

Genetic Architecture Modulates Diet-Induced Hepatic mRNA and miRNA Expression Profiles in Diversity Outbred Mice

Excel Que,^{*,†,1} Kristen L. James,^{†,1} Alisha R. Coffey,[†] Tangi L. Smallwood,[‡] Jody Albright,[§]
M. Nazmul Huda,^{*,†} Daniel Pomp,^{**} Praveen Sethupathy,^{††,2} and Brian J. Bennett^{*,†,2}

^{*}Western Human Nutrition Research Center, Agricultural Research Service, US Department of Agriculture, Davis, California 95616,

[†]Department of Nutrition, University of California, Davis, California 95616, [‡]Curriculum in Genetics and Molecular Biology, and

^{**}Department of Genetics, University of North Carolina at Chapel Hill, North Carolina 27599, [§]Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, North Carolina, 28081 and ^{††}Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853

ORCID IDs: 0000-0002-3704-3560 (K.L.J.); 0000-0003-2775-4470 (M.N.H.); 0000-0001-6609-9195 (D.P.); 0000-0002-7983-239X (P.S.); 0000-0002-0766-3195 (B.J.B.)

ABSTRACT Genetic approaches in model organisms have consistently demonstrated that molecular traits such as gene expression are under genetic regulation, similar to clinical traits. The resulting expression quantitative trait loci (eQTL) have revolutionized our understanding of genetic regulation and identified numerous candidate genes for clinically relevant traits. More recently, these analyses have been extended to other molecular traits such as protein abundance, metabolite levels, and miRNA expression. Here, we performed global hepatic eQTL and microRNA expression quantitative trait loci (mirQTL) analysis in a population of Diversity Outbred mice fed two different diets. We identified several key features of eQTL and mirQTL, namely differences in the mode of genetic regulation (*cis* or *trans*) between mRNA and miRNA. Approximately 50% of mirQTL are regulated by a *trans*-acting factor, compared to ~25% of eQTL. We note differences in the heritability of mRNA and miRNA expression and variance explained by each eQTL or mirQTL. In general, *cis*-acting variants affecting mRNA or miRNA expression explain more phenotypic variance than *trans*-acting variants. Lastly, we investigated the effect of diet on the genetic architecture of eQTL and mirQTL, highlighting the critical effects of environment on both eQTL and mirQTL. Overall, these data underscore the complex genetic regulation of two well-characterized RNA classes (mRNA and miRNA) that have critical roles in the regulation of clinical traits and disease susceptibility.

KEYWORDS Quantitative trait loci; Multiparental models; eQTL; mirQTL

THE advent of genome-wide investigation of DNA variants and gene expression have revolutionized our understanding of biology. Systems genetic approaches often utilize

variation in DNA and RNA to identify genes and pathways associated with clinical traits (Rockman and Kruglyak 2006). These approaches have been used in studies of plants (Fu *et al.* 2009), flies (Ayroles *et al.* 2009), yeast (Storey *et al.* 2005), mice (Doss *et al.* 2005), and humans (Schadt *et al.* 2003, 2008). Analyzing variation in gene expression in a segregating population for genetic regulation is a critical aspect of systems genetics. Originally, genomic positions that were found to regulate quantitative traits were deemed quantitative trait loci (QTL); correspondingly, loci that regulate the messenger RNA (mRNA) transcript levels are called expression quantitative trait loci (eQTL). QTL and eQTL analyses have identified numerous candidate genes for obesity, diabetes, and cardiovascular disease, indicating that certain genetic variants interact critically with environmental factors that predispose an organism to disease.

Copyright © 2020 Que *et al.*

doi: <https://doi.org/10.1534/genetics.120.303481>

Manuscript received November 22, 2019; accepted for publication July 27, 2020; published Early Online August 6, 2020.

Available freely online through the author-supported open access option.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material available at figshare: <https://doi.org/10.25386/genetics.12597794>.

¹These authors contributed equally to this work.

²Corresponding authors: Cornell University College of Veterinary Medicine, T7 006D Veterinary Research Tower, Box 17, Ithaca, NY 14853-6401. E-mail: praveens@cornell.edu; and Obesity and Metabolism Research Unit, US Department of Agriculture Western Human Nutrition Research Center, 430 West Health Sciences Drive, Davis, CA 95616-5270. E-mail: brian.bennett@usda.gov

The initial studies investigating eQTL were performed in model organisms and the identified eQTL were split into two distinct classes based on the location of the associated single nucleotide polymorphism (SNP). These classes were defined as *cis*-eQTL and *trans*-eQTL based on whether the SNP resided close to or far from the regulated mRNA's gene of origin. In the case of a *cis*-eQTL, a genomic variant located in proximity to the gene in question is associated with the gene's expression. The associated SNP is often thought to be in linkage disequilibrium (LD) with the functional genomic variant affecting gene expression. For *cis*-eQTL a number of possibilities exist, such as promoter and enhancer variants. *Trans*-eQTL refer to SNPs associated with gene expression that are located distal to the locus containing the gene in question. *Trans*-eQTL can be located on a different chromosome or at a distal location on the same chromosome, and suggest an alternative regulatory mechanism. More recently, many groups have characterized eQTL and the genetic architecture of gene expression in humans. Several important findings have emerged, including the discovery that *cis*-eQTL, sometimes referred to as "local eQTL," are pervasive and often enriched for known genome-wide association study (GWAS) loci (Civelek *et al.* 2017). *Trans*-eQTL, also referred to as "distant eQTL," have been found to affect hundreds of genes, are often regulated by the same transcription factors (Brynedal *et al.* 2017), and at times may be transcription factors themselves (Albert *et al.* 2018). Additionally, there may be complex relationships between *cis*-acting and *trans*-acting eQTL, as there is evidence that SNPs regulating genes as *cis*-eQTL can also mediate the effects of *trans*-eQTL (Yao *et al.* 2017).

eQTL studies have also begun to investigate the genetic regulation of noncoding RNA. MicroRNAs (miRNAs) are noncoding RNAs that regulate gene expression at the post-transcriptional level and have been implicated in a range of diseases, including cardiovascular diseases and metabolic syndromes (Rottiers and Naar 2012). The identification of microRNA expression QTL (mirQTL) has complemented GWASs and further emphasizes the complex genetic regulation of both miRNA and mRNA. For example, recent data suggests that miRNA are associated with cardiometabolic traits, and Mendelian randomization analysis has demonstrated that these miRNA associations may be causal (Nikpay *et al.* 2019). Determining the underlying mechanisms of eQTL and mirQTL, the interaction of these two classes of RNA, and how eQTL and mirQTL respond to environmental perturbation remains to be fully elucidated.

Human GWAS have been impressive in their ability to associate variants with disease traits, but are limited in their ability to quantify and control environmental effects, their ability to probe tissues, and by the sample sizes needed for robust power. To address these limitations, model organisms can be used to identify eQTL and mirQTL in multiple tissues, or under a variety of conditions, such as diet, associating the results with clinical traits and disease susceptibility (Mehrabian *et al.* 2005; Bhasin *et al.* 2008; Yang *et al.* 2009).

Recently, genetic mapping panels have been developed which incorporate variation from more than two parental strains. Using complex breeding strategies, these multiparent advanced generation intercross panels have high genetic resolution and have been employed in a number of model organisms, including *Arabidopsis* (Kover *et al.* 2009), *Drosophila* (Mackay *et al.* 2012), and mice (Aylor *et al.* 2011). In mice, two related multiparent advanced generation intercross populations exist, the Collaborative Cross (CC) and the Diversity Outbred (DO), and both are generated from eight mouse strains, of which five are classical inbred (A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, and NZO/HILtJ) and three are wild-derived inbred (CAST/EiJ, PWK/PhJ, and WSB/EiJ) strains. In comparison to traditional approaches in mice, the inclusion of the wild-derived strains increases genetic diversity and reduces identical-by-descent "blind spots" (Yang *et al.* 2007). The primary difference between CC and DO populations is that the CC is a recombinant inbred panel while the DO is maintained as an outbred population using a randomized breeding scheme. The DO contains ~45 million SNPs (Yang *et al.* 2011; Churchill *et al.* 2012) and displays a wide variation in phenotypes mimicking the variation observed in humans. Because of the DO population's allelic diversity, the DO have been utilized in high-resolution genetic mapping of complex traits, including atherosclerosis (Smallwood *et al.* 2014), response to toxicants (French *et al.* 2015), the microbiome (Kemis *et al.* 2019), and diabetes (Keller *et al.* 2018), as well as studies focused on gene expression (Munger *et al.* 2014; Tyler *et al.* 2017), metabolomics, and proteomics (Chick *et al.* 2016). Collectively, these studies utilizing DO mice have provided additional insight into the complex genetic architecture that underlies clinical traits and disease susceptibility (Svenson *et al.* 2012; Chesler *et al.* 2016; Winter *et al.* 2017).

Little is known about the genetic regulation of hepatic mRNA and miRNA under different dietary conditions; exploring these trends in the liver is highly relevant due to the liver's roles in cardiometabolic diseases. We have previously identified novel interactions between eQTL and clinical traits in the DO, including atherosclerosis (Smallwood *et al.* 2014), the atherosclerosis-associated plasma metabolite trimethylamine N-oxide (Coffey *et al.* 2019), and plasma cystatin C (Huda *et al.* 2020). In this study, we investigate the global regulation of hepatic mRNA and miRNA expression in DO mice fed two different diets: a high-fat cholic acid (HFCA) diet designed to trigger atherosclerosis, and a high-protein (HP) diet. We identify eQTL and mirQTL, classify them as *cis* or *trans*, and characterize the QTL by defining heritability, genetic resolution, and effect size in multiple QTL models. Finally, we assess the effect of diet on each of these components and identify eQTL driven by different founder alleles as a result of diet.

Materials and Methods

Experimental animals and diets

Details of the mouse experiments have been reported previously (Smallwood *et al.* 2014; Coffey *et al.* 2017, 2019). In

brief, 292 female DO mice (J:DO, Jackson Laboratory stock number 009376, outbreeding generation 11) were obtained from the Jackson Laboratory (Bar Harbor, ME) as 146 full sibling pairs at 4 weeks of age. Mice were housed in groups of five per cage in a HEPA-filtered, climate-controlled, facility under a 12-hr light-dark cycle and provided with nonirradiated pine bedding and free access to sterile water. Mice were maintained on defined synthetic diet, AIN-76A, until 6 weeks of age, to control for variability in the components of standard chow (D10001; Research Diets, New Brunswick, NJ). Afterward, one sibling from each of the 146 sibling pairs was randomly assigned to one of the diets for a total of 18 weeks (Figure 1). Thus 146 mice were transferred to a synthetic HFCA diet composed of 20.0% fat, 1.25% cholesterol, and 0.5% cholic acid, to induce atherosclerotic lesions, while the remaining 146 mice were transferred to a nonatherogenic HP diet composed of 5.0% fat and 20.3% protein (D12109C and D12083101, respectively; Research Diets). All procedures were approved by the Institutional Animal Care and Use Committee at University of North Carolina at Chapel Hill (protocol number 11-299).

RNA isolation

Livers were flash-frozen in liquid nitrogen and subsequently stored at -80° until total RNA was isolated using Norgen Total RNA Purification Kit (Norgen, ON, Canada). RNA integrity was determined by Bioanalyzer (Bio-Rad, Hercules, CA) and high-quality RNA (with RNA Quality Index > 7.5) isolated from livers of 268 of the 292 DO mice was processed and hybridized to Affymetrix Mouse Gene 2.1 ST 96-Array Plate (Thermo Fisher Scientific, Waltham, MA) using the GeneTitan Affymetrix instrument, according to the standard manufacturer's protocol. All probes containing known SNPs from the eight founder inbred mouse strains of the DO mouse population were masked (165,204 probes) during normalization by downloading the SNPs from the Sanger sequencing website (<http://www.sanger.ac.uk/science/data/mouse-genomes-project>) and overlapping them with probe sequences. To ensure integrity of downstream qualitative analyses, annotation data from the Affymetrix mogene 21 annotation database, Bioconductor version, release 3.7 (<http://www.bioconductor.org/packages/release/data/annotation/html/mogene21sttranscriptcluster.db.html>) was used to filter the remaining transcript cluster IDs for those with reliable and complete annotations. We removed transcript cluster IDs identified as cross-hybridizing ($n = 4954$), associated with unlocalized sequences (e.g., chr1_GL456210_random), residing on the Y chromosome or mitochondria ($n = 69$), or without ENSEMBL or ENTREZ annotations ($n = 4766$). The total number of unique probes postfilter was 24,004, corresponding to 23,626 genes. All transcript cluster IDs were validated by programmatically querying the ENTREZ IDs against the NCBI Gene Database for chromosome and position. Microarray data are available on the Gene Expression Omnibus repository under accession number GSE99561.

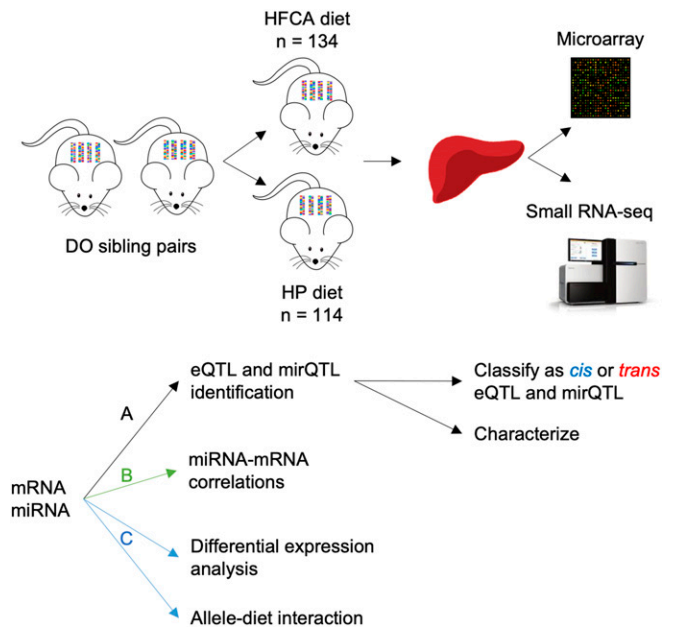


Figure 1 Experimental and analytical design classifies, characterizes, and explores the relationships between hepatic mRNA and miRNA in the context of diet. A total of 292 female DO mice were obtained from Jackson Laboratory as 146 full sibling pairs. After initial maintenance on a synthetic diet (AIN-76A), one sibling from each pair was randomly assigned either an HFCA or HP diet for 18 weeks. Extracted hepatic mRNA was hybridized to the Affymetrix Mouse Gene array and miRNA was sequenced via Illumina HiSeq. mRNA and miRNA expression values were used as phenotypes in four QTL genome scans. A: The four models were diet as an additive covariate ($n = 248$), diet as an interactive covariate ($n = 248$), subset of HFCA-fed mice ($n = 134$), and subset of HP-fed mice ($n = 114$). eQTL and mirQTL found in each model were classified as *cis* or *trans* and characterized. B: Interactions between miRNA and mRNA were assessed and compared between models. C: The role of diet was explicitly analyzed by differential expression analysis and by identifying eQTL with significant allele-diet interactions from regression analysis.

