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], Heather A Lawson[1,*]

CL St. Pierre: stpierrec@wustl.edu
JF Macias-Velasco: juanfmacias@wustl.edu
JP Wayhart: jwayhart@genetics.wustl.edu
L Yin: liyin@wustl.edu
CF Semenkovich: csemenko@wustl.edu
HA Lawson: hlawson@genetics.wustl.edu

[1] Department of Genetics, Washington University School of Medicine, 660 South Euclid Ave, Saint Louis, MO, USA
[2] Department of Medicine, Washington University School of Medicine, 660 South Euclid Ave, Saint Louis, MO, USA

*Corresponding author
660 South Euclid Ave
Campus Box 8232
Saint Louis, MO, 63110
ph: 314-362-7269, fax: 314-362-7855

Running Title: Metabolic tissues and environments influence allele-specific expression patterns

Keywords: allele-specific expression, parent-of-origin effects, gene-by-environment effects, *cis*-regulation, tissue-specificity, mouse model, RNA-Seq

ABSTRACT

Allele-specific expression (**ASE**) is a phenomenon where one allele is preferentially expressed over the other. Genetic and epigenetic factors cause ASE by altering the final allelic composition of a gene's product, leading to expression imbalances that can have functional consequences on phenotypes. Environmental signals also impact allele-specific gene regulation, but how they contribute to this crosstalk remains understudied. Here, we explored how allelic genotype, parent-of-origin, tissue type, sex, and dietary fat simultaneously influence ASE biases in a F_1 reciprocal cross mouse model. Male and female mice from a F_1 reciprocal cross of the LG/J and SM/J strains were fed a high-fat or low-fat diet. We harnessed strain-specific variants to distinguish between two classes of ASE: parent-of-origin dependent (unequal expression based on an allele's parental origin) and sequence dependent (unequal expression based on an allele's nucleotide identity). We present a comprehensive genome-wide map of ASE patterns across three metabolically-relevant tissues and nine environmental contexts. We find that both ASE classes are highly dependent on tissue type and environmental context. They vary across metabolic tissues, between males and females, and in response to dietary fat levels. Surprisingly, we also find several genes with inconsistent ASE biases that switched direction across tissues and/or contexts (e.g. SM/J biased in one cohort, LG/J biased in another). Together, our results provide novel insights into how genetic, epigenetic, and environmental mechanisms govern allele-specific gene regulation, which is an essential step towards deciphering the genotype to phenotype map.

INTRODUCTION

Deciphering the genotype to phenotype map remains a fundamental quest in biology. Gene expression is a promising focal point, as it is an intermediate step between DNA sequence and gross phenotype. Gene expression itself is a complex trait that is regulated by genetic, epigenetic, and environmental factors (Pastinen 2010). Recent efforts to characterize how genes are regulated have often focused on mapping expression quantitative trait loci (**eQTLs**), which identify genetic variants linked to changes in gene expression at the population level (Cookson et al. 2009). However, these studies can only interrogate total gene expression and assume that genes are biallelically expressed (i.e. both alleles are equally expressed), which may mask underlying regulatory mechanisms. Furthermore, epigenetic changes and gene-by-environment interactions can 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. 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 where a gene's expression diverges from biallelic. In diploid organisms, one allele is preferentially expressed over the other allele. 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; Castel et al. 2019; Keane et al. 2011). Depending on a gene's function, these expression imbalances can lead to phenotypic variation with functional consequences. To detect ASE, single nucleotide polymorphisms (**SNPs**) and other genetic variants are used to distinguish between alleles in heterozygotes and map RNA-sequencing reads to their chromosome of origin (Heap et al. 2010). Allele-specific analyses are a powerful way to exploit a within-sample control (the other allele) to gauge how genetic and epigenetic variation in *cis*-regulatory elements shapes gene expression (Pastinen 2010).

ASE can be divided into two classes: sequence dependent or parent-of-origin dependent (**Figure 1**). Sequence dependent ASE refers to cases where the two 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 (Cavalli et al. 2016; Wienholz et al. 2010).

In contrast, parent-of-origin dependent ASE refers to cases where the alleles are differentially expressed based on which parent contributed it, regardless of the underlying sequence. These patterns fall under the umbrella 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 where one parent's allele is completely silenced via selective DNA methylation (Barlow and Bartolomei 2014). Known imprinted genes comprise ~1% of the human and mouse genomes, yet play important roles in development, metabolism, cognition, and other complex traits (Reik and Walter 2001).

Comprehensive atlases of how both classes of ASE vary between tissues and developmental stages have been generated in human (Leung et al. 2015; Castel et al. 2019) and mouse models (Babak et al. 2015; Andergassen et al. 2017). These studies reveal that both parent-of-origin and sequence dependent ASE patterns are not 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 (Buil et al. 2015; Knowles et al. 2017; Moyerbrailean et al. 2016) as well as in rice models (Shao et al. 2019). Imprinted genes are also known to be responsive to environmental exposures (such as teratogenic agents and maternal nutrition) during fetal development (Kappil et al. 2015). Together, these findings suggest that a complicated crosstalk among 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 allele-specific expression patterns in a F_1 reciprocal cross of the LG/J and SM/J inbred mouse strains. These strains have been extensively used in gene-by-environment studies due to their divergent genetic backgrounds and variable responses to dietary nutrition (Ehrich et al. 2003; Nikolskiy et al. 2015; Lawson et al. 2011b, 2010, 2011a; Carson and Lawson 2020; Miranda et al. 2019). We found that both parent-of-origin and sequence dependent ASE patterns are highly dependent on tissue type and environmental context. They vary across metabolic tissues, between males and females, and in response to dietary fat, thus providing novel insights into how gene-by-environmental effects influence complex traits. Untangling the genetic, epigenetic, and environmental mechanisms that govern allele-specific gene regulation is crucial to improving our ability to predict phenotypes from genotypes.

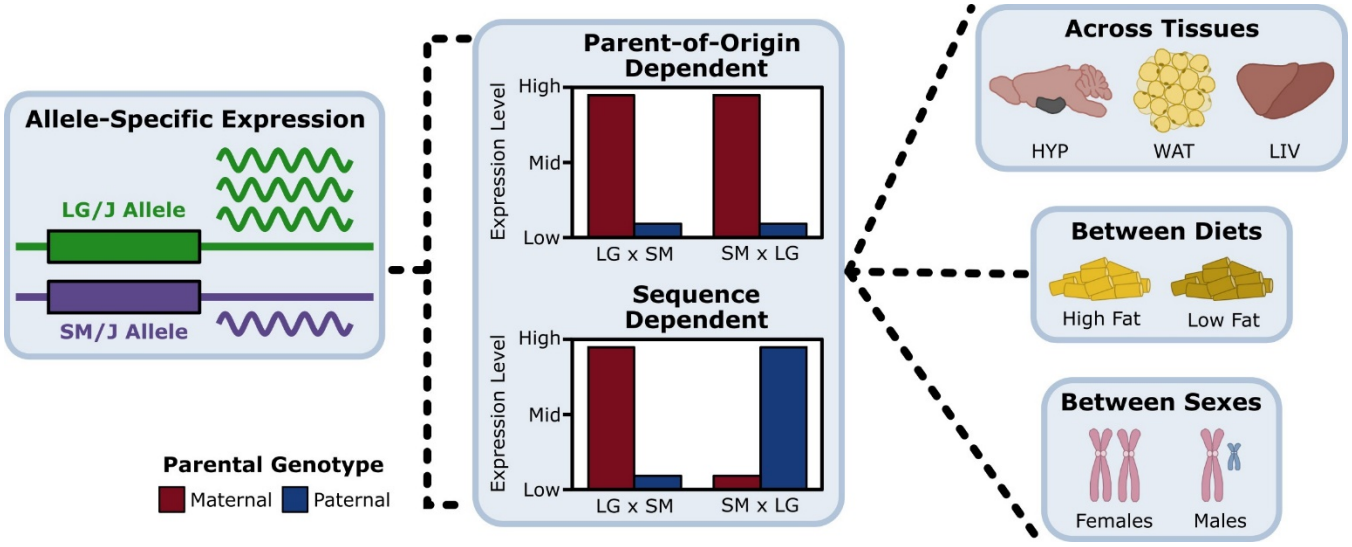

Figure 1: Evaluating parent-of-origin and sequence dependent allele-specific expression across tissues and environmental contexts. We decomposed allele-specific expression (ASE) into its parent-of-origin and sequence effects in a F_1 reciprocal cross. An example of parent-of-origin dependent ASE is when 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. Once we identified significant ASE genes, we compared how their expression patterns changed across metabolic tissues (HYP, WAT, LIV), in response to different diets (high fat, low fat), and between sexes (females, males).

