Qu, J., Liang, D., Zhao, T., Zeng, J., Cheng, H. (2026) Quantifying annotation-stratified pleiotropy and co-polygenicity between complex traits, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 24241: https://doi.org/10.31274/wcgalp.24241
©2026 *Qu, et al.* *CC BY-NC-ND*

Quantifying annotation-stratified pleiotropy and co-polygenicity between complex traits

J. Qu[1*], D. Liang[1], T. Zhao[2], J. Zeng[3], H. Cheng[1]

[1]University of California, Davis, Davis, CA, USA; [2]University of Nebraska–Lincoln, Lincoln, NE, USA; [3]University of Queensland, Brisbane, Queensland, Australia; [*]Corresponding author: jyqqu@ucdavis.edu

Introduction

Understanding the shared genetic architecture of complex traits is fundamental for advancing both precision medicine in humans and genetic selection in livestock. Genome-wide association studies (GWAS) have revealed extensive polygenicity and pleiotropy across human traits (Visscher et al., 2017; Solovieff et al., 2013), and large-scale genomic analyses in livestock are enabling parallel investigations of correlated traits of economic importance. Evidence from functional genomics further indicates that genetic contributions are not uniformly distributed across the genome, motivating methods that quantify enrichment of functional annotations (Finucane et al, 2015). However, analyses in animal populations must account for distinct population structures and longer-range linkage disequilibrium (LD) than in humans, which can affect inference on shared genetic variation. To address these challenges, we develop SBayesAPP, a unified Bayesian framework applicable to both humans and animals for quantifying the extent and sources of shared genetic variation between traits, with an explicit focus on annotation-specific shared genetic architecture.

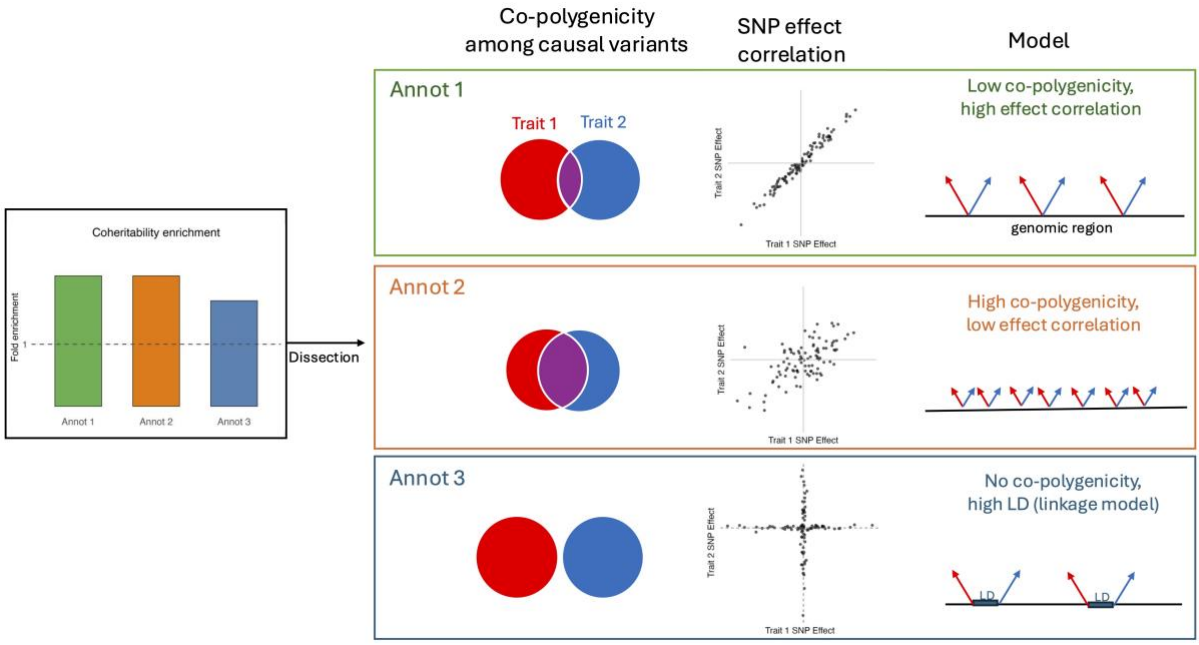

Figure 1. Illustration of three annotations showing enriched coheritability arising from different architectures. Enrichment is decomposed into co-polygenicity (proportion of shared causal variants) and effect-size correlation among shared variants.