Small RNA sequencing

High-quality RNA from livers of the DO mice was also used for small RNA sequencing (smRNA-seq). Libraries were created using NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biosciences), and 50-bp single-read sequencing was carried out on the Illumina HiSeq platform, resulting in an average of over 16 million reads per sample. miRquant 2.0 (Kanke *et al.* 2016) was used to trim off adapter sequences, align reads to the mouse genome, and quantify miRNAs and their isoforms (termed isomiRs). A previous study in mice from the CC mouse panel has shown that miRNAs do not contain variants across founder strains within their seed regions, so reads were aligned to the mm9 mouse genome (Rutledge *et al.* 2015). Reads were normalized to reads per millions mapped to miRNAs (RPMs). An expression threshold of at least 50 RPMs in about one-quarter of all samples was set to filter out the lowly expressed miRNAs, which resulted in a set of 246 robustly expressed miRNAs.

Genotyping

DNA was extracted and purified from tail biopsies taken from 6-week-old mice using the QIAGEN DNeasy kit, following the manufacturer's instructions. Genotyping was performed using the Mega Mouse Universal Genotyping Array (MegaMUGA) by GeneSeek (Neogen, Lincoln, NE) (Svenson *et al.* 2012).

QTL mapping

QTL mapping was performed using the R (v3.5.1) package QTL2 (v0.20). The genotypes from the MegaMUGA array are assigned DO founder strain probabilities using a hidden Markov model (Broman and Sen 2009) and haplotypes are defined as previously reported (Broman 2012a,b). The genotype probabilities are reduced to the eight founder allele probabilities and used to generate a kinship matrix using the "leave-one-chromosome-out" method to reduce bias from same chromosome SNPs (Yang *et al.* 2014). A genome scan is performed using a linear mixed model, which regresses the microarray gene expression phenotype matrix against the allele probabilities at each marker, using the kinship matrix to control for population structure.

Using the scan1 function in R/qtl2, four separate genome scans were performed to assess how genotype, diet, and gene-by-diet interactions affect eQTL. The four QTL models are as follows: diet as an additive covariate, diet as an interactive covariate, subset of mice fed an HFCA diet ($n = 134$), and subset of mice fed an HP diet ($n = 114$). The phenotypes for the genome scans were the individual expression data for each transcript cluster ID. To assess similar effects on mirQTL, the same models were performed on the miRNA abundance data. The log of the odds ratio (LOD), which describes the log-scaled likelihood difference of the full and null genome scan models, is generated for each SNP on the MegaMUGA ($n = 70,339$) genotyping array. Permutation analysis was used for subsequent filtering (described below), reducing the initial output to a set of high-confidence QTL. Confidence intervals for QTL were calculated by determining the upper and lower physical distances from the peak where the LOD dropped by 1.8 or more, approximating a 95% coverage of true positives for an intercross population (Manichaikul *et al.* 2006). We defined eQTL as *cis*-eQTL if the marker with the maximum LOD score was within 4 Mb of the transcription start site (TSS), and called the remaining eQTL *trans*-eQTL (Keller *et al.* 2018).

Permutation analysis

The significance threshold at $P < 0.05$ of each QTL was empirically determined via permutation analysis to minimize type 1 and type 2 errors (Huda *et al.* 2020). The rows of the genotype data were randomized with respect to the expression data (miRNA or mRNA) and a maximum LOD score was produced. This process was repeated 1000 times to create a null distribution of maximum LOD scores for a particular expression trait (Churchill and Doerge 1994). The 95th quantile of a

phenotype's null distribution was considered as its individual significance threshold at $P = 0.05$. For our analysis with diet as an additive covariate, all transcripts ($n = 24,004$) were subjected to full permutation analysis.

A modified approach, designed to reduce the computational complexity of performing 1000 QTL permutations analysis on 24,004 transcripts three times, was utilized to identify eQTL for the interactive and diet-specific subsets. After the initial eQTL model was fit for each transcript, we completed 50 permutations to generate an initial null distribution for each candidate, calculating a conservative threshold by taking the 90th quantile and subtracting the quantile SE (Cox and Hinkley 1979). The transcripts with LOD scores above this threshold (90th quantile minus the quantile SE) were subjected to full 1000 permutation testing. Transcripts with LOD scores above the 95% significance threshold, as determined by 1000 permutations, were considered significant.

We initially validated this method with diet as an additive covariate, which had 1000 permutations performed for each transcript cluster ID. Using the 50-permutation method, we found 8867 transcripts to be above our modified threshold. We subjected these 8867 transcripts to permutation testing (1000 permutations) and identified the same transcripts with significant eQTL ($n = 6276$) as our initial test of all 24,004 transcripts.

Genome scan coefficients

The contribution of each founder strain genotype at each QTL was determined using the best linear unbiased predictor (BLUP) for each QTL model (additive, HFCA diet, and HP diet) using a similar mixed effect model that treated the allele probabilities as random effects. BLUP scans show contributions from the eight DO founder strains at each SNP and are extracted from the peak SNP of each significant eQTL or mirQTL.

Phenotypic variance explained

For each significant eQTL or mirQTL, the RNA expression and genotype probabilities at the peak SNP were extracted and fit to full and null Haley–Knott regression models. The differences in R^2 values between the full model, with genotypes and covariates included, and null model with only covariates included, were taken as the phenotypic variance explained by the QTL.

Heritability

To determine the extent to which phenotypic variation is influenced by genotypic variation, we used a linear mixed effect model to estimate narrow-sense heritability (h^2) scores of transcript cluster IDs or miRNAs. This was performed using the function `est_herit` in R/qtl2, submitting a single, square, kinship matrix and the expression value of individual transcript cluster IDs or miRNA with eQTL or mirQTL.

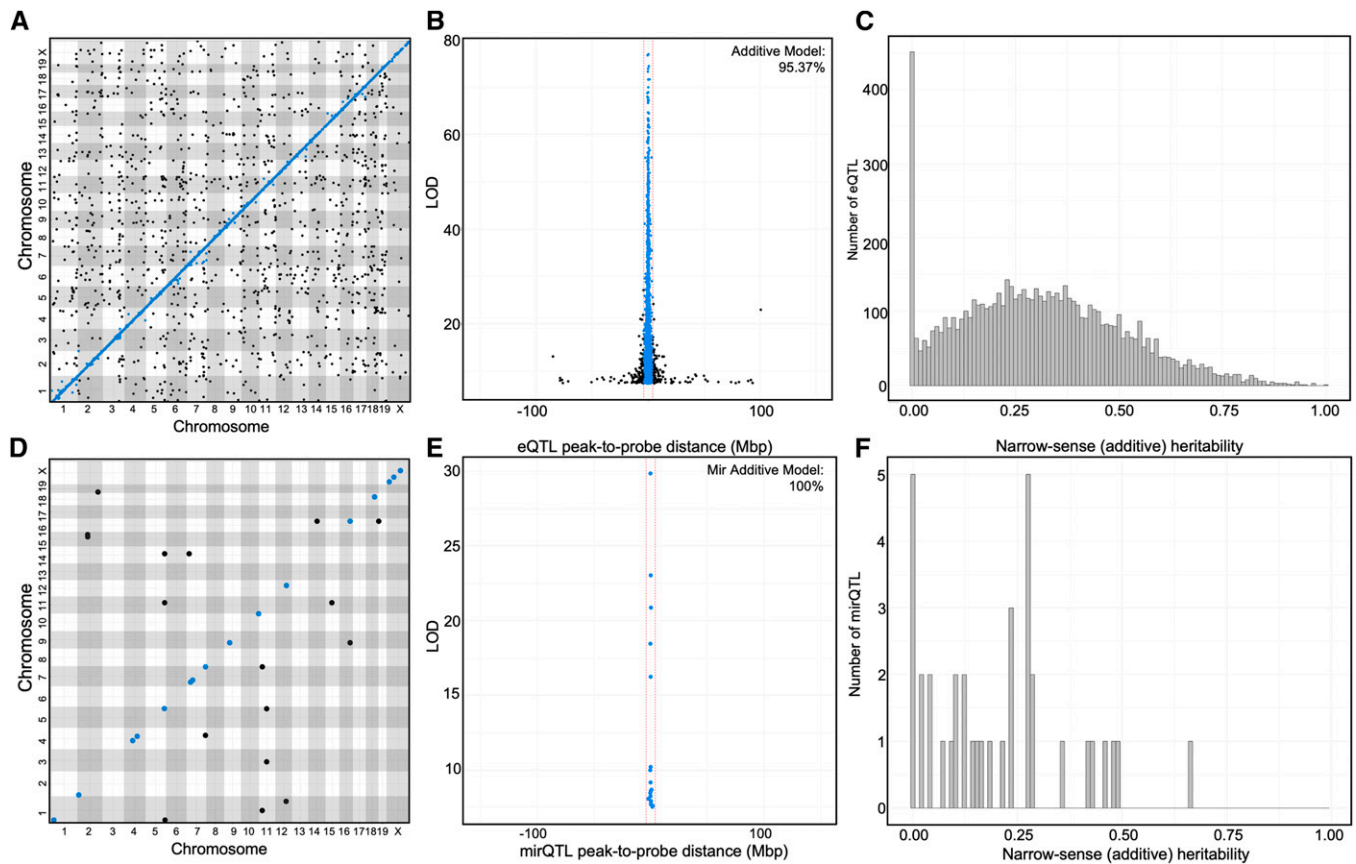


Figure 2 High-resolution hepatic eQTL and mirQTL in Diversity Outbred mice demonstrate complex regulation of expression traits. Results from expression quantitative trait loci (eQTL) analyses using a genome scan model treating the high-fat cholic acid (HFCA) diet and high-protein (HP) diet groups as an additive covariate. (A) An eQTL plot visualizes all significant associations between gene expression and structural SNP variants from a genome scan. The absolute genomic positions of the SNP and gene transcript probe set in the mouse genome are shown on the x- and y-axes, respectively. The blue line represents *cis*-eQTL, positions where gene expression variance is associated with a proximal (± 4 Mb) SNP variant. (B) Resolution of *cis*-eQTL is determined by the distance from the SNP to the transcript probe set (x-axis); resolution across all 5460 eQTL-probe pairs on the same chromosome shows 95.4% of *cis*-eQTL are within ± 4 Mb of their probe set and tend to have highly significant log of the odds scores (LOD, y-axis). *Cis*-eQTL resolution in the DO population is estimated to be 0.32 Mb. (C) Narrow-sense heritability is calculated for the probe set expression of each significant eQTL in the additive QTL analysis. (D) mirQTL plot of the results from the genome scan of miRNA transcription data with diet as an additive covariate. (E) Resolution of 18 *cis*-mirQTL, showing 100% of *cis*-mirQTL are within ± 4 Mb. *Cis*-mirQTL resolution in the DO population is estimated to be 0.25 Mb. (F) Distribution of narrow-sense heritability for the miRNA expressions with significant mirQTL in the additive genome scan.

Enrichr

Using the Ma'ayan Laboratory's Enrichr Tool (<https://amp.pharm.mssm.edu/Enrichr/>), we used the httr R library to programmatically submit gene sets to all gene-set libraries available on Enrichr via its Application Programming Interface (<https://amp.pharm.mssm.edu/Enrichr/help#api>). We then filtered the results to only include those with adjusted $P < 0.05$.

Correlation patterns in miRNA and mRNA

Pairwise Spearman correlation was performed on every sequenced miRNA ($n = 246$) and annotated mRNA probe ($n = 24,004$) pair. The outputs were combined and filtered for significance at a Benjamini–Hochberg (BH)-adjusted $P < 0.05$, resulting in 1,424,040 mRNA-miRNA correlations. Correlations were divided by mRNA into three categories based on the given model eQTL results: mRNA with significant eQTL were subdivided into “*cis*” and “*trans*,” and mRNA with

no significant eQTL were designated “no eQTL.” To better assess the differences between *cis*- and *trans*-eQTL, mRNA with multiple mappings (additive = 331; HP diet = 120; HFCA diet = 251) were removed for this analysis. Pairwise Wilcoxon rank-sum tests were performed on the absolute values of Spearman's rho for the three groups. Fisher's exact test was performed to test for dependence of mapping status and known miRNA-mRNA interactions. The *Mus musculus* miRNA catalog was downloaded from miRTarBase release 7.0 (Chou *et al.* 2018), and used to identify known miRNA-gene interactions.

mRNA eQTL and miRNA mirQTL colocalization

The confidence intervals of each eQTL and mirQTL on the same chromosome were tested for overlap. For those eQTL-mirQTL pairs with overlapping confidence intervals, a physical distance between peak markers was calculated. Pairs with a distance of zero were mapped to the same SNP.

Table 1 Means and ranges of permutation thresholds by significance level

	Additive model		Interactive model		High-fat cholic acid diet		High-protein diet	
	Mean	Range	Mean	Range	Mean	Range	Mean	Range
mRNA								
$P < 0.05$	7.51	7.14–10.11	10.87	10.27–18.65	7.65	6.98–10.02	7.74	7.08–13.77
$P < 0.63$	6.21	6.08–7.09	9.32	8.97–14.12	6.35	5.76–7.11	6.41	5.94–8.38
miRNA								
$P < 0.05$	7.63	7.41–8.41	11.67	10.54–14.48	8.1	7.58–9.38	7.81	7.61–8.34
$P < 0.63$	6.25	6.15–6.54	9.82	9.05–11.42	6.53	6.33–7.03	6.44	6.31–6.75

Differential gene expression

To determine the effect of diet on hepatic mRNA and miRNA, we performed differential gene expression analysis. Transcripts with a robust multiarray average value of four or greater in at least 25% of samples were included (Coffey *et al.* 2017), which resulted in 11,377 transcripts and 246 miRNA. The Wilcoxon rank-sum test was performed for both miRNA and mRNA data and P -values were corrected using the BH method. Significance was set at $P < 0.05$.

