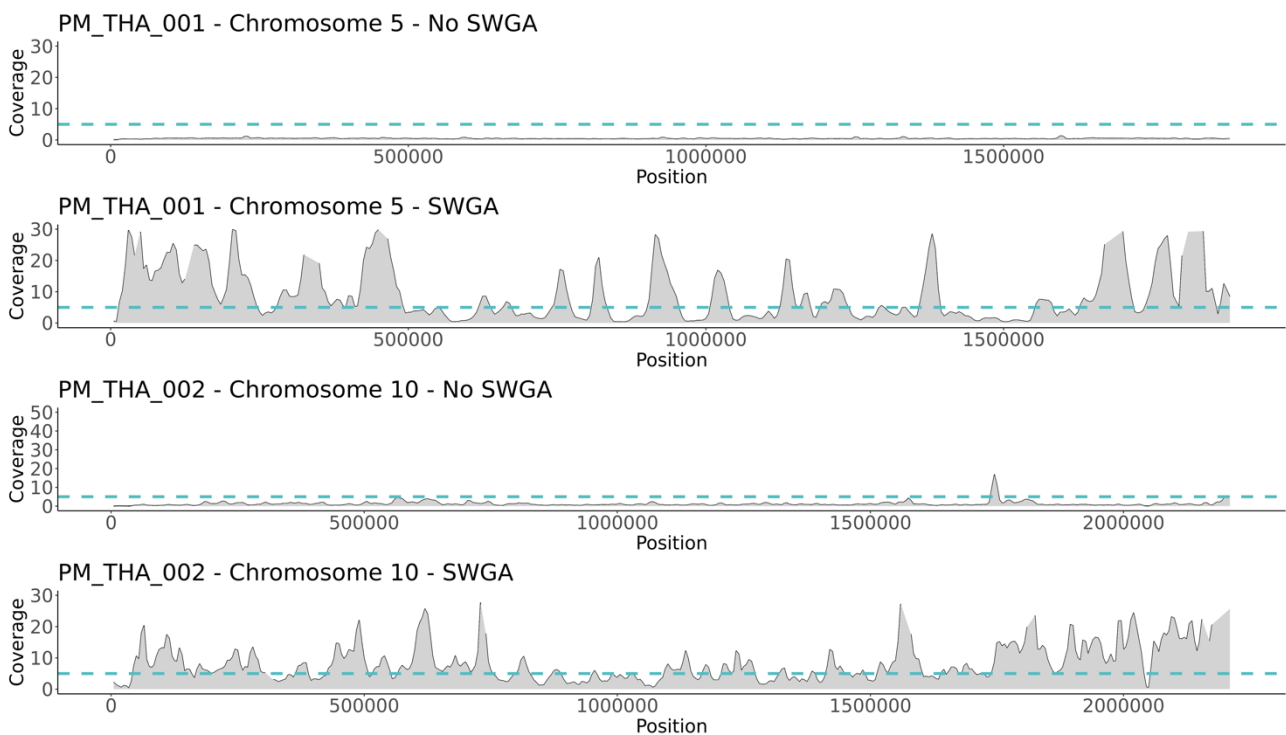


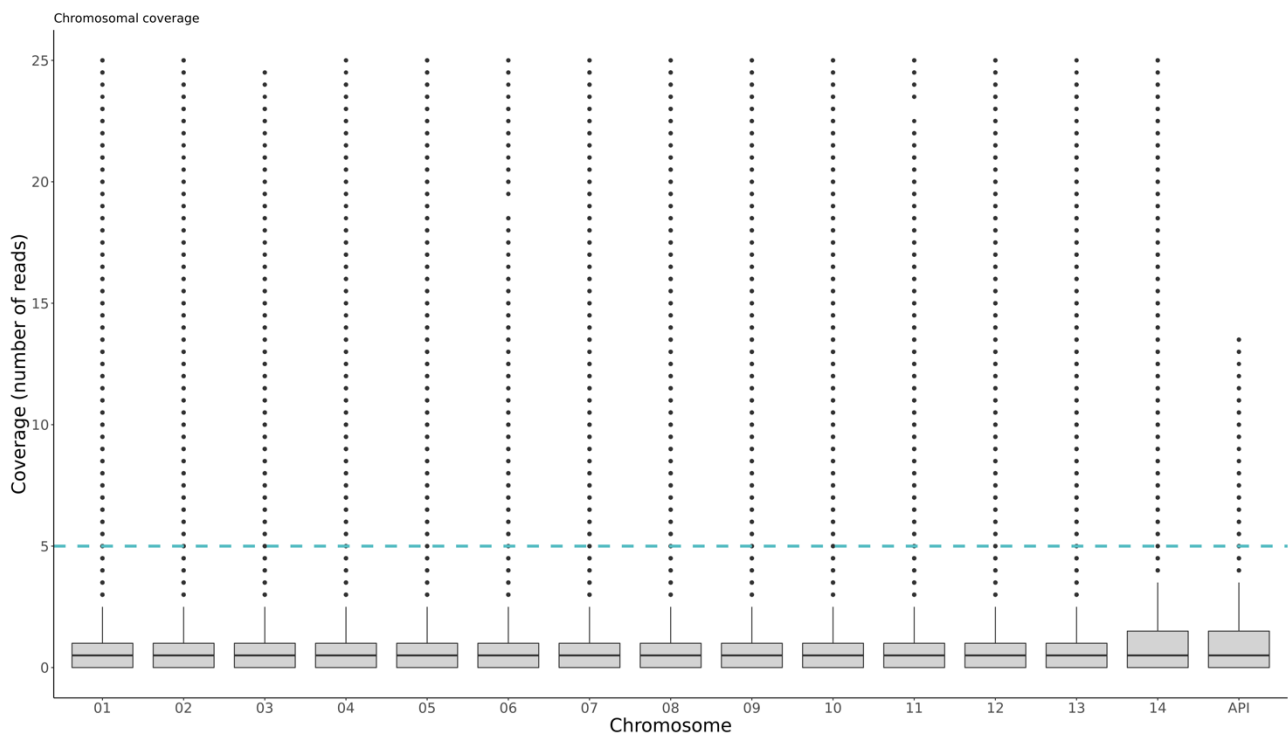
Selective Whole Genome Amplification of *Plasmodium malariae* DNA from clinical samples reveals insights into population structure

Amy Ibrahim¹, Ernest Diez Benavente¹, Debbie Nolder^{1,2}, Stephane Proux³, Matthew Higgins¹, Julian Muwanguzi^{1,2}, Paula Josefina Gomez Gonzalez¹, Hans-Peter Fuehrer⁴, Cally Roper¹, Francois Nosten^{3,5}, Colin Sutherland¹, Taane G. Clark^{1,6}, Susana Campino¹

