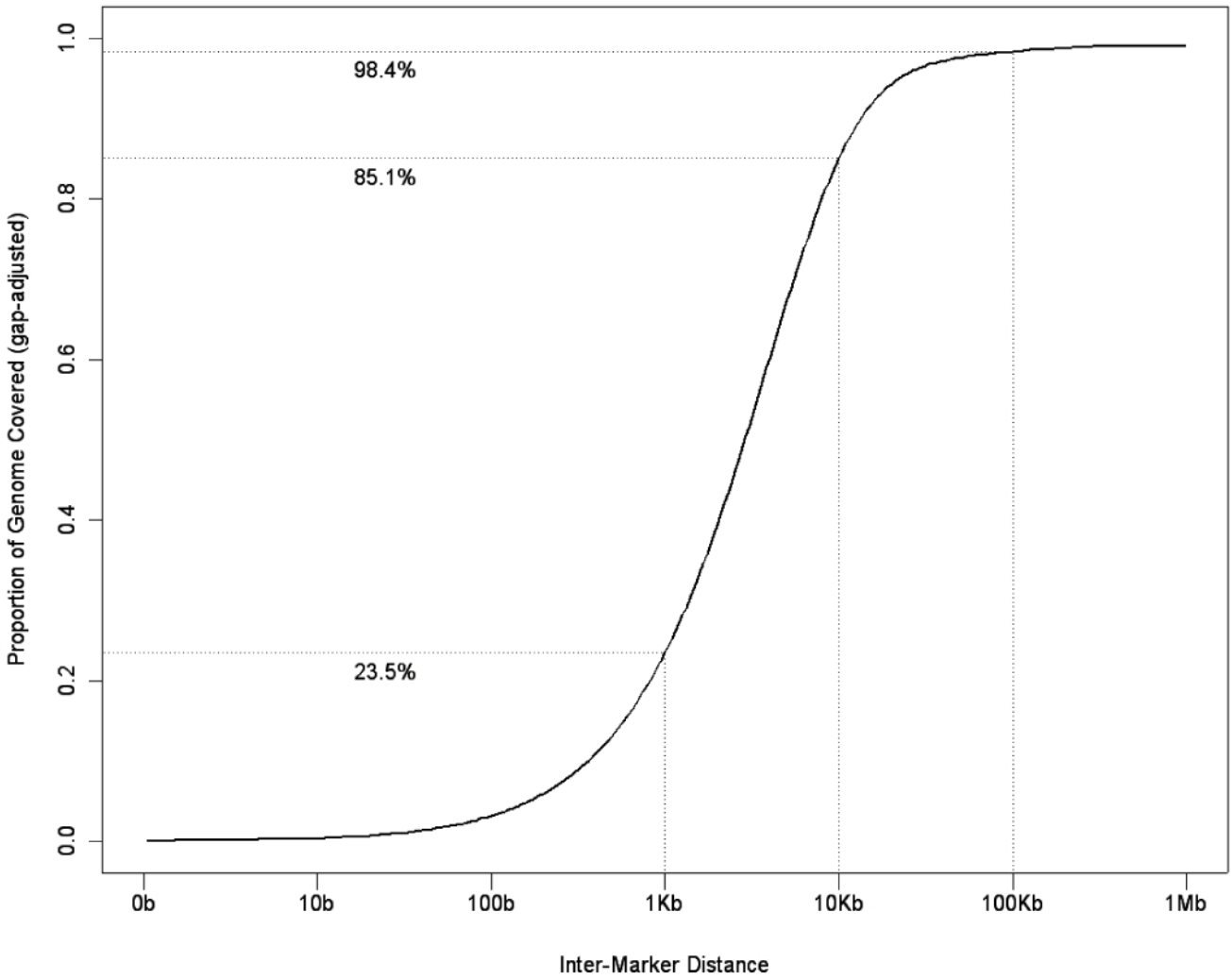
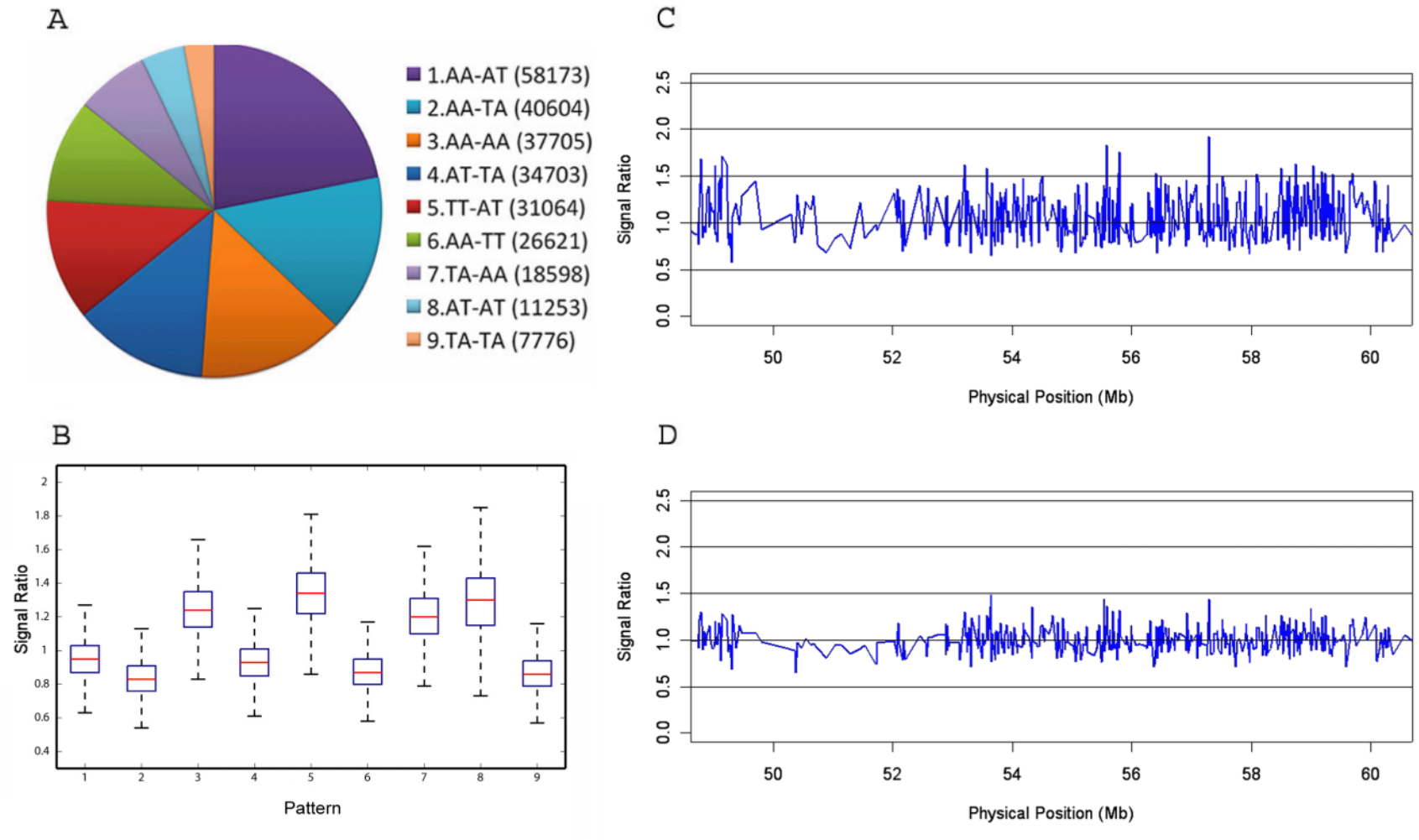
Supplementary Figure 1



