

Quantifying pleiotropy between complex traits

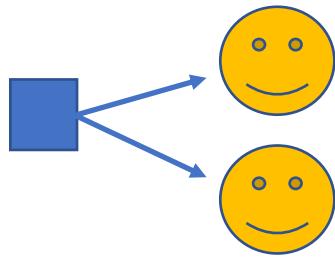
Luke J O'Connor

Price lab

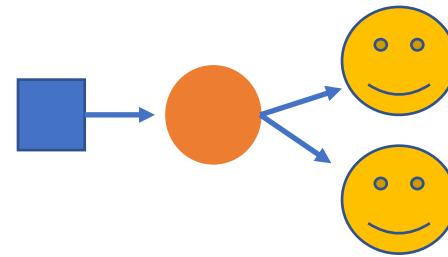
PGSG 3/8/19

Pleiotropy: genome-wide overlap between genetic associations for two traits

1. Independent mechanisms,
uncorrelated effects
("horizontal")



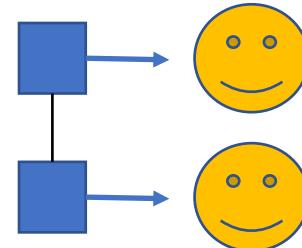
2. Shared
mechanism,
correlated effects



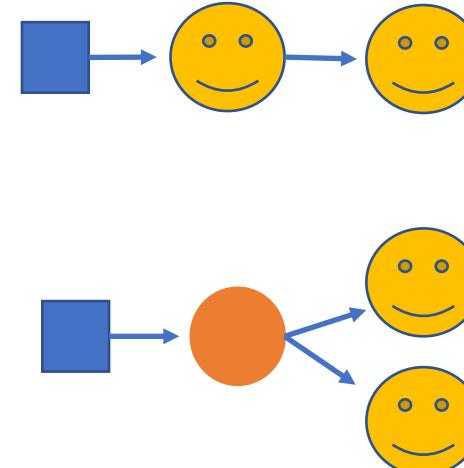
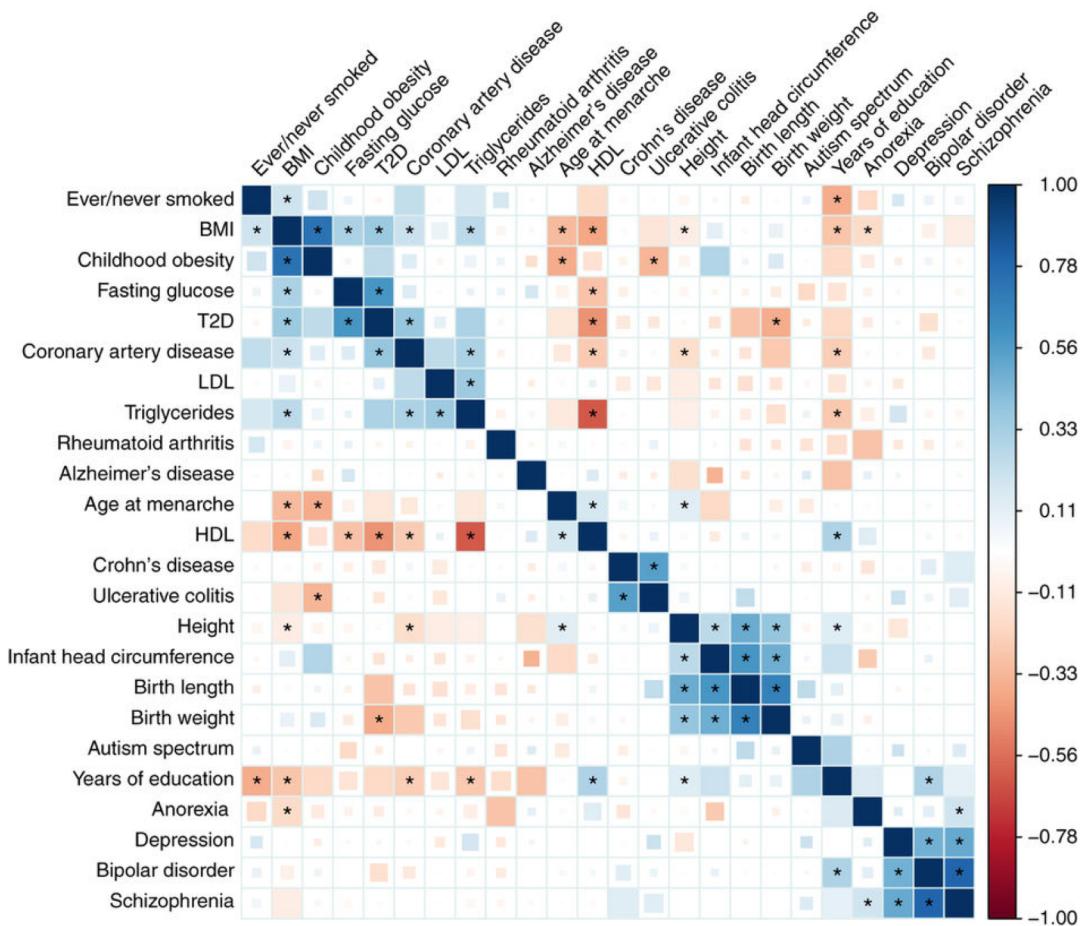
3. Causal relationship
("vertical")



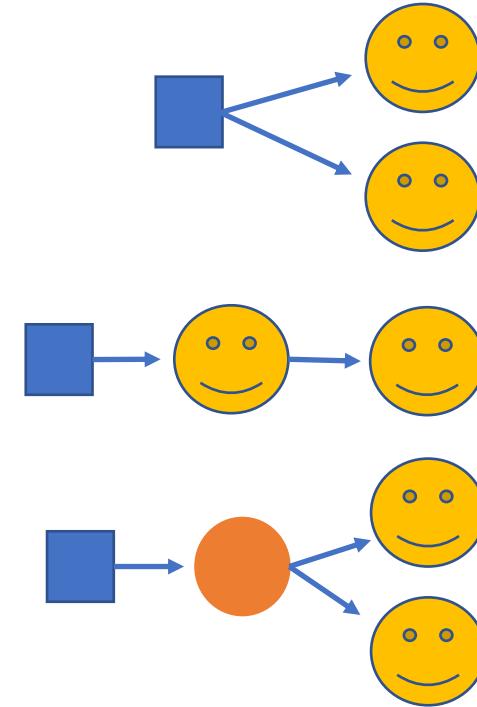
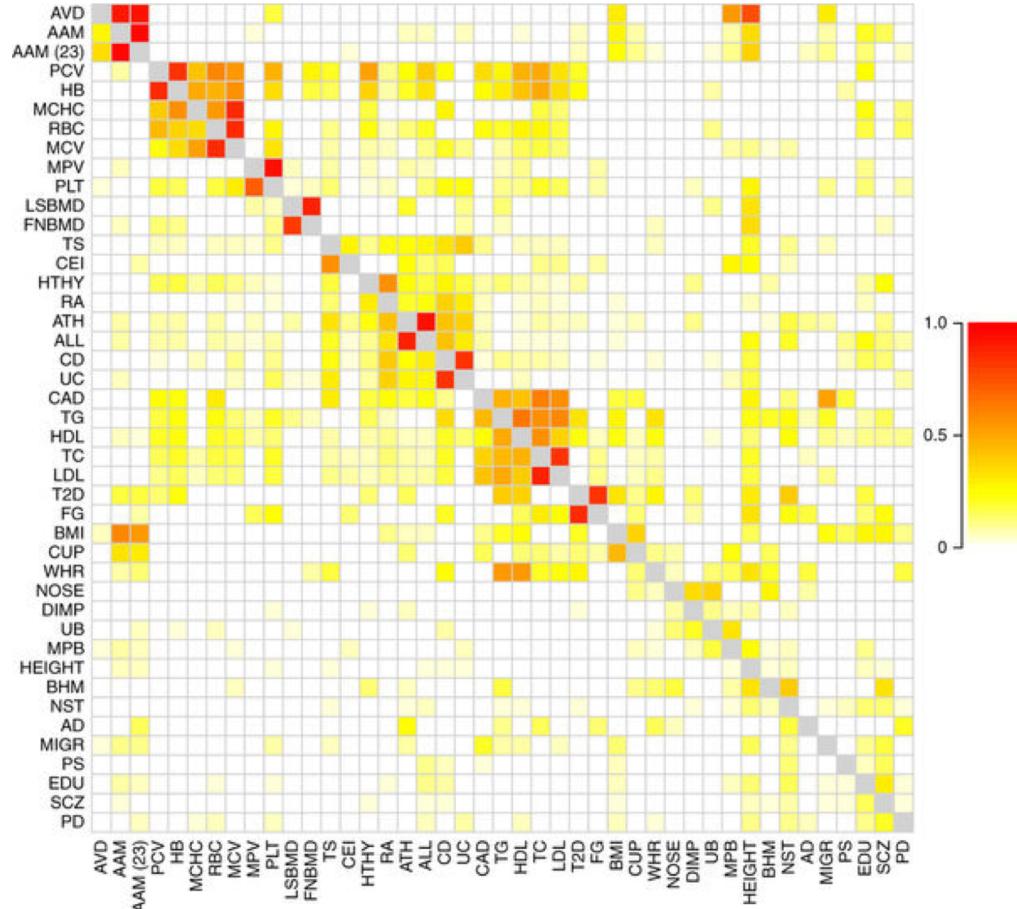
4. Different causal
SNPs in LD



