

Draft: Investigation of High Heterozygosity Variants in Freeze2

AACHueva

July 14, 2016

Files that were used in the analysis:

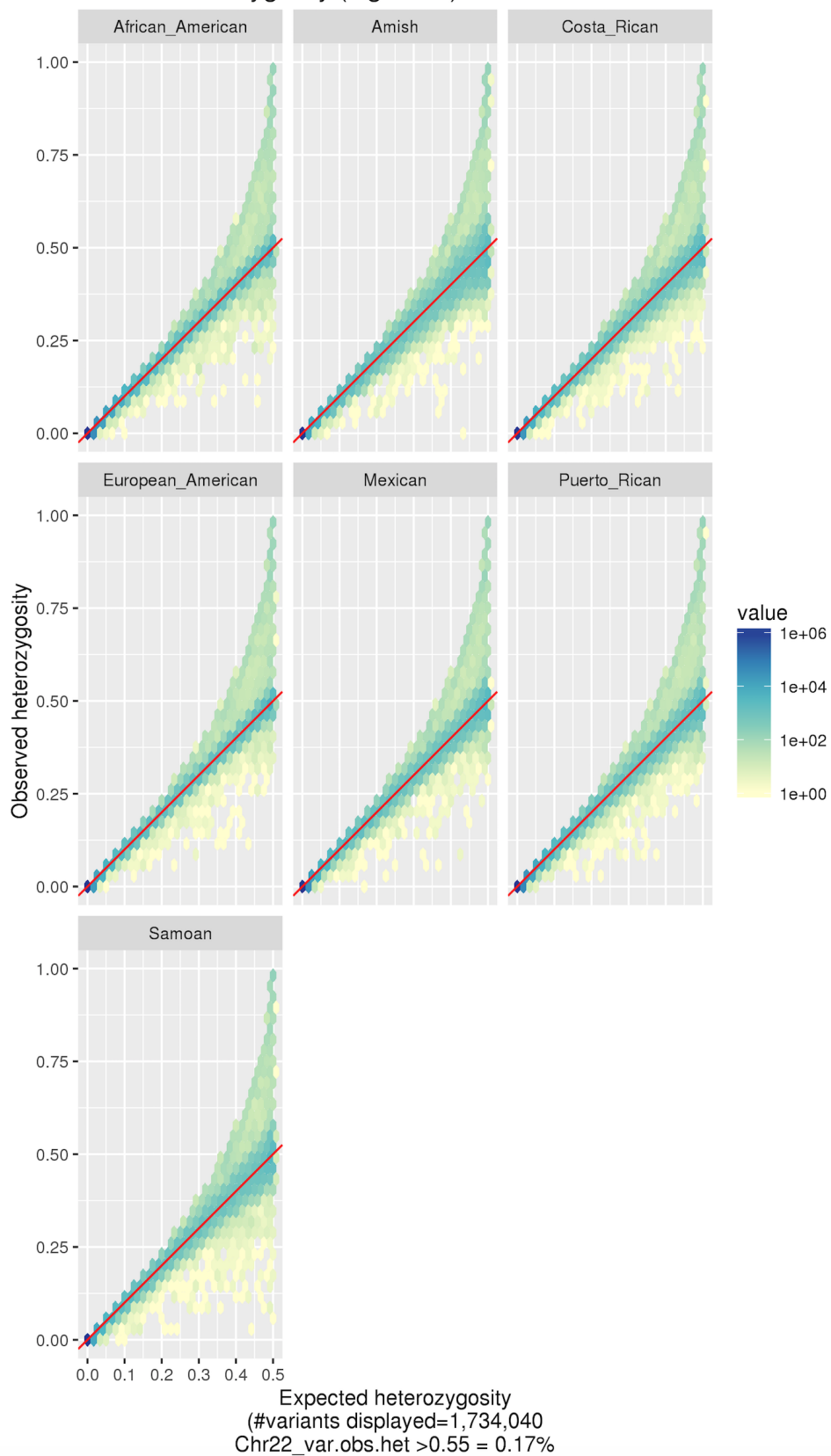
1. File with TOPMed InDel Annotations (courtesy of Xiaoming Liu)
2. Feeeze 2 GDS GT only (includes all chromosomes)
3. HWE results for each ancestry (courtesy of Stephanie G.)

