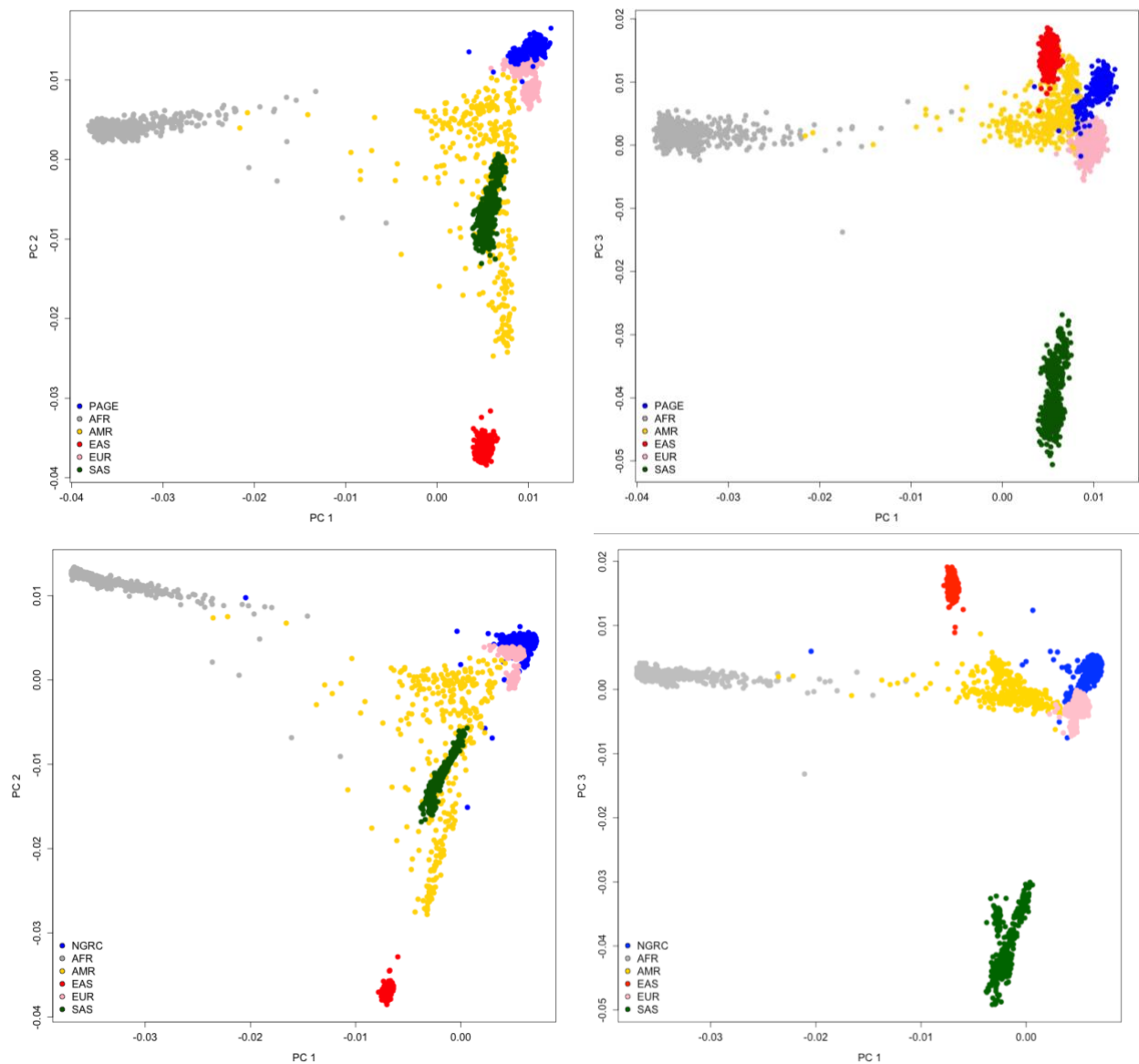


Figure e-2. Principal Component Analysis (PCA) plots.

A. NGRC and PAGE cluster with Europeans in 1000G_Phase_3.

PC1 vs. PC2 (left panel) and PC1 vs. PC3 (right panel) were plotted for NGRC participants with PD (top row, N=1950, using 100K pruned SNPs), and for subset of PAGE participants for whom PCs could be calculated (bottom row, N=396, using 20K pruned SNPs).

Blue: NGRC (first row) or PAGE (second row). Pink: European. Yellow: Americas. Green: South Asia. Red: East Asia. Grey: African.

