Nothing to disclose.

Modeling tissue and gene co-regulation reveals causal tissues for disease

Tiffany Amariuta, PhD

Assistant Professor
University of California San Diego
Halıcıoğlu Data Science Institute / Department of Medicine
October 28, 2022

Preprint of presented work: Amariuta 2022 bioRxiv

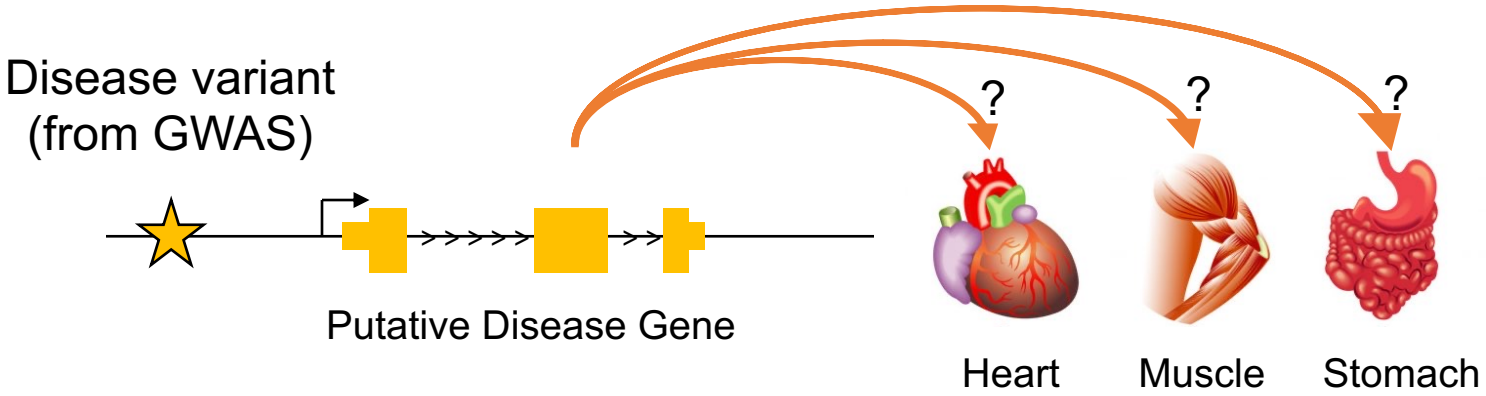
Outline

1. Background
2. Method and Simulations
3. Results on Real Traits

Outline

1. **Background**
2. Method and Simulations
3. Results on Real Traits

Inferring causal tissues is an important goal



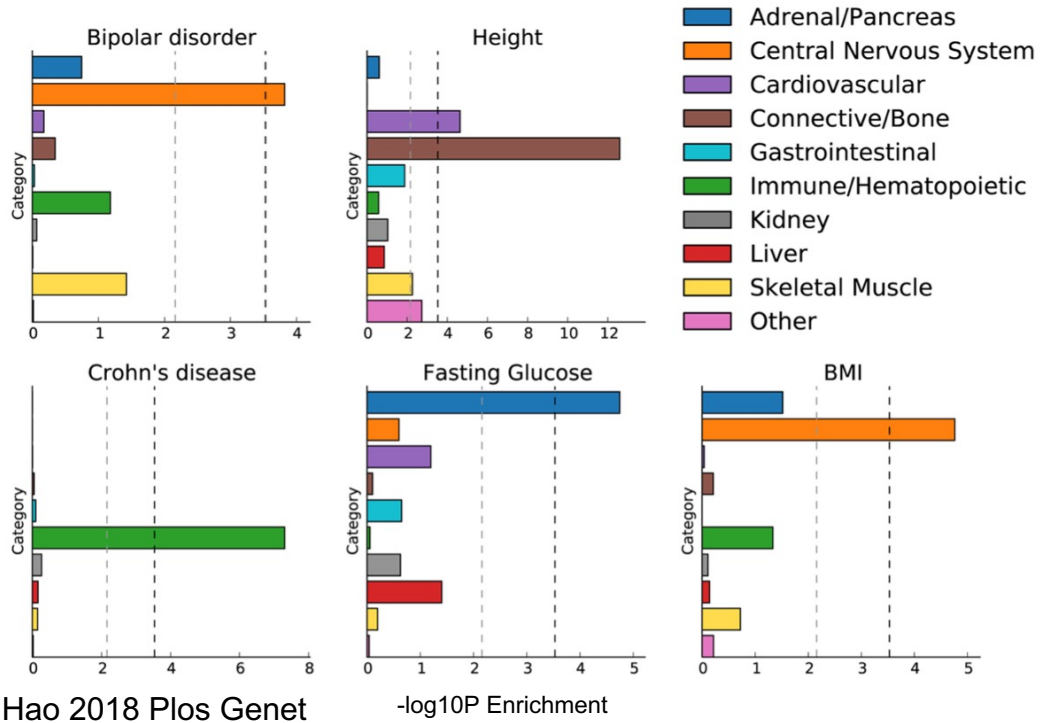
Knowing the causal disease tissue is crucial to

1) Describing the biological mechanism of disease
2) Selecting the cellular context in which to perform experiments
3) Determining the physiological target of pharmaceuticals

Claussnitzer 2020 Nature, Umans 2020 Trends in Genetics, Hekselman 2020 Nat Rev Genet

