Leveraging GTEx Data to Dissect Existing GWAS

Hae Kyung Im, PhD

GWAS subgroup call
November 12, 2014

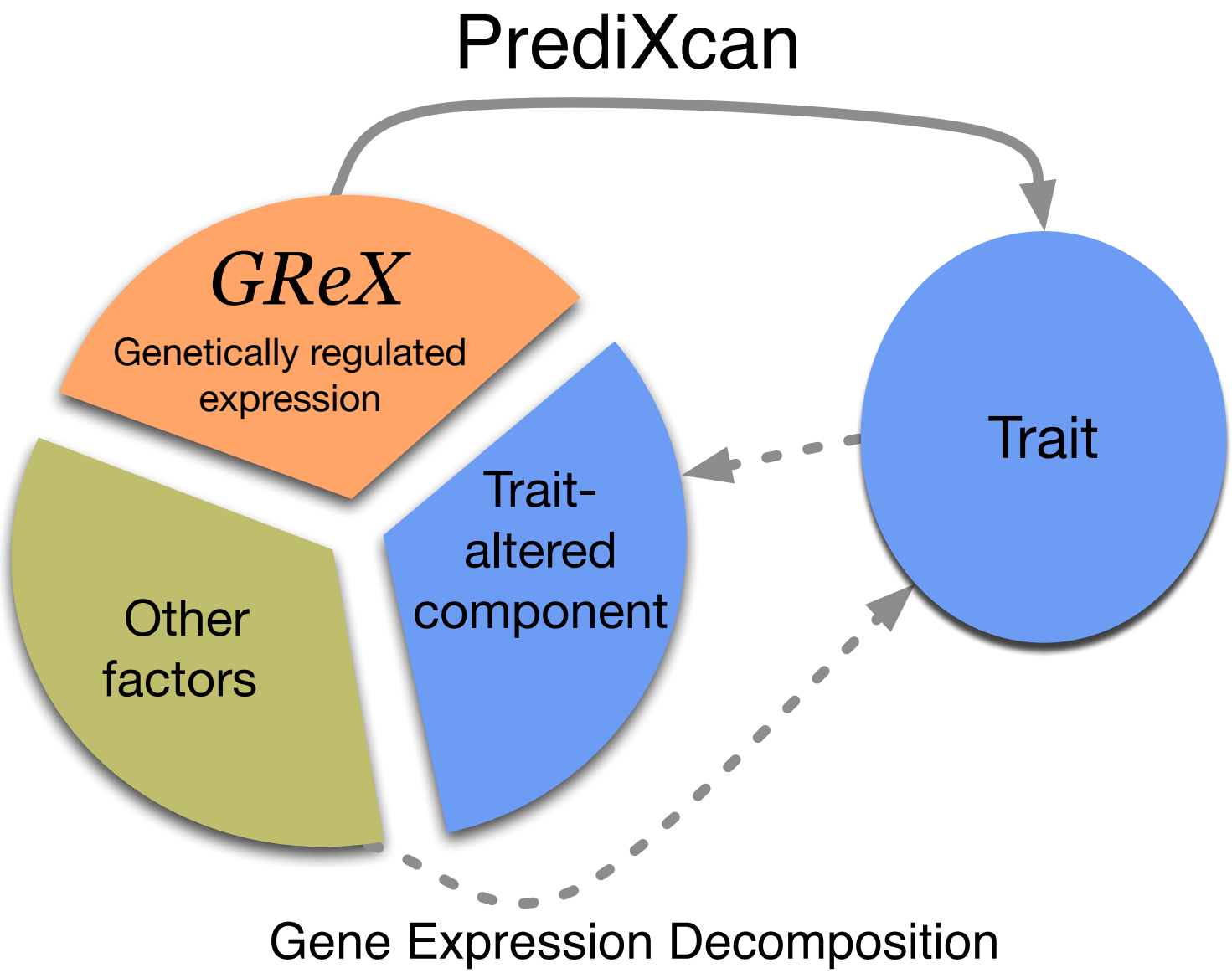
THE UNIVERSITY OF CHICAGO

Leveraging GTEx and Existing GWAS

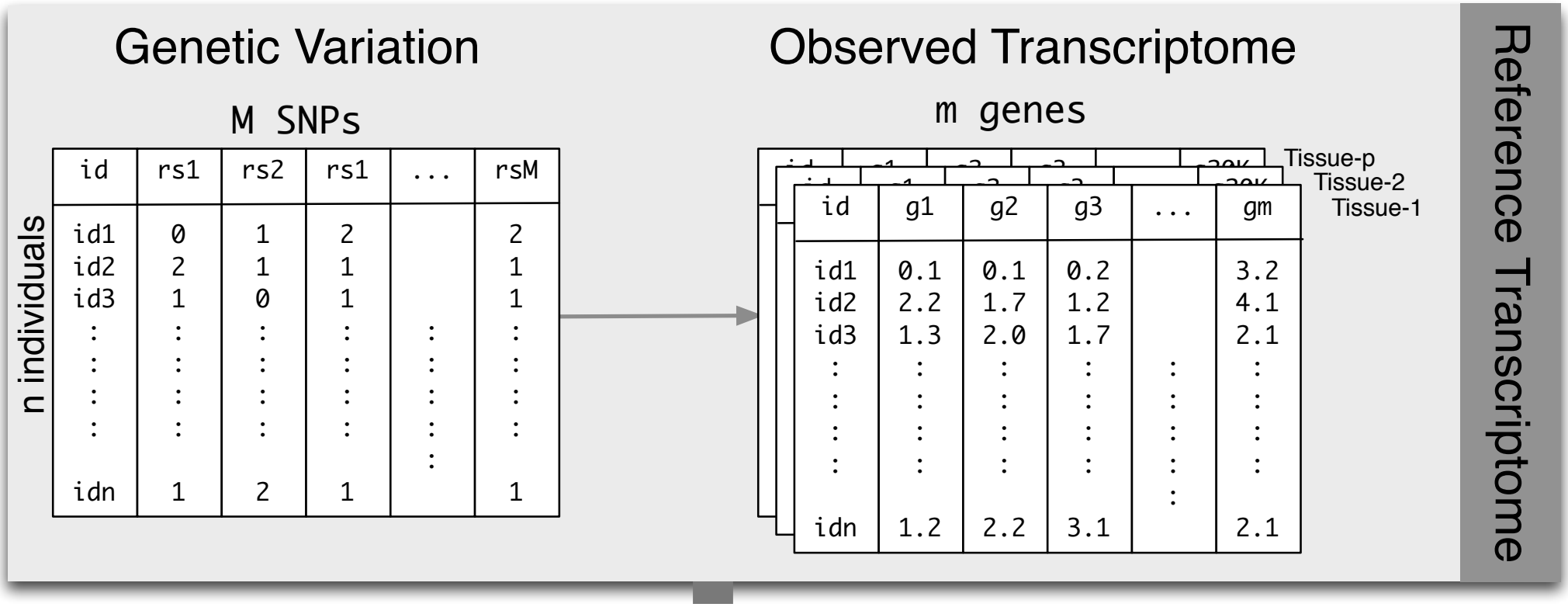
- PrediXcan
  - Gene-based test that tests tissue-specific mediating effects of gene expression traits on phenotypes
  - Imputation quality improvement with option 1 data
  - Application to drug-induced peripheral neuropathy data