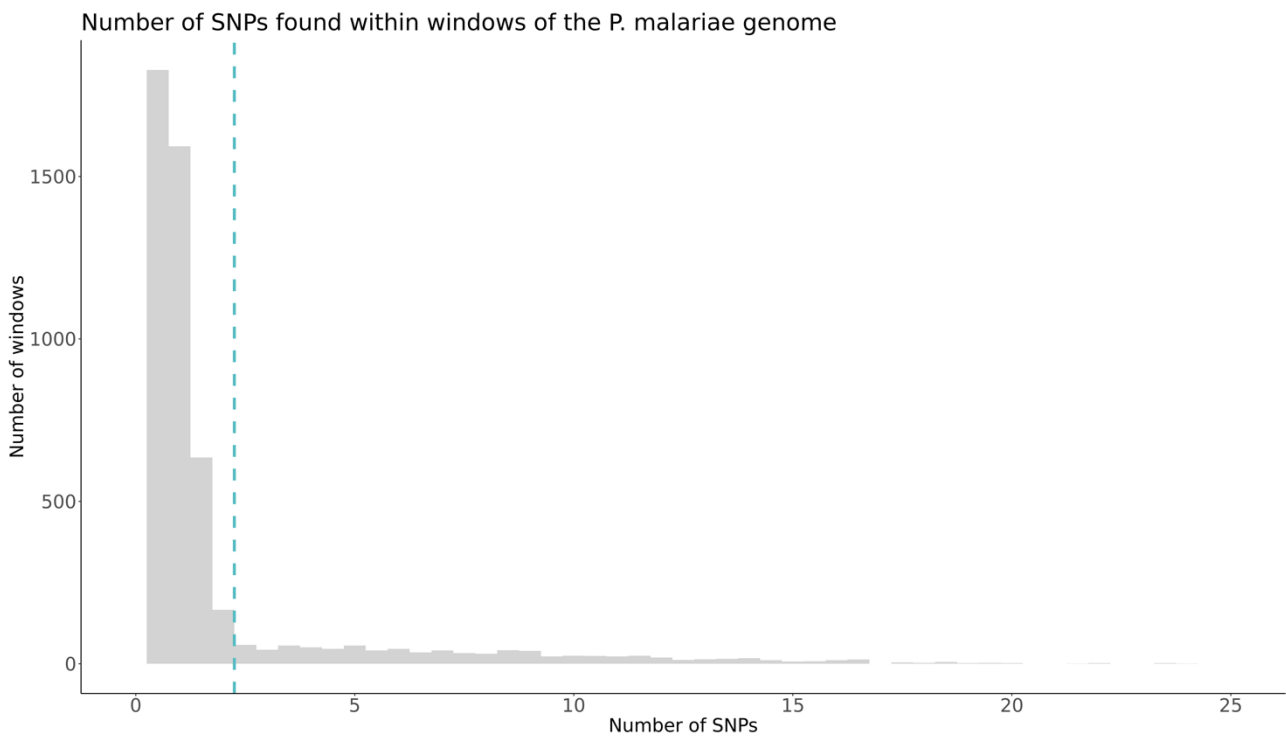
SUPPLEMENTARY INFORMATION



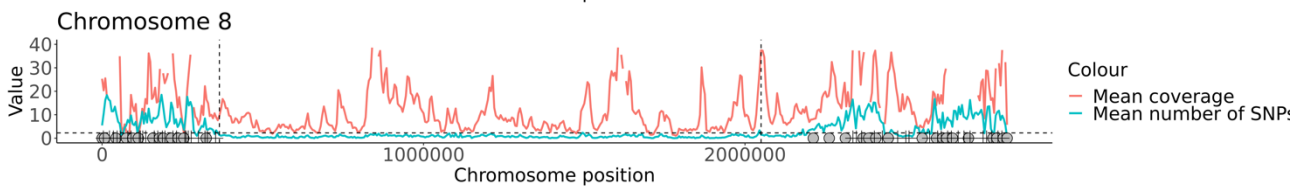
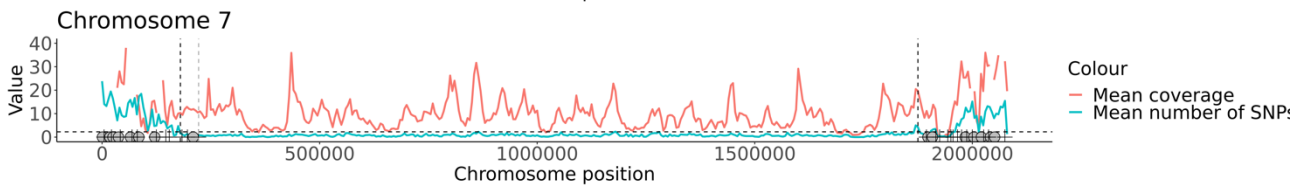
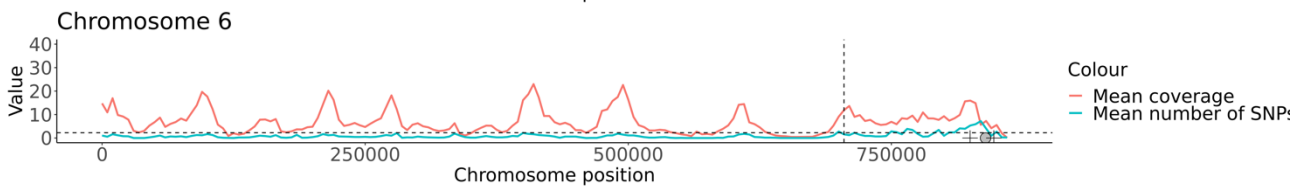
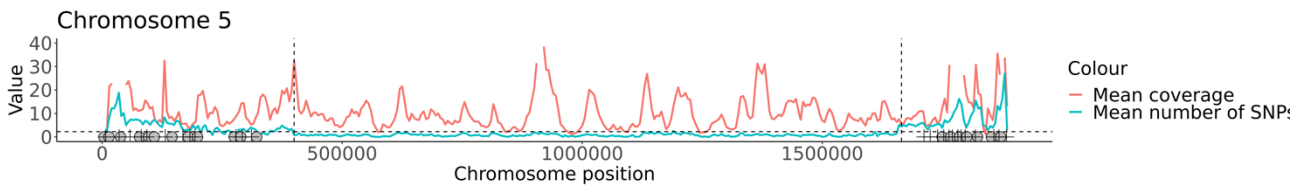
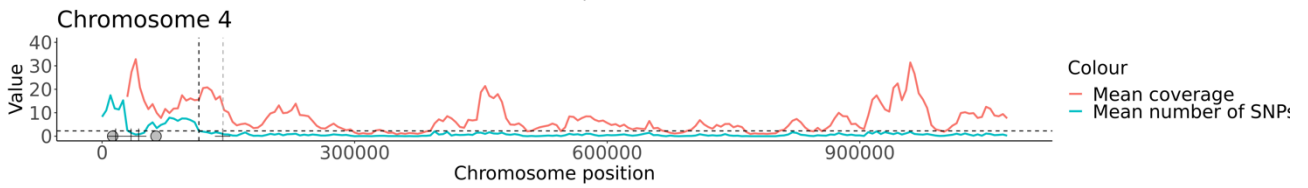
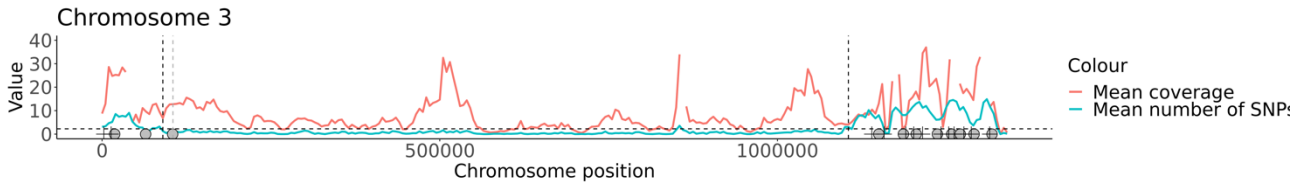
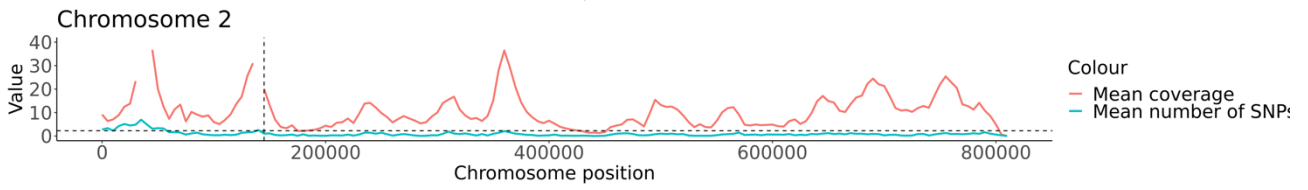
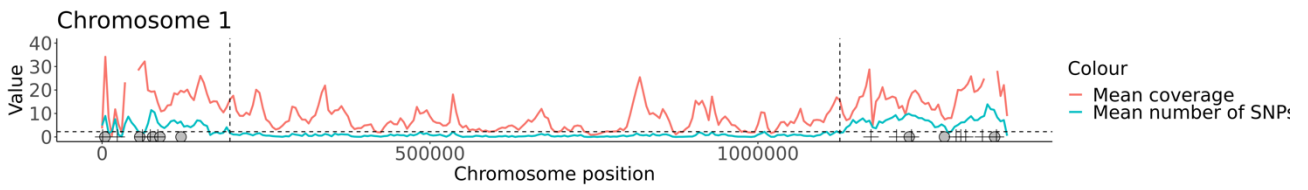
S1 Fig. Genome coverage is increased after SWGA. The top two plots show the coverage (as measured by the number of sequencing reads at each position) of chromosome 5 for sample PM_THA_001 before (No SWGA) and after amplification with Pmset1 (SWGA). The same comparison is shown below with chromosome 10 for sample PM_THA_002. Coverage was calculated using a sliding window of 10 kb. The horizontal y-intercept indicates a coverage of 5-fold which is recommended for confident SNP calling.