For most diseases, disease-associated cell types have been identified

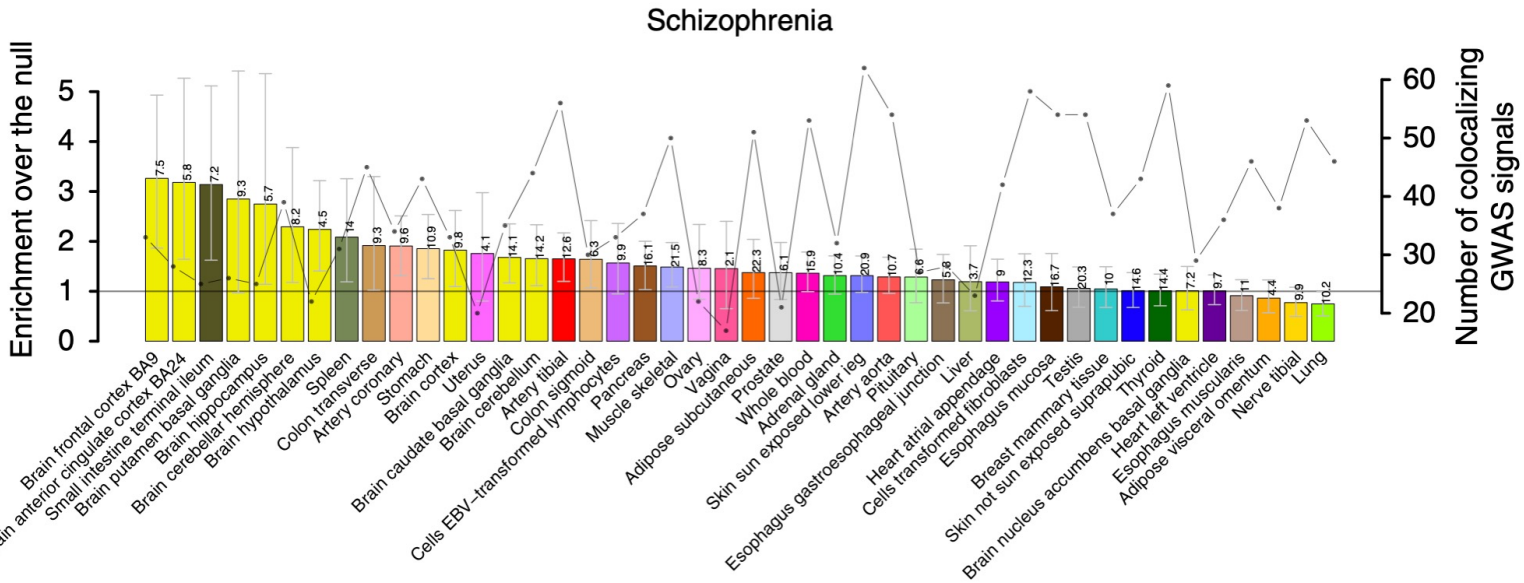
Using chromatin accessibility, bone chondrocytes are the most strongly associated cell type for height.



Finucane 2015 Nat Genet
Also see, Trynka 2013 Nat Genet, Hao 2018 Plos Genet

For most diseases, disease-associated cell types have been identified

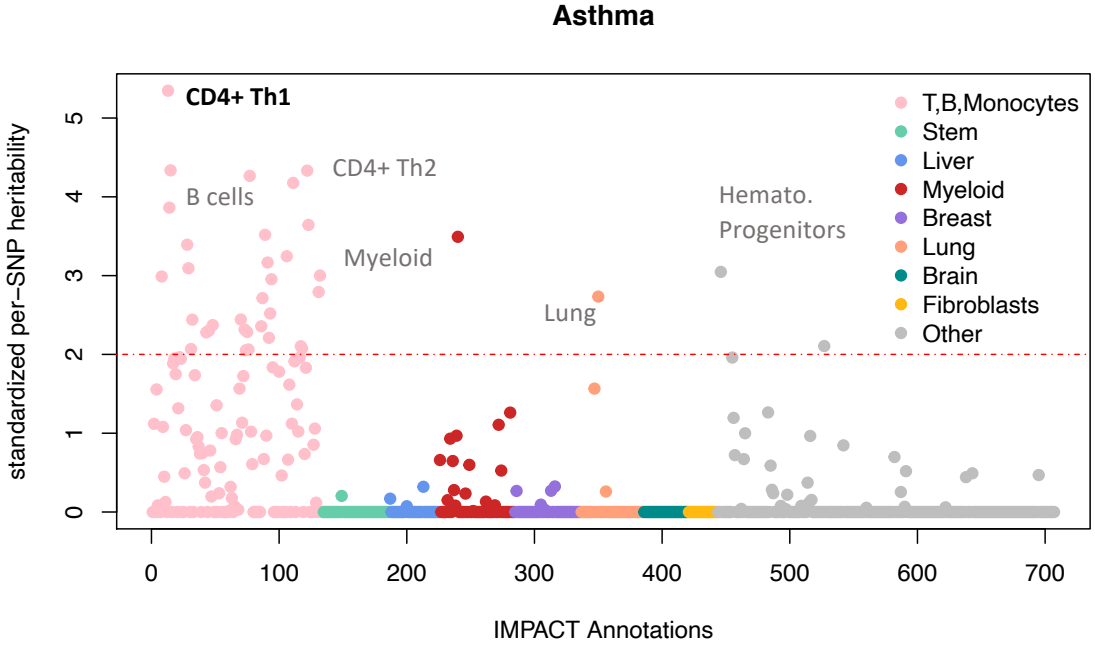
Using gene expression, the frontal cortex (BA9) of the brain is the most strongly associated tissue for schizophrenia.



Ongen 2017 Nat Genet
Also see, Finucane 2018 Nat Genet, Gamazon 2018 Nat Genet, Hormozdiari 2018 Nat Genet, Arvanitis 2022 AJHG

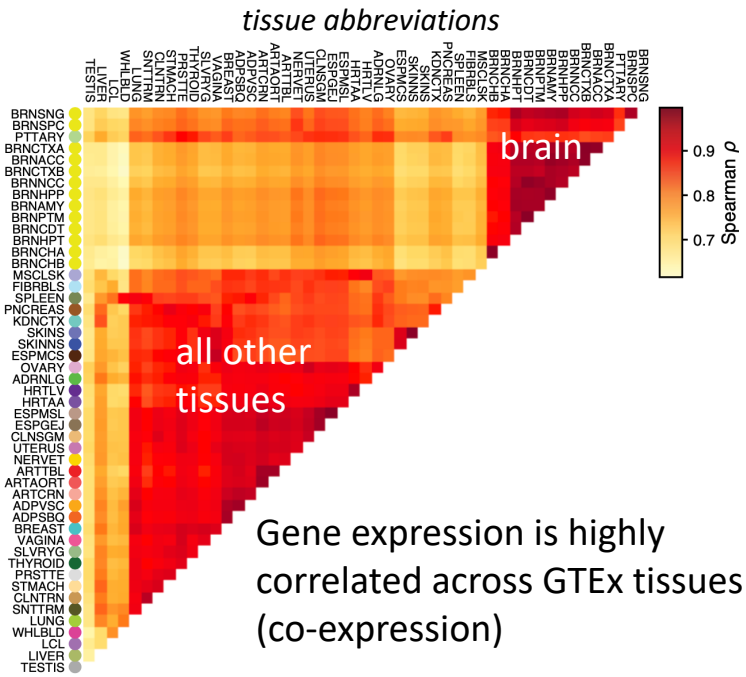
For most diseases, disease-associated cell types have been identified

Using enhancer predictions, CD4+ Th1 is the most strongly associated cell type for asthma.