- Regulability
  - Heritability of the regulome
  - Application to WTCCC

Mechanisms Tested by PrediXcan

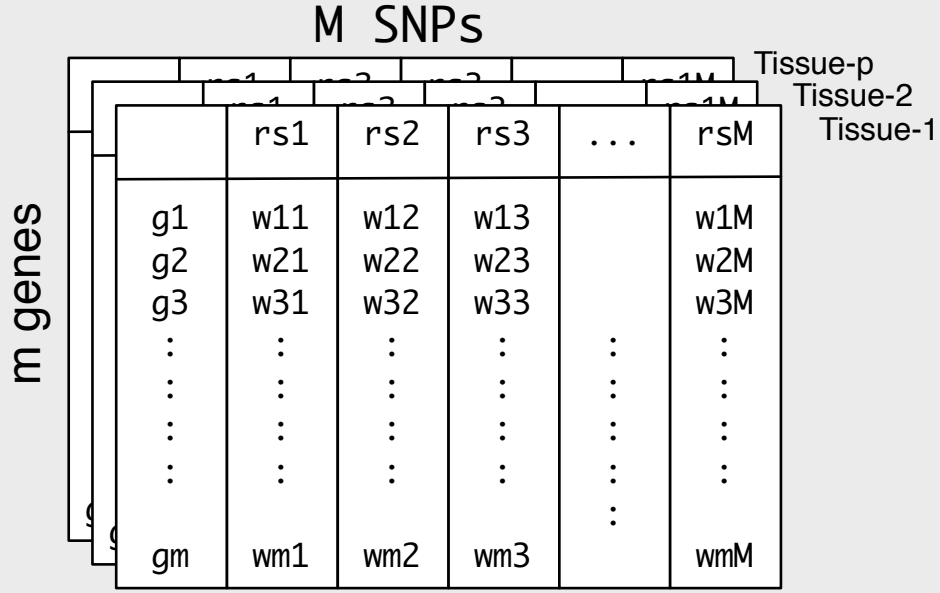


PrediXcan uses Reference Transcriptome



PredictDB: Public Database of Weights for GReX

PredictDB: Database of Prediction Models