Allele-diet interaction identification

The interactive model generated using R/qtl2 suggested that many eQTL have unique diet effects from permutation analysis of LOD scores, but did not allow us to identify eQTL with a significant allele-by-diet interaction term within the model. To achieve this, we performed an ANOVA of the allele-diet interaction against the expression of the transcript with a significant eQTL in the interactive model. Interactive eQTL with a significant allele-diet interaction term (BH-adjusted $P < 0.05$) were subset to eQTL present in both the HFCA and HP diet models. For this subset, we calculated the founder allele effects using the R/qtl2 BLUP scan method in each diet model. We calculated the percent difference of founder effects of the HFCA diet model, using HP diet as a reference, at each eQTL. Finally, we summed the absolute value of the change across all of the founders for each gene, identifying those whose founder effects are most variable as a result of diet.

Data availability

Microarray and smRNA-seq data are available on the Gene Expression Omnibus repository under accession number GSE99561. Genotypes are available at DRYAD: <https://doi.org/10.25338/B87K75>. Supplemental material include all significant QTL identified in this study. Supplemental material available at figshare: <https://doi.org/10.25386/genetics.12597794>.

Results

Genetic architecture of hepatic mRNA and miRNA expression vary within the DO population

To identify the global regulation of hepatic gene expression, we performed a genome scan on all 248 female HFCA- or HP-fed DO mice (134 HFCA-fed and 114 HP-fed), treating diet

as an additive covariate. We identified eQTL and mirQTL in the full DO cohort. For eQTL, the median LOD score that corresponded to the 95% of genome-wide type 1 error rate was 7.51 (Table 1). The distribution of LOD scores significant at $P < 0.05$ is shown in Supplemental Material, Figure S1. Using permuted LOD thresholds for individual transcripts, we identified a total of 6692 significant eQTL, representing 6321 unique transcript cluster IDs that map to 6276 unique genes (Figure 2A). We next identified the 5460 eQTL associated with SNPs residing on the same chromosome as the gene probe and calculated the distance between the SNP and the gene's TSS. The mean absolute distance between the TSS and the peak SNP's physical location was 1.25 Mb and the median distance was 0.32 Mb (Figure 2B). To allow for variable recombination rates across the genome, we used a genomic interval of 4 Mb to classify transcripts as high-confidence *cis*-eQTL. Using this metric, we classified 5207 eQTL as *cis*-eQTL (Table S1). To support the decision to classify transcripts using a 4-Mb genomic interval, we tested a 1-Mb interval, which produced similar distributions of *cis* and *trans* due to the high resolution afforded in the DO (Table S2); therefore, all subsequent analyses were performed using the 4-Mb interval.

We treated all 1485 eQTL that did not fall into the *cis*-eQTL category as *trans*-eQTL, where a distal SNP (either >4 Mb on the same chromosome or on a different chromosome) is regulating the expression of a gene (Figure 2A; Table 2). When the position of the peak QTL SNP was plotted against the QTL probe position, we observed several vertical lines that may represent *trans*-eQTL bands. A total of 362 transcript cluster IDs exhibited multiple mappings, 31 of these *trans*-eQTL were the only genetic regulation observed, while 331 had both *cis*- and *trans*-eQTL signals (Table 3). All significant *trans*-eQTL are shown in Table S3.

We then identified mirQTL in the full DO cohort, treating diet as an additive covariate, and determined the threshold of significance by performing 1000 permuted genome scans on all 246 robustly expressed miRNAs. The median LOD score for a $P < 0.05$ threshold was 7.58. We identified a total of 37 miRNA with significant mirQTL (Figure 2D). From these, we identified 18 miRNAs whose peak eQTL was regulated by a SNP residing on the same chromosome as the miRNA's physical position. This distance ranged between 2.87 kb and 2.03 Mb, with median resolution of

Table 2 Summary of eQTL results by model

Run ^a	Additive model	Interactive model	High-fat cholic acid diet	High-protein diet
mRNA				
<i>Cis</i> (% total)	5207 (77.81)	4415 (71.73)	3738 (67.11)	2098 (61.43)
<i>Trans</i> (% total)	1485 (22.19)	1740 (28.27)	1832 (32.89)	1317 (38.57)
Total	6692	6155	5570	3415
miRNA				
<i>Cis</i> (% total)	18 (49)	9 (39)	8 (47)	1 (8)
<i>Trans</i> (% total)	19 (51)	14 (61)	9 (53)	11 (92)
Total	37	23	17	12

^a *Cis* and *trans* designations are defined as QTL peaks whose probes are within \pm 4Mb of the starting site or not, respectively.

0.25 Mb, which corresponds approximately to the same resolution calculated for mRNA (0.32 Mb) and allows us to classify all of them as *cis*-mirQTL. The *cis*-mirQTL are shown in Table S4 and *trans*-mirQTL are shown in Table S5.

Notably, the distribution of *trans* and *cis* were different for eQTL and mirQTL, as *cis*-eQTL made up 78% of the eQTL compared to only 49% of the mirQTL. The *trans*-mirQTL ($n = 19$) were found on chromosomes different from the physical location of their miRNA (Table 2). We identified four miRNA that had more than one mirQTL.

Heritability and effect size of eQTL and mirQTL in the DO population

To further explore the genetic architecture of hepatic mRNA and miRNA, we sought to determine the variance explained by each significant eQTL and the narrow-sense heritability (h^2) of each eQTLs associated gene. Among the genes with eQTL, the heritability ranged between 0.0 and 1.0, with a median h^2 of 0.295 (Figure 2C). The variance of expression levels explained by the peak eQTL SNP ranged between 0.027 and 0.776, with a median of 0.173. There was a significant positive correlation ($\rho = 0.52$, $P < 2.2 \times 10^{-16}$) between narrow-sense heritability of gene expression and variance explained by eQTL (Figure 3A). The allele probabilities of the peak SNP associated with *cis*-eQTL explained 20.2% of the variation in gene expression, while the peak SNP associated with *trans*-eQTL explained significantly less variation (Wilcoxon rank-sum, $P < 2.2 \times 10^{-16}$), at 13.61% (Figure 3C).

We next assessed the heritability of mirQTL, which were generally less heritable than eQTL. Among the miRNA with mirQTL, the heritability ranged between 0.0 and 0.66, with a median h^2 of 0.18 (Figure 2F). The variance explained by the peak mirQTL SNP ranged between 0.091 and 0.358, with a median of 0.14. Similar to the finding in eQTL, there was also a correlation between narrow-sense heritability of gene expression and variance explained by mirQTL (Figure 3B). The allele probabilities of the peak mirQTL SNPs associated with *cis*-mirQTL explained 14.32% of the variation in gene expression. *Trans*-mirQTL explained 12.98% but this difference was not found to be statistically significant ($P = 0.071$; Figure 3D).

Wild-derived alleles influence eQTL and mirQTL

A significant feature of the DO population is that allele distributions are composed of, on average, 12.5% from each of the eight founder strains at any given locus, meaning 37.5% of the genome is inherited from the three wild-derived founder strains (CAST/EiJ, PWK/PhJ, and WSB/EiJ) (Chesler *et al.* 2016). These three strains contribute a significant portion of the genetic variants in the DO population. Thus, we sought to determine if any of these strains' individual allele effects disproportionately contribute to eQTL and mirQTL. To do so, we mean-center-scaled the BLUP allele effects for the eight founder strains at the peak SNP of each eQTL (Figure 4A) and mirQTL (Figure 4C). The scaled founder effect sizes from all significant eQTLs demonstrated an enriched signal for alleles contributed by CAST/EiJ and PWK/PhJ strains (Kruskal–Wallis on the effect sizes, $P < 2.2 \times 10^{-16}$) (Figure 4B). Dunn's *post hoc* analysis, with BH multiple comparison correction, confirmed that CAST/EiJ and PWK/PhJ allele effects were significantly larger than all other founder strains ($P < 1.0 \times 10^{-16}$). Testing for this same effect in mirQTL was inconclusive; Kruskal–Wallis one-way ANOVA did not indicate statistically significant differences among founder strain mirQTL contributions ($P < 0.069$; Figure 4D).

Characterization of hepatic mRNA and miRNA interactions

We investigated if similar genetic architecture affects hepatic mRNA and miRNA expression by examining eQTL and mirQTL with overlapping confidence intervals (*Materials and Methods*). We identified 33 mirQTL whose confidence interval overlapped with the confidence interval of 817 eQTL. The miRNAs were approximately evenly distributed between *cis* and *trans* regulation, with 16 miRNAs being regulated in *trans* and 18 in *cis*. mRNAs were regulated more frequently by local variants as there were 643 *cis*-eQTL and only 174 *trans*-eQTL. The distance between the SNPs regulating mRNA and miRNA averaged 4.86 Mb and ranged between 0 and 161 Mb (Table S6). Notably, several mRNA and miRNA were regulated by the same SNP and in these cases all were *cis*-eQTL and *cis*-mirQTL. The functional relevance of having *cis*-eQTL and *cis*-mirQTL co-regulated at the same locus remains to be fully elucidated.

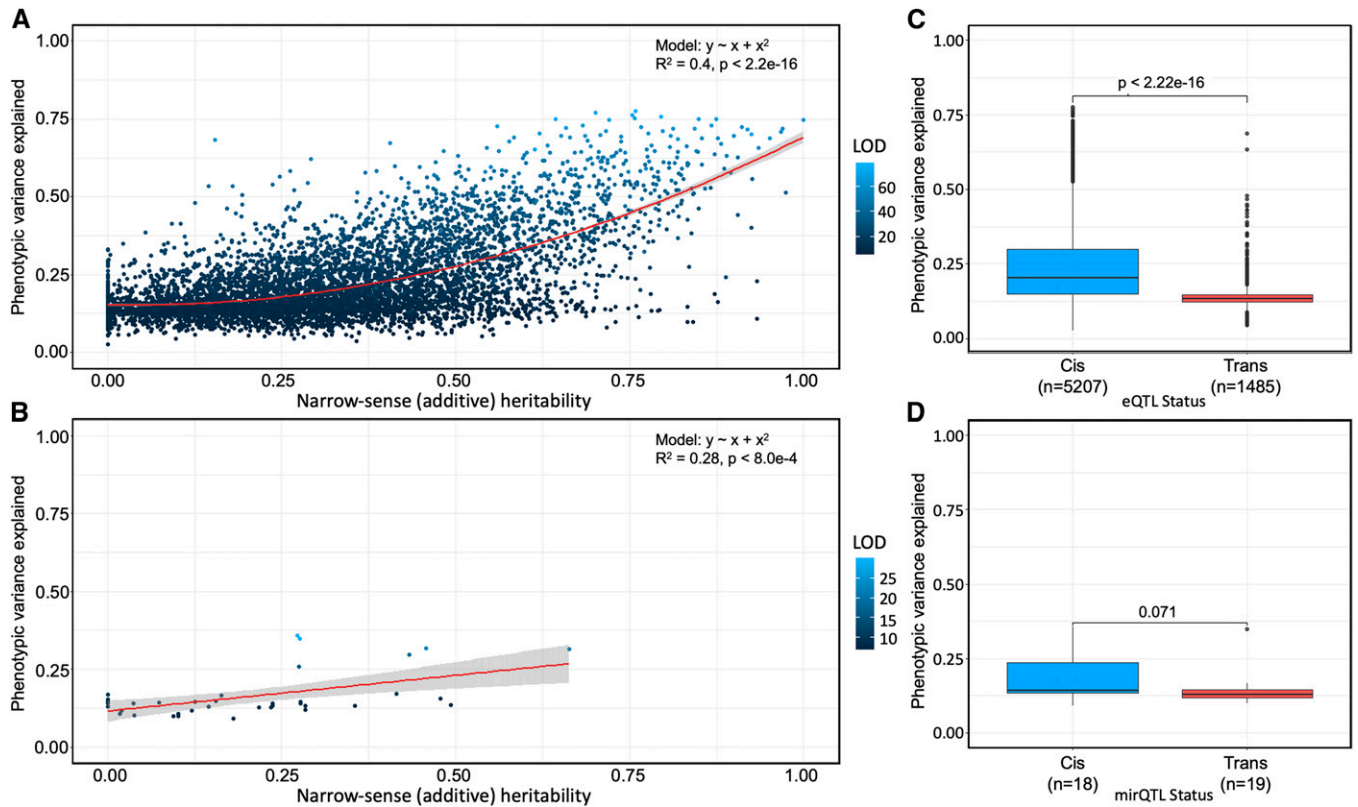


Figure 3 eQTL and mirQTL differ in heritability and effect size in Diversity Outbred mice. (A and B) Narrow-sense heritability scores and R^2 differences between full and null Haley–Knott regression models (phenotypic variance explained) were calculated for every significant eQTL and mirQTL in the additive genome scan models and colored by their LOD score. Polynomial and simple regression models were fit to the mRNA and miRNA data, respectively, to understand the association between heritability (x -axis) and phenotypic variance explained by the eQTL (y -axis). (C and D) Wilcoxon rank-sum tests reveal a significant difference in the phenotypic variance explained between *cis*- and *trans*-eQTL and a similar but insignificant ($P \geq 0.05$) difference in the phenotypic variance explained between the *cis* and *trans*-mirQTL.

A single miRNA can affect the expression of multiple mRNAs and thus play critical roles in the regulation of gene expression. So, in addition to shared genetic regulation, one can hypothesize that the correlation structure between miRNA and mRNA will result in significant interactions. We correlated expression levels of each miRNA with each mRNA and generated 5,904,492 correlations, of which 1,424,040 were significant after BH multiple comparison correction. The direction of the correlation may inform the biological association between the miRNA and mRNA and we observed a similar number of positive correlations ($n = 729,631$) and negative correlations ($n = 694,409$). We next categorized each of the correlations into three categories: miRNA:*cis*-eQTL, miRNA:*trans*-eQTL, and miRNA:no-eQTL. There were 376,051 significant miRNA:*cis*-eQTL correlations, 94,407 significant miRNA:*trans*-eQTL correlations, and 953,582 significant miRNA:no-eQTL correlations (Figure S2). The median (\pm SEM) absolute Spearman's rho for mRNA with *trans*-eQTL was 0.234 ± 0.00016 and was significantly different from both *cis*-eQTL and no-eQTL (*cis*-eQTL: $P < 2.2 \times 10^{-16}$, 0.233 ± 0.00035 ; no-eQTL: $P < 1.1 \times 10^{-5}$, 0.232 ± 0.00011). Although statistically significant, it is difficult to conclude whether the absolute expression of miRNA favors one form of genetic regulation.