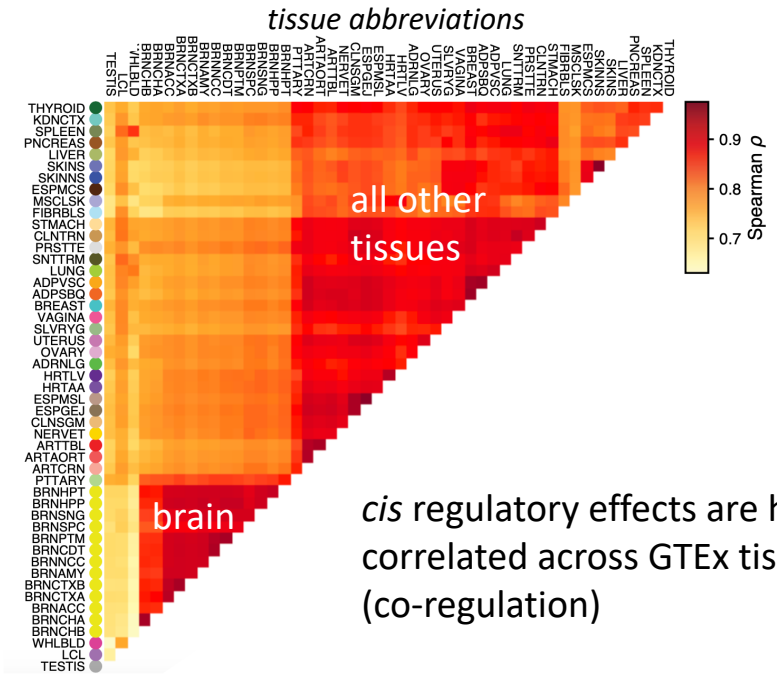


Amariuta, Ishigaki 2020 Nat Genet
Also see, Amariuta 2019 AJHG, Reshef 2018 Nat Genet

High co-regulation across tissues means that many disease-associated tissues may not be causal



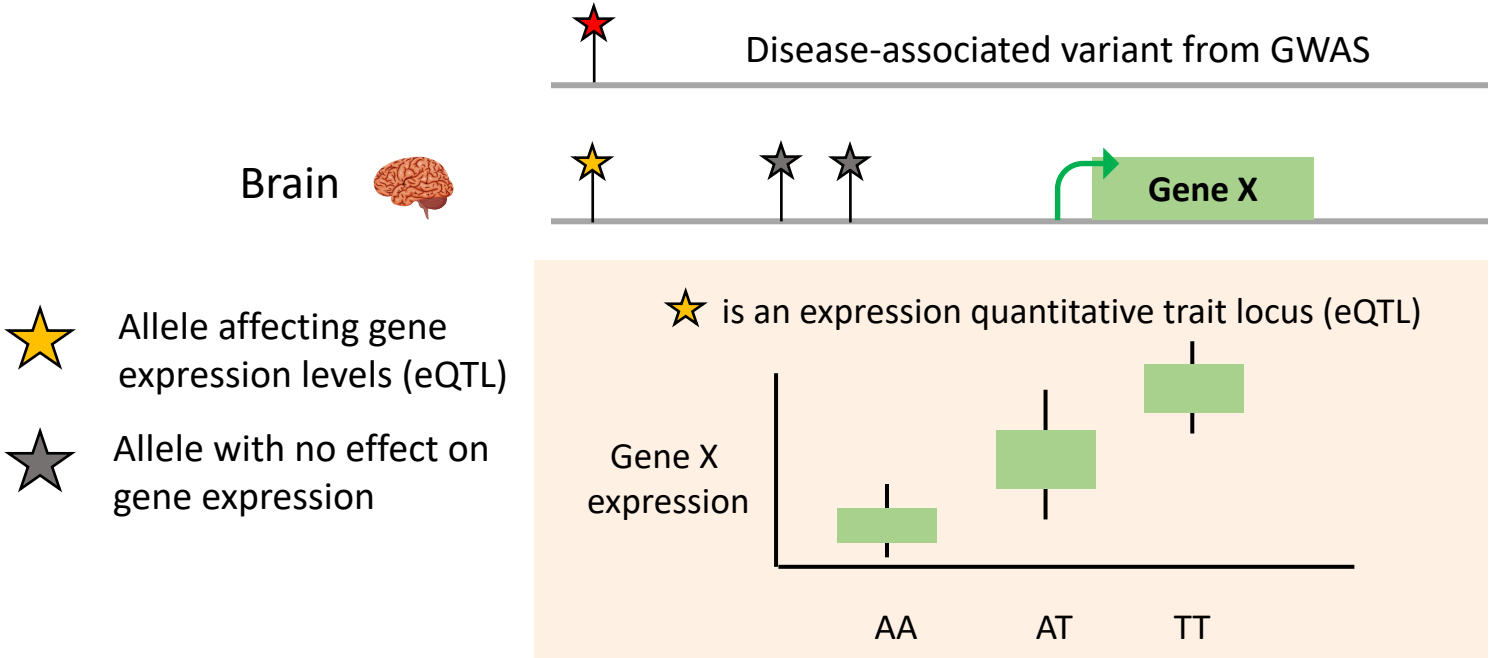
Gene expression is highly correlated across GTEx tissues (co-expression)



cis regulatory effects are highly correlated across GTEx tissues (co-regulation)

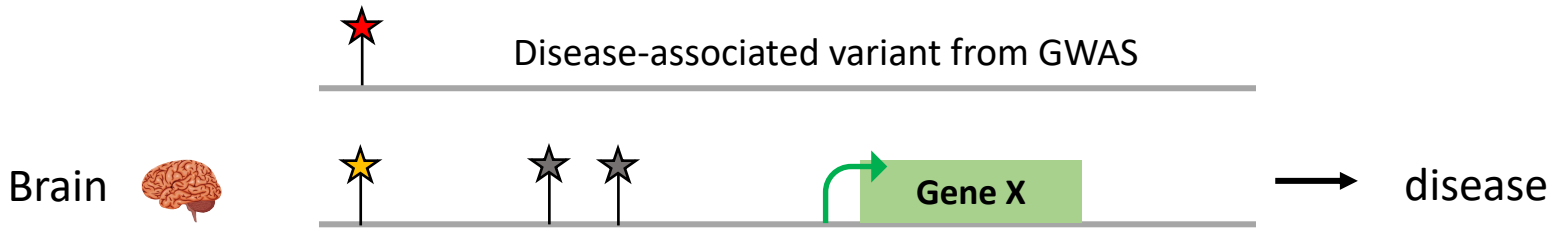
GTEx Consortium 2020 Science
Also see Ongen 2017 Nat Genet, Wainberg 2019 Nat Genet, Arvanitis 2022 AJHG

Colocalization of eQTLs with GWAS variants can implicate disease-critical genes and tissues