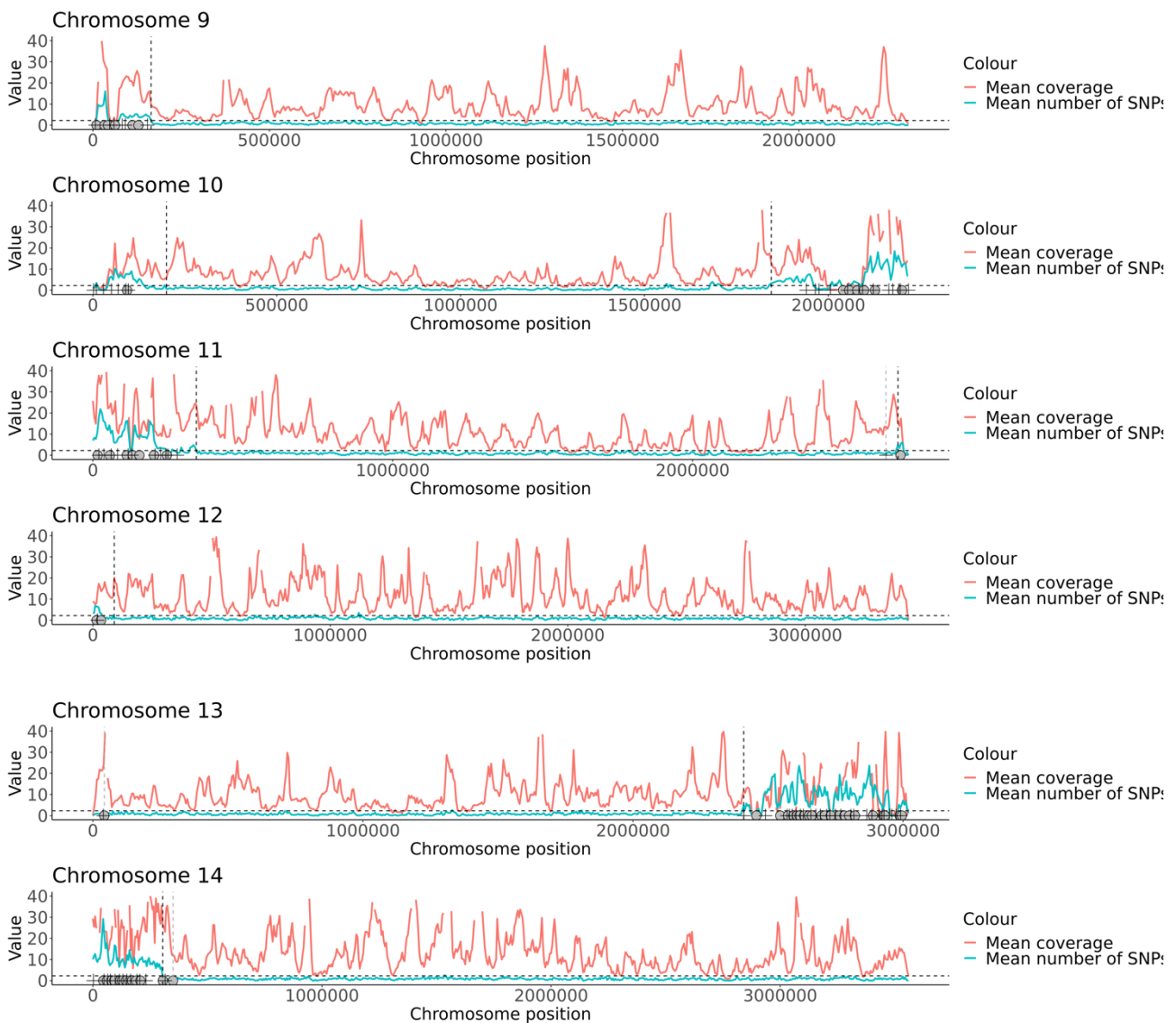


S2 Fig. Sequencing coverage by chromosome without amplification. The distribution of coverage for each position within the genome between two unamplified samples (PM_THA_001 and PM_THA_002), among the 14 nuclear chromosomes and the organellar apicoplast genome. The blue horizontal dashed line represents a cut-off of coverage of 5-fold which is recommended for SNP calling. Outliers are denoted by a black circle, many individual positions will have a coverage of 5-fold, but the majority of the genome falls below this.

