Submitter Name: Jack Pattee
Submitter email: jack.pattee@cuanschutz.edu
PI Name (if different): Laura Saba
PI email (if different): laura.saba@cuanschutz.edu

Modeling recommendations and simulated power for the analysis of expression quantitative trait loci in the Hybrid Rat Diversity Panel.

Jack Pattee[1], Lauren Vanderlinden[2], and Laura Saba[3]

[1]Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; [2]Department of Epidemiology, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; [3]Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

The Hybrid Rat Diversity Panel (HRDP) is a panel of 99 inbred rat strains comprised of two recombinant inbred panels and a group of divergent classic inbred strains. The HRDP was designed to maximize power and resolution for investigating the genetic basis of complex traits, and in particular of quantitative trait loci (QTL). This paper examines methodological approaches for the analysis of expression (e)QTL traits in the HRDP for a subset of 43 strains. Statistical analyses must account for the high degree of relatedness among the HRDP subjects, the nuance of estimating relatedness structure from genomic data, and the relatively small sample size. We characterized our approaches in terms of type I error and power, both in the real data and in simulation. We conducted a full GWAS for all 99,376 brain gene expression traits and demonstrate that simple linear regression and by-panel meta-analysis are suboptimal. Further comparison of the real data results and extensive simulation identify a set of plausible approaches, which can be characterized as permutations of a mixed model approach implemented via the GEMMA software. Which among these plausible approaches is recommended depends on the nuance of an analysis, as they exhibit tradeoffs in power and false positive rate. We generate simulated power curves for these plausible approaches applied to the HRDP data under a variety of underlying genetic architectures, demonstrating their good power and aiding study design for the investigation of new variants. Supported by NIDA (P30 DA044223) and NIAAA (R24 AA013162).