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

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

JF Macias-Velasco: juanfmacias@wustl.edu
CL St. Pierre: stpierrec@wustl.edu
JP Wayhart: jwayhart@genetics.wustl.edu
L Yin: liyin@wustl.edu
L Spears: larry.spears@stemcell.com
MA Miranda: mario.miranda@wustl.edu
C Carson: caryn.carson@wustl.edu
K Funai: kfunai@health.utah.edu
JM Cheverud: jcheverud@luc.edu
CF Semenkovich: csemenko@wustl.edu
HA Lawson: hlawson@genetics.wustl.edu

[1] Department of Genetics, Washington University School of Medicine, 660 South Euclid Ave, Saint Louis, MO, USA
[2] Department of Medicine, Washington University School of Medicine, 660 South Euclid Ave, Saint Louis, MO, USA
[3] Diabetes and Metabolism Research Center, University of Utah, 15 N 2030 E, Salt Lake City, UT, USA
[4] Department of Biology, Loyola University, 1032 W Sheridan Rd, Chicago IL, USA

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

Running Title: Parent-of-origin effects on metabolic traits

Keywords: Parent-of-origin effect, epistasis, adipose, mouse, RNA sequencing, metabolism, complex traits

ABSTRACT

Parent-of-origin effects are unexpectedly common in complex traits, including metabolic and neurological diseases. 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_{16} 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_0, F_1, F_2, F_{16}) 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 3 imprinted and 6 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 2 genes, *Nnat* and *F2r*, whose interaction associates with serum glucose levels across generations in high fat-fed females. Single-cell RNAseq reveals that *Nnat* and *F2r* are negatively correlated in pre-adipocytes along an adipogenic trajectory, a result that is consistent with our observations in bulk white adipose tissue.

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. 2013a). 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, 2011b, 2011a). 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 identify 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* and *F2r* are negatively correlated 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 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, 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.

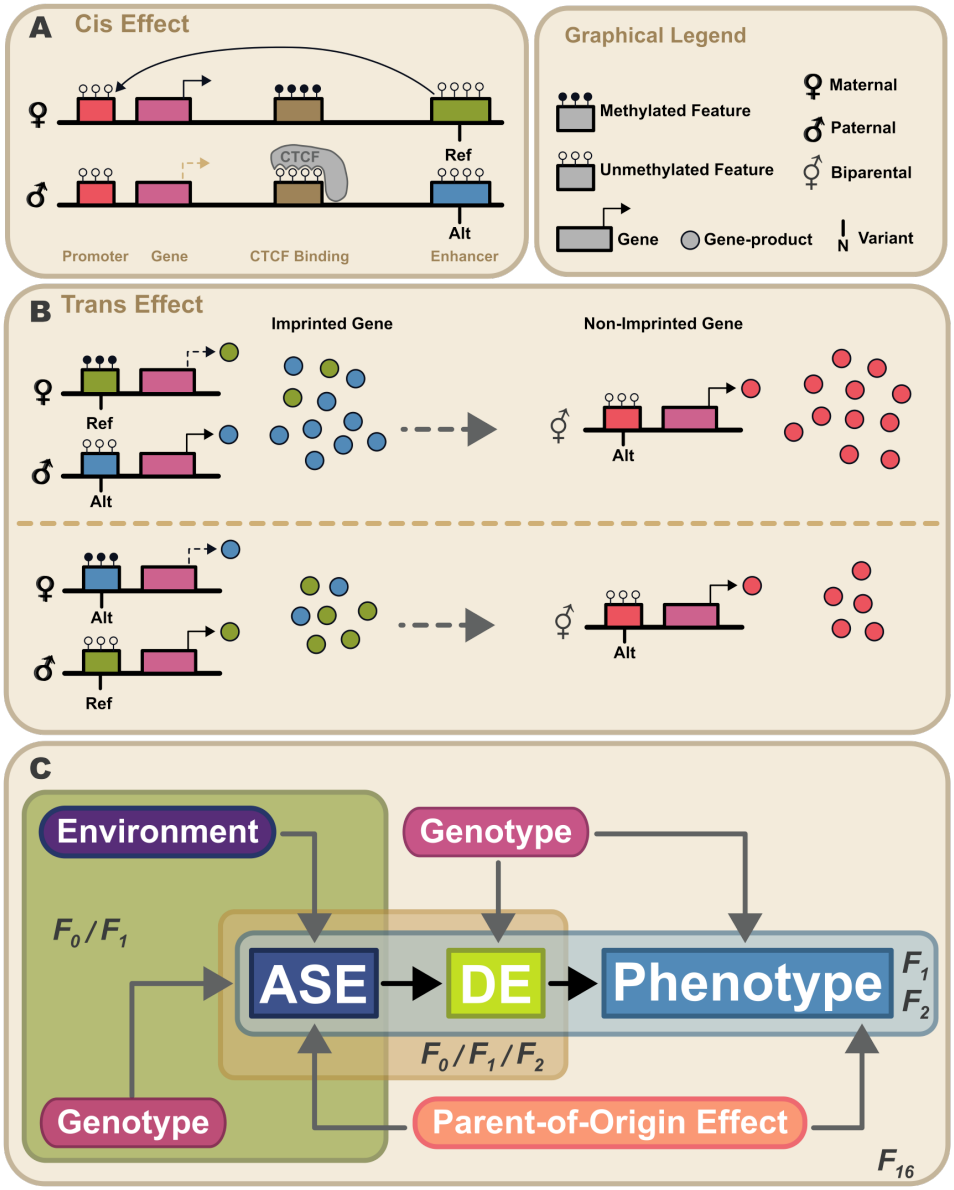

Figure 1: Proposed model for propagation of parent-of-origin effects through gene-gene interactions. Parent of origin effects should be partitioned into *cis* mechanisms and *trans* mechanisms **A.** An example of a *cis* parent-of-origin effect is a system with three regulatory elements: promoter, insulator, 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, 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).

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**; **Supplemental Table 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**; **Supplemental Table 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 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**)(**Supplemental Table 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 1C), are both canonically imprinted and differentially expressed by cross (**Supplemental Table 1**).

Functional over representation analysis was performed and seven terms were significantly overrepresented at an FDR<=0.05 (**Figure 2C**) (Zhang et al. 2005). Enriched terms suggest this network plays a role in signaling and genetic imprinting (**Supplemental Table 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 (**Supplemental Table 5**).