Our global analysis of miRNA and mRNA was broad; therefore, we sought to characterize the miRNA-mRNA correlations based on known interactions. Of the 1,424,040 significant correlations, 1916 had validated miRNA-mRNA interactions on miRTarBase (Table S7) and 60% were positively correlated. We next characterized these high confidence miRNA-mRNA interactions based on the genetic regulation of the mRNA associated with the miRNA. Fisher's exact test on the dependence of mapping status (mapping: 470,458; no mapping: 953,582) and validated miRTarBase interaction (interaction: 1928; no interaction: 1,422,112) reveals that validated miRNA-mRNA interactions were roughly 11% less likely (odds ratio 0.89, 95% CI 0.81–0.98, $P < 0.02$) to be found between miRNA and mRNA with a mapping eQTL (Table 4). This finding supports that miRNA-mRNA interactions are more likely observed in the absence of a strong genetic signal.

Diet has profound effects on the genetic regulation of hepatic mRNA and miRNA expression

Having classified and characterized hepatic mRNA and miRNA while controlling for diet, we next sought to investigate the specific effects of diet on gene and miRNA expression (Figure 1C). We performed differential expression analysis

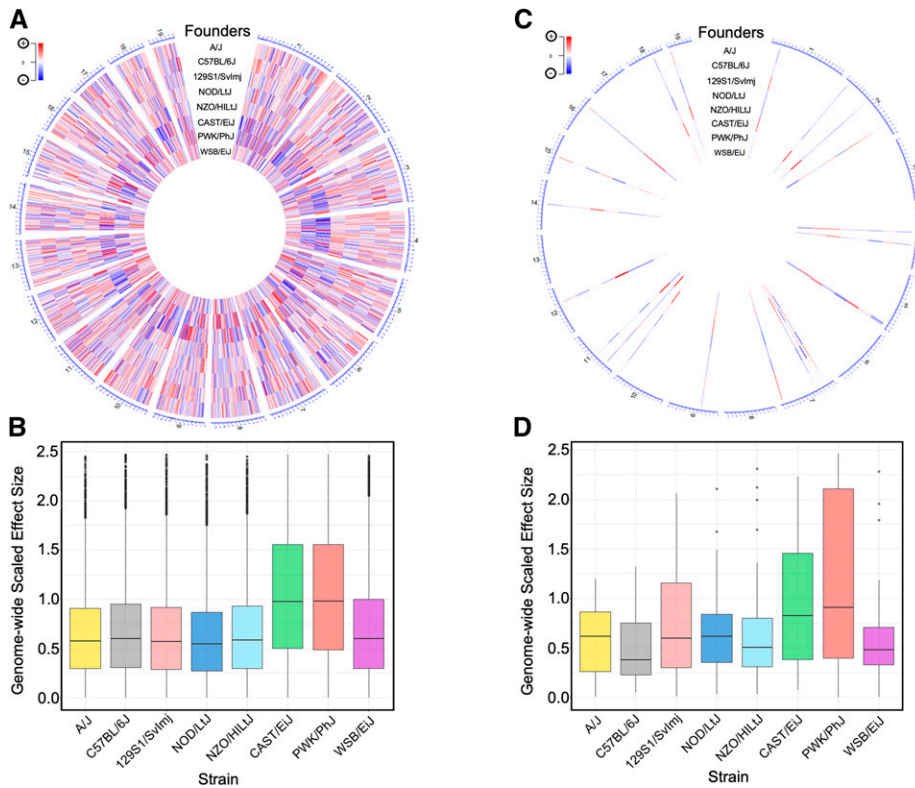


Figure 4 Alleles from wild-derived strains contribute to eQTL and mirQTL. (A) Center-scaled best linear unbiased predictor (BLUP) coefficients from eight DO founder strains are taken at the peak SNP of all significant eQTL from the additive model. Their relative effect sizes are visualized and mapped to the eQTL's physical location on a circularized mouse genome. The sixth and seventh tracks reveal a clear pattern of peak contribution (dark red and blue bands) from wild-type founder strains CAST/EiJ and PWK/PhJ that is consistent across the genome. (B) Boxplot representation of center-scaled BLUP coefficients from all significant eQTL in the additive model. Kruskal-Wallis rank-sum test indicates significant differences ($P < 2.2 \times 10^{-16}$) among founder effects. Dunn's *post hoc* test confirms significantly (BH-adjusted $P < 1.0 \times 10^{-4}$) larger effects from the CAST/EiJ and PWK/PhJ strains relative to all other strains. (C) BLUP coefficients from eight DO founder strains are taken at the peak SNP of all significant mirQTL from the additive model. Their relative effect sizes are visualized and mapped to the mirQTL's physical location on a circularized mouse genome. Patterns of peak contribution by any particular strain is visually unclear. (D) Boxplot of center-scaled BLUP coefficients from

all significant mirQTL in the additive model. Kruskal-Wallis rank-sum test does not indicate significant differences ($P < 0.064$) among founder effects. Dunn's *post hoc* test shows a single significant (BH-adjusted $P \leq 0.05$) difference, with PWK/PhJ showing higher contributions over NOD/LiJ.

by diet using the mRNA and miRNA that passed the robust multiarray average and RPMMM thresholds (*Materials and Methods*). A total of 8674 (76%) mRNA and 195 (79%) miRNA were differentially expressed, indicating a predominant effect of diet on the overall transcriptional profile (Tables S8 and S9). We classified the differentially expressed genes by their eQTL status in the additive model and found that 42% had an eQTL. In comparison, only 13% of differentially expressed miRNA had a mirQTL (Table 5). Using Fisher's exact test, we calculated that mRNA are less likely to have an eQTL when they are differentially expressed (odds ratio 0.79, 95% CI 0.73–0.87, $P < 0.001$). A stronger but nonsignificant trend was observed with miRNA.

We then analyzed the relationship between effect size, or phenotypic variance explained, and differential gene expression for the peak SNP of *cis*- and *trans*-eQTL in the additive model. The median phenotypic variance explained by the peak SNP associated with the *cis*-eQTL for a differentially expressed transcript was 0.186 (range, 0.027–0.751). In contrast, for genes without a statistically significant effect of diet on transcript levels, the median phenotypic variance explained was higher at 0.238 (range, 0.124–0.761). Thus, genes affected by diet had less of their variation explained by the peak SNP associated with an eQTL, underscoring the influence of the environment on this subset.

Further, analysis of the mode of regulation of eQTL (*cis* or *trans*) coupled with the differential expression results

paralleled our previous finding that *trans*-eQTL tended to explain less variance than *cis*-eQTL. For differentially expressed transcripts with a *trans*-eQTL, the variance in expression explained by the eQTL was 0.124 (range, 0.045–0.478), whereas the variance for transcripts not affected by diet was greater at 0.143 (range, 0.104–0.689). Likewise, for differentially expressed transcripts with a *cis*-eQTL, the variance explained was 0.186 (range, 0.027–0.750), which was less than the variance of nondifferentially expressed transcripts at 0.248 (range, 0.124–0.760). In summary, for both *cis* and *trans*, the median variance explained was significantly greater for mRNA that were not differentially expressed (Figure 5, A and B). This pattern was also observed for miRNA; however, the results were not significant, which may reflect the relatively low number of mirQTL (Figure 5, C and D).

To further characterize the behavior of differentially expressed genes and miRNA, we returned to our miRNA-mRNA interaction analysis. We assessed the magnitude of the miRNA-mRNA correlation by their mapping (miRNA:*cis*-eQTL, miRNA:*trans*-eQTL, and miRNA:no-eQTL) and differential expression status and found that in all instances, differentially expressed mRNA and miRNA have a statistically higher ($P < 2.22 \times 10^{-16}$) absolute rho than their nondifferentially expressed counterparts (Figure 6A). Further classification of the relationship between miRNA and mRNA based on differentially expressed gene status and classification of eQTL indicated that there was no difference in the

Table 3 Summary of eQTLs with multiple mappings by model

Run ^a	Additive model	Interactive model	High-fat cholic acid diet	High-protein diet
mRNA				
Mixed (% total)	331 (91.4)	265 (78.6)	251 (78.4)	120 (75.9)
<i>Trans</i> only (% total)	31 (8.6)	72 (21.4)	69 (21.6)	38 (24.1)
Total	362	337	320	158
miRNA				
Mixed (% total)	3 (75)	1 (100)	0 (0)	0 (0)
<i>Trans</i> only (% total)	1 (25)	0 (0)	0 (0)	0 (0)
Total	4	1	0	0

^a mRNA probes that appear as eQTLs at multiple locations on the genome have multiple mappings. There were no multiple mapping eQTLs that all mapped to the same chromosome (*cis* only), all others were a mix of *cis* and *trans* (mixed) or only *trans* (*trans* only).

miRNA-mRNA correlation structure based on eQTL classification if a transcript was not differentially expressed (Figure 6C). There was a statistically significant but relatively small effect of eQTL classification in transcripts that were differentially expressed by diet (Figure 6B). Thus, the effect of environmental factors such as diet have critically important effects on the relationship between hepatic mRNA and miRNA.

Diet alone affects *cis*- and *trans*-eQTL and mirQTL architecture

We have previously reported that diet significantly affects miRNA and mRNA expression in the DO population (Coffey *et al.* 2017). Moreover, having identified relationships between expression and genetic architecture, we sought to explore the effect of diet by comparing the *cis* and *trans* architecture of eQTL from separate genome scans of HFCA-fed mice ($n = 134$) and HP-fed mice ($n = 114$). The number of total eQTL found differed between diet groups, with 5570 eQTL in the HFCA-fed mice and 3415 in the HP-fed mice. We identified 3738 *cis*-eQTL and 1832 *trans*-eQTL in the HFCA-fed mice and 2098 *cis*-eQTL and 1317 *trans*-eQTL in the HP-fed mice (Table 2). The relative proportion of *trans*-eQTL was similar between diets: ~32% of eQTL were *trans*-acting in the HFCA diet model results, compared to 38% in the HP diet model results. The *cis*-eQTL of the HFCA- and HP-fed mice overlapped by 46% and 81%, while the *trans*-eQTLs overlapped by 1.9% and 2.7%, respectively (Figure 7, A–F). mirQTL analysis within the HFCA and HP diets, using individually permuted significance thresholds, yielded few mirQTL. There were 17 significant mirQTL in the HFCA diet, 8 (47%) of which were *cis*-mirQTL, and 12 significant mirQTL in the HP diet, only 1 (8%) of which was a *cis*-mirQTL. Overlap analysis revealed that the lone *cis*-mirQTL in the HP diet was represented among the *cis*-mirQTL of the HFCA-fed mice. Conversely, the HFCA- and HP-fed mice overlapped by a single *trans*-mirQTL.

We then classified the eQTL results from the diet specific QTL analysis as to their status relative to the differential expression results. Similar to the trend observed in the additive model, mRNA and miRNA were less likely to have an eQTL or mirQTL when they were differentially expressed (Tables S10 and S11). Furthermore, the odds of an eQTL having a differentially expressed gene were similar between HFCA and

HP diet models (HFCA diet: 0.83, 95% CI 0.76–0.91, $P < 0.001$; HP diet: 0.79, 95% CI 0.71–0.87, $P < 0.001$). To further explore the role of diet on *cis* and *trans* architecture, we calculated the phenotypic variance explained by eQTL and mirQTL in each diet model. In both HFCA-fed mice and HP-fed mice, the median variance explained for *cis*-eQTL was similar, at 0.349 and 0.356, respectively. *Trans*-eQTL explained 0.244 of the variance in the HFCA-fed mice and 0.282 in the HP-fed mice (Figure S3). This result follows the trend observed in the additive model that *cis*-eQTL explain more variance than *trans*-eQTL; however, the variance explained from the diet models is greater than that of the additive model, suggesting an environmental perturbation effect.

Overall, 33% of eQTL in the HFCA-fed mice were replicated in the HP-fed mice. This corresponded to 1836 eQTL in total, which were predominately *cis*. We identified 1701 concordant *cis*-eQTL and 35 concordant *trans*-eQTL overlapping between models. Power differences between the HFCA- and HP-fed mice resulted in differences in resolution and led to different classifications of 100 eQTL (Figure 7, E and F), guiding our choice to only discuss eQTL ($n = 1736$) that agreed in their classifications. We hypothesized that concordant eQTL would have similar effect sizes despite the different diets. The correlation of effect size between HFCA and HP diet models is 0.59 for *cis*-eQTL and 0.50 for *trans*-eQTL (Figure 8, A and B), suggesting a congruent relationship. Next, we calculated the effects of the eight founder strains at the peak SNP for each eQTL, using BLUP coefficients in each diet model. We correlated the allelic effects and assessed how their correlation related to the distance between the peak SNPs of the eQTL (Figure 8, C and D). We observed that allele effects underlying eQTL were highly correlated in concordant eQTL in both *cis*-eQTL and *trans*-eQTL.

We previously identified that alleles from the wild-derived strains CAST/EiJ and PWK/PhJ significantly contributed to eQTL and mirQTL in the full cohort with diet as an additive covariate. In both HFCA and HP diet models, CAST/EiJ and PWK/PhJ remained prominent contributors to eQTL and mirQTL, despite reduced power as a result of dichotomizing the cohort by diet (Figure S4).

We next investigated the relationships between mRNA and miRNA in each diet subset and established the median

Table 4 Fisher's exact test reveals that significantly correlated pairs of miRNA and mRNA are ~11% less likely to have known interactions when the mRNA has a significant eQTL (odds ratio 0.889, 95% CI 0.81–0.98, $P = 0.02$)

		Interaction	
		Yes	No
Mapping	Yes	588	469,870
	No	1340	952,242

Fisher's exact test did not show significant associations between known interactions and *cis* vs. *trans* mapping, $P = 0.15$.

absolute magnitude of the miRNA-mRNA correlations. Of all correlations, miRNA:*trans*-eQTL in the HFCA diet had the highest median absolute Spearman's rho of 0.249 ± 0.00034 , suggesting a coordinated response between the RNA classes in the extreme HFCA diet. The median value for *cis*-eQTL and no-eQTL were similar at 0.231 ± 0.0019 and 0.231 ± 0.0001 , respectively. In the HP diet, the median rho for *trans* and *cis* were similar at 0.230 ± 0.00025 and 0.230 ± 0.00035 , respectively, while transcripts without an eQTL had a higher median correlation coefficient of 0.232 ± 0.0001 (Figure S5, A and B). We acknowledge that the number and distributions of *cis*- and *trans*-eQTL are different between the diets. Therefore, the differences in the correlation results between diets may in part be attributed to power.

