

# Parent-of-origin effects propagate through networks to shape metabolic traits

Juan F Macias-Velasco<sup>1</sup>, Celine L St Pierre<sup>1</sup>, Jessica P Wayhart<sup>1</sup>, Li Yin<sup>2</sup>, Larry Spears<sup>2</sup>, Mario A Miranda<sup>1</sup>, Caryn Carson<sup>1</sup>, Katsuhiko Funai<sup>3</sup>, James M Cheverud<sup>4</sup>, Clay F Semenkovich<sup>2</sup>, Heather A Lawson<sup>1\*</sup>

<sup>1</sup>Department of Genetics, Washington University School of Medicine, Saint Louis, United States; <sup>2</sup>Department of Medicine, Washington University School of Medicine, Saint Louis, United States; <sup>3</sup>Diabetes and Metabolism Research Center, University of Utah, Salt Lake City, United States; <sup>4</sup>Department of Biology, Loyola University, Chicago, United States

**Abstract** Parent-of-origin effects are unexpectedly common in complex traits, including metabolic and neurological traits. Parent-of-origin effects can be modified by the environment, but the architecture of these gene-by-environmental effects on phenotypes remains to be unraveled. Previously, quantitative trait loci (QTL) showing context-specific parent-of-origin effects on metabolic traits were mapped in the F<sub>16</sub> generation of an advanced intercross between LG/J and SM/J inbred mice. However, these QTL were not enriched for known imprinted genes, suggesting another mechanism is needed to explain these parent-of-origin effects phenomena. We propose that non-imprinted genes can generate complex parent-of-origin effects on metabolic traits through interactions with imprinted genes. Here, we employ data from mouse populations at different levels of intercrossing (F<sub>0</sub>, F<sub>1</sub>, F<sub>2</sub>, F<sub>16</sub>) of the LG/J and SM/J inbred mouse lines to test this hypothesis. Using multiple populations and incorporating genetic, genomic, and physiological data, we leverage orthogonal evidence to identify networks of genes through which parent-of-origin effects propagate. We identify a network comprised of three imprinted and six non-imprinted genes that show parent-of-origin effects. This epistatic network forms a nutritional responsive pathway and the genes comprising it jointly serve cellular functions associated with growth. We focus on two genes, *Nnat* and *F2r*, whose interaction associates with serum glucose levels across generations in high-fat-fed females. Single-cell RNAseq reveals that *Nnat* expression increases and *F2r* expression decreases in pre-adipocytes along an adipogenic trajectory, a result that is consistent with our observations in bulk white adipose tissue.

\*For correspondence:  
lawson@wustl.edu

**Competing interest:** The authors declare that no competing interests exist.

**Funding:** See page 18

**Preprinted:** 11 August 2021

**Received:** 12 August 2021

**Accepted:** 25 March 2022

**Published:** 31 March 2022

**Reviewing Editor:** David B Allison, Indiana University, United States

© Copyright Macias-Velasco et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Editor's evaluation

We all learned simple Mendelian Punnett Squares in Junior High or earlier when studying simple Mendelian traits. But we also all know that the world is so much richer and more complex than this. The current article explores some of that complexity, opening rich insights into intergenerational effects, offering the opportunity for mathematical thinking and further hypothesis testing, and opening up exciting new hypotheses to test. As Professor Stephen Stearns wrote, "Many of us do not do science only, or even primarily, to achieve practical results. We do it because we are fascinated with neat ideas. Evolutionary medicine is full of them, including parent-of-origin pattern." Let us enjoy the wonder.

## Introduction

Parent-of-origin effects, where the phenotypic effect of an allele depends on whether the allele is inherited maternally or paternally, are epigenetic phenomena associated with a wide range of complex traits and diseases (Lawson et al., 2013). Thus, the functional impact of a specific genetic variant can depend on its parental origin. The best characterized parent-of-origin effect is genomic imprinting, an epigenetic process in which either the maternally or paternally inherited allele is silenced, typically through DNA methylation. In humans, there are 107 verified imprinted genes and in mice there are 124, of which ~ 70% overlap (Jirtle, 2012). Despite the rarity of imprinted genes, parent-of-origin effects on complex traits and diseases are relatively common, suggesting that canonical imprinting mechanisms are not sufficient to account for these phenomena (Mozaffari et al., 2019; Zeng et al., 2019). With so few imprinted genes, what mechanisms underlie these parent-of-origin effects? We hypothesize that a small number of imprinted genes can generate a large number of parent-of-origin effects through interactions with non-imprinted genes.

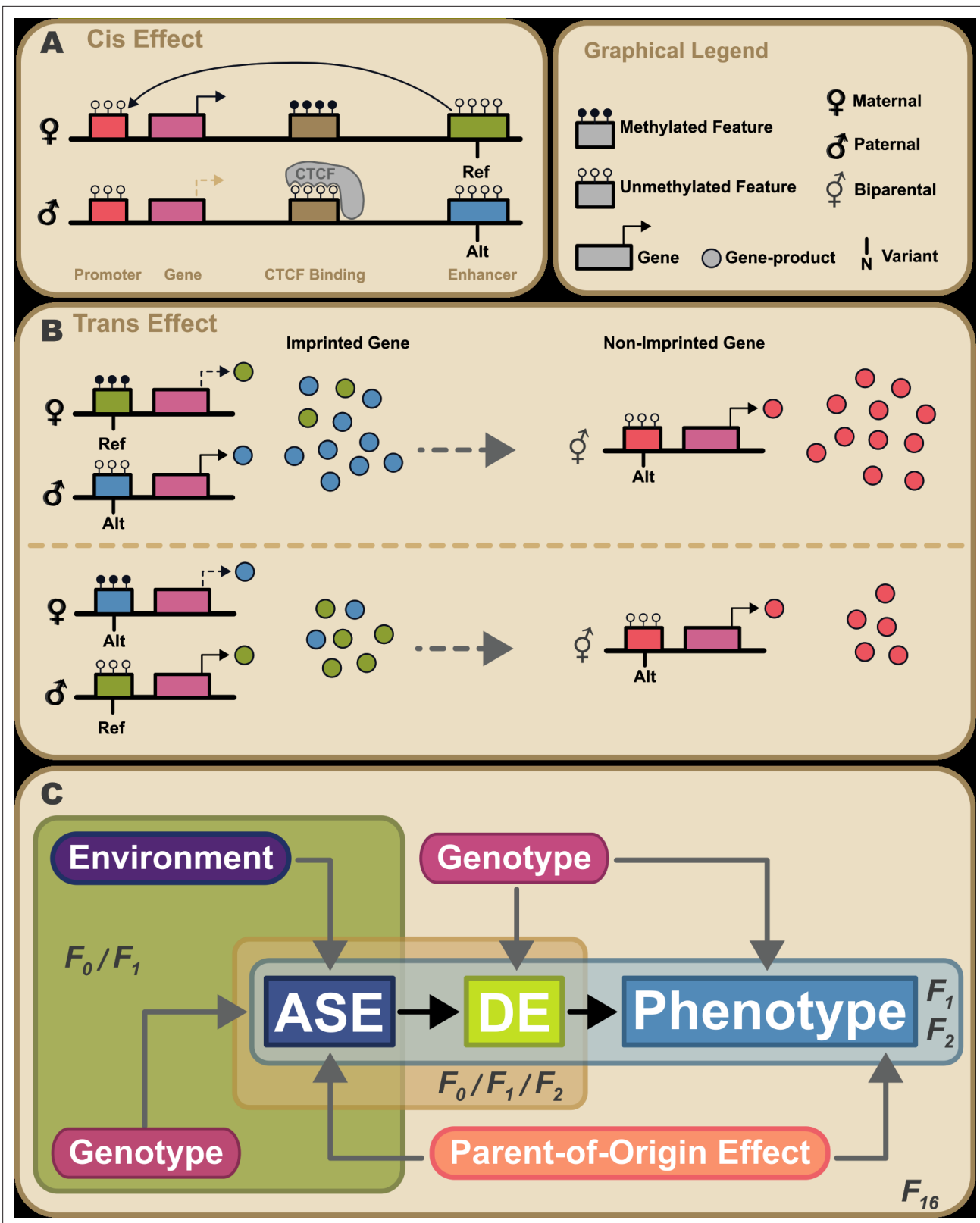
In this study, we use four populations at different levels of intercrossing of the LG/J and SM/J inbred mouse lines to test the hypothesis that non-imprinted genes can contribute to parent-of-origin effects on metabolic phenotypes through epistatic interactions with imprinted genes. Multiple populations ( $F_0$ ,  $F_1$ ,  $F_2$ ,  $F_{16}$ ) allow us to refine our search space and provide orthogonal evidence supporting putative networks of interacting genes. Metabolic traits were previously mapped in a  $F_{16}$  generation of an advanced intercross between LG/J and SM/J (Cheverud et al., 2011; Lawson et al., 2010; Lawson et al., 2011a; Lawson et al., 2011b). We generated visceral white adipose tissue gene expression profiles from 20 week-old  $F_1$  animals in order to match the age of the  $F_{16}$  LG/J x SM/J advanced intercross population.  $F_1$  reciprocal cross (LxS and SxL) mice were subjected to the same high and low-fat diets and phenotyping protocols as the previously-studied  $F_{16}$  mice to keep environmental contexts consistent. We identified genes showing parent-of-origin-dependent allele-specific expression (ASE), characterized interactions among these genes and biallelic genes that are differentially expressed by reciprocal cross (DE), and correlated interacting ASE and DE gene pairs with metabolic phenotypes in the  $F_1$  population. Pairs that significantly associated with phenotypic variation were tested for epistasis on correlated traits in the  $F_{16}$  population.

We identified an epistatic network that forms a nutritional environment responsive pathway mediated through calcium signaling. This network contributes to metabolic variation by balancing proliferation, differentiation, and apoptosis in adipocytes. The genes comprising this network jointly serve functions associated with growth in multiple tissues, which is consistent with the evolutionary hypothesis that sexual conflict underlies some parent-of-origin effects (Mochizuki et al., 1996). We focus on two key interacting genes: *Nnat* (neuronatin), a canonically imprinted gene, and *F2r* (coagulation factor II receptor), a biallelic gene showing significant DE by cross in  $F_1$  high-fat-fed female animals. Co-expression of these two genes associates with variation in basal glucose levels, and this association persists across generations. Further, single-cell RNAseq reveals that *Nnat* expression increases and *F2r* expression decreases in pre-adipocytes along an adipogenic trajectory, a pattern consistent both with their expression in bulk white adipose tissue and with their respective roles in adipogenesis. Our results demonstrate that incorporating orthogonal lines of evidence including genotype, allele specific expression, total gene expression, single-cell expression, and phenotype from different populations varying in their degree of intercrossing is a powerful way to identify putative mechanisms and test hypotheses underlying parent-of-origin effects on phenotype.

## Results

### Non-imprinted genes interact with imprinted genes and effect metabolic phenotypes

We test the hypothesis that non-imprinted genes can mediate complex parent-of-origin effects on phenotypes through genetic interactions with imprinted genes using a  $F_1$  reciprocal cross model of the LG/J and SM/J inbred mice (LxS and SxL). In this model the effects of parental origin on an allele can be tested directly and isolated from sequence dependent *cis*-regulatory differences. We validated our findings in LG/J and SM/J parentals ( $F_0$ ) as well as in  $F_2$  and  $F_{16}$  intercrosses of LGxSM (Figure 1). The parental  $F_0$  animals serve to anchor variation in allele-specific expression that is a function of allelic identity (L or S). Incorporating the  $F_2$  and  $F_{16}$  populations into our validations ensures



**Figure 1.** Proposed model for propagation of parent-of-origin effects through gene-gene interactions. Parent-of-origin effects should be partitioned into *cis* mechanisms and *trans* mechanisms (A) An example of a *cis* parent-of-origin effect is a system with three regulatory elements: promoter, insulator, and enhancer. Activation of transcription requires the enhancer to act upon the promoter. Enhancer activity is blocked by the insulator when it has been bound by CTCF. CTCF cannot be bound when methylated. In this system, the insulator is selectively methylated when inherited maternally, Figure 1 continued on next page

## Figure 1 continued

so methylation of the maternally inherited insulator blocks CTCF binding, allowing the enhancer to activate transcription. Because the paternally inherited insulator is not methylated, it is bound by CTCF which blocks enhancer activity, silencing transcription. This canonical genomic imprinting mechanism interacts with genetic variation in the three regulatory features. For example, if one allele produces stronger enhancer activity (Alt) than the other, individuals inheriting the Alt allele maternally would have elevated expression compared to those that inherit the same allele paternally. These *cis* genetic effects do not occur in isolation. Due to the highly interconnected nature of biological systems, there are downstream effects. We refer to these as *trans* parent-of-origin effects. (B) An example of a *trans* parent-of-origin effect is a system with two genes each having its own promoter. The first gene is canonically imprinted, and the activity of the gene promoter is blocked by DNA methylation. The imprinted gene's promoter is methylated when inherited maternally. Consequently, the paternally inherited allele is almost exclusively expressed. As before, when genetic variation in a regulatory feature interacts with these epigenetic mechanisms, we see parent-of-origin effects on expression of the imprinted gene. In this example, the imprinted gene regulates expression of a non-imprinted gene. Despite the non-imprinted gene being agnostic to parental origin, its expression nonetheless depends on the parental origin of alleles at the imprinted locus. (C) Summary of our experimental design. Expression patterns of genes showing allele-specific expression (ASE) such as imprinted genes are shaped by parental genotypes and environment (e.g. nutrition). Downstream gene expression is a function of their genotype and the expression of upstream ASE genes. Altered parent-of-origin dependent total gene expression of ASE genes leads to differential expression of downstream genes varying only in allelic parent-of-origin (DE). Phenotype is most directly affected by expression of DE genes. Variation in DE gene expression leads to corresponding variation in phenotype. Mouse populations used to probe parts of this model are labeled  $F_0$  (inbred lines),  $F_1$  (reciprocal cross of inbred lines),  $F_2$  (intercross of  $F_1$  mice), and  $F_{16}$  (advanced intercross of inbred lines).

that the interactions we observe are not solely a function of linkage in the  $F_1$  animals. We generated mRNA expression profiles in white adipose tissue from 20-week-old  $F_1$  reciprocal cross animals. These animals were subjected to the same high and low-fat diets and phenotyping protocols as the previously studied  $F_{16}$  animals (Cheverud et al., 2011; Lawson et al., 2011a, Lawson et al., 2010; Carson et al., 2020; Miranda et al., 2020). We identified two classes of genes: (1) imprinted genes and (2) non-imprinted genes with parent-of-origin effects on total expression.