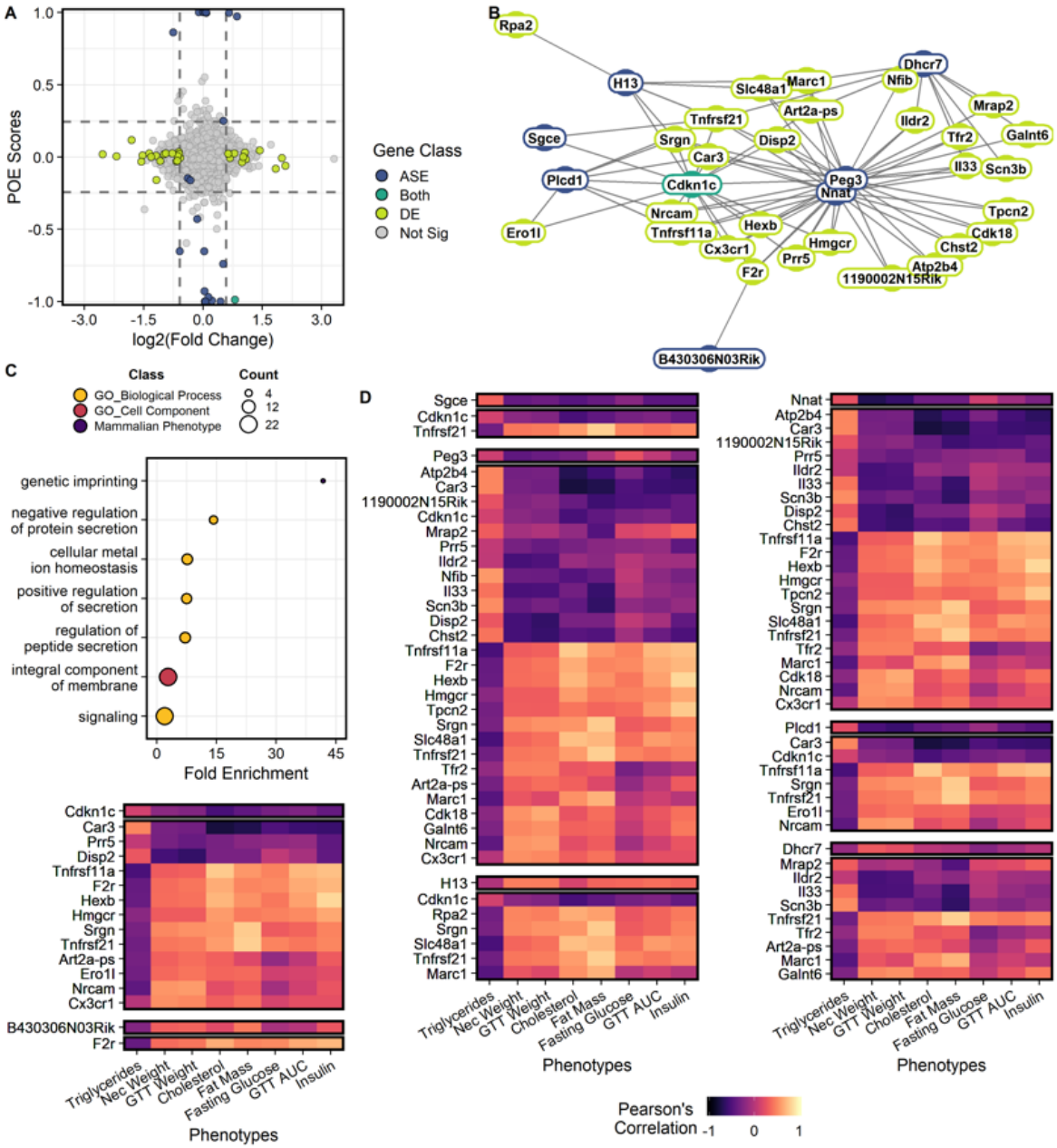

Figure 2: Genes showing parent-of-origin effects at the allele specific and/or total expression levels covary with each other and with metabolic traits. A. Mean POE score across contexts. Effect size of ASE is calculated as the mean allelic bias (L / L+S) of SxL animals minus LxS animals. Effect size 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 POE 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 grey. The two genes showing significant ASE but falling short of POE score requirements are a case of context dependent bipolar POE scores (i.e. paternally expressed in one context and maternally expressed in its opposite). **B.** POE network constructed from ASE and DE gene pairs. **C.** Significantly overrepresented ontologies after multiple tests correction in POE 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 POE 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.

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. 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 (FDR≤0.1; **Figure 3A**; **Supplemental Table 6**). These interactions were comprised of three ASE genes showing parent-of-origin (*Cdknlc*, *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 9 genes form a putative diet-responsive network affecting adipogenesis (**Figure 3B**).

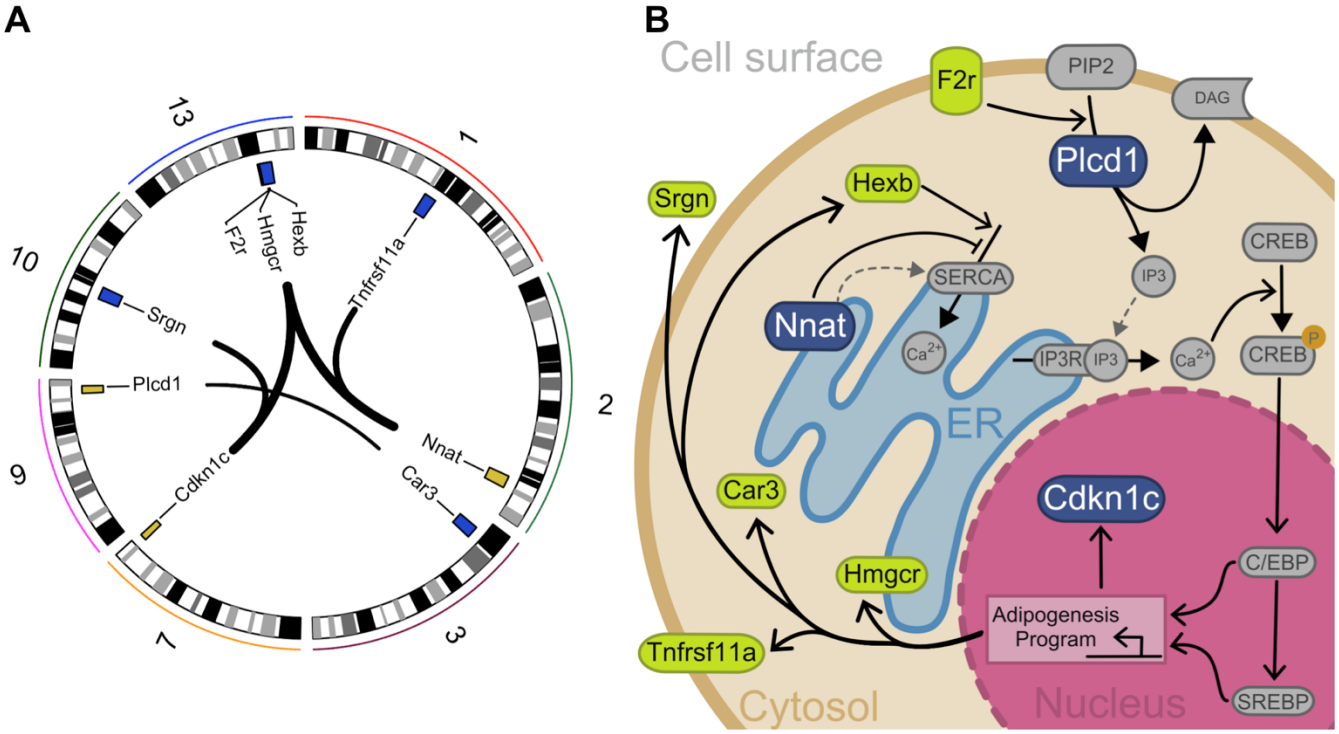

Figure 3: A. There are nine significant imprinting-by-imprinting epistatic ASE/DE/phenotype sets in the F_{16} advanced intercross population. Interactions are shown as lines connecting ASE (yellow) and DE genes (purple). Chromosome number is shown around the plot. **B.** The epistatic POE network is composed 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 network can be broken down into signal potentiation, transduction, and response. *Nnat* (neuronatin) 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 second messenger systems, one of which is mediated by *Plcd1* (Nakamura et al. 2005; McDonald and Mararack 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

257 expression of *Cdkn1c* (cyclin dependent kinase inhibitor 1C), *Hmgcr* (3-hydroxy-3-methylglutaryl-CoA
258 reductase), *Car3* (carbonic anhydrase 3), *Tnfrsf11a* (TNF receptor superfamily member 11a), and *Srgn*
259 (serglycin).