RESULTS

Allele-specific expression can be decomposed into parent-of-origin and sequence effects

We measured ASE in a F_1 reciprocal cross of the LG/J and SM/J inbred mouse 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-fat or low-fat diet. We obtained RNA-Seq data from three metabolically-relevant tissues: hypothalamus (**HYP**), white adipose (**WAT**), and liver (**LIV**). To explore how different environmental contexts (dietary fat and/or sex) impact ASE patterns, we analyzed nine separate cohorts per tissue: high-fat fed diet (**H**), low-fat fed diet (**L**), females (**F**), males (**M**), high-fat fed females (**HF**), high-fat fed males (**HM**), low-fat fed females (**LF**), low-fat fed males (**LM**), and all contexts collapsed (**All**) (**Figure 1**).

We harnessed the >6 million SNPs and indels between the LG/J and SM/J genomes (Nikolskiy et al. 2015) to map sequencing reads to their chromosome of origin. Overall, 9,016 protein-coding genes and noncoding RNAs had detectable ASE in at least one tissue-by-context analysis (~31% of all expressed genes, **Supplemental Figure S1**). Next, we identified genes with significant ASE biases and classified them into two patterns: parent-of-origin dependent (unequal expression based on the allele's parental origin) and sequence dependent (unequal expression based on the allele's nucleotide identity) (**Figure 1, Supplemental Figure S2**).

Parent-of-origin and sequence dependent ASE patterns are prevalent and distinct

Across our 27 tissue-by-context analyses, we identified 271 genes with significant parent-of-origin dependent ASE. HYP had the greatest number of significant genes (n = 229), followed by WAT (n = 35), then LIV (n = 27). 14 genes were expressed in multiple tissues, but the majority were tissue-specific (**Figure 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 (**Figure 2B**). 214 genes (79%) were paternally biased, 55 genes (20%) were maternally biased, and 2 genes (1%) switched their expression bias direction across the cohorts (**Figure 2C**). This heavy skew is driven by a 672 kb cluster of 171 paternally-biased small nucleolar RNAs (snoRNAs) located within the Prader-Willi/Angelman syndrome (**PWS/AS**) orthologous domain on mouse chromosome 7 that are only expressed in HYP (**Supplemental Figure S3**). Parent-of-origin dependent ASE genes also included protein-coding genes, microRNAs, long noncoding RNAs, long interspersed noncoding RNAs, and pseudogenes (**Figure 2D**).

We also identified 2,673 genes with significant sequence dependent ASE across our 27 tissue-by-context analyses. WAT had the greatest number of significant genes (n = 1,495), followed by LIV (n = 1,486), then HYP (n = 1,264). While some genes' biases were tissue-specific, 1,657 genes (62%) were biased in multiple tissues (**Figure 2E**). In each tissue, the most genes were detected in the "All" context, then the diet- or sex-specific contexts, and finally the diet-by-sex-specific contexts, likely reflecting the sample sizes and power available in each cohort (**Figure 2F**). 1,218 genes (46%) were SM/J biased, 1,410 genes (53%) were LG/J biased, and 45 genes (2%) switched their expression bias direction across the cohorts (**Figure 2G**). Sequence dependent ASE genes were predominantly classified as protein-coding genes (~80%) in each tissue, but also included pseudogenes, immunoglobulins, and various non-coding RNAs (long non-coding, long interspersed non-coding, micro, ribosomal, small interfering, and small nucleolar) (**Figure 2H**).

Parent-of-origin and sequence dependent ASE patterns were often mutually exclusive. In all three tissues, genes with extreme parental biases typically had weak sequence biases and vice versa (**Figure 2I, Supplemental Figure S4)**. However, 61 genes in HYP and 3 genes in WAT showed both parental and sequence biases, likely due to epigenetic regulatory mechanisms affected by haplotype variation/allelic identity. Both ASE patterns occurred genome-wide (**Supplemental Figure S5**). Parent-of-origin dependent ASE genes tended to cluster in well-known imprinted domains, such as the *Ube3a/Snrpn*, *Meg3/Gtl2*, *Peg3/Usp29*, and *H13/Mcts2* domains (**Supplemental Figures S6**). Sequence dependent ASE genes were spread more diffusely across chromosomes, but occasionally clustered in regions potentially controlled by the same regulatory element (**Supplemental Figures S7**).

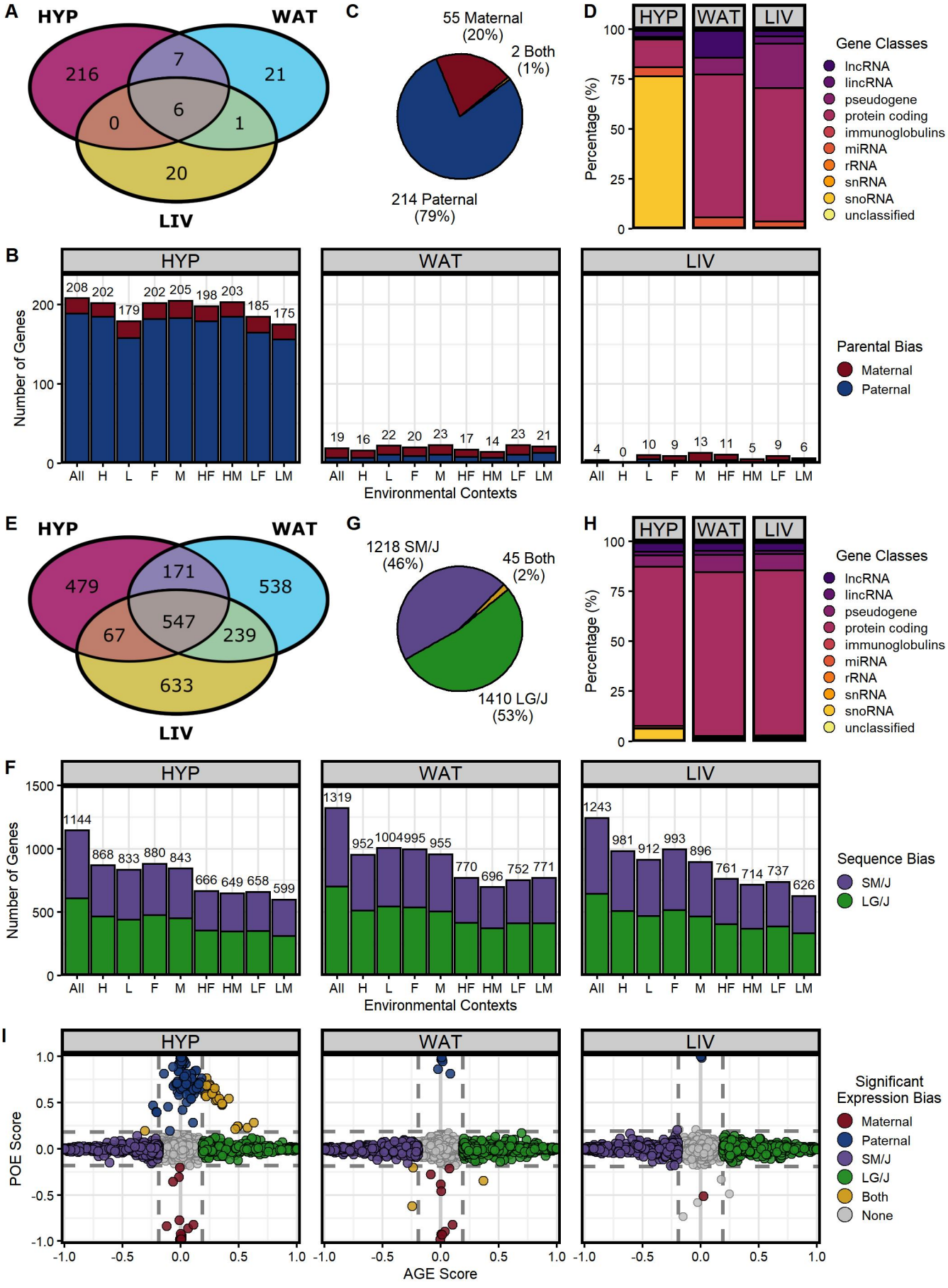

Figure 2: Parent-of-origin and sequence dependent allele-specific expression patterns are prevalent and distinct. (A) Venn diagram of the total parent-of-origin dependent ASE genes across HYP, LIV, and WAT (all contexts collapsed). **(B)** Number of genes with significant parental biases in each tissue-by-context analysis (maternal = red, paternal = blue): all contexts collapsed (All), high-fat diet (H), low-fat diet (L), females (F), males (M), high-fat fed females (HF), high-fat fed males (HM), low-fat fed females (LF), and low-fat fed males (LM). **(C)** Summary of expression bias directions across all analyses: paternally biased (blue), maternally biased (red), and genes that switch bias direction depending on the cohort (yellow). **(D)** Gene class proportions of significant parent-of-origin dependent ASE genes in each tissue. **(E)** Venn diagram of the total sequence dependent ASE genes across HYP, LIV, and WAT (all contexts collapsed). **(F)** Number of genes with significant sequence biases in each tissue-by-context analysis (SM/J = purple, LG/J = green). **(G)** Summary of expression bias directions across all analyses: SM/J biased (purple), LG/J biased (green), and genes that switch bias direction depending on the cohort (yellow). **(H)** Gene class proportions of significant sequence dependent ASE genes in each tissue. **(I)** Parent-of-Origin Effect (POE) versus Allelic Genotype Effect (AGE) scores in the "All" context of each tissue. Genes with significant parental biases (red = maternal, blue = paternal) have extreme POE scores, but weak AGE scores. Conversely, genes with significant sequence biases (purple = SM/J, green = LG/J) have extreme AGE scores, but weak POE scores. Some genes have both parental and sequence biases (yellow). Most genes have no bias (gray). Dashed lines indicate effect score thresholds.