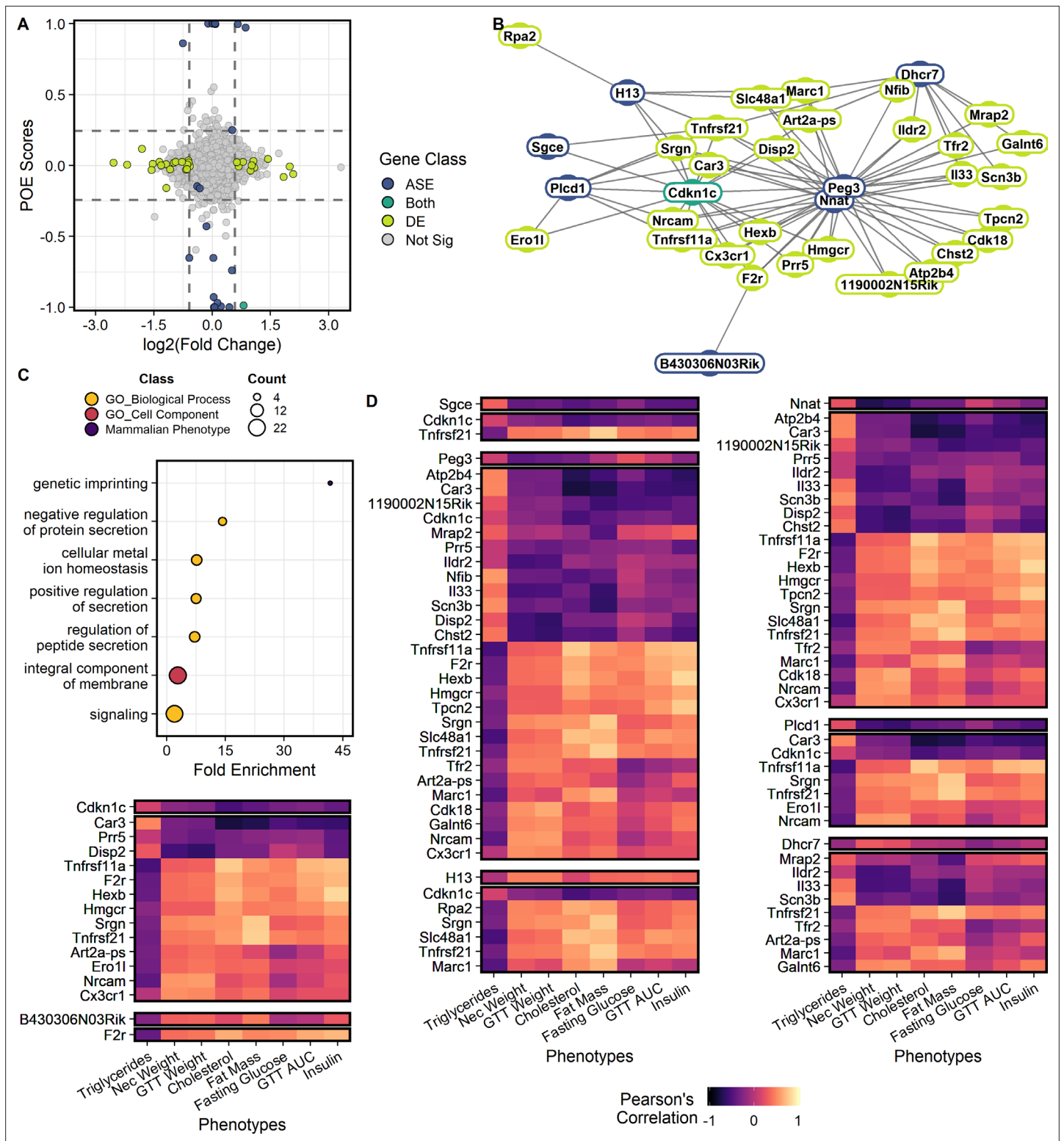
To test our model, we identified genes showing parent-of-origin dependent allele specific expression (ASE). We identified 23 genes showing significant ASE (Figure 2A; Supplementary file 1). Of these 23 genes, 17 are canonically imprinted genes, two are not reported as imprinted genes but are located in known imprinted domains, and four are novel. Next, we identified genes showing differential total expression between individuals varying only in allelic parent-of-origin (DE between reciprocal crosses, SxL vs LxS). We identified 33 genes that are significantly DE in at least one sex or dietary context (Figure 2A; Supplementary file 2). A larger set of genes show signatures of parent-of-origin effects at the total gene expression level, but do not meet the statistical rigor demanded by the massive multiple tests burden incurred by a genome-wide scan accounting for sex, diet, and parent-of-origin (see Materials and methods).

To identify interactions between gene sets, we constructed a network comprised of genes that could initiate a parent-of-origin effect on phenotype (ASE) and genes that may mediate the effect onto phenotype (DE). Interacting gene pairs were predicted by modeling the expression of biallelic genes that are significantly DE by reciprocal cross as a function of the expression of genes showing significant parent-of-origin-dependent ASE, their allelic bias ( $L_{bias}$ ), diet, sex, and the diet-by-sex interaction. Genes showing parent-of-origin effects form a highly interconnected network comprised of 52 genes forming 217 gene pairs (Figure 2B)(Supplementary file 3). Most of these interactions are *trans*-chromosomal. We identified two genes that could serve as initiation points of propagating parent-of-origin effects through this network. These two genes, *Nnat* (neuronatin) and *Cdkn1c* (cyclin dependent kinase inhibitor 1 C), are both canonically imprinted and differentially expressed by reciprocal cross (Supplementary file 1).

Functional over representation analysis was performed and seven terms were significantly over-represented at an  $FDR \leq 0.05$  (Figure 2C; Zhang et al., 2005). Enriched terms suggest this network plays a role in signaling and genetic imprinting (Supplementary file 4). In order to identify which phenotypes might be affected by genes in this network, gene expression was correlated with metabolic phenotypes collected for the  $F_1$  animals (Figure 2D). Seventy-four ASE/DE/phenotype sets were identified as candidates for subsequent testing (Supplementary file 5).

## Epistasis in an $F_{16}$ advanced intercross identifies a diet-responsive network affecting adipogenesis

To validate the interactions we identified in  $F_1$  animals, we tested for imprinting-by-imprinting epistasis in an  $F_{16}$  population. Imprinting-by-imprinting epistasis occurs when the parent-of-origin effect at



**Figure 2.** Genes showing parent-of-origin effects at the allele specific and/or total expression levels covary with each other and with metabolic traits. (A) Mean parent-of-origin effect score across contexts. Effect size of ASE is calculated as the mean allelic bias ( $L / L + S$ ) of SxL animals minus LxS animals. Effect size of DE is measured by  $\log_2$ (Fold Change) between LxS and SxL crosses. The single context with largest magnitude fold change is plotted for each gene. Dashed lines represent minimum acceptable effect size cut-offs within a context. Genes showing significant ASE and sufficiently large parent-of-origin effect score are shown in blue. Genes showing significant DE and sufficiently large fold change in some sex or dietary context are shown in lime. Genes showing both ASE and DE are shown in teal. Genes not meeting cut-offs are shown in gray. The two genes showing significant

Figure 2 continued on next page

Figure 2 continued

ASE but falling short of parent-of-origin effect score requirements are a case of context dependent bipolar parent-of-origin effect scores (i.e. paternally expressed in one context and maternally expressed in its opposite). **(B)** Parent-of-origin effect network constructed from ASE and DE gene pairs. **(C)** Significantly overrepresented ontologies after multiple tests correction in the parent-of-origin effect network. Terms are color coded by ontology domain. GO biological process (yellow), GO cellular component (orange), and Mammalian phenotype (purple). Circle size denotes the number of genes with each term. **(D)** Correlation of parent-of-origin effect network genes with metabolic traits. Only genes and phenotypes with at least one significant correlation after multiple test corrections are shown. The heatmap is broken up into subnetworks with the ASE gene as the first separated row followed by associated DE genes in subsequent rows. Columns correspond to metabolic traits. Coloration of each cell denotes the Pearson's correlation coefficient value.

The online version of this article includes the following figure supplement(s) for figure 2:

**Figure supplement 1.** RNAseq libraries are sufficiently complex to detect allele specific expression.

**Figure supplement 2.** Number of reads mapped to LG/J x SM/J pseudo-genome.

**Figure supplement 3.** Stable null permutation plots for allele-specific expression.

**Figure supplement 4.** Multiple tests correction of ASE detection.

**Figure supplement 5.** Volcano plots of parent-of-origin dependent allele-specific expression.

**Figure supplement 6.** Multiple tests correction of DE detection.

**Figure supplement 7.** Volcano plots of differentially expressed genes.

**Figure supplement 8.** Stable null permutation plots for network pairs.

**Figure supplement 9.** Multiple tests correction of pairwise network construction.

**Figure supplement 10.** Volcano plots of network construction.

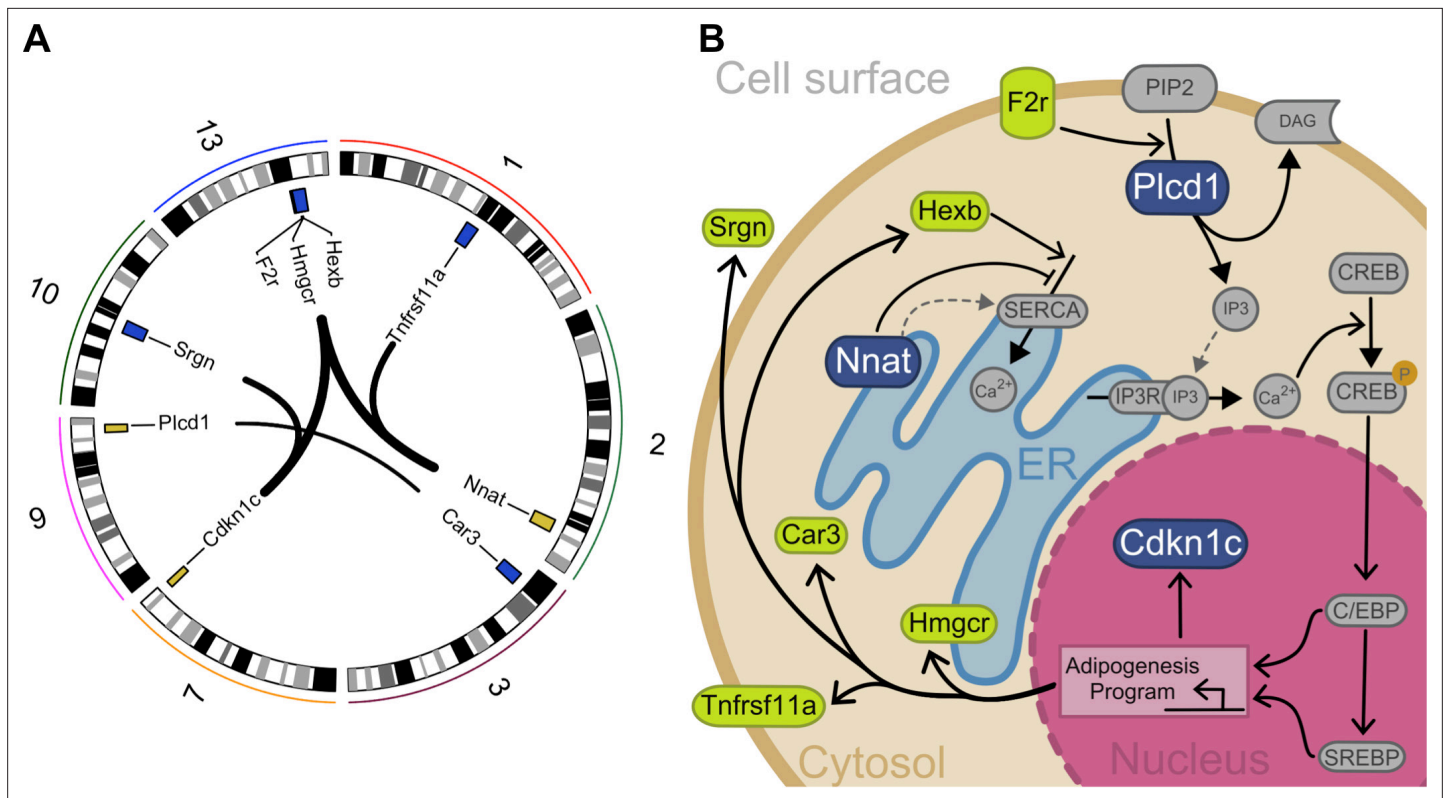
**Figure supplement 11.** Example transformation of  $F_1$  phenotypes.

