Project 2: Genetic Associations

Introduction

Genetic epidemiology accommodates different viewpoints to look at "disease". Unraveling important functional determinants to or causal factors for complex diseases requires a systems biology view, combining evidences from different data sources, involving the genome, the transcriptome, and epigenome, amongst others.

In this project (Project 2), you will select a human "complex disease" or "complex trait" of interest (take one for which evidence of disease-associated epigenetic mechanisms exist as well – cfr Project 1, Cortessis et al. 2012) and will search the "Catalog of GWAS" for published genome-wide association studies for this disease/trait. The link to search this catalog is http://www.genome.gov/gwastudies/#searchForm.
Then select one of the most recently published genome-wide association studies obtained via your search, upload the associated publication, and discuss using the specific questions below.

Obviously, you are not restricted to this publication only to build up your story.

Specific questions

- What is a genetic association study? How does it fit into the general work flow within genetic epidemiology?

For the selected study:

- Describe the biological question(s).
- What is the design of the study? (markers, subjects) Is it different from the designs seen in class? If so, what was the motivation to select a different design?
- Which quality control procedures have been put in place? Are they in line with the Travemunde criteria? If not, was there a motivation given in the paper for adopting a different criterium, or can you come up with a motivation yourself? Be critical.
- How did one make use of the concept of LD (linkage disequilibrium)? Was it used to reduce the number of tests? Was it used after the analysis to get closer to causal variants?
- What type of association test was carried out? Single locus at the time? Haplotype-analysis was considered as well? What is the possible advantage of performing a haplotype analysis? What are the drawbacks?
- Was there a need to correct for population stratification? What is population stratification? How did one correct for it? Are there other ways?
- Were the genetic association results supported by a replication analysis or a validation analysis? If so, what did it involve? What is the difference between the two?
- What are the final conclusions of the study and how much trust can be given to them (when looking at the replication/validation results)?
- What type of follow-up analyses do the authors advocate? Can you place these in the context of (modern) "genetic epidemiology"?