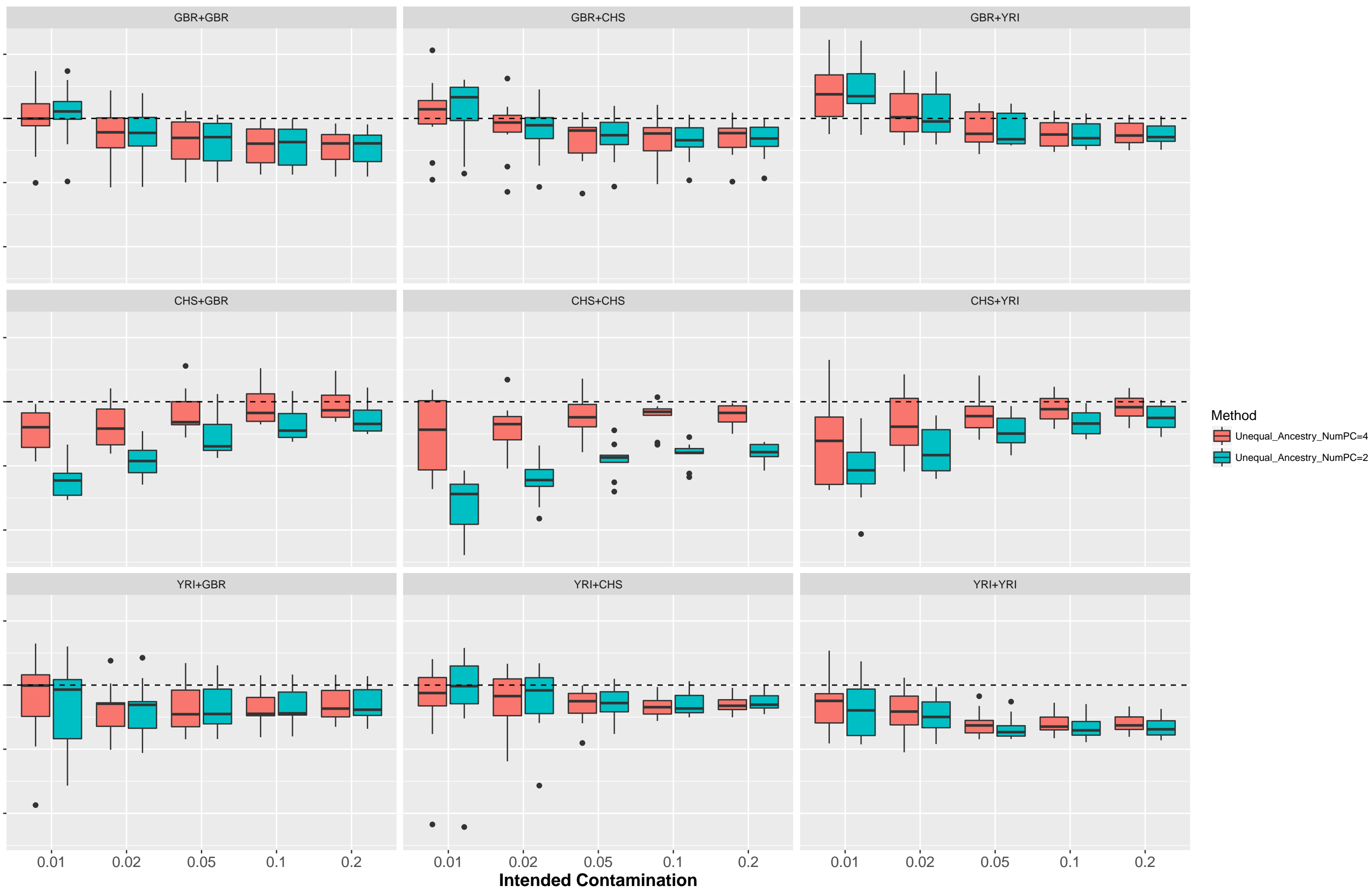


Estimated/Intended Contamination



**Supplemental Figure S1.** Impact of the number of principal components (PCs) on estimated DNA sample contamination using *verifyBamID2*. Horizontal (x) axis shows intended contamination rate, vertical (y) axis shows the ratio of estimated to intended contamination rates. Each color represents different number of PCs (PC=2 or PC=4) used to estimate contamination rates. We used Unequal\_Ancestry mode as in Figure 4, which performed the best among the options considered. With Unequal\_Ancestry, *verifyBamID2* ran allowing different genetic ancestry between intended and contaminating sample. Each panel represents different combinations of intended (row) and contaminating (column) populations, consistent with Figure 4.