**Figure supplement 12.** Multiple tests correction of phenotype correlations.

**Figure supplement 13.** Volcano plots of phenotypes correlated with POE net gene expression.

a locus is dependent on the parent-of-origin of alleles at another locus. This allowed us to determine if the effect of parent-of-origin at DE genotype on phenotype is dependent upon the parent-of-origin at ASE genotype. This orthogonal approach allows us to connect genotype at these loci to phenotype as predicted in the  $F_1$  candidates. Nine epistatic interactions replicated in the  $F_{16}$  population ( $n = 1002$  animals,  $FDR \leq 0.1$ ; **Figure 3A**; **Supplementary file 6**). These interactions were comprised of three ASE genes showing parent-of-origin (*Cdkn1c*, *Nnat*, *Plcd1*), six genes that are DE by cross (*Car3*, *F2r*, *Hexb*, *Hmger*, *Srgn*, *Tnfrsf11a*) and four phenotypes (basal glucose level, AUC calculated from a glucose tolerance test, serum cholesterol, necropsy weight). Together, these nine genes form a putative diet-responsive network affecting adipogenesis (**Figure 3B**).

The network can be broken down into signal potentiation, transduction, and response. *Nnat* (neurotatin) and *Hexb* (beta-hexosaminidase subunit beta) fall into the potentiation group. These genes play a role in managing the availability and accumulation of calcium necessary for signal transduction. *Nnat* is a paternally expressed canonically imprinted gene which encodes a proteolipid protein that localizes to the ER (**Li et al., 2010**). *Nnat* is diet-responsive and its overexpression in 3T3L1 pre-adipocytes promotes adipogenesis through increased free cytosolic calcium (**Young et al., 2005**). In pre-neural stem cells, *Nnat* binds sarco/endoplasmic reticulum  $Ca^{2+}$ -ATPase (SERCA) to block  $Ca^{2+}$  uptake into the ER thereby increasing cytosolic  $Ca^{2+}$  levels (**Lin et al., 2010**). In addition to *Nnat*, *Hexb* regulates the uptake and accumulation of  $Ca^{2+}$  in the ER via SERCA (**Pelled et al., 2003**). Upon the arrival of a signal, *F2r* (coagulation factor II receptor) and *Plcd1* (1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase delta-1) in the transduction group initiate the adipogenesis cellular program. *F2r* is a G-protein-bound receptor that promotes phosphoinositide hydrolysis (**Soh et al., 2010**). Variation in the human F2R gene is associated with obesity (**Kichaev et al., 2019**). G-protein-coupled receptors transmit external signals into the cell where they are then propagated by secondary messenger systems, one of which is mediated by *Plcd1* (**Nakamura et al., 2005**; **McDonald and Mamrack, 1995**). The downstream effect of *Plcd1*-mediated signaling is the efflux of calcium into the cytosol from the ER, thereby increasing cytosolic  $Ca^{2+}$  levels (**Thatcher, 2010**; **Berridge, 2016**). Increased cytosolic  $Ca^{2+}$  in pre-adipocytes promotes phosphorylation of cAMP-response element-binding protein (CREB), which promotes activity of CCAAT/enhancer-binding protein (C/EBP) transcription factors, activating adipogenesis, altering the expression of *Cdkn1c* (cyclin dependent kinase inhibitor



**Figure 3.** Interacting genes form a diet-responsive network affecting adipogenesis.

(A) There are nine significant imprinting-by-imprinting epistatic ASE/DE/phenotype sets in the F<sub>16</sub> advanced intercross population (n = 1002). Interactions are shown as lines connecting ASE (yellow) and DE genes (purple). Chromosome number is shown around the plot. (B) The epistatic parent-of-origin effect network is comprised of key steps in a putative pathway regulating differentiation and survival of adipocytes. This pathway was constructed by incorporating previously published cellular functions. The pathway members are color coded in blue for ASE genes (*Plcd1*, *Nnat*, and *Cdkn1c*) and green for DE genes (*F2r*, *Hexb*, *Hmgcr*, *Car3*, *Tnfrsf11a*, and *Srgn*). The network breaks down into potentiation, transduction, and response. *Nnat* and *Hexb* potentiate signaling by managing availability and accumulation of calcium necessary for signal transduction. Once a signal is received, *F2r* and *Plcd1* transduce it by activating second messengers to initiate a response. This response initiates an adipogenesis cellular program that affects expression of *Cdkn1c*, *Hmgcr*, *Car3*, *Tnfrsf11a*, and *Srgn*.

The online version of this article includes the following figure supplement(s) for figure 3:

**Figure supplement 1.** Example transformation of F<sub>16</sub> phenotypes.

**Figure supplement 2.** Stable null permutations plot for epistasis.

**Figure supplement 3.** Representative multiple tests correction of imprinting:imprinting epistasis.

1 C), *Hmgcr* (3-hydroxy-3-methylglutaryl-CoA reductase), *Car3* (carbonic anhydrase 3), *Tnfrsf11a* (TNF receptor superfamily member 11 a), and *Srgn* (serglycin).

*Cdkn1c* is a canonically imprinted maternally expressed gene that inhibits cell proliferation (Kang et al., 2008). Increased expression of *Cdkn1c* is protective against diet-induced obesity in mice (Van de Pette et al., 2018), and in humans increased caloric intake results in decreased CDKN1C expression (Franck et al., 2011). *Hmgcr* is the rate-limiting enzyme in cholesterol biosynthesis (Burg and Espenshade, 2011; Jo and Debose-Boyd, 2010) and converts HMG-CoA into mevalonate, which is essential for adipocyte survival (Yeh et al., 2018). *Srgn* is an adipocytokine thought to be part of a feedback loop with *Tnfa* (tumor necrosis factor alpha), mediating paracrine cross-talk between macrophages and adipocytes (Lemire et al., 2007; Imoto-Tsubakimoto et al., 2013; Schick et al., 2001; Zernichow et al., 2006). *Srgn* is known to play a role in osteoblast-mediated bone mineralization (Bigdeli et al., 2010), which along with osteoclast-driven bone deconstruction drives bone remodeling (Aubin, 1992). Osteoblasts share a lineage with adipocytes, and the quantity of osteoblasts is inversely proportional to that of marrow adipose tissue (Rodríguez et al., 2008; Prockop, 1997; Ali et al., 2005; Akune et al., 2004; Cho et al., 2011; Rosen and Bouxsein, 2006; Turner et al., 2018).

*Tnfrsf11a* is a cell surface protein that regulates differentiation of osteoclasts (Nakagawa et al., 1998). Osteoprotegerin (OPG) is a decoy receptor for TNFRSF11A thereby inhibiting osteoclastogenesis and bone resorption (Matsuo et al., 2020). OPG is expressed during differentiation of 3T3L1 adipocytes (An et al., 2007). Expression of OPG is induced by *Tnfa* in 3T3L1 adipocytes and is associated with obesity in humans (Holecki et al., 2007; Erol et al., 2016; Zaky et al., 2022).

The exact function of OPG/*Tnfrsf11a* outside of osteoclastogenesis is unknown, but the function of osteoclasts is to break down bone tissue during bone resorption. Bone resorption regulates the level of blood calcium. The bioavailability of calcium in the blood potentially alters ER calcium stores, creating cross-talk between bone cells and white adipose tissue calcium signaling. Osteoclasts break down bone by acidifying mineralized bone, orchestrated by osteoblasts that have become embedded in the matrix they produce (osteocytes). Oxidative stress on osteocytes from the bone acidification process is prevented by *Car3*. *Car3* is an enzyme that catalyzes the conversion of carbonic acid to CO<sub>2</sub> and water. Its expression in white adipose is negatively correlated with, and responsive to, long-term obesity in mice and humans (Stanton et al., 1991; Font-Clos et al., 2017). *Car3* does not protect against diet-induced obesity and is not necessary for fatty acid synthesis (Renner et al., 2017). As such its exact function in adipocytes is unknown.

### **Nnat and F2r covary in white adipose tissue and their interaction associates with variation in basal glucose levels across generations**

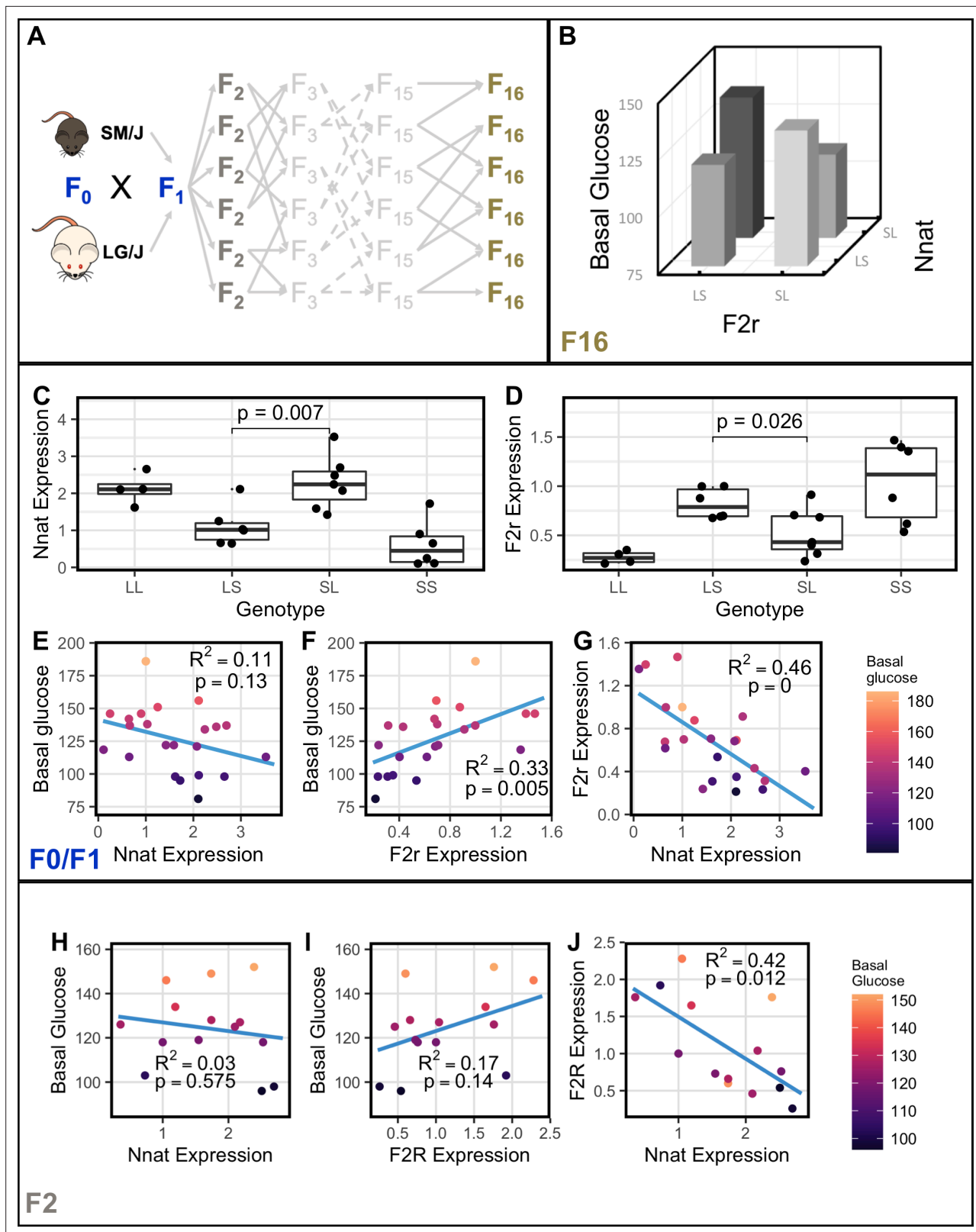
To better understand how these interactions affect phenotype, we focused on the negative correlation of the imprinted gene, *Nnat*, and the biallelic gene, *F2r*, in the above network in high fat-fed females, the cohort with the most significant covariation in the F<sub>1</sub> animals (FDR = 0.05). *Nnat* and *F2r* show significant imprinting-by-imprinting epistasis for basal glucose levels in the F<sub>16</sub> population (FDR = 6.00e<sup>-16</sup>; Figure 4A and B). To validate gene expression patterns, we combined F<sub>1</sub> biological replicates and F<sub>0</sub> high fat-fed female animals (F<sub>1</sub> n = 13 and F<sub>0</sub> n = 12) and again observe that *F2r* and *Nnat* are each significantly differentially expressed between reciprocal heterozygotes, that is by cross (Figure 4C and D). Further, the co-expression of *Nnat* and *F2r* also persists in the F<sub>0</sub>/F<sub>1</sub> population (Figure 4E).

A limitation of identifying covariation patterns in F<sub>1</sub> and F<sub>0</sub> populations is that all loci are linked. This makes it difficult to determine which ASE genes truly co-express with DE genes. While incorporation of orthogonal F<sub>16</sub> genotypes and phenotypes helps reduce false discoveries, a population with randomized genetic background for which we have expression data is needed to replicate these results. To that end, F<sub>2</sub> animals were generated and *Nnat* and *F2r* gene expression levels were measured via qPCR (n = 14). We found that *F2r* and *Nnat* are significantly co-expressed in high-fat-fed female F<sub>2</sub> animals (Figure 4H).

