Coordinated Epistasis Reveals Pathway Architecture
(and method considerations)

Jolien Rietkerk

Helmholtz Pioneer Campus, Cai lab
WCPG 2023, Montréal, Canada

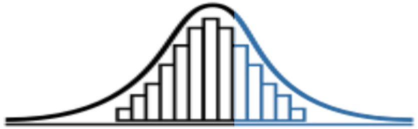
Disclosure

Nothing to disclose

Pathways leading to diseases

Unknown diagnosis Y

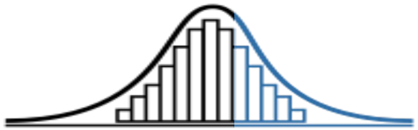
Data



N = 100K	Status
40K	Controls
70K	Cases

We think Y is actually Alzheimers + Type 2 Diabetes

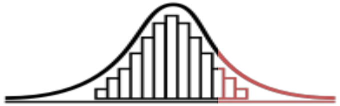
Data



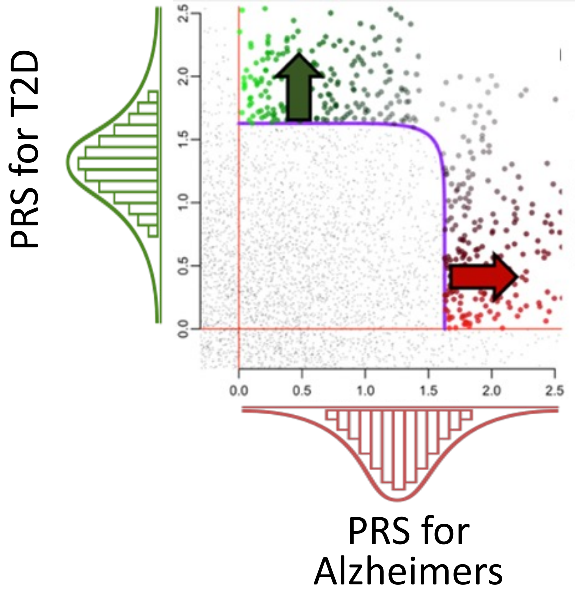
N = 100K	Status
40K	Controls
40K	Alzheimers
30K	Type 2 Diabetes

Alzheimers

Type 2 Diabetes



If we were right



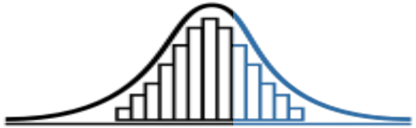
$$Y \sim \alpha_i PRS_i + \alpha_j PRS_j + \boldsymbol{\gamma_{i,j}}\, PRS_i{}^* \, PRS_j$$

$$\boldsymbol{\gamma_{i,j}} < 0$$

Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

What if Y is a psych disorder

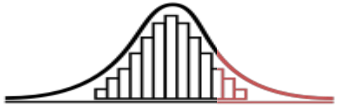
Major Depressive Disorder



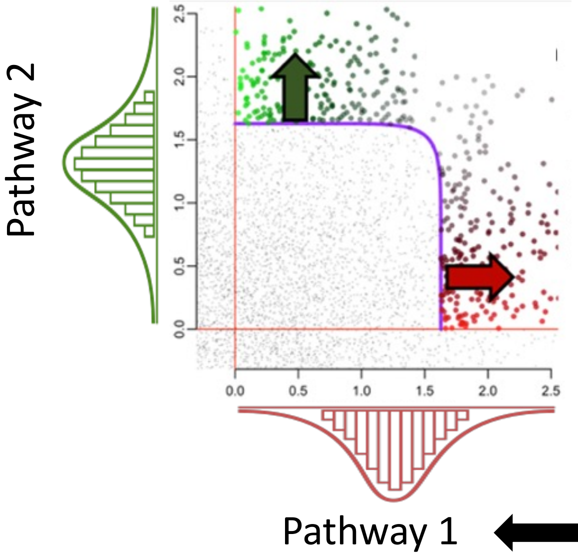
N = 100K	Status
40K	Controls
??K	?? Subtype A
??K	?? Subtype B

