

¹ Transcripts with high distal heritability mediate genetic effects on
² complex metabolic traits

³

⁴ Anna L. Tyler, J. Matthew Mahoney, Mark P. Keller, Candice N. Baker, Margaret Gaca, Anuj Srivastava,
⁵ Isabela Gerdes Gyuricza, Madeleine J. Braun, Nadia A. Rosenthal, Alan D. Attie, Gary A. Churchill and
⁶ Gregory W. Carter

⁷ **Abstract**

⁸ Although many genes are subject to local regulation, recent evidence suggests that complex distal regulation
⁹ may be more important in mediating phenotypic variability. To assess the role of distal gene regulation in
¹⁰ complex traits, we combined multi-tissue transcriptomes with physiological outcomes to model diet-induced
¹¹ obesity and metabolic disease in a population of Diversity Outbred mice. Using a novel high-dimensional
¹² mediation analysis, we identified a composite transcriptome signature that summarized genetic effects on
¹³ gene expression and explained 30% of the variation across all metabolic traits. The signature was heritable,
¹⁴ interpretable in biological terms, and predicted obesity status from gene expression in an independently
¹⁵ derived mouse cohort and multiple human studies. Transcripts contributing most strongly to this composite
¹⁶ mediator frequently had complex, distal regulation distributed throughout the genome. These results suggest
¹⁷ that trait-relevant variation in transcription is largely distally regulated, but is nonetheless identifiable,
¹⁸ interpretable, and translatable across species.

¹⁹ **Introduction**

²⁰ Evidence from genome-wide association studies (GWAS) suggests that most heritable variation in complex
²¹ traits is mediated through regulation of gene expression. The majority of trait-associated variants lie
²² in gene regulatory regions^{1–7}, suggesting a relatively simple causal model in which a variant alters the
²³ homeostatic expression level of a nearby (local) gene which, in turn, alters a trait. Statistical methods such
²⁴ as transcriptome-wide association studies (TWAS)^{8–11} and summary data-based Mendelian randomization
²⁵ (SMR)¹⁰ have used this idea to identify genes associated with multiple disease traits^{12–15}. However, despite

26 the great promise of these methods, explaining trait effects with local gene regulation has been more difficult
27 than initially assumed^{16;17}. Although trait-associated variants typically lie in non-coding, regulatory regions,
28 these variants often have no detectable effects on gene expression¹⁶ and tend not to co-localize with expression
29 quantitative trait loci (eQTLs)^{17;18}. These observations suggest that the relationship among genetic variants,
30 gene expression, and organism-level traits is more complex than the simple, local model.

31 In recent years the conversation around the genetic architecture of common disease traits has been addressing
32 this complexity, and there is increased interest in ~~distal more distant (distal) genetic~~ effects as potential
33 drivers of trait variation^{18–20;15;21}. In ~~particular, the omnigenic model~~general, distal effects are defined as
34 being greater than 4 or 5Mb away from the transcription start site of a given gene. We use the terms local
35 and distal rather than cis and trans because cis and trans have specific biochemical meanings²², whereas
36 local and distal are defined only by genomic position. The importance of distal genetic effects is proposed
37 in the omnigenic model, which posits that trait-driving genes are cumulatively influenced by many distal
38 variants. In this view, the heritable transcriptomic signatures driving clinical traits are an emergent state
39 arising from the myriad molecular interactions defining and constraining gene expression. Consistent with
40 this view, it has been suggested that part of the difficulty in explaining trait variation through local eQTLs
41 may arise in part because gene expression is not measured in the appropriate cell types¹⁶, or cell states²³,
42 and thus local eQTLs influencing traits cannot be detected in bulk tissue samples. This context dependence
43 emphasizes the essential role of complex regulatory and tissue networks in mediating variant effects. The
44 mechanistic dissection of complex traits in this model is more challenging because it requires addressing
45 network-mediated effects that are weaker and greater in number. However, the comparative importance of
46 distal effects over local effects is currently only conjectured and extremely challenging to address in human
47 populations.

48 To assess the role of wide-spread distal gene regulation in the genetic architecture of complex traits, we used
49 genetically diverse mice as a model system. In mice we can obtain simultaneous measurements of the genome,
50 transcriptome, and phenotype in all individuals. We used diet-induced obesity and metabolic disease as an
51 archetypal example of a complex trait. In humans, these phenotypes are genetically complex with hundreds of
52 variants mapped through GWAS^{24;25} that are known to act through multiple tissues^{26;27}. Likewise in mice,
53 metabolic traits are also genetically complex²⁸ and synteny analysis implicates a high degree of concordance
54 in the genetic architecture between species^{28;12}. Furthermore, in contrast to humans, in mice we have access
55 to multiple disease-relevant tissues in the same individuals with sufficient numbers for adequate statistical
56 power.

57 We generated two complementary data sets: a discovery data set in a large population of Diversity Outbred

58 (DO) mice²⁹, and an independent validation data set derived by crossing inbred strains from the Collaborative
59 Cross (CC) recombinant inbred lines³⁰ to form CC recombinant inbred intercross (CC-RIX) mice. Both
60 populations were maintained on a high-fat, high-sugar diet to model diet-induced obesity and metabolic
61 disease¹².

62 The DO population and CC recombinant inbred lines were derived from the same eight inbred founder
63 strains: five classical lab strains and three strains more recently derived from wild mice²⁹, representing three
64 subspecies and capturing 90% of the known variation in laboratory mice³¹. The DO mice are maintained
65 with a breeding scheme that ensures equal contributions from each founder across the genome thus rendering
66 almost the whole genome visible to genetic inquiry and maximizing power to detect eQTLs²⁹. The CC mice
67 were initially intercrossed to recombine the genomes from all eight founders, and then inbred for at least 20
68 generations to create recombinant inbred lines^{30;32;31}. Because these two populations have common ancestral
69 haplotypes but highly distinct kinship structure, we could directly and unambiguously compare the local
70 genetic effects on gene expression at the whole-transcriptome level while varying the population structure
71 driving distal regulation.

72 In the DO population, we paired clinically relevant metabolic traits, including body weight and plasma levels
73 of insulin, glucose and lipids¹², with transcriptome-wide gene expression in four tissues related to metabolic
74 disease: adipose tissue, pancreatic islets, liver, and skeletal muscle. We measured similar metabolic traits
75 in a CC-RIX population and gene expression from three of the four tissues used in the DO: adipose tissue,
76 liver, and skeletal muscle. Measuring gene expression in multiple tissues is critical to adequately assess the
77 extent to which local gene regulation varies across the tissues and whether such variability might account for
78 previous failed attempts to identify trait-relevant local eQTLs. The CC-RIX carry the same founder alleles
79 as the DO. Thus, local gene regulation is expected to match between the populations. However, because
80 the alleles are recombined throughout the genome, distal effects are expected to vary from those in the DO,
81 allowing us to directly assess the role of distal gene regulation in driving trait-associated transcript variation.
82 To mechanistically dissect distal effects on metabolic disease, we developed a novel dimension reduction
83 framework called high-dimensional mediation analysis (HDMA) to identify the heritable transcriptomic
84 signatures driving trait variation, which we compared between mouse populations and to human data sets
85 with measured adipose gene expression. Together, these data enable a comprehensive view into the genetic
86 architecture of metabolic disease.

87 **Results**

88 **Genetic variation contributed to wide phenotypic variation**

89 Although the environment was consistent across the DO mice, the genetic diversity present in this population
90 resulted in widely varying distributions across physiological measurements (Fig. 1). For example, body
91 weights of adult individuals varied from less than the average adult C57BL/6J (B6) body weight to several
92 times the body weight of a B6 adult in both sexes (Males: 18.5 - 69.1g, Females: 16.0 - 54.8g) (Fig. 1A).
93 Fasting blood glucose (FBG) also varied considerably (Fig. 1B), although few of the animals had FBG levels
94 that would indicate pre-diabetes (19 animals, 3.8%), or diabetes (7 animals, 1.4%) according to previously
95 developed cutoffs (pre-diabetes: $\text{FBG} \geq 250 \text{ mg/dL}$, diabetes: $\text{FBG} \geq 300 \text{ mg/dL}$)³³. Males had higher
96 FBG than females on average (Fig. 1C) as has been observed before suggesting either that males were more
97 susceptible to metabolic disease on the high-fat, high-sugar (HFHS) diet, or that males and females may
98 require different thresholds for pre-diabetes and diabetes.

99 Body weight was strongly positively correlated with food consumption (Fig. 1D $R^2 = 0.51, p < 2.2 \times 10^{-16}$)
100 and FBG (Fig. 1E, $R^2 = 0.21, p < 2.2 \times 10^{-16}$) suggesting a link between behavioral factors and metabolic
101 disease. However, the heritability of this trait and others (Fig. 1F) indicates that genetics contribute
102 substantially to correlates of metabolic disease in this population.

103 The trait correlations (Fig. 1G) showed that most of the metabolic trait pairs were only modestly correlated,
104 which, in conjunction with the trait decomposition (Supp. Fig. S1), suggests complex relationships among
105 the measured traits and a broad sampling of multiple heritable aspects of metabolic disease including overall
106 body weight, glucose homeostasis, and pancreatic function.

107 **Distal Heritability Correlated with Phenotype Relevance**

108 To comprehensively assess the genetic control of gene expression in metabolic disease we measured overall
109 gene expression via bulk RNA-Seq in adipose, islet, liver, and skeletal muscle in the DO cohort (Supp. Fig.
110 S2). We performed eQTL analysis using R/qtL2³⁴ (Methods) and identified both local and distal eQTLs for
111 transcripts in each of the four tissues (Supp. Fig. S2B-E). Significant local eQTLs far outnumbered distal
112 eQTLs (Supp. Fig. S2F) and tended to be shared across tissues (Supp. Fig. S2G) whereas the few significant
113 distal eQTLs we identified tended to be tissue-specific (Supp. Fig. S2H)

114 We ~~calculated~~ estimated the heritability of each transcript in terms of local and all non-local (distal) genetic
115 factors (Methods). Overall, local and distal genetic factors contributed approximately equally to transcript
116 abundance. In all tissues, both local and distal factors explained between 8 and ~~1817~~ 1817% of the variance in the