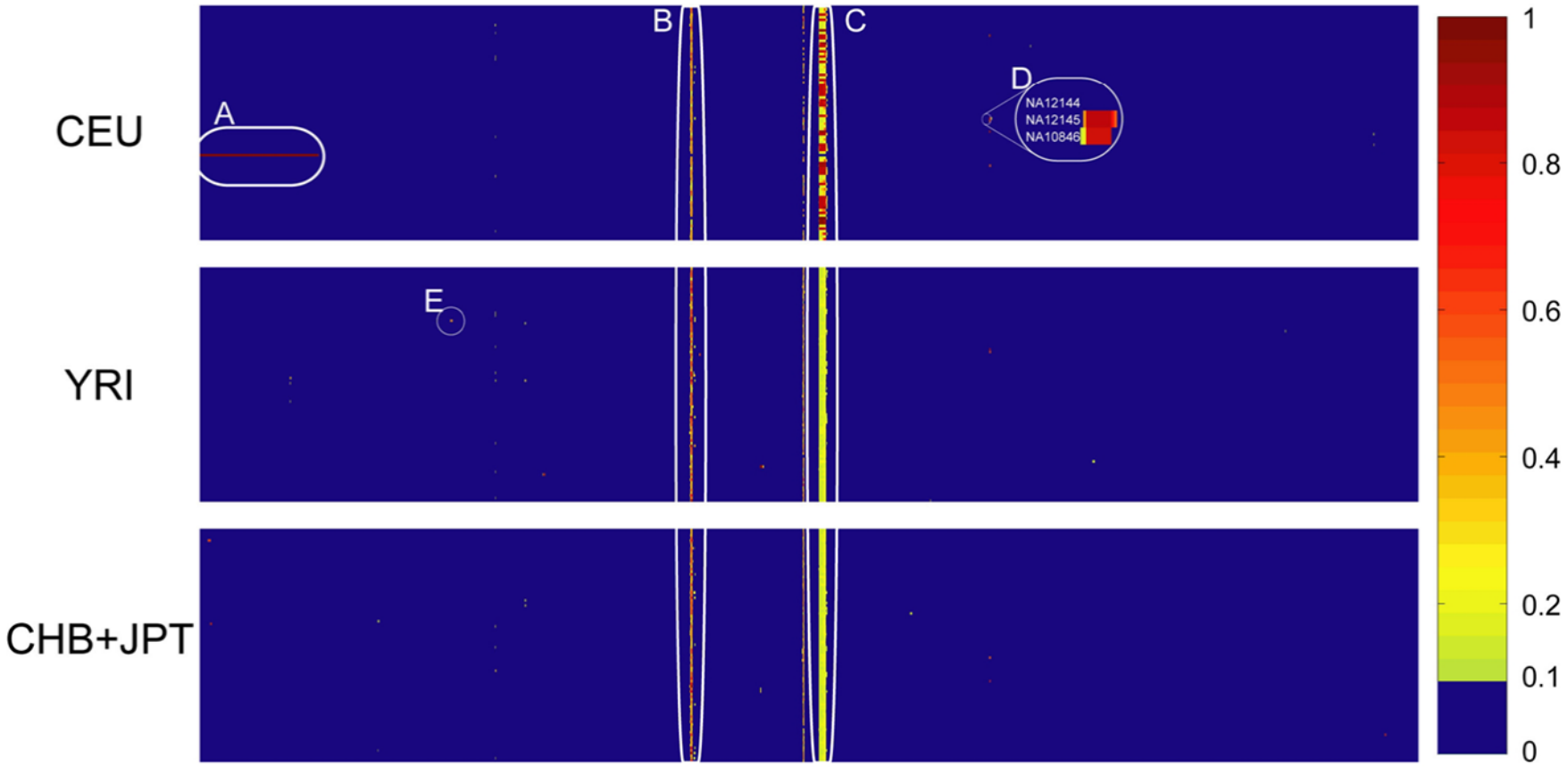
Supplementary Figure 1. Genome coverage of 500K EA probes selected for CNV analysis. 474,642 SNPs were selected from the original set of 534,500 resulting in an increase of the median inter-probe distance from 2,377bp to 2,709bp. In this figure, the X-axis shows the distance between any given point in the gap-adjusted genome and the closest marker. For the 474,642 selected SNPs, 85.1% of the genome is less than 10kb away from a SNP (compared to 87.3% for the non-selected set of 534,500 SNPs) while 98.4% (vs 98.7%) of the genome has a marker within 100kb.

