# Genetic, epigenetic, and environmental mechanisms govern allele-specific gene expression

Celine L. St. Pierre,[1] Juan F. Macias-Velasco,[1] Jessica P. Wayhart,[1] Li Yin,[2] Clay F. Semenkovich,[2] and Heather A. Lawson[1]

[1]Department of Genetics, [2]Department of Medicine, Washington University School of Medicine, St. Louis, Missouri 63110, USA

Allele-specific expression (ASE) is a phenomenon in which one allele is preferentially expressed over the other. Genetic and epigenetic factors cause ASE by altering the final composition of a gene's product, leading to expression imbalances that can have functional consequences on phenotypes. Environmental signals also impact allele-specific expression, but how they contribute to this cross talk remains understudied. Here, we explored how genotype, parent-of-origin, tissue, sex, and dietary fat simultaneously influence ASE biases. Male and female mice from a $F_1$ reciprocal cross of the LG/J and SM/J strains were fed a high or low fat diet. We harnessed strain-specific variants to distinguish between two ASE classes: parent-of-origin-dependent (unequal expression based on parental origin) and sequence-dependent (unequal expression based on nucleotide identity). We present a comprehensive map of ASE patterns in 2853 genes across three tissues and nine environmental contexts. We found that both ASE classes are highly dependent on tissue and environmental context. They vary across metabolically relevant tissues, between males and females, and in response to dietary fat. We also found 45 genes with inconsistent ASE biases that switched direction across tissues and/or environments. Finally, we integrated ASE and QTL data from published intercrosses of the LG/J and SM/J strains. Our ASE genes are often enriched in QTLs for metabolic and musculoskeletal traits, highlighting how this orthogonal approach can prioritize candidate genes. Together, our results provide novel insights into how genetic, epigenetic, and environmental mechanisms govern allele-specific expression, which is an essential step toward deciphering the genotype-to-phenotype map.

[Supplemental material is available for this article.]

Deciphering the genotype-to-phenotype map remains a fundamental quest in biology. Gene expression is a promising focal point, because it is an intermediate step between DNA sequence and gross phenotype. Gene expression itself is a complex trait regulated by genetic, epigenetic, and environmental factors (Pastinen 2010). Recent efforts to characterize gene expression have often only interrogated changes in total expression levels and assumed that genes are biallelically expressed (i.e., both alleles are equally expressed), which may mask underlying regulatory mechanisms. Epigenetic changes and gene-by-environment interactions can also alter gene expression patterns in a dynamic and tissue-specific manner without modifying the underlying nucleotide sequence; these effects are missed in a typical sequence-based method. Thus, an allele-specific approach is required to directly measure how *cis*-regulatory variation impacts gene expression and to tease it apart from *trans*-acting factors that affect both chromosomes (Bonasio et al. 2010). Allele-specific expression (ASE) is a phenomenon in which a gene's expression diverges from biallelic. In diploid organisms, one allele is preferentially expressed over the other. Previous findings estimate that 30%–56% of genes show evidence of allelic imbalance, indicating allele-specific effects have widespread impacts on gene regulation (Ge et al. 2009; Keane et al. 2011; Castel et al. 2020). Depending on a gene's function, these expression imbalances can lead to phenotypic variation with functional consequences.

ASE can be divided into two classes: sequence or parent-of-origin (Fig. 1). Sequence-dependent ASE occurs when the alleles are differentially expressed based on their haplotype or nucleotide identity. These patterns are thought to be largely driven by *cis*-acting genetic variants in coding and noncoding regions, such as a premature stop codon that truncates one allele's transcript or a variant in a promoter region that prevents transcription factors from binding (Keane et al. 2011; Rivas et al. 2015). They can also occur further away from the gene yet still impact its expression, such as motif variations for long-range enhancers or in the sequence context of DNA methyltransferase substrates (Wienholz et al. 2010; Cavalli et al. 2016). Stochastic transcriptional bursts can also cause transient fluctuations in which allele is expressed, resulting in random monoallelic expression that is inconsistent across cells or organisms (Deng et al. 2014; Reinius and Sandberg 2015).

In contrast, parent-of-origin-dependent ASE occurs when the alleles are differentially expressed based on which parent contributed it, regardless of the underlying sequence. These patterns are examples of parent-of-origin effects (POEs), a broader class of epigenetic phenomena that manifest as phenotypic differences according to maternal or paternal inheritance (Lawson et al. 2013). The best characterized POE mechanism is genomic imprinting, an extreme case of parent-of-origin-dependent ASE in which one parent's allele is silenced via selective epigenetic marks, such as DNA methylation and histone modifications (Umlauf et al. 2004; Barlow and Bartolomei 2014; Inoue et al. 2017). Known imprinted genes compose ~1% of the human and mouse genomes, yet play important roles in development, metabolism, and cognition (Reik and Walter 2001). Partial imprinting is another, more subtle, form of parent-of-origin-dependent ASE that has been
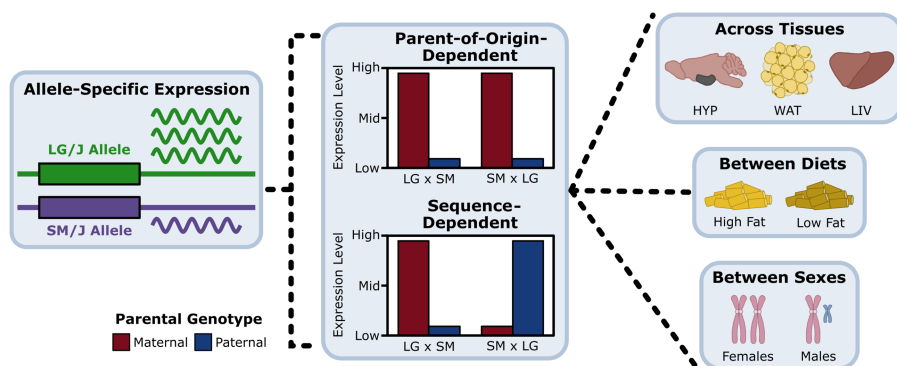
**Figure 1.** Evaluating ASE across tissues and environmental contexts. We partitioned ASE into its parent-of-origin and sequence effects, then compared ASE patterns across metabolic tissues, in response to dietary fat and between sexes. An example of parent-of-origin-dependent ASE is when the maternal allele (red) is preferentially expressed over the paternal allele (blue), regardless of which haplotype contributed it. An example of sequence-dependent ASE is when the LG/J allele is preferentially expressed over the SM/J allele, regardless of which parent contributed it.

documented but is less understood (Wolf et al. 2008; Morcos et al. 2011). Random monoallelic expression can also produce transient and nonheritable parental biases in individual cells through different mechanisms than imprinting (Gimelbrant et al. 2007; Morcos et al. 2011). Other sources of POEs include maternal effects and gene-specific trinucleotide expansions, but these do not involve ASE (Hager et al. 2008).

Comprehensive atlases of how ASE patterns vary between tissues and developmental stages have been generated in human and mouse models (Babak et al. 2015; Leung et al. 2015; Andergassen et al. 2017; Castel et al. 2020). These studies reveal that both parental and sequence biases are not always consistent between tissues, indicating that tissue-specific genetic and epigenetic features can mediate allelic imbalances. Additionally, *trans*-acting environmental factors have been shown to interact with *cis*-regulatory variants to modulate the magnitude of ASE effects in human and rice models (Buil et al. 2015; Moyerbrailean et al. 2016; Knowles et al. 2017; Shao et al. 2019). Imprinted genes can also be responsive to environmental exposures during fetal development (Kappil et al. 2015). Together, these findings suggest that a complicated cross talk between genetic variants, epigenetic changes, and environmental signals underlies allele-specific gene regulation, but this model needs to be further investigated.

Here, we explored how allelic genotype, parent-of-origin, tissue type, sex, and dietary fat simultaneously work together to influence ASE patterns in a $F_1$ reciprocal cross of the LG/J and SM/J inbred mouse strains. These strains are uniquely suited for gene-by-environment studies owing to their divergent genetic backgrounds and variable responses to dietary nutrition (Ehrich et al. 2003; Nikolskiy et al. 2015; Miranda et al. 2019; Carson and Lawson 2020). They are also the progenitors of the longest-running advanced intercross line (AIL). An AIL is a multigenerational population produced by intercrossing two inbred strains beyond the $F_2$ generation while minimizing relatedness between mates (Darvasi and Soller 1995). Since its inception, the LG/J x SM/J AIL has been extensively used to map quantitative trait loci (QTL) related to metabolic, musculoskeletal, behavioral, and physiological traits (Ehrich et al. 2005; Lawson et al. 2010, 2011a,b; Lionikas et al. 2010; Cheverud et al. 2011; Parker et al. 2014; Gonzales et al. 2018; Cordero et al. 2019). To probe whether ASE imbalances contribute to complex traits, we integrated our ASE re-

sults with these AIL mapping studies and highlighted how this orthogonal approach can prioritize candidate genes. Untangling the genetic, epigenetic, and environmental mechanisms that govern allele-specific gene regulation is crucial to improving our ability to predict phenotypes from genotypes.

## Results

### Partitioning allele-specific expression into parent-of-origin and sequence effects

We measured ASE in a $F_1$ reciprocal cross of the LG/J and SM/J inbred strains. Briefly, LG/J mothers were mated with SM/J fathers and vice versa, resulting in $F_1$ offspring who are genetically equivalent but differ in the allelic direction of inheritance. Male and female $F_1$ mice were fed either a high or low fat diet. We obtained RNA-seq data from three metabolically relevant tissues: hypothalamus (HYP), white adipose (WAT), and liver (LIV). We analyzed nine environmental contexts per tissue: high fat (H), low fat (L), females (F), males (M), high fat females (HF), high fat males (HM), low fat females (LF), low fat males (LM), and all contexts collapsed (All). Together, these 27 "tissue-by-context" analyses (three tissues × nine contexts) allowed us to explore how tissue type and environmental signals influence ASE (Fig. 1).

The LG/J and SM/J genomes differ by >5 million single-nucleotide polymorphisms (SNPs) and indels (Nikolskiy et al. 2015). To be informative for ASE, a strain-specific variant must be located within a gene's exome. In this $F_1$ model, informative variants are located in 55% of expressed genes in our three tissues and 47% of all genes in the mm10 reference genome (Supplemental Fig. S1). We harnessed those strain-specific variants to map sequencing reads to their chromosome of origin. Overall, 9016 genes had detectable ASE in at least one tissue-by-context analysis (~37% of all expressed genes) (Supplemental Fig. S2). Next, we identified 2853 genes with significant ASE biases (~6%) and classified them into two patterns: parent-of-origin-dependent (unequal expression based on parental origin) and sequence-dependent (unequal expression based on nucleotide identity) (Fig. 1; Supplemental Table S1).