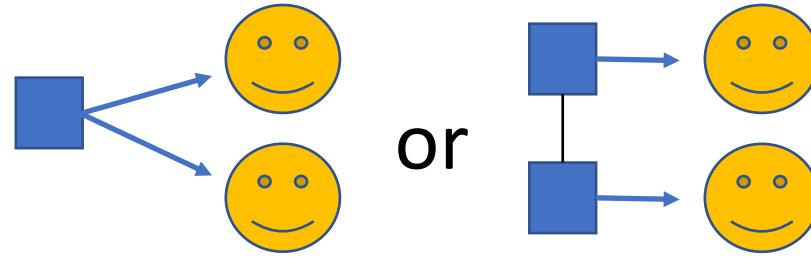
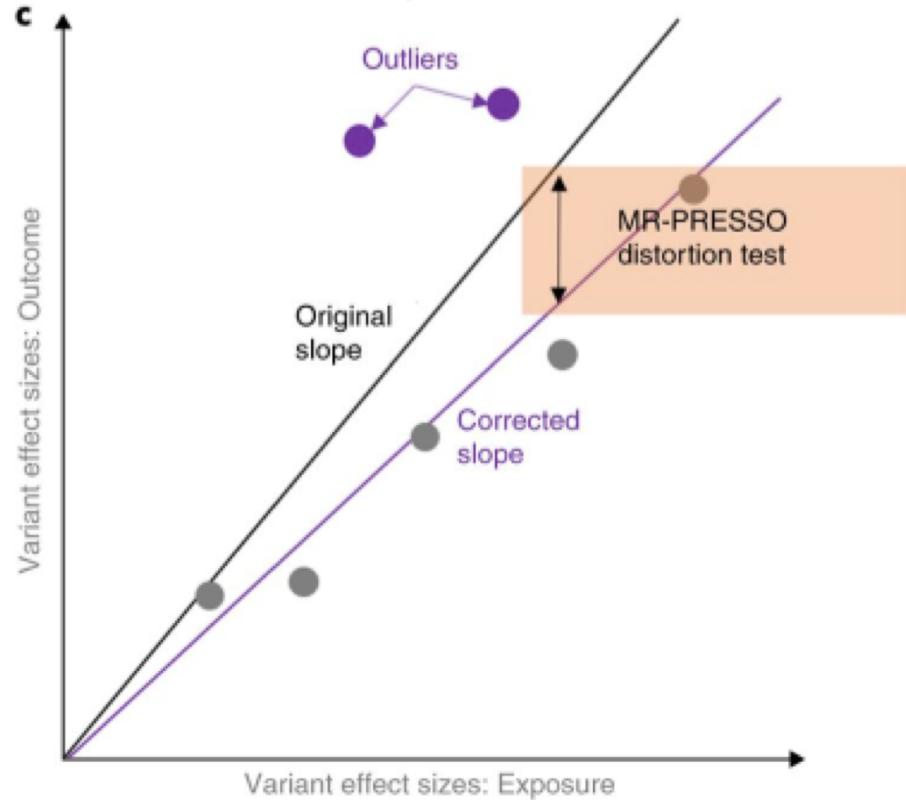
Quantifying pleiotropy: genetic correlation



Quantifying pleiotropy: joint fine mapping top loci



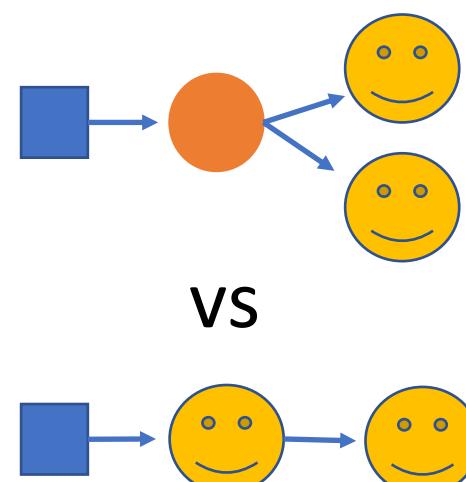
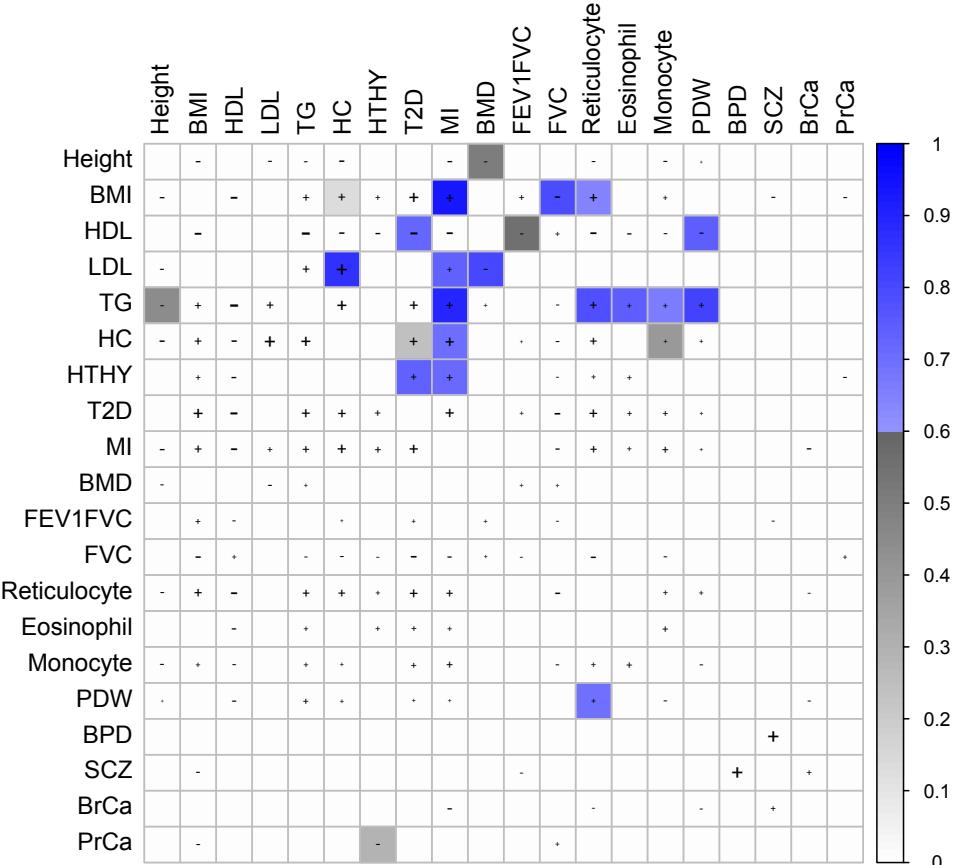
Distinguishing types of pleiotropy: mendelian randomization



VS



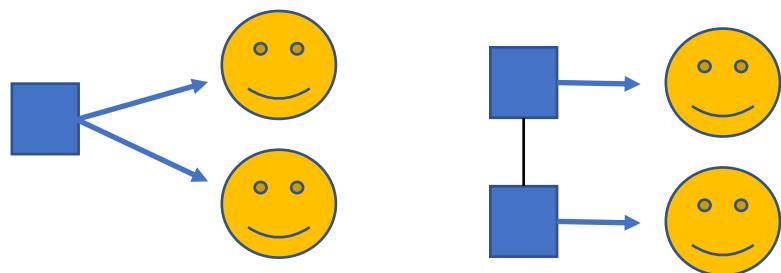
Distinguishing types of pleiotropy: latent causal variable model