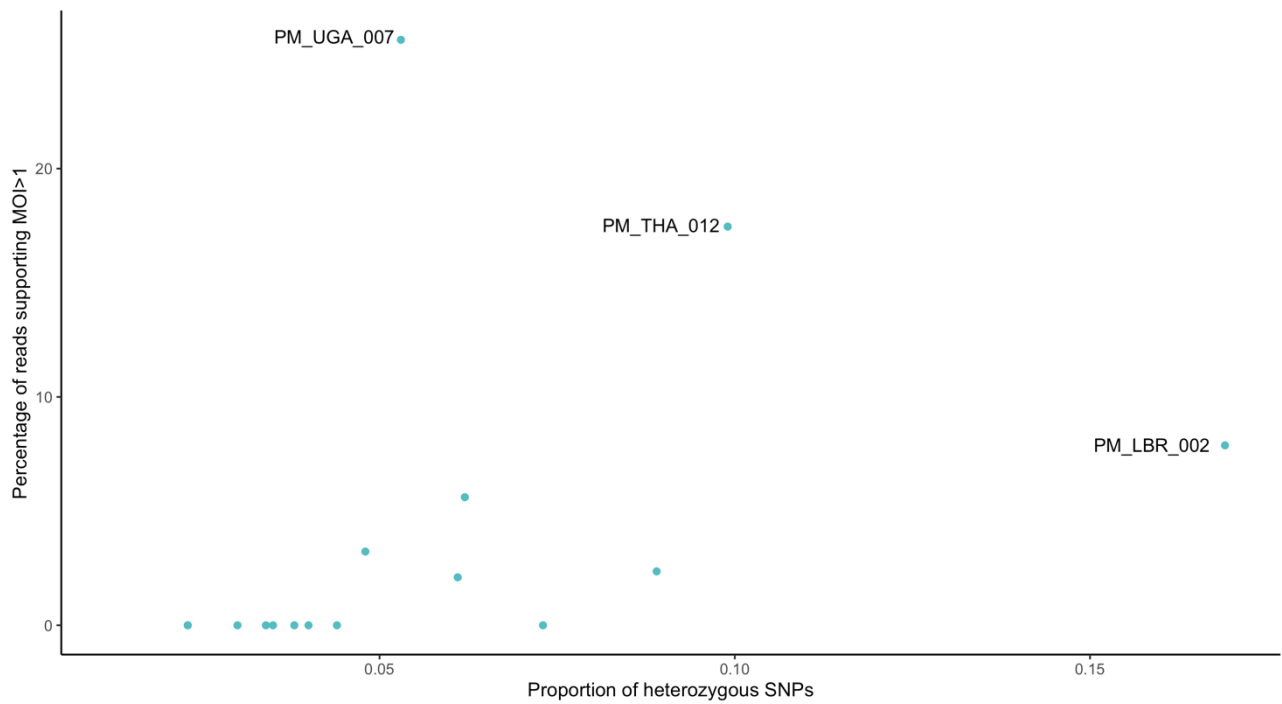


S3 Fig. Number of SNPs seen within 5 kb windows of the *P. malariae* genome. An average of the number of SNPs seen within all 18 samples was generated for 5 kb sliding windows across the entire genome. The distribution of SNPs within each window is shown. The blue dashed x-intercept is at 2.25 SNPs which is used as the limit of SNPs per window for the core genome.

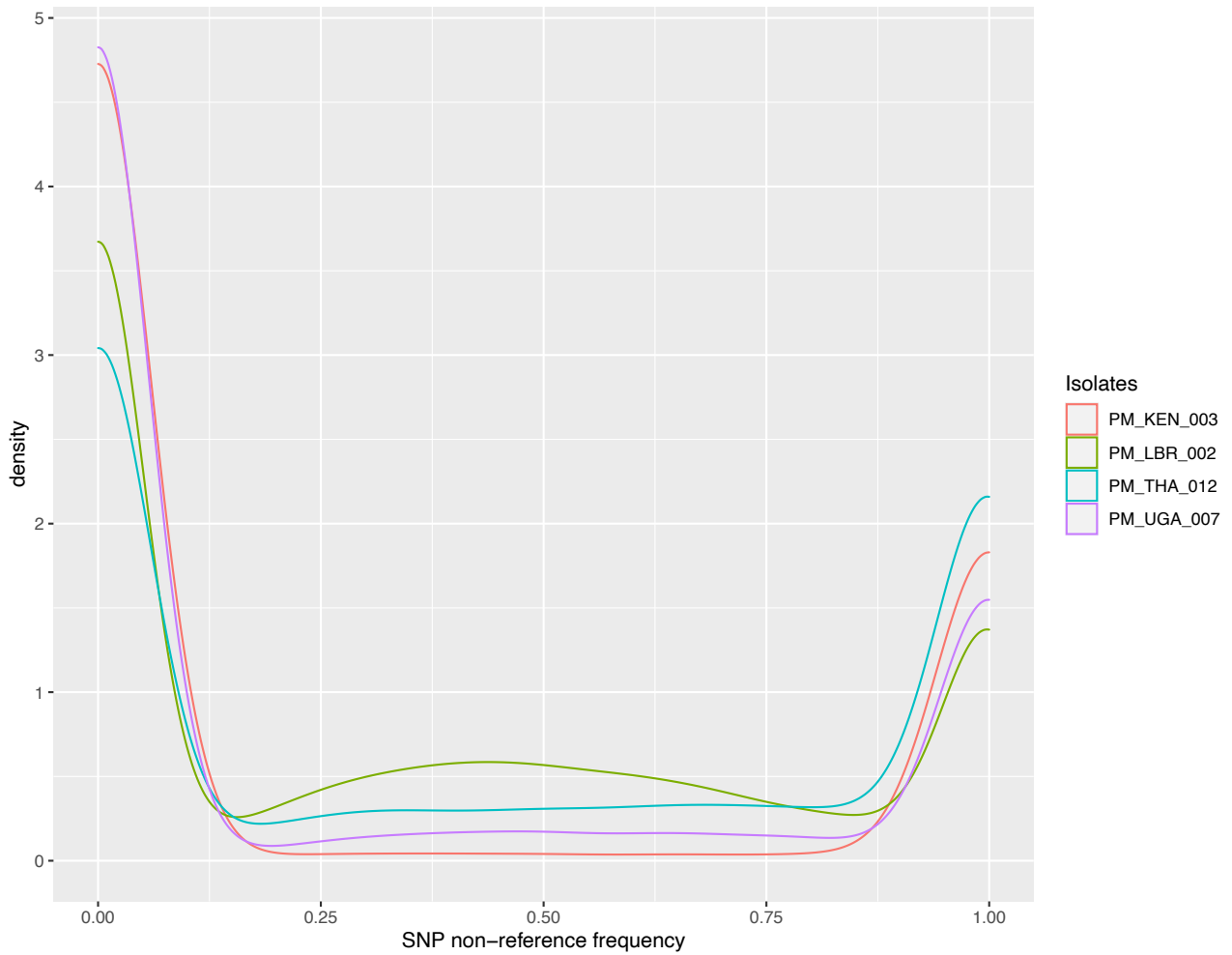




S4 Fig. Determining and excluding hypervariable subtelomeric regions. The average coverage (pink) and number of SNPs (blue) between all 18 samples for each nuclear chromosome. The black dashed horizontal line demonstrates the previously chosen SNP limit per 5 kb window (as defined in **S3 Fig.**). Black dashed lines are placed at the suggested hypervariable region cut-off points, where clusters of windows demonstrating >2.25 SNPs are seen. The midpoints of *Pm-fam* gene families are annotated; *Pm-fam-I* gene positions are denoted by a black plus, whilst *Pm-fam-m* gene positions are denoted with a grey circle. There are cases where *Pm-fam* gene families are located outside of the hypothesised hypervariable regions using the 2.25 SNP limit per window, and in these instances, the terminal or initial position of the *Pm-fam* gene is located with a grey dashed line and this is used as the amended hypervariable cut-off position (see **S3 Table** for coordinates).



S5 Fig. Determining multiplicity of infection (MOI). EstMoi was ran for all sequenced samples, values were obtained for 18 samples (not including PM_KEN_001, PM_LBR_003 and PM_UGA_001 which did not have sufficient SNP density for estMOI) (**S2 Table**). Three samples appear to separate from the main cluster of samples (PM_LBR_002, PM_THA_012 and PM_UGA_007).



S6 Fig. Density plot of non-reference MAF distribution for isolates with MOI>1.

Distribution (density plot) of non-reference minor allele reference (MAF) for isolates with MOI>1 according to estMOI (PM_LBR_002, PM_THA_012, PM_UGA_007). Isolate PM_KEN_003 did not present evidence of MOI>1 and has been included as reference. The density (y-axis) is scaled so that the area under the curve adds up to 1.

