## GWA Meta-Analysis METAL practical

In this practical we will run a GWA meta-analysis using METAL <a href="http://www.sph.umich.edu/csg/abecasis/metal/">http://www.sph.umich.edu/csg/abecasis/metal/</a> Documentation can be found at the metal wiki: <a href="http://genome.sph.umich.edu/wiki/Metal\_Documentation">http://genome.sph.umich.edu/wiki/Metal\_Documentation</a>

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

mkdir metal-practical

```
cp /faculty/meike/2013/metal-practical/* metal-practical/
```

```
cd metal-practical
```

### METAL is flexible. It can run:

- fixed effects meta-analysis
- heterogeneity tests
- effect size, sample size, or weighted meta-analysis

#### 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 2013

MARKER SNP ALLELE A1 A2 PVALUE P EFFECT log(OR) STDERR SE	specifies column names
PROCESS results1.txt PROCESS results2.txt	processes two results files
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 PROCESS results2.txt	processes two results files
OUTFILE meta_res_SE .txt	Output file naming
ANALYZE	Conducts effect size meta-analysis

#### **Step 3. Run METAL**

metal < metal\_run\_file</pre>

*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

Load your results into HAPLOVIEW. The METAL output file have a different header (*marker* instead of *SNP*), so we have to change the header In the same directory run:

./reformat.sh

This changes 1<sup>st</sup> column name to SNP

We can then load the meta-analysis results files into haploview

- Load in the meta\_res\_Z1.txt
- Make sure to include the bim file that Jeff used earlier (gwas-example.bim)

Linkage Format								
Haps Format								
HapMap Format	Results File:	ETAL_prac\test\meta_res_Z1.txt Browse						
HapMap PHASE	Man Cilar	TAL ana disationer sussels his						
HapMap Download	Map Hie:	TAL_prac(test(gwas-example.blm browse						
PLINK Format	🔲 Integrated Map Info 📃 Non-SNP							
Only load results from Chromosome Select Columns								
Igno	re pairwise compa	risons of markers > 500 kb apart.						
	Exclude individuals	with $> 50$ % missing genotypes.						

## We would like to plot the results in a Manhattan plot

MARKER POSITION	Allele1	Allele2	Weight	Zscore	P-value	Direction		
😝 🔿 🔿 Plot Options			2.0	-0.524	0.6004	-+	<b>m</b>	
			2.0	-0.881	0.3783		Ψ	
Title: New Manhattan Plot			2.0	-0.95	0.3421	+-	- 11	
			2.0	1.354	0.1756	++	- 11	
X-Axis: Chromosomes 🛟	Scale: Untran	sformed 🛟	2.0 2.0 2.0	0.578	0.5633 0.9285 0.3246	-+ -+ ++		
Y-Axis: P-value	Scale: -log10							
			Y-axis	P-value	e scaled	-log10		
Suggestive (Blue Line) >	5	Y-Axis ‡				0		
Significant (Red Line) 🕞 🗧	7.3	Y-Axis 🕴	Sugge	stive at	: 5			
			Significant at 7.3					
Data Point Size: Small	Color Key:	\$	Signin	cant at	7.5			
Show Gridlines Width: 75	0 Height:	300						
	Remuse		2.0	-1.049	0.294		- 11	
Export to SV	u: biomac		2.0	-0.509	0.611		1	
			2.0	-0.656	0.5116	+-	*	
OK	Cancel		2.0	-0.785	0.4323			
99991 results								
Chr: + Start kb: End kb: Filter: + + Filter								
Specify Marker: Prune Table Remove Column:  (Remove								
Additional Results	Combine P-Values	Ð	Plot	Res	iet (	View Active Filte	rs	

\* Which SNP has the lowest p-value?