Genome-wide association studies

Assignment I and Organization

Select one of the provided papers and discuss it following the layout of "critical evaluation of a paper" (see class material). Provide a written reflection of your work via a comprehensive slides presentation, in which you have woven in the guiding questions that match the paper of your choice. These slides will contribute to your "written exam".

Select a portion of your slides, addressing 3 series of questions in depth, name it "Bioinformatics Applications Assignment I _your name" and submit it to Prof Van Steen (kristel.vansteen@uliege.be) with the subject title "Bioinformatics Applications Assignment I", by Sunday night March 28. During the class of March 31 you are expected to present your slides in ~15 minutes. Thereafter, your presentation will be taken as a starting point for Q&A and a discussion to clarify unknown concepts, while making links to the course materials and the latest developments in the field.

Papers

Example case studies on asthma:

https://www.nejm.org/doi/10.1056/NEJMoa0906312

https://www.biorxiv.org/content/10.1101/195933v3.full

An adapted list will be provided after class 1 – based on individual interests.

---

10 series of guiding questions

1. Describe the biological/research question(s) and put them in context.
2. 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?
3. 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.
4. 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 identify causal variants? Other uses?
5. 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?
6. Was there a need to correct for population stratification? What is population stratification? How did one correct for it? Are there other ways?
7. Were the genetic association results supported by a replication analysis or a validation analysis? If so, what did it involve? What are the factors causing a non-replication? May it also be the existence of gene-gene interactions?
8. What are the final conclusions of the study and how much trust can be given to them (when looking at the replication/validation results)?
9. What type of follow-up analyses do the authors advocate? Do these analyses involve multi-omics data? Why or why not?
10. Can you situate the study in the context of (modern) "genetic epidemiology"? What does modern genetic epidemiology stand for? What is the link with bioinformatics?