*F2r* expression significantly positively correlates with basal glucose levels in the RNA-sequenced high-fat-fed female F<sub>1</sub> animals ( $r = 0.514$ , FDR = 0.01; Supplementary file 5). *F2r* expression is also significantly positively correlated with basal glucose in the combined F<sub>0</sub>/F<sub>1</sub> population (Figure 4G). A negative trend between *Nnat* expression and basal glucose level is observed but not statistically significant in the combined F<sub>0</sub>/F<sub>1</sub> animals (Figure 4F). Correlation of *F2r*'s and *Nnat*'s individual expression with basal glucose in F<sub>2</sub> mice follows the same pattern as in the F<sub>0</sub>/F<sub>1</sub>'s. Bootstrapping to calculate confidence intervals shows that the correlation differences between F<sub>0</sub>/F<sub>1</sub> and F<sub>2</sub> are not significant (Figure 4I and J; Figure 4—figure supplement 1). However, the product of *Nnat* and *F2r* expression (*Nnat* x *F2r*) is significantly predictive of basal glucose ( $p = 0.045$ ,  $R^2 = 0.29$ ). This indicates that expression of *Nnat* and *F2r*, as a function of their genotypes and allelic parent-of-origin, are not individually sufficient to explain variation in basal glucose levels. But together they are able to explain a significant amount of phenotypic variation. This is precisely what our epistatic model would predict.

Finally, studying the F<sub>2</sub> animals allows us to determine if maternal mitochondrial ancestry contributes significantly to *Nnat* or *F2r* expression or to variation in basal glucose. We find mitochondrial genome identity does not significantly covary with *F2r* expression ( $p = 0.198$ ), *Nnat* expression ( $p = 0.365$ ), or basal glucose ( $p = 0.388$ ).





**Figure 4.** *Nnat* and *F2r* covary across generations.

(A) Breeding scheme for the F<sub>16</sub>. Advanced Intercross between the LG/J and SM/J inbred strains. (B) Significant imprinting-by-imprinting epistasis associated with variation in basal glucose (n = 1002). The parent-of-origin effects of *F2r* on basal glucose depend on the parent-of-origin effects at *Nnat*. (C) Expression of *Nnat* across genotypes in a combined F<sub>0</sub>/F<sub>1</sub> population of high fat-fed females (n = 25). (D) Expression of *F2r* across genotypes

Figure 4 continued on next page

Figure 4 continued

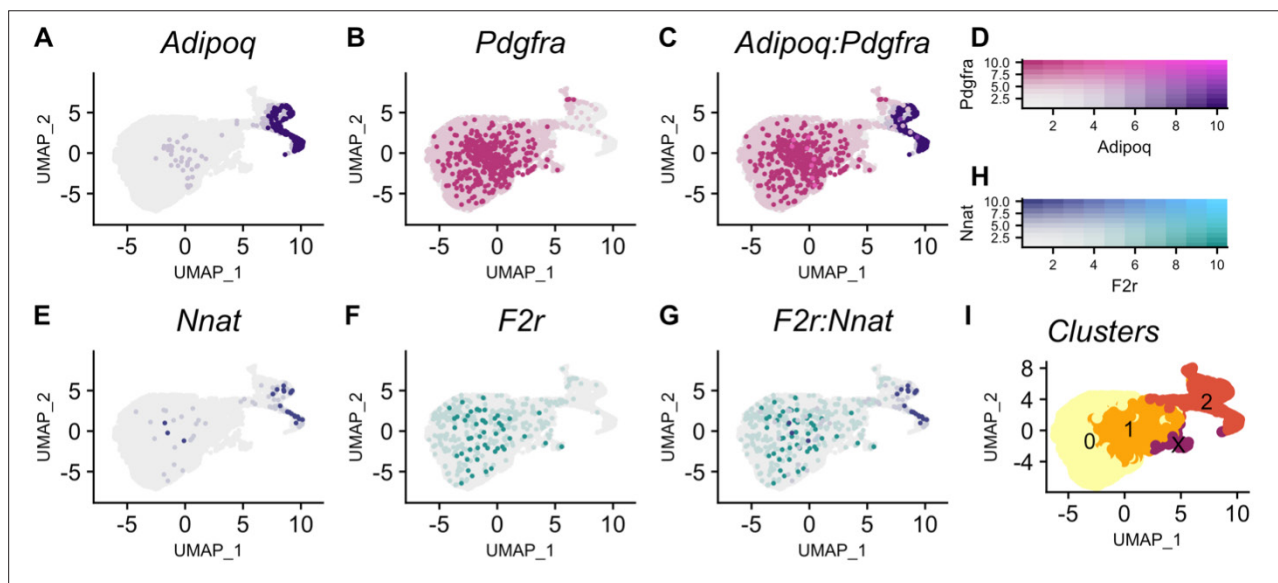
in a combined  $F_0/F_1$  population of high-fat-fed females ( $n = 25$ ). (E) Significant correlation between *Nnat* and *F2r* expression in the  $F_0/F_1$  mice ( $F_1$   $n = 13$ ;  $F_0$   $n = 12$ ). (F) and (G) Correlations between basal glucose and *Nnat* and *F2r* in the  $F_0/F_1$  mice ( $F_1$   $n = 13$ ;  $F_0$   $n = 12$ ). (H) Significant correlation between *Nnat* and *F2r* expression in the high fat-fed female  $F_2$  mice ( $n = 14$ ). (I) and (J) Correlations between basal glucose and *Nnat* and *F2r* are not individually significant in the  $F_2$  mice. Alleles are ordered maternal | paternal within the genotype classes.

The online version of this article includes the following figure supplement(s) for figure 4:

**Figure supplement 1.** Pearson's correlation coefficient confidence intervals.

## Single-cell RNAseq reveals that *Nnat* expression increases and *F2r* expression decreases in pre-adipocytes along an adipogenic trajectory

To determine what cell types express *Nnat* and *F2r* and whether the directionality of the *Nnat* imprinted  $\rightarrow$  *F2r* target correlation persists along the adipogenic trajectory, we turned to single-cell RNAseq. We used publicly available scRNAseq data collected from stromal vascular cells isolated from C57BL/6 J epididymal adipose tissue (Burl et al., 2018). Cell type identity was assigned using previously reported markers for this data set (Adipoq = differentiating mesenchymal stem cells; Pdgfra = mesenchymal stem cells; Csf1r = macrophage; Cdh5 = vascular endothelial cells; Acta2 = vascular smooth muscle cells; Cd2 = B cells) (Supplementary file 7; Figure 5—figure supplement 1). The adipogenic trajectory refers to cells transitioning from pre-adipocytes (mesenchymal stem cells) to cells differentiating into adipocytes. Clusters along this trajectory were identified by the opposing expression patterns of *Pdgfra* and *Adipoq* (Figure 5A-D and I). We found that *Nnat* expression increases along the trajectory while *F2r* expression decreases (Figure 5E-F and H). Further there



**Figure 5.** *Nnat* expression increases and *F2r* expression decreases in pre-adipocytes along an adipogenic trajectory.

(A) *Adipoq* is a marker of adipocytes whose expression (purple) increases along the trajectory. (B) *Pdgfra* is a marker of mesenchymal stem cells whose expression (pink) decreases along the trajectory. (C) Cells in clusters expressing one or both *Adipoq* and *Pdgfra* fall along an adipogenic trajectory. (D) Intensity of expression of *Adipoq* and *Pdgfra* indicated by coloration. (E) *Nnat* expression (blue) increases along the trajectory. (F) *F2r* expression (teal) decreases along the trajectory. (G) Negative association between *Nnat* and *F2r* expression within adipocytes along the trajectory. (H) Intensity of expression of *Nnat* and *F2r* indicated by coloration. (I) The adipogenic trajectory is broken into subclusters of cells with no *Adipoq* expression (cluster 0) to high *Adipoq* expression (cluster 2).

The online version of this article includes the following figure supplement(s) for figure 5:

**Figure supplement 1.** Cell types were defined using canonical markers.

**Figure supplement 2.** Candidate genes are differentially expressed across adipogenic trajectory.

**Figure supplement 3.** Single cell quality was controlled.

**Figure supplement 4.** Determining the resolution for clustering.

**Figure supplement 5.** Adipocyte clustering resolution was selected to minimize *Adipoq* variation.

is a negative association between *Nnat* and *F2r* expression within adipocytes along the trajectory (**Figure 5G**). This pattern is consistent with the negative correlation we observe between *Nnat* and *F2r* in the bulk white adipose tissue. Because available scRNAseq data do not match the exact sex/ diet/ genetic background contexts of the LGxSM mice, there will be unaccounted for differences between the data sets. However, the observed consistent pattern indicates that the pathway structure persists across sex/ diet/ genetic backgrounds.

In addition to interrogating *Nnat* and *F2r* in single cells along an adipogenic trajectory, we found that eight of the nine genes comprising the epistatic parent-of-origin effect network described above are differentially expressed along the trajectory, and they associate with cell types that are consistent with their respective roles in adipose tissue (**Figure 5—figure supplement 2**).

## Discussion

Epistatic interactions between imprinted and non-imprinted genes can influence complex traits when the genotypic effects of one gene depends on the parent-of-origin of alleles at another (**Lawson et al., 2013; Wolf and Cheverud, 2009**). Here, we examined epistatic interactions associated with parent-of-origin effects on dietary-obesity traits in white adipose using a simple yet powerful  $F_1$  reciprocal cross mouse model. Although these parent-of-origin dependent allele-specific expression biases are consistent with imprinting mechanisms, we cannot rule out that maternal and/or paternal effects also contribute to the phenomena we observe (**Hager et al., 2008**).

Interactions between imprinted and non-imprinted genes have previously been shown to contribute to variation in metabolic phenotypes. For example, the maternally expressed transcription factor KLF14 (kruppel-like factor 14) regulates biallelic gene expression related to adiposity (**Parker-Katirae et al., 2007; Small et al., 2011**). Mapping studies have identified two SNPs (rs4731702, rs972283) upstream of KLF14 associated with type II diabetes and cholesterol levels (**Voight et al., 2010; Teslovich et al., 2010**). Both variants have maternally restricted *cis*-regulatory associations with KLF14 expression in adipose tissue (**Kong et al., 2009**). eQTL analysis found that rs4731702 is also enriched for *trans*-associations with KLF family transcription factor binding sites in subcutaneous white adipose tissue, suggesting that KLF14 may be a master transcriptional regulator in adipose tissue (**Small et al., 2011**). Whether additional pairs of imprinted and biallelic genes are similarly co-expressed and associate with phenotypic variation remains an open question that has not been thoroughly investigated in large landmark functional genomics studies including ENCODE, GTEx, and GWAS, leaving a significant gap in our knowledge. Interactions between imprinted and biallelic genes could explain some of the observed parent-of-origin effect patterns associated with regions lacking obvious candidate genes, as described in a recent survey of 97 complex traits measured in outbred mice (**Mott et al., 2014**).

Our model asserts that parent-of-origin effects start at ASE genes and are transduced through DE genes onto phenotype. This is illustrated in the interaction between *Nnat* and *F2r*. If a *cis*-regulatory effect interacts with epigenetic modifications (i.e. imprinting) at *Nnat*, then *Nnat* expression of genotypic classes are affected by paternal allele identity (**Lawson et al., 2013**). Between the LG/J and SM/J alleles at *Nnat*, the LG/J allele is more highly expressed. If our model is correct, the downstream DE gene should show a corresponding pattern (**Figure 1B**). In the case of *Nnat* and *F2r*, which have strong negative correlated expression, when the LG/J allele is inherited paternally at *Nnat*, the higher expression of *Nnat* should correspond with lower expression of *F2r*. This is what we observe (**Figure 4**). If this relationship is true, we should see persistent co-expression of *Nnat* and *F2r* across genetic backgrounds ( $F_0$ ,  $F_1$ ,  $F_2$ ), which we do (**Figure 4**). This supports a biologically meaningful relationship between *Nnat* and *F2r*. Our model further predicts that the DE genes should more closely affect phenotype (**Pierce et al., 2014; Shan et al., 2019; Lutz and Hokanson, 2015**). In the case of *Nnat* and *F2r*, we expect *F2r* to more strongly associate with basal glucose levels than *Nnat*, which we observe (**Figure 4**).

There is a clear relationship between *Nnat* and *F2r* in adipogenesis, but the specifics of how this relationship extends to glucose homeostasis are unclear. One possibility is that by altering SERCA function, *Nnat* affects not only the formation of new adipocytes, but also the being of adipocytes. The SERCA channel is uncoupled in beige adipocytes as part of a UCP1-independent form of non-shivering thermogenesis (**Ikeda and Yamada, 2020**). Non-shivering thermogenesis consumes a significant amount of energy, thereby altering glucose homeostasis (**Carson et al., 2020**). This hypothesis

links these genes to physiological processes that are consistent with evolutionary hypotheses about the prevalence of parent-of-origin effects. *Nnat* and *F2r* are members of a putative network we identified that is coordinated by interactions between ASE and DE genes. From the literature, we found that the genes in this epistatic network function in key steps in a pathway regulating differentiation and survival of adipocytes in response to nutritional environment (**Figure 3B**). Specifically, there is evidence that it plays a critical role in the induction of adipogenesis. This alone demonstrates how parent-of-origin effects can move through networks along molecular pathways. Beyond proof-of-principle this network provides a clue to the puzzle of the prevalence of parent-of-origin effects.