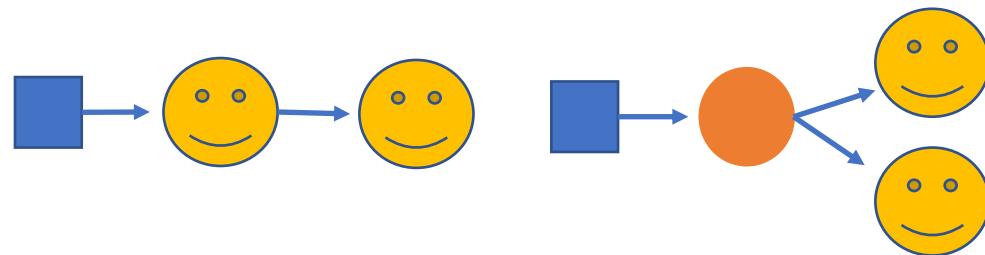
VS

Goals of this work

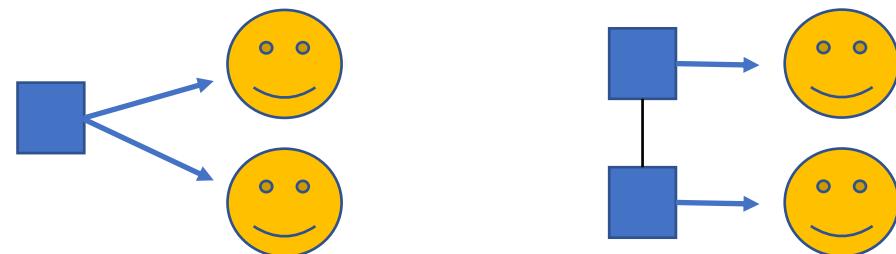
1. Quantify total pleiotropy



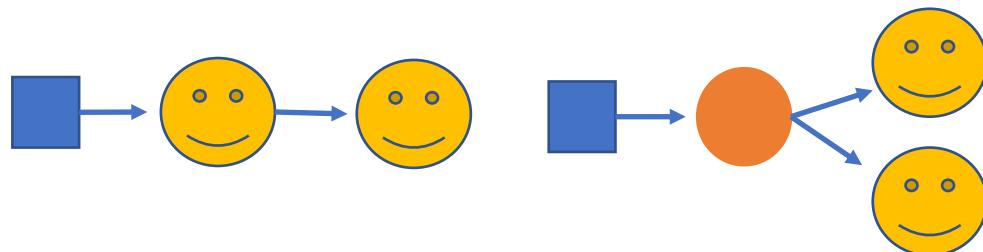
and



2. Quantify uncorrelated pleiotropy



vs



Outline

- Definition of pleiotropy: “excess heritability overlap”
- Estimation: cross-trait LD 4th moments regression
- Simulations
- Results on real traits

Outline

- **Definition of pleiotropy: “excess heritability overlap”**
- Estimation: cross-trait LD 4th moments regression
- Simulations
- Results on real traits

Definition of genetic correlation

$$r_g = \frac{E(\beta_1 \beta_2)}{\sqrt{E(\beta_1^2) E(\beta_2^2)}}$$

Definition of heritability and polygenicity

$$h^2 = ME(\beta^2)$$

$$M_e = \frac{3ME(\beta^2)^2}{E(\beta^4)}$$

Definition of polygenicity: *excess heritability overlap*

$$Q_e = \frac{k_{12}}{k_1 k_2}$$

$$k_1 = E(\beta_1^4) - 3E(\beta_1^2)^2$$

$$k_{12} = E(\beta_1^2 \beta_2^2) - E(\beta_1^2)E(\beta_2^2) - 2E(\beta_1 \beta_2)^2$$

LD-dependent definition of Q_e measures
overlap between marginal effects

β : causal effect sizes

α : marginal effect sizes, inclusive of LD

$$Q_e = \frac{k_{12}}{k_1 k_2}$$

$$k_{12} = 3E(\alpha_1^2 \beta_2^2) - 2E(\beta_1^2 \beta_2^2) - E(\alpha_1^2)E(\beta_2^2) - 2E(\alpha_1 \beta_2)^2$$

$$k_1 = 3E(\alpha_1^2 \beta_1^2) - 2E(\beta_1^4) - E(\alpha_1^2)E(\beta_1^2) - 2E(\alpha_1 \beta_1)^2$$

LD-dependent definition of Q_e measures
overlap between marginal effects

β : causal effect sizes

α : marginal effect sizes, inclusive of LD

$$Q_e = \frac{k_{12}}{k_1 k_2}$$

$$k_{12} = 3E(\alpha_1^2 \beta_2^2) - 2E(\beta_1^2 \beta_2^2) - E(\alpha_1^2)E(\beta_2^2) - 2E(\alpha_1 \beta_2)^2$$

$$k_1 = 3E(\alpha_1^2 \beta_1^2) - 2E(\beta_1^4) - E(\alpha_1^2)E(\beta_1^2) - 2E(\alpha_1 \beta_1)^2$$

Q_e can be conditioned on functional annotations

$$Q_e = \frac{k_{12}}{k_1 k_2}$$

$$k_1 = E(\beta_1^4) - 3E(E(\beta_1^2 | \text{model})^2)$$

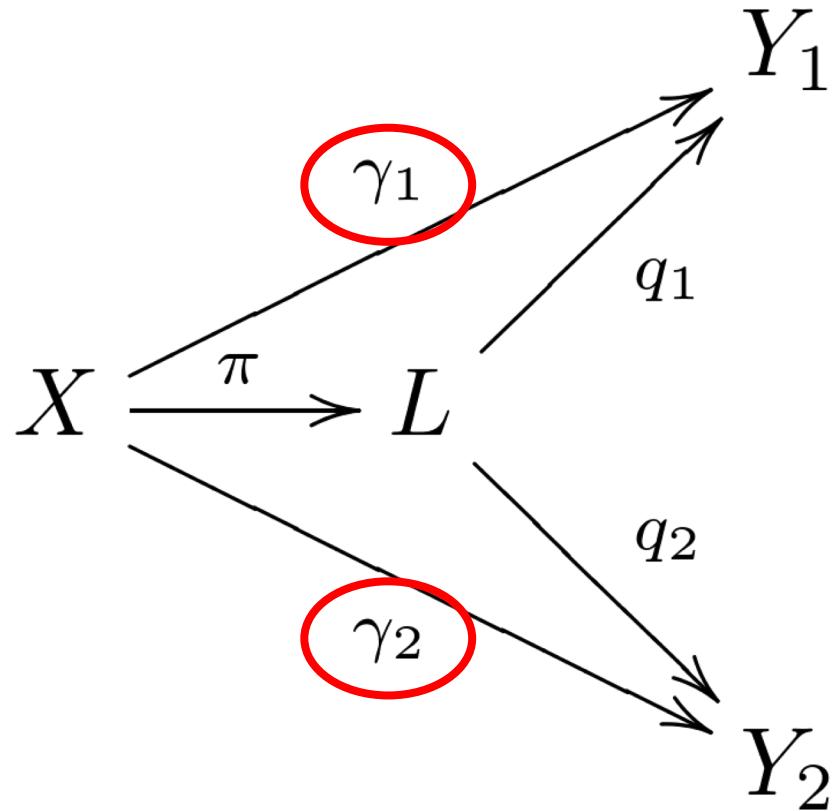
$$k_{12} = E(\beta_1^2 \beta_2^2) - E(E(\beta_1^2 | \text{model})E(\beta_2^2 | \text{model})) - 2E(E(\beta_1 \beta_2 | \text{model})^2)$$

Conditioning on functional annotations

Model 1: $E(\beta^2)$ depends on MAF (Q_e)

Model 2: $E(\beta^2)$ depends on baseline-LD model ($Q_{\text{baseline-LD}}$)

Q_{LCV} can isolate uncorrelated pleiotropic effects



Q_{LCV} measures overlap between γ_1 and γ_2

Outline

- Definition of pleiotropy: “excess heritability overlap”
- **Estimation: cross-trait LD 4th moments regression**
- Simulations
- Results on real traits

Outline

- Definition of pleiotropy: “excess heritability overlap”
- **Estimation: cross-trait LD 4th moments regression**
- Simulations
- Results on real traits

Estimation approach

- Regress products of χ^2 statistics (estimates of $\alpha_1^2 \alpha_2^2$) on stratified LD 4th moments (sums of r^4 across SNPs)
- Cross-trait LDSC: regress products of Z scores on stratified LD scores (sums of r^2) (Bulik-Sullivan et al. 2015b NG)

Cross-trait LD4M regression equation

$$E(\alpha_1^2 \alpha_2^2) \approx E(\alpha_1^2)E(\alpha_2^2) + 2E(\alpha_1 \alpha_2)^2 + l^{(4)} k_{12}$$

$l^{(4)}$: LD 4th moment of regression SNP

k_{12} : proportional to Q_e

Stratified case: $l_i^{(4)}$ and K are $1 \times P$ and $P \times 1$ vectors

Relies on an LD approximation ("LD between causal SNPs is either strong or weak")