### Both classes of ASE patterns are prevalent and distinct

Across our 27 tissue-by-context analyses, we identified 271 genes with significant parent-of-origin-dependent ASE (Supplemental Tables S2–S4). HYP had the greatest number of genes (229), followed by WAT (35), then LIV (27). Fourteen genes were expressed in multiple tissues, but the majority were tissue-specific (Fig. 2A). In HYP, the most genes were detected in the "All" context. However, in WAT and LIV, more genes were only detected in a specific diet, sex, and/or diet-by-sex context; those expression biases were missed when contexts were collapsed (Fig. 2B). Two hundred fourteen genes (79%) were paternally biased, 55 genes (20%) were maternally biased, and two genes (1%) switched their ASE bias direction across the cohorts (Fig. 2C). This heavy paternal skew is driven by a 672-kb cluster of 171 small nucleolar RNAs
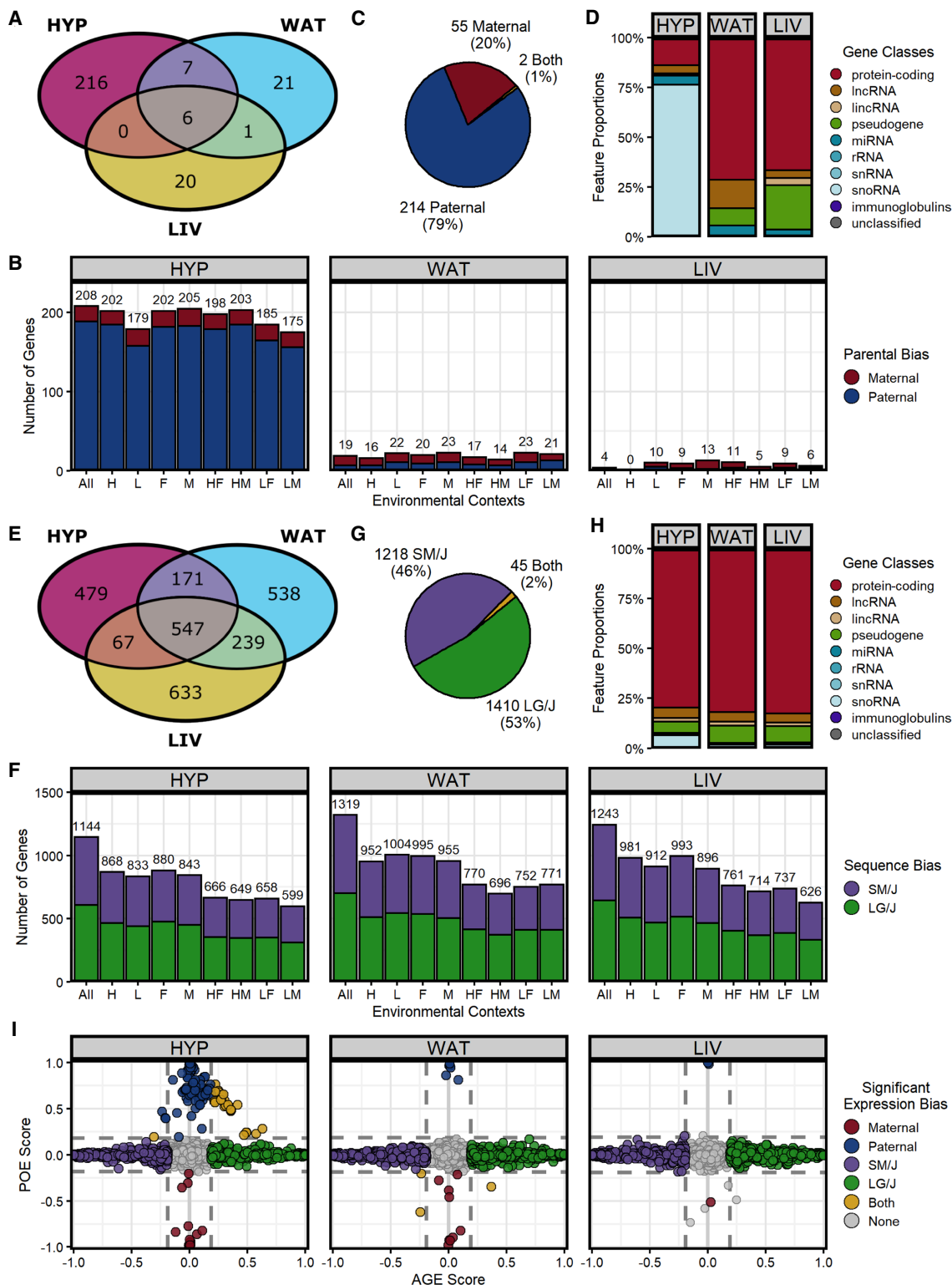
**Figure 2.** Both classes of ASE patterns are prevalent and distinct. (*A*) Venn diagram of all parent-of-origin-dependent ASE genes across tissues. (*B*) Number of significant parentally biased genes in each tissue-by-context analysis (maternal, red; paternal, blue): (All) all contexts; (H) high fat; (L) low fat; (F) females; (M) males; (HF) high fat females; (HM) high fat males; (LF) low fat females; and (LM) low fat males. (*C*) Summary of ASE biases across all analyses. (*D*) Proportions of gene classes in each tissue. (*E*) Venn diagram of all sequence-dependent ASE genes across tissues. (*F*) Number of significant sequence-biased genes in each tissue-by-context analysis (SM/J, purple; LG/J, green). (*G*) Summary of ASE biases across all analyses. (*H*) Proportions of gene classes in each tissue. (*I*) Parent-of-origin effect (POE) versus allelic genotype effect (AGE) scores in the "All" context of each tissue. Dots represent individual genes and are color coded by their ASE bias direction: (red) maternal; (blue) paternal; (purple) SM/J; (green) LG/J; and (yellow) both ASE classes. Most genes have no bias (gray). Dashed lines indicate significant score thresholds.

(snoRNAs) in the Prader-Willi/Angelman syndrome (PWS/AS) domain on Chromosome 7 that are only expressed in HYP (Supplemental Fig. S3). Parentally biased ASE genes also included protein-coding genes, microRNAs, long noncoding RNAs, long interspersed noncoding RNAs, and pseudogenes (Fig. 2D).

We identified 2673 genes with significant sequence-dependent ASE across our 27 tissue-by-context analyses (Supplemental Tables S2–S4). WAT had the greatest number of genes (1495), followed by LIV (1486), then HYP (1264). Although some genes' biases were tissue-specific, 1657 genes (62%) were biased in multiple tissues (Fig. 2E). In all tissues, the most genes were detected in the "All" context, then the diet or sex contexts, and finally the diet-by-sex contexts, likely reflecting the sample sizes and power available in each cohort (Fig. 2F). As shown in Figure 2G, 1218 genes (46%) were SM/J biased, 1410 genes (53%) were LG/J biased, and 45 genes (2%) switched their expression bias direction across the cohorts. Sequence-biased ASE genes were predominantly protein-coding genes (~80%) but also included pseudogenes, immunoglobulins, and various noncoding RNAs (long noncoding, long interspersed noncoding, micro, ribosomal, small interfering, and small nucleolar) (Fig. 2H).

Parent-of-origin and sequence-dependent ASE patterns were often mutually exclusive. In each tissue, genes with extreme parental biases typically had weak sequence biases and vice versa (Fig. 2I; Supplemental Fig. S4). However, 61 genes in HYP and three genes in WAT had both parental and sequence biases, likely because of epigenetic regulatory mechanisms affected by allelic variation. Both ASE patterns occurred genome wide (Supplemental Fig. S5). Parentally biased genes tended to cluster in known imprinted regions, like the *Ube3a/Snrpn*, *Meg3*, *Peg3/Usp29*, and *H13/Mcts2* domains (Supplemental Fig. S6). Sequence-biased genes were spread diffusely across the genome, but occasionally clustered in discrete regions potentially controlled by the same regulatory element (Supplemental Fig. S7).

## Parent-of-origin-dependent ASE recapitulates canonical imprinting patterns

To evaluate whether parent-of-origin-dependent ASE is influenced by tissue or environmental context, we characterized the expression profiles of those 271 genes across our 27 tissue-by-context analyses. For each analysis, we quantified the direction and magnitude of each gene's parental bias by calculating a POE score from the mean allelic bias of each $F_1$ reciprocal cross. POE scores ranged from −1 (completely maternally expressed) to +1 (completely paternally expressed); a score of 0 indicated biallelic expression (Methods). In each analysis, a gene could be expressed in three possible ways: significant parental bias, biallelic (expressed but with no allelic bias), or not expressed. We sorted the 271 parentally biased ASE genes into three expression profiles: tissue-independent, tissue-dependent, and context-dependent (Fig. 3A–D; Supplemental Fig. S3).

We identified 183 tissue-independent genes (67%), defined as a consistent parental bias in every tissue they were expressed. Within a given tissue, they were parentally biased in most environmental contexts (Fig. 3B). Next, we identified 15 genes (6%) as tissue-dependent. These genes showed a parental bias across most contexts in one or two tissues, but were biallelically expressed (no bias) across contexts in the other tissue(s) (Fig. 3C). Here, we distinguish between tissue-specific gene expression (not expressed in certain tissues) and tissue-dependent ASE (biased expression only in certain tissues).

All 198 tissue-independent/-dependent genes were related to genomic imprinting, the best characterized mechanism of parent-of-origin-dependent ASE. Twenty-nine genes were canonically imprinted (bolded in Fig. 3, defined in Methods), and the other 169 genes were various noncoding RNAs located within known imprinted domains. For example, 156 genes were in a cluster of 171 paternally biased snoRNAs in the PWS/AS domain on Chromosome 7 (Supplemental Fig. S3). As expected with imprinting, these genes were extremely biased toward one parent's allele; their mean POE scores were −0.82 and 0.71 for maternally and paternally biased genes, respectively (Fig. 3E). HYP had the highest proportion of these genes (85%) among our three adult tissues, even if the aforementioned HYP-specific snoRNA cluster is excluded (52%). WAT had the second highest proportion (49%), but few of these genes were expressed in LIV (26%) (Fig. 3A). These findings are consistent with the previously reported dichotomy of imprinting levels between neural and non-neural adult tissues (Babak et al. 2015) as well as imprinting's role in development and cognition (Barlow and Bartolomei 2014). Furthermore, our data replicated 25 of 96 genes previously found to be imprinted in these tissues (Babak et al. 2015), comprising nearly all bolded genes in Figure 3B–D. Five genes had sequence biases instead and 11 genes were biallelically expressed, potentially a result of methodology differences between the studies. We could not compare the imprinting status of the remaining 55 genes because 34 had no informative variants in our $F_1$ model and 21 no longer appeared in the mm10 reference assembly.

We validated the expression profiles of two canonically imprinted genes (*Peg3* and *Grb10*) by pyrosequencing (Supplemental Fig. S8). Paternally expressed 3 gene (*Peg3*) was paternally biased in all three tissues, regardless of context. Its locus had 60 strain-specific variants, but none were predicted as functional. *Peg3* functions as a DNA-binding transcriptional repressor to control fetal growth rates, maternal caring behaviors, and tumor growth. It is only expressed from the paternal allele in most tissues, especially the placenta and brain (Thiaville et al. 2013; He and Kim 2014). Growth factor receptor bound protein 10 (*Grb10*) was paternally biased in HYP but maternally biased in WAT and LIV. Its locus had 154 strain-specific variants, but none were predicted as functional. *Grb10* encodes an adapter protein that interacts with receptor tyrosine kinases to impact insulin signaling and growth hormone pathways (He et al. 1998). It has a documented pattern of maternal expression in most adult mouse tissues, but paternal expression in the brain (Plasschaert and Bartolomei 2015).

## Dietary environment and sex influence parent-of-origin-dependent ASE in a partially imprinted manner

Finally, we classified 73 genes (27%) as context-dependent, meaning they had a parental bias only in certain environmental contexts within a tissue, but biallelic expression (no bias) in other contexts and/or tissues (Fig. 3D). Only three genes were canonically imprinted, and 18 genes were noncoding RNAs located in known imprinted domains (including 14 snoRNAs in the PWS/AS domain). The remaining 52 genes had no clear connection to genomic imprinting, yet they showed significant parent-of-origin-dependent ASE in certain context(s). These genes had more subtle allelic biases than the tissue-independent/-dependent genes; their mean POE scores were −0.39 and 0.39 for maternally and paternally biased genes, respectively (Fig. 3E). These patterns are consistent with partial imprinting, in which the two parental alleles are differentially expressed in a less extreme manner than
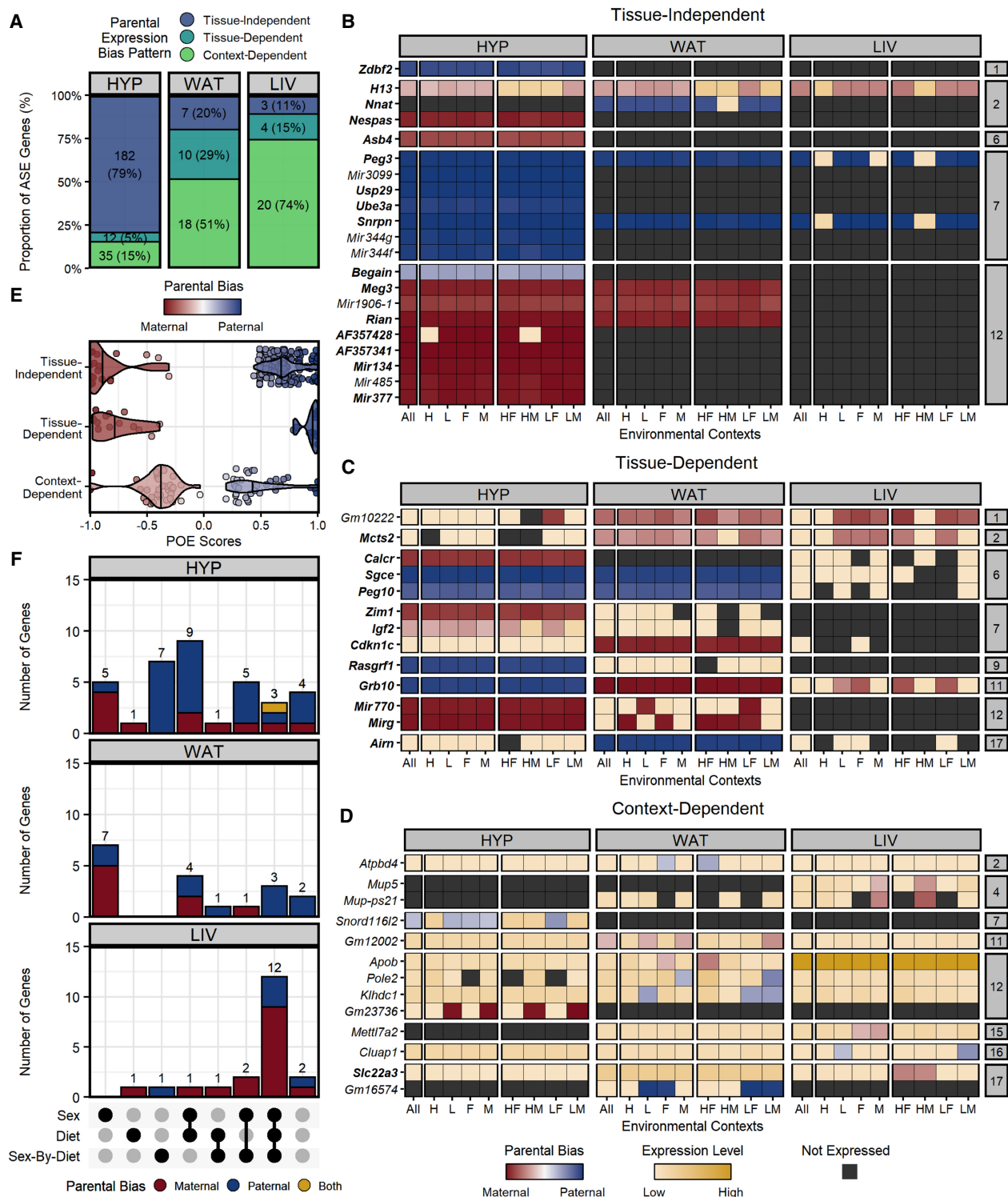
**Figure 3.** Parent-of-origin-dependent ASE patterns fall into three expression profiles. (*A*) Proportion of ASE genes per tissue with each parental bias profile. Heatmaps of ASE profiles across analyses: (*B*) tissue-independent, (*C*) tissue-dependent, and (*D*) context-dependent. A subset of the 271 genes is shown, including those validated with pyrosequencing. Genes are color coded by their expression pattern in each tissue-by-context analysis. Shades of red and blue indicate their degree of maternal or paternal bias, respectively (POE scores). If genes are not biased, shades of yellow indicate their biallelic expression levels (log-transformed total counts). Black indicates genes are not expressed. Bolded genes are canonically imprinted. The *y*-axis is grouped and sorted by chromosomal position. Supercolumns denote tissues: (HYP) hypothalamus; (WAT) white adipose; and (LIV) liver. Subcolumns denote environmental contexts: (All) all contexts; (H) high fat; (L) low fat; (F) females; (M) males; (HF) high fat females; (HM) high fat males; (LF) low fat females; and (LM) low fat males. (*E*) POE scores for each parental bias profile. Vertical lines indicate mean POE scores. Dots represent individual ASE genes. (*F*) UpSet plots of the significant sex, diet, and/or sex-by-diet effects of context-dependent genes in each tissue. Bar height and color indicate how many genes with each parental bias: (red) maternal; (blue) paternal; and (yellow) direction switching.