260 *Cdkn1c* is a canonically imprinted maternally expressed gene that inhibits cell proliferation (JW et
261 al. 2008). Increased expression of *Cdkn1c* is protective against diet-induced obesity in mice (Pette et al.
262 2018), and in humans increased caloric intake results in decreased CDKN1C expression (Franck et al.
263 2011). *Hmgcr* is the rate limiting enzyme in cholesterol biosynthesis (JS and PJ 2011; Y and RA 2010)
264 and converts HMG-CoA into mevalonate, which is essential for adipocyte survival (Yeh et al. 2018). *Srgn*
265 is an adipocytokine thought to be part of a feedback loop with *Tnfα* (tumor necrosis factor alpha), mediating paracrine cross-talk between macrophages and adipocytes (JM et al. 2007; H et al. 2013; BP
267 et al. 2001; L et al. 2006). *Srgn* is known to play a role in osteoblast-mediated bone mineralization (Bigdeli
268 et al. 2010), which along with osteoclast-driven bone deconstruction drives bone remodeling (Aubin
269 1992). Osteoblasts share a lineage with adipocytes, and the quantity of osteoblasts is inversely
270 proportional to that of marrow adipose tissue (Rodriguez et al. 2008; Prockop 1997; Ali et al. 2005; Akune
271 et al. 2004; Cho et al. 2011; Rosen and Bouxsein 2006; RT et al. 2018). *Tnfrsf11a* is a cell surface protein
272 that regulates differentiation of osteoclasts (N et al. 1998). Osteoprotegerin (OPG) is a decoy receptor
273 for TNFRSF11A thereby inhibiting osteoclastogenesis and bone resorption (FS et al. 2020). OPG is
274 expressed during differentiation of 3T3L1 adipocytes (An et al. 2007). Expression of OPG is induced by
275 *Tnfα* in 3T3L1 adipocytes and is associated with obesity in humans (M et al. 2007; Erol et al. 2016; Zaky
276 et al.).

277 The exact function of OPG/*Tnfrsf11a* outside of osteoclastogenesis is unknown, but the function
278 of osteoclasts is to break down bone tissue during bone resorption. Bone resorption regulates the level
279 of blood calcium. The bioavailability of calcium in the blood potentially alters ER calcium stores, creating
280 cross-talk between bone cells and white adipose tissue calcium signaling. Osteoclasts break down bone
281 by acidifying mineralized bone, orchestrated by osteoblasts that have become embedded in the matrix
282 they produce (osteocytes). Oxidative stress on osteocytes from the bone acidification process is

283 prevented by *Car3*. *Car3* is an enzyme that catalyzes the conversion of carbonic acid to CO_2 and water.
284 Its expression in white adipose is negatively correlated with and responsive to long term obesity in mice
285 and humans (LW et al. 1991; Font-Clos et al. 2017). *Car3* does not protect against diet induced obesity
286 and is not necessary for fatty acid synthesis (Renner et al. 2017). As such its exact function in adipocytes
287 is unknown.
288
289 *Nnat and F2r covary in white adipose tissue and their interaction associates with variation in basal*
290 *glucose levels across generations*
291 To better understand how these interactions affect phenotype, we focused on the negative
292 correlation of *Nnat* and *F2r* in the above network. In white adipose tissue, *Nnat* expression significantly
293 covaries with *F2r*, a biallelic gene showing significant DE by cross in F_1 high fat-fed females (FDR=0.05).
294 *Nnat* and *F2r* show significant imprinting-by-imprinting epistasis for basal glucose levels in the F_{16}
295 population (FDR=$6.00e^{-16}$; **Figure 4A and B**). To validate gene expression patterns, we combined F_1
296 biological replicates and F_0 high fat-fed female animals (F_1 n=13 and F_0 n=12) and again observe that
297 *F2r* and *Nnat* are each significantly differentially expressed between reciprocal heterozygotes, i.e. by
298 cross (*F2r* p=0.007 and *Nnat* p=0.026; **Figure 4C and D**). Further, the co-expression of *Nnat* and *F2r*
299 also persists in the F_0/F_1 population (p=$3.00e^{-4}$; **Figure 4E**).
300 A limitation of identifying covariation patterns in F_1 and F_0 populations is that all loci are linked.
301 This makes it difficult to determine which ASE genes truly co-express with DE genes. While incorporation
302 of orthogonal F_{16} genotypes and phenotypes helps reduce false discoveries, a population with
303 randomized genetic background for which we have expression data is needed to replicate these results.
304 To that end, F_2 animals were generated and *Nnat* and *F2r* gene expression levels were measured via
305 qPCR (n=14). We found that *F2r* and *Nnat* are significantly co-expressed in high fat-fed female F_2 animals
306 (p= 0.012; **Figure 4H**).
307 *F2r* expression significantly correlates with basal glucose levels in the RNA-sequenced F_1 animals
308 (r=0.514, FDR=0.01; **Supplemental Table 5**). *F2r* expression is also significantly correlated with basal

glucose in the combined F_0/F_1 population (p= 0.005; **Figure 4F**). A trend between *Nnat* expression and basal glucose level is observed but not statistically significant in the F_0/F_1 animals (p= 0.130; **Figure 4G**). Correlation of *F2r*'s and *Nnat*'s individual expression with basal glucose in F_2 mice follows the same pattern as in the F_0/F_1's. Bootstrapping to calculate confidence intervals shows that the correlation differences between F_0/F_1 and F_2 are not significant (**Figure 4I and J; Supplemental Figure 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_2 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).

