{fh}rp3: implications for malaria rapid diagnostic
918 tests. *PLoS One* **5**, e8091 (2010).
- 919 11. Cheng, Q. *et al.* *Plasmodium falciparum* parasites lacking histidine-rich protein 2
920 and 3: a review and recommendations for accurate reporting. *Malar. J.* **13**, 283
921 (2014).
- 922 12. Grignard, L. *et al.* A Novel Multiplex qPCR Assay for Detection of *Plasmodium*
923 *falciparum* with Histidine-rich Protein 2 and 3 (p_{fh}rp2 and p_{fh}rp3) Deletions in
924 Polyclonal Infections. *EBioMedicine* **55**, 102757 (2020).
- 925 13. Plucinski, M. M. *et al.* Screening for P_{fh}rp2/3-Deleted *Plasmodium falciparum*,
926 Non-*falciparum*, and Low-Density Malaria Infections by a Multiplex Antigen
927 Assay. *J. Infect. Dis.* **219**, 437–447 (2019).
- 928 14. Parr, J. B., Anderson, O., Juliano, J. J. & Meshnick, S. R. Streamlined,

- 929 PCR-based testing for pfhrp2- and pfhrp3-negative *Plasmodium falciparum*.
930 *Malar. J.* **17**, 137 (2018).
- 931 15. Berhane, A. *et al.* Major Threat to Malaria Control Programs by *Plasmodium*
932 *falciparum* Lacking Histidine-Rich Protein 2, Eritrea. *Emerg. Infect. Dis.* **24**,
933 462–470 (2018).
- 934 16. Akinyi, S. *et al.* Multiple genetic origins of histidine-rich protein 2 gene deletion in
935 *Plasmodium falciparum* parasites from Peru. *Sci. Rep.* **3**, 2797 (2013).
- 936 17. Parr, J. B. *et al.* Pfhrp2-Deleted *Plasmodium falciparum* Parasites in the
937 Democratic Republic of the Congo: A National Cross-sectional Survey. *J. Infect.*
938 *Dis.* **16**, 36–44 (2017).
- 939 18. Sepúlveda, N. *et al.* Global analysis of *Plasmodium falciparum* histidine-rich
940 protein-2 (pfhrp2) and pfhrp3 gene deletions using whole-genome sequencing
941 data and meta-analysis. *Infect. Genet. Evol.* **62**, 211–219 (2018).
- 942 19. Pearson, R. D., Amato, R., Kwiatkowski, D. P. & MalariaGEN *Plasmodium*
943 *falciparum* Community Project. An open dataset of *Plasmodium falciparum*
944 genome variation in 7,000 worldwide samples. Pre-print at *bioRxiv* 824730
945 (2019) doi:10.1101/824730.
- 946 20. Gibbons, J. *et al.* Lineage-specific expansion of *Plasmodium falciparum*
947 parasites with pfhrp2 deletion in the Greater Mekong Subregion. *J. Infect. Dis.*
948 (2020) Published advance access Oct 1, 2020; doi:10.1093/infdis/jjaa250.
- 949 21. Otto, T. D. *et al.* Long read assemblies of geographically dispersed *Plasmodium*
950 *falciparum* isolates reveal highly structured subtelomeres. *Wellcome Open Res*
951 **3**, 52 (2018).
- 952 22. Verma, A. K., Bharti, P. K. & Das, A. HRP-2 deletion: a hole in the ship of
953 malaria elimination. *Lancet Infect. Dis.* **18**, 826–827 (2018).
- 954 23. World Health Organization. *False-negative RDT results and implications of new*
955 *P. falciparum* *histidine-rich protein 2/3 gene deletions*. (WHO, Geneva, 2016).
- 956 24. Menegon, M. *et al.* Identification of *Plasmodium falciparum* isolates lacking
957 histidine-rich protein 2 and 3 in Eritrea. *Infect. Genet. Evol.* **55**, 131–134 (2017).
