

Assignment 2 – Genome-wide Association Studies

Deadline: 10/01/2021

Researchers are interested in detecting new SNPs associated with bmi. To do so, they performed a GWAS using DNA information about 425 individuals. Genotype information is available in plink format (files '`coronary.bed`', '`coronary.bim`', '`coronary.fam`') while phenotypic information can be found in the file '`coronary.txt`'.

- Read genotypes using `snpStats` library.
- Verify that both data sets are in the same order (NOTE: '`id`' variable in the file '`coronary.txt`' must be used since it corresponds to the unique patient number).
- Perform some QC filtering.
- Remove those SNPs that do not pass QC.
- Assess association between '`bmi`' and the SNPs adjusting for population stratification (variables '`ev3`' and '`ev4`' in the file '`coronary.txt`') (NOTE: remember that you are analyzing a quantitative trait).
- Create a Manhattan plot.
- Select a region from the genome (about 200 SNPs) and create a plot using Locus Zoom software. Indicate why this region has been selected