Giambartolomei 2014 Plos Genet, Hormozdiari 2016 AJHG

Colocalization analysis is complicated by co-regulation

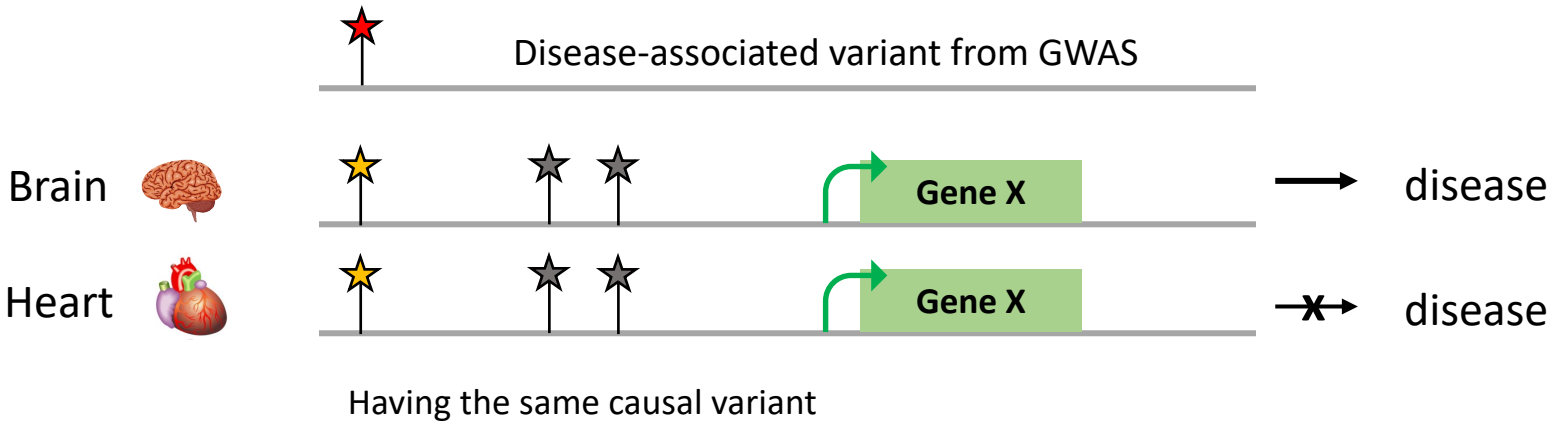


Allele affecting gene expression levels (eQTL)

Allele with no effect on gene expression

Wainberg 2019 Nat Genet
Siewert-Rocks 2022 AJHG

Colocalization analysis is complicated by co-regulation

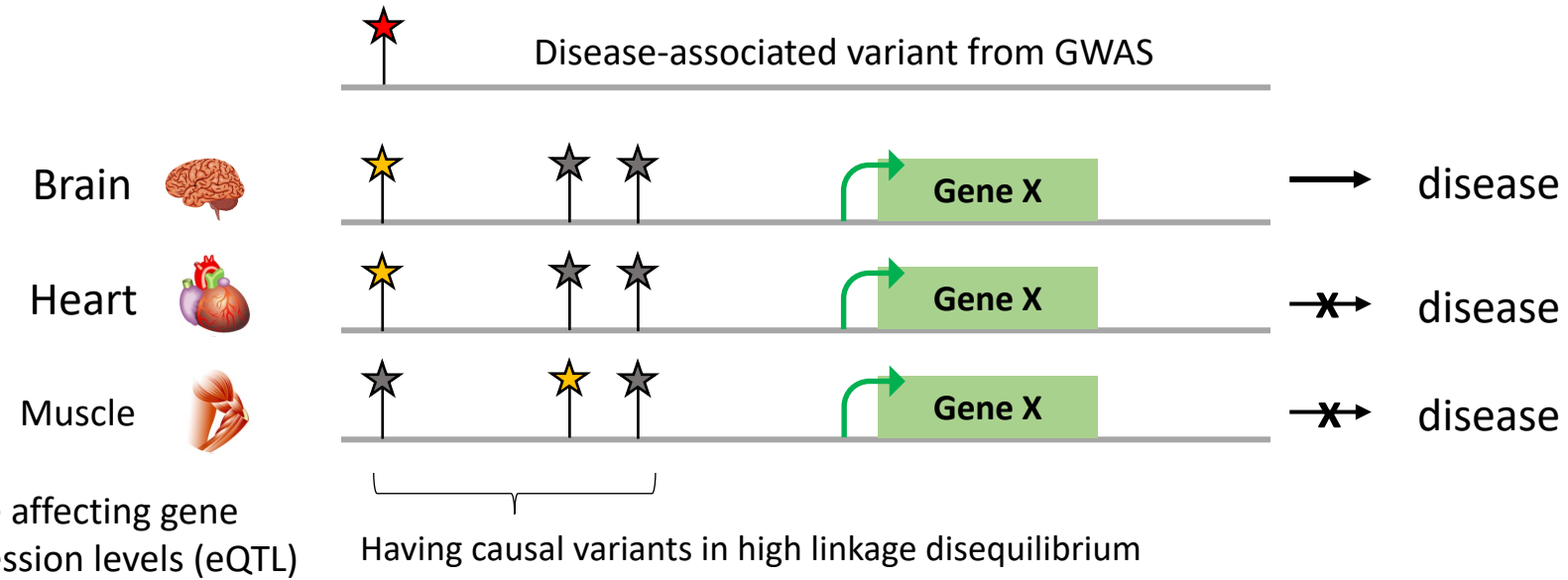


Allele affecting gene expression levels (eQTL)

Allele with no effect on gene expression

Wainberg 2019 Nat Genet
Siewert-Rocks 2022 AJHG

Colocalization analysis is complicated by co-regulation



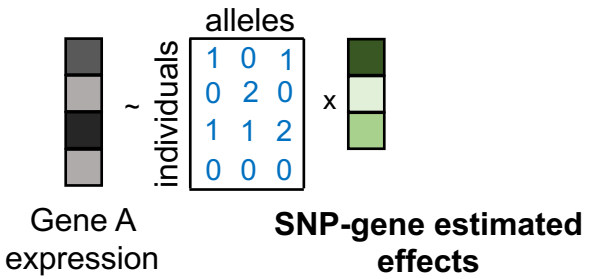
Allele affecting gene expression levels (eQTL)

