Supplementary Material

Supplementary material includes data that support and expand some of the interpretations and conclusions drawn in the main text.

Table S1. The computational time (seconds) of the estimated haplotype frequencies across statistical methods and number of SNPs

	Gibbs	R-EM	PLINK	fastPHASE	PHASE	MACH1	IMPUTE2	BEAGLE
2 SNPs	4.68	0.02	0.61	3.73	2.05	2.66	0.88	1.77
3 SNPs	6.57	0.02	0.60	7.25	3.54	4.42	1.23	1.88
4 SNPs	9.53	0.04	0.57	10.41	6.26	5.71	1.57	1.92
5 SNPs	16.84	0.07	0.59	15.07	10.73	7.42	1.95	2.00
6 SNPs	20.73	0.11	0.59	22.64	22.58	10.45	2.73	2.27
7 SNPs	30.56	0.16	0.58	23.90	37.19	11.87	2.88	2.14
8 SNPs	37.83	0.20	0.55	27.30	70.42	13.68	3.45	2.20
9 SNPs	65.25	0.25	0.58	31.22	135.13	15.57	4.00	2.30
10 SNPs	88.30	0.40	0.58	30.84	208.76	16.64	4.02	2.27

EM, EM algorithm; Gibbs, Gibbs sampler; SNPs, Single Nucleotide Polymorphisms

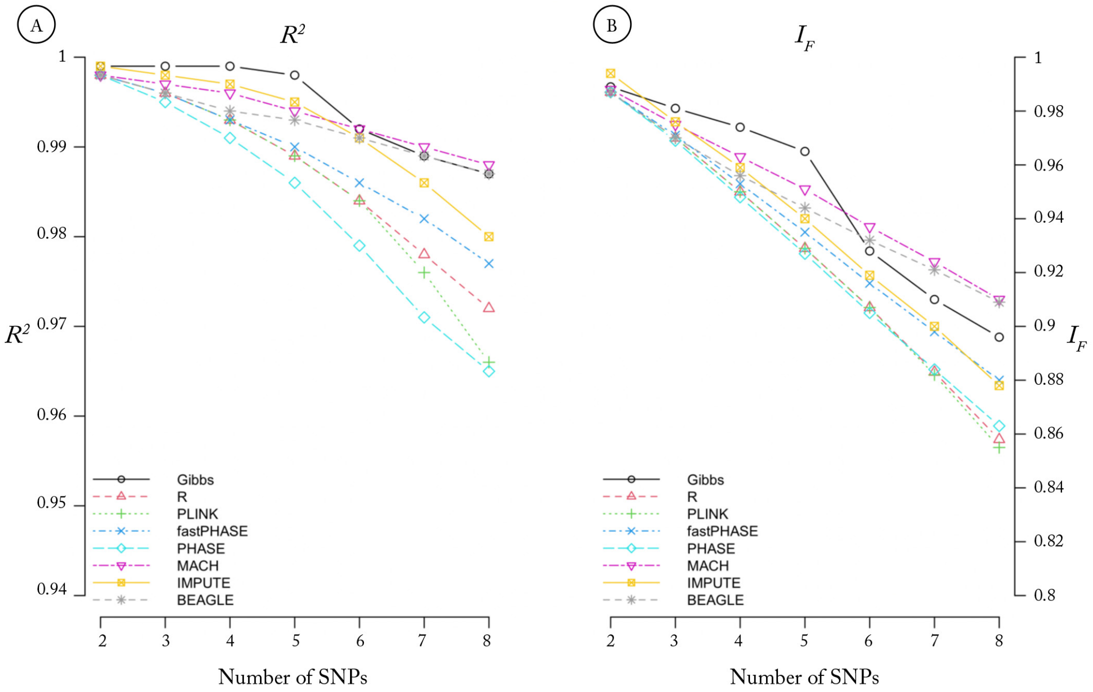

Figure S1. The correlation (R^2) and similarity index (IF) of the estimated haplotype frequencies with simulated population haplotype frequencies across statistical methods among 1000 individuals
EM, EM algorithm; Gibbs, Gibbs sampler; SNPs, Single Nucleotide Polymorphisms.