The constituents of this single network appear to play vastly different physiological roles depending on the tissue. In white adipose the network appears to play some role in balancing proliferation, differentiation, and apoptosis as we describe above. In pancreatic  $\beta$ -cells, members of this network affect insulin secretion (**Millership et al., 2018**). In bone, members of this network affect the balance of cartilage/bone growth and reabsorption. These three physiological processes may at first seem unrelated, but they share one key commonality – they are jointly critical to growth. This is consistent with the sexual conflict hypothesis attributed to parent-of-origin effects (**Patten et al., 2014; Babak et al., 2015**). The size of progeny in placental mammals can have opposing fitness consequences for mothers/ maternal relatives and fathers/ paternal relatives. The fitness of fathers and paternal relatives, particularly in the case of multi-paternity litters, is improved with larger progeny (**Mochizuki et al., 1996; Babak et al., 2015; Fowden and Moore, 2012; Patten et al., 2014; Wilkins and Haig, 2003; Haig, 1997**). This comes at a fitness disadvantage to the mother. The fitness of mothers is improved by progeny of a manageable size, allowing her to produce multiple litters.

According to this model, imprinting evolved in part to allow one parental lineage to hijack parts of a nutritional environment response pathway driving growth in a direction favorable to maximize the fitness of that lineage. Key processes in such a pathway driving growth would include the secretion of growth factors, construction of cartilage and bone, and the accumulation of energy stores. We present a network that appears to play a role in all three processes. If the sexual conflict hypothesis is true, then the most parsimonious place for imprinting to evolve would be in key regulatory points that affect as many aspects of growth as possible. This is consistent with the network we identified, a single pathway affecting many aspects of growth. This hints at the possibility that parent-of-origin effects are common because of the multi-purpose nature of the pathways in which genomic imprinting manifests and parent-of-origin effects propagate.

By leveraging the reciprocal  $F_1$  hybrids, we are able to integrate parent-of-origin-dependent allele specific expression and parent-of-origin-dependent differential expression with  $F_{16}$  phenotypes. By doing so, we identify plausible candidates for functional validation and describe discrete molecular networks that may contribute to the observed parent-of-origin effects on metabolic phenotypes. The genes and interactions we present here represent a set of actionable interacting candidates that can be probed to further identify the machinery driving these phenomena and make predictions informed by genomic sequence. The frameworks we have developed account for the genetic, epigenetic, and environmental components underlying these parent-of-origin effects, thereby improving our ability to predict complex phenotypes from genomic sequence. We focused on metabolic phenotypes in this study, but the patterns we identified may translate to other complex traits where parent-of-origin effects have been implicated.

## Materials and methods

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Other	High fat diet	Teklad	TD88137	42% kcal from fat
Other	Low-fat diet	Research Diets	D12284	15% kcal from fat
Commercial assay, kit	RNeasy Lipid Tissue Kit	QIAgen	74,804	
Commercial assay, kit	RiboZero kit	Illumina	20040529	
Commercial assay, kit	DNA 1000LabChip	Agilent	5067–1504	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay, kit	High-Capacity cDNA Reverse Transcription Kit	Thermo Fisher	4368814	
Sequence-based reagent	<i>Nnat</i> forward primer	This paper		Detailed information is found in the methods section
Sequence-based reagent	<i>Nnat</i> reverse primer	This paper		Detailed information is found in the methods section
Sequence-based reagent	<i>F2r</i> forward primer	This paper		Detailed information is found in the methods section
Sequence-based reagent	<i>F2r</i> reverse primer	This paper		Detailed information is found in the methods section
Sequence-based reagent	<i>L32</i> forward primer	This paper		Detailed information is found in the methods section
Sequence-based reagent	<i>L32</i> reverse primer	This paper		Detailed information is found in the methods section
Software, algorithm	R	R	3.6.1	
Software, algorithm	STAR	STAR	DOI: <a href="https://doi.org/10.1093/bioinformatics/bts635">10.1093/bioinformatics/bts635</a>	
Software, algorithm	FASTQC	FASTQC	other	<a href="https://www.bioinformatics.babraham.ac.uk/projects/fastqc/">https://www.bioinformatics.babraham.ac.uk/projects/fastqc/</a>
Software, algorithm	EdgeR	CRAN	DOI: <a href="https://doi.org/10.1093/bioinformatics/btp616">10.1093/bioinformatics/btp616</a>	
Software, algorithm	WEB-based Gene Set Analysis Toolkit	WEB-based Gene Set Analysis Toolkit	DOI: <a href="https://doi.org/10.1093/nar/gkz401">10.1093/nar/gkz401</a>	
Software, algorithm	Seurat	Seurat	DOI: <a href="https://doi.org/10.1038/nbt.3192">10.1038/nbt.3192</a>	
Strain, strain background ( <i>Mus musculus</i> )	SM/J	The Jackson Laboratory	000687	
Strain, strain background ( <i>Mus musculus</i> )	LG/J	The Jackson Laboratory	000675	

## Mouse husbandry and phenotyping

LG/J and SM/J founders ( $F_0$ ) were obtained from The Jackson Laboratory (Bar Harbor, ME).  $F_1$  reciprocal cross animals were generated by mating LG/J mothers with SM/J fathers (LxS) and the inverse (SxL).  $F_2$  intercrossed animals were generated by mating LxS mothers with SxL fathers and the inverse. After weaning at 21 days, animals were separated into sex-specific cages of 3–5 animals per cage and randomly placed on high-fat (42% kcal from fat; Teklad TD88137) or low-fat (15% kcal from fat; Research Diets D12284) isocaloric diets. Feeding was ad libitum. There were 96 experimental  $F_1$  animals in total, with 48 animals for each cross (LxS and SxL). Within each cross, there were 24 high-fat-fed animals (12 males; 12 females) and 24 low-fat-fed animals (12 males; 12 females). The  $F_2$  animals were generated for a different study, following the same weaning protocol and diets, and we used data from the high fat-fed females ( $n = 14$ ) for validation in the the current study (Carson *et al.*, 2020). Additionally, we used data generated from founder  $F_0$  (LG/J ( $n = 6$ ) and SM/J (6)) high fat-fed female animals, also generated for a different study and subjected to the same weaning protocol and diets (Carson *et al.*, 2020). The barrier animal facilities at WUSM follow a 12/12 hr light/dark schedule, all water is autoclaved and changed weekly, and all cages are changed weekly.

All animals were weighed weekly from three weeks of age until sacrifice. At 19 weeks of age, body composition was determined by MRI and a glucose tolerance test was performed after a 4 hr fast. At 20 weeks of age, animals were given an overdose of sodium pentobarbital after a 4 hr fast and blood was collected via cardiac puncture. Euthanasia was achieved by cardiac perfusion with phosphate-buffered saline. After cardiac perfusion, the reproductive fat pad was harvested, flash frozen in liquid nitrogen, and stored at  $-80$  °C.

## Study design

The weaning, phenotyping protocols, and diets were chosen to reproduce the protocols and diets used in studies of the F16 Advanced Intercross of the LG/J x SM/J inbred mouse lines that were used in previously published mapping studies ([Cheverud et al., 2011](#); [Lawson et al., 2011a](#); [Lawson et al., 2011b](#), [Lawson et al., 2010](#)). The experimental unit for the current study is the individual mouse. For the RNA sequencing, a single animal was randomly chosen from each cage using a random number generator in R. All other animals served as biological replicates. Mice from multiple cages representing different crosses, generations, diets, and sexes, were necropsied at the same time to avoid batch effects. Library prep and RNA sequencing was performed blinded by the WUSM Genome Technology and Access Center.

## Genomes and annotations

LG/J and SM/J indels and SNVs were leveraged to construct strain-specific genomes using the GRC38.72-mm10 reference as a template ([Nikolskiy et al., 2015](#)). This was done by replacing reference bases with alternative (LG/J | SM/J) bases using custom python scripts. Ensembl R72 annotations were adjusted for indel-induced indexing differences for both genomes.

## RNA sequencing

Total RNA was isolated from adipose tissue using the RNeasy Lipid Tissue Kit (QIAGEN) (n = 32, 4 animals per sex/diet/cross cohort). RNA concentration was measured via NanoDrop and RNA quality/integrity was assessed with a BioAnalyzer (Agilent). RNA-Seq libraries were constructed using the RiboZero kit (Illumina) from total RNA samples with RIN scores > 8.0. Libraries were checked for quality and concentration using the DNA 1000LabChip assay (Agilent) and quantitative PCR, according to manufacturer's protocol. Libraries were sequenced at 2 × 100 paired end reads on an Illumina HiSeq 4,000. After sequencing, reads were de-multiplexed and assigned to individual samples. RNAseq data are available through the NCBI-SRA, accession: PRJNA753198.

## Library complexity

Complexity was measured by fitting a beta-binomial distribution to the distribution of  $L_{\text{bias}}$  values using the VGAM package ([Yee, 2010](#)). The shape parameters ( $\alpha$ ,  $\beta$ ) of beta-binomial distributions were estimated and used to calculate dispersion ( $\rho$ ). Dispersion values less than 0.05 indicate our libraries are sufficiently complex ([Figure 2—figure supplement 1](#)).

$$\rho_s = \frac{1}{1 + \alpha_s + \beta_s}$$

One library was found to have insufficient complexity and was removed from the analyses.

## Allele-specific expression

FASTQ files were filtered to remove low quality reads and aligned against both LG/J and SM/J pseudo-genomes simultaneously using STAR with multimapping disallowed ([Dobin et al., 2013](#)). Read counts were normalized via upper quartile normalization and a minimum normalized read depth of 20 was required. Alignment summaries are provided in [Supplementary file 8](#) and [Figure 2—figure supplement 2](#).

For each gene in each individual, allelic bias ( $L_{\text{bias}}$ ) was calculated as the proportion of total reads for a given gene with the LG/J haplotype. Parent-of-origin-dependent allele-specific expression was detected by ANOVA using one of a number of models in which  $L_{\text{bias}}$  is responsive to cross and the interaction of cross with some combination of sex and diet:

$$\text{Model} \left\{ \begin{array}{l} \text{if each Cross context has } \geq 2 \text{ samples, } L_{\text{bias}} \sim \text{Cross} \\ \text{if each Cross: Sex context has } \geq 2 \text{ samples, } L_{\text{bias}} \sim \text{Cross} + \text{Cross: Sex} \\ \text{if each Cross: Diet context has } \geq 2 \text{ samples, } L_{\text{bias}} \sim \text{Cross} + \text{Cross: Diet} \\ \text{if each context has } \geq 2 \text{ samples, } L_{\text{bias}} \sim \text{Cross} + \text{Cross: Sex} + \text{Cross: Diet} + \text{Cross: Sex: Diet} \end{array} \right.$$

Accurately estimating the significance of these effects and correcting for multiple tests is challenging for two reasons: (1) the complexity of the many environmental contexts and (2) the correlation of allelic bias within and between imprinted domains breaks assumptions of independence. A permutation approach is an effective way to overcome these challenges. The context data was randomly shuffled for each gene independently and analyses were rerun to generate a stable null distribution of p-values (**Figure 2—figure supplement 3**). False discovery rates were estimated for a given significance threshold as the proportion of significant tests under the permuted null model relative to significant tests under the real data model. A value of 1 meaning that 100% of tests at a given significance threshold are likely false positives. An  $FDR \leq 0.1$  was considered significant (**Supplementary file 1, Figure 2—figure supplement 4**).

To determine the parental direction and size of expression biases, a parent-of-origin effect score was calculated as the difference in mean  $L_{bias}$  between reciprocal crosses (LxS or SxL). Parent-of-origin effect scores range from completely maternally expressed (−1), to biallelic (0), to completely paternally expressed (+1). Parent-of-origin effect score thresholds were calculated from a critical value of  $\alpha = 0.01$ , determined from a null distribution created by permutation. Genes with significant allele-specific expression and parent-of-origin scores beyond the critical value were considered to have significant parent-of-origin-dependent allele-specific expression (**Figure 2—figure supplement 5**).

## Differential expression

Differential expression by reciprocal cross was determined by first aligning reads against the LG/J and SM/J genomes simultaneously with multimapping permitted. Reads were normalized by Trimmed mean of M-values (TMM) normalization, which estimates scale factors among samples to allow for differences in RNA composition (**Robinson and Oshlack, 2010**). A minimum normalized read count of 10 was required. Generalized linear models accounting for diet, sex, and diet-by-sex were fit in EdgeR (**Robinson et al., 2010**). Differential expression was detected by a likelihood ratio test. Significance was determined for five models for each gene:

1.  $Expression \sim Cross$
2.  $Expression \sim Cross: Sex$
3.  $Expression \sim Cross: Diet$
4.  $Expression \sim Cross: Sex: Diet$
5.  $Expression \sim Cross + Cross: Sex + Cross: Diet + Cross: Sex: Diet$

Multiple test corrections were performed by implementing the ‘qvalue’ R package to estimate false discovery rates (**Figure 2—figure supplement 6**). Genes with a  $FDR \leq 0.1$  and a  $|fold\ change| \geq 1.5$  were considered significantly differentially expressed by reciprocal cross (**Figure 2—figure supplement 7** and **Supplementary file 2**).