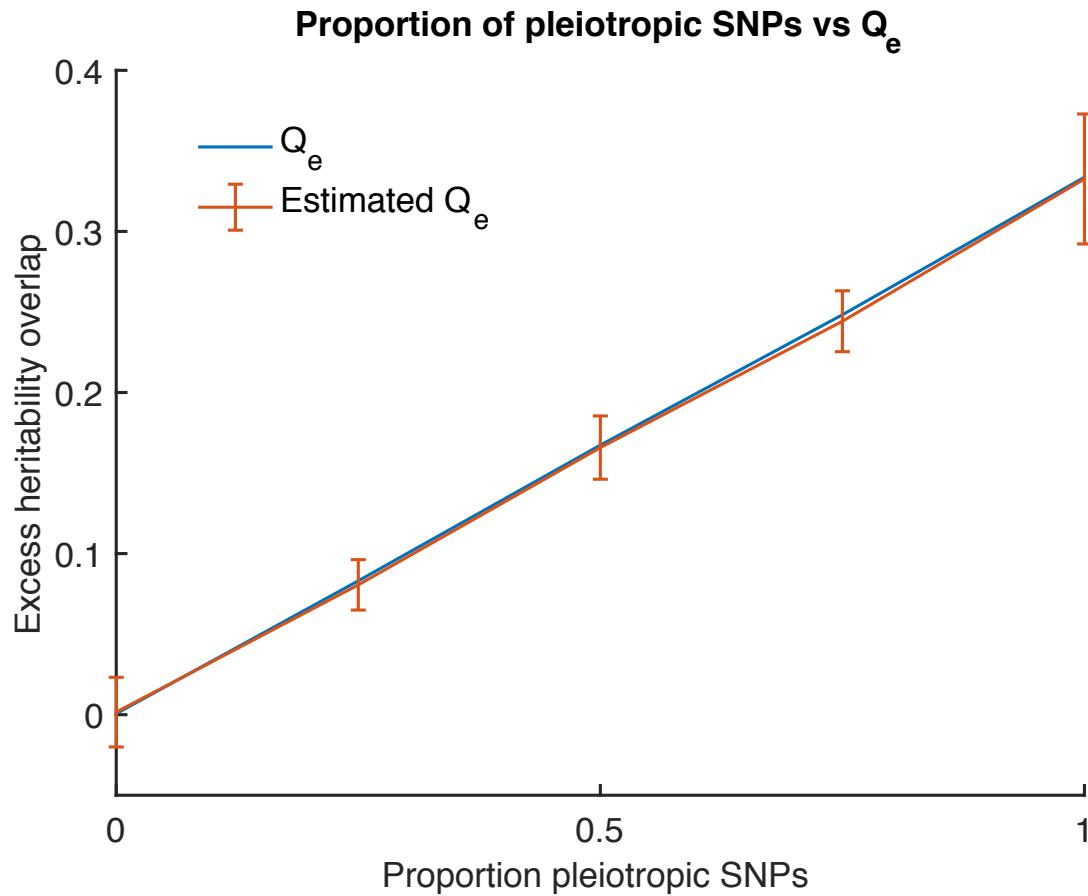
Outline

- Definition of pleiotropy: “excess heritability overlap”
- Estimation: cross-trait LD 4th moments regression
- **Simulations**
- Results on real traits

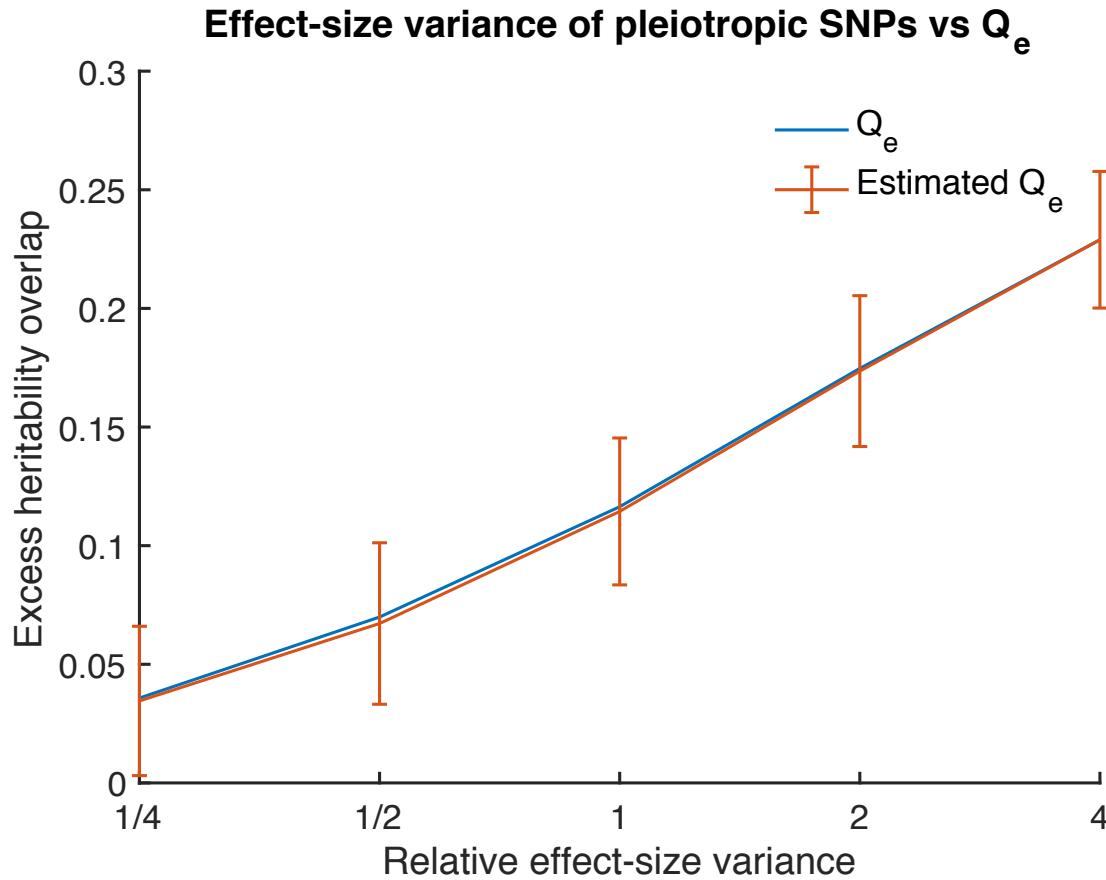
Simulations

- Real UKB LD: 100k common SNPs on chromosome 1, N=100k
- Mixture of multivariate normal distributions:
 - Null SNPs
 - SNPs affecting trait 1 only
 - SNPs affecting trait 2 only
 - SNPs affecting both traits, $r=0$
 - SNPs affecting both traits, $r=1$
- 1 simulated functional category with different mixture weights

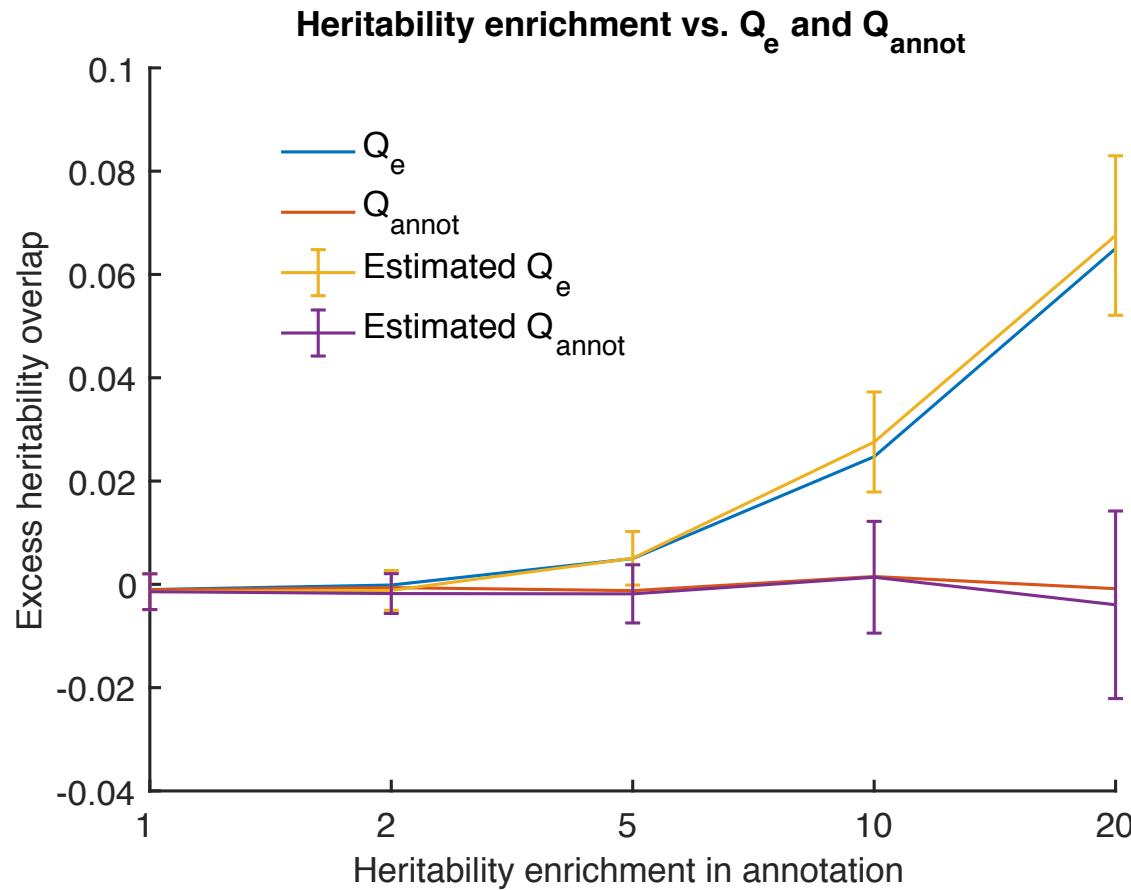
Simulations with no enrichment or genetic correlation