Melancholic

Early Onset



Hard to do this

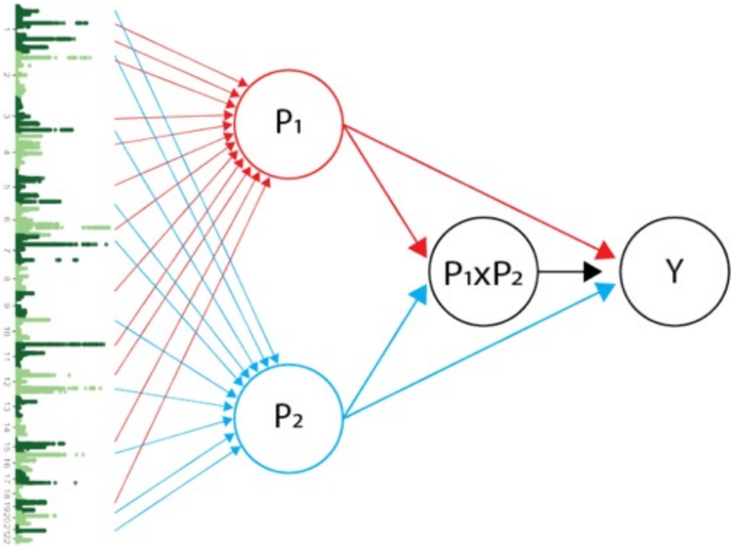


What should this even be?

Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

Coordinated Epistasis for assessing pathway architecture

Coordinated Epistasis (CE)

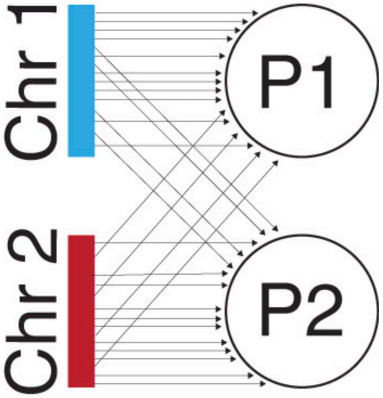


- $P_1 \approx$ PRS pathway 1
- $P_2 \approx$ PRS pathway 1

- Pathway specific variants are distributed all across the genome

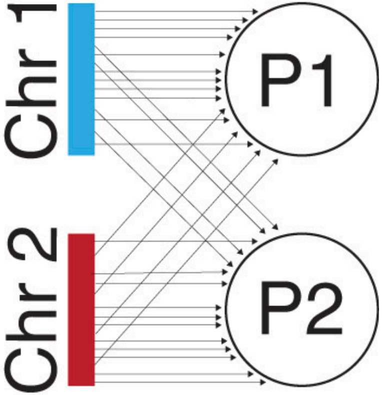
Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