Supplementary Figure 2



Supplementary Figure 2. Median scaling across Sty I recognition sequences results in a dramatic improvement in intensity variability. There are two semi-degenerate bases (W=A or T) in the six-base Sty I enzyme recognition sequences rendering sixteen possible combinations of restriction fragment ends. Nine of these sixteen combinations are represented on the 250K Sty I array. (A) The frequency of SNPs in each subset of the Sty I restriction fragments based on the sequence of the recognition sites. Each category is labeled with four bases, the first two indicating which bases are used in the WW degenerate nucleotides at the 5' end of the sense strand of the restriction fragment and the last two indicating which bases are used at the 5' end of the anti-sense strand. The numbers in parentheses lists the absolute frequency of SNPs residing on each restriction fragment class. (B) The intensity ratio box-plots for each type of fragment from a pair-wise comparison between samples NA10851 versus NA18503. The numbers on the X-axis corresponds to the category index in panel A. (C) and (D) show the intensity ratio profile of NA10851 and NA18503 before (C) and after (D) the median scaling normalization based on Sty I recognition sequence for chromosome 1: 49067660-60248385. The standard deviation of the intensity ratios decreases from 0.236 to 0.144.

Supplementary Figure 3



Supplementary Figure 3. Density distribution plot for CNVs called from all 270 HapMap samples on chromosome 17. Blue indicates no CNVs are called, and red indicates a CNV is called in 100% of pair-wise comparisons. (A) A putative cell line artifact (based on its size) observed in NA12056. (B) A common CNV observed in all populations. (C) Population specific CNV observed only in the CEU population (which skews the density plot to yellow in non-CEU populations being compared to CEU individuals). (D) A Mendelian inherited CNV in NA10846 (offspring) is inherited from NA12145. (E) Singleton CNVs found in only one individual. 11 singletons are located on chromsome 17 and are visible as individual dots in the density plot.