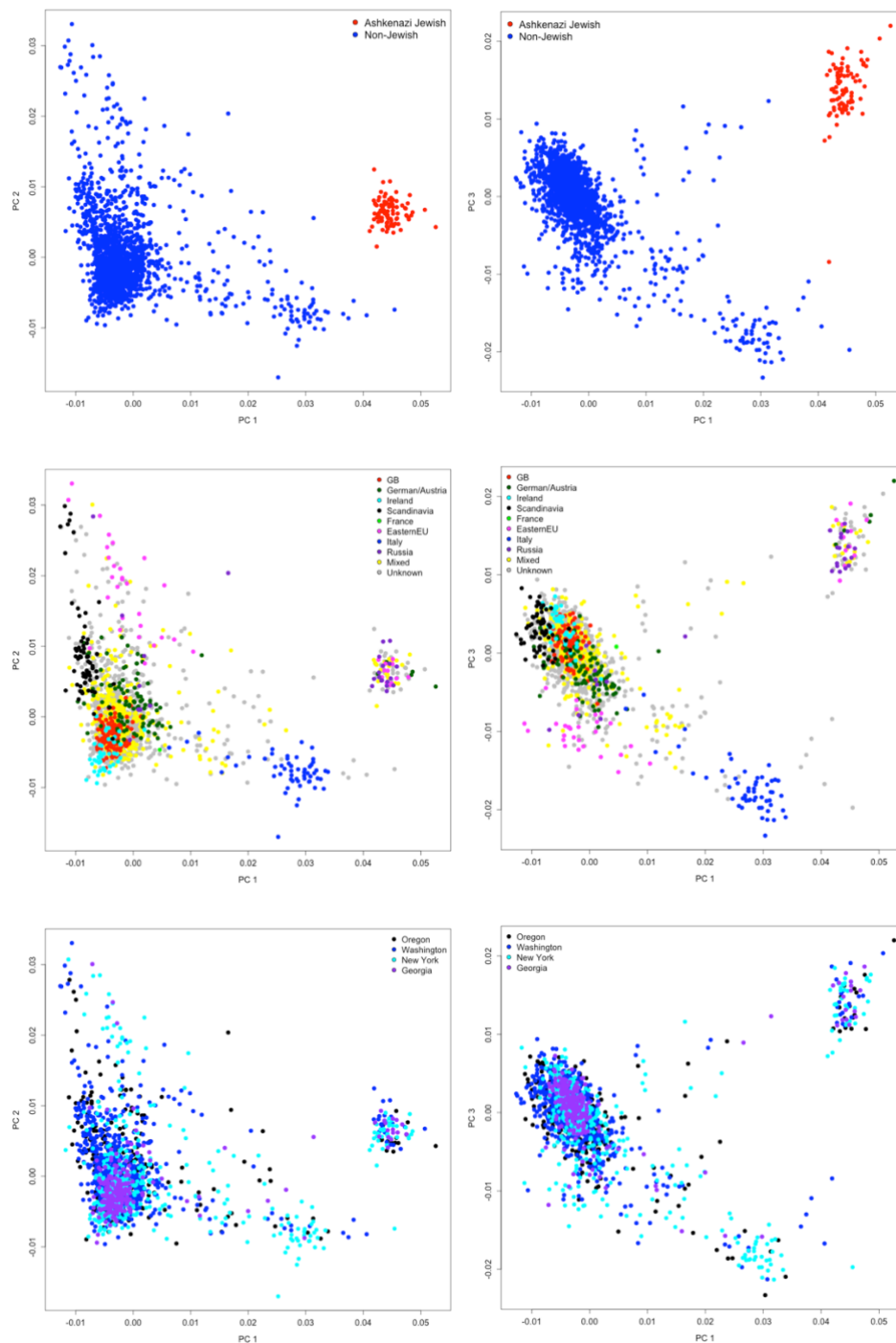


B. PCA of NGRC by Jewish ancestry, European country of origin, and US enrollment site.

First row shows clustering of Jewish ancestry (red) in NGRC. The core of the Jewish cluster was defined within $0.04 \leq PC1 \leq 0.055$ and $0.001 \leq PC2 \leq 0.013$.¹³

Second row is the self-reported European country of ancestry, which follows the map of Europe.

In the third row, NGRC participants are colored according to the state where they were enrolled, showing no particular clustering for enrollment site.



C. Sub-haplotypes & ancestral origins. As there seems to be an ancestral effect, we check whether different sub-haplotypes are present in these blocks for the different ancestral groups in NGRC participants with PD. Block 1 (left triangle in the haploview) is a single block (LD>0.9) represented by rs73495940. Block 2 (right triangle in the haploview) is composed of several sub-haplotypes (LD≤0.7). To tag these haplotypes, while incorporating functional information, we used the 5 variants in block 2 that had significant evidence for functional relevance (either HiC FDR<1E-6 and enhancer in brain, or CADD>10, or eQTL with FDR=4E-4 see **table 4**). As shown in the LD table below, 3 of the 5 variants are in close LD ($r^2>0.95$) which can be tagged by one SNP (rs17763929). We defined 3 sub-haplotypes in block 2 represented by rs17763929, rs6118842, and rs149155028. We identified the individuals who carried the minor allele of each variant, and mapped them to PC1 vs. PC2 (top row) and PC1 vs. PC3 (bottom row). There is no pattern that would suggest sub-haplotypes cluster with different ancestral group.

LD in Block 2	rs17763929	rs61188842	rs117058418	rs117314512	rs149155028
rs17763929	-	0.58	0.96	0.96	0.66
rs61188842		-	0.57	0.57	0.39
rs117058418			-	1	0.7
rs117314512				-	0.7
rs149155028					-

