



Figure S1 The informativeness for assignment for all SNPs (A, B) and five-SNP haplotypes (C, D) genome-wide based on (A, C) $K = 2$ and (B, D) $K = 6$. The horizontal dashed line is the 95th percentile. The grey vertical dashed lines delineate the two high F_{ST} regions based on comparison between the Eastern and Western populations. The green vertical dashed lines indicate the centromeric regions. Note the scale of y-axis in panel D.