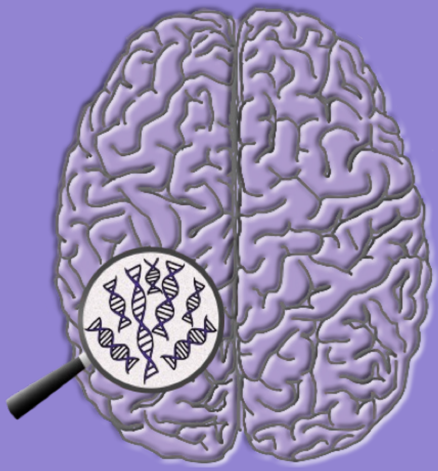
Genome-Wide Association Studies (GWAS)

Yuetiva Robles, PhD

Basics of Genetic Analysis: SNP analysis and GWAS

July 25, 2025



Refresher: SNP Analysis

- Usually logistic or linear regression
- SNP usually coded as an additive 0, 1, 2 based on number of minor alleles a person has

Outcome = Intercept + SNP + covariates

Refresher: GWAS – multiple testing

- Individual SNP tests repeated hundreds of thousands to millions
- Traditional Bonferroni threshold 5×10^{-8} based on assumption of 1 million independent tests
- Alternatives available including FDR, permutation, etc

PLINK

PLINK – files

- Genetic files:
  - bed (binary/compressed)
  - bim (text-based)
  - fam (text-based)
- Best practice – ONLY use PLINK to modify
- Input files (can be combined):
  - Outcome
  - Covariates
  - Missing data = -9 default; only numeric data (“dummy variables”)

PLINK commands deconstructed

```
$./plink --bfile ADNI1.impQC.rs --adjust --linear hide-covar --covar ADNI1_demographics_immersive.txt --covar-name AGE,SEX,C1,C2,C3,C4 --pheno ADNI1_demographics_immersive.txt --pheno-name Whole_hippocampus --pfilter 1 --maf 0.02 --out ADNI1_HV_GWAS_p1_maf02
```

*specific to project
*optional, recommended

PLINK commands deconstructed

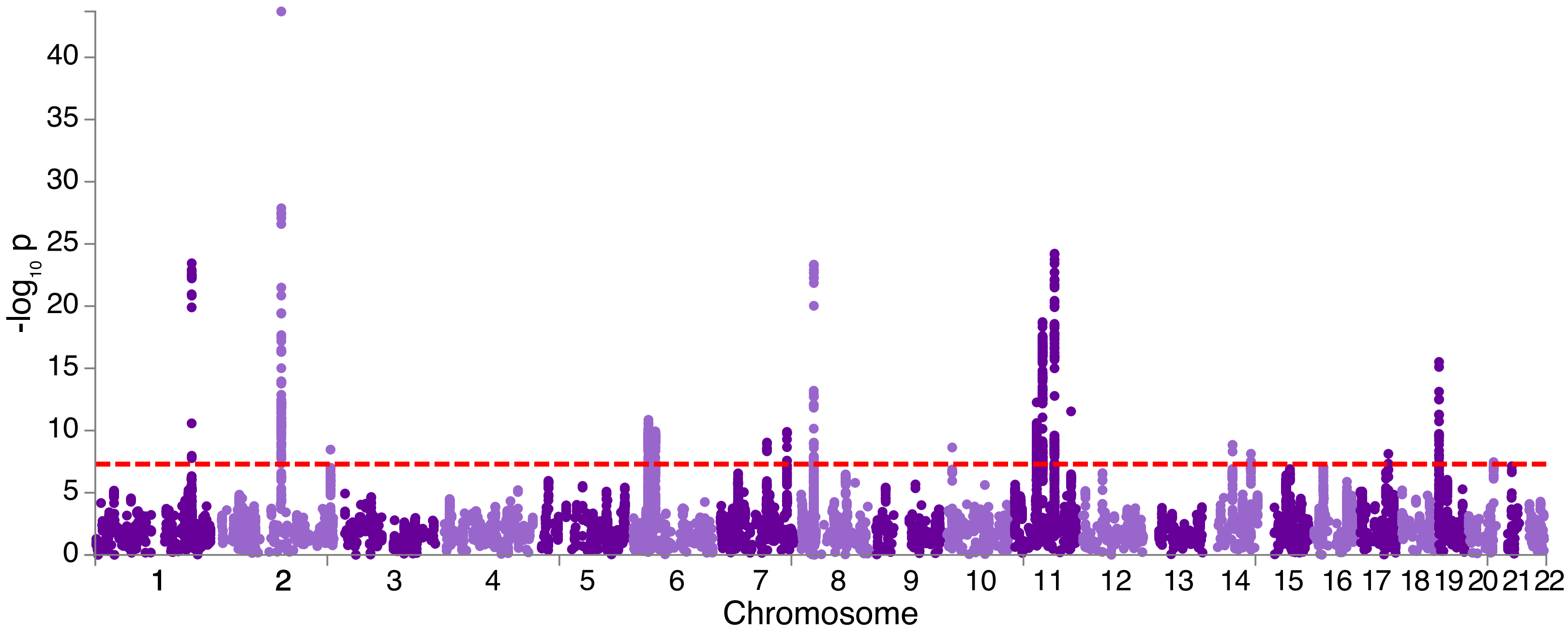
```
$ ./plink \

$ --bfile ADNI1.impQC.rs \

$ --adjust \

$ --linear hide-covar \     (or --logistic hide-covar \)

$ --covar ADNI1_demographics_immersive.txt \

$ --covar-name AGE,SEX,C1,C2,C3,C4 \

$ --pheno ADNI1_demographics_immersive.txt \

$ --pheno-name Whole_hippocampus \

$ --pfilter 1 \

$ --maf 0.02 \

$ --out ADNI1_HV_GWAS_p1_maf02
```

*specific to project
*optional, recommended

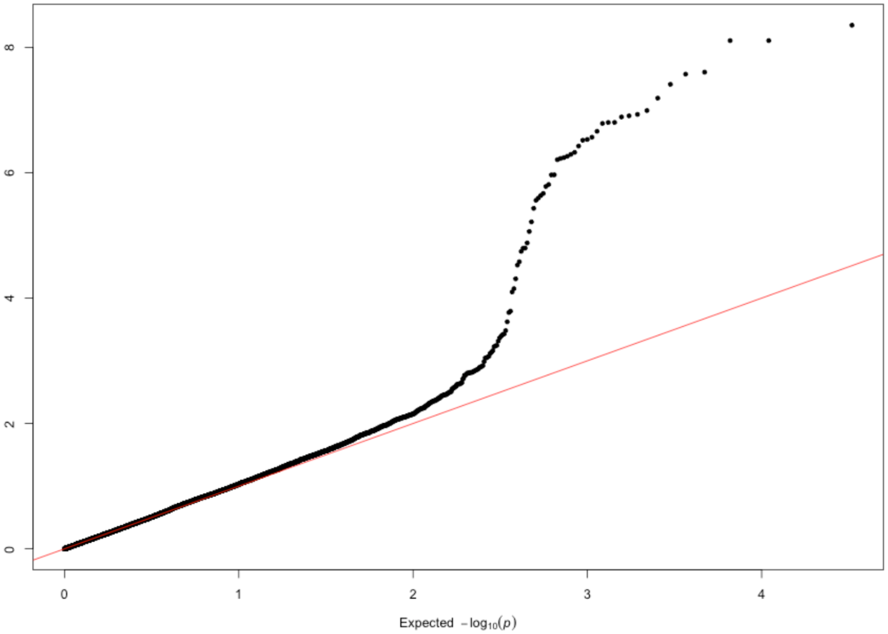
Post-GWAS reporting

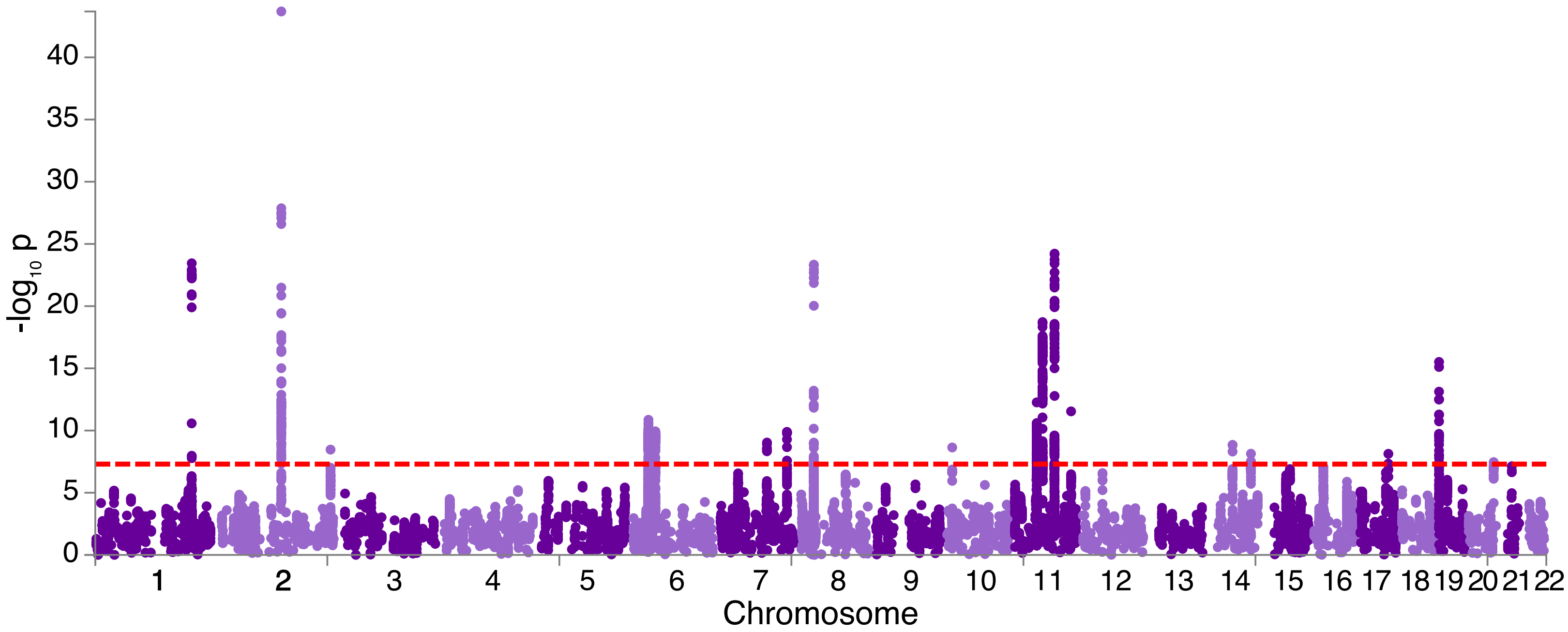
Basic plotting and reporting