Chromosome as partitions



Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

Chromosome as partitions



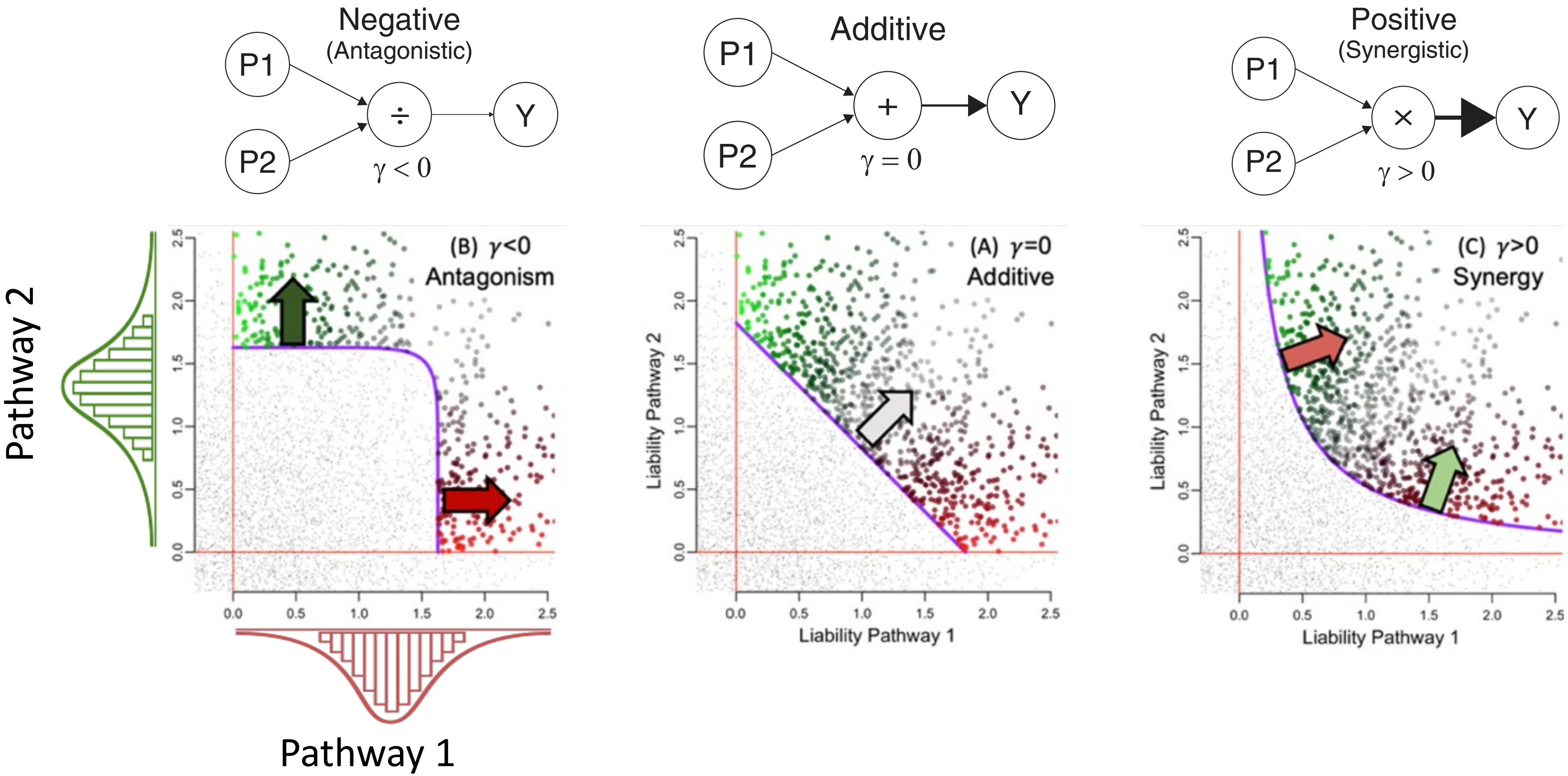
$y \sim \alpha_i PRS_i + \alpha_j PRS_j + \boldsymbol{\gamma_{i,j}}\, PRS_i{}^{*}\, PRS_j$