Gene-by-diet interactions reveal the significance of diet on eQTL and mirQTL

Given our previous reports of diet effects on hepatic gene expression (Coffey *et al.* 2017, 2019) and our observation of differing correlative structures between homologous GWAS candidate genes and traits in the DO population, we hypothesized that diet could affect the overall structure of eQTL and mirQTL. Thus, we performed eQTL and mirQTL analysis with diet as an interactive covariate and identified 6155 significant eQTL and 23 mirQTL (Figure S6, A and B), after completing individual permutation thresholds (Table 1). Of these, 4415 eQTL were *cis*-acting and 1740 were *trans*-acting, while nine mirQTL were *cis*-acting and 14 mirQTL were *trans*-acting (Table 2). As expected, the resolution of *cis*-eQTL and mirQTL identified in the interactive model were similar to the additive model, with the median distances between the TSS and the peak SNP's physical location being 0.28 and 0.31 Mb, respectively.

Next, we compared the relative distribution of *cis*- and *trans*-acting eQTL in both the additive and interactive models. Overlap between significant eQTLs using diet as an interactive or additive covariate revealed large architectural differences in *trans*-eQTLs compared to *cis*-eQTLs. Large overlap between *cis*-eQTLs was observed between the models with 4321 transcripts with clear *cis*-acting regulation, and 83% and 98% of *cis*-eQTLs overlapping from the additive and interactive models, respectively. Conversely, only 28% and 24% of *trans*-eQTLs overlapped between the additive and interactive models (Figure S6E). mirQTL architecture was

Table 5 Fisher's exact test reveals that mRNA are less likely to have an eQTL when they correspond to a differentially expressed gene (odds ratio 0.79, 95% CI 0.73–0.87, $P < 0.001$)

		Differential expression		
		mRNA		eQTL
		Yes	No	
eQTL	Yes	3541	1253	42%
	No	5133	1450	58%
		miRNA		mirQTL
		Yes	No	
mirQTL	Yes	24	9	13%
	No	171	42	87%

A stronger but nonsignificant trend is observed with miRNA.

also affected by the type of QTL model, as only 8 out of 19 *cis*-mirQTL (42%) were similar between the additive and interactive models. *Trans*-mirQTL had model overlap similar to *trans*-eQTL, with 9 out of 24 (37%) represented in both models (Figure S6F). The sparse overlap of *trans*-eQTL may be due in part to their smaller effect sizes relative to *cis*-eQTL yet may suggest that genes regulated in *trans* are more sensitive to gene-by-diet interactions than those regulated by *cis* factors.

Concordant eQTL are driven by different founder alleles in response to diet

The genetic architecture differed between the additive and interactive models, indicating that a subset of eQTL could have a distinct response to diet. In support of this hypothesis, we identified that the allele effects of concordant eQTL in the HFCA and HP diet models were highly correlated; however, this was not universal as the correlation of the allele effects was often below one (Figure 8, C and D). We hypothesized that a subset of eQTL present in both the HFCA and HP diet models could be affected by alternative founder alleles. To test this, we began by identifying eQTL with significant allele-diet interactions in the interactive model. To accomplish this, we performed an ANOVA of the allele-diet interaction against the expression of the transcript with a significant eQTL. Of the 6155 eQTL identified in the interactive model, 1460 had a significant allele-diet interaction ($P < 0.05$) and 390 were significant after BH correction (BH-adjusted $P < 0.05$) (Table S12). We then subset the eQTL present in both the HFCA and HP diet models to those with a significant (BH-adjusted $P < 0.05$) allele-diet interaction, which represented 67 eQTL. We determined the effect of diet by calculating the percent difference in the founder allele effects of the HFCA-fed mice compared to the reference, HP-fed mice (Table S13). We summed the absolute value of the change across all of the founders for each gene and highlight the top 20 in Figure 9A and provide an example (Figure 9, C and D). The *cis*-acting mRNA corresponding to Zinc Finger Protein 982 (ZFr982) demonstrates a distinct response to diet; in the HP diet, the eQTL is driven by A/J, whereas in the HFCA diet, the same eQTL is driven by C57BL/6J (Figure 9, C and D).

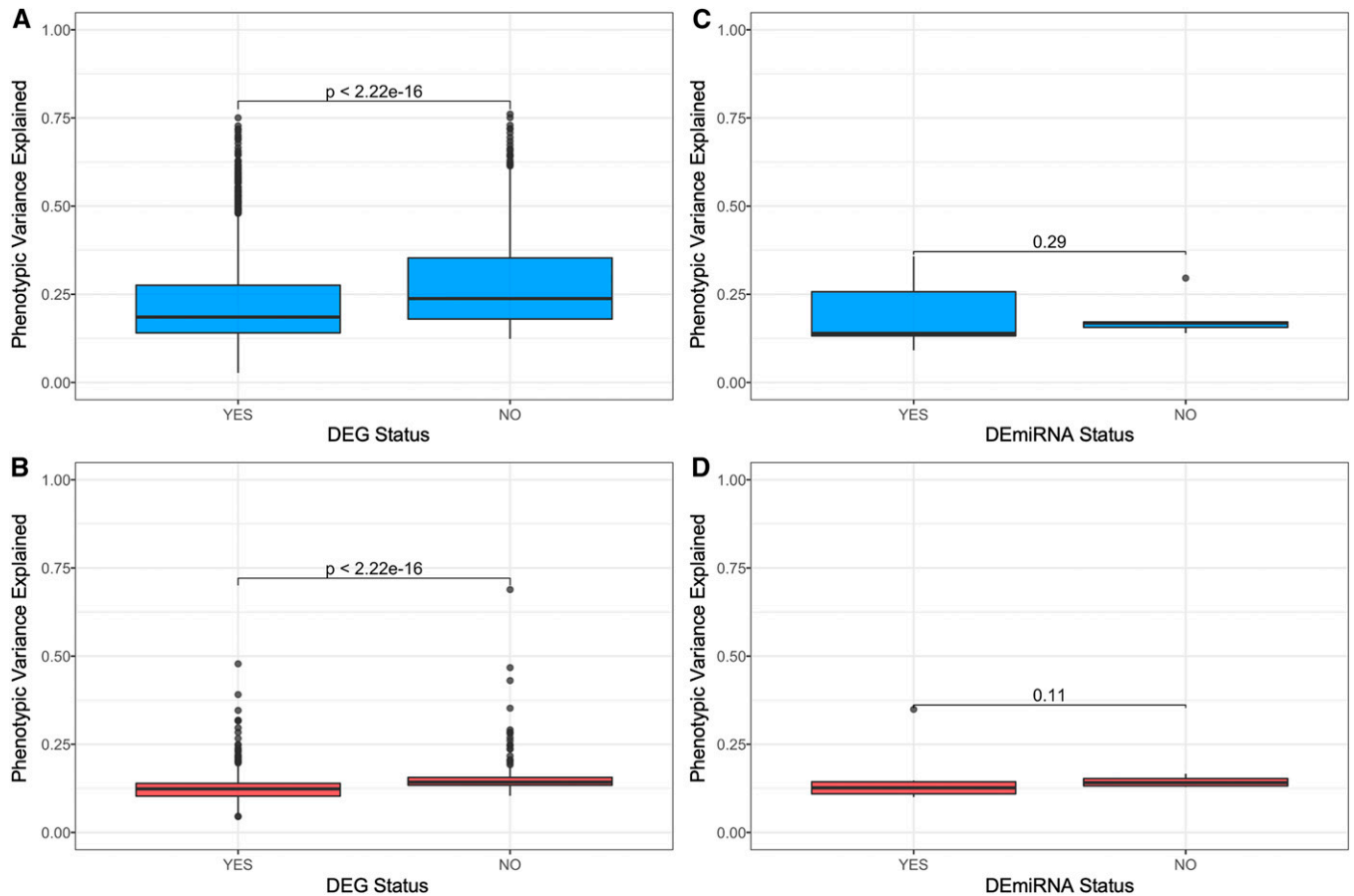


Figure 5 Effect size of eQTL is lower in genes and miRNA that are differentially expressed. The allelic effect size is significantly lower ($P < 2.22 \times 10^{-16}$) in genes that are differentially expressed in both *cis*- and *trans*-eQTL. (A) For *cis*-eQTL, the median variance explained when the transcript is differentially expressed is 0.186 compared to 0.238 when it is not differentially expressed. (B) For *trans*-eQTL, the median variance explained when the transcript is differentially expressed is 0.124 compared to 0.143 when it is not differentially expressed. (C) The allelic effect size is not significantly different in *cis* or *trans*-mirQTL; however, the median of differentially expressed *cis*-mirQTL is 0.139 compared to 0.169 when not differentially expressed. (D) Similarly, differentially expressed mirQTL in *trans* had a smaller median than those not differentially expressed at 0.127 and 0.141, respectively.

Discussion

GWASs have revolutionized genetic studies and expanded our understanding of complex trait genetics and genomic variants associated with disease susceptibility. Numerous studies have demonstrated that expression traits (eQTL) can be used as quantitative traits, and integrative studies of eQTL provide improved candidate selection at GWAS loci (Zhu *et al.* 2016). Despite these successes, much remains to be determined in regard to the genetic architecture. Identification of missing heritability (Edwards *et al.* 2013) and gene-by-environment interactions (Tan *et al.* 2012) are of particular interest. In the current study, we reported the global regulation of hepatic mRNA and miRNA expression using a genetically diverse population of mice named the DO. Our studies identified several important results. First, we classified (*cis* and *trans*) the genetic regulation of mRNA and miRNA in DO mice and proceeded to characterize the eQTL and mirQTL by describing heritability, phenotypic variance explained, and allelic contributions. Second, by coupling our correlation and differential expression analyses, we showed increased correlation

between miRNA-mRNA pairs differentially expressed by diet and most importantly increased correlation between miRNA-mRNA pairs without an eQTL. Finally, we reported that several eQTL are driven by different founder alleles as a result of diet. We discuss each of these in turn.

Classification of genetic architecture of hepatic mRNA and miRNA

A primary goal of our study was to describe the genetic architecture of hepatic mRNA and miRNA expression profiles. Our initial analysis identified eQTL and mirQTL, using diet as an additive covariate. We classified these as *cis* or *trans*-acting and used the *cis*-acting mirQTL and eQTL to estimate resolution. The DO mice have high-resolution mapping as demonstrated by the fact that 95–99% mirQTL and eQTL with *cis*-acting variants exert their effect within 1 Mb of the physical position of the mRNA or miRNA.

One of the key strengths of a “systems genetic” approach is the ability to resolve the genetic architecture of different scales of data such as mRNA and miRNA expression. In these

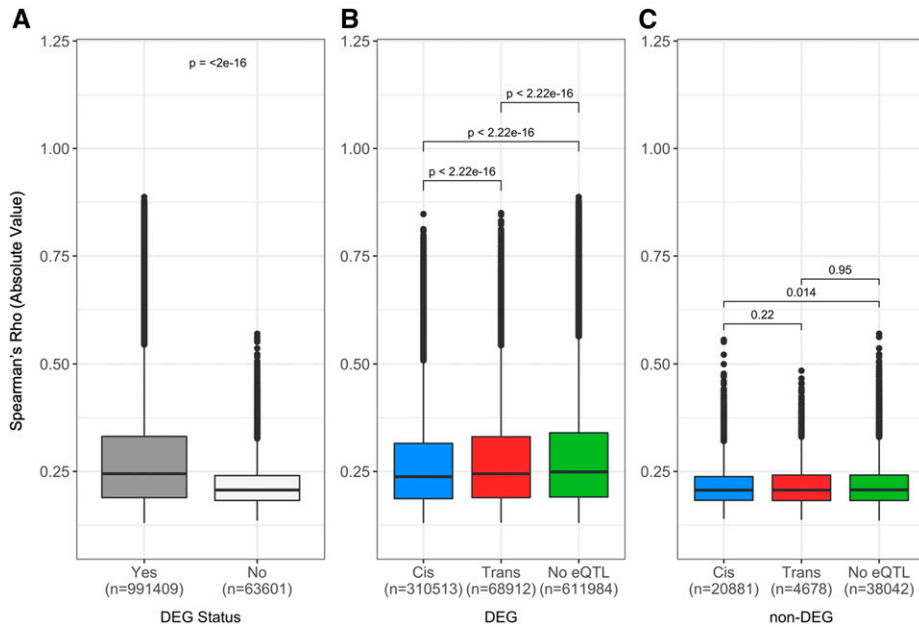


Figure 6 Magnitude of correlation is similar between miRNA and mRNA with *cis*, *trans*, and no-eQTL and reduced when mRNA are not differentially expressed. Expression data were used to generate correlations between pairs of miRNA ($n = 246$) and mRNA ($n = 24,004$) and the results were filtered for false-discovery-rate-corrected significance ($\alpha \leq 0.05$), resulting in 1,424,040 significant correlations. miRNA-mRNA pairs were classified by their eQTL status in the additive model. A total of 1,055,010 eQTL with transcripts that passed the quality threshold for the DE analysis were included. (A) Spearman's rho was plotted against the DE status of the mRNA from the additive genome model. Median \pm SE for DEG and non-DEG include 0.245 ± 0.0001 and 0.207 ± 0.0002 , respectively. (B and C) Spearman's rho was plotted against the eQTL status of mRNA (B) that were differentially expressed by diet and (C) that were not differentially expressed by diet. For differentially expressed mRNA with *cis*, *trans*, and no-eQTL, the median \pm SE are 0.238 ± 0.0002 , 0.245 ± 0.0005 , and 0.249 ± 0.001 , respectively. For nondifferentially expressed mRNA with *cis*, *trans*, and no-eQTL, the median \pm SE are 0.207 ± 0.003 , 0.207 ± 0.0007 , and 0.207 ± 0.002 , respectively. DEG, Differentially Expresses Gene.

studies, we identified thousands of genes regulated by eQTL but only dozens of miRNA regulated by mirQTL. The sheer difference in numbers between these two scales of data are notable but coincides with previous reports. The lower number of mirQTL has been previously reported in the Framingham cohort, where 76 miRNAs were regulated by mirQTL (Huan *et al.* 2015). We acknowledge that the differences (or lack of statistically significant differences) reported in subsequent analyses related to the distributions of eQTL and mirQTL could reflect the much smaller number of miRNA quantitated in the current study.

