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

Jiayi Qu[4*], Tianjing Zhao[5*], Tian Lin[1], Ang Li[1], Shouye Liu[1], Solal Chauquet[1], Peter M. Visscher[1,2], Naomi R. Wray[1,3], Loic Yengo[1], Jian Zeng[1†], Hao Cheng[4,6,7†]

[1] Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia
[2] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[3] Department of Psychiatry, University of Oxford, Oxford, UK
[4] Department of Animal Science, University of California, Davis, Davis, CA, USA
[5] Department of Animal Science, University of Nebraska–Lincoln, Lincoln, NE, USA
[6] Graduate Group in Biostatistics, University of California, Davis, Davis, CA, USA
[7] Graduate Group in Integrative Genetics and Genomics, University of California, Davis, Davis, CA, USA

Correspondence: Jian Zeng <j.zeng@uq.edu.au> and Hao Cheng <qtlcheng @ucdavis.edu>
* These authors contributed equally to this work.
† These authors jointly supervised this work.

Abstract

Understanding shared genetic architecture is essential to interpreting disease comorbidities and trait correlations. We introduce SBayesAPP, a Bayesian model that integrates GWAS summary statistics with functional annotations to jointly estimate annotation-stratified SNP effect-size correlation and pleiotropic variant proportion (co-polygenicity) between traits, dissecting genetic correlation and coheritability enrichment across annotations. Simulations and real data analyses show improved accuracy and interpretability over existing methods. In type 2 diabetes analyses with 15 traits, SBayesAPP reveals clear tissue- and cell-type-specific enrichment and distinguishes mechanisms driven by few large-effect variants versus many modest-effect variants. The analysis of smoking and lung cancer prioritizes lung and immune cells, and identifies cell-type-specific genetic correlations driven by either pleiotropic or lung-cancer-specific variants, consistent with a causal relationship model. For schizophrenia and educational attainment, despite near-zero genome-wide genetic correlation, cell-type-specific correlations range from -0.20 to 0.21, with strong (co)heritability enrichment and high co-polygenicity found in dopaminergic neurons and oligodendrocytes. These results highlight the ability of SBayesAPP to resolve annotation-specific genetic sharing and uncover biological mechanisms across complex traits.