## ABSTRACT

DECIPHERING THE GENETIC BASIS FOR COMPLEX TRAIT VARIATION: UTILIZING ALTERNATIVE GENOME-WIDE ASSOCIATION METRICS AND MOLECULAR PHENOTYPES

By

Scott A Funkhouser

Within any population, complex trait variation can be attributed to an impressive number of genetic factors. Identification of such factors has been made possible, in part, by large biomedical datasets comprised of genotypes and phenotypes for hundreds of thousands of individuals. Furthermore, understanding the biological mechanisms through which genetic variation creates complex trait variation has been facilitated by high-throughput sequencing technology, used to quantify molecular, intermediate phenotypes. Despite such datasets being widely available, we lack understanding of the full spectrum of genetic effects, including gene-by-sex (G×S) interactions. We also have yet to uncover various molecular phenotypes that may "link" genetic variation to complex trait variation. To address these gaps in knowledge, the following chapters will 1) develop and utilize statistical methodology for mapping G×S interactions among human traits, and 2) utilize a pig model to characterize RNA editing—a relatively understudied form of transcriptional regulation—and evaluate its potential to link genetic variation with complex trait variation.

Growing evidence from genome-wide parameter estimates suggest males and females from human populations possess differing genetic architectures. Despite this, mapping G×S interactions remains challenging, suggesting that the magnitude of a typical G×S interaction is exceedingly small. We have developed a local Bayesian regression (LBR) approach to estimate sex-specific single nucleotide polymorphism (SNP) marker effects after fully accounting for local linkage-disequilibrium (LD) patterns. This provided means to infer G×S interactions either at the SNP level, or by aggregating multiple sex-specific SNP effects to make inferences at the level of small, LD-based regions. In simulations, LBR provided greater power and resolution to detect G×S interactions than the traditional approach to genome-wide association (GWA), single-marker regression (SMR).

When using LBR to analyze human traits from the UK Biobank (N ~ 250,000) including height, BMI, bone-mineral density, and waist-to-hip ratio, we find evidence of novel G×S interactions where sex-specific effects explain a very small proportion of phenotypic variance ($R^2 < 1x10^{-4}$) but are enriched in expression quantitative trait loci (eQTL). By leveraging large datasets and powerful metrics, we are providing evidence that G×S interactions may influence phenotypic variance for a variety of human complex traits.

Adenosine to inosine (A-to-I) RNA editing impacts gene function by converting adenosine to inosine molecules within specific regions of the transcriptome and is catalyzed by adenosine deaminase acting on RNA (ADAR). High-throughput sequencing studies, most of which utilizing human models, have found thousands of A-to-I edited loci commonly located within repetitive elements such as the primate-specific Alu element. Here, we utilized matched whole-genome sequencing and RNA sequencing from the same animal to demonstrate that widespread RNA editing occurs within pig transcriptomes, largely within pig-specific repetitive elements known as PRE-1.

The degree that sites in the transcriptome are edited by ADAR—the "editing level"—has been observed to vary within populations but it is largely unknown how genetic variation as whole influences editing level variation. Using 168 $F_2$ pigs with SNP genotyping data and RNA sequencing from skeletal muscle, we identified five RNA editing sites across four genes whose editing level variation was significantly attributed to the additive effects of all observed SNP markers (estimated genomic heritability $\hat{h}^2_g$ = 0.31–0.56; $p$-value = $8.2x10^{-5}$–$8.8x10^{-4}$). We then used bivariate models to estimate how genetics influences covariance between site-specific RNA editing levels and complex traits in pigs. We found modest evidence that SNPs near *ADAR* contribute to covariance in RNA editing activity and numerous growth traits such as average daily gain (local genetic correlation $\hat{\rho}_{g_{local}}$ [SE] = -0.87 [0.16]; $p$-value = 0.029). These results suggest potential pleiotropic effects between RNA editing activity and complex traits and encourages further use of multi-variate mixed models determine if RNA editing can "link" genetic variation with complex trait variation.