
Table S6: **Computational time for PCA computation.** Computation time, in seconds, to perform PCA for different number of simulated sites (S) and sample size (N) for each of the 3 populations, at 2X sequencing depth. 'Genotype p.p.' includes computing genotype posterior probabilities. 'Frequency p.p.' includes estimating the SFS and computation sample allele frequency posterior probabilities and it is required to perform PCA as in 'w/o GC (2)'. 'GC' refers to estimate C from called genotypes. 'w/o GC (1)' and 'w/o GC (2)' estimate C without calling genotypes. 'w/o GC (2)' also weights each site by its probability of being variable. Computations refer to estimation of the reduced matrix C as in Equation 18 and do not include the eigenvector decomposition. Calculations were run on a Unix desktop machine, Intel Core 2 Duo CPU E8600 @ 3.33GHz x 2. Maximum memory usage was $< 0.1G$.

S	N	Simulation	Genotype p.p.	Frequency p.p.	GC	w/o GC (1)	w/o GC (2)
10k	20	4	1	76	< 1	< 1	< 1
10k	40	8	3	143	2	3	2
50k	20	20	6	408	2	3	3
50k	40	41	12	561	11	17	18