

AdjHE: An efficient way to estimate heritability

Christian Coffman

Advised by: Dr. Saonli Basu
University of Minnesota Twin Cities
Division of Biostatistics

October 28, 2022



- 1 General trait influencers**
- 2 Methods for detecting role of genetics**
- 3 AdjHE increases computational efficiency of heritability estimation**
- 4 Simulations**

General influences on traits



Traits are determined by different contributions of genetics and environmental influencers.

Which traits are dictated by which set of influencers?

Image credit:

<https://blogs.kcl.ac.uk/editlab/2019/05/07/if-something-is-genetic-it-can-still-be-influenced-by-the-environment/>

GWAS

- Genome Wide Association studies (GWAS)

$$Y = X\beta + \epsilon$$

- Hypothesis testing done on the β (sometimes millions)
- Great for highly influential SNP's
- Low power for causality spread across multiple SNP's

GRM based heritability estimation

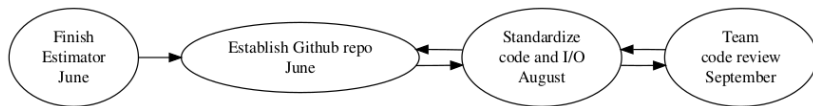
- Describe variation in phenotype as random effect (LMM)

$$Y = \epsilon, \epsilon \sim N(0, \sigma_g^2 GRM + \sigma_e^2 I)$$

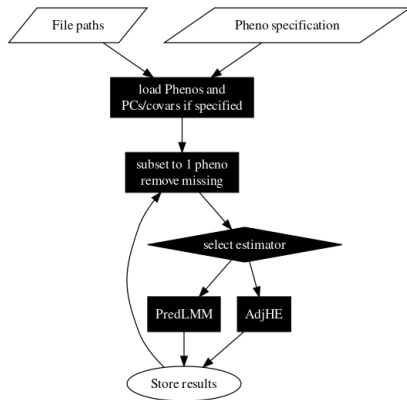
- Genetic Relatedness Matrices (GRM's) describe relatedness between subjects
- Lose resolution on genome
- Gain power for dispersed genetic effects
- GCTA uses REML which can be slow with large studies ($n \times n$ matrix)
- Not efficient for exploration of mildly heritable traits (large sample sizes)

New tool: AdjHE

- Two-stage Method of Moments approach
- Assumes ethnicities and covariate effects ortho to genetic effects yielding closed form
- Much more efficient
- Benchmarked: 2x faster with 4000 subjects
- 10x faster with 45k subjects

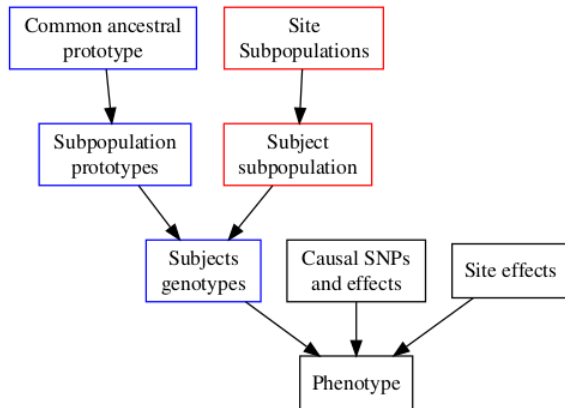


Design features



- Data reading meets and expands usable data
- Easy integration with command line interface
- Model selection with internal looping
- Automated diagnostic plots
- Multiple methods of estimation available

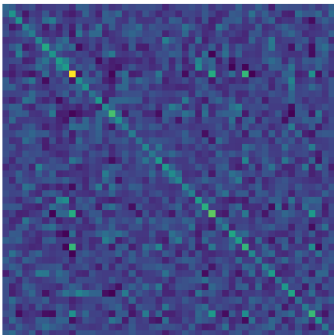
Additional Feature



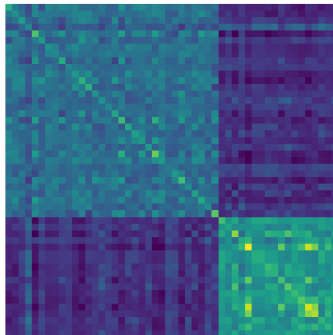
- Simulate realistically structured GRM's and phenotypes
- Determine what scenarios fit within AdjHE model

