Note taking from paper

GWAS – a research approach identify genomic variant → identify genomic variants

↳ GWAS => statistical based?

↳ statistically linked with a risk of disease or a trait.

PredixCan => gene based?

↳ estimated components of gene expression → determine by invidual genetic profile.

↳ transcrition

↳ translation

↳ correlate the "imputed" gene expression with the phenotype* under investigation

phenotype = trait = disease*

input = assign

involed = related = relevant

etiology = reason

PredixCan → identify genes that relevant with the reson of the trait

Whole genome tissue-model

↳ train with reference transcripton data

↳ predict gene expression

→ is this low gene expression prediction work before PredixCan?

PredixCan => more understand about the biology mechanism??

=> What is the benefit of it? → less testing burden.

↳ approach to the design of follow up experience

PredixCan → detect genes related with the disease

↳ provide insight into the mechanism of the association.

GWAS → successful in identifying the susceptibility loci for complete disease

↓

What GWAS have done so far

→ conduct single variant tests → test on the variant of a single gene.

→ intergorate genome in an agnostic way?

→ make the sample size larger → better inteference.

Less successful in improve on simple approaches → GWAS??

Genome-wide significant → in stats => mahathan plot?

modest = low

→ low propotion of variance in dieseas risk.

GWAS of disease susceptibility → relevant gene affect is small

pharmacologic trait → result in lim conclusive finding on gene factor contribute to complete trait

Functional significant of discovered loci => unclear

GWAS still identify susceptibility loci

→ how to understand the biology mechanism

Gene based approach + multi marke → replace single variant test → not significant def.

↳ conduct test of assosiation on biologycally informed of SNPs

↳ reduce testing penalty

The collab of omic data → extend GWAS by addressing function gap

trait / disease under significant genetic regulatio

Gene-based study ⇒ lacking

↳ improve knowledge.

⇒ suggest PrediXCan

↳ detect trait based on gene data.

PrediXCan → identify diesease based on the mechanism of gene.

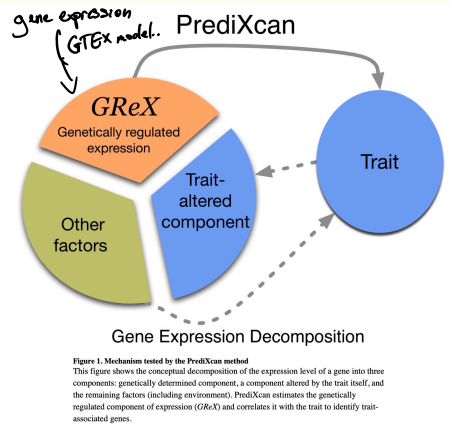


Figure 1. Mechanism tested by the PrediXcan method
This figure shows the conceptual decomposition of the expression level of a gene into three components: genetically determined component, a component altered by the trait iself, and the remaining factors (including environment). PrediXcan estimates the genetically regulated component of expression (*GReX*) and correlates it with the trait to identify trait-associated genes.

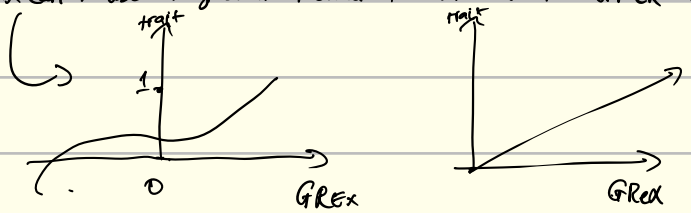
SNP = single Nucleotide Polymorphism → a type of genetic variation

Genotype impuatation work like ML.

Use prediction model to assign gene expression

↳ use PrediXCan to determine trait.

PrediXCan : use regression method to work with GReX data

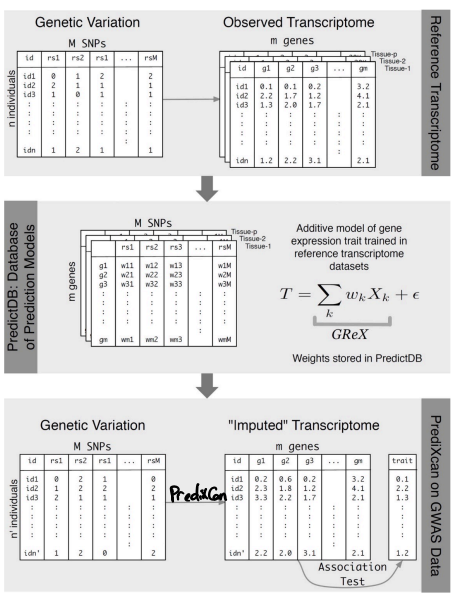


attenuation: decrease

Aware that the attenuation bias because error in prediction

↳ does not affect the analysis

Do more about GReX



SNPs → gen expression → GReX

↓

a ML model

Additive model: build a data base of prediction model

⇒ unclear

PrediXCan contribute to psychiatric + neurodevelopmental disorder

↳ identify gene target for therapy

↳ gene based, better understand the mechanism → direct insight

Pros: less test burden

more straight forward for information prior + grouping of functional unit

No transcriptome data needed? ⇒ less complex

the result is simplified

Multiple tissue can be evaluate

Can be mass apply

The prediction model was build in the DGN whole blood cohort using LASSO, elastic net ($\alpha = 0.5$), variant p-value

↳ a regression model

↓

the risk of developing trait

Asses: Judge lasso

↳ 10-fold cross-validation → use for large data set

↳

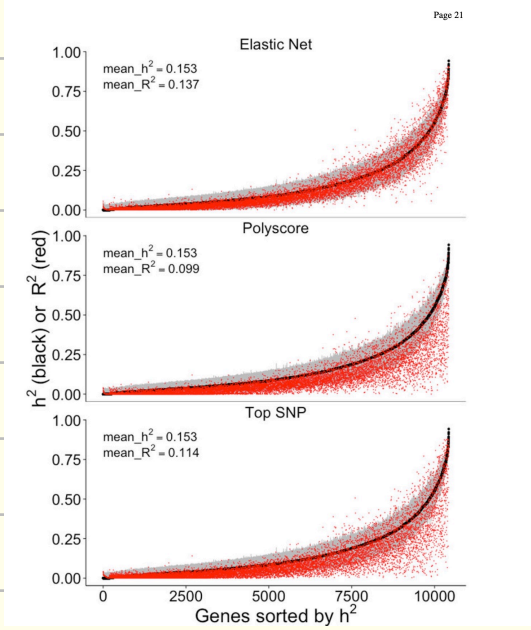
test	—	—	—
—	test	—	—
—	—	test	—
—	—	—	test

Estimate the heritability of gene expression in DGN attributable to SNPs using a mixed-effect model:

↳ only using local SNPs

↳ benchmark for prediction R^2

↳ compare the benchmark with R^2

↳ .153 & .137 ⇒ close

Benchmark = Standard