Our analyses identified differing genetic architecture affecting mRNA and miRNA but confirmed that *cis*-acting variants affecting both mRNA and miRNA explain more phenotypic variance than *trans*-acting variants. These results suggest that miRNA are more likely to be regulated by *trans*-acting factors, while mRNA are more likely to be regulated by *cis*-acting factors. Over 50% of mirQTL are *trans*-acting, whereas <25% of eQTL are *trans*-acting. Unlike a recent report of eQTL in yeast, we do not observe that most eQTL are *trans*-acting in nature (Albert *et al.* 2018), but this could be due to differences in power or model organism of each study. While the functional consequences of this prominent difference in genetic architecture of mRNA and miRNA remains unclear, similar effects have been observed in studies utilizing CC mice (Rutledge *et al.* 2015), a genetically related population to the DO mice used in this study. Although speculative, these data from the CC and DO mouse populations suggest that individual miRNAs on average may be regulated by fewer transcription factors than individual genes. In such a scenario, SNPs affecting transcription factor loci would be expected to have greater effects on miRNA expression than

on mRNA expression, which have multiple transcription factors regulating their expression.

Identifying the "optimal" definition of *cis* and *trans* (or local and distal) eQTL remains an open question. We note that we used a liberal definition of 4 Mb to define *cis*-eQTL, which reflects the mean confidence interval (~ 4 Mb) for all significant eQTL, as determined by the Bayesian credible interval (*Materials and Methods*). In addition, human GWASs have observed regions of SNPs that are physically located ~ 2 Mb apart, in LD and are associated with a complex trait, such as height (Yang *et al.* 2012). For these reasons, we chose to utilize a larger genomic interval (4 Mb) to classify the eQTL. We mention that using a 1 Mb definition of *cis*, a similar window as human GWASs, would classify 923 additional eQTL, $\sim 14\%$, as *trans* instead of *cis*. Importantly, the classification used here and within many eQTL studies is computational and not based on functional testing. Recombination rate and local LD structures will affect any global definition of *cis* or *trans*. Regardless of the classification given to a particular eQTL, understanding the genetic variant's contribution to differences in expression is ultimately the goal, but it is not yet feasible to test the mode of function or identify the causal variants for each eQTL.

Much emphasis has been placed on understanding the missing heritability observed in GWASs. Molecular traits provide a dense phenotypic space to explore these differences, and in our current studies, the median heritability of eQTL is 0.295 and mirQTL heritability is 0.18 in the additive model. The heritability estimates for eQTL are slightly higher than those recently reported in human blood cells, which was 0.089 (Lloyd-Jones *et al.* 2017), and may reflect the relatively similar environmental variation afforded by studies utilizing

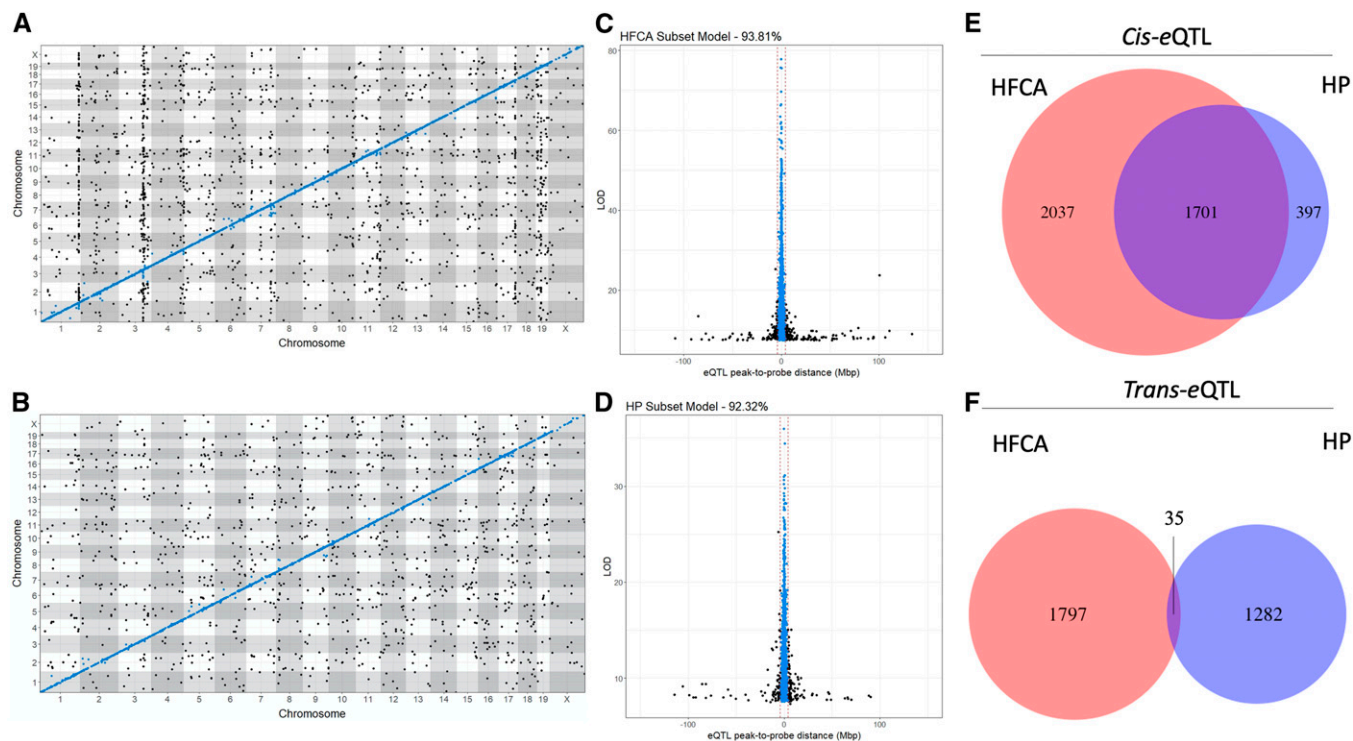


Figure 7 Analysis of genome scans from HFCA- and HP-fed mice reveal environmentally driven differences in genetic architecture. (A) Global eQTL architecture for the HFCA diet model shows numerous, dense *trans*-bands, in contrast to the (B) HP diet model, which shows notably less colocalization of *trans*-eQTL. Both diet models (C and D) show similar precisions from eQTL with probe sets on the same chromosome; resolution of these 3987 eQTL from HFCA-fed mice and 2273 eQTL from HP-fed mice are estimated to be 0.29 and 2.06 Mb, with 93.81% and 92.32% of *cis*-eQTL occurring ± 4 Mb from their probe sets, respectively. This indicates that structural variants associated with local genetic regulation tend to occur close to the gene itself, regardless of environmental perturbations. (E and F) A large number of *cis*-eQTL overlap despite environmental differences, while the lack of overlap in the *trans*-eQTL might be indicative of environmental effects.

rodents. Our heritability estimates for mirQTL are quite similar to those previously reported in humans, where heritability of miRNA ranged between 0.0 and 0.57, with an average heritability of 0.11 (Huan *et al.* 2015; Lloyd-Jones *et al.* 2017). Similar to reports in humans, a subset of both eQTL and mirQTL in the DO population have heritability at or near 0.0 (Huan *et al.* 2015). Although the heritability of eQTL and mirQTL were similar, results from the additive model indicate that the effect size of eQTL varied considerably between 0.03 and 0.8, while mirQTL were limited to a range of ~ 0.1 –0.35. These differences in effect sizes were also evident when eQTL and mirQTL were characterized by their type of regulation, *cis*- or *trans*-acting. The phenotypic variance explained by *cis*-acting eQTL was 25% higher than *trans*-acting eQTL, and a similar trend was observed in mirQTL. Furthermore, these trends proved to be robust, as *cis*-acting eQTL explained more variance than their *trans*-acting counterparts in the diet-specific analyses.

Studies utilizing the DO population often assign probabilities at each SNP to the eight founder alleles (Gatti *et al.* 2014). Three of the founder strains, CAST/EiJ, WSB/EiJ, and PWK/PhJ, are wild-derived and are more divergent from classical inbred strains. These strains contain between 900, 1000, and 5 million private SNPs (Keane *et al.* 2011), and these are passed along at relatively equal proportions in the DO

population (Chesler *et al.* 2016). Thus, the DO population contains variants that affect complex traits at loci that may not vary genetically in classical inbred strains (Keane *et al.* 2011). We hypothesized that variants from these divergent strains contribute disproportionately to eQTL and mirQTL. To test this hypothesis, we calculated the allele effects for each of the founder strains at the peak SNP for each gene or miRNA. Regardless of the model, the scaled coefficients for each founder haplotype were significantly higher for CAST/EiJ and PWK/PhJ alleles at each eQTL, indicating that a higher proportion of the variation in the expression of these eQTL and mirQTL are explained by variants from these strains. Similar effects have recently been reported in the related CC population (Keele *et al.* 2020). The authors also noted both higher magnitude of genetic effects for CAST/EiJ and PWK/PhJ alleles and relative consistency of haplotype effects derived from the eight founder strains. The exact mechanisms underlying these eQTL and the contribution of specific haplotypes remain to be determined, but are supported by large-scale resequencing. These efforts have included the eight founders of the DO and CC populations and identified a number of loci that contain novel genes or novel orthologs and are enriched for proteins associated with defense and immunity, nucleic acid binding, and transcription factors (Lilue *et al.* 2018). In particular, two of the

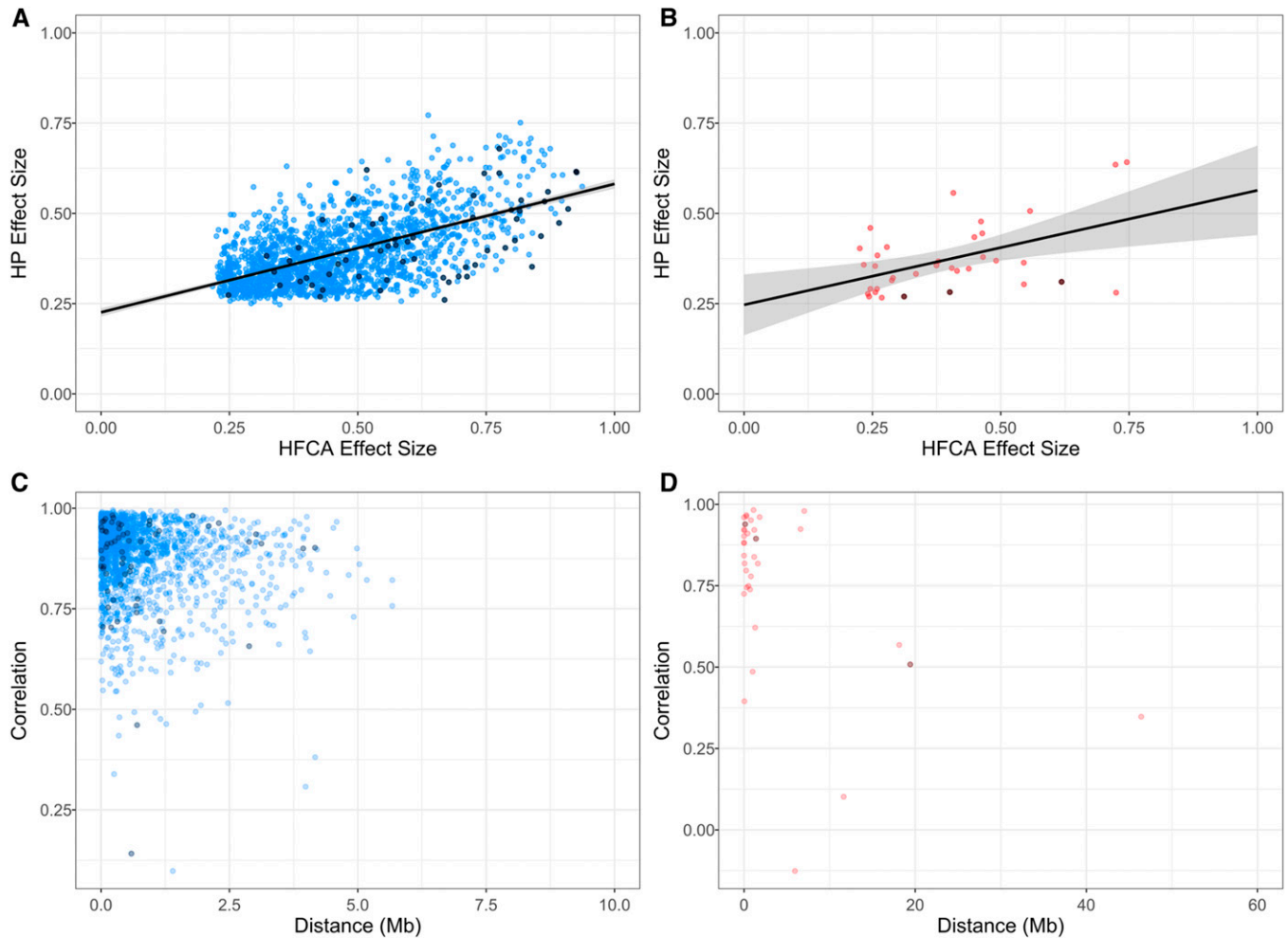


Figure 8 Comparison of concordant eQTL pairs in mice fed different diets demonstrates similarities despite different environments. Effect size of eQTL observed in both HFCA and HP diets are significantly correlated in both (A) *cis*-eQTL ($r = 0.59$) and (B) *trans*-eQTL ($r = 0.50$). Allele effects underlying concordant eQTL are highly correlated in both (C) *cis*-eQTL and (D) *trans*-eQTL despite different diet perturbations. Black dots represent eQTL with a significant allele-diet interaction (ANOVA, BH-adjusted $P < 0.05$). Color gradient represents dot density.

founder strains of the DO and CC populations, CAST/EiJ and PWK/PhJ, contained a number of loci with high sequence variation, supporting the findings of the current study.

