EVOLUTIONARY CONSTRAINT HIGHLIGHTS NONCODING VARIANT ENRICHMENT ACROSS COMPLEX TRAITS

CLARISSA LAI [1], Chenlu Di [2], Kirk E. Lohmueller [2,3]

[1] B.I.G. Summer Program, Institute for Quantitative and Computational Biosciences, UCLA
[2] Dept of Ecology and Evolutionary Biology, UCLA
[3] Dept of Human Genetics, David Geffen School of Medicine, UCLA

Background + Introduction

Big Picture

- Noncoding regions have important **regulatory functions,** influencing gene expression without altering protein sequences.
- Most genetic variants associated to human traits lie in **noncoding regions** of the genome.

Evolutionary Constraint

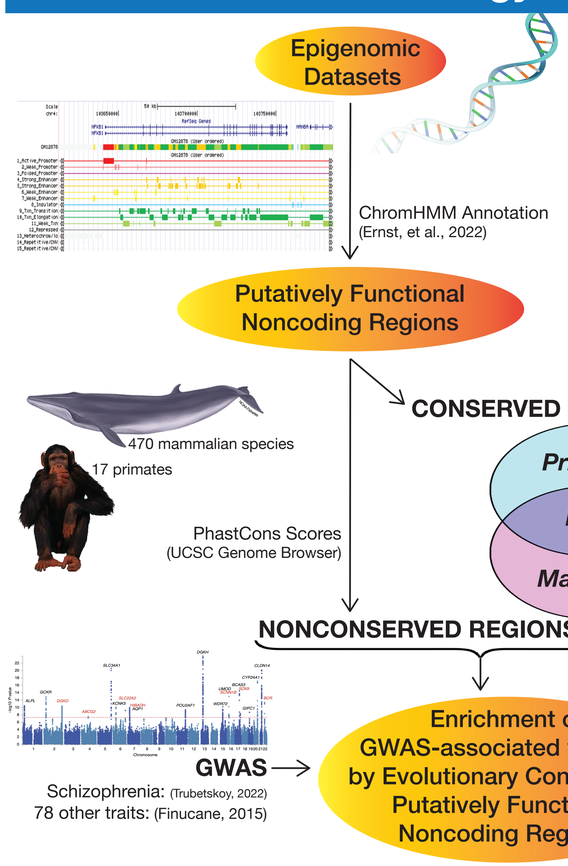
- Conserved noncoding regions (those maintained across species) are often under **negative selection,** suggesting functional importance.
- Levels of conservation can differ: **(1)** primates only
  (2) mammals only
  (3) both primates + mammals

Knowledge Gap

- While prior modeling of the distribution of fitness effects (DFE) in noncoding regions links conservation to **deleterious mutation burden,** the **phenotypic impact** of evolutionary constraint is less understood.

Research Question: ***Is GWAS variant enrichment related to evolutionary constraint in noncoding regions?***

Overview of Methodology & Data



Functional Noncoding (ChromHMM)
Keep: Active enhancers (EnhA) and promoters (TSS, PromF), etc.
Remove: Coding(GENCODE), Quiescent(Quies), Heterochromatin(HET)

Conservation bins (PhastCons)

- **Primate-conserved:** top 5% of PhastCons-17way
- **Mammal-conserved:** top 5% of PhastCons-470way
- **Conserved in both:** overlap of the two top-5% sets
- **Nonconserved:** bottom 60% of scores (=0 in mammals)

GWAS enrichment pipeline

1) Filter fine-mapped variants (Posterior Inclusion Probability ≥ 0.5)
2) Count SNPs per conservation category (primate, mammal, both)
3) Compute odds ratios for each category vs. nonconserved

Results

Figure 1.
Schizophrenia Variant Enrichment by Evolutionary Constraint Across PIP Thresholds

Odds ratio (Fisher's exact test) enrichment of Schizophrenia. Higher PIP corresponds to SNPs with stronger statistical support from fine-mapping. At PIP≥0.5, enrichment for conservation in mammals and both fall to zero, but continues to increase in primates.