i,j ∈ $\{1..22\}$, i ≠ j

In total, 231 pairs of chromosomes

Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

Gamma indicates pathway architecture

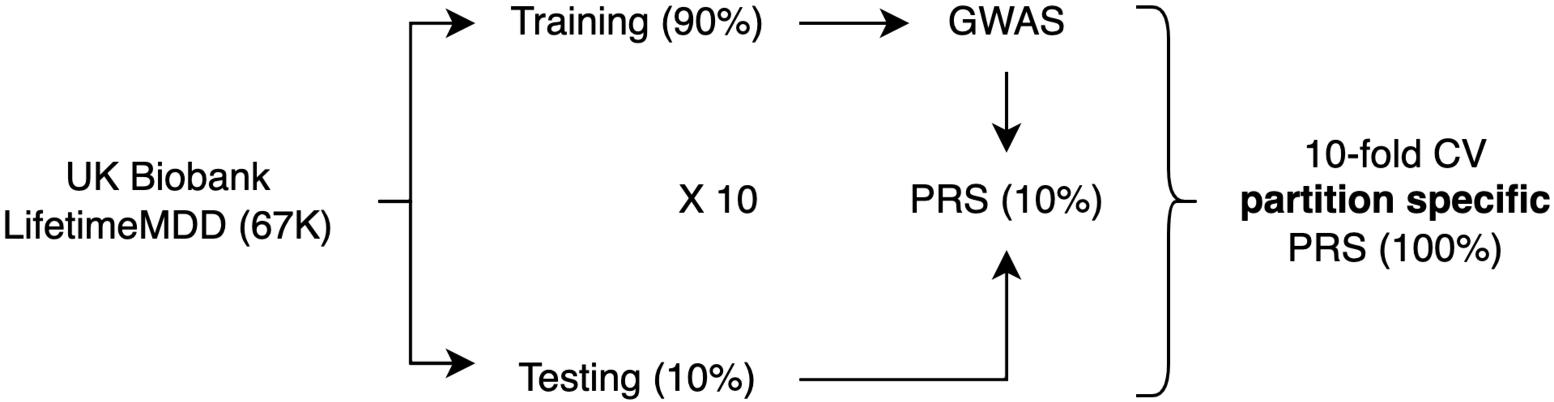


Tang D, et al. bioRxiv 2022, Sheppard, B. et al. PNAS 2022

Assess MDD pathway architecture in UKB

MDD	LifetimeMDD (67K)
Subtypes	6 clinical subtypes (N=25-65K)
Symptoms	14 symptoms CIDI-SF(N=50-100K)

10-fold Cross Validation (CV)



UK Biobank
LifetimeMDD (67K)

Training (90%)

GWAS

X 10

PRS (10%)

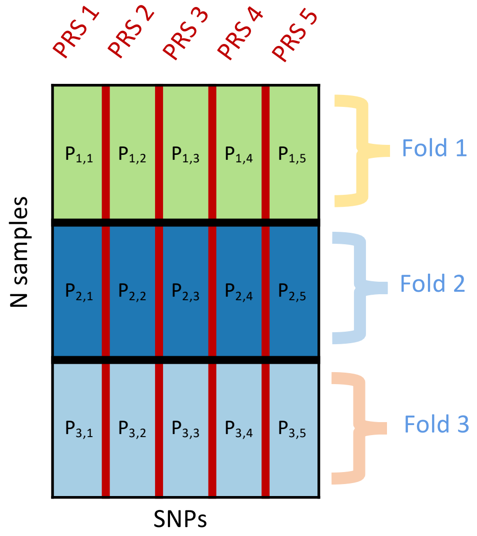
Testing (10%)

10-fold CV
partition specific
PRS (100%)