the uniparental expression associated with genomic imprinting (Wolf et al. 2008; Morcos et al. 2011). LIV had the highest proportion of these genes among the three tissues (74%). WAT had the second highest proportion (51%), whereas HYP had the lowest proportion (15%) (Fig. 3A). LIV is also a tissue where more parentally biased genes were detected in the diet- and/or sex-specific contexts than the "All" context (Fig. 2B). LIV is responsive to environmental factors, especially diet, given its central roles in digestion and detoxification (Trefts et al. 2017). Taken together, these findings suggest a mechanism of parent-of-origin-dependent ASE outside of traditional imprinting that is sensitive to environmental perturbations.

To further explore how sex and/or dietary fat can alter parental biases, we calculated individualized POE scores for each context-dependent gene and modeled how they vary across diet, sex, and diet-by-sex contexts (Supplemental Tables S5, S6). These categories were not mutually exclusive; across all three tissues, most genes were significant for more than one effect. Nonetheless, when we intersected these significant gene lists, we found that each tissue showed a distinct pattern of context-dependent parental biases (Fig. 3F; Supplemental Fig. S9). For example, LIV had similar proportions of significant sex, diet, and diet-by-sex effects (each 75%–80%); 60% of its genes were significant for all three effects. Most of WAT's genes had a significant sex effect (83%), but diet or diet-by-sex effects were much less common (44% and 28%, respectively). Finally, 63% of HYP's genes had a significant sex effect, whereas 40%–45% had significant diet or diet-by-sex effects.

We validated the expression profiles of two context-dependent genes (*Apob* and *Slc22a3*) by pyrosequencing (Supplemental Fig. S10). Apolipoprotein B (*Apob*) had a significant diet effect in WAT, reflected in maternal biases in the HF and F contexts. *Apob* was biallelically expressed in the remaining WAT contexts and all contexts in LIV and HYP. Its locus had 126 variants, including seven nonsynonymous SNPs (six, LG/J genome). *Apob* produces the main component of lipoproteins, which transport lipids in the blood (Olofsson and Borèn 2005). Maternal-specific associations between *APOB* variants and adiposity traits have been found in humans (Hochner et al. 2015). *Apob* expression also differs between high and low fructose diets in mice livers (Sud et al. 2017). Solute carrier family 22 (organic cation transporter), member 3 (*Slc22a3*) had significant diet and diet-by-sex effects in LIV, reflected in maternal biases in the HF and HM contexts. *Slc22a3* was biallelically expressed in the remaining LIV contexts and all contexts in WAT and HYP. Its locus had 285 strain-specific variants, but none were predicted as functional. *Slc22a3* is a transporter that eliminates organic cations from cells (Kekuda et al. 1998). It has been reported as maternally expressed in liver and extraembryonic tissues, but biallelic elsewhere (Babak et al. 2015). *Slc22a3* is also differentially expressed in kidneys between high fat and chow-fed mice (Gai et al. 2016).

### Sequence-dependent ASE arises from haplotype-specific genetic variation

Next, we similarly characterized the expression profiles of the 2673 sequence-biased genes across our 27 tissue-by-context analyses to evaluate whether sequence-dependent ASE is also influenced by tissue or environmental context. For each analysis, we quantified the direction and magnitude of each gene's sequence bias by calculating an allelic genotype effect (AGE) score from the mean allelic bias of each $F_1$ reciprocal cross. AGE scores ranged from –1

(completely SM/J expressed) to +1 (completely LG/J expressed); a score of 0 indicated biallelic expression (Methods). In each analysis, a gene could be expressed in three possible ways: significant sequence/allelic genotype bias, biallelic (expressed but with no bias), or not expressed. We sorted the 2673 sequence-biased ASE genes into three expression profiles: tissue-independent, tissue-dependent, and context-dependent (Fig. 4A–D).

We identified 605 genes (23%) as tissue-independent, meaning they had a consistent sequence bias in every tissue they were expressed. Within a given tissue, they had a sequence bias in most environmental contexts (Fig. 4B). These genes were strongly biased toward one strain's allele; their mean AGE scores were –0.79 and 0.78 for SM/J and LG/J biased genes, respectively (Fig. 4E). All three tissues had a similar proportion of these genes (29%–35%) (Fig. 4A). These patterns likely reflect the vast genetic variation accumulated between the LG/J and SM/J backgrounds over the decades, whereby a *cis*-acting variant impacts one strain's allelic function wherever that gene is expressed, regardless of tissue type.

We validated the expression profiles of two tissue-independent genes (*Eef1a1* and *Tubb2a*) by pyrosequencing (Supplemental Fig. S11). Eukaryotic translation had a LG/J bias in all three tissues, regardless of context. Its locus had 30 strain-specific variants, including two nonsynonymous SNPs in the SM/J genome. *Eef1a1* delivers aminoacylated transfer RNAs to the elongating ribosome during protein synthesis and has crucial roles in protein degradation and other cellular processes. It is abundantly and ubiquitously expressed in most tissues (Mateyak and Kinzy 2010; Li et al. 2013). Tubulin, beta 2A class IIA (*Tubb2a*) had a SM/J bias in all three tissues, regardless of context. Mutations in this gene are rare and often nonviable; however, its locus had one SNP in the LG/J genome. *Tubb2a* is a tubulin isoform that binds GTP to create microtubules, which are critical for the cytoskeleton organization, intracellular trafficking, and mitotic cell division. It is also ubiquitously expressed in most tissues, especially the brain (Hammond et al. 2008; Rice et al. 2008).

### Sequence-dependent ASE can be mediated by tissue-specific features

Next, we identified 684 genes (25%) as tissue-dependent. These genes showed a sequence bias across most contexts in one or two tissues, but were biallelically expressed (no bias) across contexts in the other tissue(s) (Fig. 4C). Here, we again distinguish between tissue-dependent ASE (biallelic in some tissues) and tissue-specific gene expression (not expressed in some tissues). These genes were moderately biased toward one strain's allele; their mean AGE scores were –0.57 and 0.55 for SM/J and LG/J biased genes, respectively (Fig. 4E). All three tissues also had a similar proportion of these genes (27%–32%) (Fig. 4A). These patterns show that sequence-dependent ASE is not solely a result of genetic variation; here, tissue-specific epigenetic factors likely interact with *cis*-acting variants to influence allelic frequency.

We validated the expression profiles of two tissue-dependent genes (*Upp2* and *Mettl7b*) by pyrosequencing (Supplemental Fig. S12). Uridine phosphorylase 2 (*Upp2*) had a strong SM/J bias across contexts in HYP, a weaker SM/J bias in WAT, and biallelic expression in LIV. Its locus had 1043 strain-specific variants, including five nonsynonymous SNPs in the LG/J genome. *Upp2* catalyzes the phosphorolysis of uridine into uracil and ribose-1-phosphate during nucleoside metabolism (Johansson 2003). In mice, it is predominantly expressed in liver and weakly expressed in brain (Roosild et al. 2011). Methyltransferase like 7b (*Mettl7b*) had a
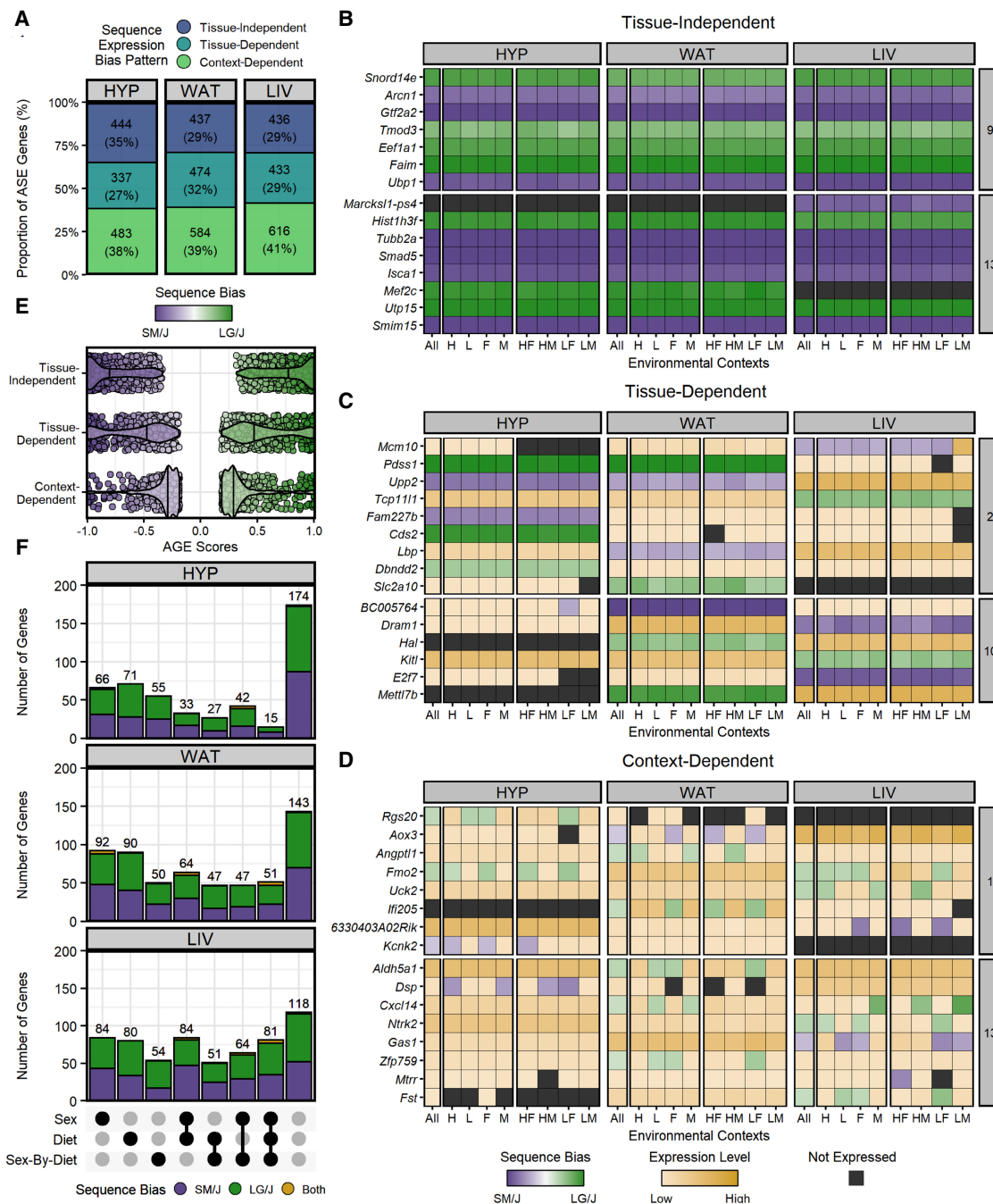
**Figure 4.** Sequence-dependent ASE patterns fall into three expression profiles. (*A*) Proportion of ASE genes per tissue with each sequence bias profile. Heatmaps of ASE profiles across analyses: (*B*) tissue-independent, (*C*) tissue-dependent, and (*D*) context-dependent. A subset of the 2673 genes are shown, including those validated with pyrosequencing. Genes are color coded by their expression pattern in each tissue-by-context analysis. Shades of purple and green indicate their degree of SM/J or LG/J bias, respectively (AGE scores). If genes are not biased, shades of yellow indicate their biallelic expression levels (log-transformed total counts). Black indicates genes are not expressed. The *y*-axis is grouped and sorted by chromosomal position. Supercolumns denote tissues: (HYP) hypothalamus; (WAT) white adipose; and (LIV) liver. Subcolumns denote environmental contexts: (All) all contexts; (H) high fat; (L) low fat; (F) females; (M) males; (HF) high fat females; (HM) high fat males; (LF) low fat females; and (LM) low fat males. (*E*) AGE scores for each sequence bias profile. Vertical lines indicate mean AGE scores. Dots represent individual ASE genes. (*F*) UpSet plots of the significant sex, diet, and/or sex-by-diet effects of context-dependent genes in each tissue. Bar height and color indicate how many genes with each sequence bias: (purple) SM/J; (green) LG/J; and (yellow) direction switching.