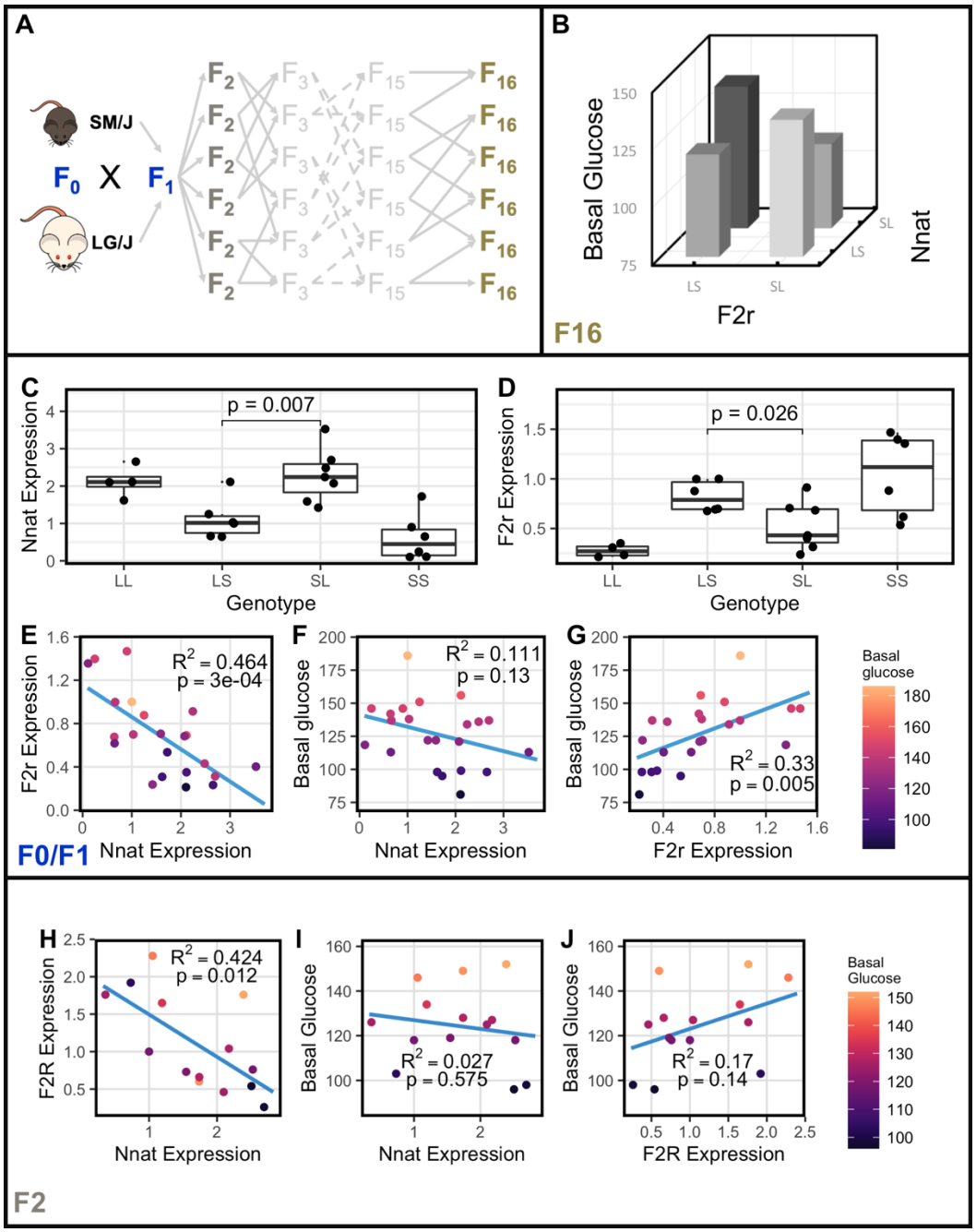

Figure 4. *Nnat* and *F2r* covary across generations. A. Breeding scheme for the F_{16} Advanced Intercross between the LG/J and SM/J inbred strains. **B.** Significant imprinting-by-imprinting epistasis associated with variation in basal glucose. The parent-of-origin effects of *F2r* on basal glucose depend on the parent-of-origin effects at *Nnat*. **C.** Expression of *Nnat* across genotypes in a combined F_0/F_1 population. **D.** Expression of *F2r* across genotypes in a combined F_0/F_1 population. **E.** Significant correlation between *Nnat* and *F2r* expression in the F_0/F_1 mice. **F and G.** Correlations between basal glucose and *Nnat* and *F2r* in the F_0/F_1 mice. **H.** Significant correlation between *Nnat* and *F2r* expression in the F_2 mice. **I and J.** Correlations between basal glucose and *Nnat* and *F2r* are not individually significant in the F_2 mice. However, their interaction significantly correlates with basal glucose in the F_2's (p=0.032), as predicted by our model of epistasis. Alleles are ordered maternal | paternal within the

genotype classes.

Single-cell RNAseq reveals that Nnat and F2r are negatively correlated in pre-adipocytes along an adipogenic trajectory

To determine what cell types express *Nnat* and *F2r* and whether the directionality of the *Nnat* imprinted → *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/6J epididymal adipose tissue (RB 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) (**Supplemental Table 7; Supplemental Figure 2**). 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 and F**). Further there 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.

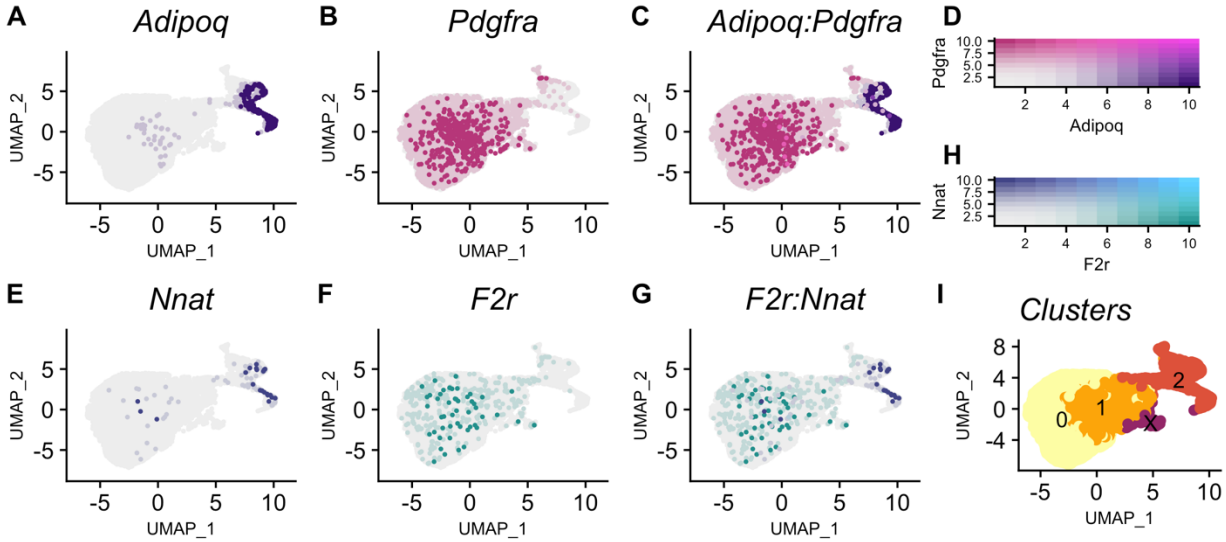

Figure 5: *Nnat* and *F2r* are negatively correlated in pre-adipocytes along an adipogenic trajectory. **A.** *Adipoq* is a marker of adipocytes whose expression (purple) increases along the trajectory. **B.** *Pdgfra* is a marker of mesenchymal stem cells whose expression (pink) decreases along the 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).

In addition to interrogating *Nnat* and *F2r* in single cells along an adipogenic trajectory, we found that 8 of the 9 genes comprising the epistatic POE 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 (**Supplemental Figure 3, Supplemental Table 8**).

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. 2013a; 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-Katiraee 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. 2013b). 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. In the case of *Nnat* and *F2r*, we expect that *F2r* more strongly associates with basal glucose than *Nnat*, which we observe (**Figure 4)**.