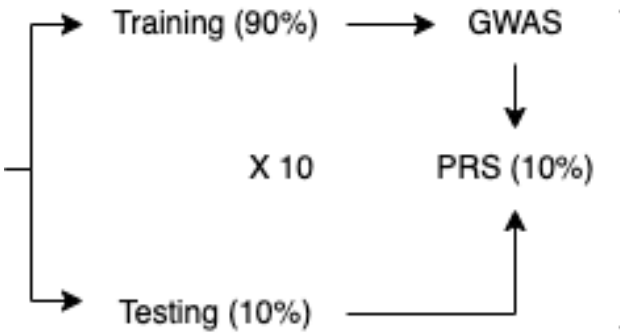
Caveats in using cross-validation PRS

CV creates batch effects

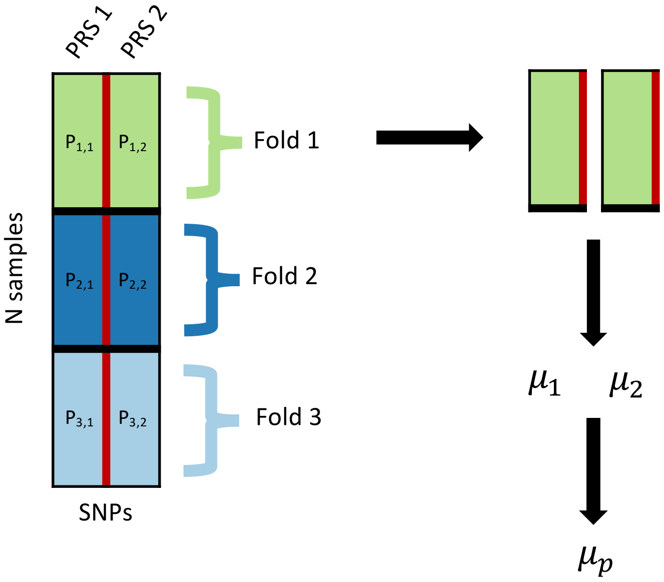
I. Clumping and thresholding



II. GWAS on varying individuals



Mundlak method corrects batch effect



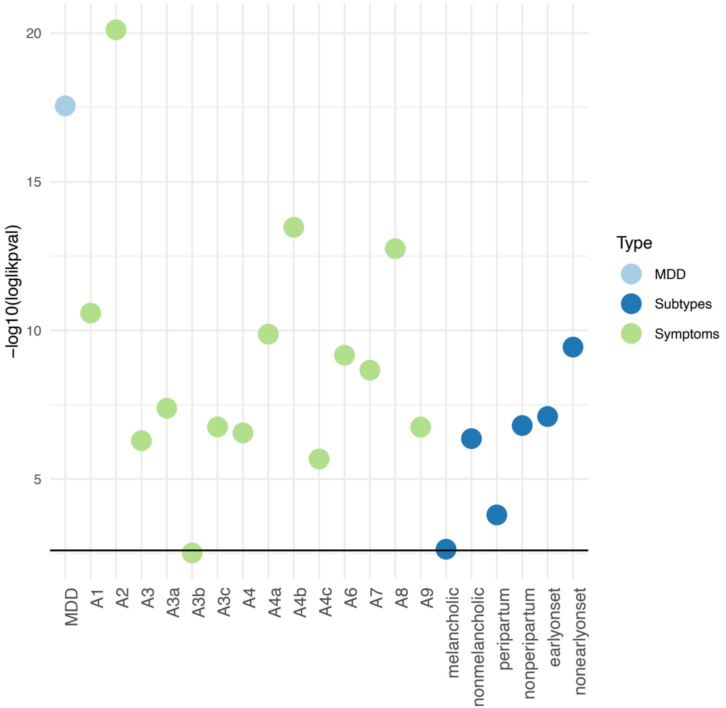
$$y \sim \alpha_1 PRS'_1 + \alpha_2 PRS'_2 + \mu_p + \gamma_{1,2} PRS'_1 * PRS'_2$$

μ_p= mean of means

PRS' = PRS - μ_p

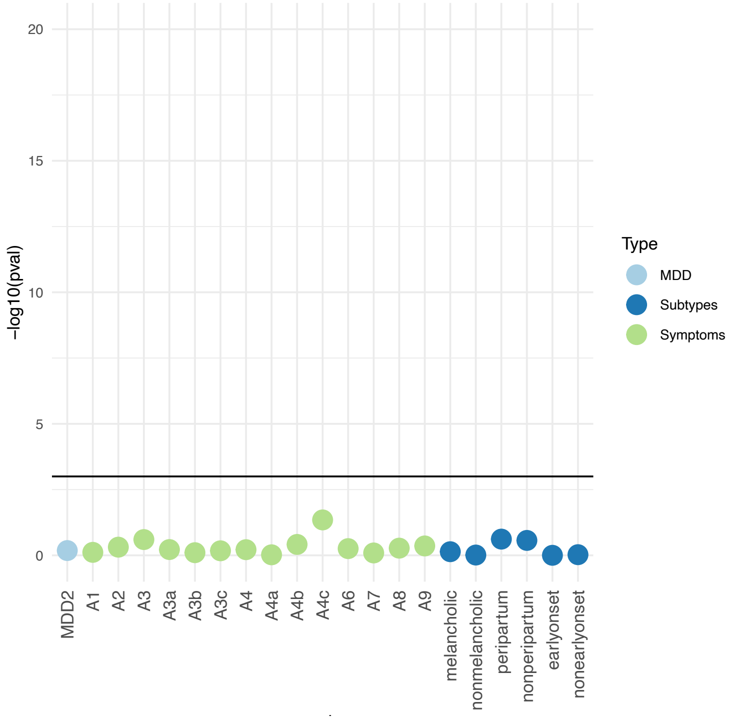
MDD analysis in UKB

MDD	LifetimeMDD (67K)
Subtypes	6 clinical subtypes (N=25-65K)
Symptoms	14 symptoms CIDI-SF(N=50-100K)



MDD analysis in UKB

MDD	LifetimeMDD (67K)
Subtypes	6 clinical subtypes (N=25-65K)
Symptoms	14 symptoms CIDI-SF(N=50-100K)



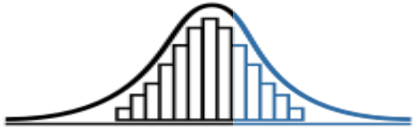
Wide-ranging implication for PRS analyses