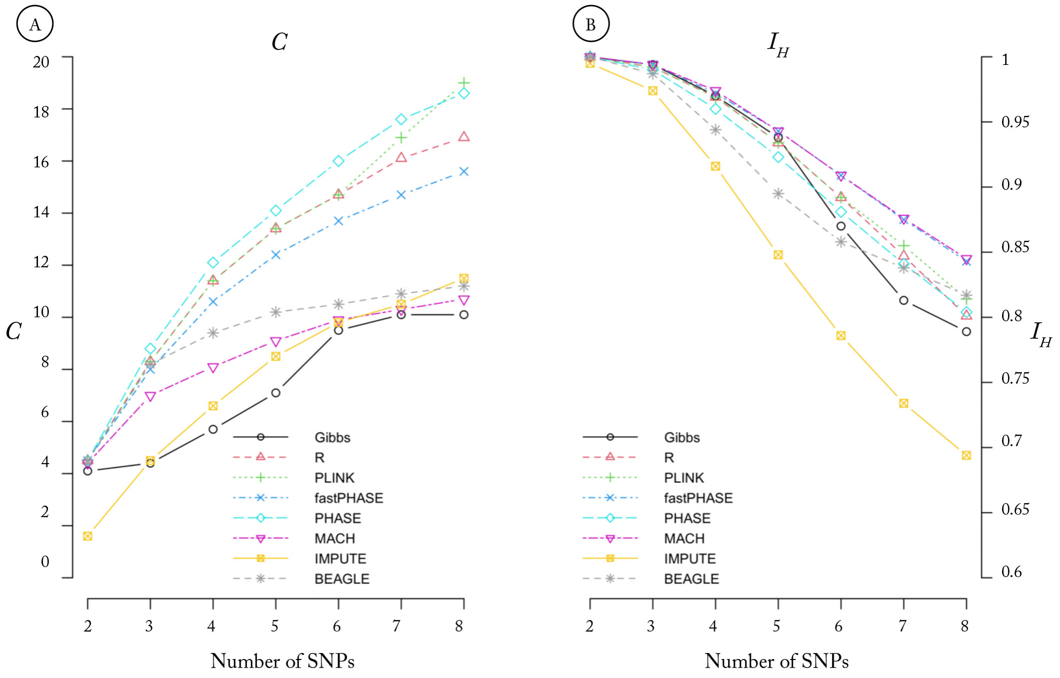

Figure S2. The average change coefficient (*C*) and haplotype identification (IH) of the estimated haplotype frequencies with simulated population haplotype frequencies across statistical methods among 1000 individuals.
Change coefficient (*C*) for haplotype frequency >5%; Gibbs, Gibbs sampler; SNPs, Single Nucleotide Polymorphisms.

