

Tutorial 7 - Honor Assignment

Additional assignment for honor students

Advanced analysis

Up until now you have analysed relatively straightforwardly what one can do given these data. However, ranking may not be the most sophisticated thing one can do. Therefore, in this third part of this assignment, we invite you to hypothesise and implement potentially interesting assumptions within the domain of epidemiology. Try to think of one or several WHAT IF scenarios and then implement each WHAT IF scenario in a script to subsequently explore the data to try answer it. Below is a four-staged format for each such exploration.

1. Come up with a new or enriched method that you could use to handle these data.
2. Perform the method.
3. Describe what you find.
4. Provide your main result in one nice picture or graph.

Finding a way

With the previous analyses you hopefully found some interesting new data that could describe overlap in SNPs between the traits (either with tutorial 7 or otherwise more sophisticatedly in the questions 1-4). Use these results to further identify biological mechanisms that could explain the overlap.

There are several databases that can help you with this. For instance you can link genetic, epigenetic, transcriptomic, and metabolomics data with the bios QTL browser at genenetwork.nl. In addition you can find sources that can annotate genetic data (for instance: SNIpa <http://snipa.helmholtz-muenchen.de/snipa3/>), which can identify pathways (for instance <http://pantherdb.org>, <https://david.ncicrf.gov>), etc.

5. Try to identify potential pathways using the results from question 1-4 or tutorial 7 to identify potential pathways.

- What do the genes for which the SNP is coding do biologically?
- Can you identify a certain pathway or is it very diverse?

Note you can put all your work in a notebook (or PDF-file) which you can submit at <http://bit.ly/infomdss-honors>.