The effect of evolutionary history and linkage disequilibrium on the genetic architecture of complex traits

Saige Daines[1], Xinjun Zhang[2], Kirk Lohmueller[2,3]

1 BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
2 Department of Ecology and Evolutionary Biology, UCLA
3 Department of Human Genetics, David Geffen School of Medicine, UCLA

Introduction

- Quantitative traits, including most disease, have a complex genetic architecture with many genetic variants (SNPs) contributing to the phenotype of the trait.
- For a given trait, genetic architecture, including the frequency of causal variants and heritability of the trait, can vary between populations. For example, the effect sizes of causal variants may correlate with the strength of selection in a population.
- It is unclear how linkage disequilibrium (LD), demographic history, and the relationship of effect sizes with selection, which are all important evolutionary forces, affect the genetic architecture of complex traits.
- Additionally, the ability to detect these causal mutations using Genome-wide association studies (GWAS) is of importance in understanding the genetic make-up and heritability of complex traits.
- However, it is also unclear how differences in genetic architecture due to LD, demographic history, and distribution of effect sizes affect GWAS's fine-mapping ability and power to detect causal variants of a complex trait.

Questions

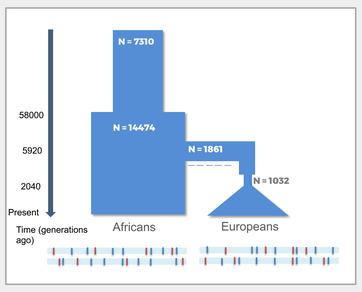
- How does linkage disequilibrium and demographic history affect the genetic architecture of quantitative traits when the effect size of causal variants is either related to or independent of purifying selection?
- How does genetic architecture affect the fine-mapping ability and power of GWAS in detecting causal mutations of a quantitative trait?

Methodology

Simulate African and European genomes with neutral and causal mutations → **Assign effect sizes to causal mutations** → **Simulate phenotype of individuals** → **Conduct GWAS analysis** • **Compute heritability** • **Analyze genetic architecture**

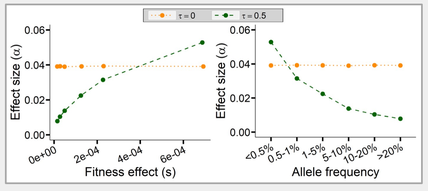
1. Simulate African and European genomes from a modern human demographic model under low and high recombination rates

Simulated African and European genomes with input of neutral and deleterious mutations under different recombination rates using a modern human demographic model



2. Model a complex trait in African and European populations :

1. Assigned **effect sizes of deleterious mutations** so heritability was ~ 0.1 under a model where:
   - Effect sizes of causal mutations are **independent of the strength of selection** acting upon them (tau=0)
   - Effect sizes are **correlated with the strength of selection** (tau=0.5)



2. **Simulated phenotype** for each individual given the genotype, effect size, and environmental variance

Results

GWAS analysis

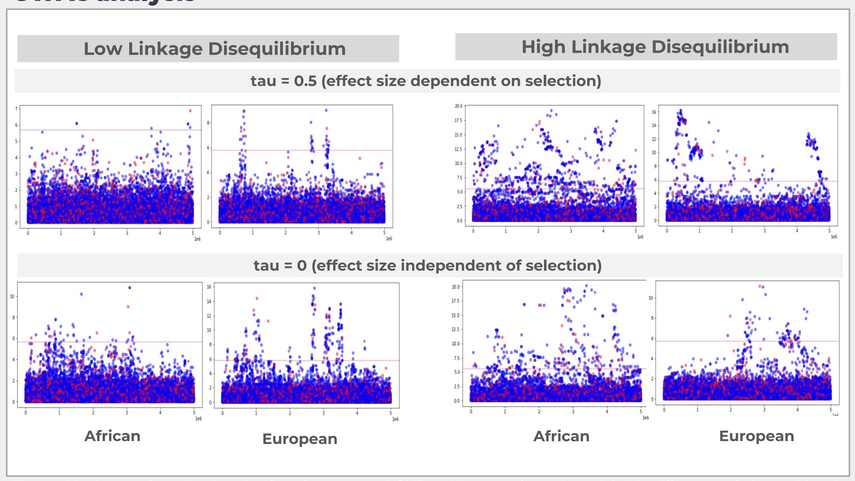


Figure: Manhattan plots showing significantly associated SNPs with a complex trait modeled in different evolutionary histories (African and European populations) under low and high Linkage disequilibrium and where effect sizes of causal mutations are (tau=0.5) or are not (tau=0) correlated with the strength of selection.