A key motivation is that genome-wide genetic correlation can conflate different mechanisms of "genetic sharing", including true pleiotropy and LD-induced correlation, and can be near zero even when traits share causal variants with heterogeneous effect directions. SBayesAPP

integrates functional annotations and mixture priors to estimate annotation-specific coheritability (co-h^2) and to dissect it into co-polygenicity and effect-size correlation among shared variants (Figure 1), helping distinguish widespread polygenic sharing from more concentrated sharing driven by fewer, more concordant variants. In this study, we focus on human analyses to demonstrate the utility of the model and provide a proof-of-concept application to public cattle data to illustrate its potential for livestock genetics.

Methods

Model description. SBayesAPP models the bivariate SNP effect at locus j as an annotation-weighted sum of annotation-specific components:

$$\begin{bmatrix}\beta_{j1}\\ \beta_{j2}\end{bmatrix} = \sum_{c=1}^{C} A_{jc} \begin{bmatrix}\beta_{jc1}\\ \beta_{jc2}\end{bmatrix}$$

where $\boldsymbol{A}_{jc}$ is the annotation value for SNP j in annotation group c. For each annotation c, the component $\boldsymbol{\beta}_{jc} = (\beta_{jc1}, \beta_{jc2})^{\top}$ follows a mixture distribution:

$$\boldsymbol{\beta}_{jc} = \pi_{0c}\delta_0 + \pi_{1c}N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{c1}^2 & 0\\ 0 & 0\end{bmatrix}\right) + \pi_{2c}N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & \sigma_{c2}^2\end{bmatrix}\right) + \pi_{12c}N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_{c1}^2 & \sigma_{c12}\\ \sigma_{c12} & \sigma_{c2}^2\end{bmatrix}\right),$$

corresponding to null, trait-1 specific, trait-2 specific, and pleiotropic effects within annotation c. The annotation-specific mixture proportions $(\pi_{0c}, \pi_{1c}, \pi_{2c}, \pi_{12c})$ and (co)variance parameters $(\sigma_{c1}^2, \sigma_{c2}^2, \sigma_{c12})$ are estimated from the data during model fitting. From these, SBayesAPP reports annotation-stratified coheritability, co-polygenicity (i.e., π_{12c}), and effect-size correlation (i.e., $\rho_c = \frac{\sigma_{c12}}{\sigma_{c1}\sigma_{c2}}$). While sharing a common modeling foundation, SBayesAPP adopts species-specific implementations that reflect differences in data availability and genomic architecture. For human analyses, SBayesAPP takes GWAS summary statistics and an LD reference as input and is implemented in an LD-block–wise manner to leverage the hierarchical LD structure of the human genome. For livestock analyses, SBayesAPP is fit to individual-level genotype and phenotype data without LD-block partitioning, better accommodating the longer-range LD and population structure in animal populations.

Data description. For human analyses, ~1 million HapMap3 SNPs and cell-type genomic annotations constructed from The Tabula Sapiens Consortium, et al. (2022) were used, yielding 155 annotations plus a background group; the largest available GWAS summary statistics for type 2 diabetes (T2D) (Mahajan et al., 2022) together with 15 comorbidity traits were analyzed. For livestock analyses, the discovery population of 80,027 cattle from Reynolds et al. (2021), with provided individual-level genotypes for nine representative genomic regions (~1.5 Mb each), was used; after standard QC, 57,386 SNPs remained and were grouped into two categories, gene-functional regions (CDS + UTR + promoter, with promoter region defined as ±5 kb from the transcription start site, TSS) versus others, for bivariate analysis of body weight (BW) and body condition score (BCS). SNPs outside these nine regions were not included.

Results & Discussion

Human analyses. To validate that SBayesAPP can recover expected enrichment patterns, we defined three GWAS-significance–based SNP sets for T2D–fasting glucose and T2D–height: a significant set ($p \leq 5 \times 10^{-5}$ in both traits), a matched control set ($p > 0.5$ in both traits), and the remaining SNPs. As expected, Figure 2 shows strong enrichment confined to the significant set: coheritability enrichment was 702× for T2D–fasting glucose versus 113× for

T2D–height, while the rest/control sets were depleted. The decomposition highlighted different sharing mechanisms: T2D–fasting glucose exhibited stronger effect concordance (effect-size correlation = 0.87 in the significant set) than T2D–height (−0.23), and among trait-associated (non-null) SNPs, the proportion of pleiotropic SNPs was much higher for T2D–fasting glucose (84%) than for T2D–height (36%), consistent with their genome-wide correlations ($r_g = 0.43$ vs -0.04). Using 155 cell-type annotations across 15 T2D–trait pairs, Figure 3 shows that annotation-stratified coheritability enrichment concentrates in a subset of T2D-relevant tissues, with most other tissues depleted; trait pairs with stronger genome-wide genetic correlation tend to show stronger enrichment and higher co-polygenicity within enriched cell types, while weakly correlated pairs often display mixed annotation-level signals.