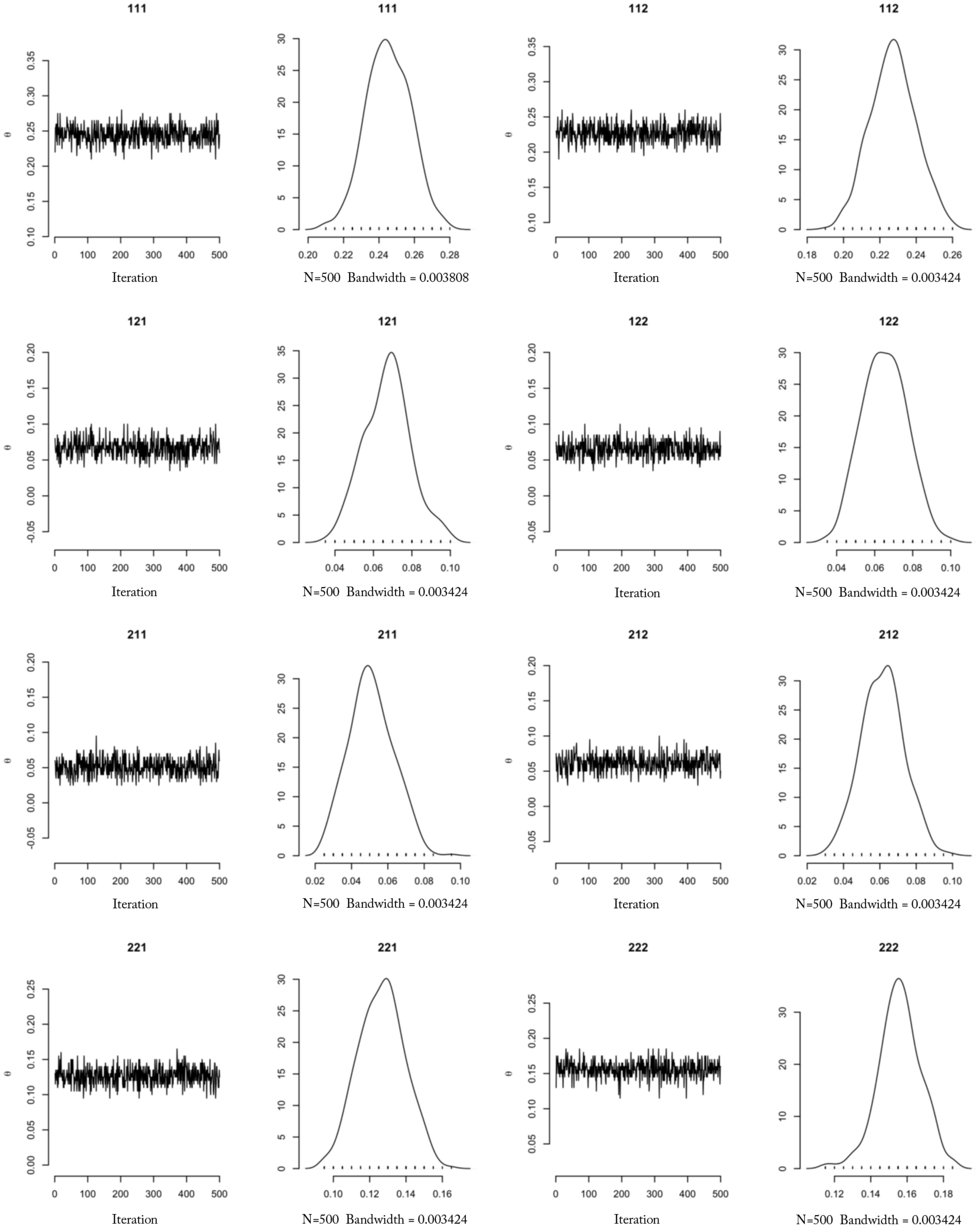

Figure S3. A trace plot and density plots of the iteration number against the value of the haplotype frequencies at 500 iterations.