Relationships between hepatic miRNA and mRNA are greater when diet affects expression

mRNA and miRNA have different genetic architecture, yet our data suggests similar correlation between mRNA and miRNA regardless of the underlying *cis* and *trans* regulation. More detailed examination of the effect of diet suggests that miRNA-mRNA pairs differentially expressed by diet were generally more correlated than pairs unaffected by diet regardless of the eQTLs mapping status. These data support that miRNA-mRNA associations act as an additional regulatory mechanism underlying an organism's response to the environment. This global perspective does not address direct interactions between mRNA and miRNA, which are important aspects of gene regulation. Approximately half of the significant correlations we observed were positive; these

scenarios may be indicative of enriched pathways as opposed to interactions between miRNA and target mRNA, which would be predicted to have negative correlations. Clearly, a number of the correlations contained in the global analysis are indirect and coincidental. When miRNA-mRNA correlations are restricted to interactions validated in miRTarBase, we observe multiple interactions associated with cardiovascular and metabolic syndromes. For example, miR-34a is negatively correlated (Spearman's $\rho = -0.760$) with Autophagy Related 9A protein, ATG9a, a gene previously shown to be involved in cardiomyocyte hypertrophy and regulated by miR-34a (Huang *et al.* 2014). miR-34a is also negatively associated with growth arrest specific 1, GAS1 (Spearman's $\rho = -0.582$), a GWAS candidate associated with plasma triglyceride levels (Rhee *et al.* 2013). In addition, we observe that miRNA with a validated miRNA-mRNA interaction are $\sim 11\%$ less likely to be correlated with a mRNA with an eQTL. We do acknowledge that we are more likely to detect QTL with inflated effect size and are unable to recognize every

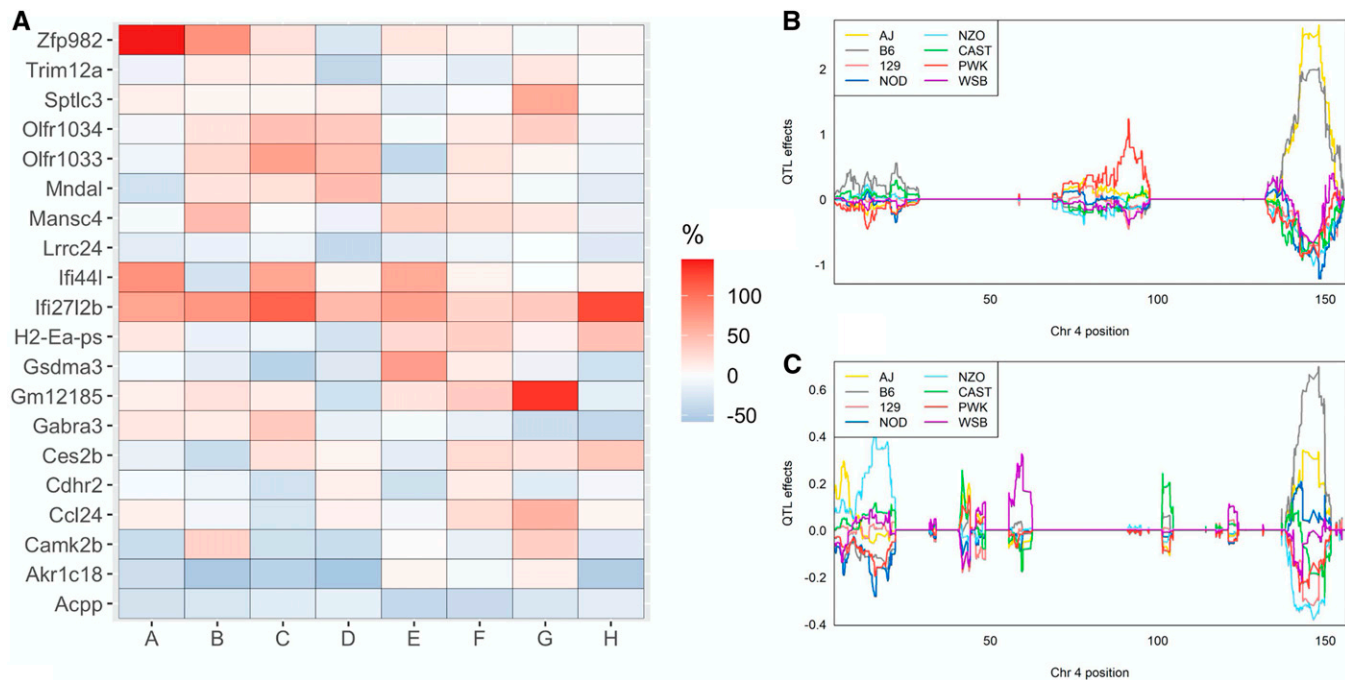


Figure 9 Significant genotype by diet interactions on the mRNA expression. (A) The top 20 genes with the greatest percent change of the best linear unbiased predictors (BLUP) founder allele coefficients for eQTLs with a significant allele-diet interaction in the HFCA diet with respect to the HP diet. (B) BLUPs coefficient plot of the eight founder mice strains to the *Zfp982* eQTL in the HFCA diet and (C) the HP diet. Color represents the eight founder mice strains as indicated.

QTL due to power (King and Long 2017); however, the observation that most miRNA-mRNA associations occur between miRNA and mRNA without an eQTL suggests that coordination between the RNA classes occurs most in the absence of a strong genetic signal. We also note that miRNA-mRNA regulatory networks are complex and could involve indirect effects outside of the classically described direct binding of a miRNA to a 3'UTR (Su *et al.* 2011).

Diet has varying effects on eQTL and mirQTL

There is much interest in determining the missing heritability observed in GWASs. One possible cause of missing heritability are differences in environmental exposures, which implicitly vary in large GWAS analysis. More importantly, the effect of environment in these large genetic studies is heterogeneous within the study population and often difficult if not impossible to accurately quantify. In the current study, we varied diet in siblings and performed the initial analysis with diet as a covariate. To further understand the role of diet, we completed differential expression analysis of mRNA and miRNA by diet. The proportions of differential expression were similar between RNA classes, with 76% of genes and 79% of miRNA differentially expressed, suggesting similar responsiveness to diet. Additionally, mRNA with a significant eQTL were less likely to be differentially expressed than nonmapping mRNA. In miRNA the pattern persisted but was not statistically significant. When we classified eQTL by their differential expression status, we clearly observed that phenotypic variance explained was greater in nondifferentially expressed

genes, perhaps indicating a stronger genetic than environmental signal regulating their expression.

Inbred strains are known to vary in their response to diet (West *et al.* 1992), and thus it is quite possible that a fraction of eQTL are influenced by diet. There have been limited investigations characterizing how diet affects eQTL status. Consomic mice have been used to confirm that alleles from AJ mice are responsible for *Xbp1* and *Socs3* diet-specific eQTL (Pasricha *et al.* 2015). Studies utilizing a different mouse population, the Hybrid Mouse Diversity Panel, have noted diet specific eQTL patterns (Parks *et al.* 2013). In the current study, we identified eQTL with a significant allele-diet interaction by regression analysis. This subset of eQTL represent genes whose founder alleles differ in response between diets. We further identified the top 20 genes whose allele effects most dramatically changed by diet. For instance, we identified *Zfp982*, Zinc Finger Protein 982, whose expression is associated with an A/J allele in the HP diet and a C57BL/6J allele in the HFCA diet. While the functional consequences of a gene-by-diet interaction for *Zfp982* remains to be determined, several of the genes with allele by diet interactions are GWAS hits. For example, variants proximal to *Camk2b* have been associated with numerous traits in humans, including diabetes (Morris *et al.* 2012). Functional studies in mice have demonstrated that *Camk2b*^{-/-} are susceptible to obesity (Bachstetter *et al.* 2014) and that the enzyme *Camk2b* is translated into, CaMKII, is involved in the hepatic insulin resistance that occurs with obesity (Ozcan *et al.* 2013). Furthermore, there is evidence that expression of *Camk2b* may

also be under genetic regulation in humans (Frau *et al.* 2017; <https://www.gtxportal.org/home/locusBrowserPage/ENSG00000058404.19>). These results provide evidence that gene-by-diet interactions affect the mRNA abundance observed in the DO population and may have important implications for disease-related phenotypes.

Additional work remains to understand the underlying mechanisms regulating both differential gene expression and allele-diet interactions. Integration of multiple scales of sequencing data may provide additional insight. For example, assaying chromatin accessibility or methylation patterns could help understand the specific variants or DNA modifications by which diet exerts its effects on gene expression. Recently, both the genetic regulation of gene expression and chromatin accessibility (assay for transposase-accessible chromatin using sequencing) was assessed in 47 strains comprising the CC mouse panel (Keele *et al.* 2020). Keele and colleagues highlight the tissue specificity of eQTL, the underlying DNA modifications corresponding to specific eQTL, and perform mediation analysis to identify potentially causal variation in chromatin accessibility responsible for specific eQTL. Diet has been shown to affect chromatin structure (Leung *et al.* 2014, 2016) and thus provides a plausible connection between diet-driven DNA modifications and eQTL. Although not tested in the current study, it is intriguing to speculate that some of the differential expression and diet-specific eQTL are due to changes in chromatin structure.

In conclusion, gene expression studies in a segregating population provide interesting insights into complex genetic regulation. In addition, assessing these patterns in the context of diet helps to reveal the relationships between genetics and the environment. In the current study, we have classified the genetic regulation and characterized the genetic architecture of hepatic mRNA and miRNA in a genetically diverse population of mice fed different diets. We observed significant differences in the regulation, effect size, and overall genetic architecture of mRNA and miRNA, and note the varying effects of diet on these trends. Overall, these key regulatory differences underscore the necessity for continued investigation into how the diverse spectrum of RNA classes are regulated, how they respond to environmental stimuli such as diet, and how dysregulation may predispose an organism to disease.

Acknowledgments

This research was supported in part by grants from the National Institutes of Health 5R01HL128572 (to B.J.B.), P30DK056350 (to B.J.B. and D.P.), and R01DK105965 (to P.S.), a pilot grant from the Nutrition Research Institute (to B.J.B. and P.S.), United States Department of Agriculture/Agricultural Research Service/Western Human Nutrition Research Center project funds (2032-51000-022-00D) and an National Science Foundation predoctoral fellowship (A.R.C.). We are grateful to Michael Vernon and the University of North Carolina at Chapel Hill Functional Genomics Core for microarrays, and Zhao Lai at the

University of Texas Health Science Center at San Antonio (TX) for small RNA sequencing. We thank John Calaway for assistance with RNA extraction from liver tissue, and Kunjie Hua, Kuo-Chen Jung, and Liyang Zhao (Department of Genetics, University of North Carolina at Chapel Hill) for assistance with husbandry and phenotyping. The USDA is an equal opportunity employer.

Literature Cited

- Albert, F. W., J. S. Bloom, J. Siegel, L. Day, and L. Kruglyak, 2018 Genetics of trans-regulatory variation in gene expression. *eLife* 7: e35471. <https://doi.org/10.7554/eLife.35471>
- Aylor, D. L., W. Valdar, W. Foulds-Mathes, R. J. Buus, R. A. Verdugo *et al.*, 2011 Genetic analysis of complex traits in the emerging Collaborative Cross. *Genome Res.* 21: 1213–1222. <https://doi.org/10.1101/gr.111310.110>
- Ayroles, J. F., M. A. Carbone, E. A. Stone, K. W. Jordan, R. F. Lyman *et al.*, 2009 Systems genetics of complex traits in *Drosophila melanogaster*. *Nat. Genet.* 41: 299–307. <https://doi.org/10.1038/ng.332>
- Bachstetter, A. D., S. J. Webster, T. Tu, D. S. Goulding, J. Haiech *et al.*, 2014 Generation and behavior characterization of CaMKII β knockout mice. *PLoS One* 9: e105191. <https://doi.org/10.1371/journal.pone.0105191>
- Bhasin, J. M., E. Chakrabarti, D. Q. Peng, A. Kulkarni, X. Chen *et al.*, 2008 Sex specific gene regulation and expression QTLs in mouse macrophages from a strain intercross. *PLoS One* 3: e1435. <https://doi.org/10.1371/journal.pone.0001435>
- Broman, K. W., 2012a Genotype probabilities at intermediate generations in the construction of recombinant inbred lines. *Genetics* 190: 403–412. <https://doi.org/10.1534/genetics.111.132647>
- Broman, K. W., 2012b Haplotype probabilities in advanced intercross populations. *G3 (Bethesda)* 2: 199–202. <https://doi.org/10.1534/g3.111.001818>
- Broman, K. W., and S. Sen, 2009 *A Guide to QTL Mapping with R/qtl*. Springer, Dordrecht, Holland. <https://doi.org/10.1007/978-0-387-92125-9>
- Brynedal, B., J. Choi, T. Raj, R. Bjornson, B. E. Stranger *et al.*, 2017 Large-scale trans-eQTLs affect hundreds of transcripts and mediate patterns of transcriptional co-regulation. *Am. J. Hum. Genet.* 100: 581–591. <https://doi.org/10.1016/j.ajhg.2017.02.004>
- Chesler, E. J., D. M. Gatti, A. P. Morgan, M. Strobel, L. Trepanier *et al.*, 2016 Diversity outbred mice at 21: maintaining allelic variation in the face of selection. *G3 (Bethesda)* 6: 3893–3902. <https://doi.org/10.1534/g3.116.035527>
- Chick, J. M., S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi *et al.*, 2016 Defining the consequences of genetic variation on a proteome-wide scale. *Nature* 534: 500–505. <https://doi.org/10.1038/nature18270>
- Chou, C. H., S. Shrestha, C. D. Yang, N. W. Chang, Y. L. Lin *et al.*, 2018 miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 46: D296–D302. <https://doi.org/10.1093/nar/gkx1067>
- Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* 138: 963–971.
- Churchill, G. A., D. M. Gatti, S. C. Munger and K. L. Svenson, 2012 The Diversity Outbred mouse population. *Mamm. Genome* 23: 713–718. <https://doi.org/10.1007/s00335-012-9414-2>
- Civelek, M., Y. Wu, C. Pan, C. K. Raulerson, A. Ko *et al.*, 2017 Genetic regulation of adipose gene expression and cardio-metabolic traits. *Am. J. Hum. Genet.* 100: 428–443. <https://doi.org/10.1016/j.ajhg.2017.01.027>