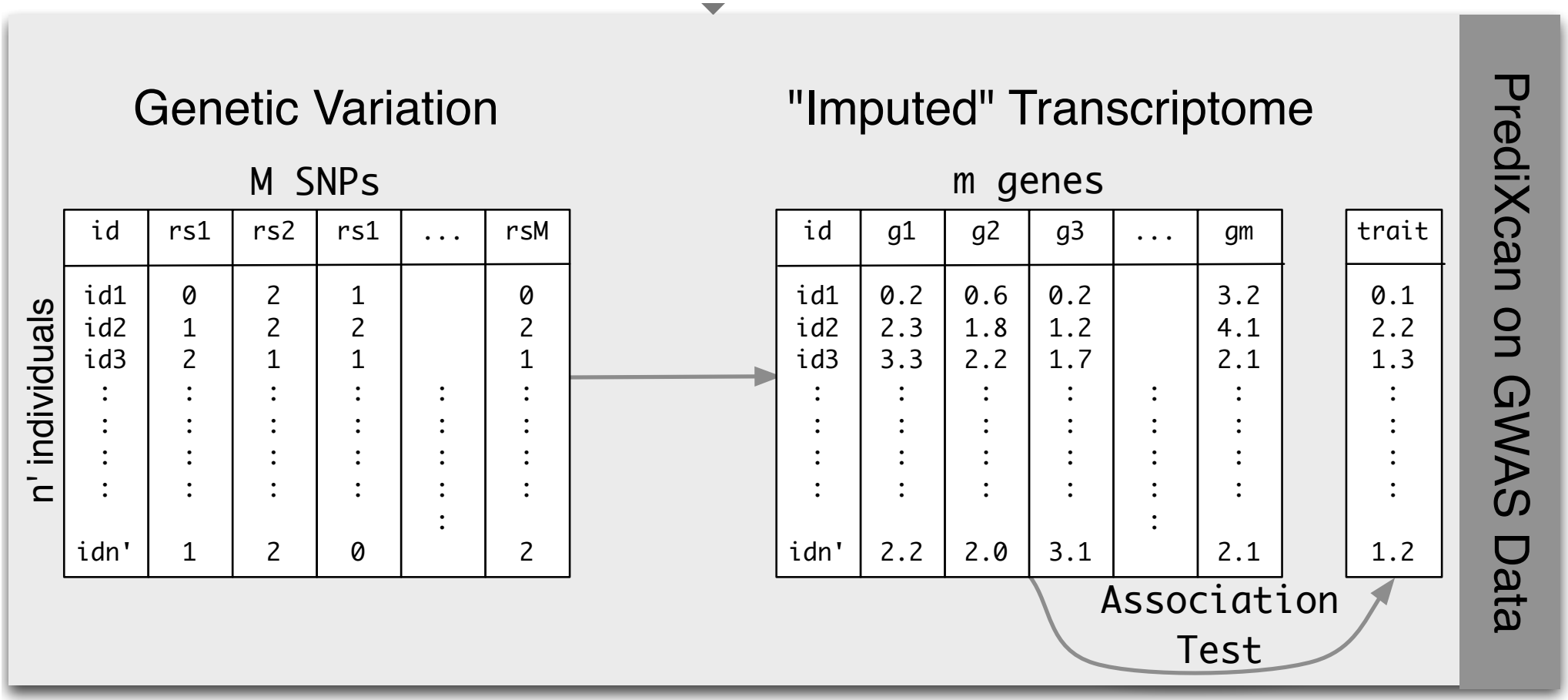


Additive model of gene expression trait trained in reference transcriptome datasets

$$T = \underbrace{\sum_k w_k X_k}_{GReX} + \epsilon$$

Weights stored in PredictDB

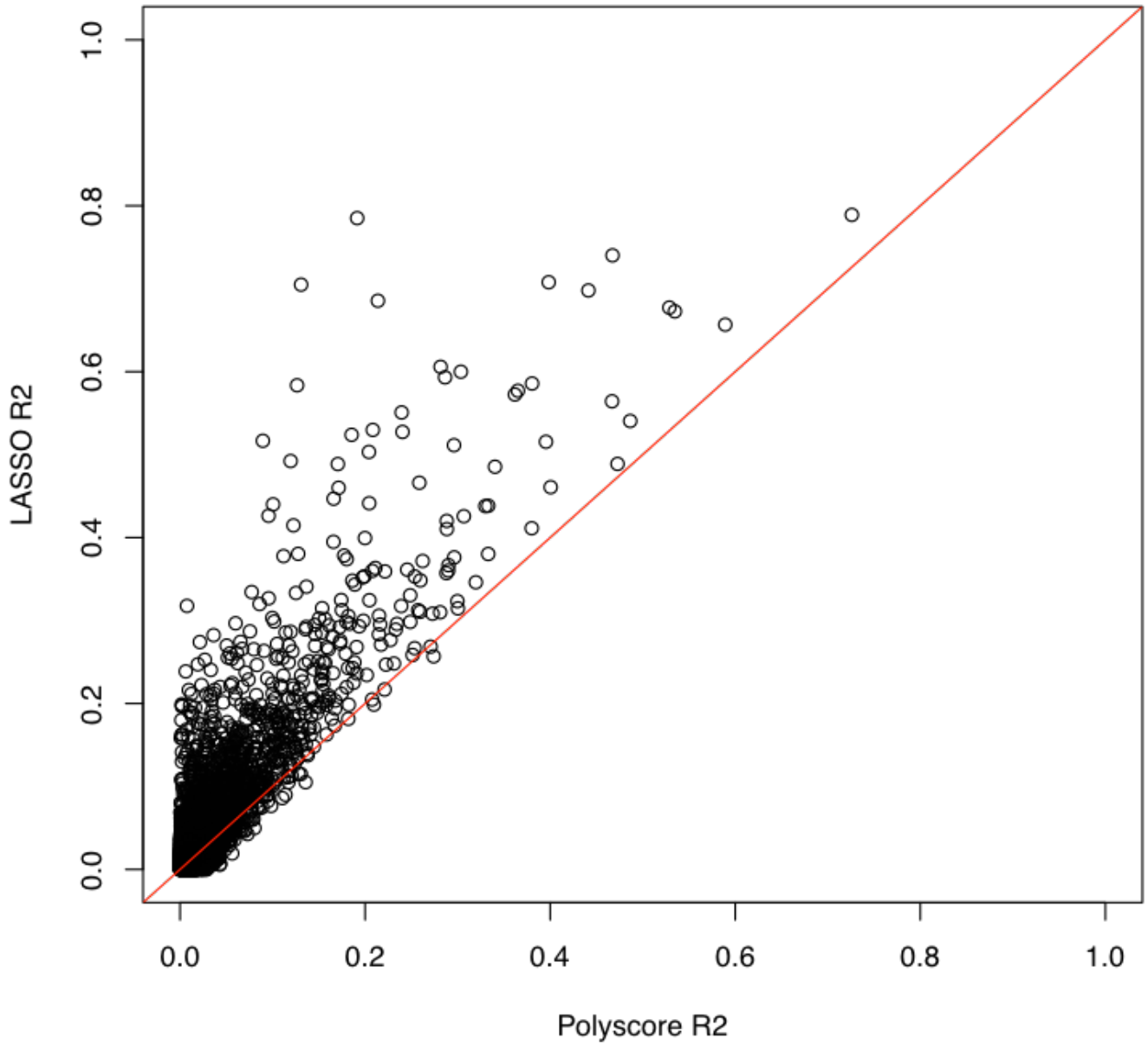
PrediXcan Imputes Transcriptome & Tests Assoc.