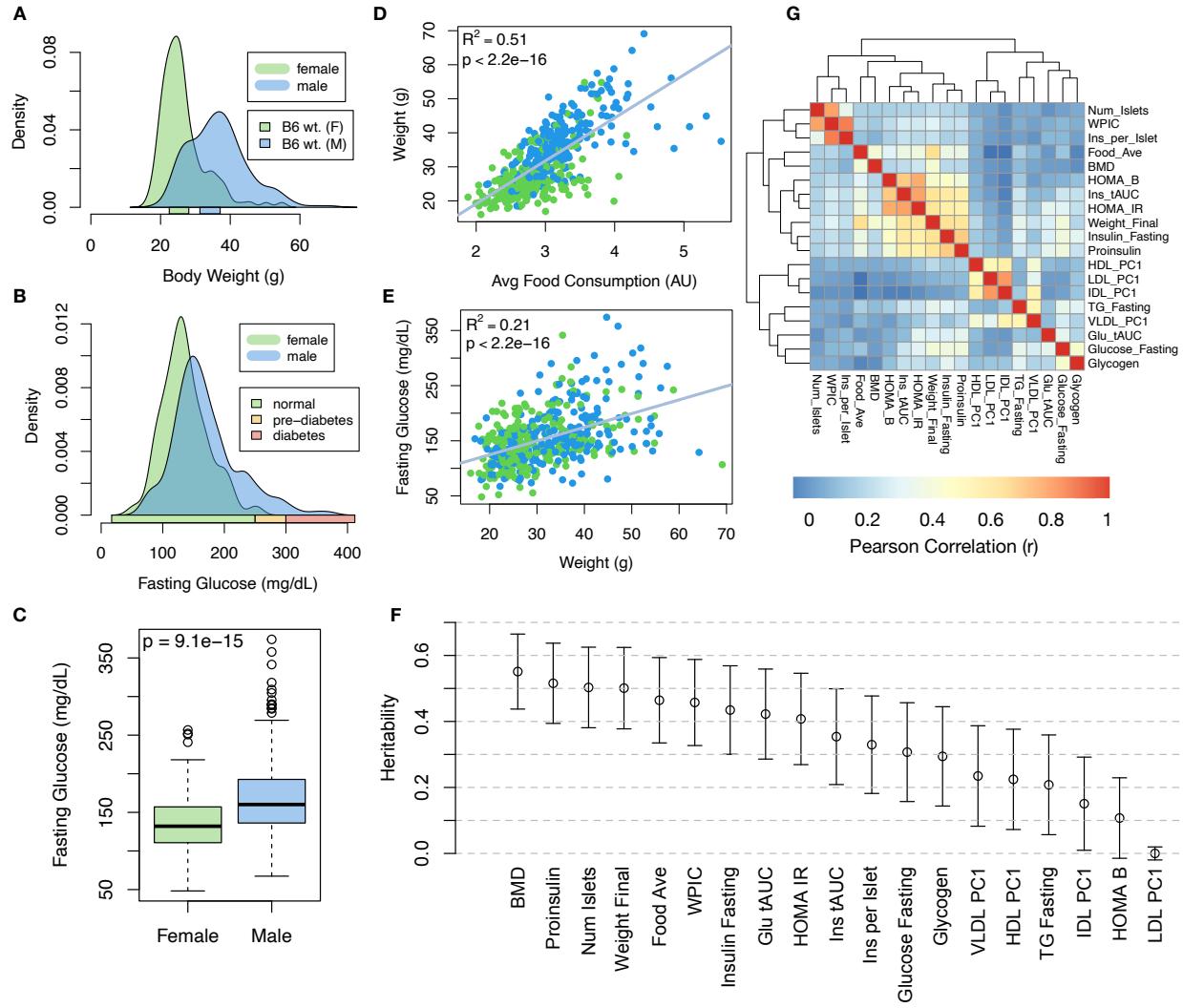


Figure 1: Clinical overview. **A.** Distributions of final body weight in the diversity outbred mice. Sex is indicated by color. The average B6 male and female adult weights at 24 weeks of age are indicated by blue and green bars on the x-axis. **B.** The distribution of final fasting glucose across the population split by sex. Normal, pre-diabetic, and diabetic fasting glucose levels for mice are shown by colored bars along the x-axis. **C.** Males had higher fasting blood glucose on average than females ($p = 9.1 \times 10^{-15}$). **D.** The relationship between food consumption and body weight for both sexes. **E.** Relationship between body weight and fasting glucose for both sexes. **F.** Heritability estimates for each physiological trait. Bars show standard error of the estimate. **G.** Correlation structure between pairs of physiological traits. BMD - bone mineral density, WPIC - whole pancreas insulin content, Glu tAUC - glucose total area under the curve, HOMA IR - homeostatic measurement of insulin resistance, HOMA B - homeostatic measure of beta cell health, VLDL - very low-density lipoprotein, LDL - low-density lipoprotein, IDL - intermediate density lipoprotein, HDL - high-density lipoprotein, TG - triglyceride.

117 median transcript (Fig, 2A).

118 The equal contribution of local and distal genetic variants to the heritability of transcript abundance contrasts
 119 with findings in humans in which local variants have been found to explain 20-30% of total heritability, while

120 distal variants explain the remaining 70-80%^{35:36}. This discrepancy may arise due to the high degree of
121 linkage disequilibrium in the DO mice compared to human populations. Each genetic marker in the mice
122 captures information from a larger genomic region than each genetic marker in human populations, and thus
123 may capture more local regulatory variants than SNPs capture in humans. It has been found that transcripts
124 with multiple local eQTL have higher local heritability than transcripts with single local eQTL³⁷. Because of
125 the high diversity in the DO and the high rates of linkage disequilibrium, it is possible that there are more
126 local variants regulating transcription creating a proportionally larger effect of local regulation.

127 To assess the importance of genetic regulation of transcript levels to clinical traits, we compared the local
128 and distal heritabilities of transcripts to their trait relevance, ~~defined as the maximum trait correlation for~~
129 ~~each transcript~~. We defined trait relevance for a transcript as its maximum absolute Spearman correlation
130 coefficient (ρ) across all traits (Methods). The local heritability of transcripts was negatively ~~correlated~~
131 ~~associated~~ with their trait relevance (Fig. 2B), suggesting that the more local genotype influenced transcript
132 abundance, the less effect this variation had on the measured traits. Conversely, the distal heritability of
133 transcripts was positively ~~correlated~~~~associated~~ with trait relevance (Fig. 2C). That is, transcripts that
134 were more highly correlated with the measured traits tended to be distally, rather than locally, heritable.
135 Importantly, this pattern was consistent across all tissues. This finding is consistent with previous observations
136 that transcripts with low local heritability explain more expression-mediated disease heritability than
137 transcripts with high local heritability¹⁹. However, the positive relationship between trait correlation and
138 distal heritability demonstrated further that there are diffuse genetic effects throughout the genome converging
139 on trait-related transcripts.

140 **High-Dimensional Mediation Analysis identified a high-heritability composite trait that was
141 mediated by a composite transcript**

142 The above univariate analyses establish the importance of distal heritability for trait-relevant transcripts.
143 However, the number of transcripts dramatically exceeds the number of phenotypes. Thus, we expect the
144 heritable, trait-relevant transcripts to be highly correlated and organized according to coherent, biological
145 processes representing the mediating endophenotypes driving clinical trait variation. To identify these
146 endophenotypes in a theoretically principled way, we developed a novel dimension-reduction technique,
147 high-dimension mediation analysis (HDMA), that uses the theory of causal graphical models to identify a
148 transcriptomic signature that is simultaneously 1) highly heritable, 2) strongly correlated to the measured phe-
149 notypes, and 3) conforms to the causal mediation hypothesis (Fig. 3). ~~HDMA projects the high-dimensional~~
150 ~~In HDMA, we first use a linear mapping called kernelization to dimension-reduce the~~ genome, transcriptome,

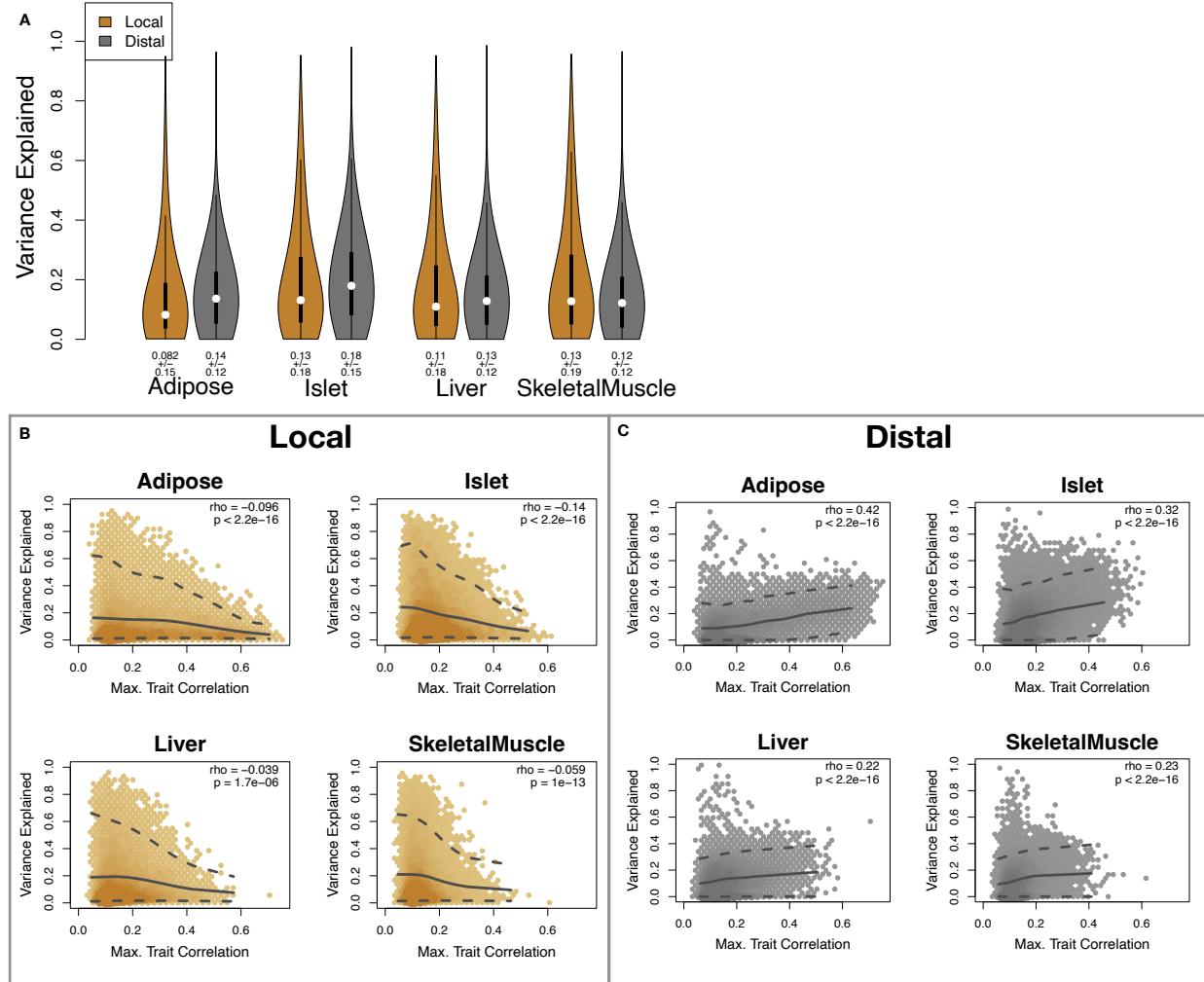


Figure 2: Transcript heritability and trait relevance. **A.** Distributions of distal and local and distal heritability of transcripts across the four tissues. Overall local and distal factors contribute contributed equally to transcript heritability. Arrows indicate the median of each distribution. The relationship between (B.) local and (C.) distal heritability and trait relevance across all four tissues. Here trait relevance is defined as the maximum correlation between the transcript and all traits. Local heritability was negatively The upper and lower dashed line in each panel show the 95th and 5th percentile correlation. The solid line shows the mean trait correlation in transcripts with increasing variance explained either locally (B) or distally (C). Transcripts that are highly correlated with trait relevance, traits tended to have low local heritability and high distal heritabilityis positively correlated with trait relevance. Pearson (r) and p values for each correlation are shown in the upper right of each panel.

151 and phenotype data to kernel matrices G_K , T_K and P_K respectively, which each have the dimensions n by n
 152 where n is the number of individuals (Methods). These kernel matrices describe the relationships among the
 153 individual mice in genome space, transcriptome space, and phenome space and ensure that these three
 154 omic spaces have the same dimensions, and thus the same weight in the analysis. If not dimension-reduced,
 155 the transcriptome would outweigh the phenome in the model. We then projected these $n \times n$ -dimensional
 156 kernel matrices onto one-dimensional scores—a composite genome score (G_C), a composite transcriptome score

157 (T_C), and a composite phenotype score (P_C)—and uses used the univariate theory of mediation to constrain these
 158 projections to satisfy the hypotheses of perfect mediation, namely that upon controlling for the transcriptomic
 159 score, the genome score is uncorrelated to the phenotype score. A complete mathematical derivation and
 160 implementation details for HDMA are available in Supp. the Methods.

