

GWA Meta-Analysis METAL practical

In this practical we will run a GWA meta-analysis using METAL

<http://www.sph.umich.edu/csg/abecasis/metal/>

Documentation can be found at the metal wiki:

http://genome.sph.umich.edu/wiki/Metal_Documentation

Start by making a working folder and copying the practical files to it. In a new terminal window:

```
mkdir metal-practical  
  
cp /faculty/meike/2019/metal/* metal-practical/  
  
cd metal-practical
```

METAL is flexible. It can run:

- By default, METAL combined p-values across studies (sample size, direction of effect)
- Alternative, standard error based weights (but beta and standard error use same units in all studies)

☐

It requires a *driver file*, which

- describes the input files
- defines meta-analysis strategy
- names output file

In the current practical we will run a meta-analysis on two files.

1. Check format of results files

- Ensure all necessary columns are available
- Modify files to include all information

METAL uses certain columns and you will use them depending on the kind on analyses method you choose. The following columns are used

- SNP
- OR
- SE [for standard error meta-analysis]
- P-value [for Z-score meta-analysis]
- N/weight column [If we had two samples of different sizes]

We will run meta-analysis based on effect size and on test statistic. For the weights of test statistic, I've assumed that the sample sizes are the same (METAL defaults to weight of 1 when no weight column is supplied).

Earlier this week you have ran your own GWA on SNPs on Chr 20 genotyped on a small number of cases and controls and we created an extra dataset. So, the current meta-analysis will be based on two datasets. In the 'real' world the number of datasets is often larger (and in that respect the preparation and QC-ing very time consuming and important).

Step 2. Prepare driver file

- Ensure headers match description
- Crosscheck each results file matches Process name

```
# PERFORM META-ANALYSIS based on effect size and on test statistic

# Loading in the input files with results from the participating samples
# Note: Order of samples is ...[sample size, alphabetic order,..]
# Phenotype is ..
# MB March 2019

MARKER SNP
ALLELE A1 A2
PVALUE P                specifies column names
EFFECT log(OR)
STDERR SE

PROCESS results1.txt    processes two results files
PROCESS results2.txt

OUTFILE meta_res_Z .txt    Output file naming

ANALYZE                Conducts Z-based meta-analysis from test statistic

CLEAR                  Clears workspace

SCHEME STDERR          Changes meta-analysis scheme to beta + SE

PROCESS results1.txt    processes two results files
PROCESS results2.txt

OUTFILE meta_res_SE .txt    Output file naming

ANALYZE                Conducts effect size meta-analysis
```

Step 3. Run METAL

```
metal < metal_run_file
```

metal is the command and *metal_run_file* is the driver file

This will output information on the running of METAL things to standard out [the terminal]

It will spawn 4 files:

- 2 results files: meta_res_Z1.txt and meta_res_SE1.txt
- 2 info files: meta_res_Z1.txt.info and meta_res_SE1.txt.info

Step 4. Looking at your results

We are going to create the manhattan plots for the two GWAs and the GWAMA
There is an R-script in your folder (script_METAL_practicum.R), that will do the trick for you.

R-script:

```
library("qqman")

#https://cran.r-project.org/web/packages/qqman/vignettes/qqman.html

#SELECT FOLDER TO PRACTICUM FILES!

#File containing info for creating MH plot
gwas_example_bim <- read.delim("gwas-example-bim.txt", header = T)

#we need the following columns
gwas_example_mh <- gwas_example_bim[c(1,2,4)]

results1 <- read.delim("results1.txt", header = T)
head(results1)
tail(results1)
colnames(results1)[3] <- "BP"

results2 <- read.delim("results2.txt", header = T, sep = " ")
head(results2)
tail(results2)
colnames(results2)[3] <- "BP"

#read in meta-analysis files
meta_SE <- read.delim("meta_res_SE1.txt", header = T, sep = "\t")
colnames(meta_SE)[1] <- "SNP"
head(meta_SE)

meta_Z <- read.delim("meta_res_Z1.txt", header= T, sep = "\t")
colnames(meta_Z)[1] <- "SNP"
head(meta_Z)

#merge file to include CHR and BP
meta_Z_MHplot <- merge(gwas_example_mh, meta_Z, by = "SNP")
colnames(meta_Z_MHplot)[8] <- "P"
head(meta_Z_MHplot)

#make MH plots

png(file = "MHplots.png", width = 8000, height = 5000, units='px', res = 300, pointsize=17)
```

```
layout(matrix(c(1,2,3), 3, 1, byrow = TRUE))
manhattan(results1, main="results1")
manhattan(results2, main="results2")
manhattan(meta_Z_MHplot, main="meta_Z_MHplot")
dev.off()
```