Pre-work that was done:

- * Created a dataframe with variant.id, chr, pos, ref, alt, MAP20 and MAP35 fields.
- * Extracted variants only for chromosome 22.

Plot 1. Expected heterozygosity versus observed heterozygosity on all variants on chr 22

Expected heterozygosity versus
observed heterozygosity (log scale) for all variants on chr 22



Fraction of high hets over all chromosomes

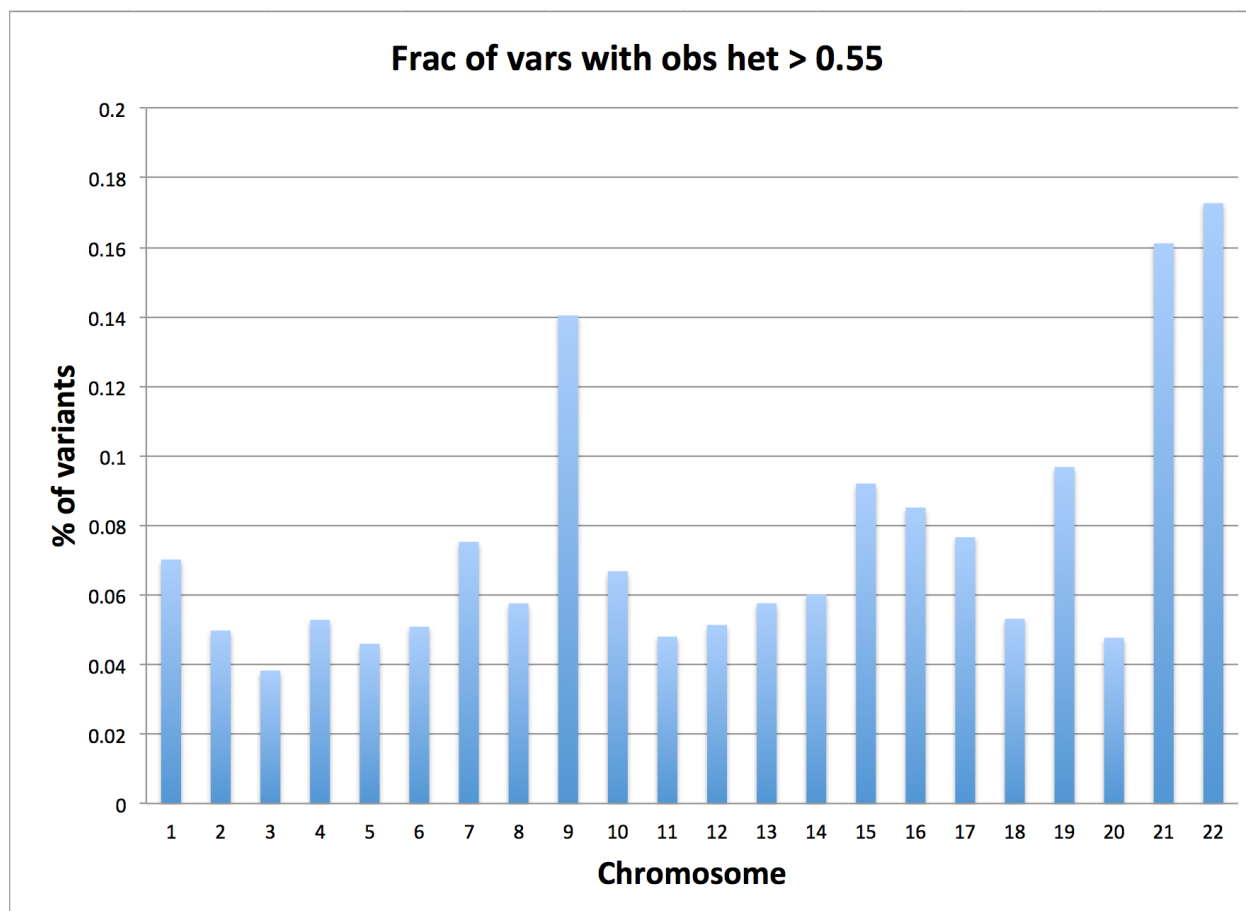


Figure 1: