Fine Mapping With TWAS Results Across Multiple Tissues

Shuai Li, Xinyu (Brian) Guo

Johns Hopkins Bloomberg School of Public Health

May 14, 2021

Section 1

Background and Methods

What is TWAS?

- TWAS: transcriptome-wide association study.
- To determine significant trait-expression associations. [1]
- This method increases the power of identifying functionally relevant loci by leveraging expression quantitative trait loci (eQTLs) from external references in relevant tissues. [2]

[1] Gusev et al. "Integrative approaches for large-scale transcriptome-wide association studies" 2016 Nature Genetics

[2] Bhattacharya et al. "A framework for transcriptome-wide association studies in breast cancer in diverse study populations" 2020 Genome Biology

TWAS/FUSION Software

- Functional Summary-based Imputation:

  FUSION is a suite of tools for performing a TWAS by predicting functional/molecular phenotypes into GWAS using only summary statistics (usually from GWAS). The goal is to identify associations between a GWAS phenotype and a functional phenotype that was only measured in reference data. [1]

[1] Gusev et al. "Integrative approaches for large-scale transcriptome-wide association studies" 2016 Nature Genetics

TWAS/FUSION Software

- Inputs:
  - GWAS summary statistics
  - Reference panels (i.e. precomputed functional weights (primarily gene expression) from multiple tissues)
  - Regerence LD data
- Outputs:
  - A data frame with corresponding z and p values for each SNPs.

[1] Gusev et al. "Integrative approaches for large-scale transcriptome-wide association studies" 2016 Nature Genetics

Bayesian Fine Mapping

- Why fine-map?
  - To find causal genes
  - To pinpoint variant
  - To understand genetic architecture
    - Gene enrichment
    - Cross-trait comparison, cross-tissue
- Bayesian fine-mapping outputs:
  - PIP: Posterior inclusion probability (the probability that a variant is causal)
  - 95% Credible Sets: Set of variants that contains $\geq 95\%$ probability

[3] ▸ Hilary Finucane, Broad Institute

Bayesian Fine Mapping

- Single-causal-variant PIPs:

$$
\begin{aligned}
PIP_j &= P(\text{j causal} \mid \text{data}) \\
&= \frac{P(\text{data} \mid \text{j causal})}{\sum_k P(\text{data} \mid \text{k causal})} \\
&= \frac{P(\text{data} \mid \text{j causal})/P(data|nocausal)}{\sum_k P(\text{data} \mid \text{k causal})/P(data|nocausal)} \\
&= \frac{\text{Bayesian Factor}_j}{\sum_k \text{Bayesian Factor}_k}
\end{aligned}
$$

[3] ▸ Hilary Finucane, Broad Institute

Bayesian Fine Mapping

- 95% Credible Sets (S):

$$P(\text{causal var is in S}) \geq 0.95$$

- Under Single-causal-variant assumption:

$$P(\text{causal var is in S}) = \sum_{j \in S} PIP_j$$

- To get the most compact credible set, add variant with highest PIPs untill sum to 0.95.

[3] ▸ Hilary Finucane, Broad Institute

Bayesian Fine Mapping

- Factors affecting Bayesian fine mapping power
  - LD
  - Sample Size
  - Effect size

[4]Schaid et al. Nat Rev Genet 2018

Bayesian Fine Mapping

- Multiple-causal-variant Fine-mapping (two approaches):
  - Divide the whole data into many pieces, and apply single-causal-variant fine-mapping in each piece
  - Jointly model Multiple-causal-variant

[3] ▸ Hilary Finucane, Broad Institute

Section 2

Data For Alzheimer's Disease

Section 3

Analysis (SuSiE)

Section 4

Reference

Reference

- Gusev et al. "Integrative approaches for large-scale transcriptome-wide association studies" 2016 Nature Genetics
- Wang, G., Sarkar, A., Carbonetto, P., & Stephens, M. (2020). A simple new approach to variable selection in regression, with application to genetic fine mapping. Journal of the Royal Statistical Society: Series B (Statistical Methodology). https://doi.org/10.1111/rssb.12388
- Schaid, D.J., Chen, W. & Larson, N.B. From genome-wide associations to candidate causal variants by statistical fine-mapping. Nat Rev Genet 19, 491–504 (2018). https://doi.org/10.1038/s41576-018-0016-z
- ▸ Hilary Finucane, Broad Institute