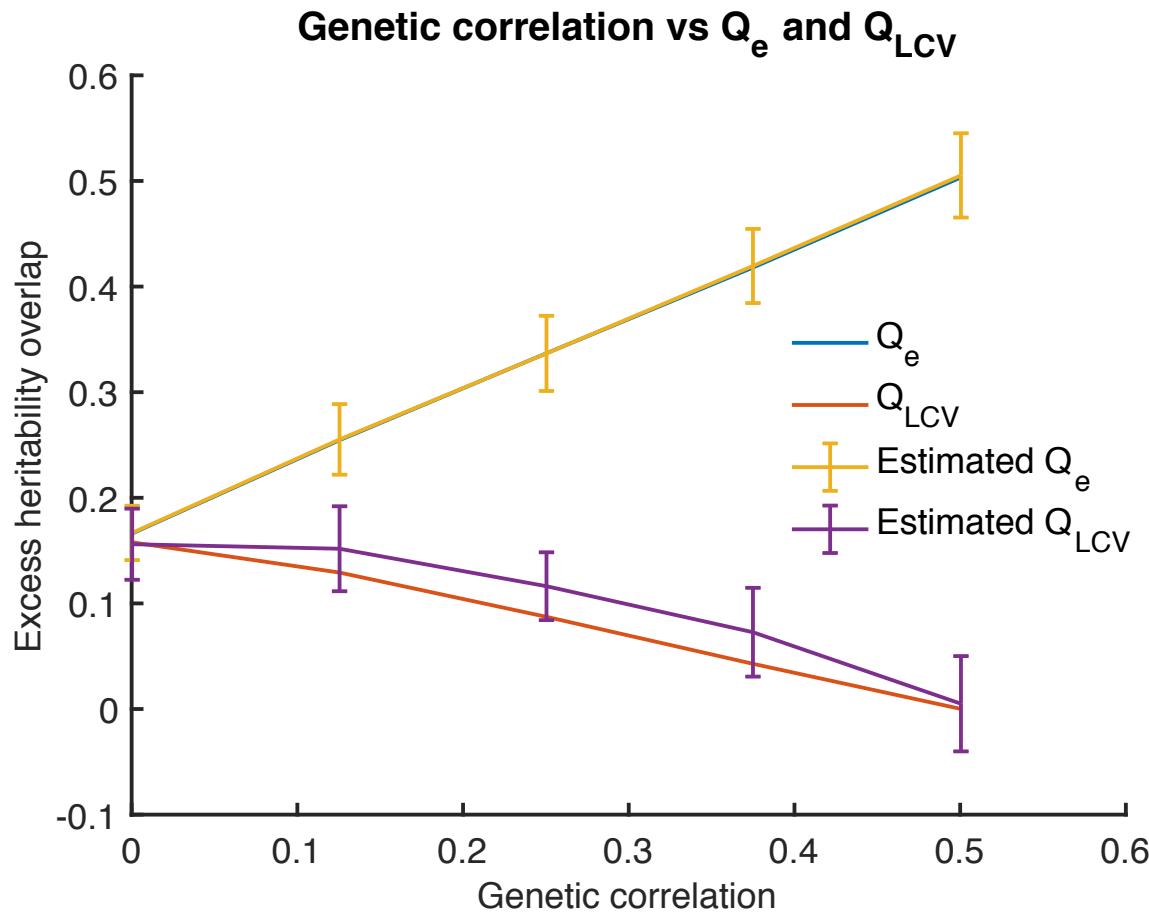
Simulations with no enrichment or genetic correlation



Simulations with increasing functional enrichment



Simulations with increasing genetic correlation



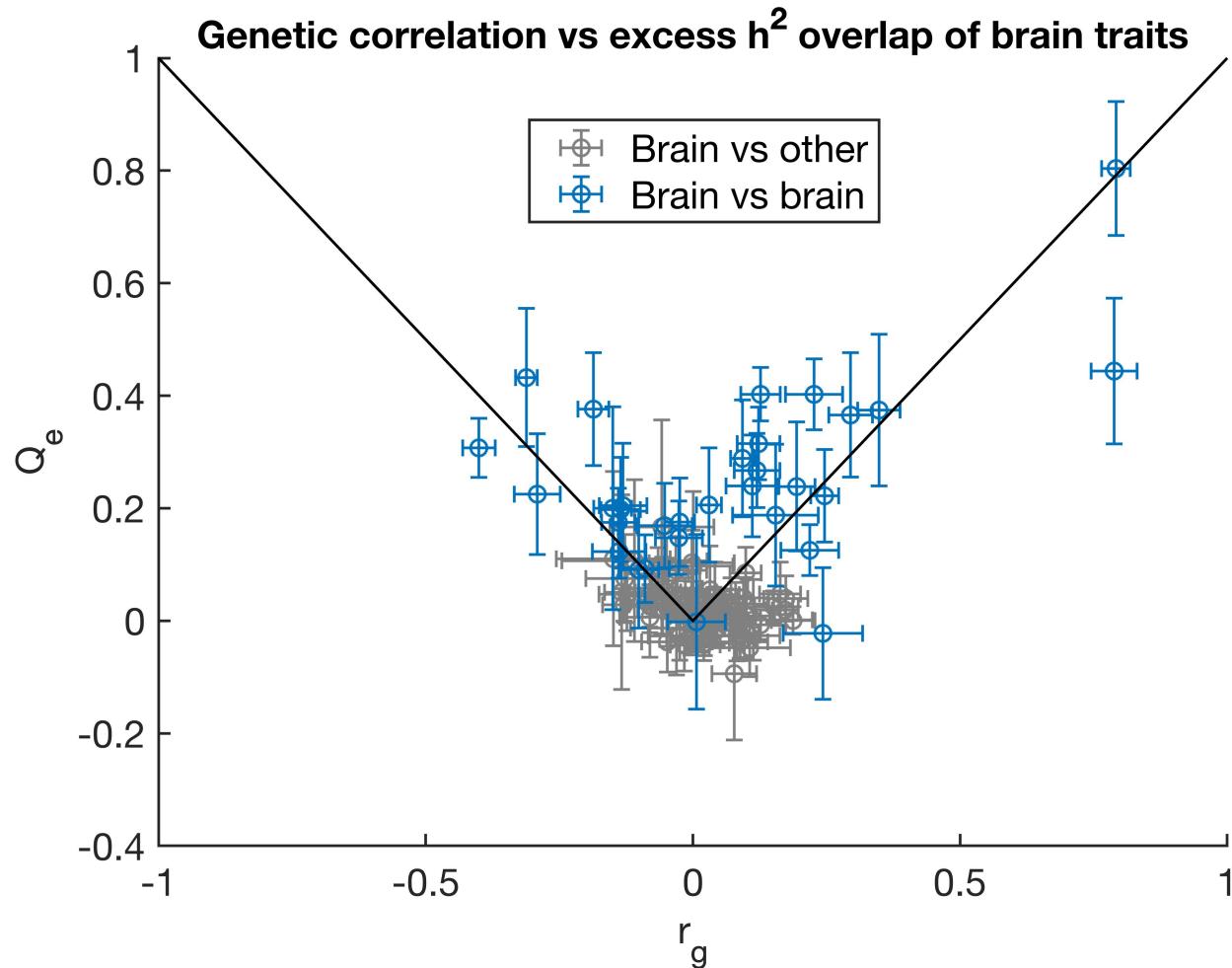
Outline

- Definition of pleiotropy: “excess heritability overlap”
- Estimation: cross-trait LD 4th moments regression
- Simulations
- **Results on real traits**