LG/J bias across contexts in WAT, biallelic expression in LIV, and no expression in HYP. Its locus had 77 strain-specific variants, including one nonsynonymous SNP in the SM/J genome. *Mettl7b* is

an akyl thiol methyltransferase implicated in several cancers (Maldonato et al. 2021). It is highly expressed in lipid droplets, particularly liver and adipose tissue (Turró et al. 2006).

## Sequence-dependent ASE is sensitive to sex and dietary environments

Finally, we classified 1384 genes (52%) as context-dependent. These genes had a sequence bias only in certain environmental contexts within a tissue, but biallelic expression (no bias) in other contexts and/or tissues (Fig. 4D). These genes had more subtle allelic biases than the other two profiles; their mean AGE scores were −0.32 and 0.34 for SM/J and LG/J biased genes, respectively (Fig. 4E). LIV had the most context-dependent genes (616) and HYP had the fewest (483), but overall all three tissues comprised a similar proportion of these genes (38%–41%) (Fig. 4A). These patterns suggest that environmental factors can interact with genetic variation to influence the final allelic composition of a gene's product.

To further explore how sex and/or dietary fat can alter sequence biases, we calculated individualized AGE scores for each context-dependent gene and modeled how they vary across diet, sex, and diet-by-sex contexts (Supplemental Tables S5, S6). When we intersected these gene lists, we found that each tissue showed a similar pattern of context-dependent sequence biases (Fig. 4F; Supplemental Fig. S13). Significant sex effects were the most prevalent in each tissue, ranging from 32% of genes in HYP, 43% in WAT, to 51% in LIV. Diet effects were slightly less common but still widespread: 30% of genes in HYP, 43% in WAT, and 48% in LIV. Finally, diet-by-sex effects were also pervasive in each tissue, comprising 29% of genes in HYP, 33% in WAT, and 41% in LIV. These categories were not mutually exclusive; most genes were significant for more than one effect across the three tissues.

We validated the expression profiles of two context-dependent genes (Ifi205 and Gas1) by pyrosequencing (Supplemental Fig. S14). Interferon activated gene 205 (Ifi205) had significant sex and diet-by-sex effects in WAT, reflected in strong LG/J biases in the three female-related (HF, LF, F) and "All" contexts. Ifi205 was biallelically expressed in the remaining WAT contexts and all LIV contexts. Its locus had 254 strain-specific variants, including 10 nonsynonymous SNPs in the SM/J genome. Ifi205 binds DNA in response to interferon signaling, which activates the immune system (Albrecht et al. 2005). Significant sex differences have been reported in mice, in which females have higher total Ifi205 expression levels than males (Cao et al. 2018). Growth arrest specific 1 (Gas1) had significant diet and sex effects in LIV, reflected in strong SM/J biases in the three low fat diet-related (LF, LM, L), female, and "All" contexts. Gas1 was biallelically expressed in the remaining high fat diet and male contexts in LIV and all contexts in HYP and WAT. Its locus had 57 strain-specific variants, although none were predicted as functional. Gas1 encodes a membrane glycoprotein that binds and regulates sonic hedgehog during development (Lee et al. 2001). Gas1 is differentially expressed between high and low selenium diets in mice ovaries, suggesting its expression may be sensitive to dietary environment (Qazi et al. 2021).

## Sequence-dependent ASE genes can switch the direction of their allelic biases

Finally, we found 45 sequence-dependent ASE genes with inconsistent patterns of allelic biases. These genes showed significant ASE in opposite directions among the tissues and/or environmental contexts (Fig. 5). For 44 of the 45 genes, such direction switching occurred at the tissue level: for example, a gene may have a LG/J bias in one tissue, a SM/J bias in another tissue, and sometimes

even no bias (biallelic) in the third tissue. Four genes had tissue-independent ASE, or a sequence bias across contexts in every expressed tissue (albeit in different directions). Nineteen genes had context-dependent ASE, or a sequence bias only in certain environmental contexts that switched direction across tissues. One gene (Cidec) had a context-dependent switch in ASE direction within the same tissue, discussed further below. The remaining 22 genes had a combination of context- and tissue-dependent ASE patterns: such genes had a sequence bias in one direction across contexts in one or two tissues, but a context-dependent sequence bias in the opposite direction in another tissue. Overall, these genes had moderate allelic biases; their mean AGE scores were −0.38 and 0.40 for SM/J and LG/J biased genes, respectively. These dynamic direction-switching patterns confirm that sequence-dependent ASE is not solely a result of genetic variation. Otherwise, the same variant causing ASE in one tissue would also be present in any other cell expressing that gene. These patterns hint at epigenetic regulatory elements interacting with genetic variation in a tissue-specific manner to influence the final allelic frequency.

We validated the expression profiles of two direction-switching genes (Stab2 and Ociad2) with pyrosequencing (Supplemental Fig. S15). Stabilin 2 (Stab2) had a LG/J bias in all contexts in LIV but a SM/J bias across most contexts in HYP and WAT. Its locus had 1019 strain-specific variants, including 14 nonsynonymous SNPs in the SM/J genome. Stab2 binds to hyaluronic acid and mediates its transportation inside the cell (Zhou et al. 2002). It is predominantly expressed in two isoforms in the liver and spleen, but weakly expressed in the brain and adipose tissue (Falkowski et al. 2003). OCIA domain containing 2 (Ociad2) had significant sex and sex-by-diet effects in WAT and LIV. These are reflected in LG/J biases in the L, F, and "All" contexts in WAT, but SM/J biases in the LF, L, F, and "All" contexts in LIV. Ociad2 was biallelically expressed in the remaining contexts of both tissues and in HYP. Its locus had 118 strain-specific variants, although none were predicted as functional. Ociad2 regulates cell migration and is moderately expressed in several tissues, including the brain and liver (Sinha et al. 2018). It is highly expressed in female ovarian tumors (Nagata et al. 2012), but otherwise sex or diet effects on Ociad2 expression have not been explored.

## ASE genes are enriched in QTLs for metabolic and anatomical traits

Mapping studies in the LG/J x SM/J advanced intercross line have identified hundreds of QTLs associated with a constellation of complex traits. Such approaches are agnostic to tissue and often implicate regions with many genes, making it difficult to pinpoint which ones are phenotypically relevant. In contrast, ASE patterns are gene-specific and vary across tissues and environments. If there is a functional difference between alleles, then an expression imbalance in the right conditions can have phenotypic consequences and manifest as QTL signals. To probe this, we integrated ASE and QTL data from populations derived from the same founder strains at various degrees of intercrossing (Fig. 6A; Supplemental Table S7).

Previously, 127 QTLs influencing metabolic traits were mapped in the AIL $F_{16}$ generation (Lawson et al. 2010, 2011b; Cheverud et al. 2011). These studies incorporated the same environmental contexts (sex and dietary fat) used here, traced the parental origin of each marker allele, and characterized the context-dependent additive and imprinting effects influencing
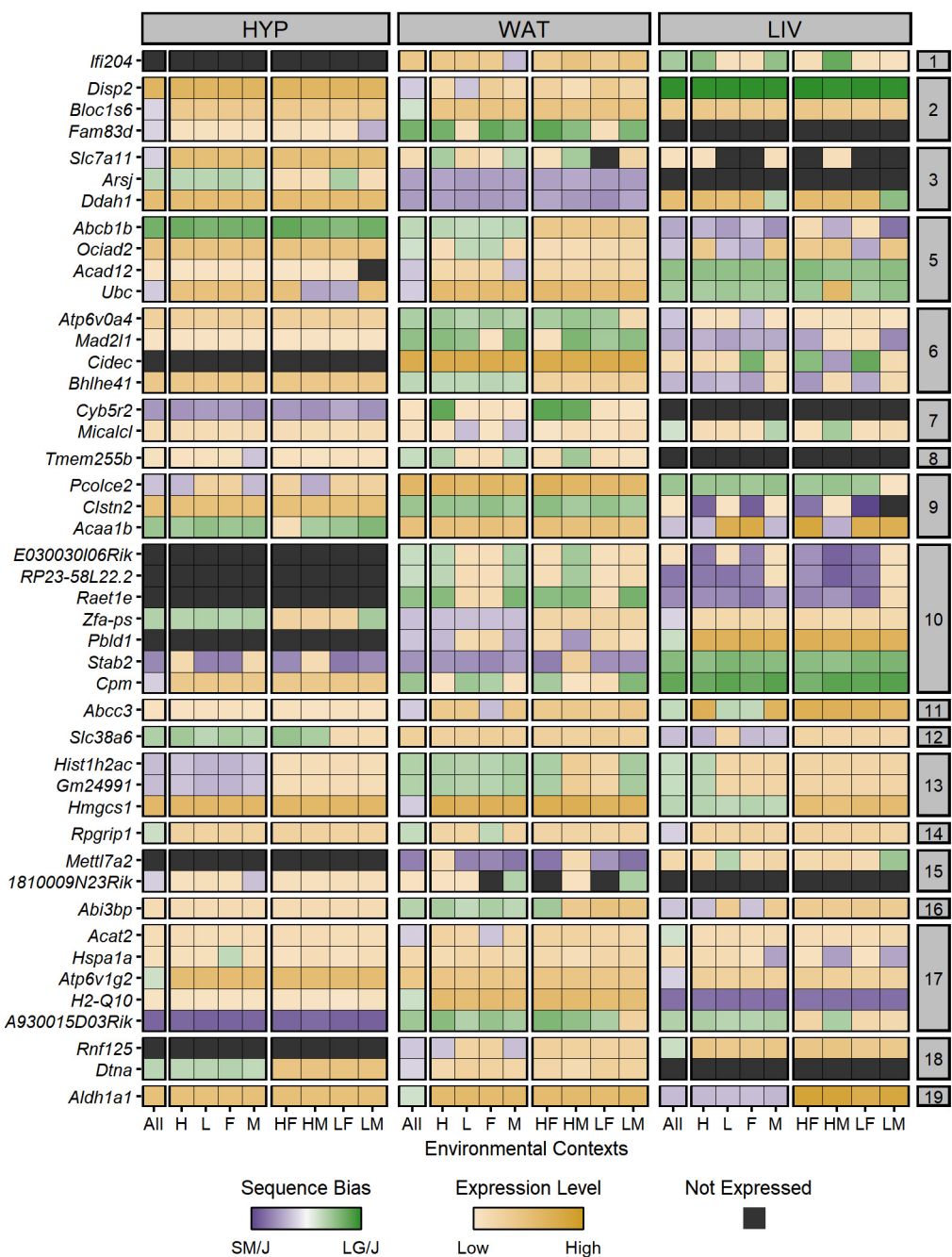
**Figure 5.** Forty-five ASE genes switch their sequence bias direction across conditions. Heatmap of ASE profiles for the 45 genes with significant sequence biases in opposite directions, including those validated with pyrosequencing. Genes are color coded by their expression pattern in each tissue-by-context analysis. Shades of purple and green indicate their degree of SM/J or LG/J bias, respectively (AGE scores). If genes are not biased, shades of yellow indicate their biallelic expression levels (log-transformed total counts). Black indicates genes are not expressed. The *y*-axis is grouped and sorted by chromosomal position. Supercolumns denote tissues: (HYP) hypothalamus; (WAT) white adipose; and (LIV) liver. Subcolumns denote environmental contexts: (All) all contexts; (H) high fat; (L) low fat; (F) females; (M) males; (HF) high fat females; (HM) high fat males; (LF) low fat females; and (LM) low fat males.