A)

```
PmUG01_05034700/1-621 1 - MEDLAD I FDIYA I CACCKVP NOGEGKKNE I FSTKTFRGLNGKGLPWKSNLSLDMKYF S VTTTYVNEKMKYKLYKREKYLEKE I SENSSTVFEN I SLLSS KLVN VVV 109
PF3D7_0417200/1-608 1 MMEQVCDVDF IYA I CA CCKVESKNEGKKEVFNNTYFRGLNGKGLVLPWK C S LDMKYF S AVTTTYVNESKYEKLYKREKYLEKLNKETVDNV - - - - - NDMPNSKLVN VVV 103

PmUG01_05034700/1-621 110 MGR S WWS I PKQYKPLNRR I NVVLSRLLKKEDEKED I F I INNMDOVL L L L KLNYYKCF I IGGA I VYKECLERN I KQ I YFTR I NNVECDVFFPE I DENVFO I TSVSDV 219
PF3D7_0417200/1-608 104 MGR T S WES I PKKFKPLSNRR I NV I LSRLLKKEDEKEDVY I INKVED I V L L GKLNYKCF I I GGSVVYQEF I LEKLL I K I YFTR I NSTYECVDFPE I INENEYQ I ISVSDV 213

PmUG01_05034700/1-621 220 YTSNCTSLDFV I I SRRK KALTOESLPHOSSGSDKGSNTSST I SNGAMSSNT I RGS I TSSSGKGGGGS I FEREYFMGDEEDDLVYFNFNKNNE - YKNAENANDFK I Y 328
PF3D7_0417200/1-608 214 YTSNNTLDFE I I YKKTNNKMLNEQNC I K - - GEEKNN - - DMP LKNDKDTCHMKKLEF - - - - YKNDVKYK I NYENDDDEEEDDFYFNFNKEKEEKNS I HPNDFQ I Y 315

PmUG01_05034700/1-621 329 NSLKF KHP EYQYLS I I YD I I MNGNKQNDRTGVGVL SKFGY I MKFNLDYFP L L TTKLFLRG I IEELLWF I RGETNGNTLLNKNVR I WEANGTREFLDNRKLFHREVND 438
PF3D7_0417200/1-608 316 NSLKYKHP EYQYLN I YD I MNGNKQSDRTGVGVL SKFGY I MKFDLSQYFP L L TTKLFLRG I IEELLWF I RGETNGNTLLNKNVR I WEANGTREFLDNRKLFHREVND 425

PmUG01_05034700/1-621 439 LGP I YGFQWRHFGAEYTNM YNEDKGV DQ LKNI I H L I KNDPTSRRI I LCAWNVKLDQMALPPCH I LCOFYVFDGKLS C I MYQRS CDLGLGVFPN I ASYS I FTHM I AQV 548
PF3D7_0417200/1-608 426 LGP I YGFQWRHFGAEYTNM YNENKGV DQ LKNI I NL I KNDPTSRRI I LCAWNVKLDQMALPPCH I LCOFYVFDGKLS C I MYQRS CDLGLGVFPN I ASYS I FTHM I AQV 535

PmUG01_05034700/1-621 549 CNLQPAQF I H I LGNAHYVNNH I DSKLVQ LNR I PYPFPTLKL NPE I KNI I E DFT I S DFT I QNYVH H K I SMDMAA 621
PF3D7_0417200/1-608 536 CNLQPAQF I H V L GNAHYVNNH I DSKL I Q LNR I PYPFPTLKL NPD I KNI I E DFT I S DFT I QNYVH H E K I SMDMAA 608
```