PrediXcan: Mechanism-driven Gene-Based Test

- Directly tests the molecular mechanism through which genetic variants affect phenotype
- Genes more attractive than genetic variants
  - A lot is know about their function
  - Follow up experiments can be easily devised
  - Reduced multiple testing burden
- Direction of effects
  - Positive effects: down regulation is therapeutic option
  - Negative effects: more likely to harbor deleterious rare variants
- No reverse causality issues
- Can be systematically applied to existing GWAS studies
- Tissue-specificity can be inferred

Simple Polygenic Score vs. LASSO

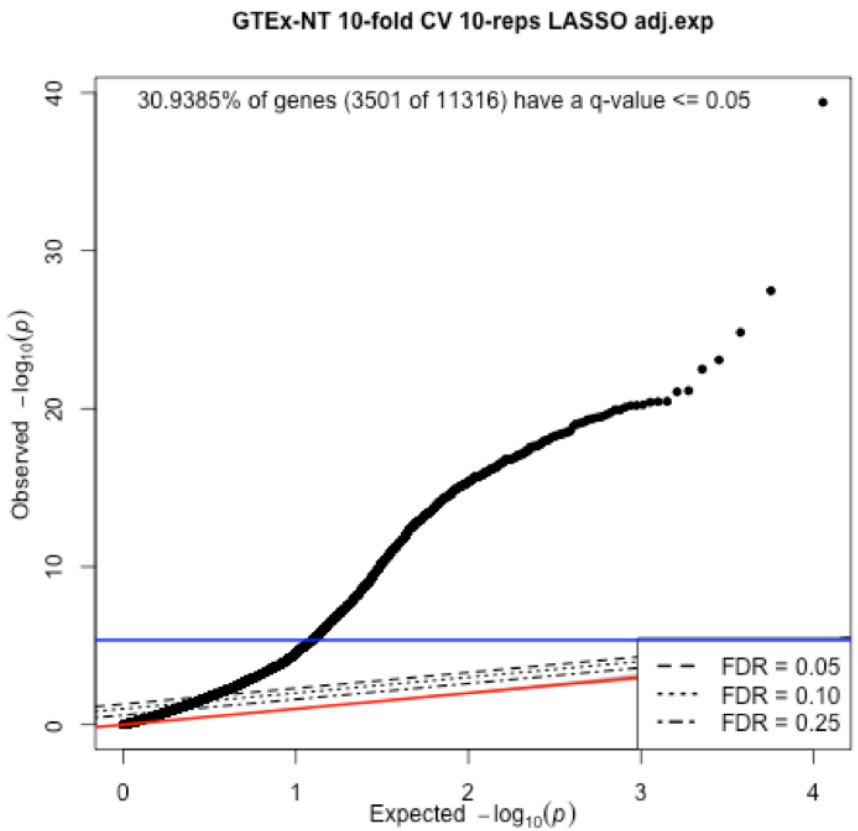


GEUVADIS

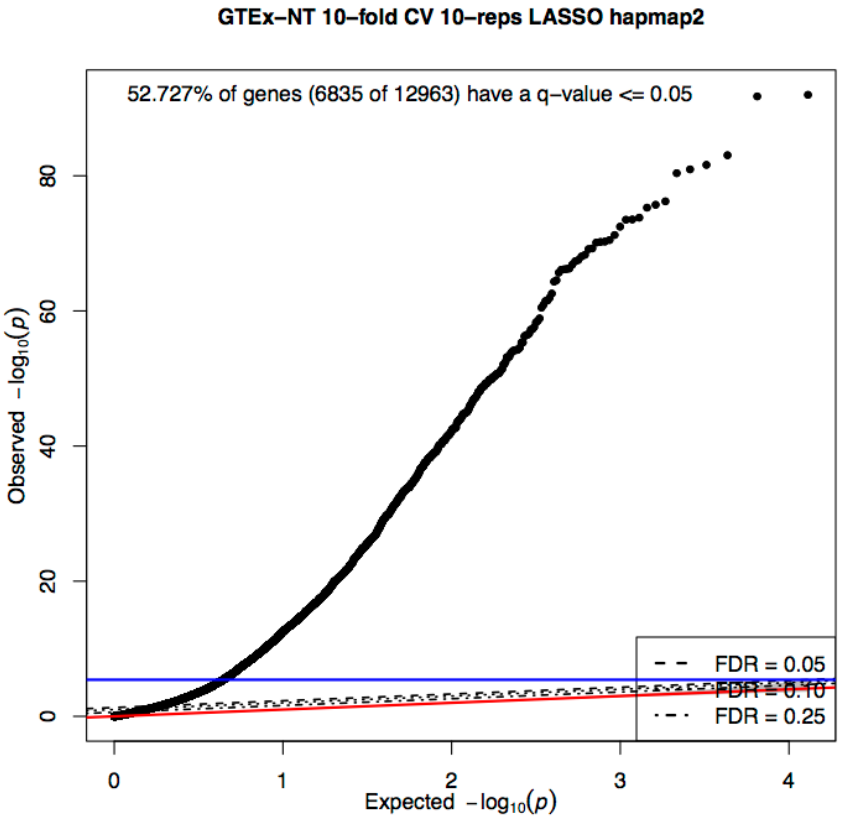
Heather Wheeler

Imputation QualityTibial Nerve

N=97
all (3.5M) SNPs
(pilot data)



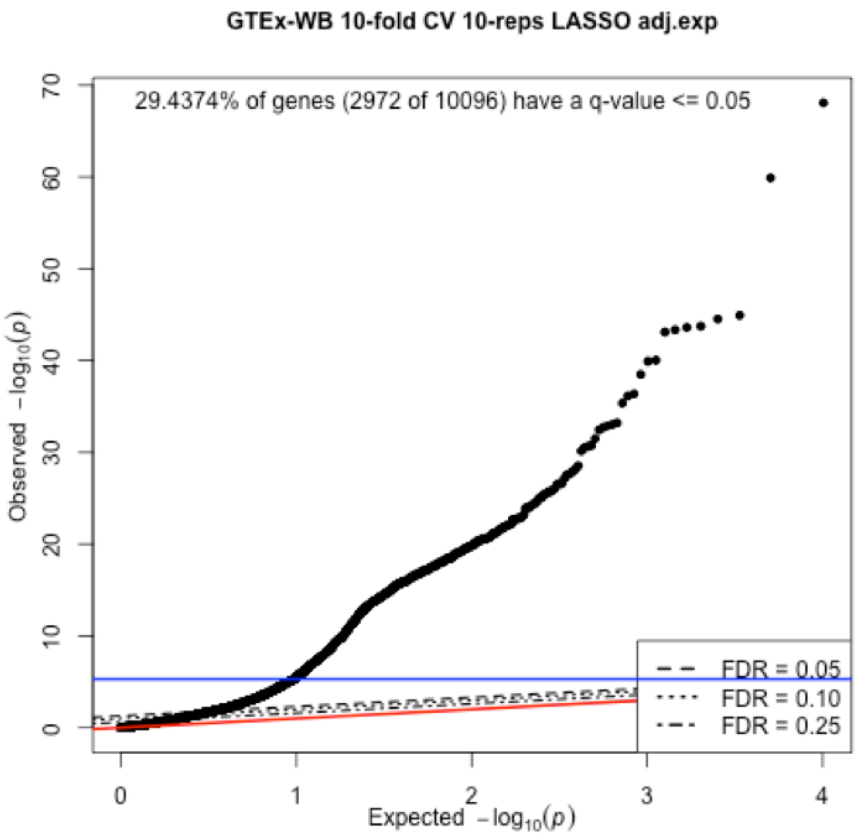
N=256
HapMap2 unamb. 2.6M SNPs
(2014-06-13 data)