## Gene-gene interactions

Networks were constructed in each tissue by pairing genes showing parent-of-origin-dependent allele-specific expression with biallelic genes that are differentially expressed by cross. Pairs were predicted by modeling the expression of biallelic genes as a function of parent-of-origin-dependent allele-specific expression,  $L_{bias}$ , sex, diet, and sex-by-diet. The strength of a prediction was measured through model fit, which was estimated as a mean test error with 10-fold cross-validation employed to prevent overfitting. Significance was estimated by likelihood ratio test using a chi-square distribution. Given the complexity of contexts, false discovery rates were determined by permuting the context and expression data to generate a stable null-distribution of p-value (**Figure 2—figure supplement 8**). Null distribution stability was evaluated by calculating the critical value for  $\alpha = 0.05$  at each genome wide iteration. The standard deviation of critical values was calculated after each iteration for the last 5 iterations. Genome-wide shuffling was done 500 times, with 759 independent randomized tests per iteration, meaning the stable null model is composed of 379,500 randomized observations. Using the null model, the ‘qvalue’ package estimated a  $\hat{\pi}_0$ . This estimate was then used to estimate false discovery rates in the real data. MTE score thresholds were calculated from a critical value of  $\alpha = 0.01$ , determined from a null distribution created by permutation (**Figure 2—figure supplement 9**).

Connections with an FDR  $\leq 0.1$  (**Supplementary file 9**) and MTE below the critical value were considered significant (**Figure 2—figure supplement 10**).

### Functional enrichment analysis

Functional enrichment of interacting genes showing parent-of-origin-dependent allele-specific expression with biallelic genes that are differentially expressed by cross was tested by over-representation analysis in the WEB-based Gene Set Analysis Toolkit v2019 (**Zhang et al., 2005**). We performed analyses of gene ontologies (biological process, cellular component, and molecular function), pathway (KEGG), and phenotype (Mammalian Phenotype Ontology). The list of all unique interacting genes was analyzed against the background of all unique genes expressed in white adipose. A Benjamini-Hochberg FDR-corrected p-value  $\leq 0.01$  was considered significant (**Supplementary file 4**).

### Phenotype correlation

In order to identify which phenotypes might be affected by genes in the parent-of-origin effects network, gene expression was correlated with metabolic phenotypes collected for F<sub>1</sub> animals with the contexts combined. Phenotypes were log transformed when necessary, as determined by Shapiro Wilkes test to approximate normality (**Figure 2—figure supplement 11**). Additionally, the effects of sex and diet were residualized out leaving only the effect of cross. This was done to mirror later residualizing of phenotypes in the F16 population when testing for epistasis. Multiple test corrections were performed by implementing the 'qvalue' R package to estimate false discovery rates (**Figure 2—figure supplement 12**). The minimum Pearson's correlation coefficient threshold was set to  $|0.5|$ . Connections with an FDR  $\leq 0.05$  (**Supplementary file 5**) and MTE below the critical value were considered significant (**Figure 2—figure supplement 13**).

### Epistasis testing

The F<sub>16</sub> LxS advanced intercross population, phenotypes, genotypes, genotypic scores, and QTL mapping methods are described elsewhere (**Cheverud et al., 2011; Lawson et al., 2011a; Lawson et al., 2011b, Lawson et al., 2010**). We tested for epistasis in interacting pairs between genes showing parent-of-origin-dependent allele-specific expression and biallelic genes that are differentially expressed by cross. We selected F<sub>16</sub> genotyped markers that fall within 1.5 Mb up- and downstream from the geometric center of each gene, defined as the genomic position halfway between the transcription start and stop position of that gene (**Supplementary file 10**). For every F<sub>16</sub> animal, an 'imprinting score' was assigned to each marker based on that animal's genotypic values (LL = 0, LS = 1, SL = -1, SS = 0; maternal allele is depicted first). Non-normally distributed phenotypes (as evaluated by a Shapiro-Wilk test) were log<sub>10</sub>-transformed to approximate normality (**Figure 3—figure supplement 1**). Because of the number of epistasis tests performed and the number of contexts represented in the data, we removed the effects of sex, diet and their interaction from each F<sub>16</sub> phenotype with a covariate screen. We tested for epistasis on the residualized data using the following generalized linear model:

$$R_{pheno} \sim BDE_{IMP} + ASE_{IMP} + BDE_{IMP}: ASE_{IMP}$$

where  $R_{pheno}$  is the residual phenotype,  $BDE_{IMP}$  is the imprinted genotypic score for the biallelic gene that is differentially expressed by cross,  $ASE_{IMP}$  is the imprinted genotypic score for the gene showing parent-of-origin-dependent allele-specific expression bias, and  $BDE_{IMP}:ASE_{IMP}$  is the interaction between the two genes' imprinted genotypic score. We employed a permutation approach to accurately estimate significance given the linkage of proximal markers. Imprinted genotypic values were randomly shuffled to generate a stable null model of p-values (**Figure 3—figure supplement 2**). False discovery rates were estimated for a given significance threshold as the proportion of significant tests under the permuted null model relative to significant tests under the real data model (**Figure 3—figure supplement 3**). An FDR  $\leq 0.1$  was considered significant. Epistasis was considered significant if the  $BDE_{IMP}: ASE_{IMP}$  interaction term met the significance threshold (**Supplementary file 6**).

### Validation of Nnat and F2r expression patterns

Expression patterns of *Nnat* and *F2r* in white adipose were validated by qRT-PCR in high-fat-fed female LG/J and SM/J mice and in biological replicates of high-fat-fed female F<sub>1</sub> reciprocal cross



animals ( $n = 6$  LG/J homozygotes,  $n = 10$  LxS and 10 SxL reciprocal heterozygotes,  $n = 6$  SM/J homozygotes). Total RNA was extracted from adipose samples using the Qiagen RNeasy Lipid Kit. High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher) was used for reverse transcription. Quantitative RT-PCR was performed with an Applied Biosystems (USA) QuantStudio 6 Flex instrument using SYBR Green reagent. Results were normalized to *L32* expression using the  $\Delta\Delta C_t$  method. *Nnat* forward primer – CTACCCCAAGAGCTCCCTTT and reverse primer – CAGCTTCTGCAGGAGTACC. *F2r* forward primer – TGAACCCCGCTCATTCTTTC and reverse primer – CCAGCAGGACGCTTTC. *L32* forward primer – TCCACAATGTCAAGGAGCTG and reverse primer – GGGATTGGTGAC TCTGATGG. Data points were considered outliers if they led to violation of normality assumptions or were considered outliers by box and whisker plots. ANOVA was used to estimate significance of differential expression by cross (1), paternal allele identity (2), mitochondrial ancestry (3).

$$\begin{aligned}
 1. \text{ Expression} \sim \text{Cross} &\in \begin{cases} LL, 0 \\ LS, -1 \\ SL, 1 \\ SS, 0 \end{cases} \\
 2. \text{ Expression} \sim \text{Paternal Allele} &\in \begin{cases} LL, 0 \\ LS, 1 \\ SL, 0 \\ SS, 1 \end{cases} \\
 3. \text{ Expression} \sim \text{Mitochondrial ancestry} &\in \begin{cases} LxS \times SxL, 0 \\ SxL \times LxS, 1 \end{cases}
 \end{aligned}$$

Expression patterns were also validated by qRT-PCR in high fat-fed female  $F_2$  animals ( $n = 14$ ). Co-expression was determined by fitting a general linear model and estimating significance using the Wald test approximation of the LR test. Correlation with basal glucose was determined by fitting a general linear model and estimating significance using the Wald test approximation of the LR test. Pearson's correlation coefficients were calculated for each gene with basal glucose. To test whether patterns in these correlations was significantly different between  $F_0/F_1$  and  $F_2$  populations, bootstrapping was used to calculate 90% confidence intervals for the Pearson's correlation coefficients. 5,000 iterations were run with 10 individuals randomly selected with replacement. scRNA analysis of *Nnat* and *F2r* scRNAseq data was downloaded from SRA: SRP145475 (Burl et al., 2018). Data were processed and aligned to the C57BL/6 J reference (mm10) using Cell Ranger (Zheng et al., 2017). Analysis and cell quality control was performed using the Seurat (3.2.2) (Stuart et al., 2019) package in R (3.6.1) (R Development Core Team, 2013). Cell quality was controlled using three metrics (Luecken and Theis, 2019): (1) number of features, (2) number of counts, (3) covariation of features and counts. High quality cells were required to have between 500 and 3000 features and read counts between 1000 and 30,000. As sequencing is a process of random sampling, the number of features and the number of counts should covary. This relationship was fit to a generalized additive model. Deviation from this relationship (residuals) were computed for each cell. High-quality cells were required to have a residual within 3 standard deviations of the mean residual of all cells (Figure 5—figure supplement 3).

Seurat normalization with a scale factor of 10,000 was performed. Dimensionality reduction (UMAP) was performed (dims = 1:10, resolution = 0.15). Resolution was chosen using the clustree (0.4.3) package (Zappia and Oshlack, 2018). A range of resolutions from 0.06 to 0.18 were tested, and the highest resolution with stable clustering was chosen (Figure 5—figure supplement 4). Cell type markers were identified by differential expression analysis using the 'MAST' hurdle-model test (Finak et al., 2015). Genes overexpressed in a given cell type relative to all other cell types were considered cell type 'markers'. Cell type identity was assigned using previously reported markers for this data set (Figure 5—figure supplement 1).

Cells along the adipogenic trajectory were subset and subjected to dimensionality reduction (UMAP, dims = 1:10, resolution = 0.17). A range of resolutions from 0.01 to 0.25 were tested. Using *Adipoq* as a marker of differentiation, we sought to identify the set of clusters that would best encapsulate the stages of differentiation. To this end for every level of resolution we calculated the mean count variance ( $\bar{C}_\sigma$ ). This is done by calculating the standard deviation ( $\sigma$ ) of *Adipoq* expression ( $E$ ) within each cluster ( $G$ ), referred to as the count variance ( $C_\sigma$ ). Cells with no expression of *Adipoq* were excluded. The mean of count variances for all clusters is calculated. This process is similar to k-means clustering, where the goal is to find that parameters which minimize the within group variation.

$$\text{Count Variance} = \frac{\sum_{G=1}^n \sigma(E_G)}{n}$$

We also calculated the percent expressing variance ( $\bar{P}_\sigma$ ). This was taken as the mean of the standard deviation in the percent of cells expressing *Adipoq*.

$$\text{Percent Expressing Variance} = \frac{\sum_{G=1}^n \sigma(\%E>0_G)}{n}$$

The resolution 0.17 was chosen as the lowest resolution where variation is minimized and no longer significantly changes (**Figure 5—figure supplement 5**). Using *Adipoq* as a marker of adipogenesis, clusters 1 and 2 were identified as pre- and post-differentiated cells, respectively. Differential expression was analyzed using the ‘MAST’ test. Expression was compared between clusters 1 and 2 only. Multiple tests correction was performed using the Bonferroni method. We required changes in expression to show either a sufficiently large fold change ( $|\log_2 \text{FoldChange}| \geq 0.3$ ) OR a sufficiently large change in the percent of cells expressing the gene in question ( $\text{pct.}\Delta \geq 0.4$ ). The change in percent of cells expressing a gene was calculated as the difference in percent of cells expressing the gene between the clusters and scaled by dividing by the larger percentage.

$$\text{pct.}\Delta = \frac{\text{pct.2} - \text{pct.1}}{\max(\text{pct.1}, \text{pct.2})}$$

Source code is available at [https://github.com/LawsonLab-WUSM/POE\\_Epistasis](https://github.com/LawsonLab-WUSM/POE_Epistasis), (copy archived at [swh:1:rev:b39046ce35f53e0c3f15bcdefa122c274aee48b7](https://www.swh.io/rev/b39046ce35f53e0c3f15bcdefa122c274aee48b7), Lawson, 2019).

## Acknowledgements

This work was supported by the Washington University Department of Genetics, the Diabetes Research Center at Washington University (grant P30DK020579), the NIH NIDDK (grant K01 DK095003), the NIEHS (grant U24ES026699), and the NIH NHGRI (grant T32-GM007067). The authors declare no conflicts of interest.

## Additional information

### Funding

Funder	Grant reference number	Author
Diabetes Research Center		Clay F Semenkovich
Diabetes Research Center		Heather A Lawson

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Juan F Macias-Velasco, Formal analysis, Methodology, Visualization, Writing - original draft, Writing - review and editing; Celine L St Pierre, Visualization, Writing - review and editing; Jessica P Wayhart, Li Yin, Larry Spears, Caryn Carson, Katsuhiko Funai, Investigation; Mario A Miranda, Validation; James M Cheverud, Clay F Semenkovich, Resources, Writing - review and editing; Heather A Lawson,

Conceptualization, Funding acquisition, Methodology, Supervision, Writing - original draft, Writing - review and editing

#### Author ORCIDs

Juan F Macias-Velasco  <http://orcid.org/0000-0003-2827-4647>

Celine L St Pierre  <http://orcid.org/0000-0001-5465-6601>

Clay F Semenkovich  <http://orcid.org/0000-0003-1163-1871>

Heather A Lawson  <http://orcid.org/0000-0002-3550-5485>

#### Ethics

Mouse colony was maintained at the Washington University School of Medicine and all experiments were approved by the Institutional Animal Care and Use Committee in accordance with the National Institutes of Health guidelines for the care and use of laboratory animals. Protocol #20-0384.

#### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.72989.sa1>

Author response <https://doi.org/10.7554/eLife.72989.sa2>

## Additional files

#### Supplementary files

- Supplementary file 1. Allele-specific expression.
- Supplementary file 2. Biallelic genes differentially expressed by cross.
- Supplementary file 3. Networks of genes showing parent-of-origin allele-specific expression interacting with biallelic genes that are differentially expressed by cross.
- Supplementary file 4. Over-representation input/output.
- Supplementary file 5. F<sub>1</sub> Phenotype correlation.
- Supplementary file 6. Epistasis results.
- Supplementary file 7. Cell cluster markers.
- Supplementary file 8. Alignment summaries.
- Supplementary file 9. List of imprinted genes queried.
- Supplementary file 10. Genome annotations.
- Transparent reporting form

#### Data availability

Sequencing data are available through the NCBI-SRA under accession code PRJNA753198.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Lawson HA	2021	LG/J and SM/J hybrid RNAseq	<a href="https://www.ncbi.nlm.nih.gov/bioproject/PRJNA753198/">https://www.ncbi.nlm.nih.gov/bioproject/PRJNA753198/</a>	NCBI BioProject, PRJNA753198

The following previously published dataset was used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Burl RB	2018	Single-cell RNA-sequencing of white adipose tissue stromal cells during CL-induced adipogenesis	<a href="https://www.ebi.ac.uk/ena/browser/view/PRJNA470640?show=reads">https://www.ebi.ac.uk/ena/browser/view/PRJNA470640?show=reads</a>	European Nucleotide Archive, PRJNA470640

## References

Akune T, Ohba S, Kamekura S, Yamaguchi M, Chung U-I, Kubota N, Terauchi Y, Harada Y, Azuma Y, Nakamura K, Kadowaki T, Kawaguchi H. 2004. PPARgamma insufficiency enhances osteogenesis through osteoblast

- formation from bone marrow progenitors. *The Journal of Clinical Investigation* **113**:846–855. DOI: <https://doi.org/10.1172/JCI19900>, PMID: 15067317
- Ali AA**, Weinstein RS, Stewart SA, Parfitt AM, Manolagas SC, Jilka RL. 2005. Rosiglitazone causes bone loss in mice by suppressing osteoblast differentiation and bone formation. *Endocrinology* **146**:1226–1235. DOI: <https://doi.org/10.1210/en.2004-0735>, PMID: 15591153
- An JJ**, Han DH, Kim DM, Kim SH, Rhee Y, Lee EJ, Lim SK. 2007. Expression and regulation of osteoprotegerin in adipose tissue. *Yonsei Medical Journal* **48**:765–772. DOI: <https://doi.org/10.3349/ymj.2007.48.5.765>, PMID: 17963332
- Aubin JE**. 1992. Osteoclast adhesion and resorption: the role of podosomes. *Journal of Bone and Mineral Research* **7**:365–368. DOI: <https://doi.org/10.1002/jbmr.5650070402>, PMID: 1609624
- Babak T**, DeVeale B, Tsang EK, Zhou Y, Li X, Smith KS, Kukurba KR, Zhang R, Li JB, van der Kooy D, Montgomery SB, Fraser HB. 2015. Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nature Genetics* **47**:544–549. DOI: <https://doi.org/10.1038/ng.3274>, PMID: 25848752
- Berridge MJ**. 2016. The Inositol Trisphosphate/Calcium Signaling Pathway in Health and Disease. *Physiological Reviews* **96**:1261–1296. DOI: <https://doi.org/10.1152/physrev.00006.2016>, PMID: 27512009
- Bigdeli N**, de Peppo GM, Lennerås M, Sjövall P, Lindahl A, Hyllner J, Karlsson C. 2010. Superior osteogenic capacity of human embryonic stem cells adapted to matrix-free growth compared to human mesenchymal stem cells. *Tissue Engineering. Part A* **16**:3427–3440. DOI: <https://doi.org/10.1089/ten.tea.2010.0112>, PMID: 20653416
- Burg JS**, Espenshade PJ. 2011. Regulation of HMG-CoA reductase in mammals and yeast. *Progress in Lipid Research* **50**:403–410. DOI: <https://doi.org/10.1016/j.plipres.2011.07.002>, PMID: 21801748
- Burl RB**, Ramseyer VD, Rondini EA, Pique-Regi R, Lee YH, Granneman JG. 2018. Deconstructing Adipogenesis Induced by  $\beta$ 3-Adrenergic Receptor Activation with Single-Cell Expression Profiling. *Cell Metabolism* **28**:300–309. DOI: <https://doi.org/10.1016/j.cmet.2018.05.025>, PMID: 29937373
- Carson C**, Macias-Velasco JF, Gunawardana S, Miranda MA, Oyama S, St Pierre CL, Schmidt H, Wayhart JP, Hanson HA. 2020. Brown Adipose Expansion and Remission of Glycemic Dysfunction in Obese SM/J Mice. *Cell Reports* **33**:108237. DOI: <https://doi.org/10.1016/j.celrep.2020.108237>, PMID: 33027654
- Cheverud JM**, Lawson HA, Fawcett GL, Wang B, Pletscher LS, R Fox A, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, Semenkovich CF. 2011. Diet-dependent genetic and genomic imprinting effects on obesity in mice. *Obesity (Silver Spring, Md.)* **19**:160–170. DOI: <https://doi.org/10.1038/oby.2010.141>, PMID: 20539295
- Cho SW**, Yang J-Y, Her SJ, Choi HJ, Jung JY, Sun HJ, An JH, Cho HY, Kim SW, Park KS, Kim SY, Baek W-Y, Kim J-E, Yim M, Shin CS. 2011. Osteoblast-targeted overexpression of PPAR $\gamma$  inhibited bone mass gain in male mice and accelerated ovariectomy-induced bone loss in female mice. *Journal of Bone and Mineral Research* **26**:1939–1952. DOI: <https://doi.org/10.1002/jbmr.366>, PMID: 21351141
- 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 (Oxford, England)* **29**:15–21. DOI: <https://doi.org/10.1093/bioinformatics/bts635>, PMID: 23104886
- Erol M**, Bostan Gayret O, Tekin Nacaroglu H, Yigit O, Zengi O, Salih Akkurt M, Tasdemir M. 2016. Association of Osteoprotegerin with Obesity, Insulin Resistance and Non-Alcoholic Fatty Liver Disease in Children. *Iranian Red Crescent Medical Journal* **18**:41873. DOI: <https://doi.org/10.5812/ircmj.41873>, PMID: 28203453
- Finak G**, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Plic M, Linsley PS, Gottardo R. 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology* **16**:1–13. DOI: <https://doi.org/10.1186/s13059-015-0844-5>, PMID: 26653891
- Font-Clos F**, Zapperi S, La Porta CAM. 2017. Integrative analysis of pathway deregulation in obesity. *NPJ Systems Biology and Applications* **3**:18. DOI: <https://doi.org/10.1038/s41540-017-0018-z>, PMID: 28685099
- Fowden AL**, Moore T. 2012. Maternal-fetal resource allocation: co-operation and conflict. *Placenta* **33** Suppl 2:e11–e15. DOI: <https://doi.org/10.1016/j.placenta.2012.05.002>, PMID: 22652046
- Franck N**, Gummesson A, Jernäs M, Glad C, Svensson PA, Guillot G, Rudemo M, Nyström FH, Carlsson LMS, Olsson B. 2011. Identification of adipocyte genes regulated by caloric intake. *The Journal of Clinical Endocrinology and Metabolism* **96**:E413–E418. DOI: <https://doi.org/10.1210/jc.2009-2534>, PMID: 21047925
- Hager R**, Cheverud JM, Wolf JB. 2008. Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. *Genetics* **178**:1755–1762. DOI: <https://doi.org/10.1534/genetics.107.080697>, PMID: 18245362
- Haig D**. 1997. Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proceedings. Biological Sciences* **264**:1657–1662. DOI: <https://doi.org/10.1098/rspb.1997.0230>, PMID: 9404029
- Holecki M**, Zahorska-Markiewicz B, Janowska J, Nieszporek T, Wojaczyńska-Stanek K, Zak-Gołąb A, Wiecek A. 2007. The influence of weight loss on serum osteoprotegerin concentration in obese perimenopausal women. *Obesity (Silver Spring, Md.)* **15**:1925–1929. DOI: <https://doi.org/10.1038/oby.2007.229>, PMID: 17712108
- Ikeda K**, Yamada T. 2020. UCP1 Dependent and Independent Thermogenesis in Brown and Beige Adipocytes. *Frontiers in Endocrinology* **11**:498. DOI: <https://doi.org/10.3389/fendo.2020.00498>, PMID: 32849287
- Imoto-Tsubakimoto H**, Takahashi T, Ueyama T, Ogata T, Adachi A, Nakanishi N, Mizushima K, Naito Y, Matsubara H. 2013. Serglycin is a novel adipocytokine highly expressed in epicardial adipose tissue. *Biochemical and Biophysical Research Communications* **432**:105–110. DOI: <https://doi.org/10.1016/j.bbrc.2013.01.078>, PMID: 23376071
- Jirtle RL**. 2012. No Title. Geneimprint. <http://www.geneimprint.com/site/genes-by-species>

- Jo Y, Debose-Boyd RA. 2010. Control of cholesterol synthesis through regulated ER-associated degradation of HMG CoA reductase. *Critical Reviews in Biochemistry and Molecular Biology* **45**:185–198. DOI: <https://doi.org/10.3109/10409238.2010.485605>, PMID: 20482385
- Kang JW, Choi Y, Park JH, Kim JS, Park KD, Baek DH, Seong SK, Choi KS, Lim SY, Kim HS. 2008. The effects of cyclin-dependent kinase inhibitors on adipogenic differentiation of human mesenchymal stem cells. *Biochemical and Biophysical Research Communications* **366**:624–630. DOI: <https://doi.org/10.1016/j.bbrc.2007.11.080>, PMID: 18047835
- Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. 2019. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *American Journal of Human Genetics* **104**:65–75. DOI: <https://doi.org/10.1016/j.ajhg.2018.11.008>, PMID: 30595370
- Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, Jonasdottir A, Sigurdsson A, Kristinsson KT, Jonasdottir A, Frigge ML, Gylfason A, Olason PI, Gudjonsson SA, Sverrisson S, Stacey SN, Sigurgeirsson B, Benediktsdottir KR, Sigurdsson H, Jonsson T, et al. 2009. Parental origin of sequence variants associated with complex diseases. *Nature* **462**:868–874. DOI: <https://doi.org/10.1038/nature08625>, PMID: 20016592
- Lawson HA, Zelle KM, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, Semenkovich CF, Cheverud JM. 2010. Genetic, epigenetic, and gene-by-diet interaction effects underlie variation in serum lipids in a LG/JxSM/J murine model. *Journal of Lipid Research* **51**:2976–2984. DOI: <https://doi.org/10.1194/jlr.M006957>, PMID: 20601649
- 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 Genetics* **7**:e1002256. DOI: <https://doi.org/10.1371/journal.pgen.1002256>, PMID: 21931559
- Lawson HA, Lee A, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf JB, Semenkovich CF, Cheverud JM. 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. *Mammalian Genome* **22**:197–208. DOI: <https://doi.org/10.1007/s00335-010-9313-3>, PMID: 21210123
- Lawson HA, Cheverud JM, Wolf JB. 2013. Genomic imprinting and parent-of-origin effects on complex traits. *Nature Reviews. Genetics* **14**:609–617. DOI: <https://doi.org/10.1038/nrg3543>, PMID: 23917626
- Lawson HA. 2019. LawsonLab-WUSM/POE\_Epistasis. swh:1:rev:b39046ce35f53e0c3f15bcdefa122c274aee48b7. Software Heritage. [https://archive.softwareheritage.org/swh:1:dir:3e0c4952b5f93241102e0b8a1c030ac0ff41ce93;origin=https://github.com/LawsonLab-WUSM/POE\\_Epistasis;visit=swh:1:snp:532c07d4bba72b337aa77fcc9cd90553ecde1426;anchor=swh:1:rev:b39046ce35f53e0c3f15bcdefa122c274aee48b7](https://archive.softwareheritage.org/swh:1:dir:3e0c4952b5f93241102e0b8a1c030ac0ff41ce93;origin=https://github.com/LawsonLab-WUSM/POE_Epistasis;visit=swh:1:snp:532c07d4bba72b337aa77fcc9cd90553ecde1426;anchor=swh:1:rev:b39046ce35f53e0c3f15bcdefa122c274aee48b7)
- Lemire JM, Chan CK, Bressler S, Miller J, LeBaron RG, Wight TN. 2007. Interleukin-1beta selectively decreases the synthesis of versican by arterial smooth muscle cells. *Journal of Cellular Biochemistry* **101**:753–766. DOI: <https://doi.org/10.1002/jcb.21235>, PMID: 17226775
- Li X, Thomason PA, Withers DJ, Scott J. 2010. Bio-informatics analysis of a gene co-expression module in adipose tissue containing the diet-responsive gene Nnat. *BMC Systems Biology* **4**:1–11. DOI: <https://doi.org/10.1186/1752-0509-4-175>, PMID: 21187013
- Lin HH, Bell E, Uwanogho D, Perfect LW, Noristani H, Bates TJD, Snetkov V, Price J, Sun YM. 2010. Neuronatin promotes neural lineage in ESCs via Ca(2+) signaling. *Stem Cells (Dayton, Ohio)* **28**:1950–1960. DOI: <https://doi.org/10.1002/stem.530>, PMID: 20872847
- Luecken MD, Theis FJ. 2019. Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology* **15**:e8746. DOI: <https://doi.org/10.15252/msb.20188746>, PMID: 31217225
- Lutz SM, Hokanson JE. 2015. Mediation analysis in genome-wide association studies: current perspectives. *Open Access Bioinformatics* **7**:S63643. DOI: <https://doi.org/10.2147/OAB.S63643>
- Matsuo FS, Cavalcanti de Araújo PH, Mota RF, Carvalho AJR, Santos de Queiroz M, Baldo de Almeida B, Ferreira KC, Metzner RJM, Ferrari GD, Alberici LC, Osako MK. 2020. RANKL induces beige adipocyte differentiation in preadipocytes. *American Journal of Physiology. Endocrinology and Metabolism* **318**:E866–E877. DOI: <https://doi.org/10.1152/ajpendo.00397.2019>, PMID: 32315212
- McDonald LJ, Mamrack MD. 1995. Phosphoinositide hydrolysis by phospholipase C modulated by multivalent cations La(3+), Al(3+), neomycin, polyamines, and melittin. *Journal of Lipid Mediators and Cell Signalling* **11**:81–91. DOI: [https://doi.org/10.1016/0929-7855\(94\)00029-c](https://doi.org/10.1016/0929-7855(94)00029-c), PMID: 7728419
- Millership SJ, Da Silva Xavier G, Choudhury AI, Bertazzo S, Chabosseau P, Pedroni SM, Irvine EE, Montoya A, Faull P, Taylor WR, Kerr-Conte J, Pattou F, Ferrer J, Christian M, John RM, Latreille M, Liu M, Rutter GA, Scott J, Withers DJ. 2018. Neuronatin regulates pancreatic  $\beta$  cell insulin content and secretion. *The Journal of Clinical Investigation* **128**:3369–3381. DOI: <https://doi.org/10.1172/JCI120115>, PMID: 29864031
- Miranda MA, Carson C, St. Pierre CL, Macias-Velasco JF, Hughes JW, Kunzmann M, Schmidt H, Wayhart JP, Lawson HA. 2020. Spontaneous restoration of functional  $\beta$ -cell mass in obese SM/J mice. *Physiological Reports* **8**:1–12. DOI: <https://doi.org/10.14814/phy2.14573>
- Mochizuki A, Takeda Y, Iwasa Y. 1996. The evolution of genomic imprinting. *Genetics* **144**:1283–1295. DOI: <https://doi.org/10.1093/genetics/144.3.1283>, PMID: 8913768
- Mott R, Yuan W, Kaisaki P, Gan X, Cleak J, Edwards A, Baud A, Flint J. 2014. The architecture of parent-of-origin effects in mice. *Cell* **156**:332–342. DOI: <https://doi.org/10.1016/j.cell.2013.11.043>, PMID: 24439386
- Mozaffari SV, DeCara JM, Shah SJ, Sidore C, Fiorillo E, Cucca F, Lang RM, Nicolae DL, Ober C. 2019. Parent-of-origin effects on quantitative phenotypes in a large Hutterite pedigree. *Communications Biology* **2**:28. DOI: <https://doi.org/10.1038/s42003-018-0267-4>