We identified a putative network coordinated by interactions between ASE and DE genes, and from the literature we found that this epistatic network is comprised of 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 ß-cells, members of this

network affect secretion of insulin (SJ 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). This hypothesis asserts that the size of progeny in placental mammals has opposing fitness consequences for mothers and fathers. The fitness of fathers, particularly in the case of multi-paternity litters, is improved with larger progeny. 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 to allow one parent to hijack parts of a nutritional environment response pathway driving growth in a direction favorable to maximize their fitness. 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

438 phenotypes from genomic sequence. We focused on metabolic phenotypes in this study, but the patterns
439 we identified may translate to other complex traits where parent-of-origin effects have been implicated.
440
441 **METHODS**
442 *Mouse husbandry and phenotyping*
443 LG/J and SM/J founders (F_0) were obtained from The Jackson Laboratory (Bar Harbor, ME). F_1
444 reciprocal cross animals were generated by mating LG/J mothers with SM/J fathers (LxS) and the inverse
445 (SxL). F_2 reciprocal cross animals were generated by mating LxS mothers with SxL fathers and the
446 inverse. At three weeks of age, animals were weaned into same-sex cages and randomly placed on high-
447 fat (42% kcal from fat; Teklad TD88137) or low-fat (15% kcal from fat; Research Diets D12284) isocaloric
448 diets. Animals were weighed weekly until sacrifice. At 19 weeks of age, body composition was determined
449 by MRI and a glucose tolerance test was performed after a 4 hour fast. At 20 weeks of age, animals were
450 given an overdose of sodium pentobarbital after a 4 hour fast and blood was collected via cardiac
451 puncture. Euthanasia was achieved by cardiac perfusion with phosphate-buffered saline. After cardiac
452 perfusion, the reproductive fat pad was harvested, flash frozen in liquid nitrogen, and stored at -80°C.
453
454 *Genomes and annotations*
455 LG/J and SM/J indels and SNVs were leveraged to construct strain-specific genomes using the
456 GRC38.72-mm10 reference as a template (Nikolskiy et al. 2015). This was done by replacing reference
457 bases with alternative (LG/J | SM/J) bases using custom python scripts. Ensembl R72 annotations were
458 adjusted for indel-induced indexing differences for both genomes.
459
460 *RNA sequencing*
461 Total RNA was isolated from adipose and hypothalamus tissues using the RNeasy Lipid Tissue
462 Kit (QIAgen) and from liver using TRIzol (n = 32, 4 animals per sex/diet/cross cohort). RNA
463 concentration was measured via NanoDrop and RNA quality/integrity was assessed with a BioAnalyzer
464 (Agilent). RNA-Seq libraries were constructed using the RiboZero kit (Illumina) from total RNA samples
465 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 2x100 paired end reads on an Illumina HiSeq 4000. 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_{bias} values using the VGAM package (Yee 2010). The shape parameters (α, β) of beta-binomial distributions were estimated and used to calculate dispersion (ρ). Dispersion values less than 0.05 indicate our libraries are sufficiently complex (**Supplemental Figure 4**).

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

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 **Supplemental Table 9 and Supplemental Figure 5**.

For each gene in each individual, allelic bias (L_{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_{bias} is responsive to cross and the interaction of cross with some combination of sex and diet:

$$Model \begin{cases} if\ each\ Cross\ context\ has \geq 2\ samples,\ Lbias \sim Cross \\ if\ each\ Cross{:}Sex\ context\ has \geq 2\ samples,\ Lbias \sim Cross + Cross{:}Sex \\ if\ each\ Cross{:}Diet\ context\ has \geq 2\ samples,\ Lbias \sim Cross + Cross{:}Diet \\ if\ each\ context\ has \geq 2\ samples,\ Lbias \sim Cross + Cross{:}Sex + Cross{:}Diet + Cross{:}Sex{:}Diet \end{cases}$$

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 (**Supplemental Figure 6**). False discovery rates were estimated for a given significance threshold as the proportion of significant tests under the permutated 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 ≤ 0.1 was considered significant (**Supplemental Table 1**, **Supplemental Figure 7**).

To determine the parental direction and size of expression biases, a parent-of-origin effect POE score was calculated as the difference in mean L_{bias} between reciprocal crosses (LxS or SxL). POE scores range from completely maternally-expressed (-1), to biallelic (0), to completely paternally-expressed (+1). POE 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 (**Supplemental Figure 8**).

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 TMM and 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 (**Supplemental Figure 9**). Genes with a FDR of ≤ 0.1 and a $|fold\ change| \geq 1.5$ were considered significantly differentially expressed by reciprocal cross (**Supplemental Figure 10 and Supplemental Table 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 (**Supplemental Figure 11**). 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 $\widehat{\pi_0}$. This estimate was then used to estimate false discovery rates in the real data. MTE score thresholds were calculated from a critical value of α = 0.01, determined from a null distribution created by permutation (**Supplemental Figure 12**). Connections with an FDR ≤ 0.1 (**Supplemental Table 10**) and MTE below the critical value were considered significant (**Supplemental Figure 13**).

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 ≤ 0.01 was considered significant (**Supplemental Table 4**).

Phenotype correlation

In order to identify which phenotypes might be affected by genes in the POE network, gene expression was correlated with metabolic phenotypes collected for F_1 animals. Phenotypes were log transformed when necessary, as determined by Shapiro Wilkes test to approximate normality (**Supplemental Figure 14**). 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 (**Supplemental Figure 15**). The minimum Pearson's correlation coefficient threshold was set to |0.5|. Connections with an FDR ≤ 0.05 (**Supplemental Table 5**) and MTE below the critical value were considered significant (**Supplemental Figure 16**).

Epistasis testing

The F_{16} LxS advanced intercross population, phenotypes, genotypes, genotypic scores, and QTL mapping methods are described elsewhere (Cheverud et al. 2011; Lawson et al. 2011a, 2011b, 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_{16} 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 (**Supplemental Table 11**). For every F_{16} 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 $\log_{10}$-transformed to approximate normality (**Supplemental Figure 17**). 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_{16} 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 (**Supplemental Figure 18**). False discovery rates were estimated for a given significance threshold as the proportion of significant tests under the permutated null model relative to significant tests under the real data model (**Supplemental Figure 19**). An FDR $\leq$ 0.1 was considered significant. Epistasis was considered significant if the BDE_{IMP} : ASE_{IMP} interaction term met the significance threshold (**Supplemental Table 6**).

Validation of Nnat and F2r expression patterns

Expression patterns of *Nnat* and *F2r* in white adipose were validated by qRT-PCR in LG/J and SM/J mice and in biological replicates of F_1 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 ΔΔCt method. *Nnat* forward primer – CTACCCCAAGAGCTCCCTTT and reverse primer – CAGCTTCTGCAGGGAGTACC. *F2r* forward primer – TGAACCCCCGCTCATTCTTTC and reverse primer – CCAGCAGGACGCTTTCATTTTT. *L32* forward primer – TCCACAATGTCAAGGAGCTG and reverse primer – GGGATTGGTGACTCTGATGG. 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).

$$1.\, Expression \sim Cross \;\in \begin{cases} LL, 0 \\ LS, -1 \\ SL, 1 \\ SS, 0 \end{cases}$$

$$2.\, Expression \sim Paternal\ Allele \;\in \begin{cases} LL, 0 \\ LS, 1 \\ SL, 0 \\ SS, 1 \end{cases}$$

$$3.\, Expression \sim Mitochondrial\ ancestry \;\in \begin{cases} LxS\ x\ SxL, 0 \\ SxL\ x\ LxS, 1 \end{cases}$$

Expression patterns were also validated by qRT-PCR in 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 (RB et al. 2018). Data were processed and aligned to the C57BL/6J reference (mm10) using Cell Ranger (Zheng et al. 2017). Analysis and cell quality control was performed using the Seurat (3.2.2)(T et al. 2019) package in R (3.6.1)(R 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 3,000 features and read counts between 1,000 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 (**Supplemental Figure 20**).

