

TIGAR-V2: Efficient TWAS Tool with Nonparametric Bayesian eQTL Weights of 49 Tissues from GTEx V8

Jingjing Yang¹, Randy L. Parrish¹, Greg C. Gibson², Michael P. Epstein¹

1. Emory University School of Medicine; 2. Center for Integrative Genomics, Georgia Tech



TIGAR-V2 TWAS Tool: <https://github.com/yanglab-emory/TIGAR>

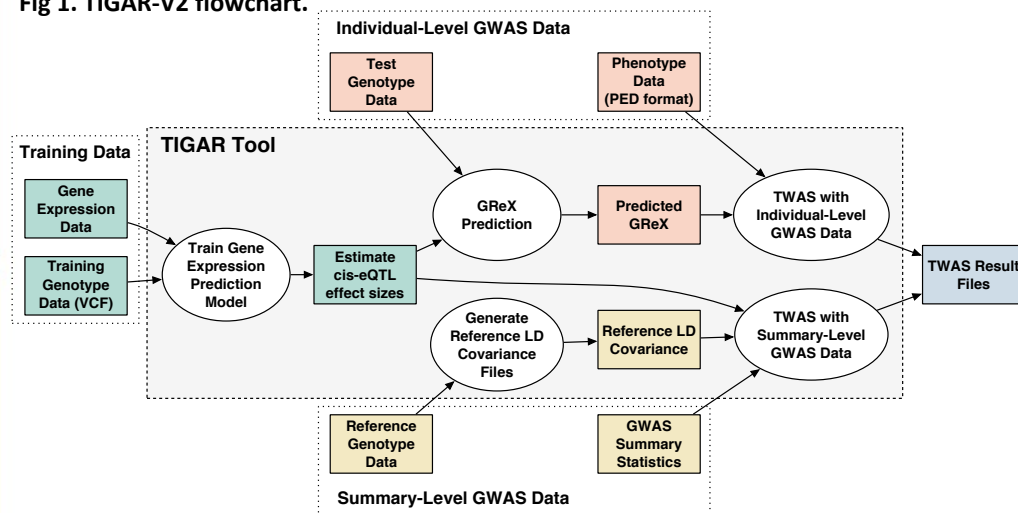
1. Fit Gene Expression Prediction Model

- Efficiently take WGS VCF genotype files (bgzipped/tabixed).
- Train both nonparametric Bayesian DPR and Elastic-Net models with 5-fold cross validation.

2. TWAS: Gene-Based Association Test with eQTL weights and GWAS data

- Burden tests: S-PrediXcan and FUSION Z-score Statistics
- Variance Component tests.

Fig 1. TIGAR-V2 flowchart.



Train Bayesian DPR Models with GTEx V8 Reference Data

- Trained gene expression prediction models (eQTL weights) for 49 human tissues with samples size >70.
- Bayesian DPR method trained significant prediction models with 5-fold cross validation $R^2 > 0.5\%$ for 2X genes than PrediXcan (predictdb.org) (Fig2A).
- Bayesian DPR model costs ~100 CPU hours per Chromosome per tissue (Fig2B).

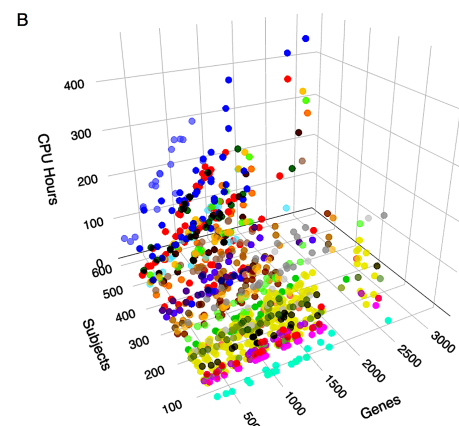
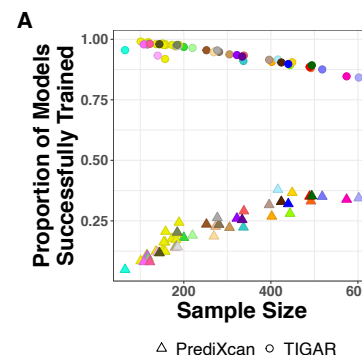


Fig 2. (A). Proportions of significant gene expression prediction models trained per tissue by Elastic-Net (PrediXcan) and Bayesian DPR methods; (B). CPU hours per tissue per Chromosome to train Bayesian DPR models.



TWAS of Breast and Ovarian Cancer

- Used Bayesian DPR eQTL weights estimated from GTEx of breast and ovarian tissues.
- Used GWAS summary statistics of Breast Cancer (BCAC) and Ovarian Cancer (OCAC).

Independently significant TWAS genes

- Breast Cancer:** 34 significant genes with 3 novel genes: KLHL25, UBE2MP1, FRG1EP (Fig3A).
- Ovarian Cancer:** 4 significant genes with 2 novel genes: UBE2MP1 (CHR 16) and FRG1EP (CHR 20) shared with Breast cancer (Fig3B).

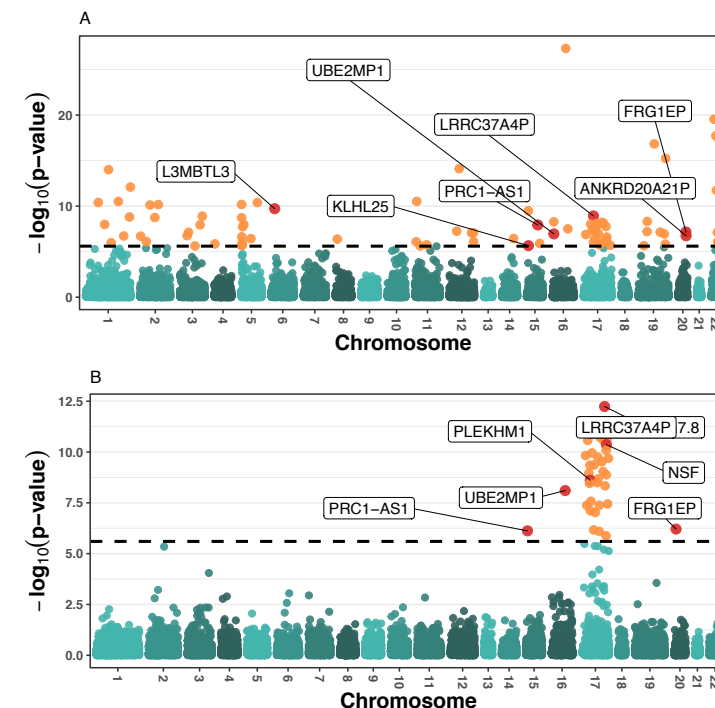


Fig 3. TWAS results of Breast Cancer (A) and Ovarian Cancer (B).