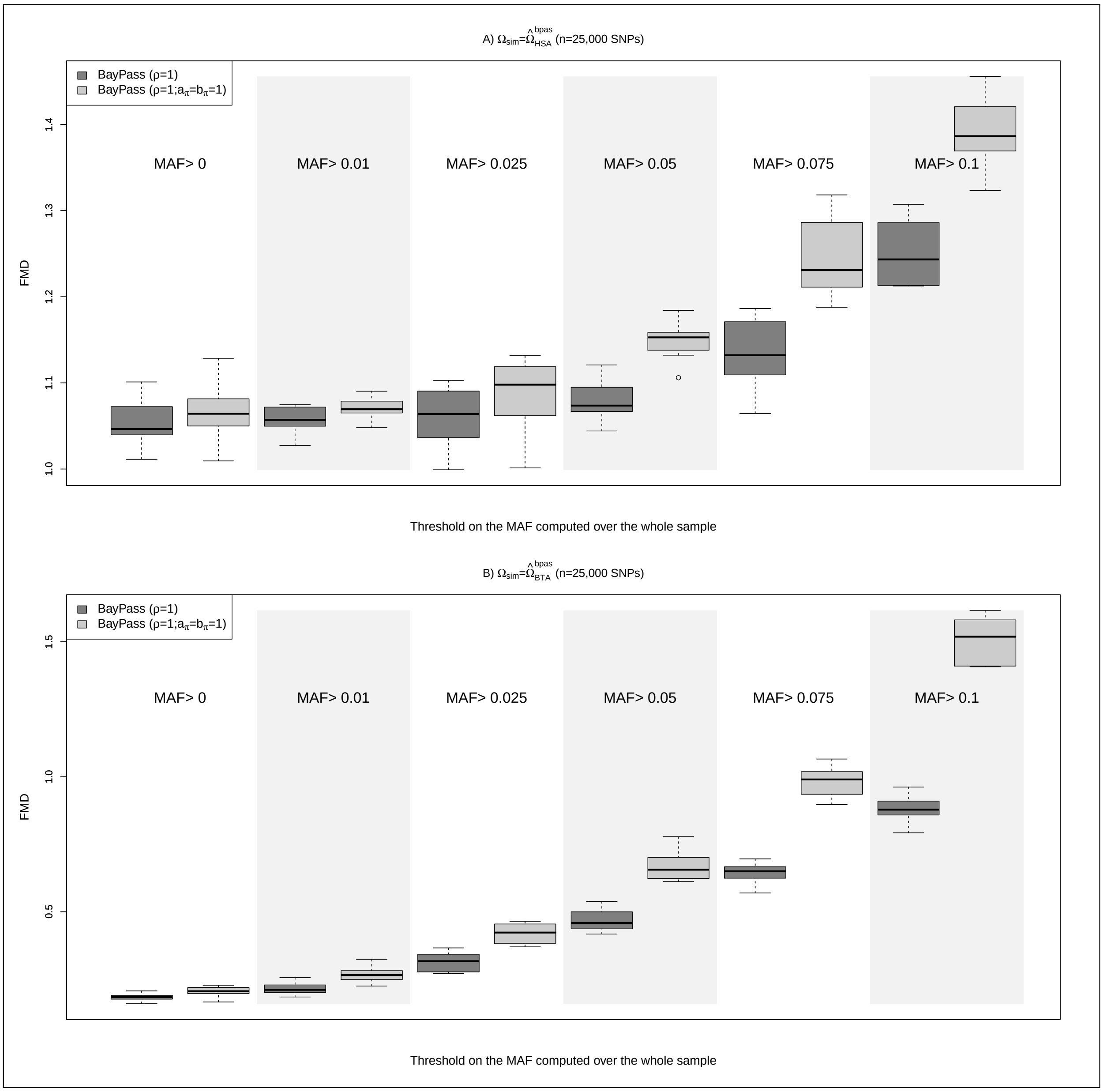

Figure S5: FMD distances (Förstner and Moonen, 2003) between the matrices used to simulate the data sets and their estimates for different SNP ascertainment scheme. Two simulation scenarios defined according to the matrix $\mathbf{\Omega}_{\text{sim}}$ used to simulated the data ($\widehat{\mathbf{\Omega}}^{\text{bpas}}_{\text{HSA}}$ in A; and $\widehat{\mathbf{\Omega}}^{\text{bpas}}_{\text{BTA}}$ in B) were considered. Ancestral allele frequencies were sampled from a Unif(0,1) distribution. Ten independent data sets of 100,000 SNPs were simulated per scenario and each divided in six subsamples by randomly sampling 25,000 SNPs with a MAF>0, >0.01, >0.025, >0.05, >0.075 and >0.10 respectively. The resulting data sets were analyzed with BayPass (assuming $\rho = 1$) by either estimating a_π and b_π or setting $a_\pi = b_\pi = 1$. Each box-plot contains 10 FMD distances computed with estimates from the 10 independent data sets.