B)

```
PmUG01_10021600/1-1461 1 M K K D Q N Y K D G S N L S I K D E V E K E L N K K G T L E L F R K I K S O N I S F F L P F K C L P S S H K K L L G I A F I C A T I S G G T L P F F I S V F G D I M K N N M L G E D I N E I I L S L V L I G I F Q F I L 109
PF3D7_0523000/1-1419 1 M G E Q K E K - K O G L N S I K E E V E K E L N K K S T A E L F R K I K N E K I S F F L P F K C L P A Q H R K L L F I S E V C A V L S G G T L P F F I S V F G V I L K N N I L G D I N P I I L S L V S I G L V F I L 108

PmUG01_10021600/1-1461 110 S F I S S F C M D V T T K I L K T L K I E F I K S V Y Q D Q G F H D N N S G S K L S D L E F Y L E Q V N A G I G T K F I F T Y A S S F L G L Y L W S F K N V R L T L C I T C V F P L I Y M C G V I C N K K V R 218
PF3D7_0523000/1-1419 109 S M I S S Y C M D V I T S K I L K T L K L E Y L R S V Y Q D Q G F H D N N P G S K L S D L E F Y L E Q V S S G I G T K F I F T Y A S S F L G L Y I W S L I K N A R L T L C I T C V F P L I Y V G V I C N K K V K 217

PmUG01_10021600/1-1461 219 N K K T S L L Y N N N T M S I I E E A L V G I R T V S Y C G E N I I L R F K L S E K M Y S K Y M L K A N F M E S L H V B L I N G F I L V S Y A F G F W Y G T R I I I S D M R K S I N N D F H G S V I S L L G V 327
PF3D7_0523000/1-1419 218 L N K K T S L L Y N N N T M S I I E E A L M G I R T V A S Y C G E K T I L N K F N L S E T F Y S K Y I L K A N F V E A L H I G L I N G L I L V S Y A F G F W Y G T R I I I S A T N Q Y P N N D F N G A S V I S I L L G V 326

PmUG01_10021600/1-1461 328 L I S M F M L T I I L P N I S E Y M K S L E A T N N L Y E I I N R K P L I E S S G N T G D V L P D I I N K I O F K N V K F H Y D T R K D V E V Y K N L N F T L T E G K T Y A F V G E S G C K S T I L K L I E R I Y D P T O 436
PF3D7_0523000/1-1419 327 L I S M F M L T I I L P N I T E Y M K A L E A T N S L Y E I I N R K P L V E N N - D D G E T L P N I K K I E F K N V R F H Y D T R K D V E I Y K D L S F T L K E G K T Y A F V G E S G C K S T I L K L I E R L Y D P T E 434

PmUG01_10021600/1-1461 437 G D I I I N D S H N L K D V N L K W W R S K I G V V S O D P L L F S N S I K N N I K Y S L S I O D L E Y I Q N G D N D H K M D N D N D E N T N K R C T N K C T K O F N D I I K T N R T S D L L E V K K A Y E S I D D S 545
PF3D7_0523000/1-1419 435 G D I I I N D S H N L K D I N L K W W R S K I G V V S O D P L L F S N S I K N N I K Y S L S L K D L E A M E N Y Y E E N - - - - - T N D T Y - - - E N K N F S L I S N S M T S N E L L E M K K E Y Q T I K D S 530

PmUG01_10021600/1-1461 546 Q V I N V S K K V L I H D F V S A L P D K Y D T M V G S N A S K L S G G Q K O R I S I A R A I I R N P K I L I D E A T S S L D N K S E Y L V O K T I N N L K G N E N R I T I I I A H R L S T I R Y A N T I F V L S N R E 654
PF3D7_0523000/1-1419 531 D V D V S K K V L I H D F V S L P D K Y D T L V G S N A S K L S G G Q K O R I S I A R A I M R N P K I L I D E A T S S L D N K S E Y L V O K T I N N L K G N E N R I T I I I A H R L S T I R Y A N T I F V L S N R E 639

PmUG01_10021600/1-1461 655 N N N M G S H S S I P Y D A A D H S S N N N P I N N L G S Y I E O G T H D S L M K N K G I Y Y S M I N N Q V S S N R S N K G S D D N D S E R K S S V Y K D S D V G N D N E S I Y N N N S M L N N A N G E G A S 763
PF3D7_0523000/1-1419 640 R S D N N N N N N - N D N N N N N N N N N K I N N E G S Y I E O G T H D S L M K N K G I Y H L M I N N Q I S N K S S N G N D N G S D N K S S A Y K D S D T G N D A N M N S L S I H - - - - - E N E N I S 741

PmUG01_10021600/1-1461 764 K N K - D M E L S T N G E E K V S F F K R I F R R K T K G P N N L R V V Y K E M F S Y R K D V F I I V L S I I V A G G L Y P L F A I L Y A K Y V T T L F D F A N L E K S N K Y S L Y I F I I A I A M F I S E T L K N 871
PF3D7_0523000/1-1419 742 N R N K N T A E N E K E E K V P F F K R M F R R K - K K A P N N L R I Y K E I F S Y K K D V T I I F S I L V A G G L Y P V F A L L Y A R Y V S T L F D F A N L E Y S N K Y S I Y I L I A I A M F I S E T L K N 849

PmUG01_10021600/1-1461 872 Y Y N N I I G E K V E K T M K L R L F E N I L H Q E I S F F D H D M N A P L L T H I N R D V H L L K T G L V N N T V I F T H F I V L F L V S M V I S F Y F C P I V A A V L T G T Y F I L M R V F A I R A R L S K S K G 980
PF3D7_0523000/1-1419 850 Y Y N N I I G E K V E K T M K R L F E N I L Y Q E S F F D O Q K N T P G V L S A H I N R D V H L L K T G L V N N I V I F S H F I M L F L V S M V M S F Y F C P I V A A V L T F I Y F I N M R V F A V R A R L T K S K E 958

PmUG01_10021600/1-1461 981 I E E K R N N Y S N T A F V Y N S D D E I F K D P N F L I Q E A F Y N M N T V I T Y G L E E Y F C K L I E K A I D Y S N K G Q K R I I V N S I L W G F S Q S A Q I F V N S F A Y W F G S L L I R R G T I E V D D F M K 1089
PF3D7_0523000/1-1419 959 I E K K E N M - - S G G V F A F S D D E M F K D R S F L I Q E A F Y N M H T V I N Y G L E O Y F C N L I E K A I D Y K N G Q K R R I I V N A A L W G F Q S A Q L I I S F A Y W F G S F L I K R G T I L V D D F M K 1065

PmUG01_10021600/1-1461 1090 S L F T F L T G S Y A G K L M S L K G D S E S A K A S F E K Y Y P L I M R K S N I D V R D V G G I R I K N I D D I K G K I E I K D V N F R Y I S R P N V P I Y K D L S F T C E S K K T T A I V G E T G S G K S T V M S I 1198
PF3D7_0523000/1-1419 1066 S L F T F I T G S Y A G K L M S L K G D S E N A K L S F E K Y P L M I R K S N I D V R D O G G I R I N K - N L I K G K V D I K D V N F R Y I S R P N V P I Y K N L S F T C D S K K T T A I V G E T G S G K S T F M N L 1173

PmUG01_10021600/1-1461 1199 L M R F Y D I K N D H I F K N D T I K N K S N D I G V T E E G A D M E G G G I P M N A D E F N V G S K E R Y V D D D D N N K Y T L F K N G S I L L D D I N D Y N L K D L R S L F S I V S Q E P M L F N M S I 1307
PF3D7_0523000/1-1419 1174 L L R F Y D L K N D H I I L K N D M T N F O D Y Q N - - - - - N N N N S L V L K N V N - - - - - E F S N Q S G S A E D Y T V E N N N G E I L L D D I N N C D Y N L R D L R L N F S I V S Q E P M L F N M S I 1265

PmUG01_10021600/1-1461 1308 Y E N I K F G K D D A T L E D V K R V C K F A A I D E F I E S L P N K Y D T N V G P Y G K L S G G Q K O R I A I A R A L L R E P K I L L D E A T S S L D S N S E K L I E K T I V D I K D R G D K T I I T I A H R I A S 1416
PF3D7_0523000/1-1419 1266 Y E N I K F G R E D A T L E E D V K R V S K F A A I D E F I E S L P N K Y D T N V G P Y G S L S G G Q K O R I A I A R A L L R E P K I L L D E A T S S L D S N S E K L I E K T I V D I K D K A D K T I I T I A H R I A S 1374

PmUG01_10021600/1-1461 1417 I K R S D R I V V L N N P D R L G S Y V Q S G T H E L L A E Q D G I Y K K Y V K L A K 1461
PF3D7_0523000/1-1419 1375 I K R S D K I V V F L N N P D R N G T F V Q S H G T H D E L L S A Q D G I Y K K Y V K L A K 1419
```

S7 Fig. Substitutions within *P. malariae* orthologs of drug resistance-associated loci in *P. falciparum*.

Amino acid sequences aligned using Clustal Omega. Regions that are conserved between the two positions are highlighted in grey. Positions associated with drug resistance mutations within *P. falciparum* are highlighted in red. Predicted amino acid substitutions found in our *P. malariae* sample set are highlighted in blue. **A)** *PMDHFR* and *PFDHFR*. Both positions 57-58 and 114 within *PMDHFR* lie close to amino acid alterations associated with drug resistance in *P. falciparum* (C59R and S108N/T). **B)** *PMMDR1* and *PFMDR1*. The amino acid substitution at position 1,063 within PmUG01_10021600.1 lies in close proximity to the N1042D substitution associated with quinine resistance.

S1 Table. Pmset1 primers. The sequences primers within Pmset1 are demonstrated. The asterisk in the primer sequence denotes the position of a phosphorothioate bond which inhibits primer degradation. Binding sites within the human genome were calculated using the human GRCh37 genome sequence available at <http://grch37.ensembl.org> and binding sites within the *P. malariae* genome were calculated using the PmUG01 reference genome available at <https://plasmodb.org>.

Primer set	Primer sequence	Binding sites in human genome	Binding sites in <i>P. malariae</i> genome
Pmset1	TATGTATA*T*T	6054	3057
	TTATTC*G*T	3475	3040
	TTCGTT*A*T	4381	3088
	TTTTTA*C*G	3991	3863
	TATTC*G*T	3961	3863
	Total Binding Sites	21862	16521

S2 Table. Sequencing statistics for all samples. Sequence analysis data for all samples. Samples highlighted in grey are comparative samples before SWGA, this is also indicated in the 'Amplification' column. Samples that were excluded due to poor quality and not used in downstream population genetics analysis are labelled with an (E). Accession codes are given for each sample and can be found using the project code (PRJEB37746). Multiplicity of infection (MOI) was calculated using estMOI and the percentage of reads supporting a MOI>1 are given.

