Assignment 1 – Genetic Association Studies

Deadline: 22/12/2019

File `asthma.txt` (tab-delimited space) contains data corresponding to 1,578 individuals belonging to a case-control study on ashtma. The file includes genetic information about 51 SNPs from selected genes (candidate approach) as well as some phenotypic information about: country, gender, age, bmi (body mass index), chronic obstructive pulmonary disease status (copd - 0: no copd, 1: mild, 2: moderate, 3: Severe, 4: Very Severe) and asthma status (caco - 0: control, 1: asthma).

There are two main aim of the study:
a) To find susceptibility genes of asthma
b) To find susceptibility genes of copd for any of its categories

This excercise tries to mimic a real situation where you are going to be involved in the near future. Let us imagine that you belong to a research group and you are in charge of data analysis. The PI (boss) wants you to address these two goals that correspond to a real problem. Therefore, you have to performe the statistical analyses you think that are necessaries to answer both research questions.

To do so:
1. Perform all the required steps we have seen in class (HWE, SNP filtering, association analysis, multiple comparisons, different genetic test, Manhattan plots, interaction analysis (not mandatory), ORs of association, adjusted analyses, ...) to perform a genetic association study.
2. Interpret your results and give some conclusions.
3. Provide the R code used to get your conclusions (Appendix).