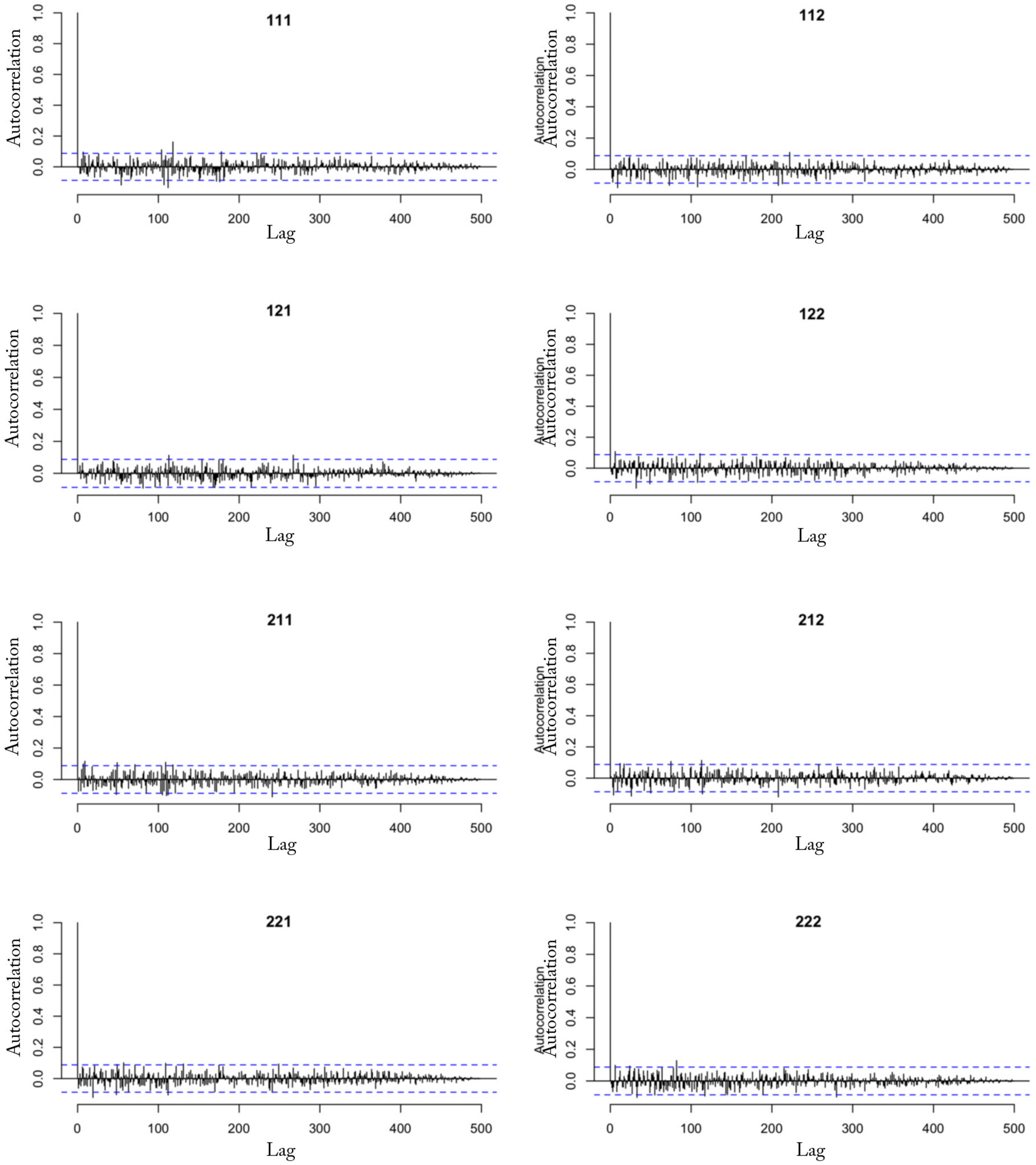

Figure S4. Autocorrelation plots to assess the autocorrelations between the haplotype frequencies of Markov chain at 500 iterations