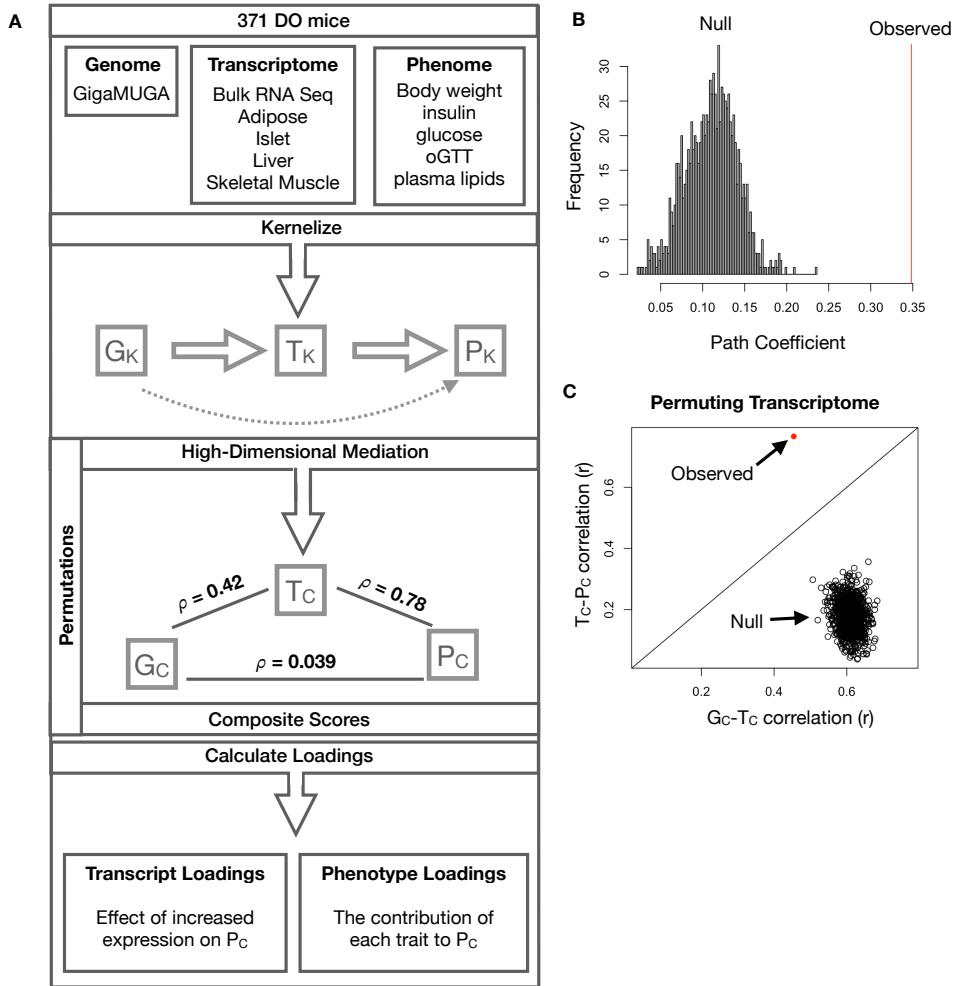


Figure 3: High-dimensional mediation. **A.** Workflow indicating major steps of high-dimensional mediation. The genotype, transcriptome, and phenotype matrices were independently normalized and converted to kernel matrices representing the pairwise relationships between individuals for each data modality ($K_G G_K$ = genome kernel, $K_T T_K$ = transcriptome kernel; $K_P P_K$ = phenotype kernel). High-dimensional mediation was applied to these matrices to maximize the direct path $G \rightarrow T \rightarrow P$, the mediating pathway (arrows), while simultaneously minimizing the direct $G \rightarrow P$ pathway (dotted line). The composite vectors that resulted from high-dimensional mediation were $G_C G_C$, T_C , and P_C . The partial correlations ρ between these vectors indicated perfect mediation. Transcript and trait loadings were calculated as described in the methodsMethods. **B.** The null distribution of the path coefficient derived from 10,000 permutations compared to the observed path coefficient (red line). **C.** The null distribution of the G_C-T_C correlation vs. the T_C-P_C correlation compared with the observed value (red dot).

161 Using HDMA we identified the major axis of variation in the transcriptome that was consistent with mediating

162 the effects of the genome on metabolic traits (Fig 3). Fig. 3A shows the partial correlations (ρ) between
163 the pairs of these composite vectors. The partial correlation between G_C and T_C was 0.42, and the partial
164 correlation between T_C and P_C was 0.78. However, when the transcriptome was taken into account, the
165 partial correlation between G_C and P_C was effectively zero (0.039). P_C captured 30% of the overall trait
166 variance, and its estimated heritability was 0.71 ± 0.084 , which was higher than any of the measured traits
167 (Fig. 1F). Thus, HDMA identified a maximally heritable metabolic composite trait and a highly heritable
168 component of the transcriptome that are correlated as expected in the perfect mediation model.

169 As discussed in ~~Supp-~~the Methods, HDMA is related to a generalized form of ~~CCA~~canonical correlation
170 analysis (CCA). Standard CCA is prone to over-fitting because in any two large matrices it can be trivial to
171 identify highly correlated composite vectors³⁸. To assess whether our implementation of HDMA was similarly
172 prone to over-fitting in a high-dimensional space, we performed permutation testing. We permuted the
173 individual labels on the transcriptome matrix 10,000 times and recalculated the path coefficient, which is the
174 correlation of G_C and T_C multiplied by the correlation of T_C and P_C . This represents the strength of the path
175 from G_C to P_C that is putatively mediated through T_C . The permutations preserved the correlation between
176 the genome and phenotype, but broke the correlations between the genome and the transcriptome, as well as
177 between the transcriptome and the phenotype. We could thus test whether, given a random transcriptome,
178 HDMA would overfit and identify apparently mediating transcriptomic signatures in random data. The null
179 distribution of the path coefficient is shown in Fig. 3B, and the observed path coefficient from the original
180 data is indicated by a red line. The observed path coefficient was well outside the null distribution generated
181 by permutations ($p < 10^{-16}$). Fig. 3C illustrates this observation in more detail. Although we identified
182 high correlations between G_C and T_C , and modest correlations between T_C and P_C in the null data (Fig 3C),
183 these two values could not be maximized simultaneously in the null data. In contrast, the red dot shows that
184 in the real data both the G_C - T_C correlation and the T_C - P_C correlation could be maximized simultaneously
185 suggesting that the path from genotype to phenotype through ~~the~~ transcriptome is highly non-trivial and
186 identifiable in this case. These results suggest that these composite vectors represent genetically determined
187 variation in phenotype that is mediated through genetically determined variation in transcription.

188 **Body weight and insulin resistance were highly represented in the expression-mediated com-**
189 **posite trait**

190 Each composite score is a weighted combination of the measured variables. The magnitude and sign of the
191 weights, called loadings, correspond to the relative importance and directionality of each variable in the
192 composite score. The loadings of each measured trait onto P_C indicate how much each contributed to the

composite phenotype. Body weight contributed the most (Fig. 4), followed by homeostatic insulin resistance (HOMA_{IR}) and fasting plasma insulin levels (Insulin_{Fasting}). We can thus interpret P_C as an index of metabolic disease (Fig. 4B). Individuals with high values of P_C have a higher metabolic disease index (MDI) and greater metabolic disease, including higher body weight and higher insulin resistance. We refer to P_C as the MDI going forward. Traits contributing the least to the MDI were measures of cholesterol and pancreas composition. Thus, when we interpret the transcriptomic signature identified by HDMA, we are explaining primarily the putative transcriptional mediation of body weight and insulin resistance, as opposed to cholesterol measurements.

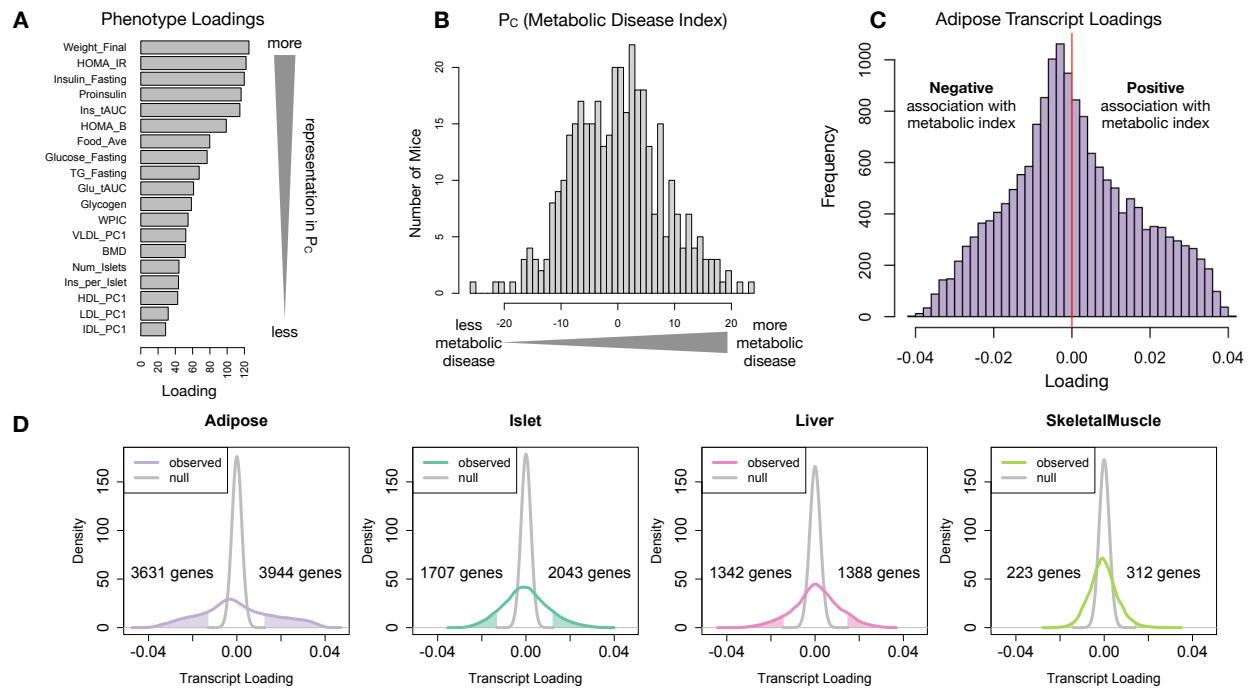


Figure 4: Interpretation of loadings. **A.** Loadings across traits. Body weight and insulin resistance contributed the most to the composite trait. **B.** Phenotype scores (MDI) across individuals. Individuals with large positive phenotype scores MDI had higher body weight and insulin resistance than average. Individuals with large negative phenotype scores MDI had lower body weight and insulin resistance than average. **C.** Distribution of transcript loadings in adipose tissue. For transcripts with large positive loadings, higher expression was associated with higher phenotype scores. For transcripts with large negative loadings, higher expression was associated with lower phenotype scores. **D.** Distribution of absolute value of transcript loadings across tissues compared to null distributions. Shaded areas represent loadings that were more extreme than the null distribution. Numbers indicate how many transcripts had loadings above and below the extremes of the null. Transcripts in adipose tissue had the largest most extreme loadings indicating that transcripts in adipose tissue gene expression was a strong mediator were the best mediators of genotype the genetic effects on body weight and insulin resistance.

201 **High-loading transcripts had low local heritability, high distal heritability, and were linked**
202 **mechanistically to obesity**

203 We interpreted large loadings onto transcripts as indicating strong mediation of the effect of genetics on ~~the~~
204 MDI. Large positive loadings indicate that higher expression was associated with a higher MDI (i.e. higher
205 risk of obesity and metabolic disease on the HFHS diet) (Fig. 4C). Conversely, large negative loadings
206 indicate that high expression of these transcripts was associated with a lower MDI (i.e. lower risk of obesity
207 and metabolic disease on the HFHS diet) (Fig. 4C). We used gene set enrichment analysis (GSEA)^{39;40} to
208 look for biological processes and pathways that were enriched at the top and bottom of this list (Methods).

209 In adipose tissue, both GO processes and KEGG pathway enrichments pointed to an axis of inflammation and
210 metabolism (Figs. S4 and S5). GO terms and KEGG pathways associated with inflammation were positively
211 associated with ~~the~~-MDI, indicating that increased expression in inflammatory pathways was associated
212 with a higher burden of disease. It is well established that adipose tissue in obese individuals is inflamed
213 and infiltrated by macrophages^{41–45}, and the results here suggest that this may be a dominant heritable
214 component of metabolic disease.

215 The strongest negative enrichments in adipose tissue were related to mitochondrial activity in general, and
216 thermogenesis in particular (Figs. S4 and ~~S4S5~~). Genes in the KEGG oxidative phosphorylation pathway
217 were almost universally negatively loaded in adipose tissue, suggesting that increased expression of these
218 genes was associated with reduced MDI (Supp. Fig. S6). Consistent with this observation, it has been shown
219 previously that mouse strains with greater thermogenic potential are also less susceptible to obesity on an
220 obesigenic diet⁴⁶.

221 Transcripts associated with the citric acid cycle as well as the catabolism of the branched-chain amino acids
222 (valine, leucine, and isoleucine) were strongly enriched with negative loadings in adipose tissue (Supp. Figs.
223 S4, S7 and S8). Expression of genes in both pathways (for which there is some overlap) has been previously
224 associated with insulin sensitivity^{12;47;48}, suggesting that heritable variation in regulation of these pathways
225 may influence risk of insulin resistance.