Sample	Accession code	Country	Amplified DNA	Pm CT-value (qPCR data)	Parasites/ul (% Parasitaemia)	Total reads	Mapped reads	% Mapped reads	% Genome coverage ≥ 5	% Genome coverage ≥ 10	Mean coverage	Mean genic coverage	Median genic coverage	Mean intergenic coverage	Median intergenic coverage	Mean mitochondria coverage	Unfiltered SNPs	Core SNPs filtered	% Genome with MOI>1
PM_THA_001 (E)	/	Thailand	No_SWGA	31.74	256 (0.0064)	6264676	1005761	16.05	0.86	0.12	0.45	0.44	0	0.47	0	/	3	/	/
PM_THA_001	ERS4423439	Thailand	SWGA	31.74	256 (0.0064)	7174584	3154695	43.97	44.17	25.28	8.57	6	3	11	4	23.78	41962	3256	0
PM_THA_002 (E)	/	Thailand	No_SWGA	27.73	576 (0.0144)	5073136	419922	8.28	5.01	0.45	1.25	1.66	1	0.89	0	/	49	/	/
PM_THA_002	ERS4423440	Thailand	SWGA	27.73	576 (0.0144)	8036914	4250708	52.89	64.06	36.16	12.59	16	7	10	6	338.52	43237	6089	0
PM_THA_003	ERS4423441	Thailand	SWGA	30.59	8096 (0.2024)	7852281	4215247	53.68	63.46	40.97	14.63	13	6	16	8	3789.74	45869	6567	2.10
PM_THA_005	ERS4423442	Thailand	SWGA	23.60	6464 (0.1616)	8028153	5735625	71.44	88.31	72.65	22.68	23	18	22	15	1533.33	57182	9998	2.36
PM_THA_006	ERS4423443	Thailand	SWGA	23.72	5312 (0.1328)	5584655	5130545	91.87	75.09	53.23	19.13	14	9	24	12	238.56	56707	6826	0
PM_THA_009a* (E)	ERS4423444	Thailand	SWGA	31.20	64 (0.0016)	3684878	2765795	75.06	44.05	25.70	9.28	7	3	11	4	/	43853	/	0
PM_THA_009b* (E)	ERS4423445	Thailand	SWGA	31.20	64 (0.0016)	2868636	2327144	81.12	27.32	15.84	6.41	4	1	8	2	/	37417	/	0
PM_THA_009_com	ERS4423446	Thailand	SWGA	31.20	64 (0.0016)	6553515	5092953	77.71	52.44	34.66	15.69	11	4	20	6	294.51	48212	4698	0
PM_THA_010	ERS4423447	Thailand	SWGA	23.28	1648 (0.0412)	6934812	4135980	59.64	83.89	62.05	16.49	16	13	17	12	884.78	43804	9042	0

PM_THA_011	ERS4423448	Thailand	SWGA	25.20	16 (0.0004)	6179776	2257408	36.53	63.10	31.03	8.47	8	7	9	6	164.40	43804	5629	0
PM_THA_012	ERS4423449	Thailand	SWGA	24.42	544 (0.0136)	6909366	3264421	47.25	78.03	50.69	12.71	13	10	13	9	414.93	52348	8147	17.46
PM_KEN_001	ERS4423432	Kenya	SWGA	/	/	5709031	3454616	60.51	72.05	44.58	12.03	11	8	13	9	59.60	50899	4820	/
PM_KEN_003	ERS4423433	Kenya	SWGA	/	/	6744405	2186998	32.43	53.11	24.02	7.44	7	5	8	5	28.76	44678	3400	0
PM_LBR_002	ERS4423434	Liberia	SWGA	/	/	7180349	4242520	59.09	86.46	66.06	17.10	17	14	17	13	309.94	63484	8474	7.88
PM_LBR_003 (E)	ERS4423435	Liberia	SWGA	/	/	9338062	5627043	60.26	0.07	0.03	0.19	0	0	0	0	/	25	/	/
PM_SDN_001	ERS4423436	Sudan	SWGA	/	/	5369713	4461837	83.09	72.52	50.32	16.96	13	8	20	11	419.30	54943	5154	0
PM_SLE_002	ERS4423437	Sierra Leone	SWGA	/	/	8039293	4286570	53.32	84.15	62.52	16.46	17	13	16	12	258.93	53739	7034	5.61
PM_SLE_004	ERS4423438	Sierra Leone	SWGA	/	/	6933075	1513093	21.82	44.96	14.67	5.35	6	4	5	4	184.48	28204	2554	0
PM_UGA_001	ERS4423451	Uganda	SWGA	/	/	2846843	1666685	58.55	46.56	17.59	5.90	5	4	6	4	50.56	37419	2534	/
PM_UGA_003	ERS4423452	Uganda	SWGA	/	/	6500133	3316826	51.03	81.84	53.76	13.22	13	11	13	10	373.45	53571	6525	3.23
PM_UGA_007	ERS4423453	Uganda	SWGA	/	/	6673599	2481168	37.18	58.43	32.44	9.48	9	6	10	6	87.20	48414	3836	25.64

S3 Table. Identification of core nuclear genome positions. Not all chromosomes have annotated *Pm-fam* genes or hypervariable regions at both ends of the chromosomes, these chromosomes are indicated with an asterisk.

Chromosome	Start	End	Classification
PmUG01_01_v1	0	194999	Hypervariable
PmUG01_01_v1	195000	1125000	Core
PmUG01_01_v1	1125001	1381517	Hypervariable
PmUG01_02_v1*	0	59999	Hypervariable
PmUG01_02_v1*	60000	813519	Core
PmUG01_03_v1	0	89999	Hypervariable
PmUG01_03_v1	90000	1105000	Core
PmUG01_03_v1	1105001	1341933	Hypervariable
PmUG01_04_v1*	0	143680	Hypervariable
PmUG01_04_v1*	143681	1076824	Core
PmUG01_05_v1	0	399999	Hypervariable
PmUG01_05_v1	400000	1665000	Core
PmUG01_05_v1	1665001	1887011	Hypervariable
PmUG01_06_v1*	0	705000	Core
PmUG01_06_v1*	705001	862289	Hypervariable
PmUG01_07_v1	0	221927	Hypervariable
PmUG01_07_v1	221928	1875000	Core
PmUG01_07_v1	1875001	2080590	Hypervariable
PmUG01_08_v1	0	364999	Hypervariable
PmUG01_08_v1	365000	2050000	Core
PmUG01_08_v1	2050000	2818517	Hypervariable
PmUG01_09_v1*	0	164999	Hypervariable
PmUG01_09_v1*	165000	2312276	Core
PmUG01_10_v1	0	199999	Hypervariable
PmUG01_10_v1	200000	1854000	Core
PmUG01_10_v1	1854001	2219074	Hypervariable
PmUG01_11_v1	0	344999	Hypervariable
PmUG01_11_v1	345000	2644881	Core
PmUG01_11_v1	2644882	2721161	Hypervariable
PmUG01_12_v1*	0	89999	Hypervariable
PmUG01_12_v1*	90000	3436769	Core
PmUG01_13_v1	0	42699	Hypervariable
PmUG01_13_v1	42700	2410000	Core
PmUG01_13_v1	2410001	3023685	Hypervariable
PmUG01_14_v1	0	350186	Hypervariable
PmUG01_14_v1	350187	3564280	Core

S4 Table. Pairwise genome-wide nucleotide diversity (π ; $\times 10^{-4}$). Nucleotide diversity values are coloured to demonstrate samples with high or low diversity in comparison to other samples. Low diversity is shaded in yellow and pale orange, increasing to darker orange for high diversity (yellow for 0 to 1, pale orange for 1 to 2, medium orange for 1 to 2, and dark orange for 2 to 4; all $\times 10^{-4}$).