- For all analyses using CV-based PRS, where PRS are used as “phenotypes” or ”covariates”
- Eg.
  - Multi-PRS predictions
  - PRS x SNP interactions
  - PRS x PRS interactions

Applying CE to **identify heterogeneity** in MDD phenotypes

Different definitions or meta-analyses of MDD have different levels of heterogeneity

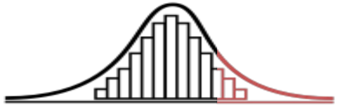
MDD



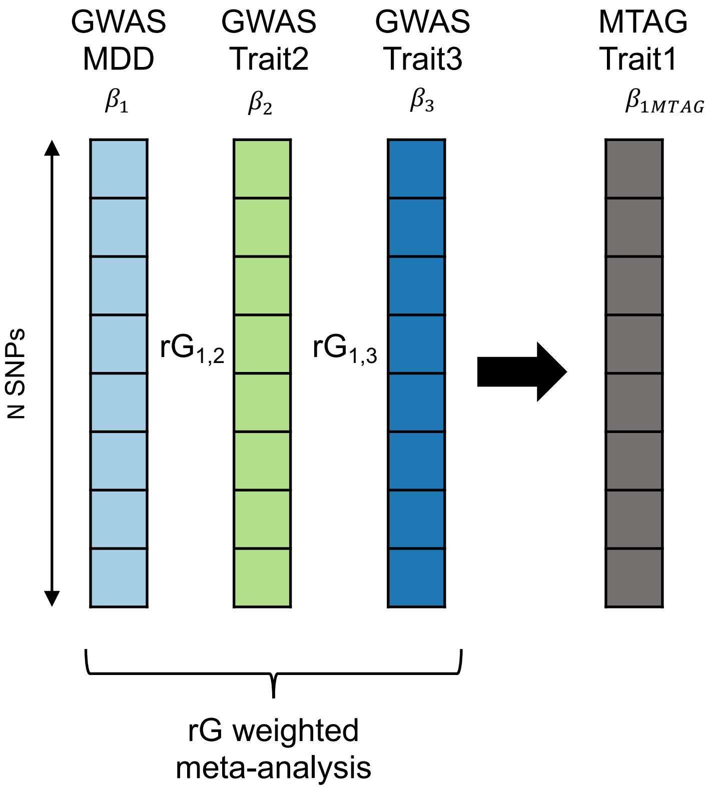
N = 100K	Status
40K	Controls
40K	Questionnaire
30K	Family history

