Part 2: Post-Genomics

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. With a different genetic background, another drug-response pattern of a diseased individual may appear (i.e., an example of a gene-environment interaction; pharmacogenetics). Although a set of genetic markers have been identified to (directly) modify gene expression (called: eQTLs) in specific tissues, their effect may be disturbed when epigenetic mechanisms are operating.

In this project, you will consider the same complex trait as selected in Part 1 and will focus on an extra level of complexity, either "heterogeneity" or "interactions". . For instance, you may wish to investigate whether there has been evidence for gene-gene interactions and gene-environment interaction, or you may wish to focus on causes of heterogeneity (linking to meta-GWA analyses).

Information about gene-gene interaction studies (methodological papers as well as applied papers) can be retrieved from the "Epistasis Blog": http://compgen.blogspot.be/2006/05/mdr-applications.html. This blog is updated on an almost daily basis by Jason Moore and is an excellent resource to stay on top of achievements in the context of gene-gene (and thus also often gene-environment) interactions.

Questions

Interactions

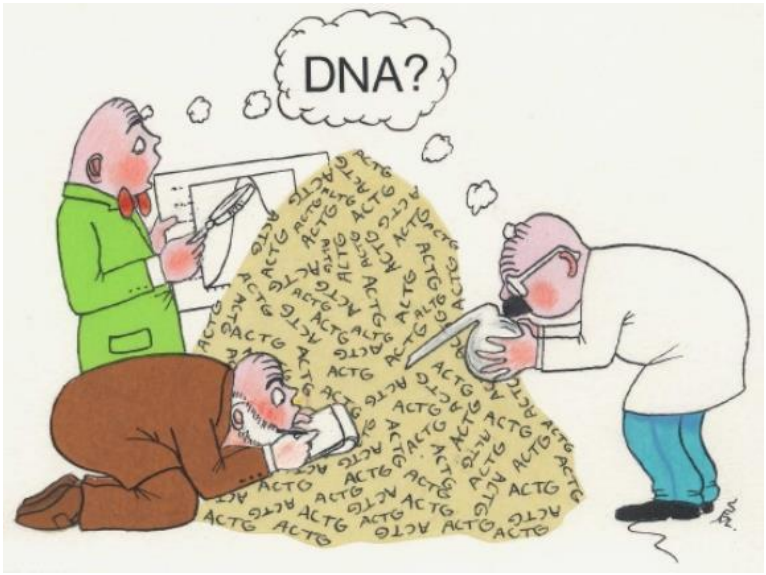
- Give some definitions of "interaction" in epidemiology. Is it different from effect modification? What does epistasis mean? Is there a difference between statistical and genetic epistasis? Is it easy to translate statistical epistasis into biological epistasis? How can this translation be facilitated?
- In what ways will a gene-environment interaction study be different (more complex? less complex?) than a gene-gene interaction study?
- What is meant by an exhaustive search? Is this feasible in the context of a genome-wide setting? Hence, are GWAI studies (genome-wide association interaction studies) at all possible? Does it make sense to investigate higher-order interactions?
- What are the criticisms to traditional regression-based approaches in the context of GWAIs and can you give alternative methods to deal with the abundance of complex data patterns? What is multifactor dimensionality reduction? What are its advantages and limitations? What are random forests? What are its advantages and limitations?
- Replication and validation are important components of any genetic association study. What would replication of a GWAI involve? Can you highlight the differences between "genomics for personalized medicine" and "public health genomics"?

Heterogeneity

- Give a definition of “heterogeneity”. Also consider allelic and genetic heterogeneity, or genetic heterogeneity when modeling gene-gene interactions. In the context of GWAs, which are the forms of heterogeneity one can encounter?
- How can heterogeneity be quantified? Can it be tested for? Can it be modelled? How is heterogeneity dealt with in several contexts within genetic epidemiology (aggregation, segregation, linkage, GWA, …)?
- What is a meta-analysis? What are the pros and cons of such an analysis? Can you list a few techniques that are common in a meta-analysis?
- Will the problem of heterogeneity become larger when multiple omics data sets are merged? Why or why not?
- Replication and validation are important components of any genetic association study. How may heterogeneity affect these components? Can you highlight the differences between “genomics for personalized medicine” and “public health genomics” and can you discuss this in view of (sources of) heterogeneity?

For the selected study:

- Nowadays, editors often ask to carry out a (at least basic) interaction study. Hence, since you have selected a recent publication on a genome-wide association analysis, was an epistasis analysis carried out? Give more details about how the interaction analysis was performed.
- Does the literature in general provide any support for the existence of gene-gene or gene-environment interactions for the trait of interest? Give examples.
- For the identified interaction studies (gene-gene / gene-environment) were there different quality control measures taken compared to a classic GWA study? Why or why not?
- Is the study a case-only study? What are the advantages / disadvantages of such a design in the context of interaction analyses?
- How do the authors “validate” their interaction findings? Comment on the adopted strategies to replicate or validate the findings.
- How do the authors derive a biological interpretation for their results? For instance, in GWA studies one often complements the analysis by investigating whether an association exists between the identified genetic markers and a gene’s expression, since this may give a clue about “functionality”. What is typically done in GWAI settings?



Useful references

- Moore JH (2005). A global view of epistasis. Nat Genet. 37(1):13-4.
- Cordell H (2009). Detecting gene-gene interactions that underlie human diseases. Nature Review Genetics 10: 392.
- Van Steen 2011. Travelling the world of gene-gene interactions. Brief Bioinform. 13(1):1-19
- Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin TE, Chatterjee N, Kraft P, Van Steen K (2012). Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Hum Genet. 131(10):1591-613.
- Martin Gögele, Cosetta Minelli*, Ammarin Thakkinstian, Alex Yurkiewich, Cristian Pattaro,
- Peter P. Pramstaller, Julian Little, John Attia, and John R. Thompson (2012) Methods for Meta-Analyses of Genome-wide Association Studies: Critical Assessment of Empirical Evidence American Journal of Epidemiology 175 (8) DOI: 10.1093/aje/kwr385