226 Looking ~~at~~ the 10 most positively and negatively loaded transcripts from each tissue, it is apparent that
227 transcripts in the adipose tissue had the largest loadings, both positive and negative (Fig. 5A bar plot). This
228 suggests that much of the effect of genetics on body weight and insulin resistance is mediated through gene
229 expression in adipose tissue. This finding does not speak to the relative importance of tissues not included
230 in this study, such as brain, in which transcriptional variation may mediate a large portion of the genetic
231 effect on obesity. The strongest loadings in liver and pancreas were comparable, and those in skeletal muscle

were the weakest (Fig. 5A), suggesting that less of the genetic effects were mediated through transcription in skeletal muscle. As expected, heritability analysis showed that transcripts with the largest loadings had higher distal heritability than local heritability (Fig. 5A heat map and box plot). ~~This pattern contrasts with transcripts nominated by TWAS~~ We also performed TWAS in this population by imputing transcript levels for each gene based on local genotype only and correlating the imputed transcript levels with each trait. In contrast to HDMA, the TWAS procedure tended to nominate transcripts with lower loadings (Fig. 5B), ~~which tended to have lower loadings,~~ higher local heritability and lower distal heritability. ~~Transcripts Finally, we focused on transcripts~~ with the highest local heritability in each tissue (Fig. 5C) ~~had the lowest loadings.~~ This procedure selected transcripts with low loadings on average, consistent with our findings above (Fig. 2B).

We performed a literature search for the genes in each of these groups along with the terms “diabetes”, “obesity”, and the name of the expressing tissue to determine whether any of these genes had previous associations with metabolic disease in the literature (Methods). Multiple genes in each group had been previously associated with obesity and diabetes (Fig. 5 bolded gene names). Genes with high loadings were most highly enriched for previous literature support. They were ~~2.4~~2.1 times more likely than TWAS hits and 3.8 times more likely than genes with high local heritability to be previously associated with obesity or diabetes.

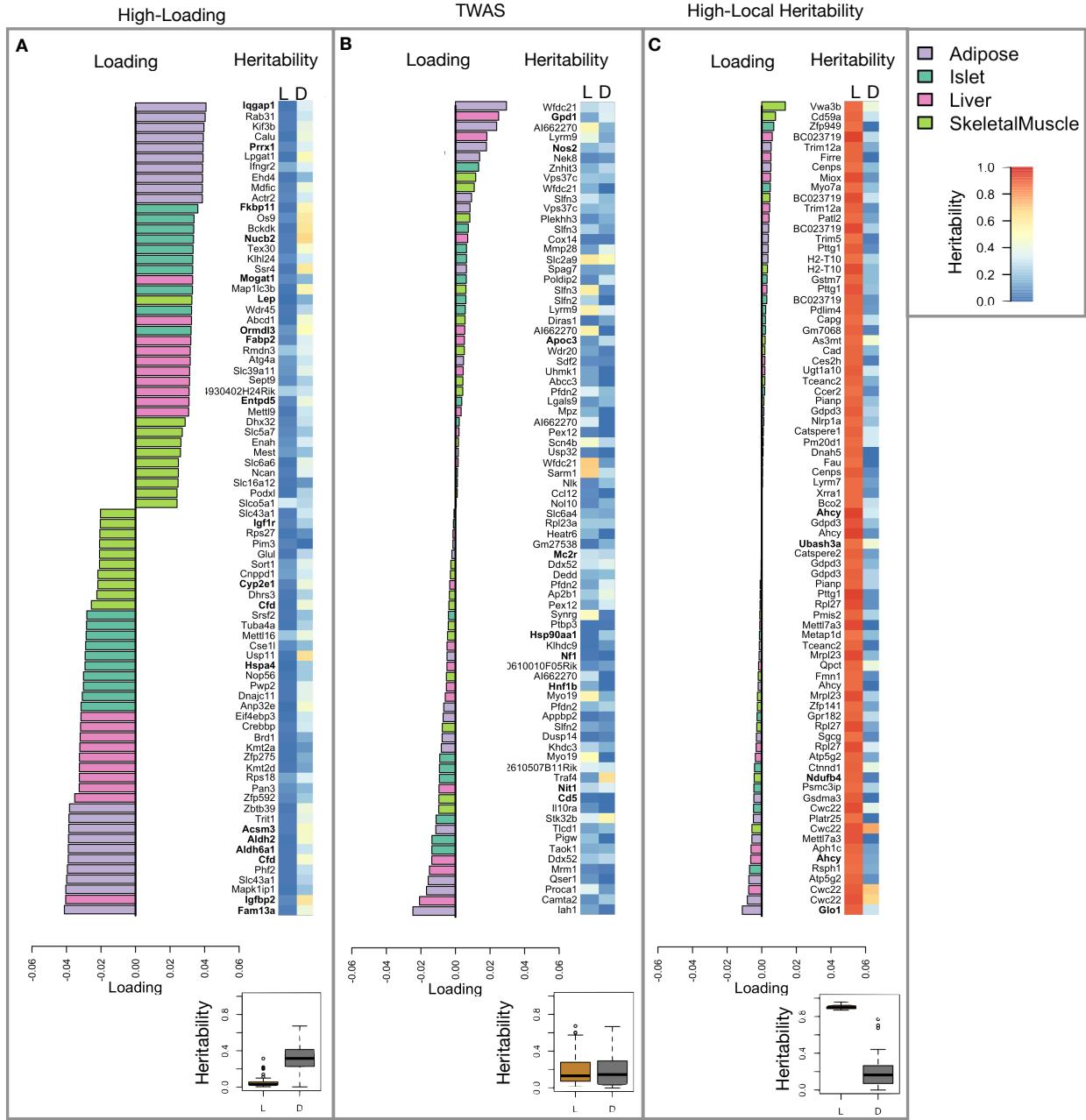


Figure 5: Transcripts with high loadings have high distal heritability and literature support ([bolded gene names](#)). Each panel has a bar plot showing the loadings of transcripts selected by different criteria. Bar color indicates the tissue of origin. The heat map shows the local (L - left) and distal (D - right) heritability of each transcript. **A.** Loadings for the 10 transcripts with the largest positive loadings and the 10 transcripts with the largest negative loadings for each tissue. [Distal heritability was significantly higher than local heritability \(t-test \$p < 2.2^{-16}\$ \)](#). **B.** Loadings of TWAS candidates with the 10 largest positive correlations with traits and the largest negative correlations with traits across all four tissues. [Local and distal heritability were not significantly different for this group \(t-test \$p = 0.77\$ \)](#). **C.** The transcripts with the largest local heritability (top 20) across all four tissues. [Local heritability was significantly higher than distal heritability of these genes \(t-test \$p < 2.2^{-16}\$ \)](#)

249 Tissue-specific transcriptional programs were associated with metabolic traits

250 Clustering of transcripts with top loadings in each tissue showed tissue-specific functional modules associated
251 with obesity and insulin resistance (Fig. 6A) (Methods). The clustering highlights the importance of immune
252 activation particularly in adipose tissue. The “mitosis” cluster had large positive loadings in three of the four
253 tissues potentially suggesting system-wide proliferation of immune cells. Otherwise, all clusters were strongly
254 loaded in only one or two tissues. For example, the lipid metabolism cluster was loaded most heavily in liver.
255 The positive loadings suggest that high expression of these genes, particularly in the liver, was associated with
256 increased metabolic disease. This cluster included the gene *Pparg*, whose primary role is in the adipose tissue
257 where it is considered a master regulator of adipogenesis⁴⁹. Agonists of *Pparg*, such as thiazolidinediones, are
258 FDA-approved to treat type II diabetes, and reduce inflammation and adipose hypertrophy⁴⁹. Consistent
259 with this role, the loading for *Pparg* in adipose tissue was negative, suggesting that higher expression was
260 associated with leaner mice (Fig. 6B). In contrast, *Pparg* had a large positive loading in liver, where it is
261 known to play a role in the development of hepatic steatosis, or fatty liver. Mice that lack *Pparg* specifically
262 in the liver, are protected from developing steatosis and show reduced expression of lipogenic genes^{50;51}.
263 Overexpression of *Pparg* in the livers of mice with a *Ppara* knockout, causes upregulation of genes involved in
264 adipogenesis⁵². In the livers of both mice and humans high *Pparg* expression is associated with hepatocytes
265 that accumulate large lipid droplets and have gene expression profiles similar to that of adipocytes^{53;54}.
266 The local and distal heritability of *Pparg* is low in adipose tissue suggesting its expression in this tissue is
267 highly constrained in the population (Fig. 6B). However, the distal heritability of *Pparg* in liver is relatively
268 high suggesting it is complexly regulated and has sufficient variation in this population to drive variation in
269 phenotype. Both local and distal heritability of *Pparg* in the islet are relatively high, but the loading is low,
270 suggesting that variability of expression in the islet does not drive variation in MDI. These results highlight
271 the importance of tissue context when investigating the role of heritable transcript variability in driving
272 phenotype.

273 Gene lists for all clusters are available in Supp. File 1.

274 Gene expression, but not local eQTLs, predicted body weight in an independent population

275 To test whether the transcript loadings identified in the DO could be translated to another population, we
276 tested whether they could predict metabolic phenotype phenotypes in an independent population of CC-RIX
277 mice, which were F1 mice derived from multiple pairings of Collaborative Cross (CC)^{55;32;56;57} strains (Fig.
278 7) (Methods). We tested two questions. First, we asked whether the loadings identified in the DO mice
279 were relevant to the relationship between the transcriptome and the phenotype in the CC-RIX. We predicted

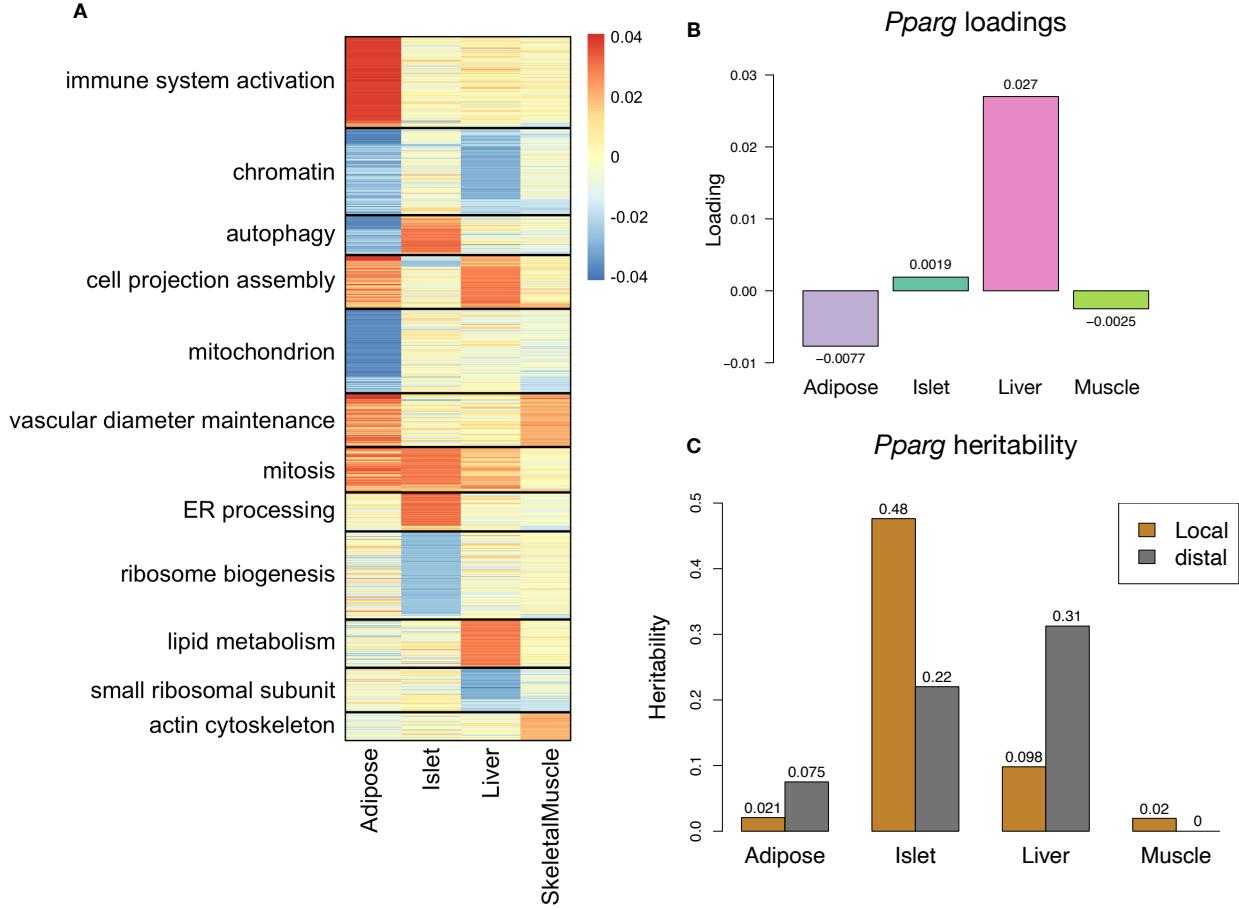


Figure 6: Tissue-specific transcriptional programs were associated with obesity and insulin resistance. **A** Heat map showing the loadings of all transcripts with loadings greater than 2.5 standard deviations from the mean in any tissue. The heat map was clustered using k medoid clustering. Functional enrichments of each cluster are indicated along the left margin. **B** Loadings for *Pparg* in different tissues. **C** Local and distal of *Pparg* expression in different tissues.