Questionnaire

Family History

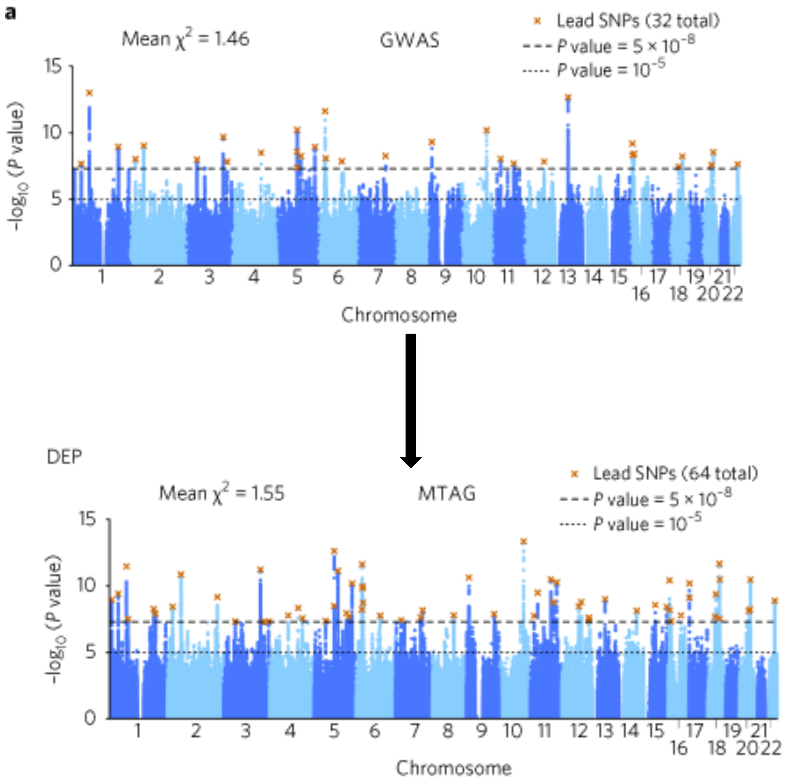


MTAG to improve effective sample size



Turley P, et al., Nat Genet 2018

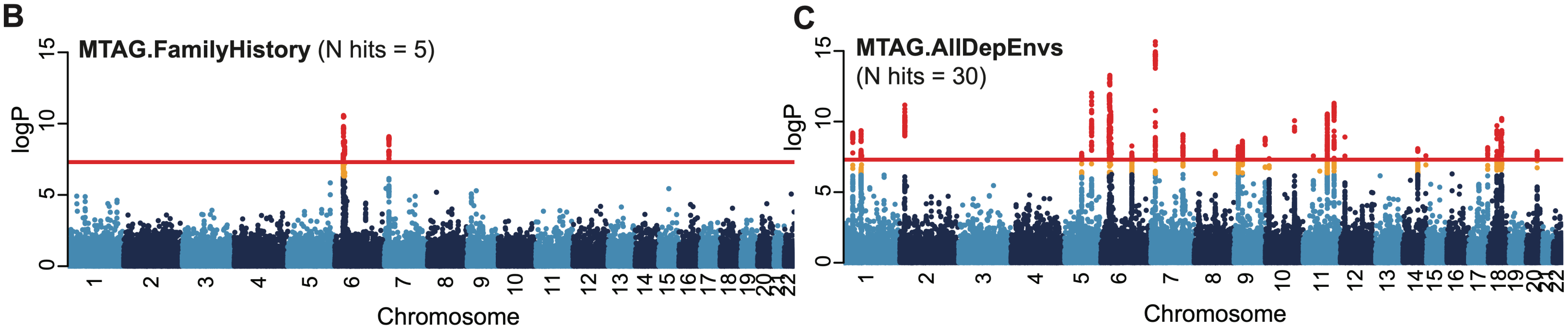
MTAG to improve effective sample size



Turley P, et al., Nat Genet 2018

MDD phenotypes from MTAG

Phenotype	Number of input GWAS	Description
MTAG.FamilyHistory	3	severe depression in family members
MTAG.ALLDep+Envs	9	Shallow MDD phenotypes and environment variables



Dahl A*, ... Cai N*. Nat Genet (2023 – in press)

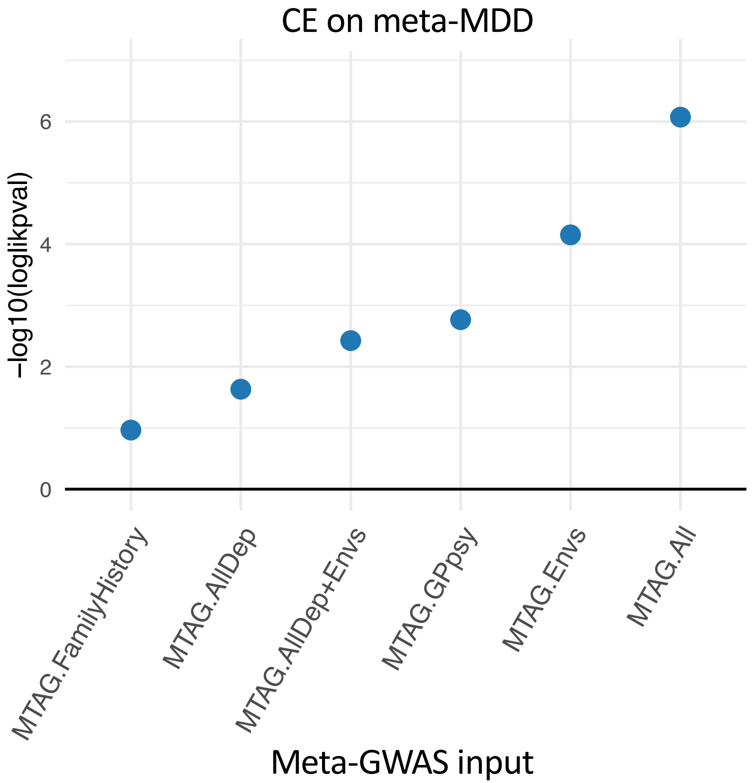
CE detects heterogeneity in pathways in MDD from MTAG