Allele with no effect on gene expression

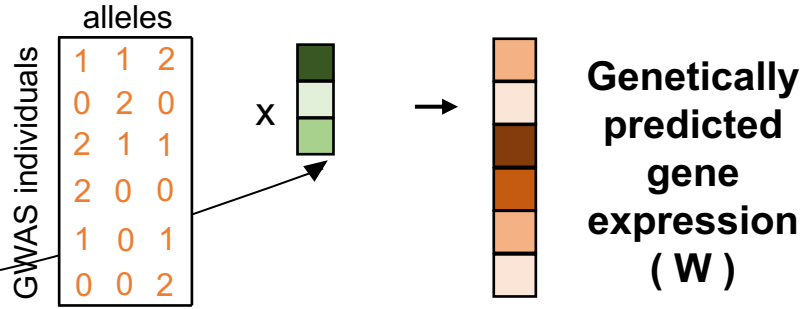
Wainberg 2019 Nat Genet
Siewert-Rocks 2022 AJHG

Transcriptome-wide association studies (TWAS) perform polygenic colocalization of genes with disease

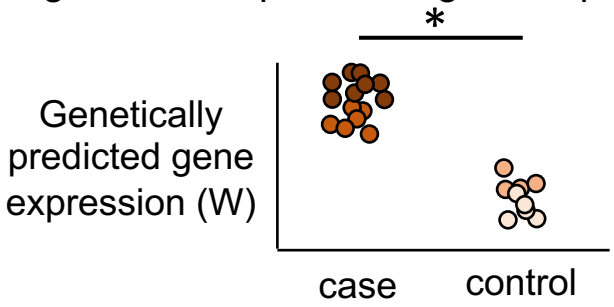
1. Learn SNP-gene weights from eQTL model.



2. Predict gene expression using new genotypes.

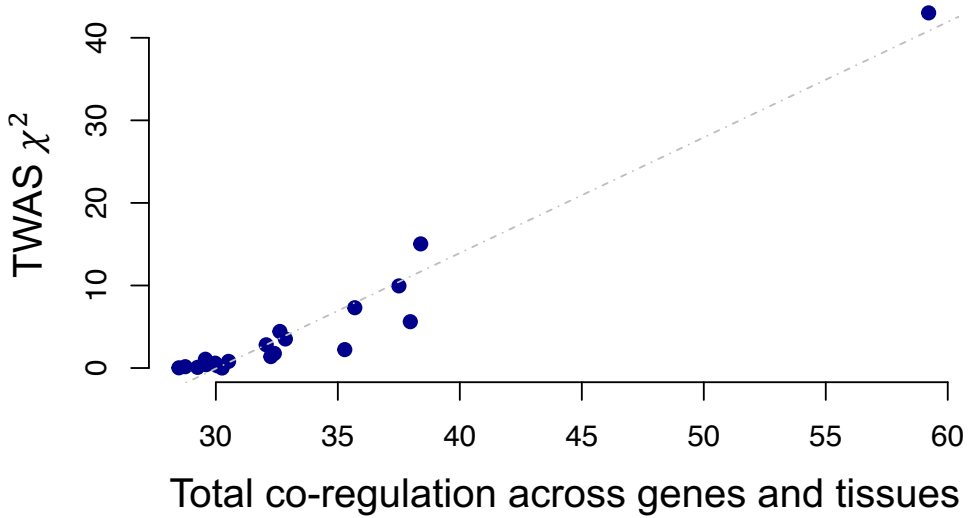


3. Compute correlation of genetic component of gene expression with disease status in GWAS cohort.



Gamazon 2015 Nat Genet, Gusev 2016 Nat Genet

TWAS association statistics are proportional to the amount of tagged causal effects due to co-regulation



TWAS statistics include direct causal effects and tagging effects of co-regulated genes and tissues.

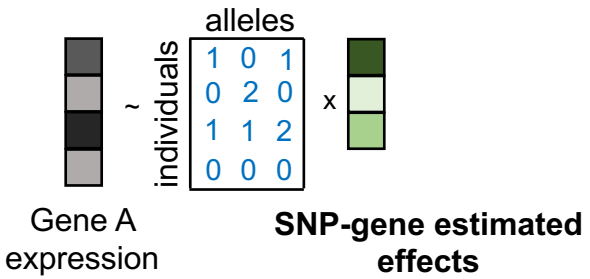
Wainberg 2019 Nat Genet
Amariuta 2022 bioRxiv

Outline

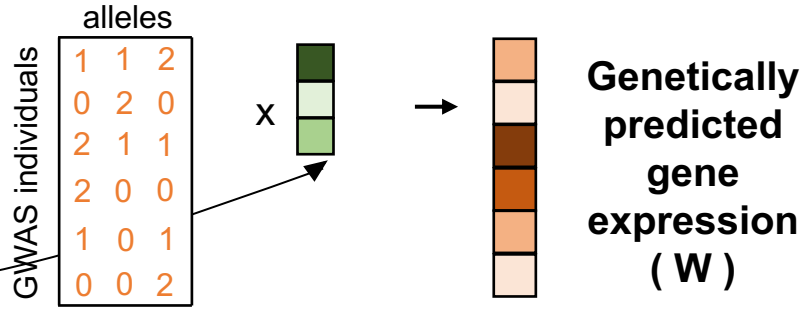
1. Background
2. **Method and Simulations**
3. Results on Real Traits

Co-regulation across tissues and genes can be estimated using gene expression prediction models and a reference panel

1. Learn SNP-gene weights from eQTL model.



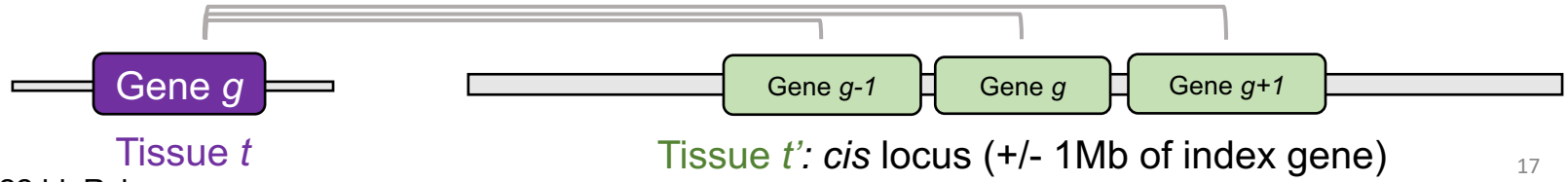
2. Predict gene expression using new genotypes.