280 body weight (a surrogate for MDI) in each CC-RIX individual using measured gene expression in each tissue
 281 and the transcript loadings identified in the DO (Methods). The predicted body weight and acutal body
 282 weight were highly correlated (Fig. 7B left column). The best prediction was achieved for adipose tissue,
 283 which supports the observation in the DO that adipose expression was the strongest mediator of the genetic
 284 effect on MDI. This result also confirms the validity and translatability of the transcript loadings and their
 285 relationship to metabolic disease.

286 The second question related to the source of the relevant variation in gene expression. If local regulation was
 287 the predominant factor influencing trait-relevant gene expression, we should be able to predict phenotype in
 288 the CC-RIX using transcripts imputed from local genotype (Fig. 7A). The DO and the CC-RIX were derived
 289 from the same eight founder strains and so carry the same alleles throughout the genome. We imputed gene

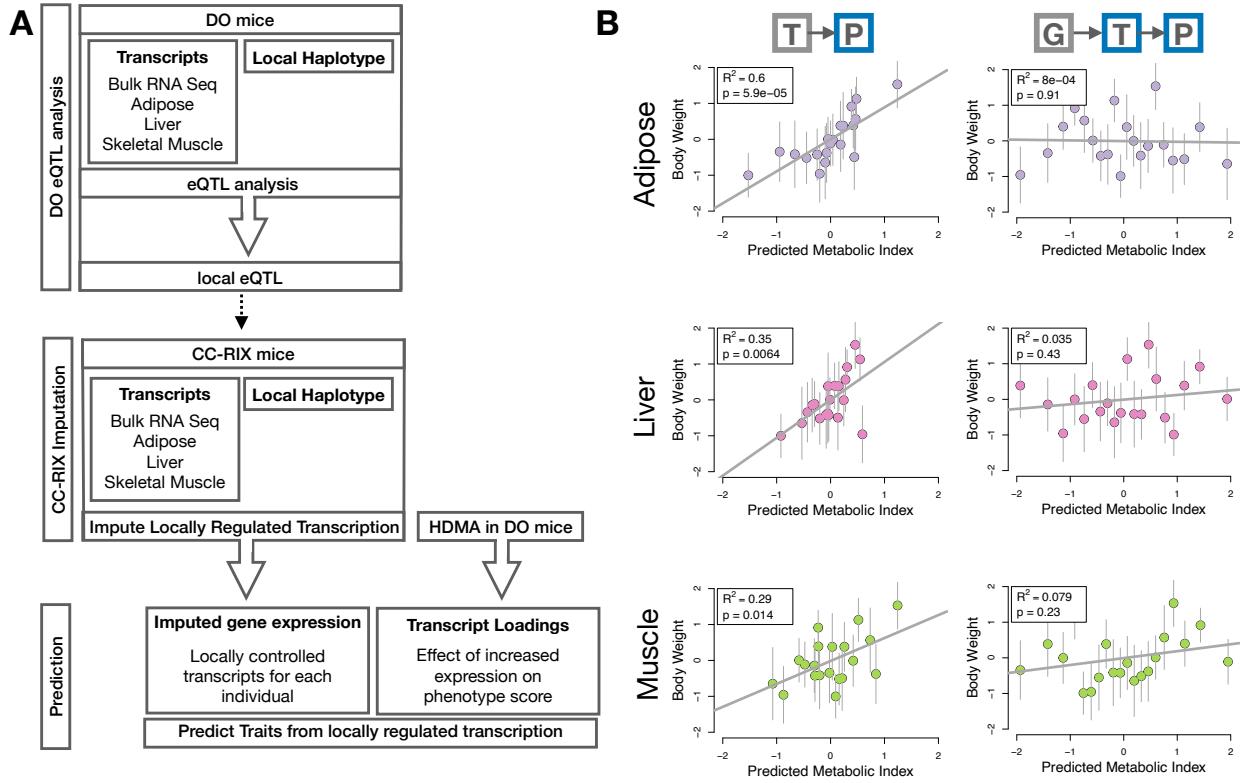


Figure 7: Transcription, but not local genotype, predicts phenotype in the CC-RIX. **A.** Workflow showing procedure for translating HDMA results to an independent population of mice. **B.** Relationships between the predicted metabolic disease index (MDI) and measured body weight [in the CC-RIX](#). The left column shows the predictions using measured transcripts. The right column shows the prediction using transcript levels imputed from local genotype. Gray boxes indicate measured quantities, and blue boxes indicate calculated quantities. The dots in each panel represent individual CC-RIX strains. The gray lines show the standard deviation on body weight for the strain.

290 expression in the CC-RIX using local genotype and were able to estimate variation in gene transcription
 291 robustly (Supp. Fig. S9). However, these imputed values failed to predict body weight in the CC-RIX when
 292 weighted with the loadings from HDMA. (Fig. 7B right column). This result suggests that local regulation of
 293 gene expression is not the primary factor driving heritability of complex traits. It is also consistent with our
 294 findings in the DO population that distal heritability was a major driver of trait-relevant gene expression and
 295 that high-loading transcripts had comparatively high distal and low local heritability.

296 **Distally heritable transcriptomic signatures reflected variation in composition of adipose tissue
 297 and islets**

298 The interpretation of global genetic influences on gene expression and phenotype is potentially more challenging
 299 than the interpretation and translation of local genetic influences, as genetic effects cannot be localized to
 300 individual gene variants or transcripts. However, there are global patterns across the loadings that can inform

mechanism. For example, heritable variation in cell type composition can be inferred from transcript loadings. We observed above that immune activation in the adipose tissue was a highly enriched process correlating with obesity in the DO population. In humans, it has been extensively observed that macrophage infiltration in adipose tissue is a marker of obesity and metabolic disease⁵⁸. To determine whether the immune activation reflected a heritable change in cell composition in adipose tissue in DO mice, we compared loadings of cell-type specific genes in adipose tissue (Methods). The mean loading of macrophage-specific genes was significantly greater than 0 ($p < 2 \times 10^{-16}$) (Fig. 8A), indicating that obese mice were genetically predisposed to have high levels of macrophage infiltration in adipose tissue in response to the HFHS diet. Loadings for marker genes for other cell types were not statistically different from zero (Adipocytes: $p = 0.08$, Progenitors: $p = 0.58$, Leukocytes: $p = 0.28$), indicating that changes in the abundance of those cell types is not a mediator of MDI.

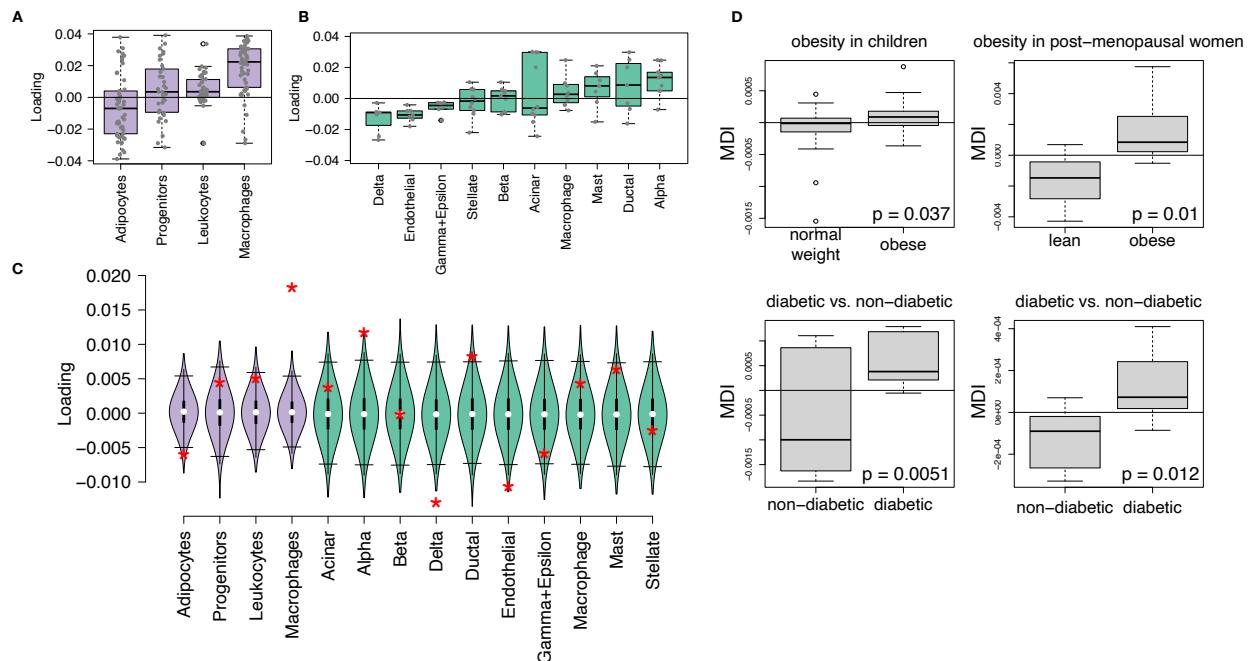


Figure 8: HDMA results translate to humans. **A.** Distribution of loadings for cell-type-specific transcripts in adipose tissue. **B.** Distribution of loadings for cell-type-specific transcripts in pancreatic islets. **C.** Null distributions for the mean loading of randomly selected transcripts in each cell type compared with the observed mean loading of each group of transcripts (red asterisk). **D.** Predictions of metabolic phenotypes in four adipose transcription data sets downloaded from GEO. In each study the obese/diabetic patients were predicted to have greater metabolic disease MDI than the lean/non-diabetic patients based on the HDMA results from DO mice.

We also compared loadings of cell-type specific transcripts in islet (Methods). The mean loadings for alpha-cell specific transcripts were significantly greater than 0 ($p = 0.002$), while the mean loadings for delta- ($p < 2 \times 10^{-16}$) and endothelial-cell ($p = 0.01$) specific genes were significantly less than 0 (Fig. 8B). These

315 results suggest that mice with higher MDI inherited an altered cell composition that predisposed them to
316 metabolic disease, or that these compositional changes were induced by the HFHS diet in a heritable way. In
317 either case, these results support the hypothesis that alterations in islet composition drive variation in MDI.
318 Notably, the mean loading for pancreatic beta cell marker transcripts was not significantly different from zero
319 ($p=0.95$). We stress that this is not necessarily reflective of the function of the beta cells in the obese mice,
320 but rather suggests that any variation in the number of beta cells in these mice was unrelated to obesity and
321 insulin resistance, the major contributors to MDI. This is further consistent with the islet composition traits
322 having small loadings in the phenome score (Fig. 4).

323 **Heritable transcriptomic signatures translated to human disease**

324 Ultimately, the heritable transcriptomic signatures that we identified in DO mice will be useful if they inform
325 mechanism and treatment of human disease. To investigate the potential for translation of the gene signatures
326 identified in DO mice, we compared them to transcriptional profiles in obese and non-obese human subjects
327 (Methods). We limited our analysis to adipose tissue because the adipose tissue signature had the strongest
328 relationship to obesity and insulin resistance in the DO.

329 We calculated a predicted MDI for each individual in the human studies based on their adipose tissue gene
330 expression (Methods) and compared the predicted scores for obese and non-obese groups as well as diabetic
331 and non-diabetic groups. In all cases, the predicted MDIs were higher on average for individuals in the
332 obese and diabetic groups compared with the lean and non-diabetic groups (Fig. 8D). This indicates that
333 the distally heritable signature of MDI identified in DO mice is relevant to obesity and diabetes in human
334 subjects.