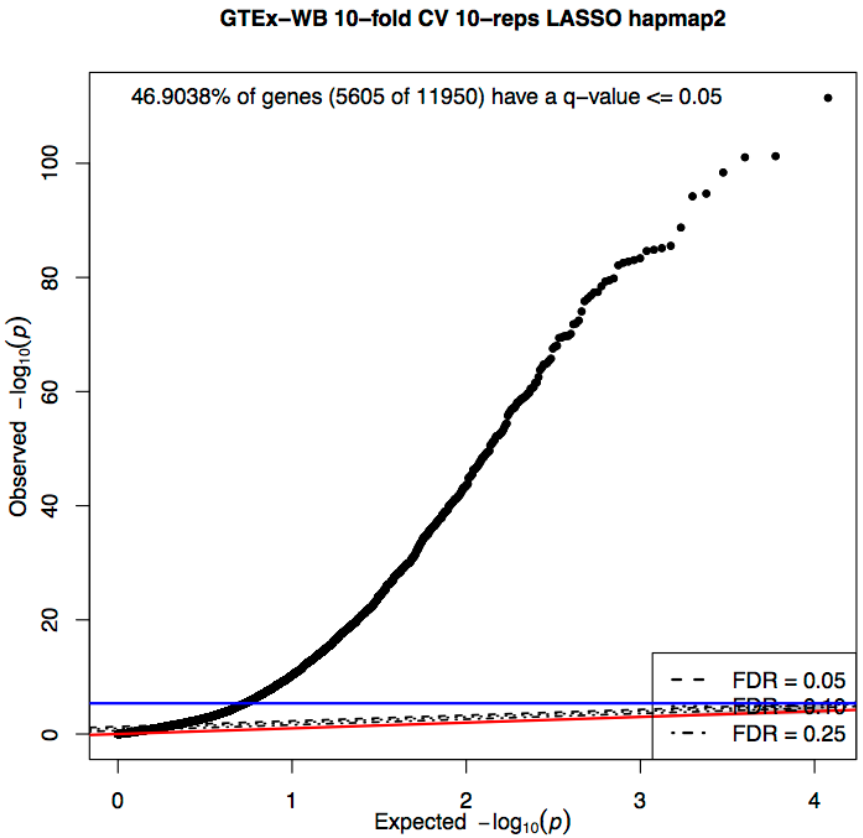
Heather Wheeler

Imputation Quality Whole Blood

N=169
all (3.5M) SNPs
(pilot data)



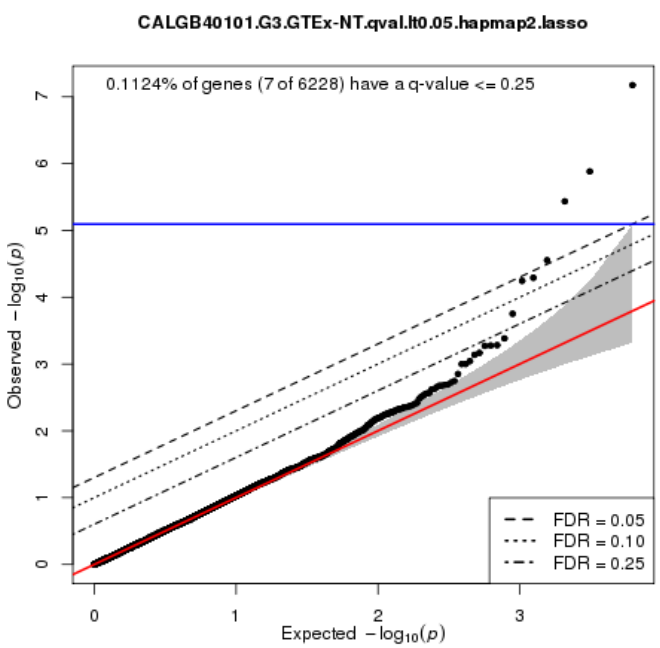
N=339
HapMap2 unamb. 2.6M SNPs
(2014-06-13 data)