Region		Thailand									East Africa						West Africa		
	Samples	PM_THA_001	PM_THA_002	PM_THA_003	PM_THA_005	PM_THA_006	PM_THA_009com	PM_THA_010	PM_THA_011	PM_THA_012	PM_KEN_001	PM_KEN_003	PM_UGA_001	PM_UGA_003	PM_UGA_007	PM_SDN_001	PM_LBR_002	PM_SLE_002	PM_SLE_004
Thailand	PM_THA_001	/	2.1	2.23	2.29	2.04	1.75	1.99	1.64	1.87	2.75	2.79	2.35	2.91	2.64	2.72	3	3.05	2.41
	PM_THA_002	2.1	/	0.36	2.52	2.16	2.13	2.32	2.04	1.94	3.06	3.1	2.59	3.19	2.91	3.08	3.26	3.33	2.75
	PM_THA_003	2.23	0.36	/	2.52	2.25	2.18	2.34	2.1	1.98	3.12	3.12	2.67	3.21	2.96	3.12	3.25	3.3	2.74
	PM_THA_005	2.29	2.52	2.52	/	2.38	2.21	2.49	2.18	2.1	3.25	3.26	2.95	3.53	3.17	3.37	3.64	3.68	3.06
	PM_THA_006	2.04	2.16	2.25	2.38	/	1.99	2.38	2.17	2.17	2.91	2.98	2.55	3.05	2.85	2.9	3.26	3.27	2.68
	PM_THA_009com	1.75	2.13	2.18	2.21	1.99	/	2.13	1.4	2.07	2.8	2.88	2.43	2.98	2.82	2.83	3.11	3.14	2.62
	PM_THA_010	1.99	2.32	2.34	2.49	2.38	2.13	/	1.94	2.23	3.22	3.24	2.78	3.39	3.11	3.26	3.44	3.55	2.89
	PM_THA_011	1.64	2.04	2.1	2.18	2.17	1.4	1.94	/	2.03	2.91	2.92	2.55	3.07	2.78	2.89	3.08	3.26	2.57
	PM_THA_012	1.87	1.94	1.98	2.1	2.17	2.07	2.23	2.03	/	3.08	3.06	2.67	3.2	2.91	3.09	3.23	3.38	2.72
East Africa	PM_KEN_001	2.75	3.06	3.12	3.25	2.91	2.8	3.22	2.91	3.08	/	0.27	1.87	2.45	2.09	2.26	2.48	2.43	1.89
	PM_KEN_003	2.79	3.1	3.12	3.26	2.98	2.88	3.24	2.92	3.06	0.27	/	1.93	2.4	2.08	2.33	2.5	2.47	1.89
	PM_UGA_001	2.35	2.59	2.67	2.95	2.55	2.43	2.78	2.55	2.67	1.87	1.93	/	2.07	1.86	1.94	2.11	2.17	1.69
	PM_UGA_003	2.91	3.19	3.21	3.53	3.05	2.98	3.39	3.07	3.2	2.45	2.4	2.07	/	2.2	2.45	2.71	2.78	2.19
	PM_UGA_007	2.64	2.91	2.96	3.17	2.85	2.82	3.11	2.78	2.91	2.09	2.08	1.86	2.2	/	2.15	2.34	2.38	1.88
	PM_SDN_001	2.72	3.08	3.12	3.37	2.9	2.83	3.26	2.89	3.09	2.26	2.33	1.94	2.45	2.15	/	2.47	2.59	2.01
West Africa	PM_LBR_002	3	3.26	3.25	3.64	3.26	3.11	3.44	3.08	3.23	2.48	2.5	2.11	2.71	2.34	2.47	/	2.71	2.09
	PM_SLE_002	3.05	3.33	3.3	3.68	3.27	3.14	3.55	3.26	3.38	2.43	2.47	2.17	2.78	2.38	2.59	2.71	/	2.06
	PM_SLE_004	2.41	2.75	2.74	3.06	2.68	2.62	2.89	2.57	2.72	1.89	1.89	1.69	2.19	1.88	2.01	2.09	2.06	/

S5 Table. Drug resistance associated genes in *P. falciparum* and their known orthologs in *P. malariae*.

Gene	<i>P. falciparum</i> (3D7)		<i>P. malariae</i> (UG01)	
	Gene ID	Paper code	Gene ID	Paper code
<i>crt</i>	PF3D7_070900	<i>Pfcrt</i>	PmUG01_01020700	<i>Pmcrt</i>
<i>dhfr-ts</i>	PF3D7_0417200	<i>Pfdhfr</i>	PmUG01_05034700	<i>Pmdhfr</i>
<i>pppk-dhps</i>	PF3D7_0810800	<i>Pfdhps</i>	PmUG01_14045500	<i>Pmdhps</i>
<i>mdr1</i>	PF3D7_052300	<i>Pfmdr1</i>	PmUG01_10021600	<i>Pmmdr1</i>
<i>kelch13</i>	PF3D7_1343700	<i>Pfk13</i>	PmUG01_12021200	<i>Pmk13</i>

S6 Table. WGS data from mixed infections. Four samples with mixed infections (multiple species) were sequenced and mapped to their respective reference genomes. The proportion of each parasite present was analysed using qPCR and the resulting cycle thresholds (CTs) are given. The following reference genomes are used: *P. malariae* (PmUG01), *P. vivax* (PvP01, https://plasmodb.org/common/downloads/Current_Release/PvivaxP01/fasta/data/) *P. ovale curtisi* (PocGh01, https://plasmodb.org/common/downloads/Current_Release/PovalecurtisiGH01/fasta/data/), *P. ovale wallikeri* (PowCR01 <https://www.ebi.ac.uk/ena/data/view/ERS1452913>).

Sample	Mapped_to	Pf_CT	Po_CT	Pm_C_RT	Pv_C_RT	Parasites/ul (%Parasitemia)	Total_reads	Map_reads	%_Mapped_reads	%_Genome_cov5	%_Genome_cov10
PM_THA_004	PmUG01	/	/	23.48	31.02	1520 (0.038)	7644777	4619930	60.433	85.004	66.919
PM_THA_004	PvP01	/	/	23.48	31.02	1520 (0.038)	7644777	37351	0.489	0.015	0.010
PM_THA_008	PmUG01	/	29.33	31.12	29.26	192 (0.0048)	3679905	305249	3.993	0.072	0.023
PM_THA_008	PocGH01	/	29.33	31.12	29.26	192 (0.0048)	3679905	1540021	41.849	25.825	15.677
PM_THA_008	PowENA	/	29.33	31.12	29.26	192 (0.0048)	3679905	117162	3.184	0.169	0.100
PM_THA_008	PvP01	/	29.33	31.12	29.26	192 (0.0048)	3679905	242561	6.592	0.046	0.024
PM_THA_013	PmUG01	/	33.43	26.31	/	4560 (0.114)	7533858	3928080	52.139	83.985	61.240
PM_THA_013	PocGH01	/	33.43	26.31	/	4560 (0.114)	7533858	106192	1.410	0.035	0.014
PM_THA_013	PowENA	/	33.43	26.31	/	4560 (0.114)	7533858	48638	0.646	0.034	0.018
PM_THA_020	PmUG01	/	26.72	28.33	26.70	4880 (0.122)	6917577	236151	3.414	0.048	0.020
PM_THA_020	PocGH01	/	26.72	28.33	26.70	4880 (0.122)	6917577	2605168	37.660	23.871	20.314
PM_THA_020	PowENA	/	26.72	28.33	26.70	4880 (0.122)	6917577	895565	12.946	28.435	12.095
PM_THA_020	PvP01	/	26.72	28.33	26.70	4880 (0.122)	6917577	391155	5.655	0.046	0.038