Abstract/Session Information for Program Number 1483F

Print Close window

Session Information

Session Title: Statistical Genetics and Genetic Epidemiology **Session Type:** Poster

Session Location: Exhibit Hall, Lower Level South, Moscone Center **Session Time:** Fri 7:00AM-4:30PM

Abstract Information

Program Number: 1483F **Presentation Time:** Fri, Nov 9, 2012, 2:15PM-3:15PM

Keywords: Statistical Genetics and Genetic Epidemiology, KW084 - haplotype, KW008 - bioinformatics, KW031 - computational tools, KW057 - evolutionary genetics, KW079 - genome sequencing

Abstract Content

Haplotype clusters of rare variants in Korean genomes. *S. Hochreiter, G. Klambauer, G. Povysil, D.-A. Clevert*
Institute of Bioinformatics, Johannes Kepler University Linz, Linz, upper Austria, Austria.

A haplotype cluster corresponds to a DNA segment that is descended from a single founder and, therefore, some haplotypes are similar to each other in this segment. The advent of new sequencing technologies facilitates the identification of rare variants and therefore haplotype clusters containing rare variants (“rare haplotype clusters”). However, LD-related methods fail to extract rare haplotype clusters because of the high variance of LD measures for rare variants. IBD detection methods require sufficiently long IBD regions to avoid high false positive rates. We propose identifying rare haplotype clusters by HapFABIA which uses biclustering to combine LD information across individuals and IBD information along the chromosome. To identify rare haplotype clusters in the Korean population, we applied HapFABIA to data from the Korean Personal Genome Project (KPGP). Genotyping data from the KPGP was combined with those from the 1000-Genomes-Project leading to 1,131 individuals and 3.1 million single nucleotide variants (SNVs) on chromosome 1. HapFABIA identified 113,963 different rare haplotype clusters marked by tagSNVs that have a minor allele frequency of 5% or less. The rare haplotype clusters comprise 680,904 SNVs; that is 36.1% of the rare variants and 21.5% of all variants. The vast majority of 107,473 haplotype clusters contains Africans, while only 9,554 and 6,933 contain Europeans and Asians, respectively. We characterized haplotype clusters by matching with archaic genomes. Haplotype clusters that match the Denisova or the Neandertal genome are significantly enriched by Asians and Europeans. Interestingly, haplotype clusters matching the Denisova or the Neandertal genome contain, in some cases exclusively, Africans. Our findings indicate that the majority of rare haplotype clusters from chromosome 1 are based on ancient founder segments from times before humans migrated out of Africa. The enrichment of Koreans in Neandertal haplotype clusters (odds ratio 10.6 of Fisher’s exact test) is not as high as for Han Chinese from Beijing, Han Chinese from South, and Japanese (odds ratios 23.9, 19.1, 22.7 of Fisher’s exact test). In contrast to these results, the enrichment of Koreans in Denisova haplotype clusters (odds ratio 36.7 of Fisher’s exact test) is is higher than for Han Chinese from Beijing, Han Chinese from South, and Japanese (odds ratios 7.6, 6.9, 7.0 of Fisher’s exact test).

You may contact the first author (during and after the meeting) at hochreit@bioinf.jku.at

Print Close window