Heritability

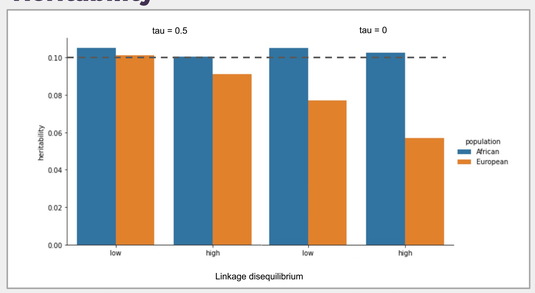


Figure: Barplot comparing the genetic heritability of quantitative traits modeled with effect size correlated with selection (tau=0.5) or independent of selection (tau=0) and under low and high LD. In low LD, tau=0.5 heritability was simulated to be 0.1, represented by the dashed line.

- In all scenarios, Europeans have lower heritability than Africans
- When effect size is independent of selection (tau=0) and LD is high, European heritability further decrease

Genetic architecture and GWAS power:

		---------- mean frequency of SNPs in population ----------						
		causal hits	all causal SNPs	neutral hits	all neutral SNPs	# GWAS hits	% causal : neutral hits	% causal hits, all causal muts
tau 0	African low LD	0.257353	0.029624	0.249000	0.055626	264	25.118483	0.190394
	European low LD	0.386870	0.023177	0.410461	0.045761	665	28.875969	0.409161
	African high LD	0.254657	0.019376	0.256124	0.031666	1109	22.541436	0.832993
	European high LD	0.372567	0.014655	0.390156	0.024091	634	22.157996	0.339063
tau 0.5	African low LD	0.192682	0.029624	0.253723	0.055626	41	36.666667	0.039516
	European low LD	0.350608	0.023177	0.415051	0.045761	113	28.409091	0.068651
	African high LD	0.226732	0.019376	0.225130	0.031666	1600	22.982321	1.220906
	European high LD	0.366273	0.014655	0.394992	0.024091	678	26.728972	0.421617

Figure: Comparison of the mean frequency of causal and neutral SNPs in African and European populations and of causal and neutral associated GWAS hits in a trait modeled under low or high LD and where effect sizes of causal mutations are (tau=0.5) or are not (tau=0) correlated with the strength of selection.

- Linkage disequilibrium increases the percentage of causal variants detected by GWAS in all models
- Linkage disequilibrium decreases the number of causal hits compared to neutral hits detected by GWAS
- The mean frequency of causal hits decreases when effect size depends on selection strength (tau=0.5)
- When tau=0.5, a higher proportion of causal versus neutral mutations are detected than when effect sizes are independent of selection (tau=0)

Conclusion

In this project, we simulated two types of quantitative traits where the effect size was either related to or independent of the strength of purifying selection. Under each scenario we varied the LD and demographic model.

We have shown that the genetic architecture of quantitative traits, including the frequency of causal mutations, their effect sizes, and heritability, varies between populations with different demographic histories, under varying levels of linkage disequilibrium, and with different selection models.

Additionally, our results indicate that LD decreases the ability of GWAS to fine-map causal mutations and increases GWAS's power to detect causal mutations regardless of the selection or demographic model, while a correlation between selection and effect size distribution increases fine-mapping and decreases power.

These results suggest the importance of taking into consideration demographic history, linkage disequilibrium, and tau when analyzing complex traits and interpreting GWAS results of traits in different genomic regions within populations and between populations.

Special thanks to Kirk Lohmueller and Xinjun Zhang for their guidance and mentorship!