*more about post-GWAS after PRS tutorial

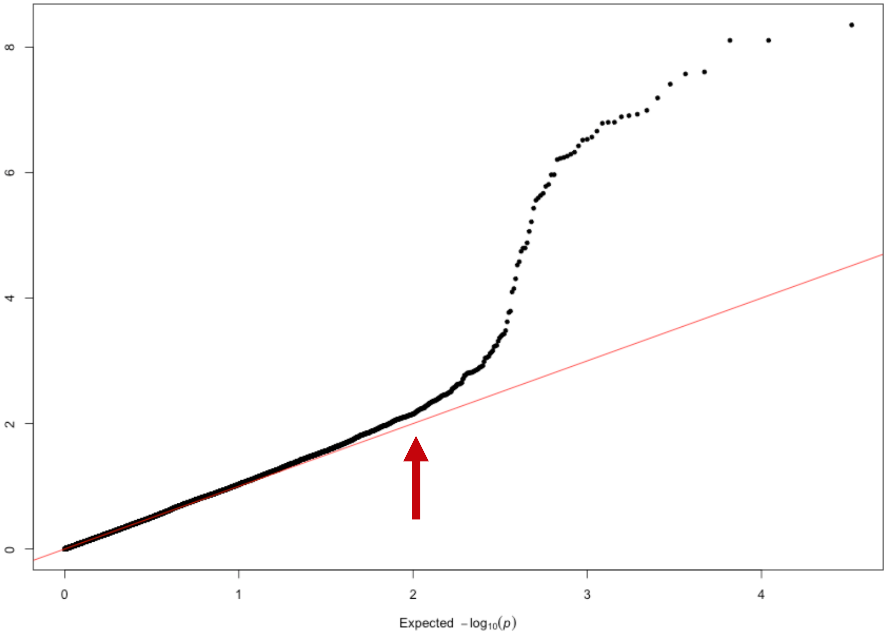


- Created using `qqman` R package; other packages available
- Can also use online tools like FUMA