3. Compute co-regulation score for a gene-tissue pair with a tissue t'.

r^2 = squared correlation of **W** between two genes.

$$r^2_{g;g-1} + r^2_{g;g} + r^2_{g;g+1}$$



Amariuta 2022 bioRxiv

Tissue co-regulation score regression (TCSC) estimates tissue-specific contributions to disease

$$E\left[\chi^2_{g,t}\right] = N \sum_{t'} \left[l(g,t,t') \frac{h^2_{ge(t')}}{G_{t'}}\right] + 1$$

Gene-tissue association with disease (TWAS)

GWAS sample size

Tissue co-regulation scores

heritable genes in tissue t'

Estimand: Disease heritability explained by predicted gene expression in tissue t'

Amariuta 2022 bioRxiv

Tissue co-regulation score regression (TCSC) estimates tissue-specific contributions to disease

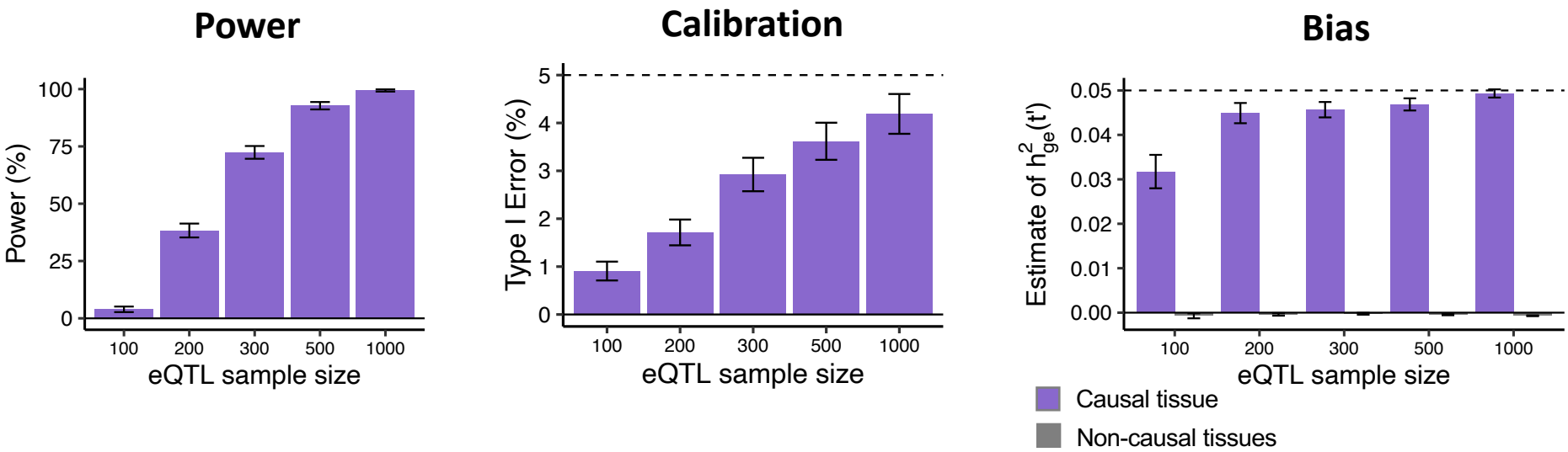
$$E\left[\chi^2_{g,t}\right] = N \sum_{t'} \left[l(g,t,t') \frac{h^2_{ge(t')}}{G_{t'}}\right] + 1$$

Our method determines that tissue t' causally contributes to disease if genes with **high co-regulation** to tissue t' have **higher TWAS χ^2 statistics** and genes with **low co-regulation** to tissue t' have **lower TWAS χ^2 statistics**.

We may identify tissue-specific contributions to the **covariance** of two diseases by regressing products of TWAS z-scores on co-regulation scores.

Amariuta 2022 bioRxiv

TCSC is powerful, well-calibrated, and unbiased in simulations



TCSC has substantially higher power than the Ongen 2017 Nat Genet method.

Bars represent 95% CI.
Mancuso Lab TWAS simulator
Amariuta 2022 bioRxiv (see manuscript for many more simulations)

Outline

1. Background
2. Method and Simulations
3. **Results on Real Traits**

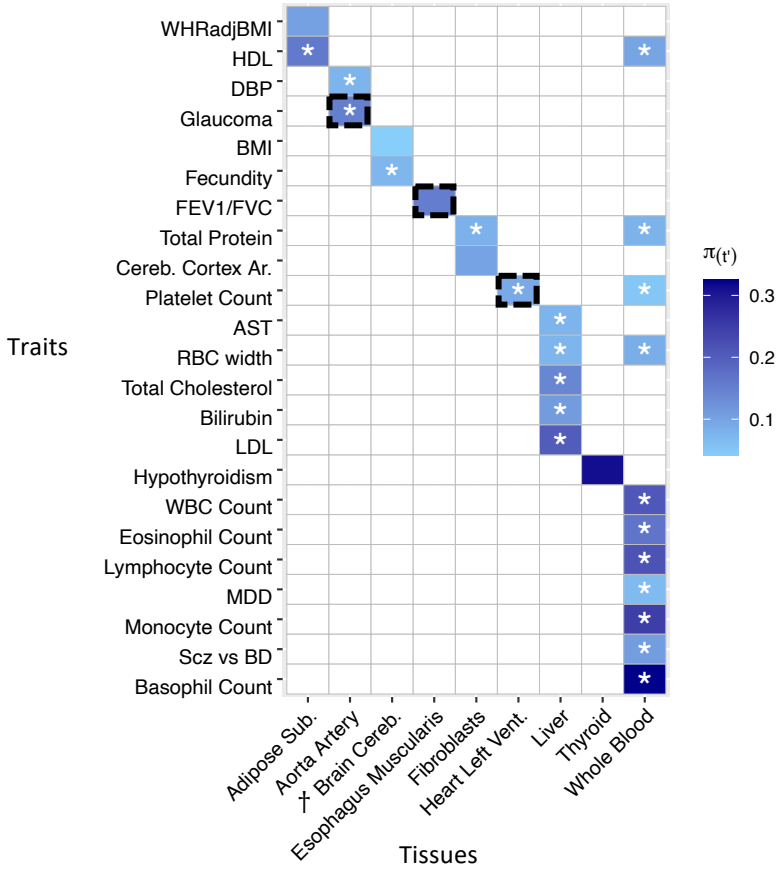
Applying TCSC to real gene expression and trait data

1. We built gene expression prediction models across 48 GTEx tissues, retaining heritable protein coding genes.

2. We analyzed 78 European GWAS summary statistics (average N = 302K).

3. TCSC
   - Finds 27 causal tissue-disease pairs at 10% FDR.
   - Increases the specificity of known tissue-disease associations.