Parent-of-origin dependent ASE recapitulates canonical imprinting patterns

To evaluate whether parent-of-origin dependent ASE is influenced by tissue or environmental context (diet and/or sex), we characterized the expression profiles of the 271 parentally biased genes across our 27 tissue-by-context analyses (3 tissues x 9 diet-by-sex contexts). For each analysis, we quantified the direction and magnitude of each gene's parental expression bias by calculating a Parent-of-Origin Effect (**POE**) score from the mean allelic bias of each F_1 reciprocal cross. POE scores range from -1 (completely maternally expressed) to +1 (completely paternally expressed); a score of 0 indicates biallelic expression (see Methods). In each tissue-by-context analysis, a gene could be expressed in one of three ways: significant parental bias, biallelic (expressed but with no allele-specific bias), or not expressed. We sorted the 271 genes with significant parent-of-origin dependent ASE into three expression profiles: tissue-independent, tissue-dependent, and context-dependent (**Figure 3A-3D**).

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 (**Figure 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) (**Figure 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. 29 genes were canonically imprinted (bolded in **Figure 3**) and the other 169 genes were various non-coding RNAs located within well-known imprinted domains. For example, 156 genes were part of a cluster of 171 paternally-biased snoRNAs in the PWS/AS domain on mouse chromosome 7 (**Supplemental Figure S3**). As expected with imprinting, these genes were extremely biased towards one parent's allele; their mean POE scores were -0.82 and 0.71 for maternally and paternally biased genes, respectively (**Figure 3E**). HYP had the highest proportion of these tissue-independent/dependent genes among our three adult tissues (85%), largely due to the snoRNA cluster in the PWS/AS domain being HYP-specific. WAT had the second highest proportion (49%), but very few of these genes were expressed in LIV (7%) (**Figure 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).

We validated the expression profiles of two canonically imprinted genes (*Peg3* and *Grb10*) by pyrosequencing (**Figure 3B-3C**, **Supplemental Figure S8**). *Peg3* (paternally expressed gene 3) had a significant paternal bias in all three tissues, regardless of context. Its locus has 60 variants between the LG/J and SM/J backgrounds, though none are predicted to be 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). *Grb10* (growth factor receptor bound protein 10) had a tissue-dependent pattern of significant paternal bias in HYP but a maternal bias in WAT and LIV. Its locus has 154 variants between the LG/J and SM/J backgrounds, but none are predicted to be 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).

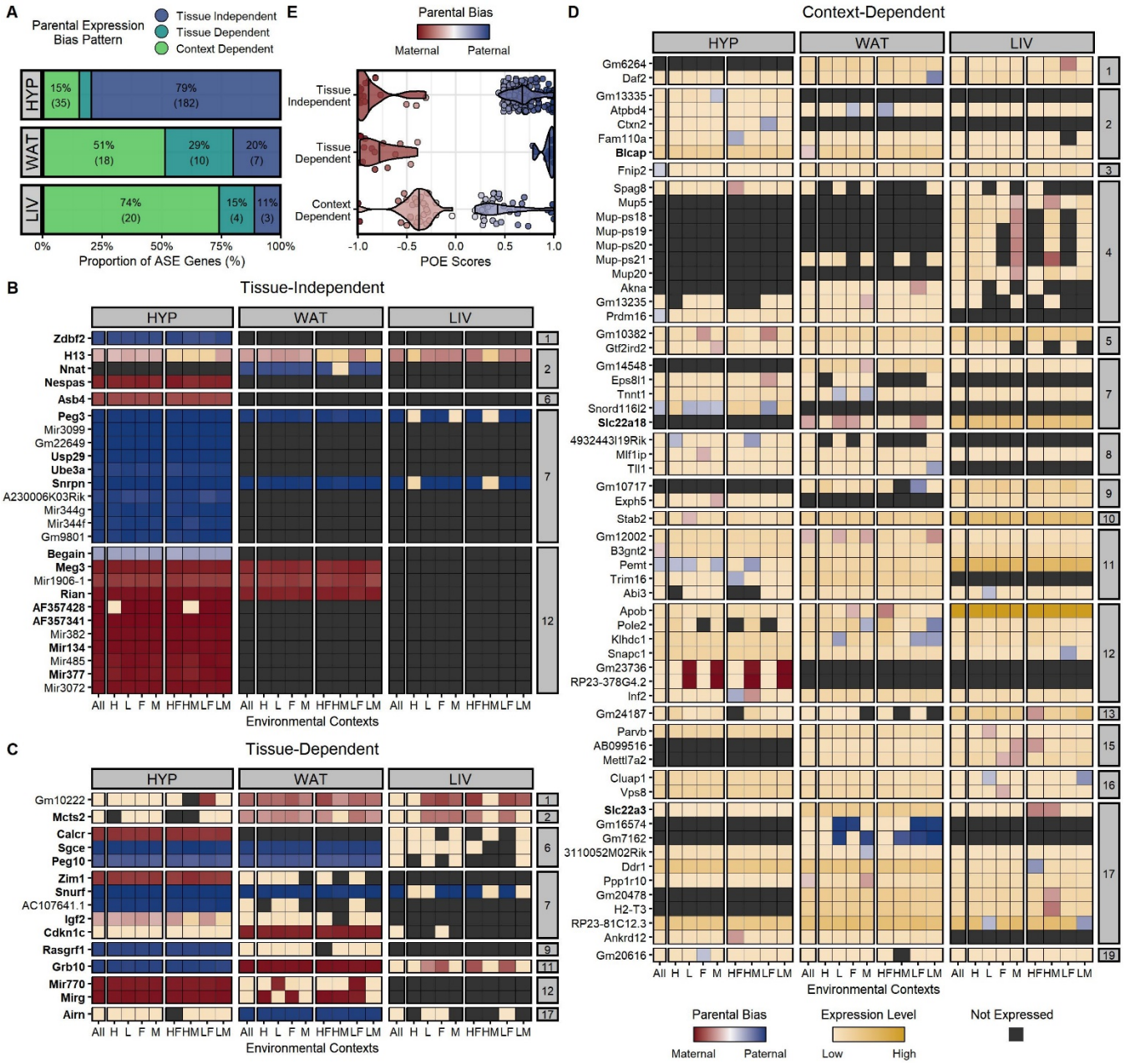

Figure 3: Atlas of parent-of-origin dependent ASE patterns across tissues and environmental contexts. (A) Number and proportion of ASE genes per tissue with each parental expression bias profile: tissue-independent (dark blue), tissue-dependent (teal), and context-dependent (green). Heatmaps of ASE profiles across tissues and contexts: **(B)** tissue-independent (parental bias wherever expressed), **(C)** tissue-dependent (parental bias in some tissues, no bias in others), and **(D)** context-dependent (parental bias only in certain diet-by-sex contexts, no bias elsewhere). All genes with significant parent-of-origin dependent ASE are shown, except the cluster of 171 snoRNAs on chr 7 (**Supplemental Figure S3**). 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). Where genes are not significantly biased, shades of yellow indicate their biallelic gene expression levels (log-transformed total counts). Black indicates genes are not expressed in that context/tissue. Bolded genes are canonically imprinted. The y-axis is grouped and sorted by chromosomal position. Each supercolumn denotes a tissue: hypothalamus (HYP), white adipose (WAT), and liver (LIV). Each subcolumn denotes an environmental context: all contexts collapsed (All), high-fat diet (H), low-fat diet (L), females (F), males (M), high-fat fed females (HF), high-fat fed males (HM), low-fat fed females (LF), and low-fat fed males

(LM). **(E)** Parent-of-Origin Effect (POE) score distributions for each parental expression bias profile. Vertical lines indicate that profile's mean POE score. Dots represent individual significant ASE genes.