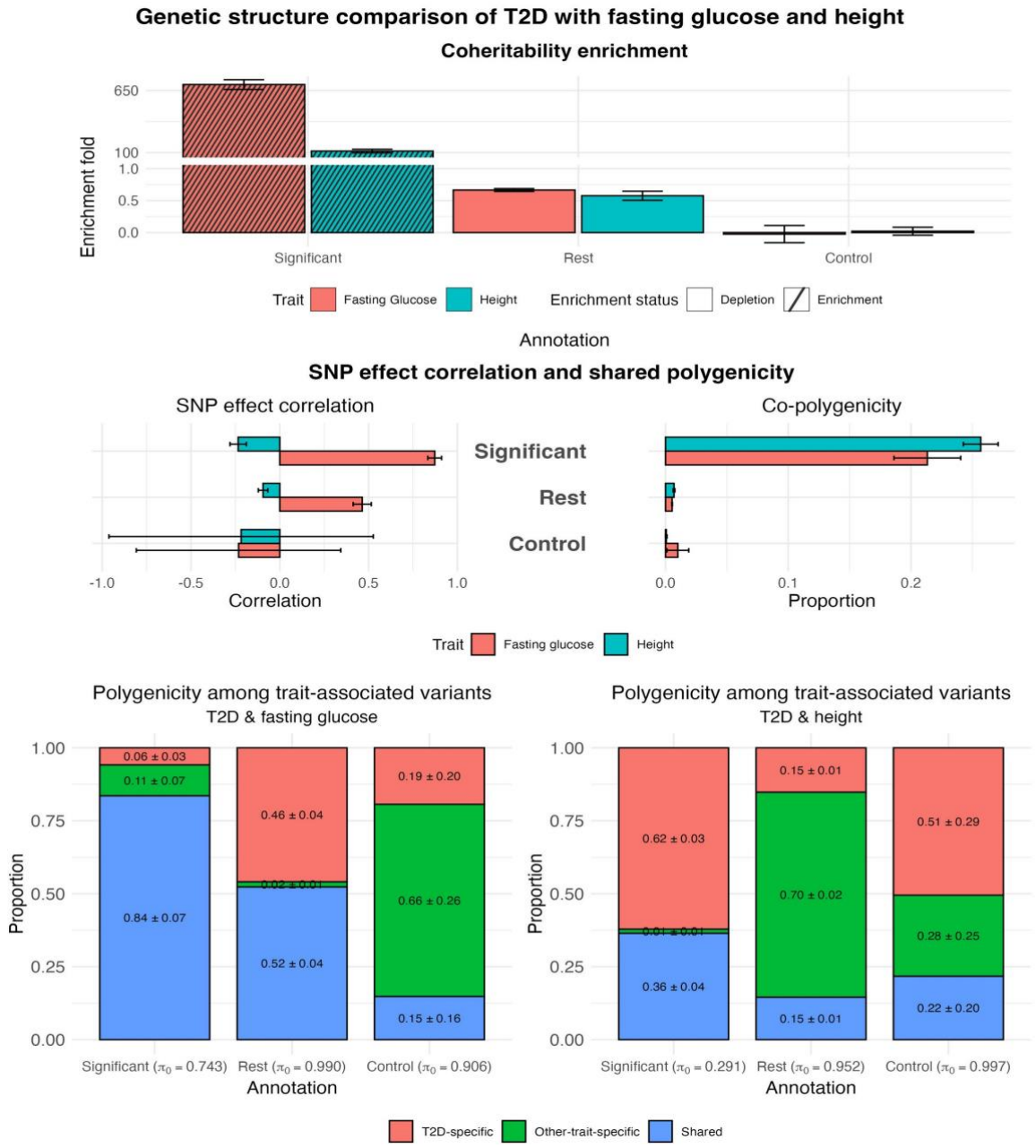

Figure 2. Genetic structure comparison of T2D with fasting glucose and height using GWAS-significance–based annotations. SBayesAPP results for three SNP sets defined by GWAS significance. Top: co-h^2 enrichment fold (error bars: posterior SD). Middle: decomposition into co-polygenicity and effect-size correlation. Bottom: trait-associated variant composition.