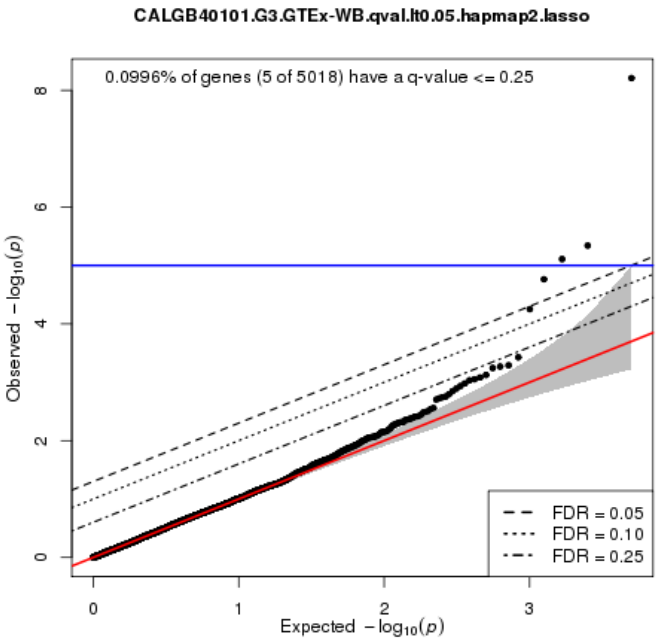
Heather Wheeler

Drug Induced Peripheral Neuropathy

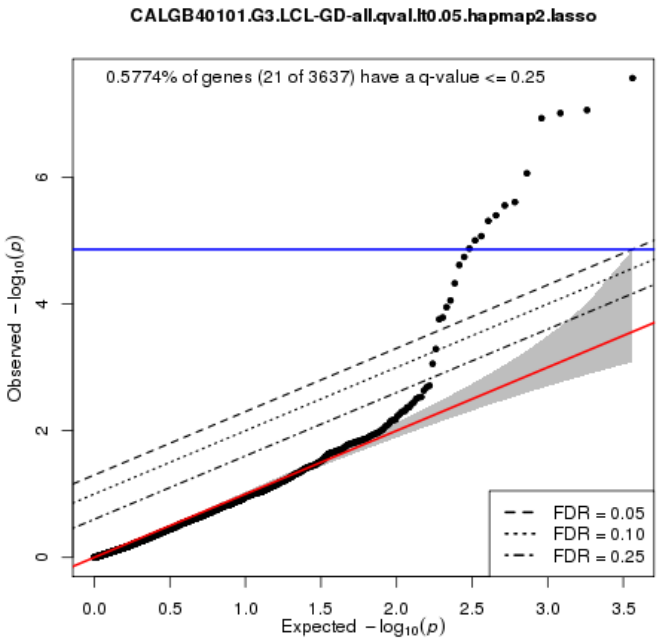
Tibial Nerve



Whole Blood



LCLs



Heather Wheeler

Chip Heritability & Regulability

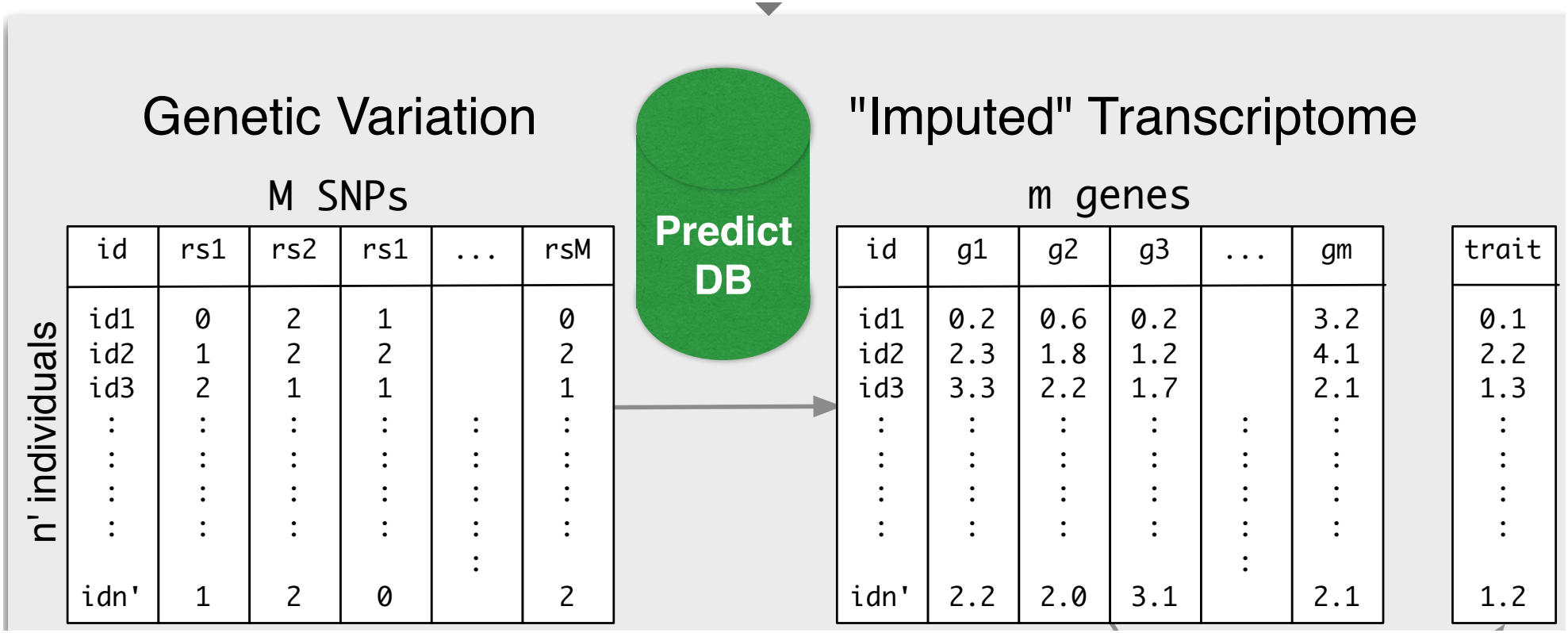
- Chip heritability: proportion of trait variability explained by genotyped variants

$$Y_i = \sum_k c_k X_{ik} + \epsilon_i$$

- Regulability: proportion of trait variability explained by the genetically predicted transcriptome —regulome