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 (**Figure 3D**). Only three of these genes were canonically imprinted and 18 genes were non-coding 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 context-dependent 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 (**Figure 3E**). These patterns are consistent with partial imprinting, where the two parental alleles are differentially expressed in a less extreme manner than the uniparental expression associated with genomic imprinting (Wolf et al. 2008; Morcos et al. 2011). Interestingly, LIV had the highest proportion of context-dependent genes among its total parent-of-origin dependent ASE genes (74%) while HYP had the lowest proportion (15%) (**Figure 3A**). LIV is also a tissue where more parent-of-origin dependent ASE genes were detected in the diet- and/or sex-specific contexts than the "All" context (**Figure 2B**). LIV is incredibly 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 intrinsic (sex) and/or extrinsic (dietary fat) environments can alter parental ASE biases, we calculated individualized POE scores for each context-dependent gene and modeled how they vary across diet, sex, and diet-by-sex contexts. These categories were not mutually exclusive; across all three tissues, most context-dependent 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 ASE biases (**Supplemental Figure S9**). For example, LIV had similar proportions of significant sex, diet, and diet-by-sex effects (each 75 – 80% of its context-

dependent genes); 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, while 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 (**Figure 3D**, **Supplemental Figure S10**). *Apob* (apolipoprotein B) had a significant diet effect in WAT, reflected in significant maternal biases in the HF and F contexts. It also had a maternal bias in the HM context, but low sample sizes excluded it from further analysis. *Apob* was biallelically expressed in the remaining contexts in WAT and all contexts in LIV and HYP. Its locus has 126 variants between the LG/J and SM/J backgrounds, including seven non-synonymous SNPs in the coding region (six = LG/J genome, one = SM/J genome). *Apob* produces the main component of lipoproteins, which transport lipids (including cholesterol) 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 levels also differ between high and low fructose diets in mice livers, suggesting its function is susceptible to dietary environment (Sud et al. 2017). *Slc22a3* (solute carrier family 22 member 3) had significant diet and diet-by-sex effects in LIV, reflected in strong maternal biases in the HF and HM contexts. *Slc22a3* was biallelically expressed in the remaining contexts in LIV and all contexts in WAT and HYP. Its locus has 285 variants between the LG/J and SM/J backgrounds, but none are predicted to be functional. *Slc22a3* is a transporter that eliminates organic cations from cells, such as monoamine neurotransmitters, cationic drugs, and xenobiotics (Kekuda et al. 1998). It has been reported to be maternally expressed in liver and extraembryonic tissues, but biallelically expressed elsewhere (Babak et al. 2015). *Slc22a3* is also significantly differentially expressed in kidneys between high-fat diet- 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 2,673 sequence biased genes across our 27 tissue-by-context analyses (3 tissues x 9 diet-by-sex contexts) to evaluate whether sequence

306 dependent ASE is also influenced by tissue or environmental context. For each analysis, we quantified
307 the direction and magnitude of each gene's sequence expression bias by calculating an Allelic Genotype
308 Effect (**AGE**) score from the mean allelic bias of each F_1 reciprocal cross. AGE scores range from -1
309 (completely SM/J expressed) to +1 (completely LG/J expressed); a score of 0 indicates biallelic
310 expression (see Methods). In each analysis, a gene could be expressed in one of three ways: significant
311 sequence/allelic genotype bias, biallelic (expressed but with no allele-specific bias), or not expressed.
312 We sorted the 2,673 genes with significant sequence dependent ASE into three expression profiles:
313 tissue-independent, tissue-dependent, and context-dependent (**Figure 4A-4D**).

314 We identified 605 genes (23%) as tissue-independent, meaning they had a consistent sequence bias in
315 every tissue they were expressed. Within a given tissue, they had a sequence bias in most environmental
316 contexts (**Figure 4B**). These tissue-independent genes were strongly biased towards one strain's allele;
317 their mean AGE scores were -0.79 and 0.78 for SM/J and LG/J biased genes, respectively (**Figure 4E**).
318 They also comprised a similar proportion of the total sequence dependent ASE genes in all three tissues
319 (29-35%, **Figure 4A**). These patterns likely reflect the vast genetic variation that has accumulated
320 between the LG/J and SM/J backgrounds over the decades, whereby a *cis*-acting variant impacts one
321 strain's allelic function wherever that gene is expressed, regardless of tissue type.

322 We validated the expression profiles of two tissue-independent genes (*Eef1a1* and *Tubb2a*) by
323 pyrosequencing (**Figure 4B**, **Supplemental Figure S11**). *Eef1a1* (eukaryotic translation elongation
324 factor 1 alpha 1) had a significant LG/J allelic bias in all three tissues, regardless of context. Its locus has
325 30 variants between the LG/J and SM/J backgrounds, including two non-synonymous SNPs in the coding
326 region of the SM/J genome. *Eef1a1* delivers aminoacylated transfer RNAs to the elongating ribosome
327 during protein synthesis and has crucial roles in protein degradation, RNA virus replication, and other
328 cellular processes. It is abundantly and ubiquitously expressed in most tissues (Mateyak and Kinzy 2010;
329 Li et al. 2013). *Tubb2a* (tubulin beta-2A Class IIa) had a significant SM/J allelic bias in all three tissues,
330 regardless of context. Mutations in this gene are rare and often non-viable; however, its locus has one
331 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).

