



Supplemental Figure S5. Comparison of proxy and original panels. **(a)** Distribution of Pearson correlation of genotypes between proxy (with and without resampling), matched and holdout panels. **(b)** Density plot shows the distribution of allele frequencies in proxy panel (y-axis) versus the original panel (x-axis). **(c)** Density plot shows the distribution of allele frequencies in holdout panel (y-axis) versus the original panel (x-axis). **(d)** Distribution of the variant-variant correlations in 500 variant windows that illustrates linkage disequilibrium patterns in the proxy panel and the original panel for matching pairs of variants. Each dot corresponds to a pair of variants whose R2 correlation is shown on x and y axes on proxy and original panels, respectively. **(e)** The comparison of variant-variant R2 in cleartext original and holdout (control) panels. **(f)** The distribution of local k-mer counts in proxy panel and matching original (original1) panel. **(g)** The distribution of local k-mer counts in proxy panel from proxy panel generated from original2 panel and the cleartext original1 panel. **(h)** The distribution of local k-mer counts in proxy panels generated from original1 and the original2 panels. **(i)** The distribution of local k-mer counts in cleartext original1 panel and the cleartext original2 panel. **(j)** Distribution of the subjects in original panel and the proxy panels from Principal component analysis (PCA). The first two components are shown in the scatter plot. Each dot is a sample and colors indicate the populations for the corresponding individuals. Proxy panels are generated with and without resampling indicated by colors. Note that when resampling is used, population is not depicted since each haplotype is a mosaic of all populations.