October XXX, 2024

Dear Editors,

My co-authors and I are writing to submit the manuscript entitled “Transcripts with high distal heritability mediate genetic effects on complex metabolic traits”, as a research article for *Nature Genetics.*

This work combines two multi-tissue, clinically relevant data sets from independent, genetically diverse mouse populations with a novel high-dimensional mediation analysis (HDMA) to investigate the relationship between expression quantitative trait loci (eQTLs) and the heritability of complex traits. We provide experimental validation for a conversation about the heritability of complex traits that has been ongoing in *Nature Genetics* and other top journals in recent years.

Evidence from GWAS suggests that genetic variants influence complex traits through gene expression regulation, however identifying disease-associated eQTLs has had minimal success.

Recent studies published in *Nature Genetics* have explored this shortcoming by critically examining the relationship between eQTLs and complex traits. Notable work has demonstrated that trait-associated variants and eQTLs tend not to co-localize (Mostavi *et al.,* 2023), and that only small amounts of trait heritability are mediated through local eQTLs (Yao *et al.* 2020). Evidence suggests, rather, that distal eQTLs may be more relevant to complex traits than local eQTLs (Vosa *et al.* 2021). As these studies were done humans, they are limited in their ability to comment on causal relationships between eQTLs and complex traits. To address this limitation, we developed two large complementary data sets in genetically diverse mice to directly assess the role of local and distal eQTL in driving complex trait variation.

The two mouse populations were derived from the Diversity Outbred (DO) mice, and the Collaborative Cross (CC) mice. Both were maintained on high-fat, high-sugar diets to model diet-induced obesity and metabolic disease. The two populations share ancestral haplotypes but have independent population structure. Thus, local genetic effects on gene expression are identical across the two populations, but distal effects are independent allowing us to differentiate the effects of local and distal eQTLs. Further, we measured genome-wide genotypes, clinically relevant phenotypes, and gene expression in four disease-relevant tissues (adipose, pancreatic islet, liver, and skeletal muscle) across hundreds of individuals. Such data collection is highly impractical in human subjects.

Initial eQTL analysis showed that transcripts confirmed what has been shown in human studies, notably that transcripts with high local heritability tended to have low trait relevance, and that transcripts with higher trait relevance tended to have high distal heritability. This finding supports earlier findings in humans that complex trait heritability is mediated primarily through distal gene regulation. We then used a novel HDMA to directly identify composite transcripts that mediated the effects of genetic background on complex metabolic traits in the DO mice. The transcripts contributing most to the composite transcripts again tended to have high distal, and low local heritability. They were also enriched in the literature as having known connections to obesity and metabolic disease. We showed that the composite transcripts we identified were highly biologically interpretable in a tissue-specific manner and highlighted known biology of metabolic disease at multiple levels of organization, from individual transcripts to cell type composition. To further test the contributions of local and distal eQTLs in driving complex trait variation, we used the composite transcripts to predict phenotypic outcomes in an independent population of mice derived from CC recombinant inbred crosses (CC-RIX). We used the composite transcripts identified in the DO mice to generate weighted gene expression vectors in the CC-RIX. When measured transcripts were used, which include both locally and distally determined components of gene expression, the weighted vectors were highly predictive of obesity in the CC-RIX. However, when only the locally determined component of gene expression was used, the prediction failed completely. This finding offers experimental confirmation that the distal component of gene regulation is highly relevant to complex traits, whereas local gene regulation contributes only minimally.

These findings have profound implications for the conversation

PMIDs: 32424349

To further investigate

Significantly...

With this manuscript, we provide a unique data set consisting of... gene expression, and multiple epigenomic modifications in hepatocytes across nine commonly used inbred strains of mouse. The raw and processed data are hosted on Gene Expression Omnibus (accession: GSE213968). These data provide an important resource that will aid in the investigation of genetically determined epigenomic variation, and the functional annotation of the mouse genome.

We believe that this manuscript will be of interest to geneticists working in model organism populations. The manuscript includes main text, seven figures, and five supplementary figures. No author has any financial, personal, or professional interests that could be construed to have influenced the paper. Thank you for your consideration of this manuscript.

Sincerely,