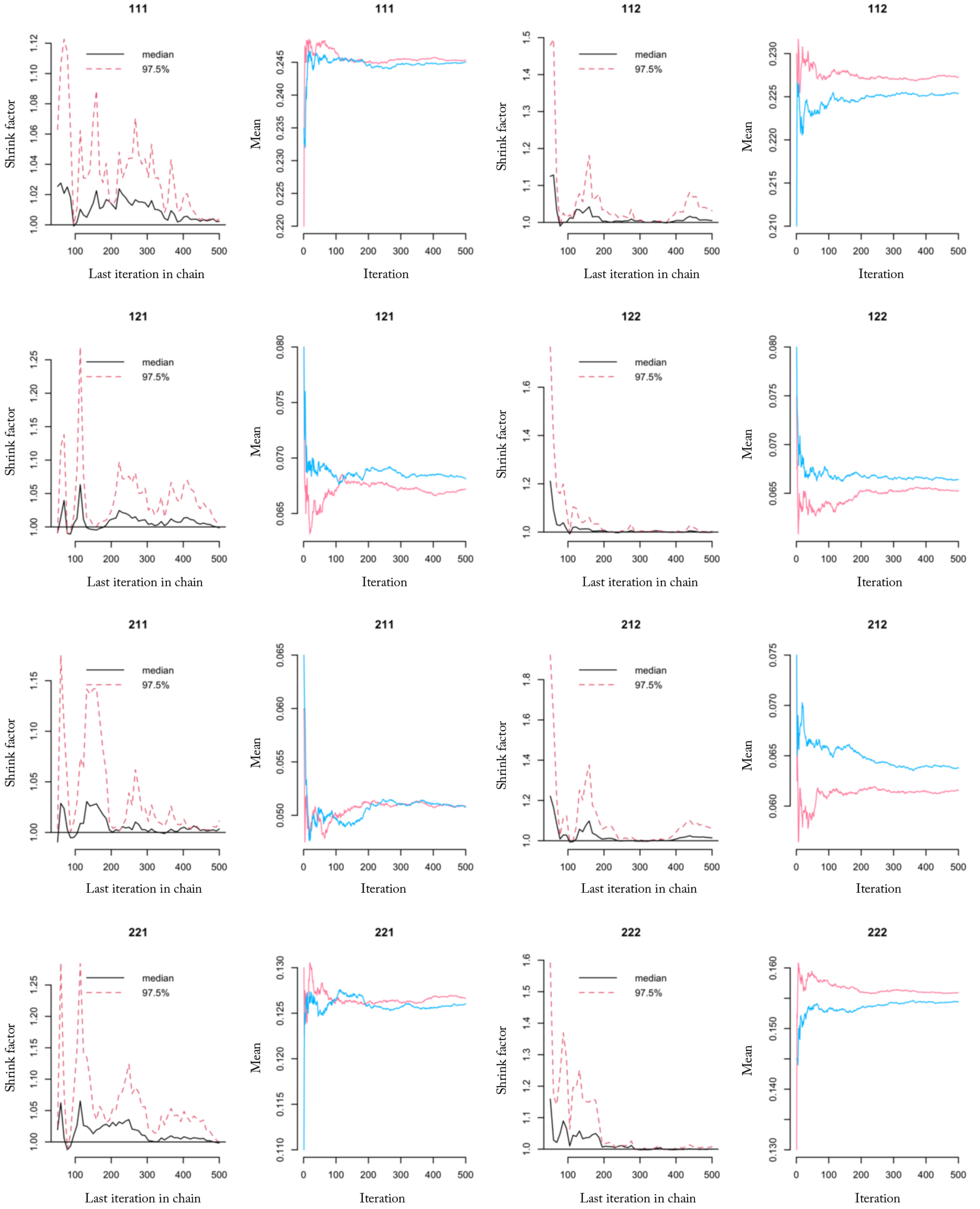

Figure S5. Gelman Rubin Brooks plot (shrink factor as the number of iterations) and mean plot (mean as the number of iterations) with 2 chains at 500 iterations.

