

# Genetic Variation

# Observed Transcriptome

Reference Transcriptome

M SNPs

	id	rs1	rs2	rs1	...	rsM
n individuals	id1	0	1	2		2
	id2	2	1	1		1
	id3	1	0	1		1
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	idn	1	2	1		1

m genes

	id	g1	g2	g3	...	gm	Tissue-p	Tissue-2	Tissue-1
	id1	0.1	0.1	0.2		3.2			
	id2	2.2	1.7	1.2		4.1			
	id3	1.3	2.0	1.7		2.1			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	idn	1.2	2.2	3.1		2.1			

PredictDB: Database of Prediction Models

M SNPs

		rs1	rs2	rs3	...	rsM	Tissue-p	Tissue-2	Tissue-1
m genes	g1	w11	w12	w13		w1M			
	g2	w21	w22	w23		w2M			
	g3	w31	w32	w33		w3M			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	⋮	⋮	⋮	⋮	⋮	⋮			
	gm	wm1	wm2	wm3		wmM			

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

# Genetic Variation

# "Imputed" Transcriptome

PrediXcan on GWAS Data

M SNPs

	id	rs1	rs2	rs1	...	rsM
n' individuals	id1	0	2	1		0
	id2	1	2	2		2
	id3	2	1	1		1
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮
	idn'	1	2	0		2

m genes

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

Association Test