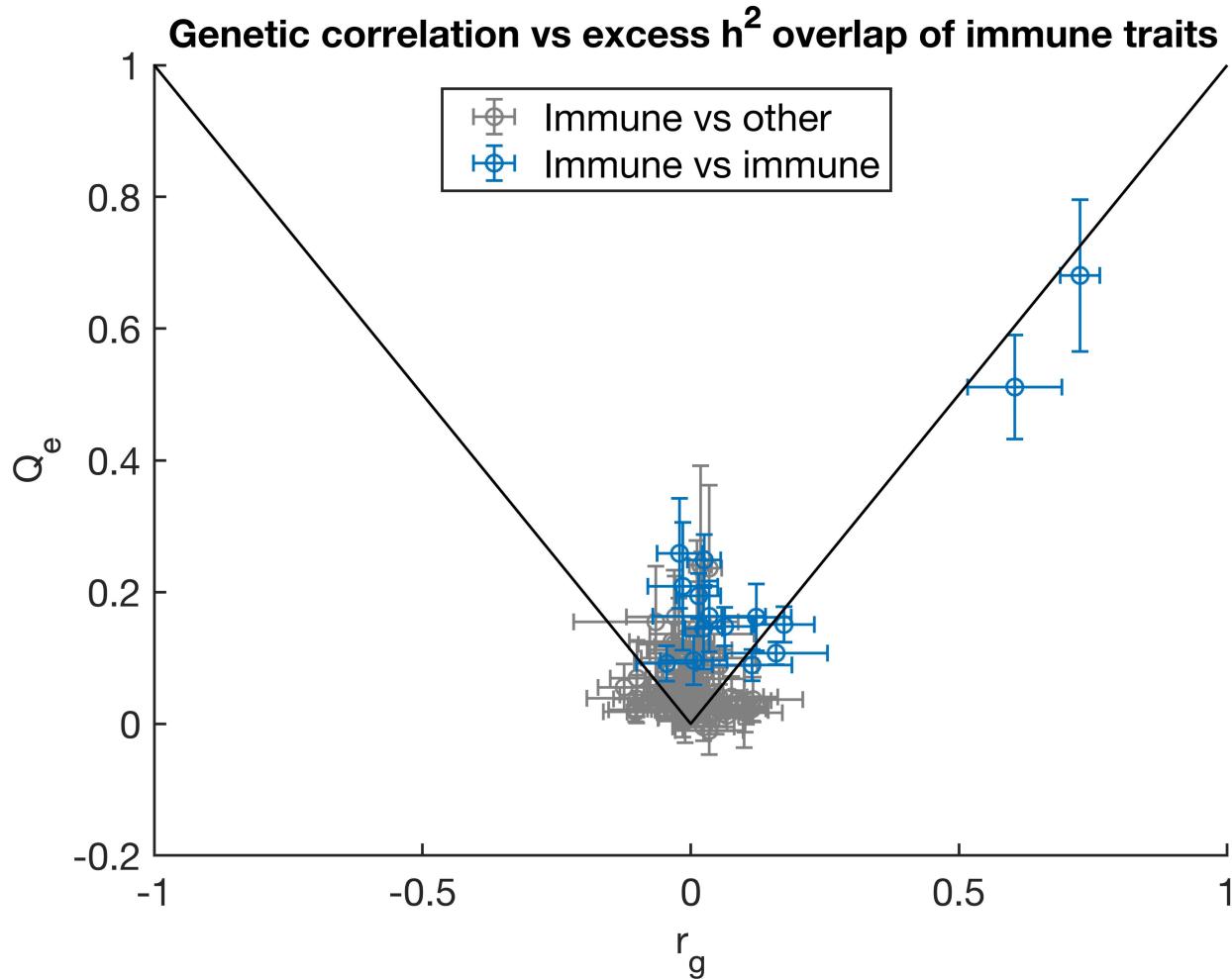
Data analyzed

- 28 phenotypes, mostly from UKB ($N \sim 460k$):
 - 12 blood traits
 - 10 brain-related traits
 - 6 immune/inflammatory traits
- LD scores + LD 4th moments with 10 common MAF bins from UK10K

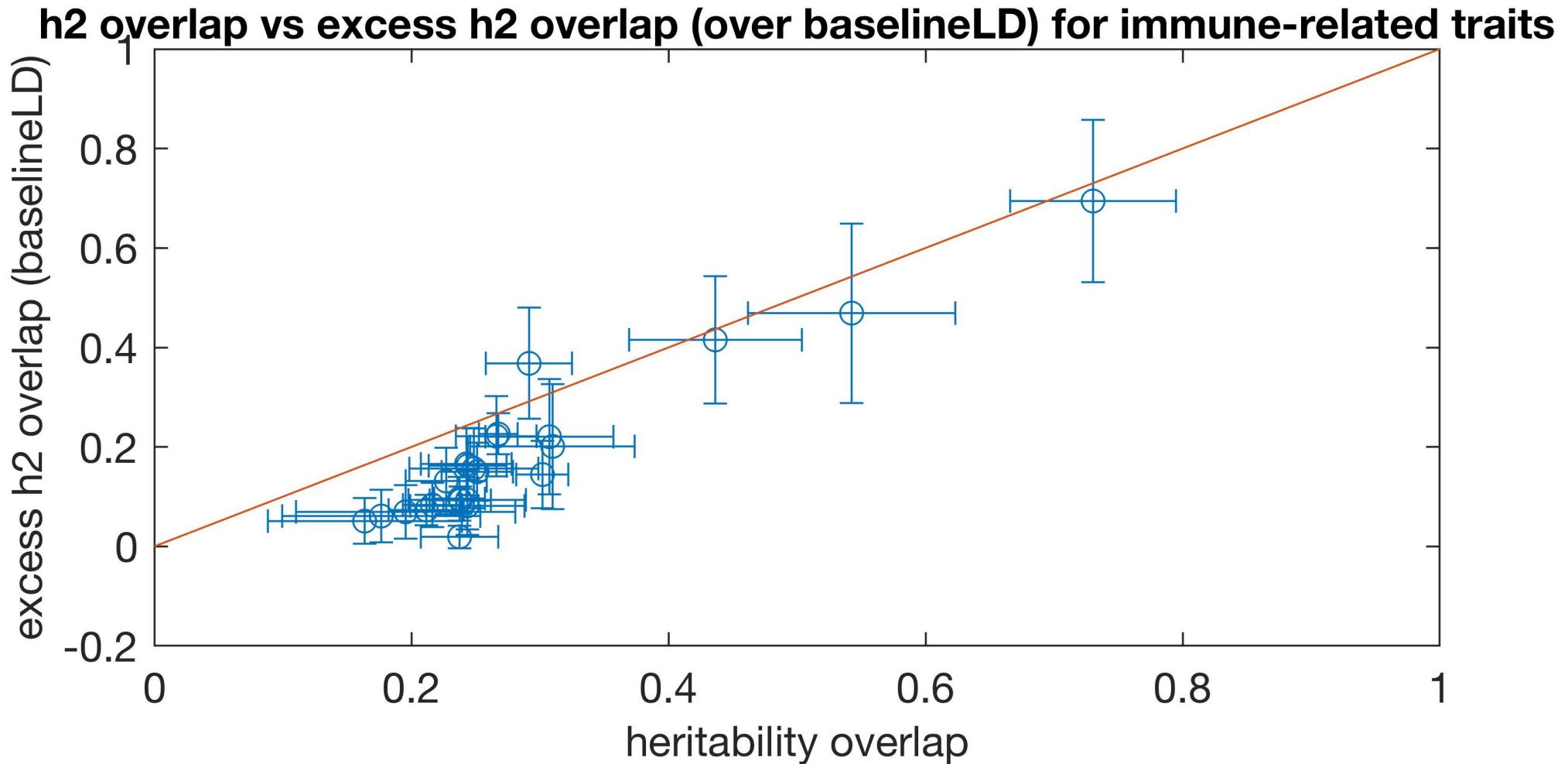
Pervasive pleiotropy among brain-related traits



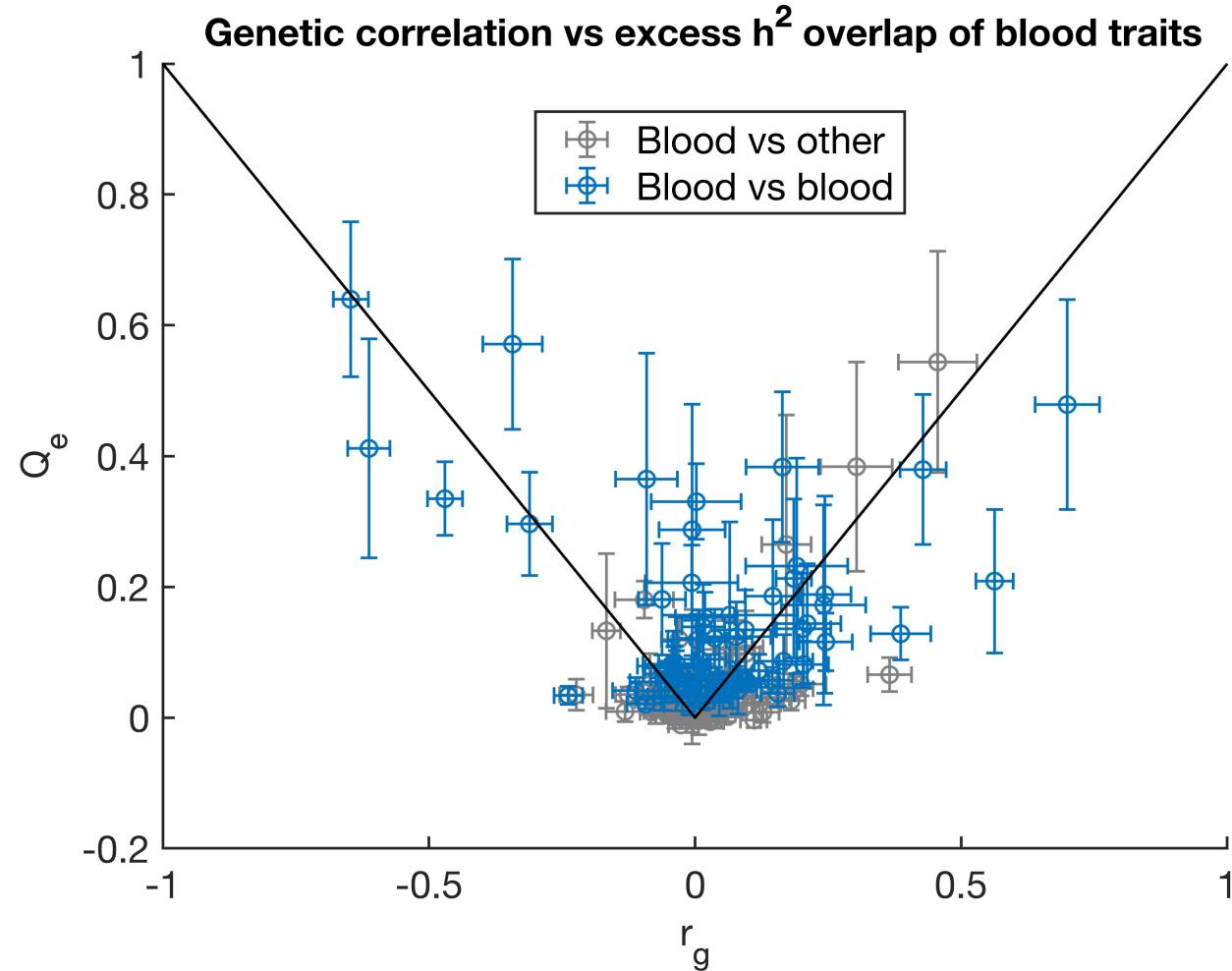
Pervasive pleiotropy among immune-related traits