335 **Existing therapies are predicted to target mediator gene signatures**

336 Another **potential** application of the transcript loading landscape is in ranking potential drug candidates for the
337 treatment of metabolic disease. Although high-loading transcripts may be good candidates for understanding
338 specific biology related to obesity, the transcriptome overall is highly interconnected and redundant. The
339 ConnectivityMap (CMAP) database [59-59;60](#) developed by the Broad Institute allows querying thousands of
340 compounds that reverse or enhance the extreme ends of transcriptomic signatures in multiple different cell
341 types. By identifying drugs that reverse pathogenic transcriptomic signatures, we can potentially identify
342 compounds that have favorable effects on gene expression. To test this hypothesis, we queried the CMAP
343 database through the CLUE online query tool (<https://clue.io/query/>, version 1.1.1.43) (Methods). We
344 identified top anti-correlated hits across all cell types (Supp. Figs S10 and S11). To get more tissue-specific

345 results, we also looked at top results in cell types that most closely resembled our tissues. We looked at
346 results in adipocytes (ASC) as well as pancreatic tumor cells (YAPC) regardless of *p* value (Supp. Figs S12
347 and S13).

348 Looking across all cell types, the notable top hits from the adipose tissue loadings included mTOR inhibitors
349 and glucocorticoid agonists (Supp. Fig. S10). It is thought that metformin, which is commonly used to
350 improve glycemic control, acts, at least in part, by inhibiting mTOR signaling^{61:62}. However, long-term use
351 of other mTOR inhibitors, such as rapamycin, are known to cause insulin resistance and β -cell toxicity⁶²⁻⁶⁴.
352 Glucocorticoids are used to reduce inflammation, which was a prominent signature in the adipose tissue,
353 but these drugs also promote hyperglycemia and diabetes^{65:66}. Acute treatment with glucocorticoids has
354 further been shown to reduce thermogenesis in rodent adipocytes⁶⁷⁻⁶⁹, but increase thermogenesis in human
355 adipocytes^{70:71}. Thus, the pathways identified by CMAP across all cell types were highly related to the
356 transcript loading profiles, but the relationship was not a simple reversal.

357 The top hit for the adipose composite transcript in CMAP adipocytes was a PARP inhibitor (Supp. Fig.
358 S12). PARPs play a role in lipid metabolism and are involved in the development of obesity and diabetes⁷².
359 PARP1 inhibition increases mitochondrial biogenesis⁷³. Inhibition of PARP1 activity can further prevent
360 necrosis in favor of the less inflammatory apoptosis⁷⁴, thereby potentially reducing inflammation in stressed
361 adipocytes. Other notable hits among the top 20 were BTK inhibitors, which have been observed to suppress
362 inflammation and improve insulin resistance⁷⁵. The CMAP database identified both known diabetes drugs
363 (e.g. sulfonylureas), as well as to reduce insulin antibodies in type I diabetes⁷⁶. IkappaB kinase (IKK) is an
364 enzyme complex involved in regulating cellular responses to inflammation⁷⁷. Inhibitors of IKK have been
365 shown to improve glucose control in type II diabetes^{78:79}.

366 Among the top most significant hits for the transcript loadings from pancreatic islets (Supp. Fig. S11),
367 was suppression of T cell receptor signaling, which is drugs that target pathways known to be involved in
368 Type 1 diabetes⁸⁰, as well as TNFR1, which has been associated with mortality in diabetes patients⁸¹.
369 Suppression of NOD1/2 signaling was also among the top hits. NOD1 and 2 sense ER stress^{82:83}, which
370 is associated with β -cell death in type 1 and type 2 diabetes⁸⁴. This cell death process is dependent on
371 NOD1/2 signaling⁸², although the specifics have not yet been worked out.

372 We also looked specifically at hits in pancreatic tumor cells (YAPC) regardless of significance level to get
373 a transcriptional response more specific to the pancreas (Supp. Fig. S13). Hits in this list included widely
374 used diabetes drugs, such as sulfonylureas, PPAR receptor agonists, and insulin sensitizers. Rosiglitazone is
375 a PPAR- γ agonist and was one of the most prescribed drugs for type 2 diabetes before its use was reduced

376 due to cardiac side effects⁸⁵. Sulfonylureas are another commonly prescribed drug class for type 2 diabetes,
377 but also have notable side effects including hypoglycemia and accelerated β -cell death⁸⁶.

378 In summary, the high-loading transcripts derived from HDMA in mice prioritized drugs with demonstrated
379 effectiveness in reducing type 2 diabetes phenotypes in humans in a tissue-specific manner. Drugs identified
380 using the islet loadings are known diabetes drugs that act directly on pancreatic function. Drugs identified
381 by the adipose loadings tended to reduce inflammatory responses and have been shown incidentally to
382 reduce obesity-related morbidity⁸⁷ diabetes pathogenesis (e.g. mTOR inhibitors). These findings help support
383 the mediation model we fit here. Although the composite variables we identified here are consistent with
384 mediation, they do not prove causality. However, the results from CMAP suggest that reversing the
385 transcriptomic signatures we found also reverses metabolic disease phenotypes, which supports a causal
386 role of the transcript levels in driving pathogenesis of metabolic disease. These results thus support the
387 mediation model we identified here and its translation to therapies in human disease.

388 Discussion

389 Here we investigated the relative contributions of local and distal gene regulation in four tissues to heritable
390 variation in traits related to metabolic disease in genetically diverse mice. We found that distal heritability
391 was positively correlated with trait relatedness, whereas high local heritability was negatively correlated with
392 trait relatedness. We used a novel high-dimensional mediation analysis (HDMA) to identify tissue-specific
393 composite transcripts that are predicted to mediate the effect of genetic background on metabolic traits. The
394 adipose-derived composite transcript robustly predicted body weight in an independent cohort of diverse
395 mice with disparate population structure, as well as to humans. It also predicted MDI in four human
396 cohorts. However, gene expression imputed from local genotype failed to predict body weight in the second
397 mouse population. Taken together, these results highlight the complexity of gene expression regulation in
398 relation to trait heritability and suggest that heritable trait variation is mediated primarily through distal
399 gene regulation.

400 Supplemental Discussion

401 Our result that distal regulation accounted for most trait-related gene expression differences is consistent
402 with a complex model of genetic trait determination. It has frequently been assumed that gene regulation in
403 *cis* is the primary driver of genetically associated trait variation, but attempts to use local gene regulation
404 to explain phenotypic variation have had limited success^{16;17}. In recent years, evidence has mounted that
405 distal gene regulation may be an important mediator of trait heritability^{19;18;87;88}. It has been observed

406 that transcripts with high local heritability explain less expression-mediated disease heritability than those
407 with low local heritability¹⁹. Consistent with this observation, genes located near GWAS hits tend to be
408 complexly regulated¹⁸. They also tend to be enriched with functional annotations, in contrast to genes with
409 simple local regulation, which tend to be depleted of functional annotations suggesting they are less likely
410 to be directly involved in disease traits¹⁸. These observations are consistent with principles of robustness
411 in complex systems in which simple regulation of important elements leads to fragility of the system^{89–91}.
412 Our results are consistent, instead, with a more complex picture where genes whose expression can drive
413 trait variation are buffered from local genetic variation but are extensively influenced indirectly by genetic
414 variation in the regulatory networks converging on those genes.

415 Our results are also consistent with the recently proposed omnigenic model, which posits that complex traits
416 are massively polygenic and that their heritability is spread out across the genome⁹². In the omnigenic model,
417 genes are classified either as “core genes,” which directly impinge on the trait, or “peripheral genes,” which
418 are not directly trait-related, but influence core genes through the complex gene regulatory network. HDMA
419 explicitly models a central proposal of the omnigenic model which posits that once the expression of the
420 core genes (i.e. trait-mediating genes) is accounted for, there should be no residual correlation between the
421 genome and the phenome. Here, we were able to fit this model and identified a composite transcript that,
422 when taken into account, left no residual correlation between the composite genome and composite phenome
423 scores (Fig. 3A, [Supp. Fig S3](#)).

424 Unlike in the omnigenic model, we did not observe a clear demarcation between the core and peripheral
425 genes in loading magnitude, but we do not necessarily expect a clear separation given the complexity of gene
426 regulation and the genotype-phenotype map⁹³.

427 An extension of the omnigenic model proposed that most heritability of complex traits is driven by weak
428 distal eQTLs that are potentially below the detection threshold in studies with feasible sample sizes⁸⁷. This
429 is consistent with what we observed here. For example, *Nucb2*, had a high loading in islets and was also
430 strongly distally regulated (66% distal heritability) (Fig. 5). This gene is expressed in pancreatic β cells and
431 is involved in insulin and glucagon release^{94–96}. Although its transcription was highly heritable in islets, that
432 regulation was distributed across the genome, with no clear distal eQTL (Supp. Fig. S14). Thus, although
433 distal regulation of some genes may be strong, this regulation is likely to be highly complex and not easily
434 localized.

435 Individual high-loading transcripts also demonstrated biologically interpretable, tissue-specific patterns. We
436 highlighted *Pparg*, which is known to be protective in adipose tissue⁴⁹ where it was negatively loaded, and

437 harmful in the liver^{50–54}, where it was positively loaded. Such granular patterns may be useful in generating
438 hypotheses for further testing, and prioritizing genes as therapeutic targets. The tissue-specific nature of
439 the loadings also may provide clues to tissue-specific effects, or side effects, of targeting particular genes
440 system-wide.

441 In addition to identifying individual transcripts of interest, the composite transcripts can be used as weighted
442 vectors in multiple types of analysis, such as drug prioritization using gene set enrichment analysis (GSEA)
443 and the CMAP database. In particular, the CMAP analysis identified drugs which have been demonstrated
444 to reverse insulin resistance and other aspects of metabolic disease. This finding supports the ~~eausal role~~
445 ~~of these full gene signatures~~ hypothesis that HDMA identified transcripts that truly mediate genetic effects
446 on traits. On its own, HDMA identifies transcriptional patterns that are consistent with a mediation model,
447 but alone does not prove mediation. However, the finding that these drugs act both on the transcriptional
448 patterns and on the desired traits support the mediation model and the hypothesis that these transcripts
449 have a causal role in pathogenesis of metabolic disease ~~and thus their utility in prioritizing drugs and gene~~
450 ~~targets as therapeutics.~~

451 Together, our results have shown that both tissue specificity and distal gene regulation are critically important
452 to understanding the genetic architecture of complex traits. We identified important genes and gene signatures
453 that were heritable, plausibly causal of disease, and translatable to other mouse populations and to humans.
454 Finally, we have shown that by directly acknowledging the complexity of both gene regulation and the
455 genotype-to-phenotype map, we can gain a new perspective on disease pathogenesis and develop actionable
456 hypotheses about pathogenic mechanisms and potential treatments.

457 Data and Code Availability

458 **DO mice:** Genotypes, phenotypes, and pancreatic islet gene expression data were previously published¹².
459 Gene expression for the other tissues can be found at the Gene Expression Omnibus <https://www.ncbi.nlm.nih.gov/geo/> with the following accession numbers: DO adipose tissue - GSE266549; DO liver tissue
460 - GSE266569; DO skeletal muscle - GSE266567. Expression data with calculated eQTLs are available at
461 Figshare https://figshare.com/articles/dataset/Data_and_code_for_High-Dimensional_Mediation_Anal
ysis_HDMA_in_diversity_outbred_mice/27066979 DOI: 10.6084/m9.figshare.27066979
464 [10.6084/m9.figshare.27066979.v1](https://doi.org/10.6084/m9.figshare.27066979.v1)

465 **CC-RIX mice:** Gene expression can be found at the Gene Expression Omnibus <https://www.ncbi.nlm.nih.gov/geo/> with the following accession numbers: CC-RIX adipose tissue - GSE237737; CC-RIX liver tissue -

467 GSE237743; CC-RIX skeletal muscle - GSE237747. Count matrices and phenotype data can be found at
468 Figshare https://figshare.com/articles/dataset/Data_and_code_for_High-Dimensional_Mediation_Anal
469 ysis_HDMA_in_diversity_outbred_mice/27066979 DOI: 10.6084/m9.figshare.27066979

470 **Code:** All code used to run the analyses reported here are available at Figshare: https://figshare.com/articles/dataset/Data_and_code_for_High-Dimensional_Mediation_Analysis_HDMA_in_diversity_outbred_mice/27066979 DOI: 10.6084/m9.figshare.27066979

473 Acknowledgements

474 This project was supported by The Jackson Laboratory Cube Initiative, as well as grants from the National
475 Institutes of Health (grant numbers.: R01DK101573, R01DK102948, and RC2DK125961) (to A. D. A.) and
476 by the University of Wisconsin-Madison, Department of Biochemistry and Office of the Vice Chancellor for
477 Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation (to M. P.
478 K.).