Simulating population structures

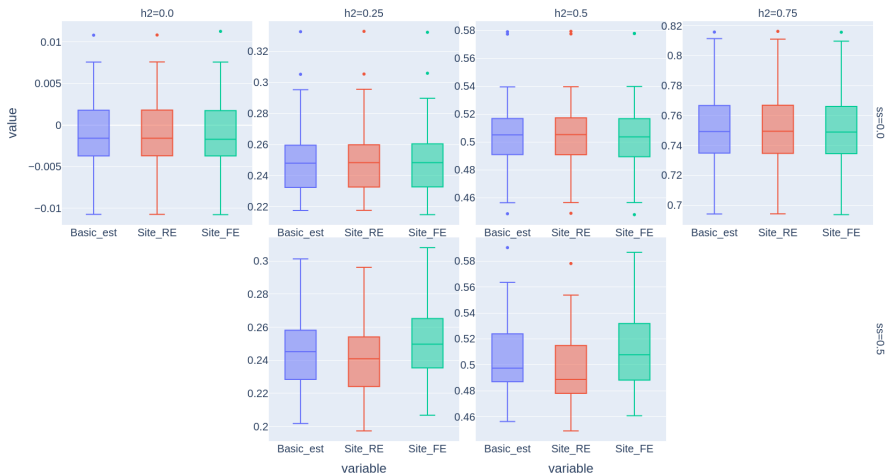
GRM: clusters = 1, Site pops = Homo



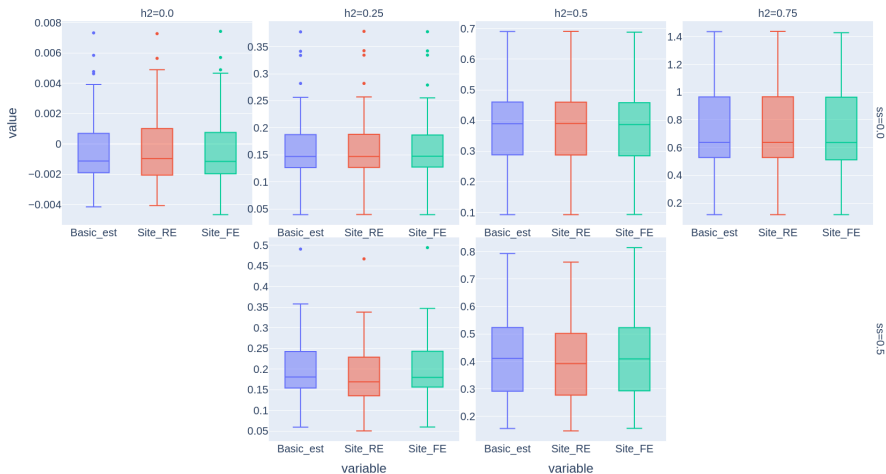
GRM: clusters = 2, Site pops = IID



Estimation on Homogeneous

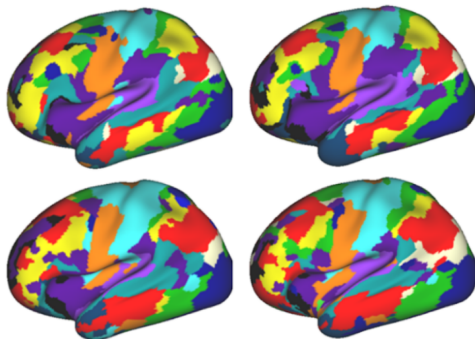


Estimation on Heterogeneous



Conclusions and future aims

- AdjHE is efficient estimator and accounts for basic effect from site
- Need to expand to estimation with imbalanced sites for multisite studies (ex. ABCD)
- Applications to functional topology



Thank you for listening
Questions?

References

Lin, Seal, and Basu. “Estimating SNP Heritability in Presence of Population Substructure in Biobank-Scale Datasets.” *Genetics* 2022

Hermosillo et al. “A Precision Functional Atlas of Network Probabilities and Individual-Specific Network Topography.” 2022
bioRxiv