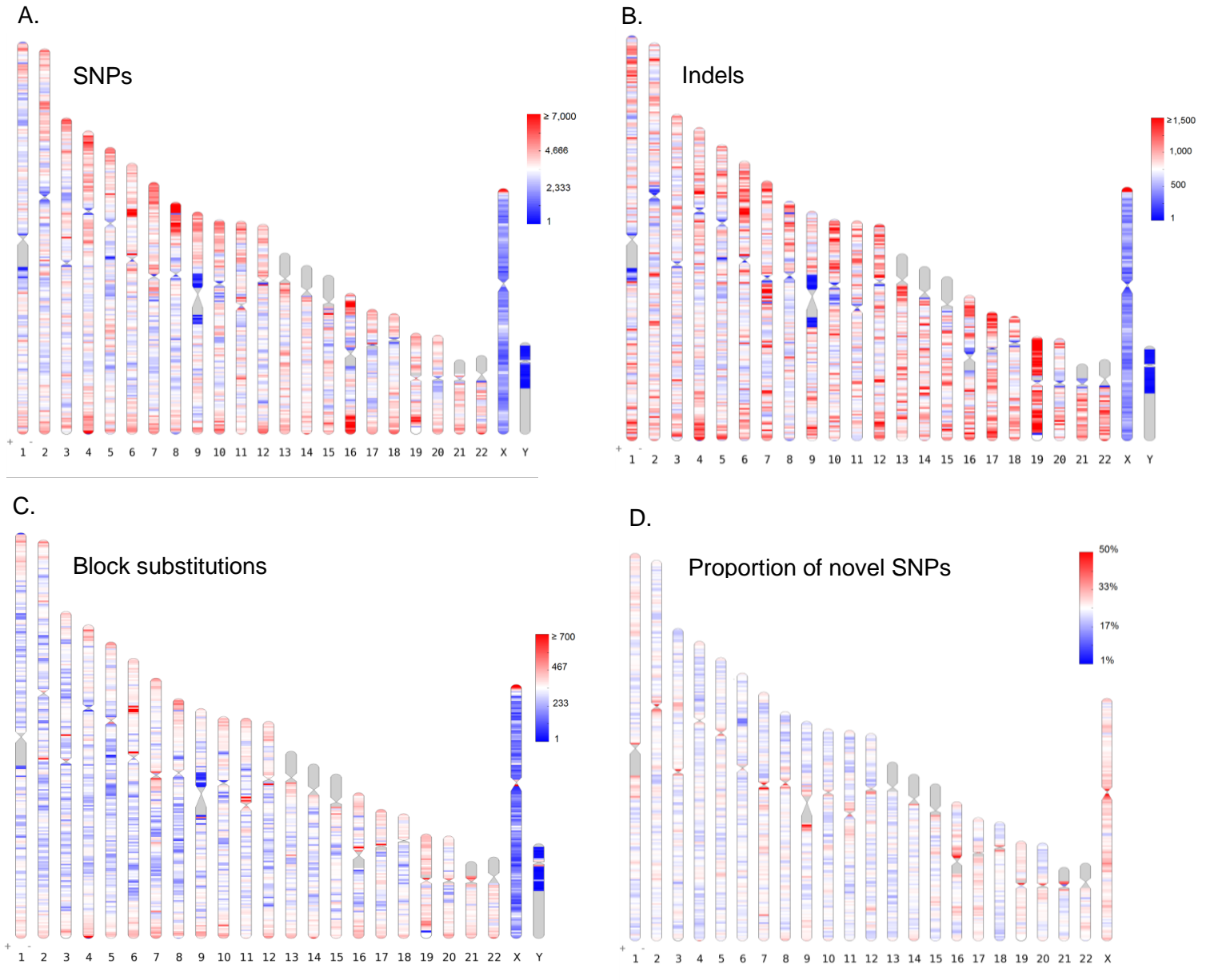


Figure S1. Genome-wide distribution of A) SNPs, B) indels, C) block substitutions, and D) proportion of novel SNPs.



Total number of SNP, indels and block substitutions were calculated in non-overlap 1-megabases (Mb) windows across the human genome. The proportion of novel SNPs in each 1Mb window was calculated as the percentage of all the identified SNPs that were not archived in dbSNP v131 or the 1000 Genomes Project Phase 1 data. Ideograms were plotted using *Idiographica* (Kin and Ono 2007). The diversities were illustrated by colors, with red indicating higher numbers and blue indicating lower numbers. Genomic regions in which no indels or block substitutions were identified or no reference sequences could be determined are shown in grey.