metabolic traits. Overall, sequence-dependent ASE genes were significantly enriched in the 83 $F_{16}$ QTLs with additive effects (*P* = 0.01) (Fig. 6B). We compared the QTL sets for each metabolic trait category with the full ASE gene sets for each tissue. QTLs for obesity traits were enriched for HYP, WAT, and LIV genes (*P* < 0.01). QTLs for diabetes traits were enriched for HYP genes (*P* = 0.01),

but not WAT or LIV. QTLs for serum lipid traits were enriched for WAT genes (*P* = 0.02), but not HYP or LIV (Supplemental Fig. S16). These findings reflect the discrete roles of each tissue in metabolism physiology and illustrate that ASE patterns can inform which tissues are phenotypically relevant. Conversely, parent-of-origin–dependent ASE genes (excluding the 171 snoRNA cluster
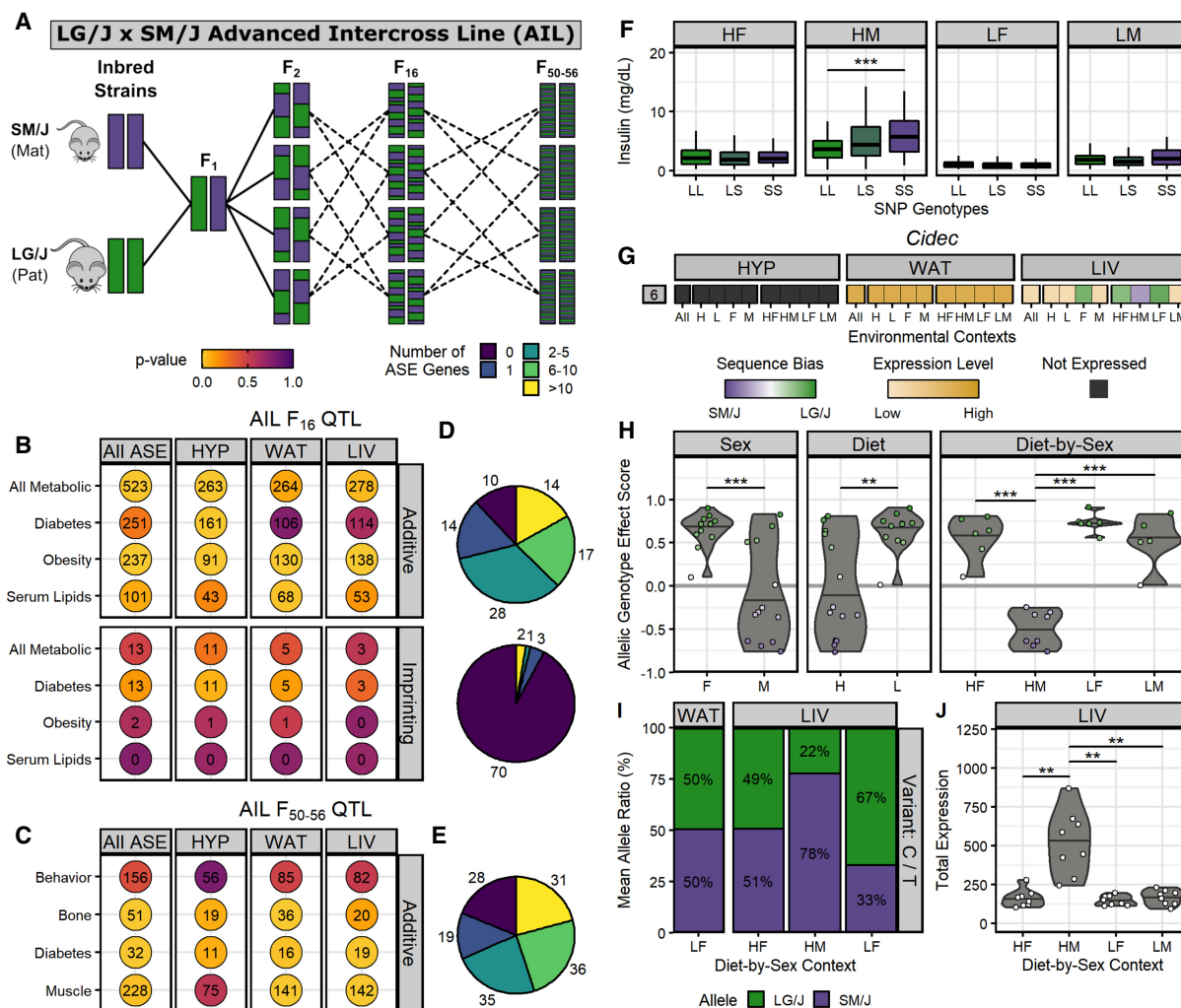
**Figure 6.** Integrating ASE and QTL data reveal *Cidec* as a candidate gene for insulin levels. (*A*) Breeding scheme for the LG/J x SM/J advanced intercross line. We calculated enrichment of tissue-specific ASE gene sets (*x*-axis) in trait-specific AIL QTL sets (*y*-axis) among sequence-dependent ASE genes in additive $F_{16}$ QTL (*top*) and parent-of-origin-dependent ASE genes in imprinting $F_{16}$ QTL (*bottom*) (*B*), and sequence-dependent ASE genes in additive $F_{50-56}$ QTL (*C*). Circle color corresponds to the enrichment *P*-value. Numbers indicate the total overlapping ASE genes in each QTL set. (*D,E*) Tally of how many AIL QTLs contain each range of ASE genes: 0 (purple) to more than 10 (yellow). Pie charts match their adjacent ASE gene and QTL set intersections. (*F*) The AIL $F_{16}$ QTL *Ddiab6d* showed context-dependent additive effects. Box plots of serum insulin levels across SNP rs6393943 genotypes in $F_{16}$ mice (LL, LG/J homozygous; LS, heterozygous; SS, SM/J homozygous). The *x*-axis is grouped by diet-by-sex context: (HF) high fat females; (HM) high fat males; (LF) low fat females; and (LM) low fat males. Horizontal bars denote mean phenotypes. (***) $P \leq 0.001$; assessed by Student's *t*-test. (*G*) *Cidec* had a context-dependent switch in sequence bias direction within LIV. ASE heatmap of *Cidec* across tissues-by-context analyses (for full description, see Fig. 4 or 5). (*H*) *Cidec* ASE biases had significant sex, diet, and diet-by-sex effects. Violin plots display individual AGE scores for *Cidec* (*y*-axis) across environmental contexts (*x*-axis). Horizontal bars denote mean AGE scores. Dots are color coded by their sequence bias. (**) $P \leq 0.01$, (***) $P \leq 0.001$; assessed by ANOVA (sex, diet) or Tukey's post hoc tests (diet-by-sex). (*I*) *Cidec* ASE profile in LIV and WAT was validated by pyrosequencing. Bar graphs denote mean allelic ratios (*y*-axis) in select cohorts (*x*-axis) and are color coded by allele (LG/J, green; SM/J, purple). (*J*) *Cidec* had significantly higher expression in HM. Violin plots of total expression levels in LIV (*y*-axis) for each diet-by-sex context (*x*-axis). (**) $P \leq 0.01$; assessed by Student's *t*-test.

in the PWS/AS domain) were not enriched in the 76 $F_{16}$ QTLs with imprinting effects ($P = 0.47$) (Fig. 6B). Tissue-specific ASE gene sets were also not enriched in the QTL sets for any metabolic trait category ($P > 0.05$) (Supplemental Fig. S16), supporting previous work showing parental biases are often not the direct source of imprinting QTL effects (Macias-Velasco et al. 2022).

Additionally, 149 QTLs influencing behavioral, musculoskeletal, and metabolic traits were mapped in later generations ($F_{50-56}$) of the same AIL population (Gonzales et al. 2018; Cordero et al. 2019). These studies did not capture environmental contexts or parental origin, but still explored how additive effects impact complex traits in a context-independent manner. Overall, sequence-dependent ASE genes were significantly enriched in additive $F_{50-56}$ QTLs for muscle weights ($P = 0.03$), bone sizes ($P = 0.01$), and diabetes-related traits ($P = 0.01$) (Fig. 6C). These results are consistent with the history of the LG/J and SM/J founder strains: they were independently selectively bred for large and small body sizes, so it is not surprising that strain-specific variation can induce ASE biases in their $F_1$ cross that often occur in/near regions relevant for morphometric traits. Furthermore, sequence-dependent ASE genes were not enriched in additive $F_{50-56}$ QTLs for behavioral traits related to locomotor activity and prepulse inhibition ($P =$

0.42) (Fig. 6C). Tissue-specific ASE gene sets were also not enriched in behavioral QTLs, even for the brain region HYP ($P = 0.79$) (Supplemental Fig. S17), suggesting the metabolic tissues profiled here are not relevant for these behavioral phenotypes.

Although sequence-biased genes were located in additive AIL QTL more often than expected, individual QTL typically contained few ASE genes (Supplemental Table S7). Additive $F_{16}$ QTLs overlapped a mean of 75 genes each, yet 51% of those QTLs contained only one to five sequence-dependent ASE genes. Thirty-seven percent of QTLs contained more than six ASE genes and merely 12% had no ASE genes (Fig. 6D). Similarly, additive $F_{50-56}$ QTLs overlapped a mean of 61 genes each, but 36% of those QTLs contained only one to five sequence-dependent ASE genes. Forty-five percent of QTLs contained more than six ASE genes and 19% had no ASE genes (Fig. 6E). In contrast, imprinting $F_{16}$ QTLs overlapped a mean of 62 genes each, but 92% of them did not contain any parent-of-origin-dependent ASE genes (Fig. 6D). Only two imprinting $F_{16}$ QTLs contained more than 10 ASE genes and they both overlapped the cluster of 171 snoRNAs on Chromosome 7.

### *Cidec* has a context-dependent switch in ASE direction and is a candidate gene for insulin levels

Because ASE genes often comprise a subset of all genes within a QTL, we propose that incorporating ASE patterns can dissect QTL etiology and prioritize candidate genes in relevant tissues. Here, we present an example of the utility of this orthogonal approach. *Ddiab6d* was an AIL $F_{16}$ QTL on Chromosome 6 associated with serum insulin, basal glucose, and glucose tolerance levels that showed additive effects in a context-dependent manner (Lawson et al. 2011b). In high fat males, AIL $F_{16}$ mice homozygous for the SM/J allele had significantly higher insulin levels than the homozygous LG/J mice ($P < 0.001$). There were no significant differences in insulin levels between homozygote genotypes in the high fat females, low fat females, or low fat male cohorts (Fig. 6F). The *Ddiab6d* locus spanned 14.7 Mb and contained 100 genes, complicating efforts to pinpoint the causal mechanisms of this additive effect.

However, only four sequence-dependent ASE genes were located in the *Ddiab6d* locus: *Brpf1*, *Cidec*, *Irak2*, and *Syn2*. Cell death-inducing DFFA-like effector c (*Cidec*) had a diet- and sex-dependent switch in bias direction within the same tissue, which was a notable exception to the tissue-level ASE direction switches reported above. In LIV, *Cidec* had a strong SM/J bias in the HM context, yet a LG/J bias in the HF, LF, and F contexts (Fig. 6G). It was biallelically expressed in the remaining LIV contexts and across all WAT contexts, but not expressed in HYP. *Cidec* ASE biases also had significant sex, diet, and diet-by-sex effects in LIV ($P < 0.01$) (Fig. 6H). We validated the context-dependent switch in ASE bias direction with pyrosequencing (Fig. 6I). *Cidec* was also expressed in LIV at a significantly higher level in HM than the other diet-by-sex contexts ($P < 0.01$) (Fig. 6J).

*Cidec* promotes and regulates lipid storage in liver and adipose tissue (Xu et al. 2012). *Cidec* expression is sensitive to diet composition and sex hormones; for example, male mice had higher total hepatic expression than females when fed a Western diet (i.e., high cholesterol and saturated fats) (Herrera-Marcos et al. 2020). In another study, high fat-fed mice had increased *Cidec* expression in liver that was also associated with improved insulin sensitivity (Iv et al. 2015). The liver plays an important role in regulating homeostatic insulin levels and removing it from the bloodstream (Tokarz et al. 2018). Hepatic insulin clearance is an emerging com-

ponent of type 2 diabetes and metabolic syndrome, as impaired hepatic insulin clearance is correlated with increased waist circumference, blood pressure, fasting glucose, triglycerides, and insulin secretion (Pivovarova et al. 2013; Najjar and Perdomo 2019). Together, these data implicate *Cidec* as a candidate gene for *Ddiab6d* whose allelic composition in the liver is sensitive to metabolically relevant environmental factors and can influence insulin levels.

## Discussion

Allele-specific expression imbalances caused by genetic and epigenetic variation have widespread functional consequences on complex traits, but how environmental signals contribute to this cross talk remains understudied. Here, we explored how allelic genotype, parent-of-origin, tissue type, sex, and dietary fat simultaneously impact ASE biases in an adult mouse $F_1$ reciprocal cross. We present a genome-wide map of parent-of-origin and sequence-dependent ASE patterns across three metabolically relevant tissues and nine environmental contexts. The granularity of our analyses revealed that both ASE classes are highly dependent on tissue and environmental context. We identified 2853 genes with significant parental and/or sequence biases and sorted them into three expression profiles: tissue-independent, tissue-dependent, and context-dependent. We also found 45 genes with inconsistent ASE biases that switched direction across tissues and/or contexts. Although the breadth of these patterns precluded a detailed discussion of each gene, we validated examples of each expression profile to show how these allelic imbalances could manifest and lead to potential functional consequences.

"Tissue-independent" genes are strongly biased wherever they are expressed. Most parent-of-origin-dependent ASE genes (67%) have this expression profile—all of which are related to genomic imprinting, a well-characterized epigenetic phenomenon that results in uniparental expression and is often conserved across tissues (Barlow and Bartolomei 2014). In contrast, only 23% of sequence-dependent ASE genes have this expression profile, likely owing to *cis*-acting genetic variants in coding regions that severely affect one allele's function whenever that gene is expressed. These patterns conform to the conventional ASE mechanism whereby the genetic and epigenetic processes that influence allelic composition will always do so wherever a gene is expressed, in a manner impervious to tissue type or environmental signals (Lo et al. 2003; Babak et al. 2015). However, we found this is not always the case.