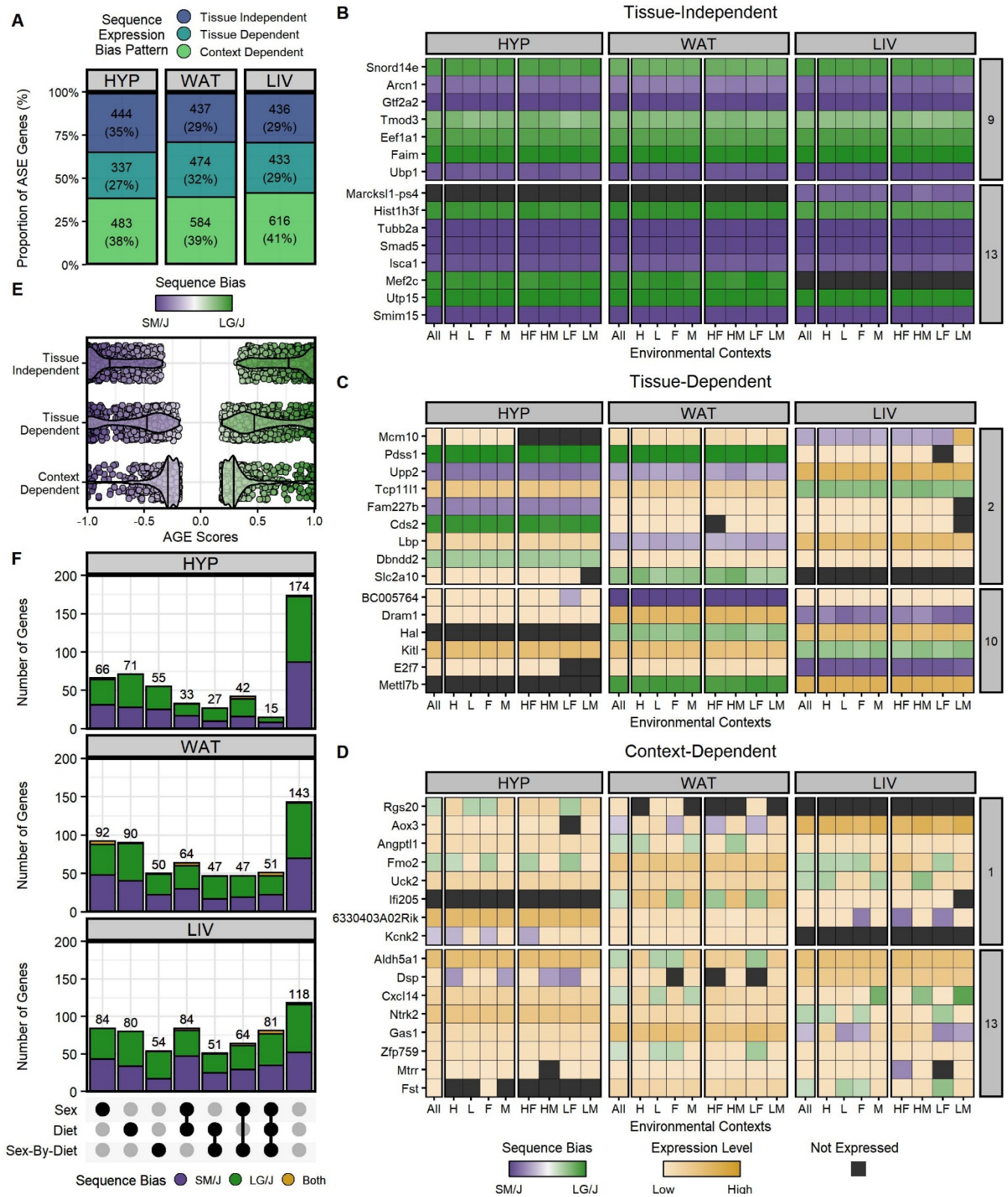

Figure 4: Atlas of sequence dependent ASE patterns across tissues and environmental contexts. **(A)** Number and proportion of ASE genes per tissue with each sequence expression bias profile: tissue-independent (dark blue), tissue-dependent (teal), and context-dependent (green). Heatmaps of ASE profiles across tissues and contexts: **(B)** tissue-independent (sequence bias wherever expressed), **(C)** tissue-dependent (sequence bias in some tissues, no bias in others), and **(D)** context-dependent (sequence bias only in certain diet-by-sex contexts, no bias elsewhere). A subset of the 2,673 genes with significant sequence dependent ASE 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 allelic bias, respectively (AGE scores). Where genes are not significantly biased, shades of yellow indicate their biallelic gene expression levels (log-transformed total counts). Black indicates genes are not expressed in that context/tissue. The y-axis is grouped and sorted by chromosomal position. Each supercolumn denotes a tissue: hypothalamus (HYP), white adipose (WAT), and liver (LIV). Each subcolumn denotes an environmental context: all contexts collapsed (All), high-fat diet (H), low-fat diet (L), females (F), males (M), high-fat fed females (HF), high-fat fed males (HM), low-fat fed females (LF), and low-fat fed males (LM). **(E)** Allelic Genotype Effect (AGE) score distributions for each sequence expression bias profile. Vertical lines indicate that profile's mean AGE score. Dots represent individual significant ASE genes. **(F)** UpSet plots for each tissue summarizing the set intersections of context-dependent genes with significant sex, diet, and/or sex-by-diet effects. Bar height and color indicate the number of genes with each sequence bias direction: SM/J biased (purple), LG/J biased (green), and those that switch expression bias direction depending on the cohort (yellow).

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) (**Figure 4C**). Here, we again distinguish between tissue-dependent ASE (biallelic expression in some tissues) and tissue-specific gene expression (simply not expressed in some tissues). These tissue-dependent genes were moderately biased towards one strain's allele; their mean AGE scores were -0.57 and 0.55 for SM/J and LG/J biased genes, respectively (**Figure 4E**). They also comprised a similar proportion of the total sequence dependent ASE genes in all three tissues (27-32%, **Figure 4A**). These patterns demonstrate that sequence dependent ASE is not solely due to 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 (**Figure 4C, Supplemental Figure S12**). *Upp2* (uridine phosphorylase 2) had a strong SM/J allelic bias across contexts in HYP, a weaker SM/J bias in WAT, and biallelic expression in LIV. Its locus has 1,043 variants between the LG/J and SM/J backgrounds, including five non-synonymous SNPs in the coding region of the LG/J genome. *Upp2* catalyzes the phosphorolysis of uridine into uracil and ribose-1-phosphate, which are used as carbon and energy sources during catabolic metabolism. In mice,

it is predominantly expressed in liver and weakly expressed in brain (Johansson 2003; Roosild et al. 2011). *Mettl7b* (methyltransferase-like 7b) had a LG/J allelic bias across contexts in WAT, biallelic expression in LIV, and was not expressed in HYP. Its locus has 77 variants between the LG/J and SM/J backgrounds, including one non-synonymous SNP in the coding region of the SM/J genome. *Mettl7b* is an akyl thiol methyltransferase that is implicated in several cancers. It is highly expressed in lipid droplets, particularly in liver and adipose tissue (Maldonato et al. 2021; Turró et al. 2006).

Sequence dependent ASE is sensitive to sex and dietary environments

Finally, we classified 1,384 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 (**Figure 4D**). These context-dependent 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 (**Figure 4E**). LIV had the most context-dependent genes (n = 616) and HYP had the fewest (n = 483), but overall they comprised a similar proportion of the total ASE genes in all three tissues (38-41%, **Figure 4A**). These patterns suggest that environmental factors can interact with genetic variation to influence the final allelic composition of a gene's product, resulting in sequence dependent ASE.

To further explore how sex and/or dietary fat can alter sequence/allelic genotype ASE 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 Table S13**). When we intersected these gene lists, we found that each tissue showed a similar pattern of context-dependent sequence ASE biases (**Figure 4F**). Significant sex effects were the most prevalent in each tissue, ranging from 156 genes in HYP (32%), to 254 genes in WAT (43%), and 313 genes in LIV (51%). Diet effects were slightly less common but still widespread: 146 genes in HYP (30%), 252 genes in WAT (43%), and 296 genes in LIV (48%). Finally, diet-by-sex effects were the least frequent yet incredibly rampant, comprising 139 genes in HYP (29%), 195 genes in WAT (33%), and 250 genes in LIV (41%). These categories were not mutually exclusive; most context-dependent genes were significant for more than one effect but in different combinations across the three tissues.

We validated the expression profiles of two context-dependent genes (*Ifi205* and *Gas1*) by pyrosequencing (**Figure 4D**, **Supplemental Figure S14**). *Ifi205* (interferon activated gene 205) had significant sex and diet-by-sex effects in WAT, reflected in strong LG/J allelic biases in the three female-related (HF, LF, F) and "All" contexts. *Ifi205* was biallelically expressed in the remaining contexts in WAT and all contexts in LIV. Its locus has 254 variants between the LG/J and SM/J backgrounds, including ten non-synonymous SNPs in the coding region of the SM/J genome. *Ifi205* binds DNA in response to interferon signaling, a cytokine family that activates the immune system. Significant sex differences have been reported in mouse *Ifi205* expression levels, where females have higher total expression levels than males (Albrecht et al. 2005; Cao et al. 2018). *Gas1* (growth arrest specific 1) had significant diet and sex effects in LIV, reflected in strong SM/J biases in the three low-fat fed diet-related (LF, LM, L), female, and "All" contexts. *Gas1* was biallelically expressed in the remaining high-fat fed diet and male contexts in LIV and all contexts in HYP and WAT. Its locus has 57 variants between the LG/J and SM/J backgrounds, though none are predicted to have functional impacts. *Gas1* encodes a membrane glycoprotein that binds and regulates sonic hedgehog during development. *Gas1* has been found to be differentially expressed between high and low selenium diets in mice ovaries, suggesting its expression may be sensitive to dietary environment (Lee et al. 2001; Qazi et al. 2021).

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

Surprisingly, 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 (**Figure 5**). For 44 of the 45 genes (98%), 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). 19 genes had context-dependent ASE, or a sequence bias only in certain diet and/or sex contexts that switched direction across tissues. One of these genes (*Cidec*) had a sex-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 most contexts in one or two

tissues, but a context-dependent sequence bias in the opposite direction in another tissue. These dynamic direction-switching patterns confirm that sequence dependent ASE is not solely due to genetic variation; otherwise, the same variant causing biased expression in one tissue would also be present in the cells of any other tissue expressing that gene. These patterns hint at epigenetic regulatory elements or post-transcriptional modifications that interact with genetic variation in a tissue-specific manner to influence the final allelic frequency.