$$Y_i = \sum_g \gamma_g T_{ig} + \eta_i$$

Regulability Calculation



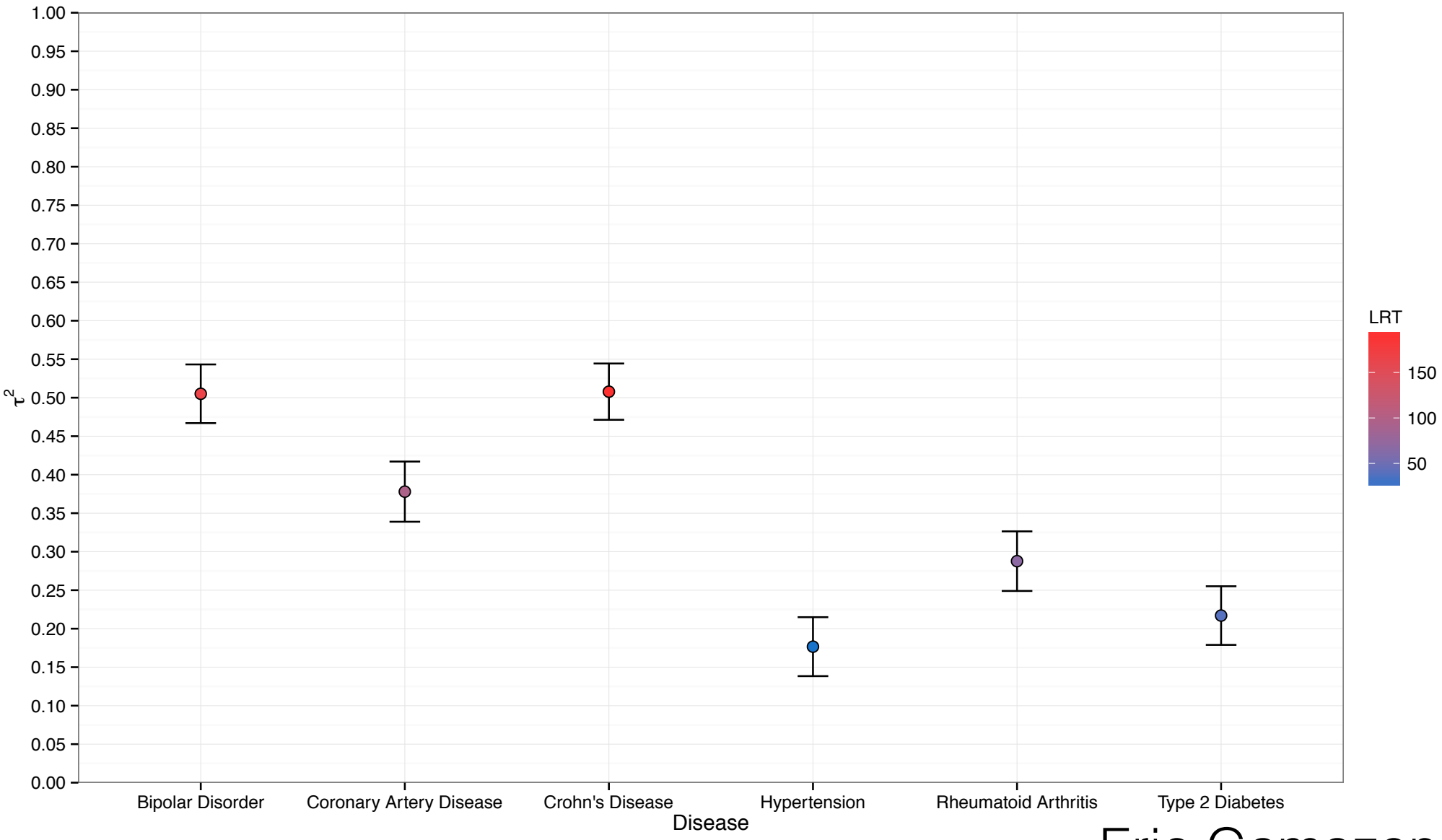
id	rs1	rs2	rs1	...	rsM
id1	0	2	1		0
id2	1	2	2		2
id3	2	1	1		1
:	:	:	:	:	:
:	:	:	:	:	:
:	:	:	:	:	:
:	:	:	:	:	:
idn'	1	2	0		2

id	g1	g2	g3	...	gm	trait
id1	0.2	0.6	0.2		3.2	0.1
id2	2.3	1.8	1.2		4.1	2.2
id3	3.3	2.2	1.7		2.1	1.3
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
idn'	2.2	2.0	3.1		2.1	1.2

- “Impute” transcriptome
- Compute covariance matrix of imputed transcriptome
- Compute REML estimate of variance

$$Y_i = \sum_g \gamma_g T_{ig} + \eta_i$$

WTCCC - Regulability Estimates using Whole Blood



Eric Gamazon

Acknowledgements

Thank you!

Contributors

- **Eric R. Gamazon**
- **Heather E. Wheeler**
- Nancy J. Cox
- Dan Nicolae
- Kaanan P. Shah
- Sahar Mozaffari
- M. Eileen Dolan

Data sources

- CALGB 90401
- WTCCC
- GEUVADIS

Funding

- HKI was funded in part by Uchicago CTSA NCI K12CA139160
- University of Chicago Diabetes Research and Training Center: P60 DK20595, P30 DK020595
- Genotype of Tissue Expression GTEx R01 MH090937 and R01 MH101820
- Pharmacogenomics of Anticancer Agents PAAR UO1GM61393
- Pharmacogenomics Research Network (PGRN) Statistical Analysis Resource (P-STAR) U19 HL065962
- Conte Center grant P50MH094267