"Tissue-dependent" genes are moderately biased in one direction in some tissues, but are biallelically expressed or biased in the opposite direction in other tissues. Tissue-specific gene expression (simply not expressed in certain tissues) is not considered here, because we are specifically interested in cases in which a gene has one allelic ratio in one tissue (e.g., 80% SM/J, 20% LG/J) and a different ratio in another tissue (e.g., 50% SM/J, 50% LG/J). Merely 6% of parent-of-origin-dependent ASE genes have this expression profile, and all are located in known imprinted domains. Twenty-five percent of sequence-dependent ASE genes have this profile, demonstrating that sequence biases are not solely attributable to genetic sources. Instead, both patterns may be explained by variation in tissue-specific regulatory features, such as enhancers or RNA-binding proteins. Such tissue-specific factors can then interact with *cis*-acting variants to produce a tissue-dependent allelic imbalance. This is likely the case for genes with inconsistent biases between tissues, for which the variation in one allele may be

favorable in certain tissues' regulatory landscape but the opposite allele is more favorable elsewhere (Leung et al. 2015; Andergassen et al. 2017).

Finally, "context-dependent" genes are subtly biased in one direction in certain environmental contexts, but are biallelically expressed or biased in the opposite direction in other contexts and tissues. We found this expression profile is quite prevalent in both classes of ASE. Twenty-seven percent of parent-of-origin-dependent ASE genes have diet- and/or sex-specific biases in a less extreme manner than the other profiles. Most genes have no clear connection to imprinting, suggesting another mechanism for parental ASE outside of traditional imprinting that is sensitive to environmental perturbations (Wolf et al. 2008; Morcos et al. 2011). More than half of sequence-dependent ASE genes (52%) have diet- and/or sex-specific biases, indicating these extrinsic and intrinsic factors interact with genetic variation to cause a sequence bias. Both patterns suggest a model in which environmental signals may interact more efficiently with one allele over the other, leading to shifting and inconsistent allelic proportions in response to environmental cues (Shao et al. 2019). Often, these genes are not significantly biased in our "all contexts" analysis, only in a specific environmental cohort. This highlights the necessity of studying gene-by-environment interactions, as such effects are obscured when multiple contexts are collapsed together.

We can only detect ASE in genes with strain-specific variants in transcribed regions, which is a fraction of all expressed genes. In particular, imprinted genes that are crucial for development may be under tight evolutionary control and not have variants; thus, we are unable to assess their ASE status. Although our exact findings are limited to this $F_1$ mouse model, the broad patterns nevertheless show the complexity of ASE and its contribution to complex traits. Traditional mapping studies connect genotypes to phenotypes, but are agnostic to tissue and often do not consider environment. Expression QTL studies connect genotypes to total gene expression, but assume biallelic expression. Tissue- and context-dependent ASE of both classes can bridge these approaches and pinpoint which tissues and/or environments are relevant for a phenotype. A gene may be expressed at the same level in two cohorts, but its allelic composition could differ owing to ASE. If there is a functional difference between its alleles, then an expression imbalance in the right conditions could have phenotypic consequences. We integrated ASE and QTL data from populations derived from the same founder strains at various degrees of intercrossing. Tissue-specific ASE genes are enriched in QTLs for metabolic and musculoskeletal traits, yet consist of a small portion of all genes within the QTL. Incorporating these dynamic ASE patterns with orthogonal evidence will help us decipher the genotype-to-phenotype map.

## Methods

### $F_1$ reciprocal cross mouse model and RNA sequencing

We obtained LG/J and SM/J founders from The Jackson Laboratory and generated $F_1$ reciprocal crosses by mating LG/J mothers with SM/J fathers (LxS) and vice versa (SxL). $F_1$ offspring were weaned into sex-specific cages at 3 wk and randomly placed on either a high fat diet (42% kcal from fat; Teklad TD88137) or an isocaloric low fat diet (15% kcal from fat; Research Diets D12284). They were fed ad libitum. The $F_1$ crosses did not differ in weight gained in any diet-by-sex cohort ($P > 0.05$) (Supplemental Fig. S18). At 20 wk, $F_1$

mice were euthanized with a sodium pentobarbital injection followed by cardiac perfusion with phosphate-buffered saline. We harvested hypothalamus (HYP), liver (LIV), and reproductive white adipose (WAT) tissue, which were flash frozen in liquid nitrogen and stored at −80°C until RNA extraction. All procedures were approved by the Institutional Animal Care and Use Committee at Washington University School of Medicine.

We sequenced 32 samples per tissue, representing four mice from each sex, diet, and $F_1$ cross cohort. We extracted total RNA from WAT and HYP using the RNeasy Lipid Tissue Mini Kit (QIAGEN) and from LIV using a standard TRIzol-chloroform procedure. Samples were selected based on sufficient NanoDrop RNA concentrations (Thermo Fisher Scientific) and RNA integrity scores ≥8.0 (Agilent). We constructed RNA-seq libraries with the Ribo-Zero rRNA Removal Kit (Illumina), checked their quality with the Bioanalyzer DNA 1000 assay (Agilent), and sequenced them at 100-bp paired-end reads on an Illumina HiSeq 400. After sequencing, reads were demultiplexed and assigned to individual samples.

### Allele-specific expression mapping

Mapping ASE in $F_1$ heterozygotes is vulnerable to reference genome alignment bias (Degner et al. 2009; Wang and Clark 2014). Previously, LG/J and SM/J reference genomes were created by combining strain-specific SNPs and indels with the GRC38.72-mm10 reference template (Nikolskiy et al. 2015). Customized gene annotations were created by adjusting Ensembl definitions (GRCm38.72) for indexing differences owing to strain-specific indels. To mitigate this concern, we combined those LG/J and SM/J genomes into a custom pseudogenome and aligned RNA-seq reads to both strains simultaneously. We mapped reads uniquely using the two-pass mapping strategy in STAR v.2.7.2b (Dobin et al. 2013). Briefly, splice junctions are collected during a first round and used to inform a second round of mapping. By not allowing multimapping, we only retained reads that uniquely cover strain-specific variants so we could assign their allelic origin; reads covering identical regions between the parental strains were discarded. Alignment summaries are provided in Supplemental Table S8 and Supplemental Figure S19.

Next, we assigned each aligned read to a gene using BEDTools v.2.27.1 (Quinlan and Hall 2010) and our strain-specific Ensembl annotations. We assessed ASE at the gene level because not all exons within a given gene contained informative variants. Gene-level allele-specific counts were then upper quartile normalized (Supplemental Fig. S20) and filtered to remove lowly expressed genes (total normalized counts < 20). We retained a total of 9171 genes in HYP, 9745 genes in WAT, and 8026 genes in LIV.

### Library complexity

Insufficient library complexity can hamper ASE detection (Wang and Clark 2014). To measure each library's complexity, we fit the distribution of LG/J allele expression biases (the proportion of total allele-specific read counts with the LG/J haplotype) to a beta-binomial distribution using the VGAM package (Yee 2010). We estimated the shape parameters ($\alpha$, $\beta$) of the beta-binomial distribution and calculated the overdispersion parameter ($\rho$) as $\rho = 1/(1 + \alpha + \beta)$. Lower values of $\rho$ (<0.075) indicate a library is sufficiently complex. One WAT and one LIV library had poor complexity and were removed from further analyses (Supplemental Fig. S21).

### Determining biased allele-specific expression

For each "tissue-by-context" analysis, we required a gene to be expressed in ≥75% of biological replicates per $F_1$ reciprocal cross. This

conservative sample size threshold retained only genes expressed in most biological replicates in a given cohort, allowing us to confidently conclude their expression status. We intentionally removed sporadically expressed genes, because they may result in false or transient biases.

We adapted a published regression model (Takada et al. 2017) to jointly estimate parent-of-origin (PO) and allelic genotype (AG) effects on ASE. First, we assigned two binary variables to each gene's allele-specific counts based on their allelic origin. For the PO term, maternal alleles received a 0 and paternal alleles received a 1; for the AG term, LG/J alleles received a 0 and SM/J alleles received a 1 (Supplemental Fig. S22). Next, we added a term representing the RNA-seq library barcodes so that both allele-specific read counts (two separate rows in our data set) for an individual mouse are treated as a pair. This paired-sample design helps distinguish between cases in which a gene is not expressed in a particular sample (i.e., both alleles' read counts are zero) and an extreme expression bias (i.e., only one allele's read counts are zero), which results in a lower biological coefficient of variation and lower overall dispersion (Supplemental Fig. S22). Our final regression model equation was

$$Allele - Specific\ Read\ Counts\ \sim\ PO + AG + Library\ Barcode.$$

Finally, we fit a negative binomial generalized linear model (GLM) and conducted likelihood ratio tests in edgeR (Robinson et al. 2010; McCarthy et al. 2012) to estimate the POE and AGE on ASE.

Next, we quantified the direction and magnitude of each gene's expression biases. For each sample, we calculated a gene's allelic bias as the proportion of total counts with the LG/J haplotype ($L_{bias}$) or SM/J haplotype ($S_{bias}$). Using the mean allelic biases of each $F_1$ cross, we constructed POE and AGE scores per gene as follows:

$$POE = mean(SxL\ L_{bias}) - mean(LxS\ L_{bias}),$$
$$AGE = \frac{(mean(LxS\ L_{bias}) + mean(SxL\ L_{bias})) - (mean(LxS\ S_{bias}) + mean(SxL\ S_{bias}))}{2}.$$

POE scores ranged from −1 (completely maternally expressed) to +1 (completely paternally expressed). Similarly, AGE scores ranged from −1 (completely SM/J expressed) to +1 (completely LG/J expressed). Scores of 0 for both indicated biallelic expression. Full summary statistics for each analysis are provided in Supplemental Tables S2–S4.

### Estimating significance thresholds

Adjusting for multiple tests in ASE analyses is challenging. Allelic biases are often correlated for genes within and between regions, breaking any independence assumptions. We used a permutation approach to estimate our significance thresholds. For each tissue-by-context analysis, we generated a permuted null distribution of both AGE and POE terms. Over several iterations, we randomly shuffled the allele-specific read counts for all genes, reran our GLM analyses, and recalculated the POE/AGE scores. Instead of choosing an arbitrary number, we continued to add new iterations until the null model was "stable." After every iteration, we added those permuted results to the overall null model and evaluated the difference in mean permuted likelihood ratios. We defined "stability" as when the difference in mean permuted likelihood ratios for both terms fluctuated by <|0.001| for 10 consecutive iterations (Supplemental Fig. S23). In other words, the null model was considered "stable" once incorporating another iteration did not shift the overall distribution of test statistics (mean iterations per analysis = 51). Summary statistic comparisons between the real and permuted data sets are provided in Supplemental Figures S24–S26.

Next, we built an empirical cumulative distribution function (ECDF) from each term's permuted $P$-values for each tissue-by-context analysis (Supplemental Fig. S27). We fit each term's raw $P$-values to its respective ECDF to compute the adjusted $P$-values, that is, the proportion of tests from the null model that are more extreme (smaller $P$-values) than a test from the real model. We also calculated the 5th and 95th quantiles of the permuted POE and AGE scores in each analysis; the more extreme value of each score became our critical thresholds. We deemed adjusted $P$-values ≤ 0.05 as statistically significant and real scores beyond their critical threshold as biologically significant. Parent-of-origin-dependent ASE genes had significant POE $P$-values and POE scores; sequence-dependent ASE genes had significant AGE $P$-values and AGE scores (Supplemental Fig. S28; Supplemental Table S1).

### Total expression mapping

During allele-specific mapping, reads were not aligned if they covered homozygous regions between the parental strains. We saved those unmapped reads and aligned them separately to the LG/J and SM/J genomes using a standard multimapping approach in STAR v.2.7.2b. We assigned each aligned read to a gene using BEDTools v.2.27.1 and our strain-specific Ensembl annotations. Next, we averaged the read counts for both genomes; because those reads covered homozygous regions, they mapped to each genome similarly. We combined the averaged "multimapped" counts with the raw "allele-specific" counts from the previous mapping step, resulting in raw "total expression" counts. Finally, gene-level counts were upper quartile normalized and filtered to remove lowly expressed genes (total counts < 10). We retained a total of 25,926 genes in HYP, 26,450 genes in WAT, and 24,076 genes in LIV.

### Characterizing ASE profiles: tissue-independent, tissue-dependent, and context-dependent

We sorted both classes of ASE genes into three expression profiles. Tissue-independent genes had a significant bias in every expressed tissue. Within a tissue, the gene must be biased in five or more of the nine environmental contexts (i.e., most, but not all, contexts). This flexibility allowed for genes that may have true ASE but were excluded because of our conservative sample size requirements; these genes could appear as biallelic because they still passed our other filtering thresholds. Tissue-dependent genes had a significant bias in some tissues and no bias in others. The gene must be biased in more than five of the nine contexts in one or two tissues, but biallelic in more than five of the nine contexts in the other tissue(s). Both profiles allowed for genes to not be expressed in some tissues, because we wanted to distinguish between tissue-specific gene expression and tissue-dependent ASE. Finally, context-dependent genes had a significant bias only in certain environmental contexts and no bias elsewhere. The gene must be biased in five or more of the nine contexts within a tissue, but biallelic in the other contexts and/or tissues. All significant ASE genes fit into one of these expression profiles; no genes were unclassified.