TCSC identifies causal tissue-disease pairs

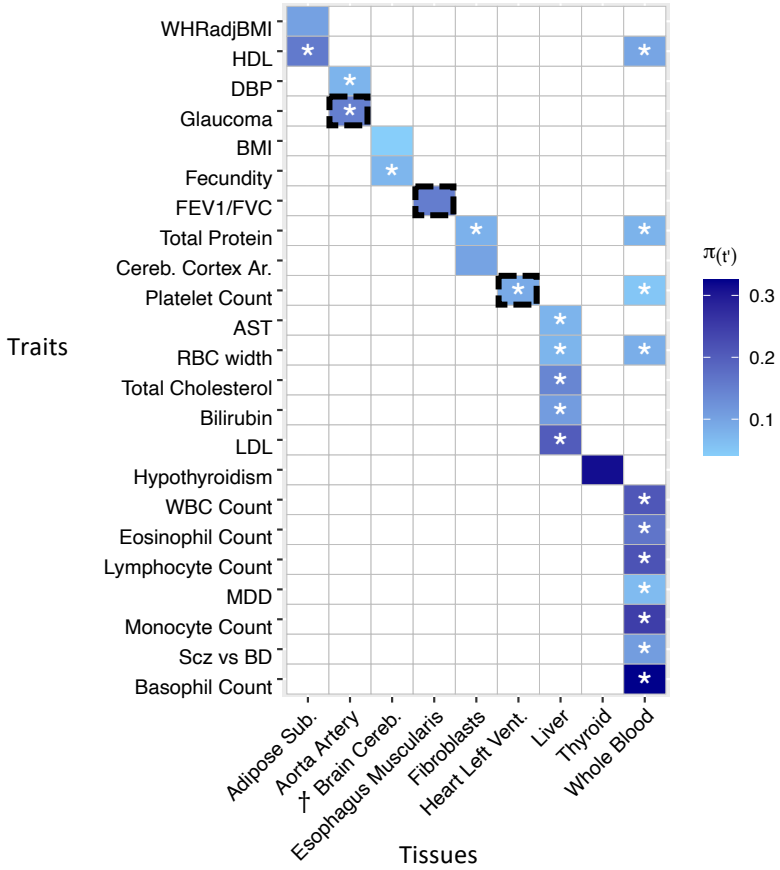


1. **Aorta artery → Glaucoma**:
   - High blood pressure is a known risk factor for glaucoma.
2. **Esophagus muscularis → FEV1/FVC**:
   - Strength of esophageal muscles likely impacts air expulsion rate (FEV1). Analysis of composite traits identified no association with lung capacity (FVC).
3. **Heart ventricle → Platelet count**:
   - Platelets cause blood clots in response to damaged blood vessels; the left ventricle pumps blood out of the heart potentially modifying platelet counts in serum.

Where $\pi(t')$ = proportion of disease heritability explained by predicted expression in tissue t'

Amariuta 2022 bioRxiv

TCSC increases specificity of known tissue-disease pairs

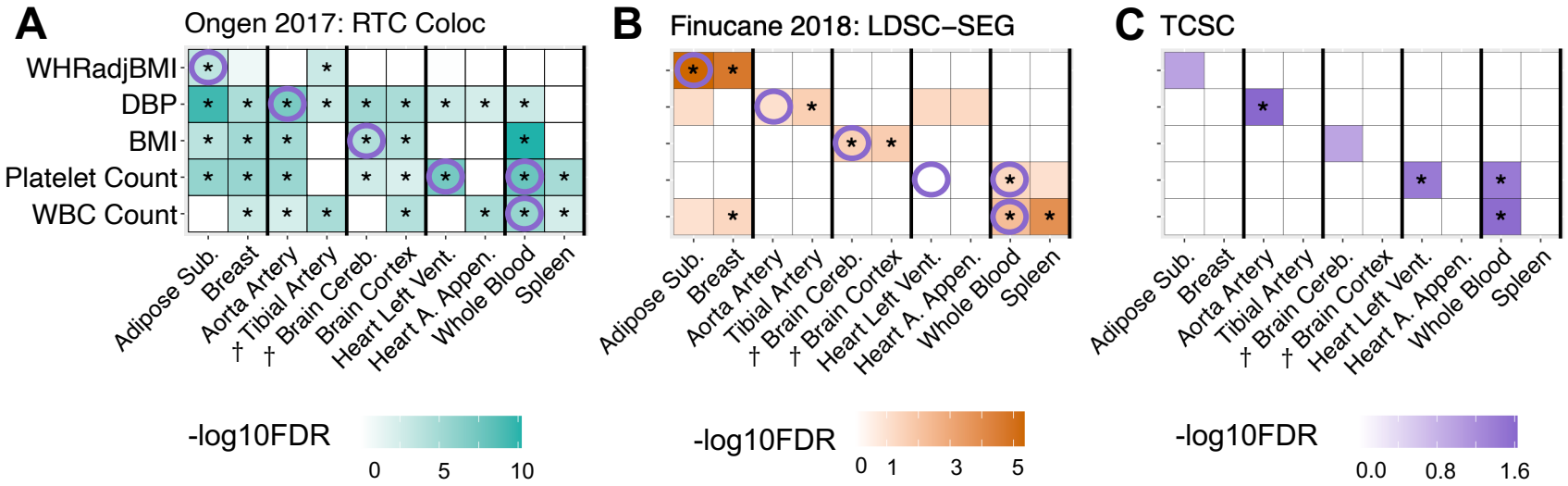


1. **Adipose (subcutaneous) → HDL**:
   - No causal link to adipose visceral omentum. Link between HDL and subcutaneous may involve adiponectin.
2. **Adipose (subcutaneous) → WHRadjBMI**:
   - No causal link with any other metabolic tissue.
3. **Brain (cerebellum) → BMI**:
   - Previous studies have identified generic associations with the central nervous system.

Where $\pi(t') =$ proportion of disease heritability explained by predicted expression in tissue t'

Amariuta 2022 bioRxiv

Other methods are less specific in implicating tissue-disease pairs



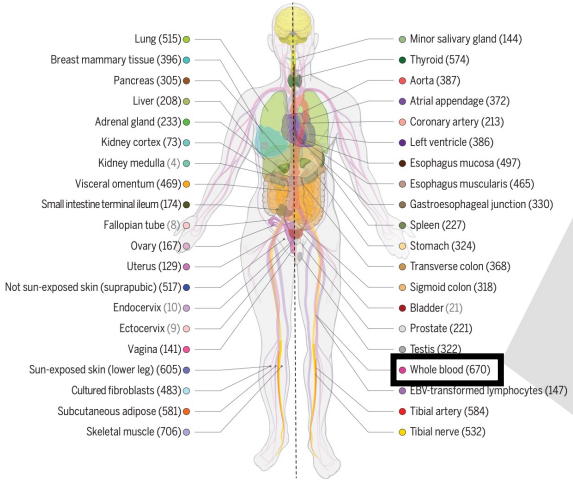
Black lines separate tissues with high genetic correlation. Purple circle is TCSC tissue-disease pair.

The remaining complex traits and diseases have similar patterns.

