

WGS data preparation

As in (Cerqueira et al. 2017), reads were aligned to the *P. falciparum* 3D7 v3 reference genome, genotypes called and sites filtered. Sites removed included heterozygous sites, indels, sites with $QUAL < 60$, $GQ < 30$, polymorphic sites located in pericentromeric, subtelomeric and hypervariable regions, and sites occurring in genes belonging to large antigenic gene families.

In addition to the filters described above, we removed 121 sites with reference or alternative allele assignments indicating potential indels, multi-allelic sites, sites with no non-reference allele counts, and sites lacking genotype calls in 20% or more of the isolates.

From the ensuing text file, which contained a total of 42879 SNPs and 180 isolates, we removed mitochondrial and apicoplast SNPs (labelled chrom 15 and 16, respectively), monomorphic non-reference SNPs, and 2 isolates from Mae Ramat.

The final text and RData files used for all downstream analyses (WGS.txt, Data_store_WGS.RData) contain data for 34911 polymorphic biallelic SNPs (Fig 1) and 178 isolates distributed over sites are years (Table 1).

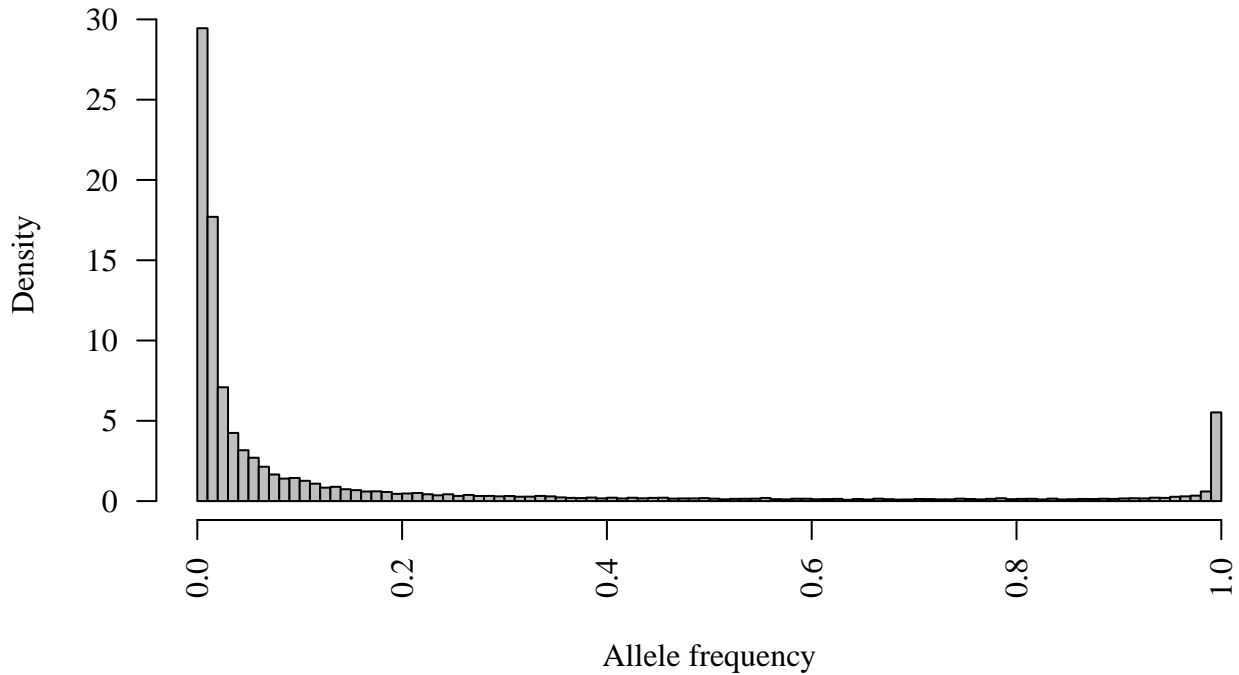


Figure 1: Allele frequency spectrum.

Table 1: Isolate counts over sites and years

	Early	Late	Middle	VeryLate	Total
MKK	0	0	0	4	4
MKT	0	0	0	16	16
MLA	33	2	12	8	55
WPA	3	48	36	16	103

References

Cerqueira, Gustavo C., Ian H. Cheeseman, Steve F. Schaffner, Shalini Nair, Marina McDew-White, Aung Pyae Phy, Elizabeth A. Ashley, et al. 2017. "Longitudinal genomic surveillance of *Plasmodium falciparum* malaria parasites reveals complex genomic architecture of emerging artemisinin resistance." *Genome Biology* 18 (78). doi:10.1186/s13059-017-1204-4.