- 958 25. Thomson, R. *et al.* Prevalence of *Plasmodium falciparum* lacking histidine-rich
959 proteins 2 and 3: a systematic review. *Bull. World Health Organ.* **98**, 558–568F
960 (2020).
- 961 26. Ethiopian Federal Ministry of Health. *National Malaria Elimination Roadmap*.
962 (Addis Ababa, 2017).
- 963 27. World Health Organization. *World Malaria Report 2019*. (WHO, Geneva 2020).
- 964 28. Taffese, H. S. *et al.* Malaria epidemiology and interventions in Ethiopia from
965 2001 to 2016. *Infect Dis Poverty* **7**, 103 (2018).

- 966 29. World Health Organization. *Template protocols to support surveillance and*
967 *research for pfhrp2/pfhrp3 gene deletions*. Available at
968 <https://www.who.int/malaria/publications/atoz/hrp2-deletion-protocol/en/> (WHO,
969 Geneva, 2020).
- 970 30. Morgan, A. P. *et al.* Falciparum malaria from coastal Tanzania and Zanzibar
971 remains highly connected despite effective control efforts on the archipelago.
972 *Malar. J.* **19**, 47 (2020).
- 973 31. Aydemir, O. *et al.* Drug-Resistance and Population Structure of Plasmodium
974 falciparum Across the Democratic Republic of Congo Using High-Throughput
975 Molecular Inversion Probes. *J. Infect. Dis.* **218**, 946–955 (2018).
- 976 32. Markwalter, C. F. *et al.* Characterization of Plasmodium Lactate Dehydrogenase
977 and Histidine-Rich Protein 2 Clearance Patterns via Rapid On-Bead Detection
978 from a Single Dried Blood Spot. *Am. J. Trop. Med. Hyg.* **98**, 1389–1396 (2018).
- 979 33. Nunes, M. C., Okada, M., Scheidig-Benatar, C., Cooke, B. M. & Scherf, A.
980 Plasmodium falciparum FIKK kinase members target distinct components of the
981 erythrocyte membrane. *PLoS One* **5**, e11747 (2010).
- 982 34. Jaskiewicz, E., Jodłowska, M., Kaczmarek, R. & Zerka, A. Erythrocyte
983 glycophorins as receptors for Plasmodium merozoites. *Parasit. Vectors* **12**, 317
984 (2019).
- 985 35. Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome
986 from haplotype structure. *Nature* **419**, 832–837 (2002).
- 987 36. Berhane, A. *et al.* Rapid diagnostic tests failing to detect Plasmodium falciparum
988 infections in Eritrea: an investigation of reported false negative RDT results.
989 *Malar. J.* **16**, 105 (2017).
- 990 37. Boush, M. A. *et al.* Plasmodium falciparum isolate with histidine-rich protein 2
991 gene deletion from Nyala City, Western Sudan. *Sci. Rep.* **10**, 12822 (2020).
- 992 38. Cunningham, J. Tracking the global distribution and prevalence of pfhrp2/3 gene
993 deletions. Presentation at Tropical Medicine 2020, the annual meeting of the
994 American Society for Tropical Medicine and Hygiene, on November 19, 2020.
995 Symposium #167. (2020).
- 996 39. Gatton, M. L. *et al.* Implications of Parasites Lacking Plasmodium falciparum
997 Histidine-Rich Protein 2 on Malaria Morbidity and Control When Rapid
998 Diagnostic Tests Are Used for Diagnosis. *J. Infect. Dis.* **215**, 1156–1166 (2017).
- 999 40. Watson, O. J. *et al.* Modelling the drivers of the spread of Plasmodium
1000 falciparum hrp2 gene deletions in sub-Saharan Africa. *Elife* **6**, (2017).
- 1001 41. Watson, O. J. *et al.* Impact of seasonal variations in Plasmodium falciparum
1002 malaria transmission on the surveillance of pfhrp2 gene deletions. *Elife* **8**,
1003 (2019).