Amariuta 2022 bioRxiv

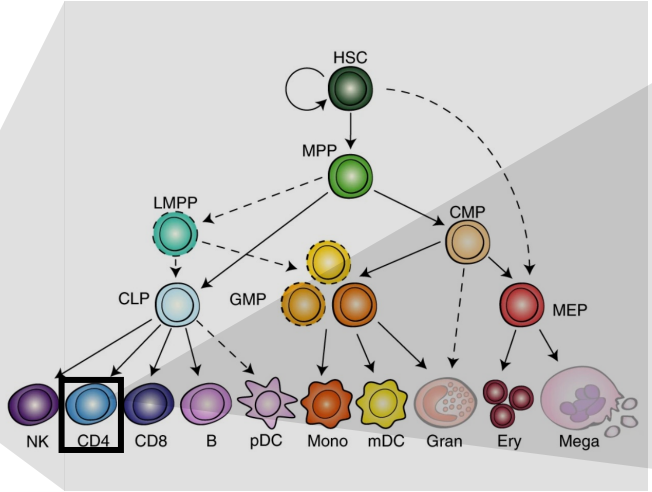
Higher resolution makes causal inference more difficult, but TCSC can be applied here

Tissues



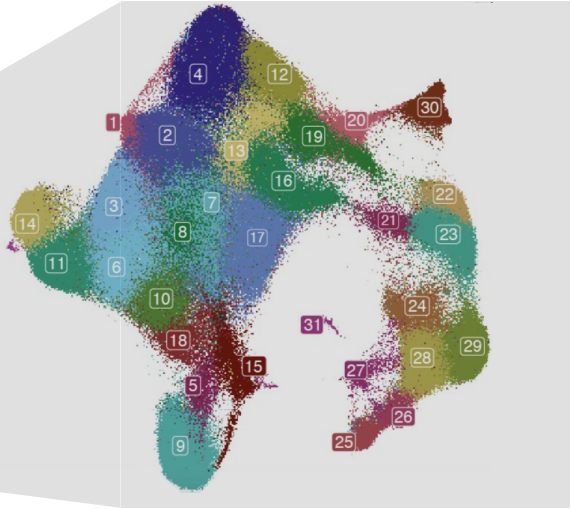
GTEx Consortium 2020 Science

Cell types



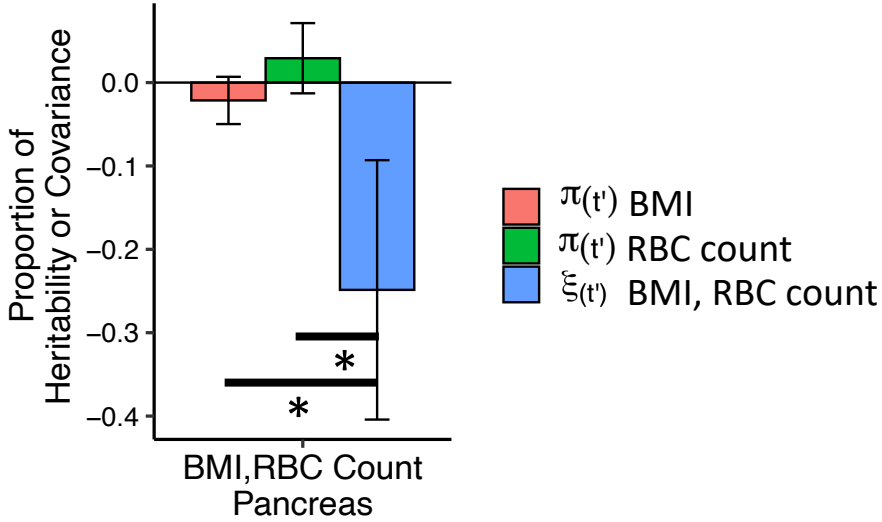
Ulirsch 2019 Nature Genetics

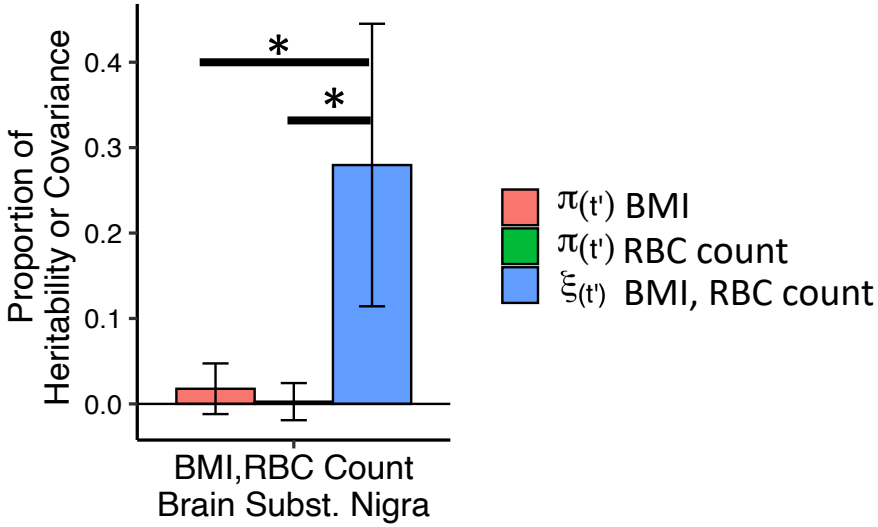
Populations of single cells

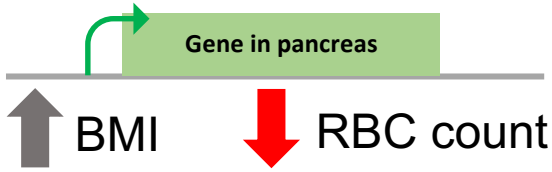


from Nathan 2022 Nature
Perez 2022 Science
Yazar 2022 Science

Cross-trait TCSC identifies tissue-disease covariance pairs







Where $\pi(t')(resp.\ \xi(t'))$ = proportion of disease heritability (resp. covariance) explained by predicted expression in tissue t'

Amariuta 2022 bioRxiv

Conclusions

1. Co-regulation scores and TWAS statistics can be used to infer the causal tissue(s) underlying disease heritability (covariance).

2. TCSC identifies new, biologically plausible tissue-disease pairs including the aorta artery and glaucoma.

3. TCSC may be more informative when applied to dynamic eQTL datasets, case/control eQTL, and single cell gene expression datasets.

Acknowledgments

Price Lab
Alkes Price
Katherine Siewert-Rocks
Huwenbo Shi
Martin Zhang

Please check out our manuscript:
Amariuta 2022 bioRxiv

I'm looking for graduate students and postdocs interested in statistical and population genetics to join my group.

website: amariutalab.org contact: tamariuta@gmail.com