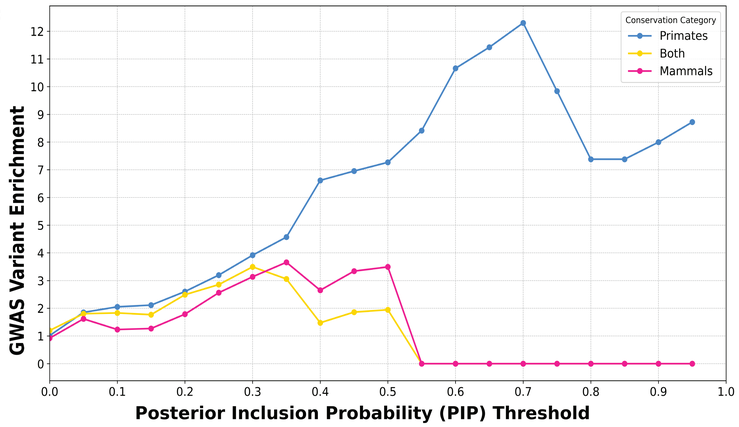
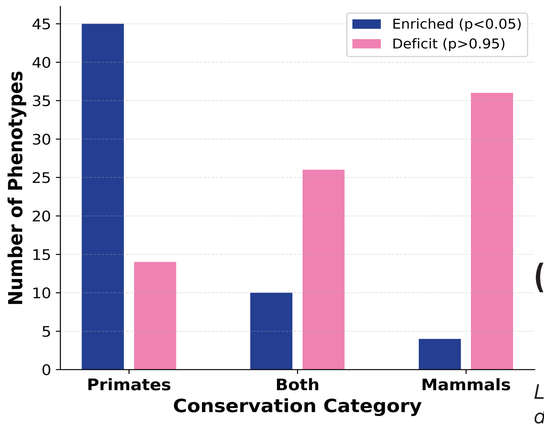


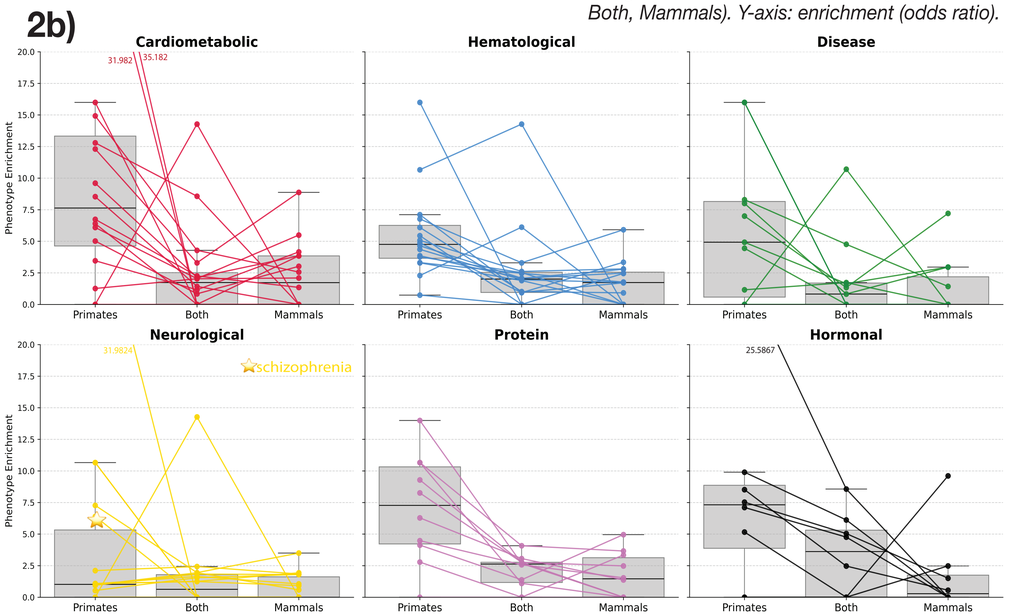
Figure 2.
(a) Number of Traits in Enrichment vs. Deficit

PIP≥0.5; GWAS-associated variants with significant enrichment is highest in primate-conserved regions, defecit in mammal-conserved and both.



(b) Evolutionary Conservation and GWAS Enrichment by Clinical Category

Lines show individual traits; boxplots show the dist. within each clinical and evolutionary category. X-axis: conservation category (Primates, Both, Mammals). Y-axis: enrichment (odds ratio).