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 (**Supplemental Figure 21**). 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 (**Supplemental Figure 2**).

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 ($\overline{C_\sigma}$). This is done by calculating the standard deviation (σ) of *Adipoq* expression (E) within each cluster (G), referred to as the count variance (C_σ). 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.

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

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

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

The resolution 0.17 was chosen as the lowest resolution where variation is minimized and no longer significantly changes (**Supplemental Figure 22**). 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 ($|\log_2 FoldChange| \geq 0.3$) *OR* a sufficiently large change in the percent of cells expressing the gene in question ($pct.\Delta \geq 0.4$). 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.

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

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 to HAL), and the NIH NHGRI (grant T32-GM007067 to JFMV).The authors declare no conflicts of interest.

Supplemental Materials:

Supplemental Figure 1: Pearson's correlation coefficient confidence intervals.
Supplemental Figure 2. Cell types were defined using canonical markers.

Supplemental Figure 3. Candidate genes are differentially expressed across adipogenic trajectory.
Supplemental Figure 4: RNAseq libraries are sufficiently complex to detect allele specific expression.
Supplemental Figure 5: Number of reads mapped to LG/J x SM/J pseudo-genome.
Supplemental Figure 6: Stable null permutation plots for allele-specific expression.
Supplemental Figure 7: Multiple tests correction of ASE detection.
Supplemental Figure 8: Volcano plots of parent-of-origin dependent allele specific expression.
Supplemental Figure 9: Multiple tests correction of DE detection.
Supplemental Figure 10: Volcano plots of differentially expressed genes.
Supplemental Figure 11: Stable null permutation plots for network pairs.
Supplemental Figure 12: Multiple tests correction of pairwise network construction.
Supplemental Figure 13: Volcano plots of network construction.
Supplemental Figure 14: Example transformation of F_1 phenotypes.
Supplemental Figure 15: Multiple tests correction of phenotype correlations.
Supplemental Figure 16: Volcano plots of phenotypes correlated with POE net gene expression.
Supplemental Figure 17: Example transformation of F_{16} phenotypes.
Supplemental Figure 18: Stable null permutations plot for epistasis.
Supplemental Figure 19: Representative multiple tests correction of imprinting:imprinting epistasis.
Supplemental Figure 20. Single cell quality was controlled.
Supplemental Figure 21. Determining the resolution for clustering.
Supplemental Figure 22. Adipocyte clustering resolution was selected to minimize Adipoq variation.
Supplemental Table1.xlsx : Allele-specific expression
Supplemental Table2.xlsx : Biallelic genes differentially expressed by cross
Supplemental Table3.xlsx : Networks of genes showing parent-of-origin allele-specific expression interacting with biallelic genes that are differentially expressed by cross
Supplemental Table4.xlsx : Over-representation input/output
Supplemental Table5.xlsx : F_1 Phenotype correlation
Supplemental Table6.xlsx : Epistasis results
Supplemental Table7.xlsx : Cell Cluster Markers
Supplemental Table8.xlsx : ASC Type 1 vs Type 2 DE
Supplemental Table9.xlsx : Alignment summaries
Supplemental Table10.xlsx : List of imprinted genes queried
Supplemental Table11.xlsx : Genome annotations

REFERENCES

Akune T, Ohba S, Kamekura S, Yamaguchi M, Chung U Il, Kubota N, Terauchi Y, Harada Y, Azuma Y, Nakamura K, et al. 2004. PPARγ insufficiency enhances osteogenesis through osteoblast formation from bone marrow progenitors. *J Clin Invest* **113**: 846–855.

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. https://academic.oup.com/endo/article/146/3/1226/2500478 (Accessed July 15, 2021).

An J-J, Han D-H, Kim D-M, Kim S-H, Rhee Y, Lee E-J, Lim S-K. 2007. Expression and Regulation of Osteoprotegerin in Adipose Tissue. *Yonsei Med J* **48**: 765. /pmc/articles/PMC2628141/ (Accessed July 15, 2021).

Aubin JE. 1992. Perspectives: Osteoclast adhesion and resorption: The role of podosomes. *J Bone Miner Res* **7**: 365–368. https://asbmr.onlinelibrary.wiley.com/doi/full/10.1002/jbmr.5650070402 (Accessed July 15, 2021).

Babak T, DeVeale B, Tsang EK, Zhou Y, Li X, Smith KS, Kukurba KR, Zhang R, Li JB, van der Kooy D, et al. 2015. Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nat Publ Gr* **47**: 544–549. http://dx.doi.org/10.1038/ng.3274%5Cnpapers3://publication/doi/10.1038/ng.3274.

Berridge MJ. 2016. The Inositol Trisphosphate/Calcium Signaling Pathway in Health and Disease. *https://doi.org/101152/physrev000062016* **96**: 1261–1296. https://journals.physiology.org/doi/abs/10.1152/physrev.00006.2016 (Accessed July 15, 2021).

Bigdeli N, Peppo GM de, 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. *https://home.liebertpub.com/tea* **16**: 3427–3440. https://www.liebertpub.com/doi/abs/10.1089/ten.tea.2010.0112 (Accessed July 15, 2021).

BP S, JF G, JD SA. 2001. Synthesis, secretion, and subcellular localization of serglycin proteoglycan in human endothelial cells. *Blood* **97**: 449–458. https://pubmed.ncbi.nlm.nih.gov/11154222/ (Accessed July 15, 2021).

Carson C, Macias-Velasco JF, Gunawardana S, Miranda MA, Oyama S, St. Pierre CL, Schmidt H, Wayhart JP, Lawson HA. 2020. Brown Adipose Expansion and Remission of Glycemic Dysfunction in Obese SM/J Mice. *Cell Rep* **33**: 108237. https://doi.org/10.1016/j.celrep.2020.108237.

736 Cheverud JM, Lawson HA, Fawcett GL, Wang B, Pletscher LS, R Fox A, Maxwell TJ, Ehrich TH,
737 Kenney-Hunt JP, Wolf JB, et al. 2011. Diet-dependent genetic and genomic imprinting effects on
738 obesity in mice. *Obesity (Silver Spring)* **19**: 160–70.
739 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3677968&tool=pmcentrez&rendertype=
740 abstract.
741 Cho SW, Yang J-Y, Her SJ, Choi HJ, Jung JY, Sun HJ, An JH, Cho HY, Kim SW, Park KS, et al. 2011.
742 Osteoblast-targeted overexpression of PPARγ inhibited bone mass gain in male mice and
743 accelerated ovariectomy-induced bone loss in female mice. *J Bone Miner Res* **26**: 1939–1952.
744 https://asbmr.onlinelibrary.wiley.com/doi/full/10.1002/jbmr.366 (Accessed July 15, 2021).
745 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR.
746 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
747 Erol M, Gayret OB, Nacaroglu HT, Yigit O, Zengi O, Akkurt MS, Tasdemir M. 2016. Association of
748 Osteoprotegerin with Obesity, Insulin Resistance and Non-Alcoholic Fatty Liver Disease in
749 Children. *Iran Red Crescent Med J* **18**: 41873. /pmc/articles/PMC5294423/ (Accessed July 15,
750 2021).
751 Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ,
752 Prlic M, et al. 2015. MAST: a flexible statistical framework for assessing transcriptional changes
753 and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol 2015 161* **16**:
754 1–13. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0844-5 (Accessed
755 July 15, 2021).
756 Font-Clos F, Zapperi S, Porta CAM La. 2017. Integrative analysis of pathway deregulation in obesity.
757 *NPJ Syst Biol Appl* **3**. /pmc/articles/PMC5493646/ (Accessed July 15, 2021).
758 Franck N, Gummesson A, Jernås M, Glad C, Svensson P-A, Guillot G, Rudemo M, Nyström FH,
759 Carlsson LMS, Olsson B. 2011. Identification of Adipocyte Genes Regulated by Caloric Intake. *J*
760 *Clin Endocrinol Metab* **96**: E413–E418. https://academic.oup.com/jcem/article/96/2/E413/2709577
761 (Accessed July 15, 2021).