- Coffey, A. R., T. L. Smallwood, J. Albright, K. Hua, M. Kanke *et al.*, 2017 Systems genetics identifies a co-regulated module of liver microRNAs associated with plasma LDL cholesterol in murine diet-induced dyslipidemia. *Physiol. Genomics* 49: 618–629. <https://doi.org/10.1152/physiolgenomics.00050.2017>
- Coffey, A. R., M. Kanke, T. L. Smallwood, J. Albright, W. Pitman *et al.*, 2019 microRNA-146a-5p association with the cardiometabolic disease risk factor TMAO. *Physiol. Genomics* 51: 59–71. <https://doi.org/10.1152/physiolgenomics.00079.2018>
- Cox, D. R., and D. V. Hinkley, 1979 *Theoretical Statistics*. Taylor & Francis, Abingdon, UK. <https://doi.org/10.1201/b14832>
- Doss, S., E. E. Schadt, T. A. Drake, and A. J. Lusis, 2005 Cis-acting expression quantitative trait loci in mice. *Genome Res.* 15: 681–691. <https://doi.org/10.1101/gr.3216905>
- Edwards, S. L., J. Beesley, J. D. French, and A. M. Dunning, 2013 Beyond GWAS: illuminating the dark road from association to function. *Am. J. Hum. Genet.* 93: 779–797. <https://doi.org/10.1016/j.ajhg.2013.10.012>
- Frau, F., D. Crowther, H. Ruetten, and K. V. Allebrandt, 2017 Type-2 diabetes-associated variants with cross-trait relevance: post-GWAS strategies for biological function interpretation. *Mol. Genet. Metab.* 121: 43–50. <https://doi.org/10.1016/j.ymgme.2017.03.004>
- French, J. E., D. M. Gatti, D. L. Morgan, G. E. Kissling, K. R. Shockley *et al.*, 2015 Diversity Outbred mice identify population-based exposure thresholds and genetic factors that influence benzene-induced genotoxicity. *Environ. Health Perspect.* 123: 237–245. <https://doi.org/10.1289/ehp.1408202>
- Fu, J., J. J. Keurentjes, H. Bouwmeester, T. America, F. W. Verstappen *et al.*, 2009 System-wide molecular evidence for phenotypic buffering in Arabidopsis. *Nat. Genet.* 41: 166–167. <https://doi.org/10.1038/ng.308>
- Gatti, D. M., K. L. Svenson, A. Shabalina, L. Y. Wu, W. Valdar *et al.*, 2014 Quantitative trait locus mapping methods for Diversity Outbred mice. *G3 (Bethesda)* 4: 1623–1633. <https://doi.org/10.1534/g3.114.013748>
- Huan, T., J. Rong, C. Liu, X. Zhang, K. Tanriverdi *et al.*, 2015 Genome-wide identification of microRNA expression quantitative trait loci. *Nat. Commun.* 6: 6601. <https://doi.org/10.1038/ncomms7601>
- Huang, J., W. Sun, H. Huang, J. Ye, W. Pan *et al.*, 2014 miR-34a modulates angiotensin II-induced myocardial hypertrophy by direct inhibition of ATG9A expression and autophagic activity. *PLoS One* 9: e94382. <https://doi.org/10.1371/journal.pone.0094382>
- Huda, M. N., M. VerHague, J. Albright, T. Smallwood, T. A. Bell *et al.*, 2020 Dissecting the genetic architecture of cystatin C in Diversity Outbred mice. *G3 (Bethesda)* 10: 2529–2541. <https://doi.org/10.1534/g3.120.401275>
- Kanke, M., J. Baran-Gale, J. Villanueva, and P. Sethupathy, 2016 miRquant 2.0: an expanded tool for accurate annotation and quantification of microRNAs and their isomiRs from small RNA-sequencing data. *J. Integr. Bioinform.* 13: 307.
- Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* 477: 289–294. <https://doi.org/10.1038/nature10413>
- Keele, G. R., B. C. Quach, J. W. Israel, G. A. Chappell, L. Lewis *et al.*, 2020 Integrative QTL analysis of gene expression and chromatin accessibility identifies multi-tissue patterns of genetic regulation. *PLoS Genet.* 16: e1008537. <https://doi.org/10.1371/journal.pgen.1008537>
- Keller, M. P., D. M. Gatti, K. L. Schueler, M. E. Rabaglia, D. S. Stapleton *et al.*, 2018 Genetic drivers of pancreatic islet function. *Genetics* 209: 335–356. <https://doi.org/10.1534/genetics.118.300864>
- Kemis, J. H., V. Linke, K. L. Barrett, F. J. Boehm, L. L. Traeger *et al.*, 2019 Genetic determinants of gut microbiota composition and bile acid profiles in mice. *PLoS Genet.* 15: e1008073. <https://doi.org/10.1371/journal.pgen.1008073>
- King, E. G., and A. D. Long, 2017 The beavis effect in next-generation mapping panels in *Drosophila melanogaster*. *G3 (Bethesda)* 7: 1643–1652. <https://doi.org/10.1534/g3.117.041426>
- Kover, P. X., W. Valdar, J. Trakalo, N. Scarcelli, I. M. Ehrenreich *et al.*, 2009 A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5: e1000551. <https://doi.org/10.1371/journal.pgen.1000551>
- Leung, A., B. W. Parks, J. Du, C. Trac, R. Setten *et al.*, 2014 Open chromatin profiling in mice livers reveals unique chromatin variations induced by high fat diet. *J. Biol. Chem.* 289: 23557–23567. <https://doi.org/10.1074/jbc.M114.581439>
- Leung, A., C. Trac, J. Du, R. Natarajan, and D. E. Schones, 2016 Persistent chromatin modifications induced by high fat diet. *J. Biol. Chem.* 291: 10446–10455. <https://doi.org/10.1074/jbc.M115.711028>
- Lilue, J., A. G. Doran, I. T. Fiddes, M. Abrudan, J. Armstrong *et al.*, 2018 Sixteen diverse laboratory mouse reference genomes define strain-specific haplotypes and novel functional loci. *Nat. Genet.* 50: 1574–1583. <https://doi.org/10.1038/s41588-018-0223-8>
- Lloyd-Jones, L. R., A. Holloway, A. McRae, J. Yang, K. Small *et al.*, 2017 The genetic architecture of gene expression in peripheral blood. *Am. J. Hum. Genet.* 100: 228–237 (erratum: *Am. J. Hum. Genet.* 100: 371). <https://doi.org/10.1016/j.ajhg.2016.12.008>
- Mackay, T. F., S. Richards, E. A. Stone, A. Barbadilla, J. F. Ayroles *et al.*, 2012 The *Drosophila melanogaster* genetic reference panel. *Nature* 482: 173–178. <https://doi.org/10.1038/nature10811>
- Manichaikul, A., J. Dupuis, S. Sen, and K. W. Broman, 2006 Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174: 481–489. <https://doi.org/10.1534/genetics.106.061549>
- Mehrabian, M., H. Allayee, J. Stockton, P. Y. Lum, T. A. Drake *et al.*, 2005 Integrating genotypic and expression data in a segregating mouse population to identify 5-lipoxygenase as a susceptibility gene for obesity and bone traits. *Nat. Genet.* 37: 1224–1233 (erratum: *Nat. Genet.* 37: 1381). <https://doi.org/10.1038/ng1619>
- Morris, A. P., B. F. Voight, T. M. Teslovich, T. Ferreira, A. V. Segre *et al.*, 2012 Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* 44: 981–990. <https://doi.org/10.1038/ng.2383>
- Munger, S. C., N. Raghupathy, K. Choi, A. K. Simons, D. M. Gatti *et al.*, 2014 RNA-seq alignment to individualized genomes improves transcript abundance estimates in multiparent populations. *Genetics* 198: 59–73. <https://doi.org/10.1534/genetics.114.165886>
- Nikpay, M., K. Beehler, A. Valsesia, J. Hager, M. E. Harper *et al.*, 2019 Genome-wide identification of circulating-miRNA expression quantitative trait loci reveals the role of several miRNAs in the regulation of cardiometabolic phenotypes. *Cardiovasc. Res.* 115: 1629–1645. <https://doi.org/10.1093/cvr/cvz030>
- Ozcan, L., J. Cristina de Souza, A. A. Harari, J. Backs, E. N. Olson *et al.*, 2013 Activation of calcium/calmodulin-dependent protein kinase II in obesity mediates suppression of hepatic insulin signaling. *Cell Metab.* 18: 803–815. <https://doi.org/10.1016/j.cmet.2013.10.011>
- Parks, B. W., E. Nam, E. Org, E. Kostem, F. Norheim *et al.*, 2013 Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metab.* 17: 141–152. <https://doi.org/10.1016/j.cmet.2012.12.007>
- Pasricha, S., J. Kenney-Hunt, K. Anderson, N. Jafari, R. A. Hall *et al.*, 2015 Identification of eQTLs for hepatic Xbp1s and Socs3 gene expression in mice fed a high-fat, high-calorie diet. *G3 (Bethesda)* 5: 487–496. <https://doi.org/10.1534/g3.115.016626>
- Rhee, E. P., J. E. Ho, M. H. Chen, D. Shen, S. Cheng *et al.*, 2013 A genome-wide association study of the human metabolome in a

- community-based cohort. *Cell Metab.* 18: 130–143. <https://doi.org/10.1016/j.cmet.2013.06.013>
- Rockman, M. V., and L. Kruglyak, 2006 Genetics of global gene expression. *Nat. Rev. Genet.* 7: 862–872. <https://doi.org/10.1038/nrg1964>
- Rottiers, V., and A. M. Naar, 2012 MicroRNAs in metabolism and metabolic disorders. *Nat. Rev. Mol. Cell Biol.* 13: 239–250. <https://doi.org/10.1038/nrm3313>
- Rutledge, H., J. Baran-Gale, F. P. de Villena, E. J. Chesler, G. A. Churchill *et al.*, 2015 Identification of microRNAs associated with allergic airway disease using a genetically diverse mouse population. *BMC Genomics* 16: 633. <https://doi.org/10.1186/s12864-015-1732-9>
- Schadt, E. E., S. A. Monks, T. A. Drake, A. J. Lusk, N. Che *et al.*, 2003 Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422: 297–302. <https://doi.org/10.1038/nature01434>
- Schadt, E. E., C. Molony, E. Chudin, K. Hao, X. Yang *et al.*, 2008 Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 6: e107. <https://doi.org/10.1371/journal.pbio.0060107>
- Smallwood, T. L., D. M. Gatti, P. Quizon, G. M. Weinstock, K. C. Jung *et al.*, 2014 High-resolution genetic mapping in the diversity outbred mouse population identifies Apoe1 as a candidate gene for atherosclerosis. *G3 (Bethesda)* 4: 2353–2363. <https://doi.org/10.1534/g3.114.014704>
- Storey, J. D., J. M. Akey, and L. Kruglyak, 2005 Multiple locus linkage analysis of genomewide expression in yeast. *PLoS Biol.* 3: e267. <https://doi.org/10.1371/journal.pbio.0030267>
- Su, W. L., R. R. Kleinhanz, and E. E. Schadt, 2011 Characterizing the role of miRNAs within gene regulatory networks using integrative genomics techniques. *Mol. Syst. Biol.* 7: 490. <https://doi.org/10.1038/msb.2011.23>
- Svenson, K. L., D. M. Gatti, W. Valdar, C. E. Welsh, R. Cheng *et al.*, 2012 High-resolution genetic mapping using the mouse Diversity Outbred population. *Genetics* 190: 437–447. <https://doi.org/10.1534/genetics.111.132597>
- Tan, A., J. Sun, N. Xia, X. Qin, Y. Hu *et al.*, 2012 A genome-wide association and gene-environment interaction study for serum triglycerides levels in a healthy Chinese male population. *Hum. Mol. Genet.* 21: 1658–1664. <https://doi.org/10.1093/hmg/ddr587>
- Tyler, A. L., B. Ji, D. M. Gatti, S. C. Munger, G. A. Churchill *et al.*, 2017 Epistatic networks jointly influence phenotypes related to metabolic disease and gene expression in Diversity Outbred mice. *Genetics* 206: 621–639. <https://doi.org/10.1534/genetics.116.198051>
- West, D. B., C. N. Boozer, D. L. Moody, and R. L. Atkinson, 1992 Dietary obesity in nine inbred mouse strains. *Am. J. Physiol.* 262: R1025–R1032.
- Winter, J. M., D. E. Gildea, J. P. Andreas, D. M. Gatti, K. A. Williams *et al.*, 2017 Mapping complex traits in a diversity outbred F1 mouse population identifies germline modifiers of metastasis in human prostate cancer. *Cell Syst.* 4: 31–45.e6. <https://doi.org/10.1016/j.cels.2016.10.018>
- Yang, H., T. A. Bell, G. A. Churchill, and F. Pardo-Manuel de Villena, 2007 On the subspecific origin of the laboratory mouse. *Nat. Genet.* 39: 1100–1107. <https://doi.org/10.1038/ng2087>
- Yang, H., J. R. Wang, J. P. Didion, R. J. Buus, T. A. Bell *et al.*, 2011 Subspecific origin and haplotype diversity in the laboratory mouse. *Nat. Genet.* 43: 648–655. <https://doi.org/10.1038/ng.847>
- Yang, J., T. Ferreira, A. P. Morris, S. E. Medland; Genetic Investigation of Anthropometric Traits (GIANT) Consortium; Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium *et al.*, 2012 Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* 44: 369–375, S1–3.
- Yang, J., N. A. Zaitlen, M. E. Goddard, P. M. Visscher, and A. L. Price, 2014 Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* 46: 100–106. <https://doi.org/10.1038/ng.2876>
- Yang, X., J. L. Deignan, H. Qi, J. Zhu, S. Qian *et al.*, 2009 Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nat. Genet.* 41: 415–423. <https://doi.org/10.1038/ng.325>
- Yao, C., R. Joeanes, A. D. Johnson, T. Huan, C. Liu *et al.*, 2017 Dynamic role of trans regulation of gene expression in relation to complex traits. *Am. J. Hum. Genet.* 100: 985–986. <https://doi.org/10.1016/j.ajhg.2017.05.002>
- Zhu, Z., F. Zhang, H. Hu, A. Bakshi, M. R. Robinson *et al.*, 2016 Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* 48: 481–487. <https://doi.org/10.1038/ng.3538>

Communicating editor: D. Nielsen