479 We thank the following scientific services at The Jackson Laboratory: Genome Technologies for the RNA
480 sequencing, necropsy services for the tissue harvests, and the Center for Biometric Analysis for metabolic
481 phenotyping.

482 **Supplemental Figures**

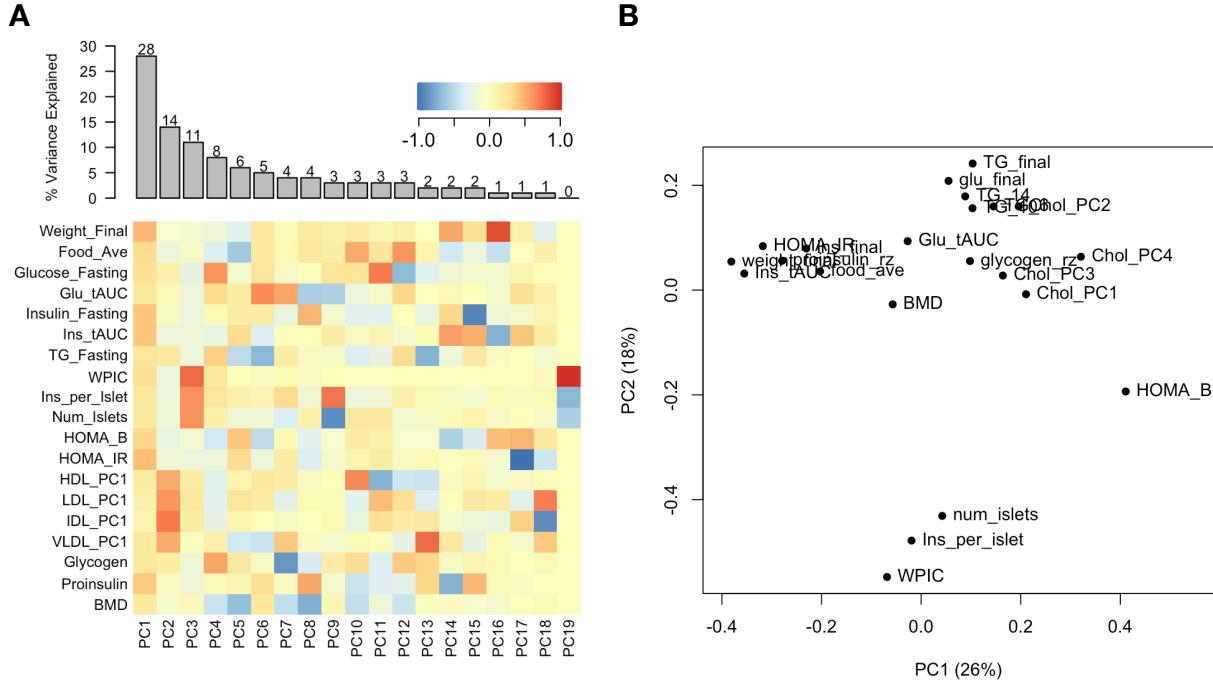


Figure S1: Trait matrix decomposition. **A** The heat map shows the loadings of each trait onto each principal component of the trait matrix. The bars at the top show the percent variance explained for each principal component. **B** Traits plotted by the first and second principal components of the trait matrix. This view shows clustering of traits into insulin- and weight-related traits, lipid-related traits, and ex-vivo pancreatic measurements.

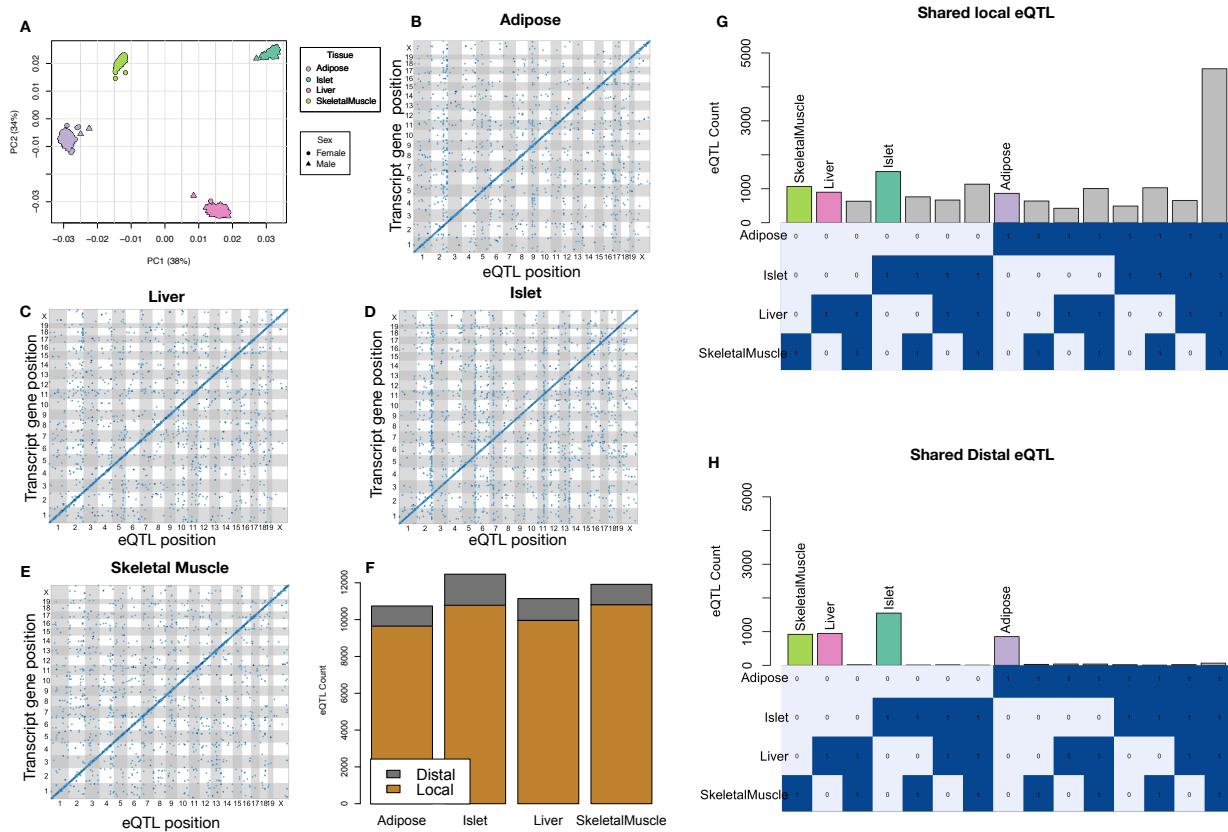


Figure S2: Overview of eQTL analysis in DO mice. **A.** RNA seq samples from the four different tissues clustered by tissue. **B.-E.** eQTL maps are shown for each tissue. The *x*-axis shows the position of the mapped eQTL, and the *y*-axis shows the physical position of the gene encoding each mapped transcript. Each dot represents an eQTL with a minimum LOD score of 8. The dots on the diagonal are locally regulated eQTL for which the mapped eQTL is at the within 4Mb of the encoding gene. Dots off the diagonal are distally regulated eQTL for which the mapped eQTL is distant from the gene encoding the transcript. **F.** Comparison of the total number of local and distal eQTL with a minimum LOD score of 8 in each tissue. All tissues have comparable numbers of eQTL. Local eQTLs are much more numerous than distal eQTL. **G.** Counts of transcripts with local eQTL shared across multiple tissues. The majority of local eQTLs were shared across all four tissues. **H.** Counts of transcripts with distal eQTL shared across multiple tissues. The majority of distal eQTL were tissue-specific and not shared across multiple tissues. For both G and H, eQTL for a given transcript were considered shared in two tissues if they were within 4Mb of each other. Colored bars indicate the counts for individual tissues for easy of visualization.

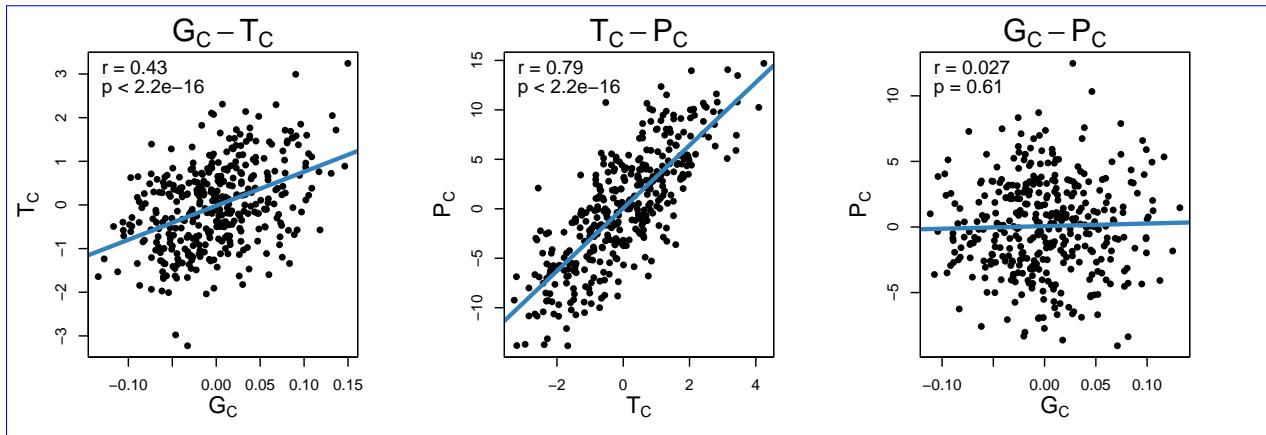


Figure S3: Scatter plots showing correlations between composite vectors for the genome (G_C), the transcriptome (T_C), and the phenome (P_C). The $G_C - T_C$ correlation is high, the $T_C - P_C$ correlation is high, and there is no significant correlation between G_C and P_C . This correlation structure is consistent with perfect mediation.

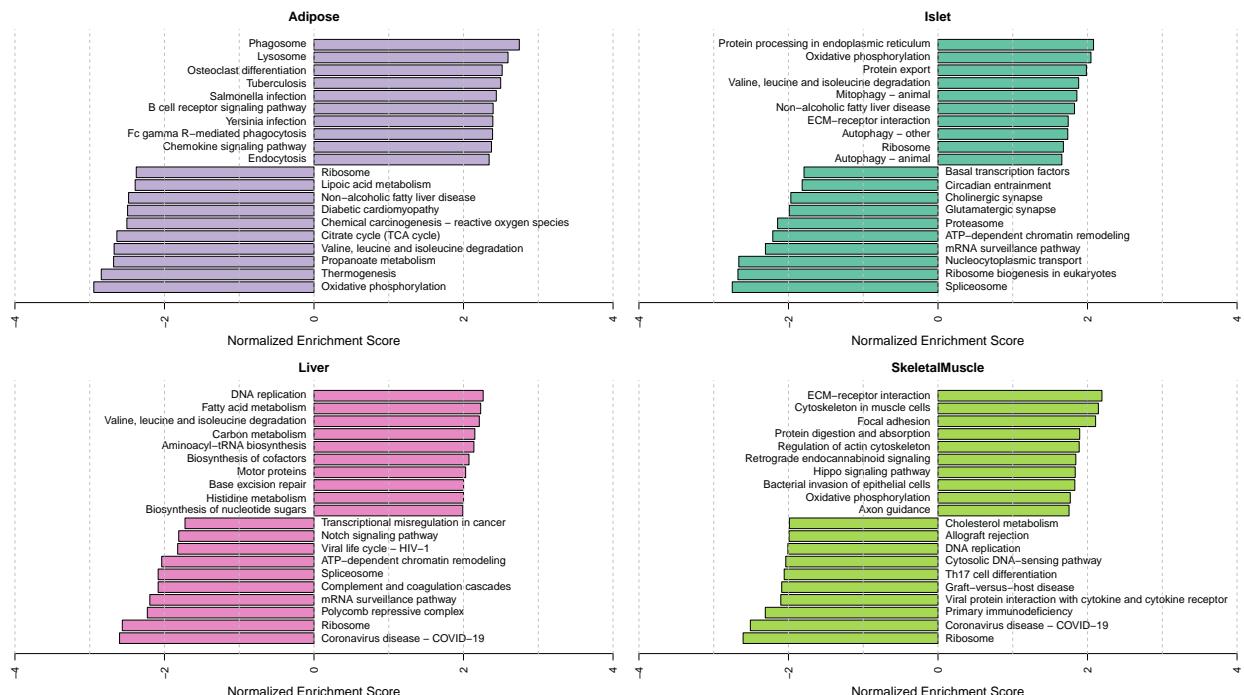


Figure S4: Bar plots showing normalized enrichment scores (NES) for KEGG pathways as determined by fast gene score enrichment analysis (fgsea). Only the top 10 positive and top 10 negative scores are shown. Colors indicate tissue. The name beside each bar shows the name of each enriched KEGG pathway.