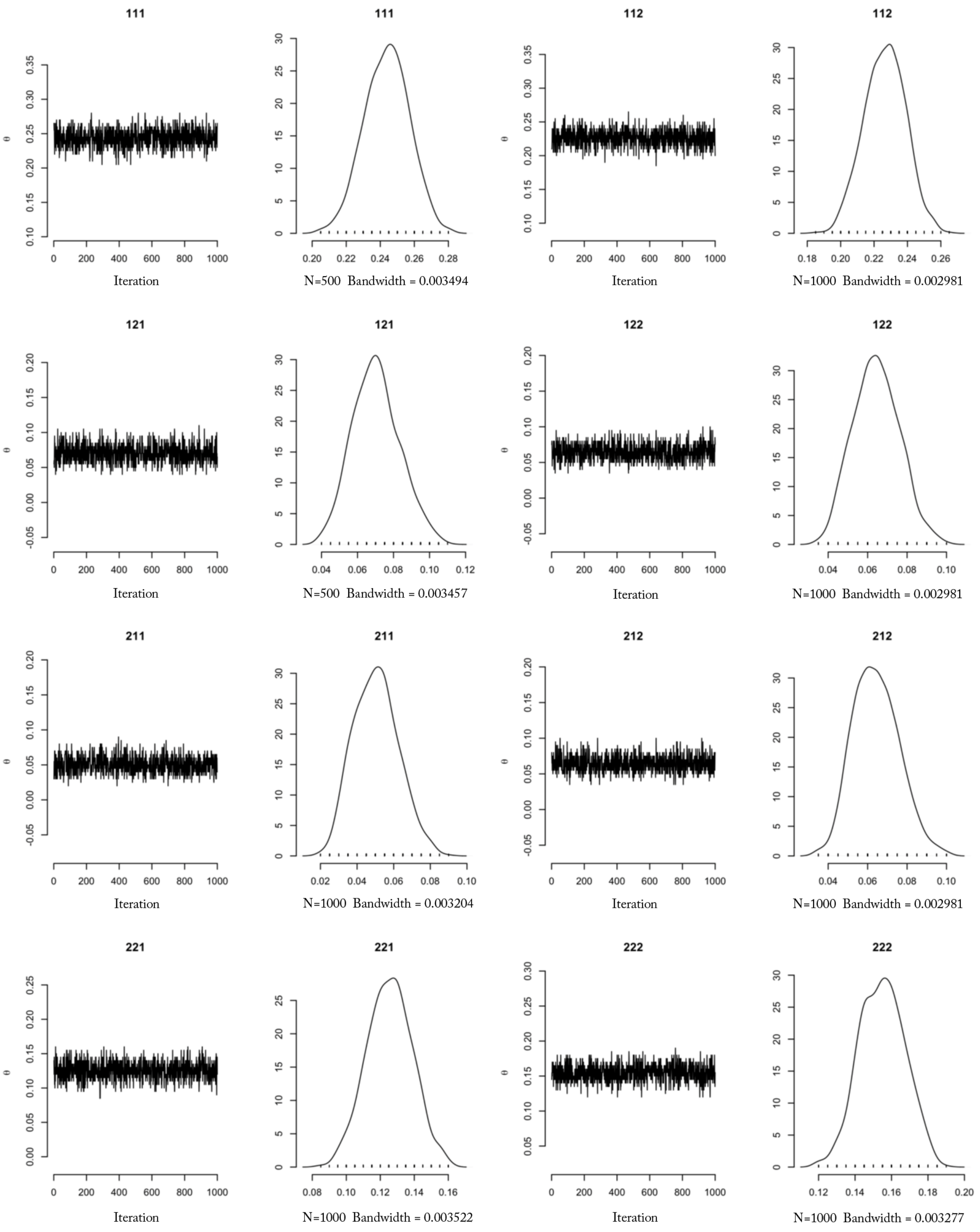

Figure S6. A trace plot and density plots of the iteration number against the value of the haplotype frequencies at 1000 iterations