Conclusions & Discussion

Key finding: Variants in constrained noncoding regions, especially *primate-conserved*, show the highest enrichment for GWAS associations.
Phenotype pattern: Enrichment clusters by clinical category.

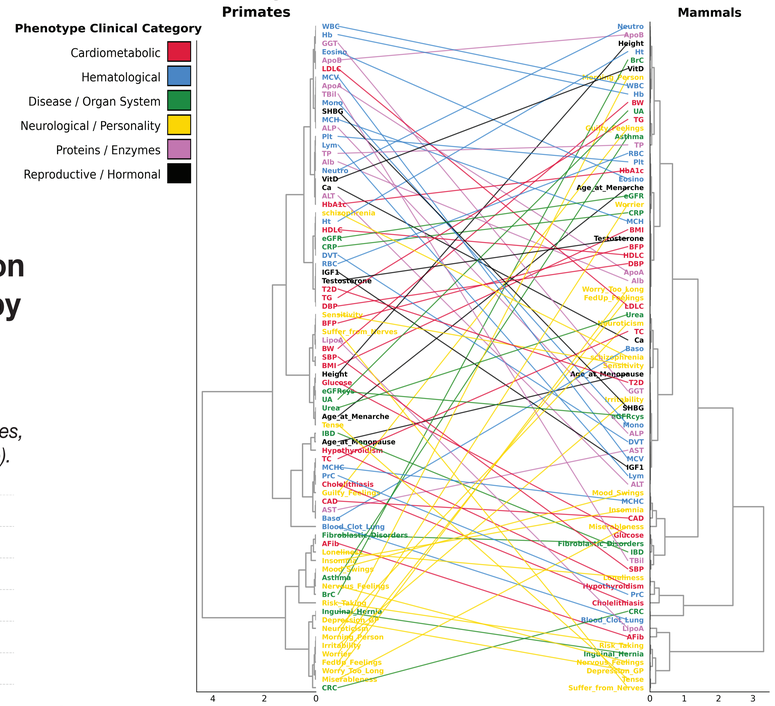
Impact:

- Provides **empirical trait-level evidence** that constraint in noncoding regions reflects functional importance and shapes the **polygenic architecture** of complex traits.
- Supports **purifying selection** as a major force acting on functional noncoding elements.
- Highlights **lineage-specific constraint,** where primate-only vs. mammal-only conservation differ in enrichment patterns.

Future Work:

- Apply pipeline to **PhyloP scores** for broader species coverage and finer evolutionary resolution.
- Expand to **more phenotypes** to test generalizability.
- Use **LD Score Regression** to estimate heritability by conservation category.
- **Weight counts by PIP** instead of thresholding to capture uncertainty.
- Report **statistical significance** (p-values, CIs) with odds ratios.

Figure 3.
Tanglegram of Trait Clustering: Primate vs. Mammal Conservation



Colored lines connect the same trait across dendrograms; colors indicate clinical category. Traits cluster more coherently in the primates (left) than mammals (right), suggesting primate-specific evolutionary constraint better captures functional relationships between traits.

Acknowledgement + References

Thank you to the Lohmueller Lab for this incredible opportunity. I am especially grateful to my direct mentor, Dr. Chenlu Di, for her patience, guidance, and mentorship, and to my PI, Dr. Kirk E. Lohmueller, for his support. I also thank the Garud Lab for welcoming me into their lab meetings and inspiring new ideas. Finally, thank you to my fellow B.I.G. Summer researchers for keeping me motivated both in and out of the lab! :D

Bruins In GeneKICKS (0-3-0)

[1] Di C, Ramesh S, Ernst J, Lohmueller KE. The landscape of fitness effects of putatively functional noncoding mutations in humans. bioRxiv. 2025. doi:10.1101/2025.05.14.654124
[2] Vu H, Ernst J. Universal annotation of the human genome through integration of over a thousand epigenomic datasets. Genome Biol. 2022;23:9. doi:10.1186/s13059-021-02572-z
[3] Trubetskoy V, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022
[4] Finucane HK, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet. 2015
[5] Kent WJ, et al. The human genome browser at UCSC. Genome Res. 2002