

Background

In this practical you will explore FUMA results for functional analysis of a GWAS of LDL-cholesterol from Willer et al 2013 Nature Genetics (PMID 24097068). The original dataset was downloaded from: <https://csg.sph.umich.edu/willer/public/lipids2013/> Because it can take some time to run FUMA, the GWAS summary data has already been formatted, uploaded to FUMA and results made public.

Go to the list of publicly available FUMA results <https://fuma.ctglab.nl/browse> And click on the results with ID 610 (title: GLGC_Willeretl_Submission2) to see the FUMA results for this GWAS.

Aim: The purpose of this practical is to get you to explore and get familiar with FUMA results by answering the questions below, so that you are able to run and interpret such analyses with your own data.

Questions

Exploring SNP2GENE Genome-wide plots

1. What is the most significantly associated gene from the gene-based analysis?
2. Based on MAGMA gene-set analysis, what processes are enriched in the GWAS risk regions?
3. Based on MAGMA gene-set analysis, which tissue is significantly enriched amongst GWAS loci, and is that expected *a priori*?

Exploring SNP2GENE Results Summary

4. How many genomics risk loci are associated with LDL-C?
5. How many genes are mapped to these regions?
6. Which genomic region (e.g. intronic, intergenic, 3'UTR etc) is significantly under-represented amongst candidate GWAS SNPs?
7. What proportion of the candidate SNPs are intronic?
8. Which genomic region is the most enriched amongst candidate SNPs?

Exploring SNP2GENE Results

9. What is the rs ID for the most significantly associated SNP?
10. Which chromosome is this SNP located on?
11. What genomic region is the SNP located in (e.g. coding region for gene X, UTR3 for gene Y, downstream of gene X or intergenic between gene X and gene Y)?
12. What is the CADD score for this SNP?
13. What is the RegulomeDB score for this SNP and what does the score mean?
14. Which genes is the above SNP an eQTL for in GTEx Liver data?
15. Of the above identified genes, which have evidence of physical interaction between the GWAS associated region and the gene's promoter (use the circus plot to determine this)?

16. Which Wiki pathways related to medication are enriched in LDL-C-associated loci?
17. Which non-cardiovascular GWAS data also show an enrichment of LDL-C-associated loci and which genes overlap with both traits?