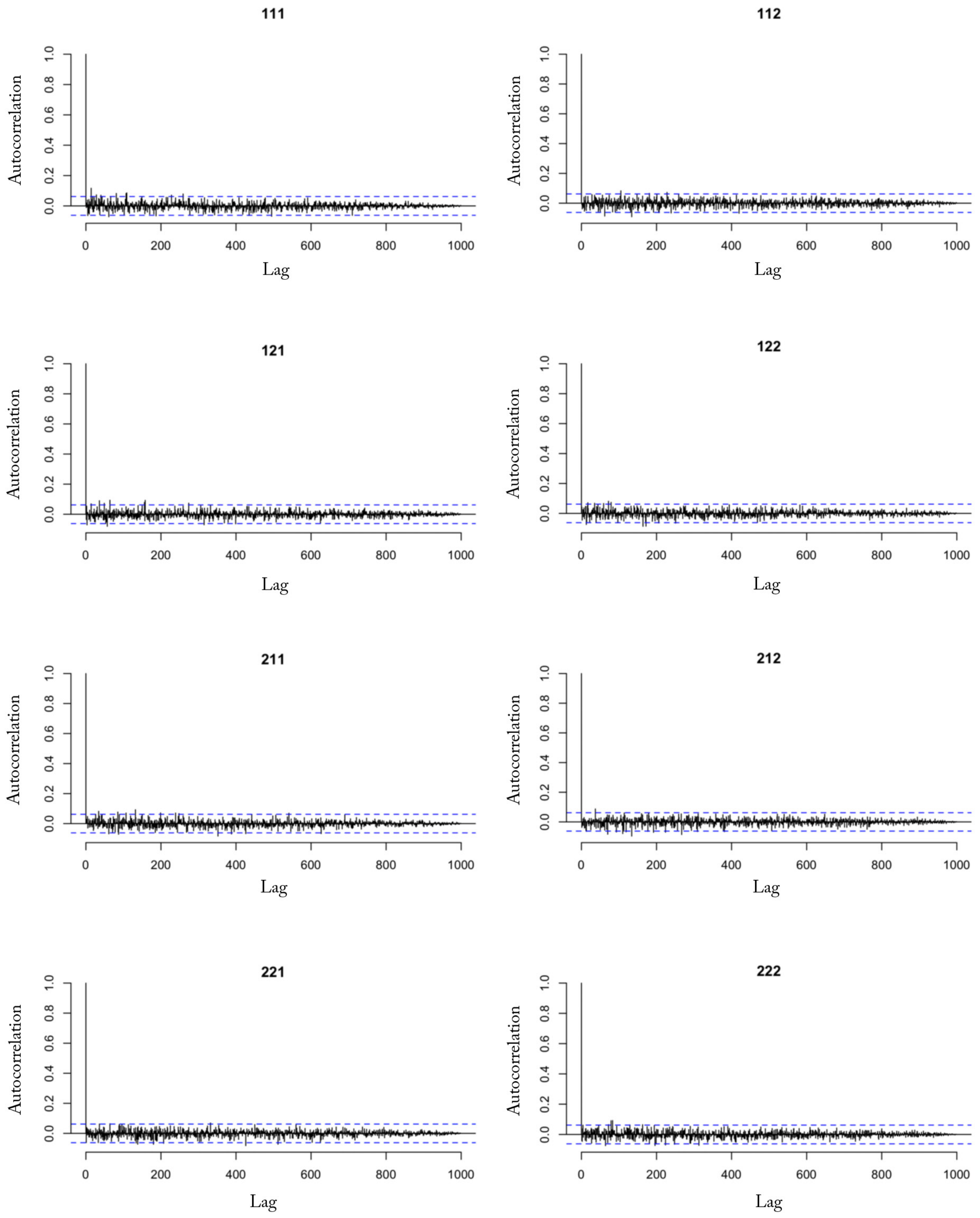

Figure S7. Autocorrelation plots to assess the autocorrelations between the haplotype frequencies of Markov chain at 1000 iterations.