762 FS M, PH C de A, RF M, AJR C, M S de Q, B B de A, KCOS F, RJM M, GD F, LC A, et al. 2020.
763 RANKL induces beige adipocyte differentiation in preadipocytes. *Am J Physiol Endocrinol Metab*
764 **318**: E866–E877. https://pubmed.ncbi.nlm.nih.gov/32315212/ (Accessed July 15, 2021).
765 H I-T, T T, T U, T O, A A, N N, K M, Y N, H M. 2013. Serglycin is a novel adipocytokine highly
766 expressed in epicardial adipose tissue. *Biochem Biophys Res Commun* **432**: 105–110.
767 https://pubmed.ncbi.nlm.nih.gov/23376071/ (Accessed July 15, 2021).
768 Hager R, Cheverud JM, Wolf JB. 2008. Maternal effects as the cause of parent-of-origin effects that
769 mimic genomic imprinting. *Genetics* **178**: 1755–1762.
770 Jirtle RL. 2012. No Title. *http://www.geneimprint.com/site/genes-by-species*.
771 JM L, CK C, S B, J M, RG L, TN W. 2007. Interleukin-1beta selectively decreases the synthesis of
772 versican by arterial smooth muscle cells. *J Cell Biochem* **101**: 753–766.
773 https://pubmed.ncbi.nlm.nih.gov/17226775/ (Accessed July 15, 2021).
774 JS B, PJ E. 2011. Regulation of HMG-CoA reductase in mammals and yeast. *Prog Lipid Res* **50**: 403–
775 410. https://pubmed.ncbi.nlm.nih.gov/21801748/ (Accessed July 15, 2021).
776 JW K, Y C, JH P, JS K, KD P, DH B, SK S, KS C, SY L, HS K. 2008. The effects of cyclin-dependent
777 kinase inhibitors on adipogenic differentiation of human mesenchymal stem cells. *Biochem*
778 *Biophys Res Commun* **366**: 624–630. https://pubmed.ncbi.nlm.nih.gov/18047835/ (Accessed July
779 15, 2021).
780 Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. 2019.
781 Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *Am J Hum Genet* **104**: 65–
782 75. https://doi.org/10.1016/j.ajhg.2018.11.008.
783 Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, Jonasdottir A,
784 Sigurdsson A, Kristinsson KT, Jonasdottir A, et al. 2009. Parental origin of sequence variants
785 associated with complex diseases. *Nature* **462**: 868–874. http://dx.doi.org/10.1038/nature08625.
786 L Z, M A, J H, M G, G P, SO K. 2006. Serglycin is the major secreted proteoglycan in macrophages
787 and has a role in the regulation of macrophage tumor necrosis factor-alpha secretion in response

788 to lipopolysaccharide. *J Biol Chem* **281**: 26792–26801. https://pubmed.ncbi.nlm.nih.gov/16807245/
789 (Accessed July 15, 2021).
790 Lawson H a, Zelle KM, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP,
791 Wolf JB, Semenkovich CF, et al. 2010. Genetic, epigenetic, and gene-by-diet interaction effects
792 underlie variation in serum lipids in a LG/JxSM/J murine model. *J Lipid Res* **51**: 2976–84.
793 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2936764&tool=pmcentrez&rendertype=
794 abstract.
795 Lawson HA, Cady JE, Partridge C, Wolf JB, Semenkovich CF, Cheverud JM. 2011a. Genetic effects at
796 pleiotropic loci are context-dependent with consequences for the maintenance of genetic variation
797 in populations. *PLoS Genet* **7**.
798 Lawson HA, Cheverud JM, Wolf JB. 2013a. Genomic imprinting and parent-of-origin effects on complex
799 traits. *Nat Rev Genet*.
800 Lawson HA, Cheverud JM, Wolf JB. 2013b. Genomic imprinting and parent-of-origin effects on complex
801 traits. *Nat Rev Genet* **14**: 609–617.
802 http://dx.doi.org/10.1038/nrg3543%5Cnpapers2://publication/doi/10.1038/nrg3543.
803 Lawson HA, Lee A, Fawcett GL, Wang B, Pletscher LS, Maxwell TJ, Ehrich TH, Kenney-Hunt JP, Wolf
804 JB, Semenkovich CF, et al. 2011b. The importance of context to the genetic architecture of
805 diabetes-related traits is revealed in a genome-wide scan of a LG/J × SM/J murine model. *Mamm*
806 *Genome* **22**: 197–208.
807 Li X, Thomason PA, Withers DJ, Scott J. 2010. Bio-informatics analysis of a gene co-expression
808 module in adipose tissue containing the diet-responsive gene Nnat. *BMC Syst Biol* **4**: 1–11.
809 Lin H-H, Bell E, Uwanogho D, Perfect LW, Noristani H, Bates TJD, Snetkov V, Price J, Sun Y-M. 2010.
810 Neuronatin Promotes Neural Lineage in ESCs via Ca2+ Signaling. *Stem Cells* **28**: 1950–1960.
811 www.ambion.com/.
812 Luecken MD, Theis FJ. 2019. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol*
813 *Syst Biol* **15**: e8746. https://www.embopress.org/doi/full/10.15252/msb.20188746 (Accessed July

15, 2021).

LW S, PA P, RT C, MA S. 1991. Expression of CA III in rodent models of obesity. *Mol Endocrinol* **5**: 860–866. https://pubmed.ncbi.nlm.nih.gov/1922100/ (Accessed July 15, 2021).

M H, B Z-M, J J, T N, K W-S, A Z-G, A W. 2007. The influence of weight loss on serum osteoprotegerin concentration in obese perimenopausal women. *Obesity (Silver Spring)* **15**: 1925–1929. https://pubmed.ncbi.nlm.nih.gov/17712108/ (Accessed July 15, 2021).

McDonald LJ, Mararack MD. 1995. Phosphoinositide hydrolysis by phospholipase C modulated by multivalent cations La3, Al3+, neomycin, polyamines, and melittin. *J Lipid Mediat Cell Signal* **11**: 81–91.

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 β-cell mass in obese SM/J mice. *Physiol Rep* **8**: 1–12.

Mochizuki A, Takeda Y, Iwasa Y. 1996. The evolution of genomic imprinting. *Genetics* **144**: 1283–1295.

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. http://dx.doi.org/10.1016/j.cell.2013.11.043.