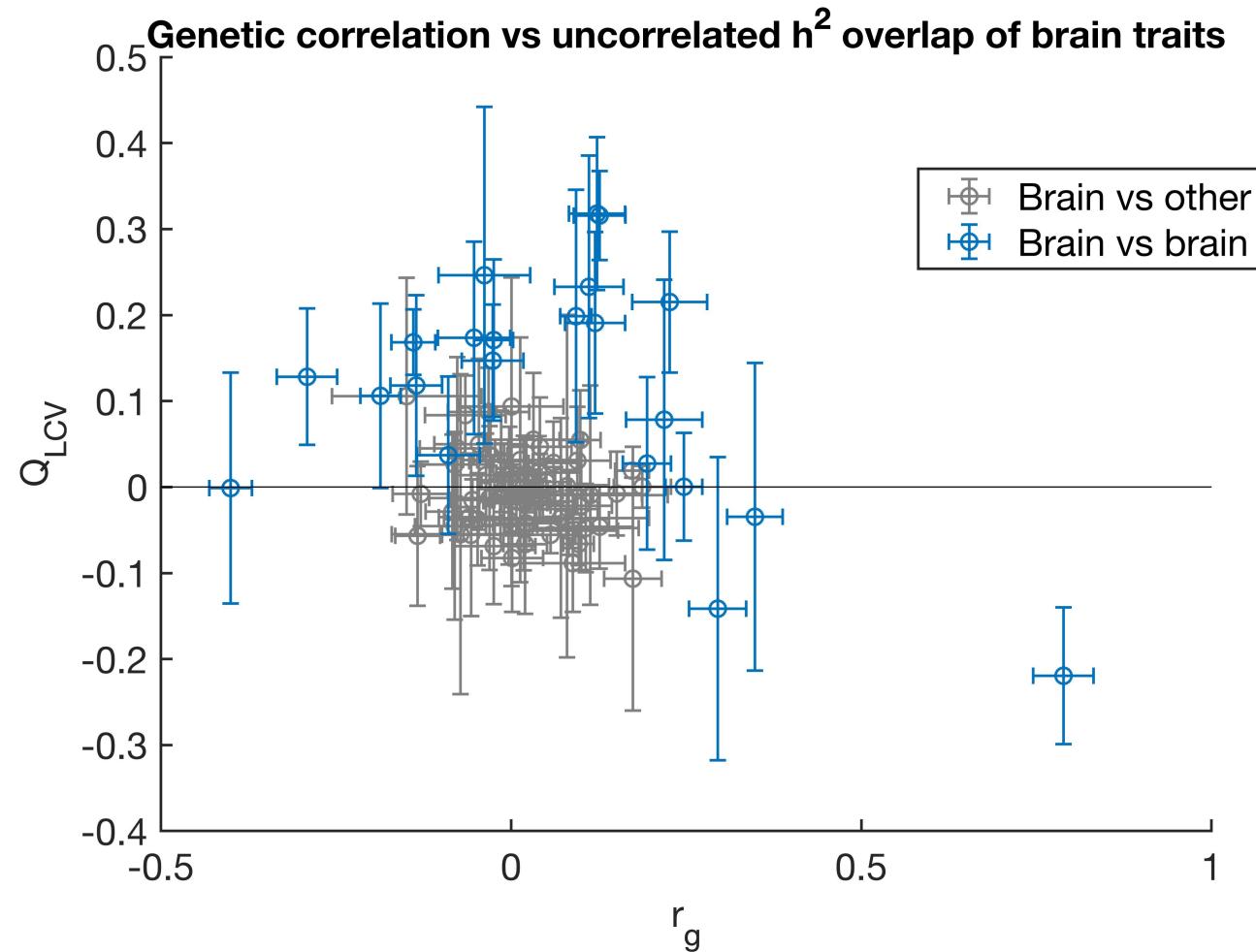
Immune trait pleiotropy not driven by shared enrichments



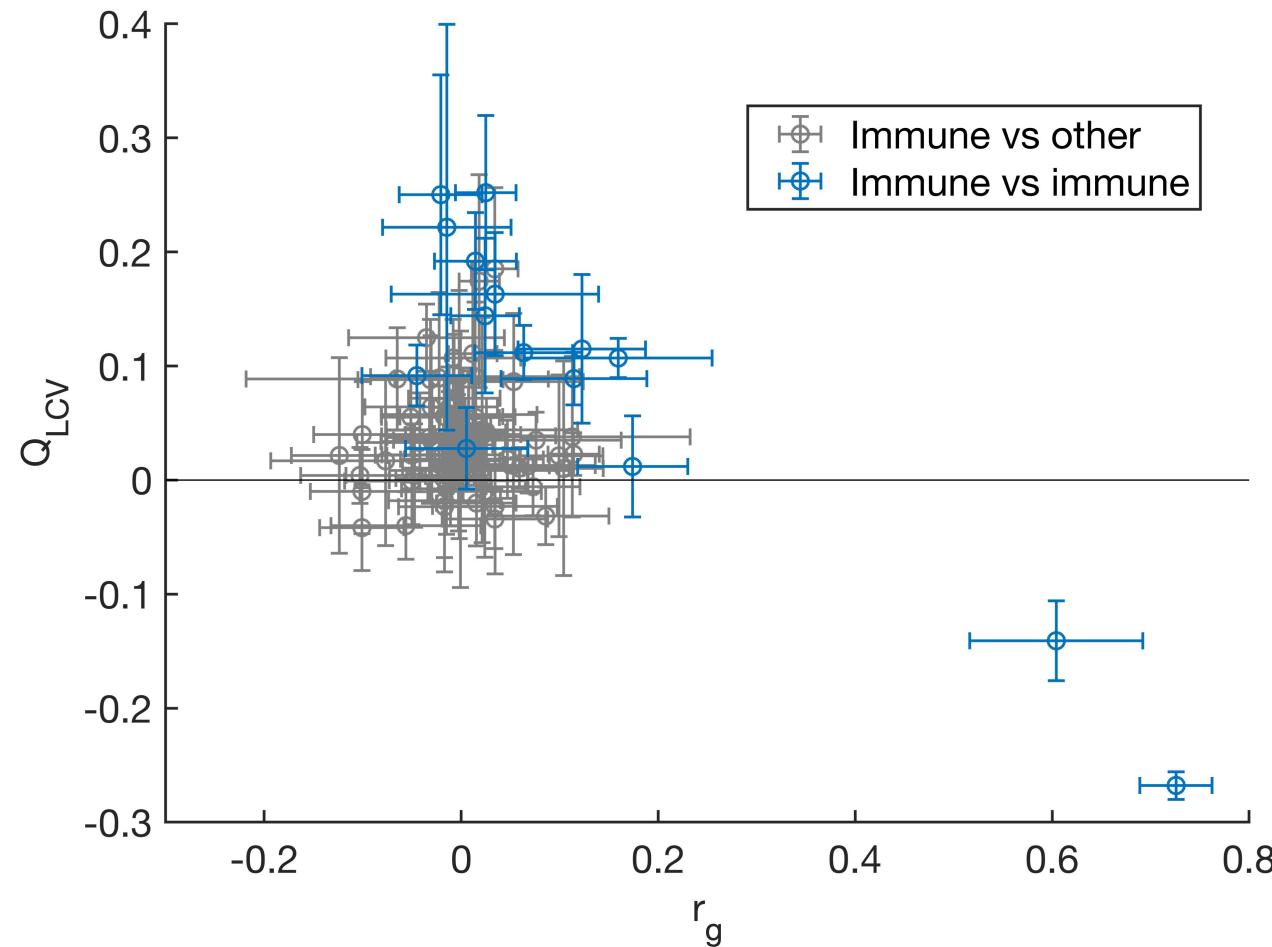
Little pleiotropy among most blood-related traits