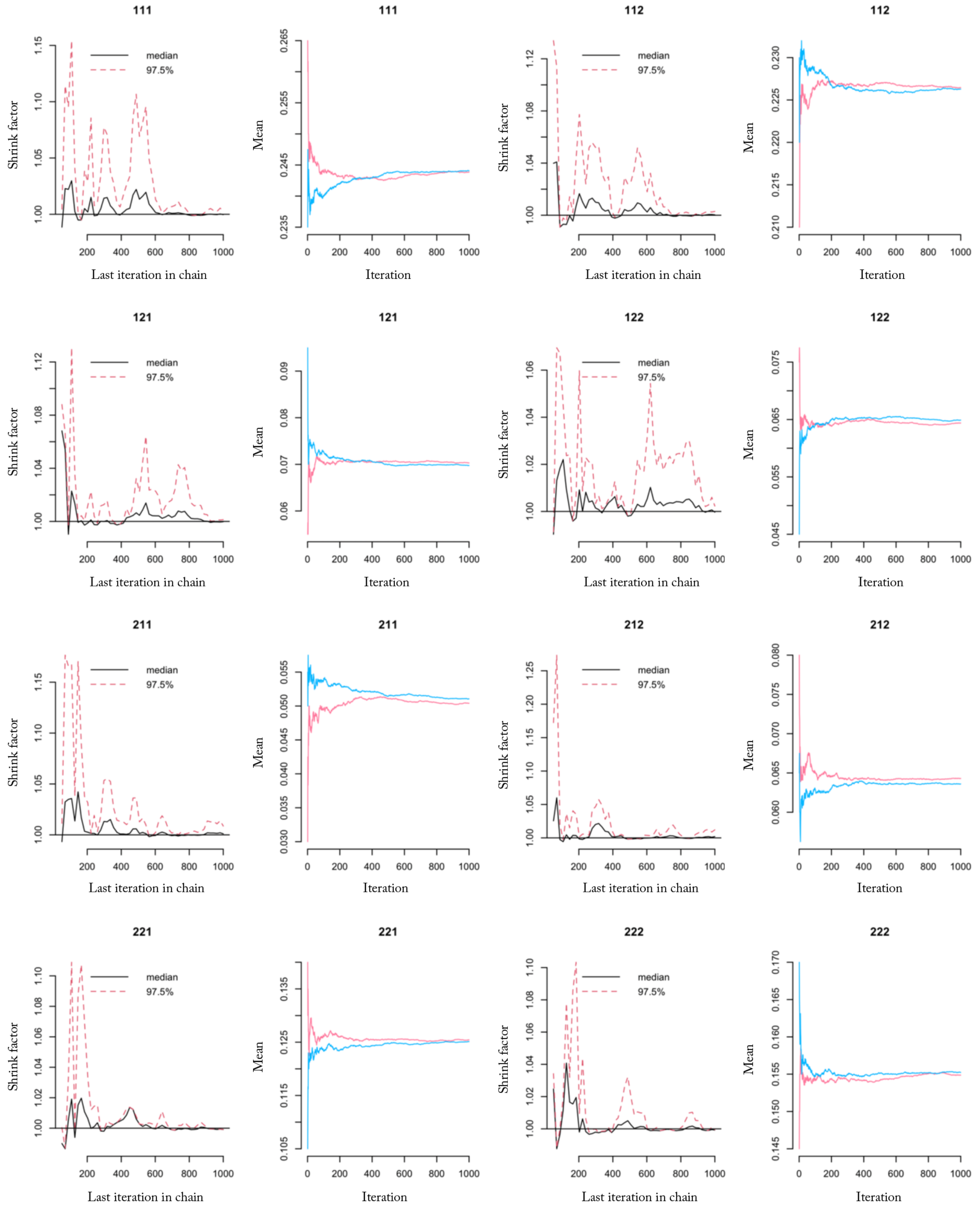

Figure S8. Gelman Rubin Brooks plot (shrink factor as the number of iterations) and mean plot (mean as the number of iterations) with 2 chains at 1000 iterations.