Phenotype	Number of input GWAS	Description
MTAG.FamilyHistory	3	severe depression in family members
MTAG.AllDep	5	Various shallow MDD phenotypes
MTAG.ALLDep+Envs	9	Shallow MDD phenotypes and environment variables
MTAG.Gppsy	1	Gppsy
MTAG.Envs	5	Related environment variables
MTAg.All	12	Shallow MDD phenotypes, enivornment variables, family history

Dahl A*, ... Cai N*. Nat Genet (2023 – in press)

CE detects heterogeneity in pathways in MDD from MTAG

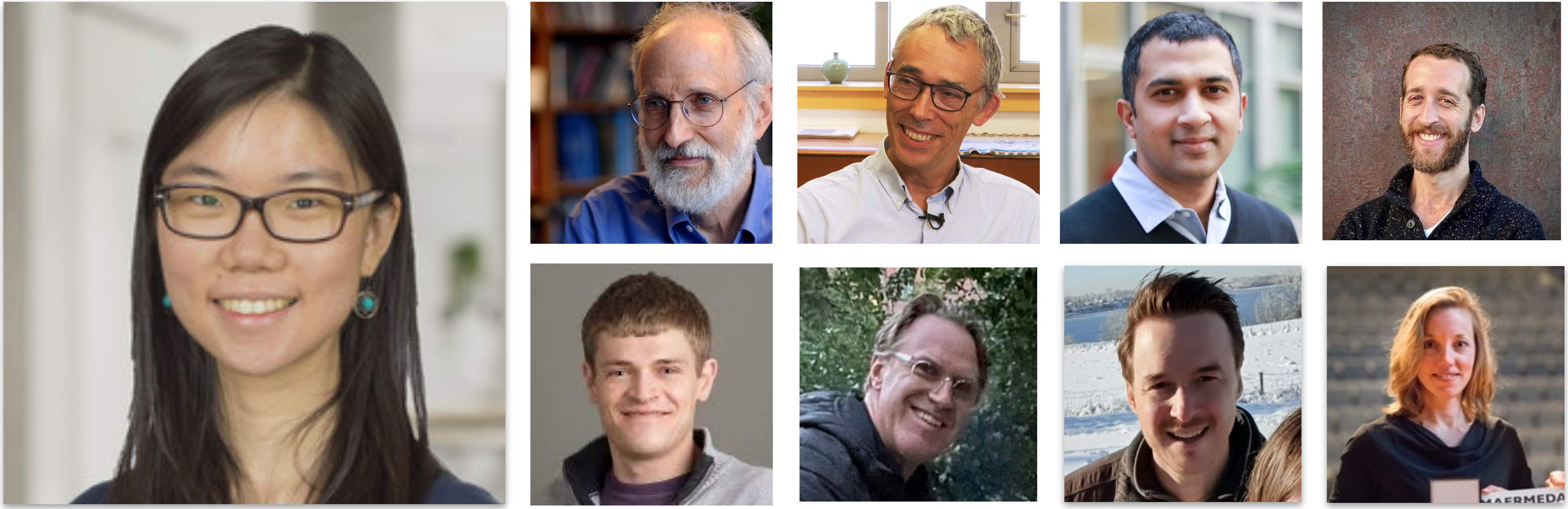
Phenotype	Number of input GWAS
MTAG.FamilyHistory	3
MTAG.AllDep	5
MTAG.ALLDep+Envs	9
MTAG.Gppsy	1
MTAG.Envs	5
MTAg.All	12



Summary

- CE detects pathway architecture
- 10x CV in PRS creates batch effect (wide-ranging implications!)
- Mundlak approach correct for batch effect in CV
- CE can detect significant pathway heterogeneity in heterogenous MDD meta-analyses

Acknowledgements



biobank uk

HELMHOLTZ MUNICH

TUM

X: @JolienRietkerk
Website: rietkerk-research.com