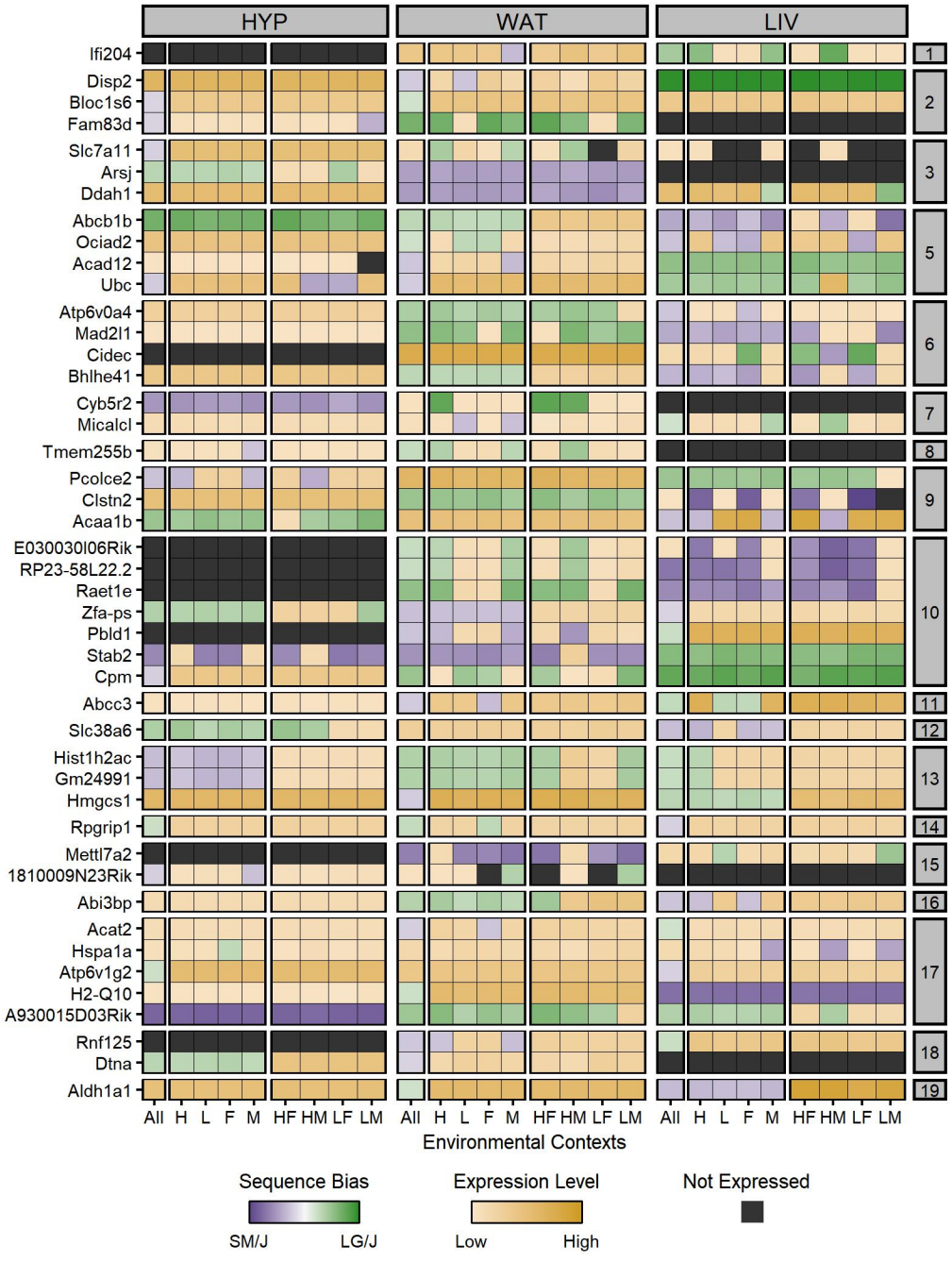

Figure 5: 45 sequence dependent ASE genes switch their expression bias directions across tissues or environmental contexts. Heatmap of ASE profiles for the 45 genes with significant sequence biases in opposite directions across tissues and/or environmental contexts, including those validated with pyrosequencing. 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 allelic bias, respectively (AGE scores). Where genes are not significantly biased, shades of yellow indicate their biallelic gene expression levels (log-transformed total counts). Black indicates genes are not expressed in that context/tissue. The y-axis is grouped and sorted by chromosomal position. Each supercolumn denotes a tissue: hypothalamus (HYP), white adipose (WAT), and liver (LIV). Each subcolumn denotes an environmental context: all contexts collapsed (All), high-fat diet (H), low-fat diet (L), females (F), males (M), high-fat fed females (HF), high-fat fed males (HM), low-fat fed females (LF), and low-fat fed males (LM).

We validated the expression profiles of three direction-switching genes (*Stab2*, *Ociad2*, and *Cidec*) with pyrosequencing (**Figure 5**, **Supplemental Figure S15**). *Stab2* (stabilin 2) had a significant SM/J bias across most contexts in HYP and WAT, but a LG/J bias in all contexts in LIV. Its locus has 1,019 variants between the LG/J and SM/J backgrounds, including 14 non-synonymous SNPs in the coding region of 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, while weakly expressed in the brain and adipose tissue (Falkowski et al. 2003). Together with the non-synonymous variants in the SM/J allele, this may explain *Stab2*'s LG/J bias in the LIV but SM/J bias in the other two tissues. *Ociad2* (ovarian cancer immunoreactive antigen domain-containing protein 2) had significant sex and sex-by-diet effects in WAT and LIV. These are reflected in strong LG/J biases in the L, F, and "All" contexts in WAT, but strong 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 has 118 variants between the LG/J and SM/J backgrounds, though none are predicted to be functional. *Ociad2* activates STAT3, regulates cell migration, and is implicated in several cancers. It is moderately expressed in several non-cancerous tissues, including the brain and liver (Sinha et al. 2018). Sex or diet differences in expression levels have not been explored in the literature, but *Ociad2* is highly expressed in female ovarian tumors (Nagata et al. 2012). Low-fat fed female WAT and LIV tissues have strong sequence biases in *Ociad2* (but in different directions), suggesting a tissue- and sex-specific expression pattern that could have functional consequences.

An interesting exception to the tissue-level switch in bias direction is the gene *Cidec* (cell death-inducing DNA fragmentation effector C), which showed a sex-dependent switch in ASE direction within the same

tissue validated by pyrosequencing (**Figure 5**, **Supplemental Figure S15**). In LIV, *Cidec* had a strong LG/J bias in the HF, LF, and F contexts, yet a SM/J bias in the HM context. *Cidec* was biallelically expressed in the remaining LIV contexts and across all WAT contexts, but was not expressed in HYP. Its locus has 46 variants between the LG/J and SM/J backgrounds, though none are predicted to be functional. *Cidec* promotes lipid droplet formation in adipocytes and regulates low-density lipoprotein maturation in hepatocytes (Xu et al. 2012). Hepatic *Cidec* expression is very sensitive to diet composition and sex hormones; in particular, female mice have lower total expression levels than males when fed a Western diet (high cholesterol and saturated fats) (Herrera-Marcos et al. 2020).

DISCUSSION

Allele-specific expression imbalances due to genetic and epigenetic variation have widespread functional consequences on complex traits, but how environmental signals contribute to this crosstalk 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 comprehensive genome-wide map of both 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 types of ASE are highly dependent on tissue type and environmental context. We identified 2,853 genes with significant parental and/or sequence biases and sorted them into three major expression profiles: tissue-independent, tissue-dependent, and context-dependent. Interestingly, we also found several genes with inconsistent ASE biases that switched direction across tissues and/or contexts (e.g. SM/J biased in one cohort, LG/J biased in another). Although the breadth of these patterns precludes a detailed discussion of each gene, we have 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 due to strain-specific *cis*-acting genetic variants in coding regions that severely affect one allele's function whenever that gene is expressed. Both patterns conform to the conventional mechanism for ASE whereby the genetic and epigenetic processes that influence the 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 (no bias) or biased in the opposite direction in other tissues. Tissue-specific gene expression (simply not expressed in certain tissues) is not considered here, as we are specifically interested in cases where a gene has one allelic ratio in one tissue (e.g. 80% SM/J, 20% LG/J) and a different allelic 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. 25% of sequence dependent ASE genes also have this expression profile, demonstrating that sequence biases are often not solely attributable to genetic variation in coding regions. 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, where the variation in one allele may be favorable in certain tissues' regulatory landscape but the opposite allele is more favorable elsewhere (Andergassen et al. 2017; Leung et al. 2015).

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 more prevalent than expected in both classes of ASE. 27% 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; Macias-Velasco et al. 2021). Surprisingly, over half of sequence dependent ASE genes (52%) have diet- and/or sex-specific biases, indicating these intrinsic and extrinsic environmental

516 factors interact with strain-specific genetic variation to cause a sequence bias. Both patterns suggest a
517 model where environmental signals may interact more efficiently with one allele over the other, leading
518 to shifting and inconsistent allelic proportions in response to environmental cues (Shao et al. 2019).
519 Often, these genes are not significantly biased in our "all contexts" analysis, only in a specific diet and/or
520 sex cohort. This highlights the necessity of studying gene-by-environment interactions, as such effects
521 are obscured when multiple contexts are collapsed together.

522 It is important to note that we can only detect ASE in genes with strain-specific variants in transcribed
523 regions, which is a fraction of the total set of expressed genes. In particular, imprinted genes that are
524 crucial for development/survival may be under tight evolutionary control and not have variants, thus we
525 are unable to assess their ASE status here. Although our exact findings are limited to this F_1 reciprocal
526 cross mouse model, the broad patterns nevertheless demonstrate the complexity of allele-specific gene
527 regulation and its contribution to complex traits. Traditional mapping studies connect genotypes to
528 phenotypes, but are agnostic to tissue and often do not consider environment. eQTL studies connect
529 genotypes to total gene expression levels, but assume biallelic expression. Tissue- and context-
530 dependent ASE of both classes (parent-of-origin and sequence dependent) can bridge these approaches
531 and pinpoint what tissues and/or environments are relevant for a phenotype. A gene could be expressed
532 at the same level in two cohorts, but its allelic composition could differ due to ASE. If there is a functional
533 difference between the two alleles, then an expression imbalance in the right tissue and/or environment
534 could lead to phenotypic consequences. Incorporating these dynamic ASE patterns into our frameworks
535 will help us decipher the genotype to phenotype map.

