The University
Of Queensland
AUSTRALIA

Methylome-wide Association Studies

Practical 1: OSCA

Overview

- This practical will cover
    - Using OSCA to perform MWAS – linear and MOA models
    - OREML and prediction using ‘omics data

- See https://yanglab.westlake.edu.cn/software/osca/ for command reference

OmicS-data-based Complex trait Analysis

- OSCA
- A software tool for the analysis of complex traits using multi-omics data and genetic analysis of molecular phenotypes
- Supports a range of methods
    - Estimating the epigenetic (or transcriptomic) relationships between individuals from genome-wide data
    - Estimating the proportion of phenotypic variance for a complex trait that can be captured by all ‘omics data
    - Mixed linear model-based approaches to test for associations between ‘omics data and a trait
    - Prediction of phenotypes
    - ...

OSCA

- Syntax similar to PLINK and GCTA.

- Data for practical found in `/data/module2/ewas`
- Three input files to describe ‘omics data
  - `ewas.bod` → binary representation of ‘omics data
  - `ewas.oii` → like PLINK .fam file. FID, IID PID, MID, sex
  - `ewas.opi` → like PLINK .map file. Chromosome, probe ID, position, gene, direction

ORM

- ‘Omics Relationship Matrix
- Measures genome-wide similarity between pairs of individuals

```
- osca --befile /data/module2/ewas/ewas --make-orm
```

- Outputs 3 files:

```
- osca.orm.bin
- osca.orm.N.bin
- osca.orm.id
```

OREML

- How much of the variation in a trait can **be captured** by ‘omics variation
- Note – **NOT** how much variation can **be explained** – there is a causality issue.

```
osca --orm ewas –reml --pheno /data/module2/ewas/age.pheno
osca --orm ewas –reml --pheno /data/module2/ewas/smoking.pheno
```

- What proportion of trait variation is captured by the ‘omics data?
- What is the likely direction of effect for the two traits?

EWAS – Linear Regression

- Simple regression without covariates

```
- osca --befile /data/module2/ewas/ewas \
       --pheno /data/module2/ewas/???.pheno \
       --linear
```

- Create a “Manhattan” and QQ plot using the “qqman” library in R
- What is the Genomic Inflation Factor?

EWAS – Linear Regression with Covariates

- **WARNING: SLOW!   May not be practical in class…**

- Add covariates

```
... --linear --covar /data/module2/ewas/ewas_covariates.txt
```

- Calculate PCs and add as covariates

```
osca --orm osca --pca 20
```

```
... --linear --covar <file> --qcovar osca.eigenvec
```

EWAS – MOA Method

- Command: `--moa`

- Plot the Manhattan and QQ plots again
- What is the GIF now?

EWAS – MOMENT Method

- **WARNING: SLOW!   Unlikely to finish running today...**

- Command: --moment

- Can try running to see the first step
- How many probes are included in each component?