- Nakagawa N**, Kinoshita M, Yamaguchi K, Shima N, Yasuda H, Yano K, Morinaga T, Higashio K. 1998. RANK is the essential signaling receptor for osteoclast differentiation factor in osteoclastogenesis. *Biochemical and Biophysical Research Communications* **253**:395–400. DOI: <https://doi.org/10.1006/bbrc.1998.9788>, PMID: 9878548
- Nakamura Y**, Hamada Y, Fujiwara T, Enomoto H, Hiroe T, Tanaka S, Nose M, Nakahara M, Yoshida N, Takenawa T, Fukami K. 2005. Phospholipase C-delta1 and -delta3 are essential in the trophoblast for placental development. *Molecular and Cellular Biology* **25**:10979–10988. DOI: <https://doi.org/10.1128/MCB.25.24.10979-10988.2005>, PMID: 16314520
- 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: <https://doi.org/10.1186/s12864-015-1592-3>, PMID: 26016481
- Parker-Katiraei L**, Carson AR, Yamada T, Arnaud P, Feil R, Abu-Amero SN, Moore GE, Kaneda M, Perry GH, Stone AC, Lee C, Meguro-Horike M, Sasaki H, Kobayashi K, Nakabayashi K, Scherer SW. 2007. Identification of the imprinted KLF14 transcription factor undergoing human-specific accelerated evolution. *PLOS Genetics* **3**:0030065. DOI: <https://doi.org/10.1371/journal.pgen.0030065>, PMID: 17480121
- Patten MM**, Ross L, Curley JP, Queller DC, Bonduriansky R, Wolf JB. 2014. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity* **113**:119–128. DOI: <https://doi.org/10.1038/hdy.2014.29>, PMID: 24755983
- Pelled D**, Lloyd-Evans E, Riebeling C, Jeyakumar M, Platt FM, Futerman AH. 2003. Inhibition of calcium uptake via the sarco/endoplasmic reticulum Ca<sup>2+</sup>-ATPase in a mouse model of Sandhoff disease and prevention by treatment with N-butyldeoxynojirimycin. *The Journal of Biological Chemistry* **278**:29496–29501. DOI: <https://doi.org/10.1074/jbc.M302964200>, PMID: 12756243
- Pierce BL**, Tong L, Chen LS, Rahaman R, Argos M, Jasmine F, Roy S, Paul-Brutus R, Westra HJ, Franke L, Esko T, Zaman R, Islam T, Rahman M, Baron JA, Kibriya MG, Ahsan H. 2014. Mediation analysis demonstrates that trans-eQTLs are often explained by cis-mediation: A genome-wide analysis among 1,800 South Asians. *PLOS Genetics* **10**:e1004818. DOI: <https://doi.org/10.1371/journal.pgen.1004818>, PMID: 25474530
- Prockop DJ**. 1997. Marrow stromal cells as stem cells for nonhematopoietic tissues. *Science (New York, N.Y.)* **276**:71–74. DOI: <https://doi.org/10.1126/science.276.5309.71>, PMID: 9082988
- R Development Core Team**. 2013. R: A Language and Environment for Statistical Computing. Vienna, Austria. R Foundation for Statistical Computing. <http://www.r-project.org>
- Renner SW**, Walker LM, Forsberg LJ, Sexton JZ, Brenman JE. 2017. Carbonic anhydrase III (Car3) is not required for fatty acid synthesis and does not protect against high-fat diet induced obesity in mice. *PLOS ONE* **12**:e0176502. DOI: <https://doi.org/10.1371/journal.pone.0176502>, PMID: 28437447
- Robinson MD**, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**:139–140. DOI: <https://doi.org/10.1093/bioinformatics/btp616>, PMID: 19910308
- Robinson MD**, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **11**:1–9. DOI: <https://doi.org/10.1186/gb-2010-11-3-r25>, PMID: 20196867
- Rodríguez JP**, Astudillo P, Ríos S, Pino AM. 2008. Involvement of adipogenic potential of human bone marrow mesenchymal stem cells (MSCs) in osteoporosis. *Current Stem Cell Research & Therapy* **3**:208–218. DOI: <https://doi.org/10.2174/157488808785740325>, PMID: 18782003
- Rosen CJ**, Bouxsein ML. 2006. Mechanisms of disease: is osteoporosis the obesity of bone? *Nature Clinical Practice. Rheumatology* **2**:35–43. DOI: <https://doi.org/10.1038/ncprheum0070>, PMID: 16932650
- Schick BP**, Gradowski JF, San Antonio JD. 2001. Synthesis, secretion, and subcellular localization of serglycin proteoglycan in human endothelial cells. *Blood* **97**:449–458. DOI: <https://doi.org/10.1182/blood.v97.2.449>, PMID: 11154222
- Shan N**, Wang Z, Hou L. 2019. Identification of trans-eQTLs using mediation analysis with multiple mediators. *BMC Bioinformatics* **20**:14–16. DOI: <https://doi.org/10.1186/s12859-019-2651-6>, PMID: 30925861
- Small KS**, Hedman AK, Grundberg E, Nica AC, Thorleifsson G, Kong A, Thorsteindottir U, Shin S-Y, Richards HB, GIANT Consortium, MAGIC Investigators, DIAGRAM Consortium, Soranzo N, Ahmadi KR, Lindgren CM, Stefansson K, Dermitzakis ET, Deloukas P, Spector TD, McCarthy MI, et al. 2011. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nature Genetics* **43**:561–564. DOI: <https://doi.org/10.1038/ng.833>, PMID: 21572415
- Soh UJK**, Dores MR, Chen B, Trejo J. 2010. Signal transduction by protease-activated receptors. *British Journal of Pharmacology* **160**:191–203. DOI: <https://doi.org/10.1111/j.1476-5381.2010.00705.x>, PMID: 20423334
- Stanton LW**, Ponte PA, Coleman RT, Snyder MA. 1991. Expression of CA III in rodent models of obesity. *Molecular Endocrinology (Baltimore, Md.)* **5**:860–866. DOI: <https://doi.org/10.1210/mend-5-6-860>, PMID: 1922100
- Stuart T**, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**:1888–1902. DOI: <https://doi.org/10.1016/j.cell.2019.05.031>
- Teslovich TM**, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, Johansen CT, Fouchier SW, Isaacs A, Peloso GM, Barbalic M, Ricketts SL, Bis JC, Aulchenko YS, Thorleifsson G, Feitosa MF, et al. 2010. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**:707–713. DOI: <https://doi.org/10.1038/nature09270>, PMID: 20686565
- Thatcher JD**. 2010. The inositol trisphosphate (IP<sub>3</sub>) signal transduction pathway. *Science Signaling* **3**:tr3. DOI: <https://doi.org/10.1126/scisignal.3119tr3>, PMID: 20424267

- Turner RT**, Martin SA, Iwaniec UT. 2018. Metabolic Coupling Between Bone Marrow Adipose Tissue and Hematopoiesis. *Current Osteoporosis Reports* **16**:95–104. DOI: <https://doi.org/10.1007/s11914-018-0422-3>, PMID: 29492879
- Van de Pette M**, Tunster SJ, John RM. 2018. Loss of Imprinting of *Cdkn1c* Protects against Age and Diet-Induced Obesity. *International Journal of Molecular Sciences* **19**:E2734. DOI: <https://doi.org/10.3390/ijms19092734>, PMID: 30213134
- Voight BF**, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, et al. 2010. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature Genetics* **42**:579–589. DOI: <https://doi.org/10.1038/ng.609>, PMID: 20581827
- Wilkins JF**, Haig D. 2003. What good is genomic imprinting: the function of parent-specific gene expression. *Nature Reviews. Genetics* **4**:359–368. DOI: <https://doi.org/10.1038/nrg1062>, PMID: 12728278
- Wolf JB**, Cheverud JM. 2009. A framework for detecting and characterizing genetic background-dependent imprinting effects. *Mammalian Genome* **20**:681–698. DOI: <https://doi.org/10.1007/s00335-009-9209-2>, PMID: 19657694
- Yee TW**. 2010. The VGAM Package for Categorical Data Analysis. *Journal of Statistical Software* **32**:i10. DOI: <https://doi.org/10.18637/jss.v032.i10>
- Yeh Y-S**, Jheng H-F, Iwase M, Kim M, Mohri S, Kwon J, Kawarasaki S, Li Y, Takahashi H, Ara T, Nomura W, Kawada T, Goto T. 2018. The Mevalonate Pathway Is Indispensable for Adipocyte Survival. *iScience* **9**:175–191. DOI: <https://doi.org/10.1016/j.isci.2018.10.019>, PMID: 30396151
- Young HS**, Won HK, Moon C, Yun HH, Eun SY, Joo HL, Joo SC, Song J, Jung MH. 2005. Ectopic expression of Neuronatin potentiates adipogenesis through enhanced phosphorylation of cAMP-response element-binding protein in 3T3-L1 cells. *Biochemical and Biophysical Research Communications* **337**:481–489. DOI: <https://doi.org/10.1016/j.bbrc.2005.09.078>
- Zaky DS**, Ali AA, Abd-Elraheem SE, Abdel-Moniem SH. 2022. Circulating Osteoprotegerin Level in Relation to Obesity in Middle Aged Females. Sapub.
- Zappia L**, Oshlack A. 2018. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *GigaScience* **7**:1–9. DOI: <https://doi.org/10.1093/gigascience/giy083>, PMID: 30010766
- Zeng Y**, Amador C, Xia C, Marioni R, Sproul D, Walker RM, Morris SW, Bretherick A, Canela-Xandri O, Boutin TS, Clark DW, Campbell A, Rawlik K, Hayward C, Nagy R, Tenesa A, Porteous DJ, Wilson JF, Deary IJ, Evans KL, et al. 2019. Parent of origin genetic effects on methylation in humans are common and influence complex trait variation. *Nature Communications* **10**:1383. DOI: <https://doi.org/10.1038/s41467-019-09301-y>, PMID: 30918249
- Zernichow L**, Åbrink M, Hallgren J, Grujic M, Pejler G, Kolset SO. 2006. Serglycin Is the Major Secreted Proteoglycan in Macrophages and Has a Role in the Regulation of Macrophage Tumor Necrosis Factor- $\alpha$  Secretion in Response to Lipopolysaccharide. *Journal of Biological Chemistry* **281**:26792–26801. DOI: <https://doi.org/10.1074/jbc.M512889200>, PMID: 16807245
- Zhang B**, Kirov S, Snoddy J. 2005. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research* **33**:W741–W748. DOI: <https://doi.org/10.1093/nar/gki475>, PMID: 15980575
- Zheng GXY**, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, et al. 2017. Massively parallel digital transcriptional profiling of single cells. *Nature Communications* **8**:1–12. DOI: <https://doi.org/10.1038/ncomms14049>, PMID: 28091601