536 **METHODS**

537 *F_1 Reciprocal Cross Mouse Model*

538 We obtained LG/J and SM/J founders from The Jackson Laboratory (Bar Harbor, ME) and generated F_1
539 reciprocal crosses by mating LG/J mothers with SM/J fathers (**LxS**) and vice versa (**SxL**). F_1 offspring
540 were weaned into sex-specific cages at three weeks of age and randomly placed on either a high-fat diet
541 (42% kcal from fat; Teklad TD88137) or an isocaloric low-fat diet (15% kcal from fat; Research Diets

D12284). They were fed *ad libitum*. F_1 mice were euthanized at 20 weeks of age 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.

RNA-Sequencing and Allele-Specific Mapping

We sequenced 32 samples per tissue, representing 4 mice from each sex (male or female), diet (high or low fat), and F_1 reciprocal cross (LxS or SxL) 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 (ThermoFisher) and RNA integrity scores ≥8.0 (Agilent). We constructed RNA-Seq libraries with the RiboZero rRNA Removal Kit (Illumina), checked their quality with the BioAnalyzer DNA 1000 assay (Agilent), and sequenced them at 100bp, paired-end reads on an Illumina HiSeq 400. After sequencing, reads were de-multiplexed and assigned to individual samples.

Quantifying ASE in a F_1 reciprocal cross is vulnerable to reference genome alignment bias. If one parental strain is more closely related to the reference genome, then their reads tend to have a higher mapping quality than reads from the other parental strain (Wang and Clark 2014; Degner et al. 2009). We mitigated this concern by aligning RNA-Seq reads to a custom LG/J x SM/J “merged genome”. Previously, LG/J and SM/J reference genomes were created by combining strain-specific SNVs and indels with the GRC38.72-mm10 reference template (Nikolskiy et al. 2015). Customized gene annotations were also created by adjusting Ensembl definitions (Mus_musculus.GRCm38.72) for indexing differences due to strain-specific indels. Here, those LG/J and SM/J genomes were combined into one pseudogenome so we could align reads to both parental 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 the first round and used to inform a second round of mapping, thus detecting more reads that span novel junctions. By not allowing multi-mapping, we only retained

reads uniquely covering strain-specific variants so we could assign reads to their allelic origin; reads covering identical regions between the parental strains were discarded. STAR alignment summaries are provided in **Supplemental Table S1** and **Supplemental Figure S16**.

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. Gene-level allele-specific read counts were then upper quartile normalized (**Supplemental Figure S17**) and filtered to remove lowly-expressed genes (total normalized read counts <20). We retained a total of 9171 genes with detectable allele-specific expression in HYP, 9761 genes in WAT, and 8082 genes in LIV.

Library Complexity

Insufficient library complexity can also hamper detecting ASE in a F_1 reciprocal cross. In lowly or moderately expressed genes, if a read fragment from only one of the two alleles is randomly subjected to a duplicating event (i.e. PCR jackpotting) during RNA-Seq library construction, then that gene may spuriously appear as monoallelically expressed, even though it is a false positive (Wang and Clark 2014). To check for this, we measured each library's complexity by fitting 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 (α, β) of the beta-binomial distribution and calculated the overdispersion parameter (ρ) as $\rho = 1 / (1 + \alpha + \beta)$. Lower values of ρ (< 0.075) indicate a library is sufficiently complex. One WAT library (CCGGACC) and one LIV library (TGATTAC) were deemed to have poor complexity and were removed from further analyses (**Supplemental Figure S18**).

Determining Biased Allele-Specific Expression

To explore how environmental context (diet and/or sex) impacts ASE patterns, we analyzed nine separate cohorts per tissue: high fat diet (**H**), low fat diet (**L**), females (**F**), males (**M**), high-fat fed females (**HF**), high-fat fed males (**HM**), low-fat fed females (**LF**), low-fat fed males (**LM**), and all contexts together (**All**) (**Figure 1A**) (3 tissues x 9 contexts = 27 tissue-by-context cohorts). For each tissue-by-context analysis, we required a gene to be expressed in ≥75% of the total sample size per F_1 reciprocal cross: 3 mice per

cross for diet-by-sex-specific contexts (N = 4), 6 mice per cross for diet- and sex-specific contexts (N = 8), and 12 mice per cross for the "All" context (N = 16).

We adapted a previously-published model (Takada et al. 2017) for jointly estimating 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 Figure S2**). Next, we implemented a paired-sample design to handle missing data in edgeR, which converts NAs to zero counts. There is a fundamental difference between a gene not being expressed in a particular sample (both allele counts are zero) and an extreme allelic expression bias (only one allele's counts are zero). Thus, we added a series of n-1 dummy variables (indicating library barcodes) to the GLM so that both allele-specific counts for each sample are treated as a pair. If a gene is not expressed in a library, then the coefficient corresponding to that sample will approach negative infinity in the fitted GLM. The missing sample is effectively removed from consideration and does not affect the PO and AG coefficient estimates for the other samples. 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 PO and AG effects on allele-specific gene expression levels.

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 read counts with the LG/J haplotype (L_{bias}) or the SM/J haplotype (S_{bias}). Using the mean allelic biases of each F_1 reciprocal cross, we constructed Parent-of-Origin Effect (**POE**) and Allelic Genotype Effect (**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 range from -1 (completely maternally expressed) to +1 (completely paternally expressed). Similarly, AGE scores range from -1 (completely SM/J expressed) to +1 (completely LG/J expressed).

Scores of 0 for both indicate biallelic expression. Full GLM summary statistics and POE/AGE scores for each tissue-by-context analysis are provided in **Supplemental Tables S2 - S4**.

Significance Thresholds

A crucial consideration for ASE analyses is how to best adjust for multiple tests. Allelic expression biases are often correlated for genes within and between imprinted domains (Edwards and Ferguson-Smith 2007) and for genes controlled by the same regulatory element with a functional variant (Cavalli et al. 2016). This ensures any tests performed on those genes are also correlated, breaking independence assumptions. We addressed this challenge by using a permutation approach (Westfall and Young 1993) to estimate our statistical and biological significance thresholds.

For each tissue-by-context cohort, we generated a stable null distribution of likelihood ratios (LR) for both AG and PO terms. We randomly shuffled the allele-specific read counts for all genes and reran our analyses over several iterations. After each iteration, we calculated the change in mean LR quantiles for both terms in the full permuted dataset. Quantiles corresponded to one percent increments of the LR distribution. We added new iterations until the null model met our “stability” criteria: the mean LR quantile differences for both AG and PO terms fluctuated by less than |0.001| for 10 consecutive iterations (mean = 51 iterations) (**Supplemental Figure S19**). Comparisons between the real and permuted datasets are provided for p-values, POE/AGE scores, and likelihood ratios (**Supplemental Figures S20 – S22**).

Next, we built an empirical cumulative distribution function (ECDF) from each term’s permuted p-values for each tissue-by-context analysis (**Supplemental Figures S23**). We fit each term’s raw p-values to its respective ECDF to compute the adjusted p-values, i.e. the proportion of tests from the permuted null model that are more extreme (smaller p-values) than the test from the real model. We also calculated the 5^{th} and 95^{th} quantiles of the permuted POE/AGE scores for each term in each tissue-by-context analysis. We set our critical threshold as the more extreme value. We deemed an adjusted p-value ≤ 0.05 as statistically significant and real POE/AGE scores beyond the critical threshold as biologically significant.

Genes with significant PO term p-values and POE scores were considered to have parent-of-origin dependent ASE. Similarly, genes with significant AG term p-values and AGE scores were considered to have sequence dependent ASE (**Supplemental Figures S24, Supplemental Table S5**).

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

To examine how ASE is influenced by tissue type and/or environmental context, we characterized the expression profiles of the significant ASE genes across our 27 tissue-by-context analyses (3 tissues x 9 diet-by-sex contexts). In each analysis, a gene could be expressed in one of three ways: significantly biased, expressed with no allele-specific bias (biallelic), or simply not expressed. We sorted the significant ASE genes of both classes into three expression profiles based on the following criteria.

Tissue-independent ASE genes have a significant bias in every tissue where they are expressed. Within a tissue, the gene must have a significant bias in ≥5 of the 9 diet-by-sex contexts (i.e. most, but not all, environmental contexts). This flexibility allows for genes that may have true ASE but were excluded due to our stringent minimum sample size requirements; these genes could appear as biallelically expressed since they still pass our minimum read depth requirements. Tissue-dependent ASE genes have a significant bias in some tissues and no bias in others. The gene must have a significant bias in >5 of the 9 contexts in one or two tissues, but biallelic expression in >5 of the 9 contexts in the other tissue(s). Both profiles allowed for genes to not be expressed in some tissues, as we wanted to deliberately distinguish between tissue-specific gene expression (not expressed in certain tissues) and tissue-dependent ASE (biased in certain tissues). Finally, context-dependent ASE genes have a significant bias only in certain environmental contexts and no bias elsewhere. The gene must have a significant bias in ≤5 of the 9 contexts within a tissue, but biallelic expression in the other contexts and/or tissues. All significant ASE genes of both classes fit into one of these expression profiles; no genes were unclassified.