Cattle analyses. Using Ensembl gene annotations, we applied SBayesAPP to a public cattle data for BW and BCS. Because the released genotypes were limited to several genomic regions, the cattle analysis should be interpreted as a proof-of-concept based on the available regions rather than a complete characterization of genome-wide shared genetic architecture. We grouped SNPs into two functional categories: near-gene functional regions and other. Near-gene functional regions showed enrichment for both heritability (~1.14–1.91× higher than the other category) and coheritability (~2.14× higher) between BW and BCS. Consistently, co-polygenicity was higher in near-gene functional regions (4.6%) than in other regions (0.4%), suggesting that shared genetic effects on growth and body composition are more concentrated near genes. Although restricted genomic coverage limits the extent to which these local patterns can be generalized to the whole genome, the contrast between near-gene functional and other categories within the analyzed high-LD regions suggests that SBayesAPP can still detect annotation-specific enrichment in livestock.

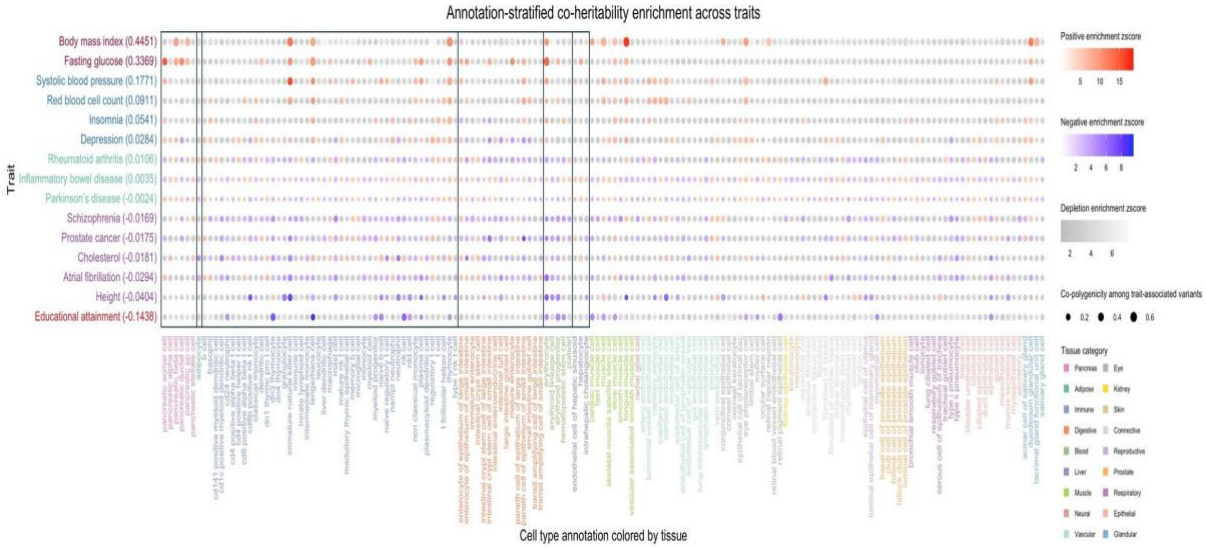

Figure 3. Annotation-stratified co-h^2 enrichment between T2D and 15 traits. Traits (y-axis) are ordered by genome-wide genetic correlation with T2D; 155 cell-type annotations (x-axis) are grouped into 18 tissue categories, with six T2D-relevant tissues outlined. Circle color shows the sign of genetic covariance and enrichment *z*-score (posterior mean/SD), and circle size indicates co-polygenicity among trait-associated variants. Gray indicates depleted annotations.

Conclusions

We developed SBayesAPP, a Bayesian framework for quantifying annotation-specific shared genetic architecture by estimating annotation-stratified coheritability, co-polygenicity, and effect-size correlation among shared variants. In human analyses, it recovers expected enrichment patterns under GWAS-significance SNP partitions and reveals structured cell-type enrichment across T2D and its comorbidity traits, illustrating that shared genetic architecture can vary substantially across annotations. A proof-of-concept cattle application further shows enrichment of coheritability and higher co-polygenicity in near-gene functional regions for BW and BCS, supporting SBayesAPP's potential for livestock genetics.

References

Finucane H.K., Bulik-Sullivan B., Gusev A., et al. (2015). Nat Genet 47(11):1228. doi.org/10.1038/ng.3404
Mahajan A., Spracklen C.N., Zhang W., et al. (2022). Nat Genet 54:560. doi.org/10.1038/s41588-022-01058-3
Reynolds E.G.M., Neeley C., Lopdell T.J., et al. (2021). Nat Genet 53:949 doi.org/10.1038/s41588-021-00872-5
Solovieff N., Cotsapas C., Lee P.H., et al, (2013). Nat Rev Genet 14(7):483-495. doi.org/10.1038/nrg3461
Tabula Sapiens Consortium et al. (2022). Science 376(6594):eabl4896. doi.org/10.1126/science.abl4896
Visscher P.M., Wray N.R., Zhang Q., et al. (2017). Am J Hum Genet 101:5. doi.org/10.1016/j.ajhg.2017.06.005