### Evaluating environmental context dependency

For each context-dependent gene, we calculated each sample's allelic bias as the proportion of total read counts with the LG/J haplotype ($L_{bias}$) or the SM/J haplotype ($S_{bias}$). We constructed

individualized POE or AGE scores per sample per gene as follows:

$$Individual\ POE = \begin{cases} mean(SxL\ L_{bias}) - L_{bias}, & cross = LxS \\ L_{bias} - mean(LxS\ L_{bias}), & cross = SxL \end{cases},$$

$$Individual\ AGE = \begin{cases} \dfrac{(L_{bias} + mean(SxL\ L_{bias})) - (S_{bias} + mean(SxL\ S_{bias}))}{2}, & cross = LxS \\ \dfrac{(mean(LxS\ L_{bias}) + L_{bias}) - (mean(LxS\ S_{bias}) + S_{bias})}{2}, & cross = SxL \end{cases}.$$

POE scores ranged from −1 (completely maternal) to +1 (completely paternal); AGE scores ranged from −1 (completely SM/J) to +1 (completely LG/J). Scores of 0 indicated biallelic expression or not expressed.

Next, we used ANOVA models to test whether a gene's allelic biases (individualized scores) were influenced by sex, diet, and/or their interaction. We considered FDR-corrected $P$-values $\leq 0.1$ to be significant (Supplemental Table S5). For genes with significant diet-by-sex interactions, we conducted Tukey's post hoc tests to identify significant differences among diet-by-sex cohorts (adjusted $P \leq 0.05$) (Supplemental Table S6).

## Chromosomal location enrichment

We performed overrepresentation analysis for chromosomal locations with the WEB-based GEne SeT AnaLysis Toolkit v2019 (Liao et al. 2019). For each tissue-by-context analysis, we analyzed the list of genes with significant biases against the background of all genes with detectable ASE. A Benjamini–Hochberg FDR-corrected $P$-value $\leq 0.1$ was considered significant.

## Canonically imprinted genes

We defined genes as "canonically imprinted" if they appeared in the Geneimprint mouse database (https://www.geneimprint .com, as of May 2020) and/or a PubMed search with imprinting-related terms.

## Pyrosequencing

We validated two to three examples of each expression profile (total = 13 genes). We sorted genes in each profile by total expression level. To minimize noise during pyrosequencing, we prioritized genes that were highly expressed and statistically significant (for context-dependent examples), but excluded those with high expression variance between biological replicates. For each gene, we identified the strain-specific SNPs within exons and designed primer sets to flank the variants using Geneious Prime 2020.0.4 (https://www.geneious.com) (Kearse et al. 2012). Wherever possible, target regions were 150- to 200-bp long and spanned an exon–exon junction to avoid genomic DNA contamination. We verified the specificity of each primer set in silico with Geneious and in vitro with PCR and Sanger sequencing. All primer sequences are provided in Supplemental Table S9.

We extracted total RNA from the HYP, WAT, and LIV of one mouse per $F_1$ cross in each diet-by-sex cohort using the RNeasy Lipid Tissue Mini Kit (QIAGEN). The cDNA of each gene target was reverse transcribed and PCR amplified with the PyroMark OneStep RT-PCR Kit (QIAGEN) using one biotinylated (reverse) and one nonbiotinylated (forward) primer. The biotinylated single-stranded PCR products were purified with Streptavidin Sepharose High Performance beads (Cytiva) and hybridized to sequencing primers (same as forward) on the PyroMark Vacuum Prep Workstation (QIAGEN). Finally, we performed pyrosequencing with the allele quantification program on the PyroMark Q24 system (QIAGEN). The PyroMark Q24 software quantified the allelic ratio of the variable position(s) in each gene's assay. We calculated the mean allelic ratios of each variant in each tissue-by-context cohort.

## Integrating ASE and QTL data

The genotypes, phenotypes, and QTL mapping methods for the LG/J x SM/J advanced intercross line are described elsewhere (Lawson et al. 2010, 2011b; Cheverud et al. 2011; Gonzales et al. 2018; Cordero et al. 2019). The $F_{16}$ QTL were originally mapped to the NCBI37/mm9 reference genome, so we converted those intervals to their GRCm38/mm10 coordinates with the UCSC Genome Browser's liftOver tool (https://genome.ucsc.edu) (Lee et al. 2022). We also extracted the ASE gene positions from the GRCm38/mm10 assembly annotation.

We analyzed the two ASE classes separately: sequence-dependent ASE genes were intersected with $F_{16}$ and $F_{50–56}$ additive QTL, and parent-of-origin-dependent ASE genes were intersected with $F_{16}$ imprinting QTL (Supplemental Table S7). We counted how many ASE genes overlapped with each QTL using BEDTools v.2.27.1. We sorted each AIL generation's QTL based on their associated trait category and compared those trait-specific QTL sets with the tissue-specific ASE gene sets. Next, we conducted enrichment analyses to determine whether ASE genes are overrepresented in AIL QTLs. We randomly shuffled the QTL windows around the genome and tallied how many ASE genes overlapped with regions of those lengths by chance. These permutations were repeated more than 10,000 iterations to establish a null model. We calculated a $Z$-score based on the null model and the real ASE/QTL overlap, then sampled a normal distribution to obtain a $P$-value. We considered $P$-values $\leq 0.05$ to be significantly enriched.

## Data access

All raw RNA sequencing data generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi .nlm.nih.gov/bioproject/) under accession number PRJNA753198.

## Competing interest statement

The authors declare no competing interests.

## Acknowledgments

## References

Albrecht M, Choubey D, Lengauer T. 2005. The HIN domain of IFI-200 proteins consists of two OB folds. *Biochem Biophys Res Commun* **327:** 679–687. doi:10.1016/j.bbrc.2004.12.056

Andergassen D, Dotter CP, Wenzel D, Sigl V, Bammer PC, Muckenhuber M, Mayer D, Kulinski TM, Theussl HC, Penninger JM, et al. 2017. Mapping the mouse allelome reveals tissue-specific regulation of allelic expression. *eLife* **6:** e25125. doi:10.7554/eLife.25125

Babak T, DeVeale B, Tsang EK, Zhou Y, Li X, Smith KS, Kukurba KR, Zhang R, Li JB, van der Kooy D, et al. 2015. Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nat Genet* **47:** 544–549. doi:10.1038/ng.3274

Barlow DP, Bartolomei MS. 2014. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol* **6:** a018382. doi:10.1101/cshperspect.a018382

Bonasio R, Tu S, Reinberg D. 2010. Molecular signals of epigenetic states. *Science* **330:** 612–616. doi:10.1126/science.1191078

Buil A, Brown AA, Lappalainen T, Viñuela A, Davies MN, Zheng HF, Richards JB, Glass D, Small KS, Durbin R, et al. 2015. Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nat Genet* **47:** 88–91. doi:10.1038/ng.3162

Cao Y, Wang L, Wang CY, Ye J, Wang Y, Li T, Garcia-Godoy F, Sun D, Gu W, Postlethwaite AE. 2018. Sex differences in correlation with gene expression levels between *Ifi200* family genes and four sets of immune disease-relevant genes. *J Immunol Res* **2018:** e1290814. doi:10.1155/2018/1290814

Carson C, Lawson HA. 2020. Genetic background and diet affect brown adipose gene coexpression networks associated with metabolic phenotypes. *Physiol Genomics* **52:** 223–233. doi:10.1152/physiolgenomics.00003.2020

Castel SE, Aguet F, Mohammadi P, Aguet F, Anand S, Ardlie KG, Gabriel S, Getz GA, Graubert A, Hadley K, et al. 2020. A vast resource of allelic expression data spanning human tissues. *Genome Biol* **21:** 234. doi:10.1186/s13059-020-02122-z

Cavalli M, Pan G, Nord H, Wallerman O, Wallén Arzt E, Berggren O, Elvers I, Eloranta ML, Rönnblom L, Lindblad Toh K, et al. 2016. Allele-specific transcription factor binding to common and rare variants associated with disease and gene expression. *Hum Genet* **135:** 485–497. doi:10.1007/s00439-016-1654-x

Cheverud JM, Lawson HA, Fawcett GL, Wang B, Pletscher LS, Fox AR, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, et al. 2011. Diet-dependent genetic and genomic imprinting effects on obesity in mice. *Obesity* **19:** 160–170. doi:10.1038/oby.2010.141

Cordero AIH, Gonzales NM, Parker CC, Sokolof G, Vandenbergh DJ, Cheng R, Abney M, Sko A, Douglas A, Palmer AA, et al. 2019. Genome-wide associations reveal human-mouse genetic convergence and modifiers of myogenesis, *CPNE1* and *STC2*. *Am J Hum Genet* **105:** 1222–1236. doi:10.1016/j.ajhg.2019.10.014

Darvasi A, Soller M. 1995. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics* **141:** 1199–1207. doi:10.1093/genetics/141.3.1199

Degner JF, Marioni JC, Pai AA, Pickrell JK, Nkadori E, Gilad Y, Pritchard JK. 2009. Effect of read-mapping biases on detecting allele-specific expression from RNA-sequencing data. *Bioinformatics* **25:** 3207–3212. doi:10.1093/bioinformatics/btp579

Deng Q, Ramsköld D, Reinius B, Sandberg R. 2014. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343:** 193–196. doi:10.1126/science.1245316

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21. doi:10.1093/bioinformatics/bts635

Ehrich TH, Kenney JP, Vaughn TT, Pletscher LS, Cheverud JM. 2003. Diet, obesity, and hyperglycemia in LG/J and SM/J mice. *Obes Res* **11:** 1400–1410. doi:10.1038/oby.2003.189

Ehrich TH, Hrbek T, Kenney-Hunt JP, Pletscher LS, Wang B, Semenkovich CF, Cheverud JM. 2005. Fine-mapping gene-by-diet interactions on chromosome 13 in a LG/J × SM/J murine model of obesity. *Diabetes* **54:** 1863–1872. doi:10.2337/diabetes.54.6.1863

Falkowski M, Schledzewski K, Hansen B, Goerdt S. 2003. Expression of stabilin-2, a novel fasciclin-like hyaluronan receptor protein, in murine sinusoidal endothelia, avascular tissues, and at solid/liquid interfaces. *Histochem Cell Biol* **120:** 361–369. doi:10.1007/s00418-003-0585-5

Gai Z, Visentin M, Hiller C, Krajnc E, Li T, Zhen J, Kullak-Ublick GA. 2016. Organic cation transporter 2 overexpression may confer an increased risk of gentamicin-induced nephrotoxicity. *Antimicrob Agents Chemother* **60:** 5573–5580. doi:10.1128/AAC.00907-16

Ge B, Pokholok DK, Kwan T, Grundberg E, Morcos L, Verlaan DJ, Le J, Koka V, Lam KCL, Gagné V, et al. 2009. Global patterns of *cis* variation in human cells revealed by high-density allelic expression analysis. *Nat Genet* **41:** 1216–1222. doi:10.1038/ng.473

Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. 2007. Widespread monoallelic expression on human autosomes. *Science* **318:** 1136–1140. doi:10.1126/science.1148910

Gonzales NM, Seo J, Cordero AIH, Pierre CLS, Gregory JS, Distler MG, Abney M, Canzar S, Lionikas A, Palmer AA. 2018. Genome wide association analysis in a mouse advanced intercross line. *Nat Commun* **9:** 5162. doi:10.1038/s41467-018-07642-8

Hager R, Cheverud JM, Wolf JB. 2008. Maternal effects as the cause of parent-of-origin effects that mimic imprinting. *Genetics* **178:** 1755–1762. doi:10.1534/genetics.107.080697

Hammond JW, Cai D, Verhey KJ. 2008. Tubulin modifications and their cellular functions. *Curr Opin Cell Biol* **20:** 71–76. doi:10.1016/j.ceb.2007.11.010

He H, Kim J. 2014. Regulation and function of the Peg3 imprinted domain. *Genomics Inform* **12:** 105–113. doi:10.5808/GI.2014.12.3.105

He W, Rose DW, Olefsky JM, Gustafson TA. 1998. Grb10 interacts differentially with the insulin receptor, insulin-like growth factor I receptor, and epidermal growth factor receptor via the Grb10 Src homology 2 (SH2)

domain and a second novel domain located between the Pleckstrin homology and SH2 domains. *J Biol Chem* **273:** 6860–6867. doi:10.1074/jbc.273.12.6860

Herrera-Marcos LV, Sancho-Knapik S, Gabás-Rivera C, Barranquero C, Gascón S, Romanos E, Martínez-Beamonte R, Navarro MA, Surra JC, Arnal C, et al. 2020. *Pgc1a* is responsible for the sex differences in hepatic *Cidec/Fsp27β* mRNA expression in hepatic steatosis of mice fed a Western diet. *Am J Physiol Endocrinol Metab* **318:** E249–E261. doi:10.1152/ajpendo.00199.2019

Hochner H, Allard C, Granot-Hershkovitz E, Chen J, Sitlani CM, Sazdovska S, Lumley T, McKnight B, Rice K, Enquobahrie DA, et al. 2015. Parent-of-origin effects of the *APOB* gene on adiposity in young adults. *PLoS Genet* **11:** e1005573. doi:10.1371/journal.pgen.1005573

Inoue A, Jiang L, Lu F, Suzuki T, Zhang Y. 2017. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547:** 419–424. doi:10.1038/nature23262