Mozaffari S V., 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. *Commun Biol*.

N N, M K, K Y, N S, H Y, K Y, T M, K H. 1998. RANK is the essential signaling receptor for osteoclast differentiation factor in osteoclastogenesis. *Biochem Biophys Res Commun* **253**: 395–400. https://pubmed.ncbi.nlm.nih.gov/9878548/ (Accessed July 15, 2021).

Nakamura Y, Hamada Y, Fujiwara T, Enomoto H, Hiroe T, Tanaka S, Nose M, Nakahara M, Yoshida N, Takenawa T, et al. 2005. Phospholipase C-1 and-3 Are Essential in the Trophoblast for Placental Development. *Mol Cell Biol* **25**: 10979–10988. https://journals.asm.org/journal/mcb.

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. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4445795&tool=pmcentrez&rendertype=abstract.

Parker-Katiraee L, Carson AR, Yamada T, Arnaud P, Feil R, Abu-Amero SN, Moore GE, Kaneda M, Perry GH, Stone AC, et al. 2007. Identification of the imprinted KLF14 transcription factor undergoing human-specific accelerated evolution. *PLoS Genet* **3**: 665–678.

Patten MM, Ross L, Curley JP, Queller DC, Bonduriansky R, Wolf JB. 2014. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity (Edinb)* **113**: 119–28. http://dx.doi.org/10.1038/hdy.2014.29.

Pelled D, Lloyd-Evans E, Riebeling C, Jeyakumar M, Platt FM, Futerman AH. 2003. Inhibition of Calcium Uptake via the Sarco/Endoplasmic Reticulum Ca2+-ATPase in a Mouse Model of Sandhoff Disease and Prevention by Treatment with N-Butyldeoxynojirimycin. *J Biol Chem* **278**: 29496–29501. http://dx.doi.org/10.1074/jbc.M302964200.

Pette M Van de, Tunster SJ, John RM. 2018. Loss of Imprinting of Cdkn1c Protects against Age and Diet-Induced Obesity. *Int J Mol Sci* **19**. /pmc/articles/PMC6164918/ (Accessed July 15, 2021).

Prockop DJ. 1997. Marrow Stromal Cells as Stem Cells for Nonhematopoietic Tissues. *Science (80-)* **276**: 71–74. https://science.sciencemag.org/content/276/5309/71 (Accessed July 15, 2021).

R Core Team. 2013. R: A Language and Environment for Statistical Computing. http://www.r-project.org/.

RB B, VD R, EA R, R P-R, YH L, JG G. 2018. Deconstructing Adipogenesis Induced by β3-Adrenergic Receptor Activation with Single-Cell Expression Profiling. *Cell Metab* **28**: 300-309.e4. https://pubmed.ncbi.nlm.nih.gov/29937373/ (Accessed July 15, 2021).

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. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0176502 (Accessed July 15,

2021).

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–40.

Rodriguez J, Astudillo P, Rios S, Pino A. 2008. Involvement of Adipogenic Potential of Human Bone Marrow Mesenchymal Stem Cells (MSCs) in Osteoporosis. *Curr Stem Cell Res Ther* **3**: 208–218.

Rosen CJ, Bouxsein ML. 2006. Mechanisms of Disease: is osteoporosis the obesity of bone? *Nat Clin Pract Rheumatol 2006 21* **2**: 35–43. https://www.nature.com/articles/ncprheum0070 (Accessed July 15, 2021).

RT T, SA M, UT I. 2018. Metabolic Coupling Between Bone Marrow Adipose Tissue and Hematopoiesis. *Curr Osteoporos Rep* **16**: 95–104. https://pubmed.ncbi.nlm.nih.gov/29492879/ (Accessed July 15, 2021).

SJ M, G DSX, AI C, S B, P C, SM P, EE I, A M, P F, WR T, et al. 2018. Neuronatin regulates pancreatic β cell insulin content and secretion. *J Clin Invest* **128**: 3369–3381. https://pubmed.ncbi.nlm.nih.gov/29864031/ (Accessed July 15, 2021).

Small KS, Hedman ÅK, Grundberg E, Nica AC, Thorleifsson G, Kong A, Thorsteindottir U, Shin S-Y, Richards HB, Soranzo N, et al. 2011. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet* **43**: 561–564. http://www.nature.com/ng/journal/v43/n6/full/ng.833.html.

Soh UJ, Dores MR, Chen B, Trejo J. 2010. Signal transduction by protease-activated receptors. *Br J Pharmacol* **160**: 191–203. www.brjpharmacol.org.

T S, A B, P H, C H, E P, WM M, Y H, M S, P S, R S. 2019. Comprehensive Integration of Single-Cell Data. *Cell* **177**: 1888-1902.e21. https://pubmed.ncbi.nlm.nih.gov/31178118/ (Accessed July 15, 2021).

Teslovich TM, Musunuru K, Smith A V, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, et al. 2010. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**: 707–13.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3039276&tool=pmcentrez&rendertype=abstract.

Thatcher JD. 2010. The Inositol Trisphosphate (IP3) Signal Transduction Pathway. *Sci Signal* **3**: tr3–tr3. https://stke.sciencemag.org/content/3/119/tr3 (Accessed July 15, 2021).

Voight BF, Scott LJ, Steinthorsdottir V, Morris ADP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, et al. 2010. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* **42**: 579–89. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3080658&tool=pmcentrez&rendertype=abstract.

Wolf JB, Cheverud JM. 2009. A framework for detecting and characterizing genetic background-dependent imprinting effects. *Mamm Genome* **20**: 681–698.

Y J, RA D-B. 2010. Control of cholesterol synthesis through regulated ER-associated degradation of HMG CoA reductase. *Crit Rev Biochem Mol Biol* **45**: 185–198. https://pubmed.ncbi.nlm.nih.gov/20482385/ (Accessed July 15, 2021).

Yee TW. 2010. The **VGAM** Package for Categorical Data Analysis. *J Stat Softw*.

Yeh Y-S, Jheng H-F, Iwase M, Kim M, Mohri S, Kwon J, Kawarasaki S, Li Y, Takahashi H, Ara T, et al. 2018. The Mevalonate Pathway Is Indispensable for Adipocyte Survival. *iScience* **9**: 175. /pmc/articles/PMC6223239/ (Accessed July 15, 2021).

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. *Biochem Biophys Res Commun*.

Zaky DS, Ali AA, Abd-Elraheem SE, Abdel-Moniem SH. Circulating Osteoprotegerin Level in Relation to Obesity in Middle Aged Females. http://article.sapub.org/10.5923.j.ijpt.20190802.02.html (Accessed July 15, 2021).

Zappia L, Oshlack A. 2018. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**: 1–9. https://academic.oup.com/gigascience/article/7/7/giy083/5052205

(Accessed July 15, 2021).

Zeng Y, Amador C, Xia C, Marioni R, Sproul D, Walker RM, Morris SW, Bretherick A, Canela-Xandri O, Boutin TS, et al. 2019. Parent of origin genetic effects on methylation in humans are common and influence complex trait variation. *Nat Commun* **10**: 1383.

Zhang B, Kirov S, Snoddy J. 2005. WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res*.

Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, et al. 2017. Massively parallel digital transcriptional profiling of single cells. *Nat Commun 2017 81* **8**: 1–12. https://www.nature.com/articles/ncomms14049 (Accessed July 15, 2021).