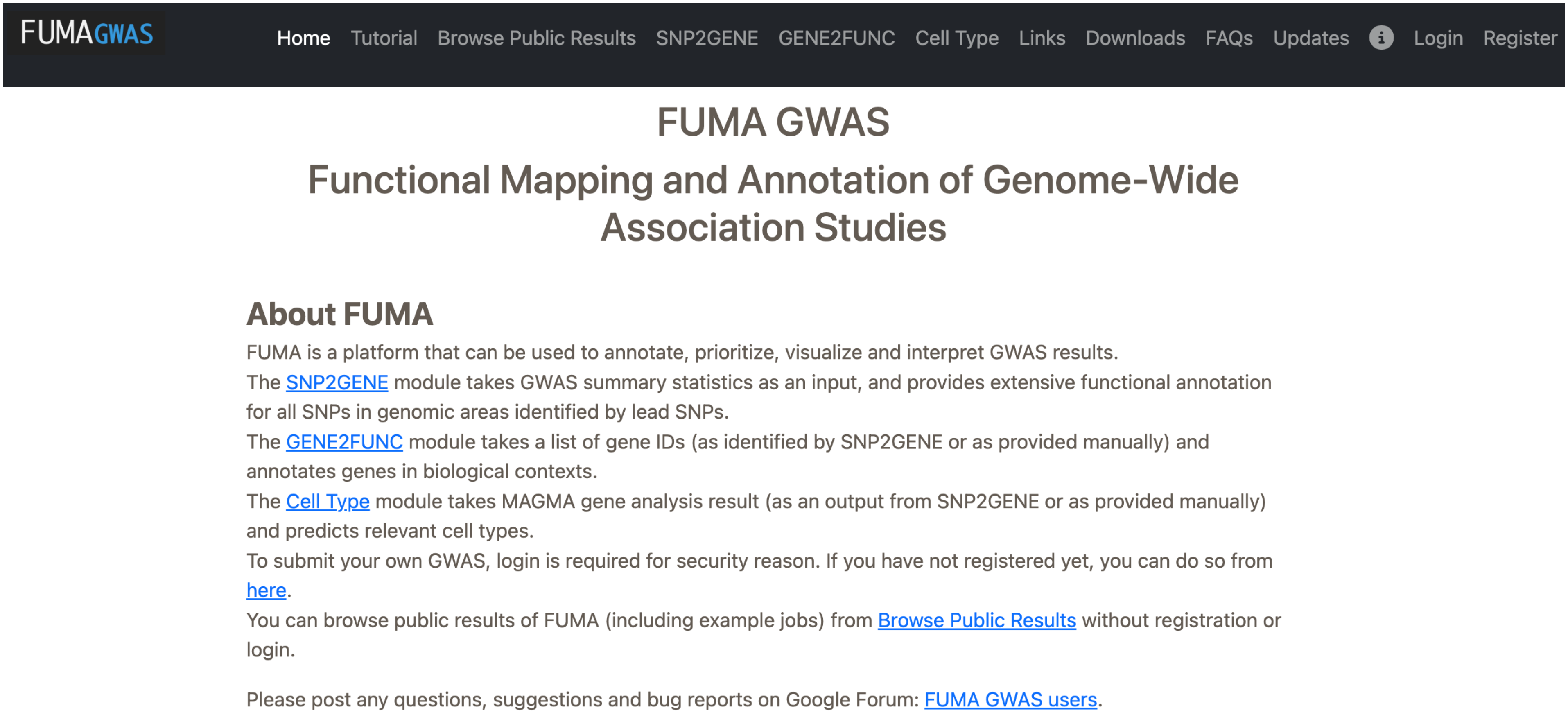


- Created using `qqman` R package; other packages available
- Can also use online tools like FUMA



GWAS reporting – FUMA

https://fuma.ctglab.nl/



FUMA GWAS

Functional Mapping and Annotation of Genome-Wide Association Studies

About FUMA

FUMA is a platform that can be used to annotate, prioritize, visualize and interpret GWAS results.
The SNP2GENE module takes GWAS summary statistics as an input, and provides extensive functional annotation for all SNPs in genomic areas identified by lead SNPs.
The GENE2FUNC module takes a list of gene IDs (as identified by SNP2GENE or as provided manually) and annotates genes in biological contexts.
The Cell Type module takes MAGMA gene analysis result (as an output from SNP2GENE or as provided manually) and predicts relevant cell types.
To submit your own GWAS, login is required for security reason. If you have not registered yet, you can do so from here.
You can browse public results of FUMA (including example jobs) from Browse Public Results without registration or login.

Please post any questions, suggestions and bug reports on Google Forum: FUMA GWAS users.

Note: I am not affiliated in anyway, just a handy tool I like