Iv THR, Banerjee S, Sharma VM, Donohue J, Couldwell S, Sosinsky A, Frulla A, Robinson A, Puri V. 2015. Effects of a high fat diet and voluntary wheel running exercise on Cidea and Cidec expression in liver and adipose tissue of mice. *PLoS One* **10:** e0130259. doi:10.1371/journal.pone.0130259

Johansson M. 2003. Identification of a novel human uridine phosphorylase. *Biochem Biophys Res Commun* **307:** 41–46. doi:10.1016/S0006-291X(03)01062-3

Kappil M, Lambertini L, Chen J. 2015. Environmental influences on genomic imprinting. *Curr Environ Health Rep* **2:** 155–162. doi:10.1007/s40572-015-0046-z

Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, Heger A, Agam A, Slater G, Goodson M, et al. 2011. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477:** 289–294. doi:10.1038/nature10413

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28:** 1647–1649. doi:10.1093/bioinformatics/bts199

Kekuda R, Prasad PD, Wu X, Wang H, Fei YJ, Leibach FH, Ganapathy V. 1998. Cloning and functional characterization of a potential-sensitive, polyspecific organic cation transporter (OCT3) most abundantly expressed in placenta. *J Biol Chem* **273:** 15971–15979. doi:10.1074/jbc.273.26.15971

Knowles DA, Davis JR, Edgington H, Raj A, Favé MJ, Zhu X, Potash JB, Weissman MM, Shi J, Levinson DF, et al. 2017. Allele-specific expression reveals interactions between genetic variation and environment. *Nat Methods* **14:** 699–702. doi:10.1038/nmeth.4298

Lawson HA, Zelle KM, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, Semenkovich CF, et al. 2010. Genetic, epigenetic, and gene-by-diet interaction effects underlie variation in serum lipids in a LG/J×SM/J murine model. *J Lipid Res* **51:** 2976–2984. doi:10.1194/jlr.M006957

Lawson HA, Lee CE, Partridge C, Wolf JB, Semenkovich CF, Cheverud JM. 2011a. Genetic effects at pleiotropic loci are context-dependent with consequences for the maintenance of genetic variation in populations. *PLoS Genet* **7:** e1002256. doi:10.1371/journal.pgen.1002256

Lawson HA, Lee A, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, Semenkovich CF, et al. 2011b. The importance of context to the genetic architecture of diabetes-related traits is revealed in a genome-wide scan of a LG/J × SM/J murine model. *Mamm Genome* **22:** 197–208. doi:10.1007/s00335-010-9313-3

Lawson HA, Cheverud JM, Wolf JB. 2013. Genomic imprinting and parent-of-origin effects on complex traits. *Nat Rev Genet* **14:** 609–617. doi:10.1038/nrg3543

Lee CS, Buttitta L, Fan CM. 2001. Evidence that the WNT-inducible *growth arrest-specific gene 1* encodes an antagonist of sonic hedgehog signaling in the somite. *Proc Natl Acad Sci* **98:** 11347–11352. doi:10.1073/pnas.201418298

Lee BT, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee CM, et al. 2022. The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res* **50:** D1115–D1122. doi:10.1093/nar/gkab959

Leung D, Jung I, Rajagopal N, Schmitt A, Selvaraj S, Lee AY, Yen CA, Lin S, Lin Y, Qiu Y, et al. 2015. Integrative analysis of haplotype-resolved epigenomes across human tissues. *Nature* **518:** 350–354. doi:10.1038/nature14217

Li D, Wei T, Abbott CM, Harrich D. 2013. The unexpected roles of eukaryotic translation elongation factors in RNA virus replication and pathogenesis. *Microbiol Mol Biol Rev* **77:** 253–266. doi:10.1128/MMBR.00059-12

Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. 2019. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res* **47:** W199–W205. doi:10.1093/nar/gkz401

Lionikas A, Cheng R, Lim JE, Palmer AA, Blizard DA. 2010. Fine-mapping of muscle weight QTL in LG/J and SM/J intercrosses. *Physiol Genomics* **42A:** 33–38. doi:10.1152/physiolgenomics.00100.2010

Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH, Lee MP. 2003. Allelic variation in gene expression is common in the human genome. *Genome Res* **13:** 1855–1862. doi:10.1101/gr.1006603

Macias-Velasco JF, St. Pierre CL, Wayhart JP, Yin L, Spears L, Miranda MA, Carson C, Funai K, Cheverud J, Semenkovich CF, et al. 2022. Parent-of-origin effects propagate through networks to shape metabolic traits. *eLife* **11:** e72989. doi:10.7554/eLife.72989

Maldonato BJ, Russell DA, Totah RA. 2021. Human METTL7B is an alkyl thiol methyltransferase that metabolizes hydrogen sulfide and captopril. *Sci Rep* **11:** 4857. doi:10.1038/s41598-021-84218-5

Mateyak MK, Kinzy TG. 2010. eEF1A: thinking outside the ribosome. *J Biol Chem* **285:** 21209–21213. doi:10.1074/jbc.R110.113795

McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* **40:** 4288–4297. doi:10.1093/nar/gks042

Miranda MA, St. Pierre CL, Macias-Velasco JF, Nguyen HA, Schmidt H, Agnello LT, Wayhart JP, Lawson HA. 2019. Dietary iron interacts with genetic background to influence glucose homeostasis. *Nutr Metab* **16:** 13. doi:10.1186/s12986-019-0339-6

Morcos L, Ge B, Koka V, Lam KC, Pokholok DK, Gunderson KL, Montpetit A, Verlaan DJ, Pastinen T. 2011. Genome-wide assessment of imprinted expression in human cells. *Genome Biol* **12:** R25. doi:10.1186/gb-2011-12-3-r25

Moyerbrailean GA, Richards AL, Kurtz D, Kalita CA, Davis GO, Harvey CT, Alazizi A, Watza D, Sorokin Y, Hauff N, et al. 2016. High-throughput allele-specific expression across 250 environmental conditions. *Genome Res* **26:** 1627–1638. doi:10.1101/gr.209759.116

Nagata C, Kobayashi H, Sakata K, Satomi K, Minami Y, Morishita Y, Ohara R, Yoshikawa H, Arai Y, Nishida M, et al. 2012. Increased expression of OCIA domain containing 2 during stepwise progression of ovarian mucinous tumor. *Pathol Int* **62:** 471–476. doi:10.1111/j.1440-1827.2012.02825.x

Najjar SM, Perdomo G. 2019. Hepatic insulin clearance: mechanism and physiology. *Physiology* **34:** 198–215. doi:10.1152/physiol.00048.2018

Nikolskiy I, Conrad DF, Chun S, Fay JC, Cheverud JM, Lawson HA. 2015. Using whole-genome sequences of the LG/J and SM/J inbred mouse strains to prioritize quantitative trait genes and nucleotides. *BMC Genomics* **16:** 415. doi:10.1186/s12864-015-1592-3

Olofsson SO, Borèn J. 2005. Apolipoprotein B: a clinically important apolipoprotein which assembles atherogenic lipoproteins and promotes the development of atherosclerosis. *J Intern Med* **258:** 395–410. doi:10.1111/j.1365-2796.2005.01556.x

Parker CC, Carbonetto P, Sokoloff G, Park YJ, Abney M, Palmer AA. 2014. High-resolution genetic mapping of complex traits from a combined analysis of F2 and advanced intercross mice. *Genetics* **198:** 103–116. doi:10.1534/genetics.114.167056

Pastinen T. 2010. Genome-wide allele-specific analysis: insights into regulatory variation. *Nat Rev Genet* **11:** 533–538. doi:10.1038/nrg2815

Pivovarova O, Bernigau W, Bobbert T, Isken F, Möhlig M, Spranger J, Weickert MO, Osterhoff M, Pfeiffer AFH, Rudovich N. 2013. Hepatic insulin clearance is closely related to metabolic syndrome components. *Diabetes Care* **36:** 3779–3785. doi:10.2337/dc12-1203

Plasschaert RN, Bartolomei MS. 2015. Tissue-specific regulation and function of Grb10 during growth and neuronal commitment. *Proc Natl Acad Sci* **112:** 6841–6847. doi:10.1073/pnas.1411254111

Qazi IH, Cao Y, Yang H, Angel C, Pan B, Zhou G, Han H. 2021. Impact of dietary selenium on modulation of expression of several non-selenoprotein genes related to key ovarian functions, female fertility, and proteostasis: a transcriptome-based analysis of the aging mice ovaries. *Biol Trace Elem Res* **199:** 633–648. doi:10.1007/s12011-020-02192-x

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842. doi:10.1093/bioinformatics/btq033

Reik W, Walter J. 2001. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* **2:** 21–32. doi:10.1038/35047554

Reinius B, Sandberg R. 2015. Random monoallelic expression of autosomal genes: stochastic transcription and allele-level regulation. *Nat Rev Genet* **16:** 653–664. doi:10.1038/nrg3888

Rice LM, Montabana EA, Agard DA. 2008. The lattice as allosteric effector: structural studies of αβ- and γ-tubulin clarify the role of GTP in microtubule assembly. *Proc Natl Acad Sci* **105:** 5378–5383. doi:10.1073/pnas.0801155105

Rivas MA, Pirinen M, Conrad DF, Lek M, Tsang EK, Karczewski KJ, Maller JB, Kukurba KR, DeLuca DS, Fromer M, et al. 2015. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* **348:** 666–669. doi:10.1126/science.1261877

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26:** 139–140. doi:10.1093/bioinformatics/btp616

Roosild TP, Castronovo S, Villoso A, Ziemba A, Pizzorno G. 2011. A novel structural mechanism for redox regulation of uridine phosphorylase 2 activity. *J Struct Biol* **176:** 229–237. doi:10.1016/j.jsb.2011.08.002

Shao L, Xing F, Xu C, Zhang Q, Che J, Wang X, Song J, Li X, Xiao J, Chen LL, et al. 2019. Patterns of genome-wide allele-specific expression in hybrid rice and the implications on the genetic basis of heterosis. *Proc Natl Acad Sci* **116:** 5653–5658. doi:10.1073/pnas.1820513116

Sinha S, Bheemsetty VA, Inamdar MS. 2018. A double helical motif in OCIAD2 is essential for its localization, interactions and STAT3 activation. *Sci Rep* **8:** 7362. doi:10.1038/s41598-018-25667-3

Sud N, Zhang H, Pan K, Cheng X, Cui J, Su Q. 2017. Aberrant expression of microRNA induced by high-fructose diet: implications in the pathogenesis of hyperlipidemia and hepatic insulin resistance. *J Nutr Biochem* **43:** 125–131. doi:10.1016/j.jnutbio.2017.02.003

Takada Y, Miyagi R, Takahashi A, Endo T, Osada N. 2017. A generalized linear model for decomposing *cis*-regulatory, parent-of-origin, and maternal effects on allele-specific gene expression. *G3 (Bethesda)* **7:** 2227–2234. doi:10.1534/g3.117.042895

Thiaville MM, Huang JM, Kim H, Ekram MB, Roh TY, Kim J. 2013. DNA-binding motif and target genes of the imprinted transcription factor PEG3. *Gene* **512:** 314–320. doi:10.1016/j.gene.2012.10.005

Tokarz VL, MacDonald PE, Klip A. 2018. The cell biology of systemic insulin function. *J Cell Biol* **217:** 2273–2289. doi:10.1083/jcb.201802095

Trefts E, Gannon M, Wasserman DH. 2017. The liver. *Curr Biol* **27:** R1147–R1151. doi:10.1016/j.cub.2017.09.019

Turró S, Ingelmo-Torres M, Estanyol JM, Tebar F, Fernández MA, Albor CV, Gaus K, Grewal T, Enrich C, Pol A. 2006. Identification and characterization of associated with lipid droplet protein 1: a novel membrane-associated protein that resides on hepatic lipid droplets. *Traffic* **7:** 1254–1269. doi:10.1111/j.1600-0854.2006.00465.x

Umlauf D, Goto Y, Cao R, Cerqueira F, Wagschal A, Zhang Y, Feil R. 2004. Imprinting along the *Kcnq1* domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* **36:** 1296–1300. doi:10.1038/ng1467

Wang X, Clark AG. 2014. Using next-generation RNA sequencing to identify imprinted genes. *Heredity (Edinb)* **113:** 156–166. doi:10.1038/hdy.2014.18

Wienholz BL, Kareta MS, Moarefi AH, Gordon CA, Ginno PA, Chédin F. 2010. DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B *in vivo*. *PLoS Genet* **6:** e1001106. doi:10.1371/journal.pgen.1001106

Wolf JB, Cheverud JM, Roseman C, Hager R. 2008. Genome-wide analysis reveals a complex pattern of genomic imprinting in mice. *PLoS Genet* **4:** e1000091. doi:10.1371/journal.pgen.1000091

Xu L, Zhou L, Li P. 2012. CIDE proteins and lipid metabolism. *Arterioscler Thromb Vasc Biol* **32:** 1094–1098. doi:10.1161/ATVBAHA.111.241489

Yee TW. 2010. The VGAM package for categorical data analysis. *J Stat Softw* **32:** 1–34. doi:10.18637/jss.v032.i10

Zhou B, Weigel JA, Saxena A, Weigel PH. 2002. Molecular cloning and functional expression of the rat 175-kDa hyaluronan receptor for endocytosis. *Mol Biol Cell* **13:** 2853–2868. doi:10.1091/mbc.02-03-0048