- 1004 42. Golassa, L., Messele, A., Amambua-Ngwa, A. & Swedberg, G. High prevalence

- 1005 and extended deletions in *Plasmodium falciparum* hrp2/3 genomic loci in
1006 Ethiopia. *PLoS One* **15**, e0241807 (2020).
- 1007 43. Pal, P. *et al.* *Plasmodium falciparum* Histidine-Rich Protein II Compromises
1008 Brain Endothelial Barriers and May Promote Cerebral Malaria Pathogenesis.
1009 *MBio* **7**, (2016).
- 1010 44. Pal, P. *et al.* *Plasmodium falciparum* histidine-rich protein II causes vascular
1011 leakage and exacerbates experimental cerebral malaria in mice. *PLoS One* **12**,
1012 e0177142 (2017).
- 1013 45. Leffler, E. M. *et al.* Resistance to malaria through structural variation of red
1014 blood cell invasion receptors. *Science* **356**, (2017).
- 1015 46. Saito, F. *et al.* Immune evasion of *Plasmodium falciparum* by RIFIN via inhibitory
1016 receptors. *Nature* **552**, 101–105 (2017).
- 1017 47. Niang, M. *et al.* STEVOR is a *Plasmodium falciparum* erythrocyte binding
1018 protein that mediates merozoite invasion and rosetting. *Cell Host Microbe* **16**,
1019 81–93 (2014).
- 1020 48. Federal Democratic Republic of Ethiopia Ministry of Health. *National Malaria*
1021 *Guidelines*. Available at
1022 [https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files](https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/eth_national_malaria_guidline_4th_edition.pdf)
1023 [s/documents/files/eth_national_malaria_guidline_4th_edition.pdf](https://www.humanitarianresponse.info/sites/www.humanitarianresponse.info/files/documents/files/eth_national_malaria_guidline_4th_edition.pdf) (Addis Ababa,
1024 2017).
- 1025 49. Plowe, C. V., Djimde, A., Bouare, M., Doumbo, O. & Wellems, T. E.
1026 Pyrimethamine and proguanil resistance-conferring mutations in *Plasmodium*
1027 *falciparum* dihydrofolate reductase: polymerase chain reaction methods for
1028 surveillance in Africa. *Am. J. Trop. Med. Hyg.* **52**, 565–568 (1995).
- 1029 50. Pickard, A. L. *et al.* Resistance to Antimalarials in Southeast Asia and Genetic
1030 Polymorphisms in *pfmdr1*. *Antimicrobial Agents and Chemotherapy* vol. 47
1031 2418–2423 (2003).
- 1032 51. *MIPTools*. (Github). Available at <https://github.com/bailey-lab/MIPTools>.
- 1033 52. Verity, R. *et al.* The impact of antimalarial resistance on the genetic structure of
1034 *Plasmodium falciparum* in the DRC. *Nat. Commun.* **11**, 2107 (2020).
- 1035 53. *MIPWrangler*. (Github). Available at <https://github.com/bailey-lab/MIPWrangler>.
- 1036 54. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel,
1037 Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss,
1038 Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau,
1039 Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. Scikit-learn: Machine
1040 Learning in Python. *J Mach Learn Res* **12**, 2825–2830 (2011).
- 1041 55. Oyola, S. O. *et al.* Whole genome sequencing of *Plasmodium falciparum* from
1042 dried blood spots using selective whole genome amplification. *Malar. J.* **15**, 597

- 1043 (2016).
- 1044 56. Miles, A. *et al.* Indels, structural variation, and recombination drive genomic
1045 diversity in *Plasmodium falciparum*. *Genome Res.* **26**, 1288–1299 (2016).
- 1046 57. Gautier, M. & Vitalis, R. rehh: an R package to detect footprints of selection in
1047 genome-wide SNP data from haplotype structure. *Bioinformatics* **28**, 1176–1177
1048 (2012).
- 1049 58. Chang, H. H. *et al.* THE REAL McCOIL: A method for the concurrent estimation
1050 of the complexity of infection and SNP allele frequency for malaria parasites.
1051 *PLoS Comput. Biol.* **13**, e1005348 (2017).