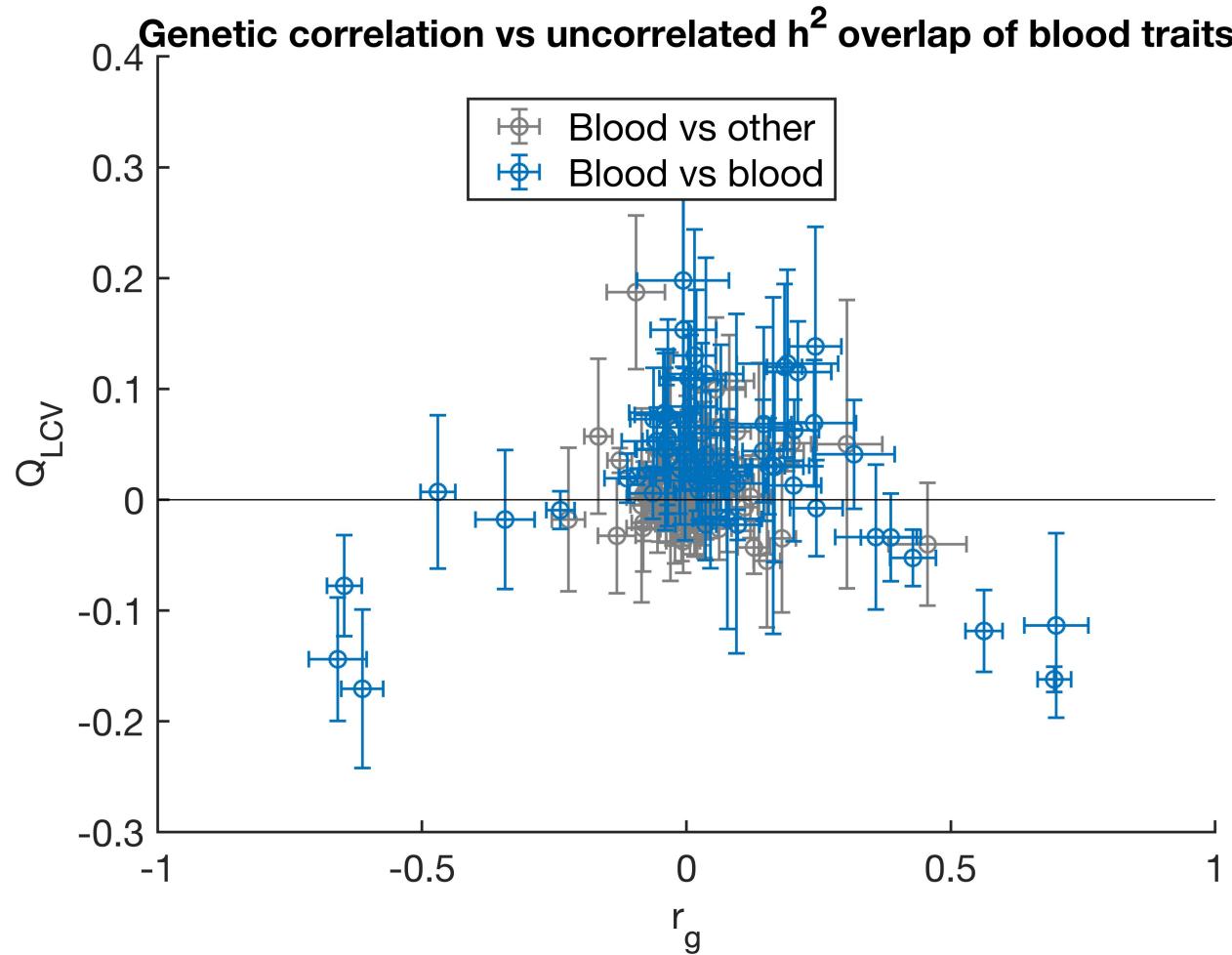
Brain trait pleiotropy largely uncorrelated



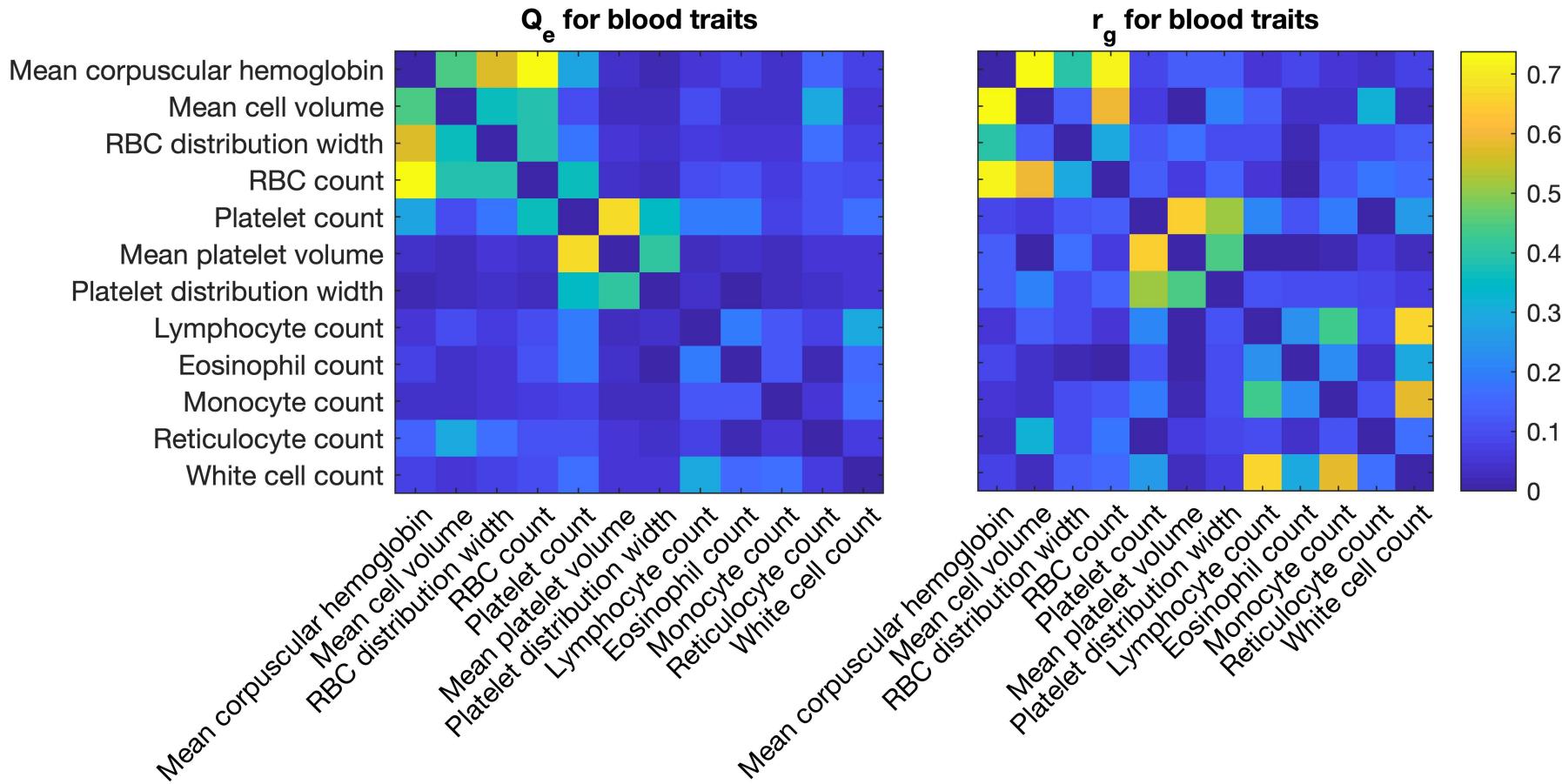
Immune trait pleiotropy largely uncorrelated



Some blood traits have uncorrelated pleiotropy



Blood traits cluster by cell type



Summary

- Excess heritability overlap (Q_e) is a natural way to quantify and distinguish different types of pleiotropy
- Cross-trait S-LD4M produces reliable estimates of Q_e
- Traits that share causal cell types almost always have high Q_e
- For brain and immune traits, cell-type-mediated pleiotropy is largely uncorrelated