PgmNr 320: IBD sharing in the 1000 Genomes Project Phase 3 data reveals relationships from Neandertals to present day families

G. Povysil; S. Hochreiter

Institute of Bioinformatics, Johannes Kepler University Linz, Linz, Austria

The 1000 Genomes Project data harbor information about a great variety of relationships which can be recovered using identity by descent (IBD) analysis. Short IBD segments convey information about events far back in time because the shorter IBD segments are, the older they are assumed to be. At the same time longer IBD segments can be used to detect more recent relationships as they occur in families. The identification of short IBD segments becomes possible through next generation sequencing (NGS), which offers high variant density and reports variants of all frequencies. However, only recently HapFABIA has been proposed as the first method for detecting very short IBD segments in NGS data. HapFABIA utilizes rare variants to identify IBD segments with a low false discovery rate.

We applied HapFABIA to the 1000 Genomes Phase 3 whole genome sequencing data to identify IBD segments which are shared within and between populations as well as with the genomes of Neandertal and Denisova. Using the proportion of IBD segments an individual shares with any other individual in the data set, we were able to discover first degree relatives that we consequently removed from further analyses. Not only are most IBD segments found in Africans, but also each African individual has about ten times more IBD segments than any East Asian, South Asian, or European individual. Furthermore, the number of IBD segments of an individual correlates with his degree of African ancestry as reported by other methods. IBD segments can be used to recover the population of origin of an individual and find individuals with wrong population labels. By comparing the rare variants that tag an IBD segment with the genome of Neandertal and Denisova, we were able to find IBD segments shared with these ancient genomes. We extracted two types of very old IBD segments that are shared with Neandertals/Denisovans: (1) longer segments primarily found in East Asians, South Asians, and Europeans that indicate introgression events outside of Africa; (2) shorter segments mainly shared by Africans that may indicate events involving ancestors of humans and other ancient hominins within Africa. Our results from the autosomes are further supported by an analysis of chromosome X, on which segments that are shared by Africans and match the Neandertal and/or Denisova genome were even more prominent.