Evaluating environmental context-dependency

Once we identified context-dependent genes, we evaluated how sex and/or dietary environment shape their ASE patterns. 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 Parent-of-Origin Effect (POE) or Allelic Genotype Effect (AGE) scores per sample per gene by modifying the above equations as follows:

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

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

Individualized POE scores range from -1 (maternally expressed) to +1 (paternally expressed). Similarly, individualized AGE scores range from -1 (SM/J expressed) to +1 (LG/J expressed). Scores of 0 for both scores indicate biallelic expression or that the gene was not expressed in that sample.

Next, we used ANOVA models to test whether a gene's allelic biases (individualized POE/AGE scores) were influenced by sex, diet, and/or their interaction. We considered FDR-corrected p-values ≤ 0.1 to be significant (**Supplemental Table S6**). 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 ≤ 0.05) (**Supplemental Table S7**).

Imprinted gene list

We defined a significant ASE gene as "canonically imprinted" if it appeared in the GeneImprint mouse database (https://www.geneimprint.com, as of May 2020) and/or a PubMed search of the gene name (or synonyms) and imprinting-related terms.

Pyrosequencing

We randomly selected 13 genes to validate based on their allele-specific expression profiles (i.e. tissue-independent, tissue-dependent, context-dependent, switches bias direction). We prioritized genes that were highly expressed and statistically significant (for context-dependent examples), but excluded those with high expression level variance between biological replicates. For each gene, we identified the strain-specific SNPs within exons and designed primer sets to flank these variants using Geneious Prime 2020.0.4 (https://www.geneious.com) (Kearse et al. 2012). Wherever possible, target regions were 150-

200bp 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 S8**.

We extracted total RNA from the HYP, WAT, and LIV of one mouse per F_1 reciprocal cross (LxS and SxL) in each diet-by-sex cohort using the RNeasy Lipid Tissue Mini Kit (QIAGEN). Next, 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 non-biotinylated (forward) primer. The biotinylated single-stranded PCR products were then 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 pyrosequencing reaction emits a light signal as it builds the DNA fragment, which appears on the pyrogram output as a peak whose height is proportional to how many nucleotides were incorporated at that base. From these peaks, 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.

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.

ACKNOWLEDGEMENTS

This work was supported by the Washington University Department of Genetics, the Diabetes Research Center at Washington University (CFS: P30DK020579), the National Science Foundation Graduate Research Fellowship (CLS: DGE1745038, DGE2139839), as well as the National Institute of Diabetes and Digestive and Kidney Diseases (HAL: K01DK095003), the National Institute of Environmental Health Sciences (HAL: U24ES026699), and the National Human Genome Research Institute (JFMV: T32GM007067) of the National Institutes of Health. The authors declare no conflicts of interest.

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.

Andergassen D, Dotter CP, Wenzel D, Sigl V, Bammer PC, Muckenhuber M, Mayer D, Kulinski TM, Theussl H-C, Penninger JM, et al. 2017. Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression ed. T.R. Gingeras. *eLife* **6**: e25125.

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.

Barlow DP, Bartolomei MS. 2014. Genomic Imprinting in Mammals. *Cold Spring Harb Perspect Biol* **6**: a018382.

Bonasio R, Tu S, Reinberg D. 2010. Molecular Signals of Epigenetic States. *Science* **330**: 612–616.

Buil A, Brown AA, Lappalainen T, Viñuela A, Davies MN, Zheng H-F, 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.

Cao Y, Wang L, Wang C-Y, 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.

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

Castel SE, Aguet F, Mohammadi P, Consortium Gte, Ardlie KG, Lappalainen T. 2019. A vast resource of allelic expression data spanning human tissues. *bioRxiv* 792911.

Cavalli M, Pan G, Nord H, Wallerman O, Wallén Arzt E, Berggren O, Elvers I, Eloranta M-L, 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.

Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. 2009. Mapping complex disease traits with global gene expression. *Nat Rev Genet* **10**: 184–194.

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.

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.

Edwards CA, Ferguson-Smith AC. 2007. Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol* **19**: 281–289.

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.

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

756 Gai Z, Visentin M, Hiller C, Krajnc E, Li T, Zhen J, Kullak-Ublick GA. 2016. Organic Cation Transporter
757 2 Overexpression May Confer an Increased Risk of Gentamicin-Induced Nephrotoxicity.
758 *Antimicrob Agents Chemother* **60**: 5573–5580.

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

762 Hammond JW, Cai D, Verhey KJ. 2008. Tubulin modifications and their cellular functions. *Curr Opin*
763 *Cell Biol* **20**: 71–76.

764 He H, Kim J. 2014. Regulation and Function of the Peg3 Imprinted Domain. *Genomics Inform* **12**: 105–
765 113.

766 He W, Rose DW, Olefsky JM, Gustafson TA. 1998. Grb10 Interacts Differentially with the Insulin
767 Receptor, Insulin-like Growth Factor I Receptor, and Epidermal Growth Factor Receptor via the
768 Grb10 Src Homology 2 (SH2) Domain and a Second Novel Domain Located between the
769 Pleckstrin Homology and SH2 Domains *. *J Biol Chem* **273**: 6860–6867.

770 Heap GA, Yang JHM, Downes K, Healy BC, Hunt KA, Bockett N, Franke L, Dubois PC, Mein CA,
771 Dobson RJ, et al. 2010. Genome-wide analysis of allelic expression imbalance in human
772 primary cells by high-throughput transcriptome resequencing. *Hum Mol Genet* **19**: 122–134.

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

777 Hochner H, Allard C, Granot-Hershkovitz E, Chen J, Sitlani CM, Sazdovska S, Lumley T, McKnight B,
778 Rice K, Enquobahrie DA, et al. 2015. Parent-of-Origin Effects of the APOB Gene on Adiposity in
779 Young Adults. *PLOS Genet* **11**: e1005573.

780 Johansson M. 2003. Identification of a novel human uridine phosphorylase. *Biochem Biophys Res*
781 *Commun* **307**: 41–46.

782 Kappil M, Lambertini L, Chen J. 2015. Environmental Influences on Genomic Imprinting. *Curr Environ*
783 *Health Rep* **2**: 155–162.

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

787 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz
788 S, Duran C, et al. 2012. Geneious Basic: An integrated and extendable desktop software
789 platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.

790 Kekuda R, Prasad PD, Wu X, Wang H, Fei Y-J, Leibach FH, Ganapathy V. 1998. Cloning and
791 Functional Characterization of a Potential-sensitive, Polyspecific Organic Cation Transporter
792 (OCT3) Most Abundantly Expressed in Placenta*. *J Biol Chem* **273**: 15971–15979.

Knowles DA, Davis JR, Edgington H, Raj A, Favé M-J, 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.

Lawson HA, Cady JE, 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.

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

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.

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.

Lee CS, Buttitta L, Fan C-M. 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.

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

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.

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.

Macias-Velasco JF, Pierre CLS, Wayhart JP, Yin L, Spears L, Miranda MA, Carson C, Funai K, Cheverud JM, Semenkovich CF, et al. 2021. *Parent-of-origin effects propagate through networks to shape metabolic traits*. https://www.biorxiv.org/content/10.1101/2021.08.10.455860v1 (Accessed August 23, 2021).

Maldonato BJ, Russell DA, Totah RA. 2021. Human METTL7B is an alkyl thiol methyltransferase that metabolizes hydrogen sulfide and captopril. *Sci Rep* **11**: 4857.

Mateyak MK, Kinzy TG. 2010. eEF1A: Thinking Outside the Ribosome *. *J Biol Chem* **285**: 21209–21213.

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.

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.

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.

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.

Nagata C, Kobayashi H, Sakata A, 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.

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.

Olofsson S-O, 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.

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

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

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.

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

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

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

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.

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

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.

Shao L, Xing F, Xu C, Zhang Q, Che J, Wang X, Song J, Li X, Xiao J, Chen L-L, 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.

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.

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.

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 Genes Genomes Genet* **7**: 2227–2234.

Thiaville MM, Huang JM, Kim H, Ekram MB, Roh T-Y, Kim J. 2013. DNA-binding motif and target genes of the imprinted transcription factor PEG3. *Gene* **512**: 314–320.

Trefts E, Gannon M, Wasserman DH. 2017. The liver. *Curr Biol* **27**: R1147–R1151.

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.

Wang X, Clark AG. 2014. Using next-generation RNA sequencing to identify imprinted genes. *Heredity* **113**: 156–166.

Westfall PH, Young SS. 1993. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. John Wiley & Sons.

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.

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.

Xu L, Zhou L, Li P. 2012. CIDE Proteins and Lipid Metabolism. *Arterioscler Thromb Vasc Biol* **32**: 1094–1098.

Yee TW. 2010. The VGAM Package for Categorical Data Analysis. *J Stat Softw* **32**: 1–34.

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.