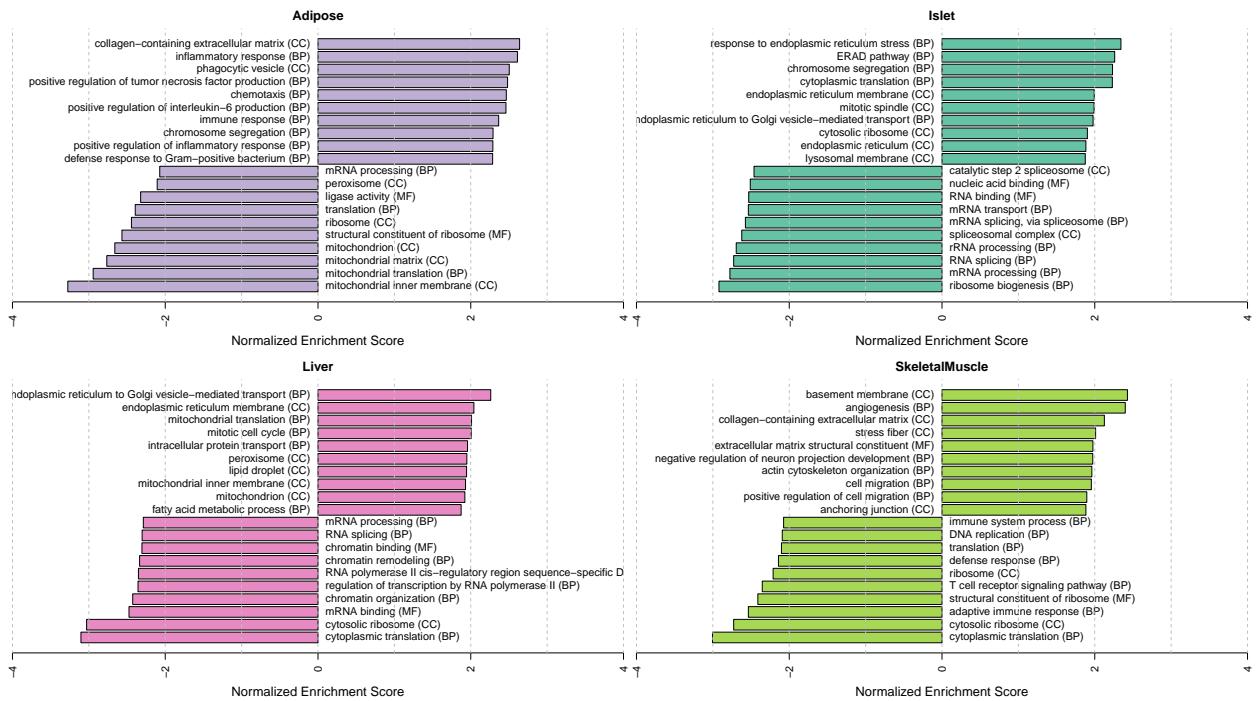


Figure S5: Bar plots showing normalized enrichment scores (NES) for GO terms as determined by fast gene score enrichment analysis (fgsea). Only the top 10 positive and top 10 negative scores are shown. Colors indicate tissue. The name beside each bar shows the name of each enriched GO term. The letters in parentheses indicate whether the term is from the biological process ontology (BP), the molecular function ontology (MF), or the cellular compartment ontology (CC).

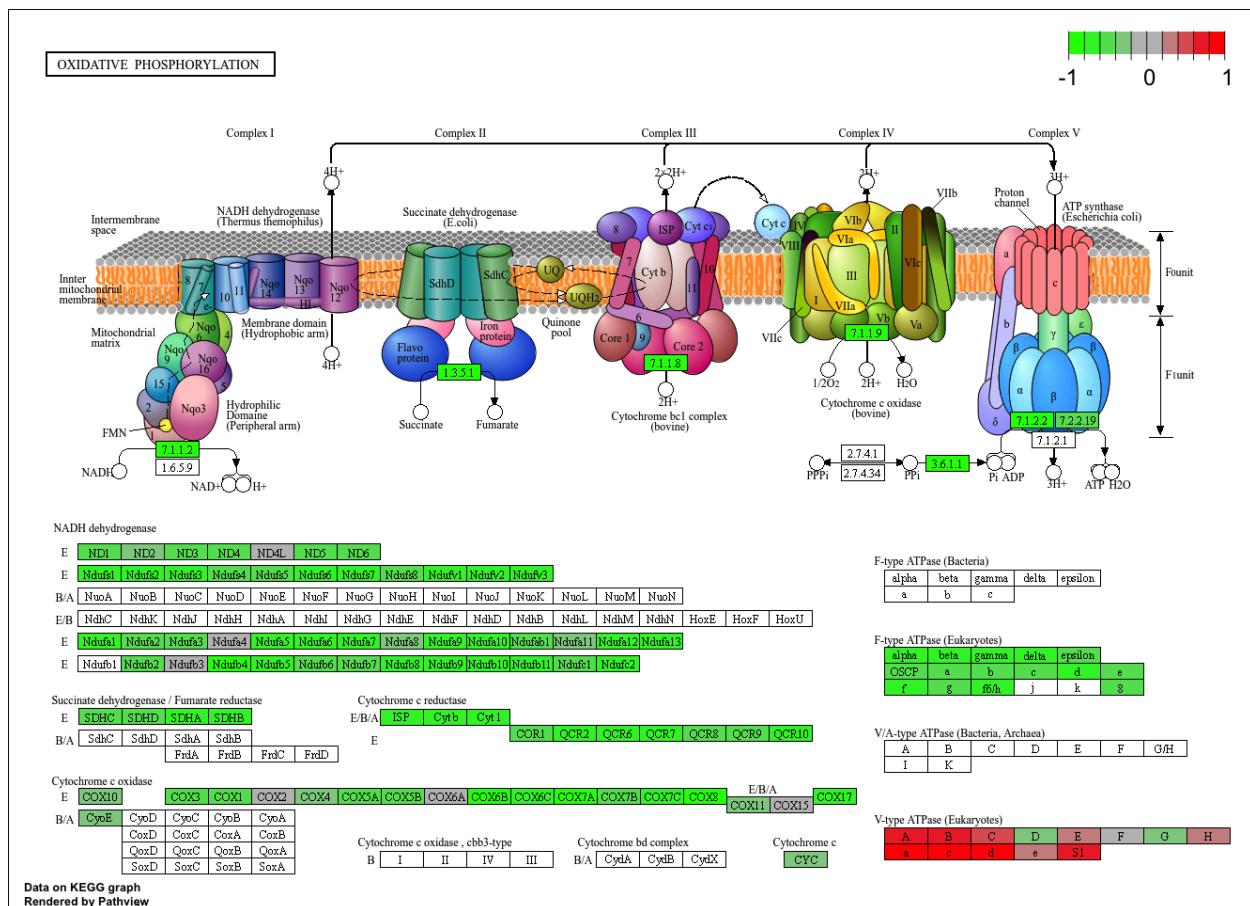


Figure S6: The KEGG pathway for oxidative phosphorylation in mice. Each element is colored based on its HDMA loading from adipose tissue normalized scaled to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Almost the entire pathway was strongly negatively loaded indicating that increased expression of genes involved in oxidative phosphorylation was associated with reduced MDI.

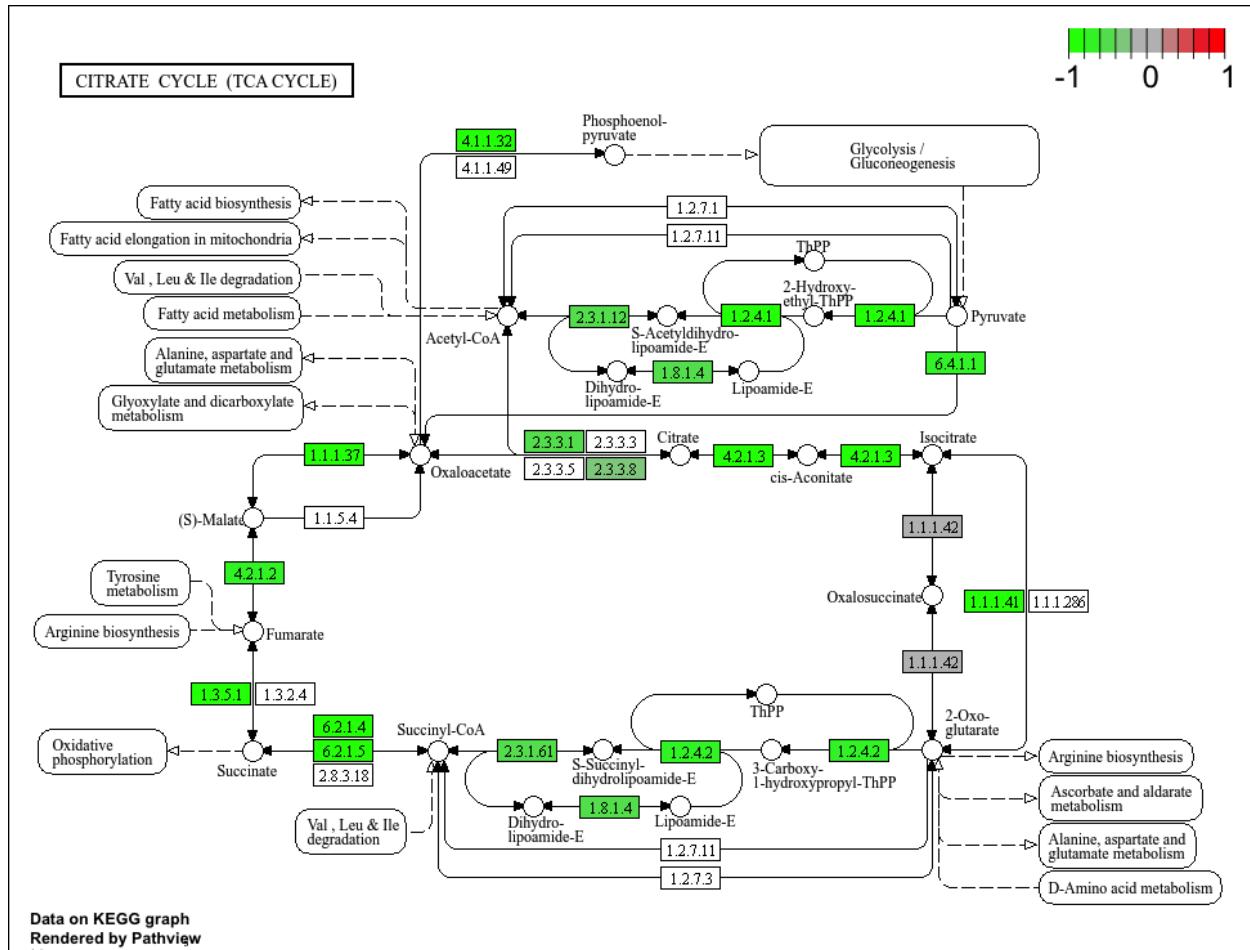


Figure S7: The KEGG pathway for the TCA (citric acid) cycle in mice. Each element is colored based on its HDMA loading from adipose tissue [normalized-scaled](#) to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Many genes in the cycle were strongly negatively loaded indicating that increased expression of genes involved in [branched-chain-amino acid degradation](#) [the TCA cycle](#) was associated with reduced MDI.

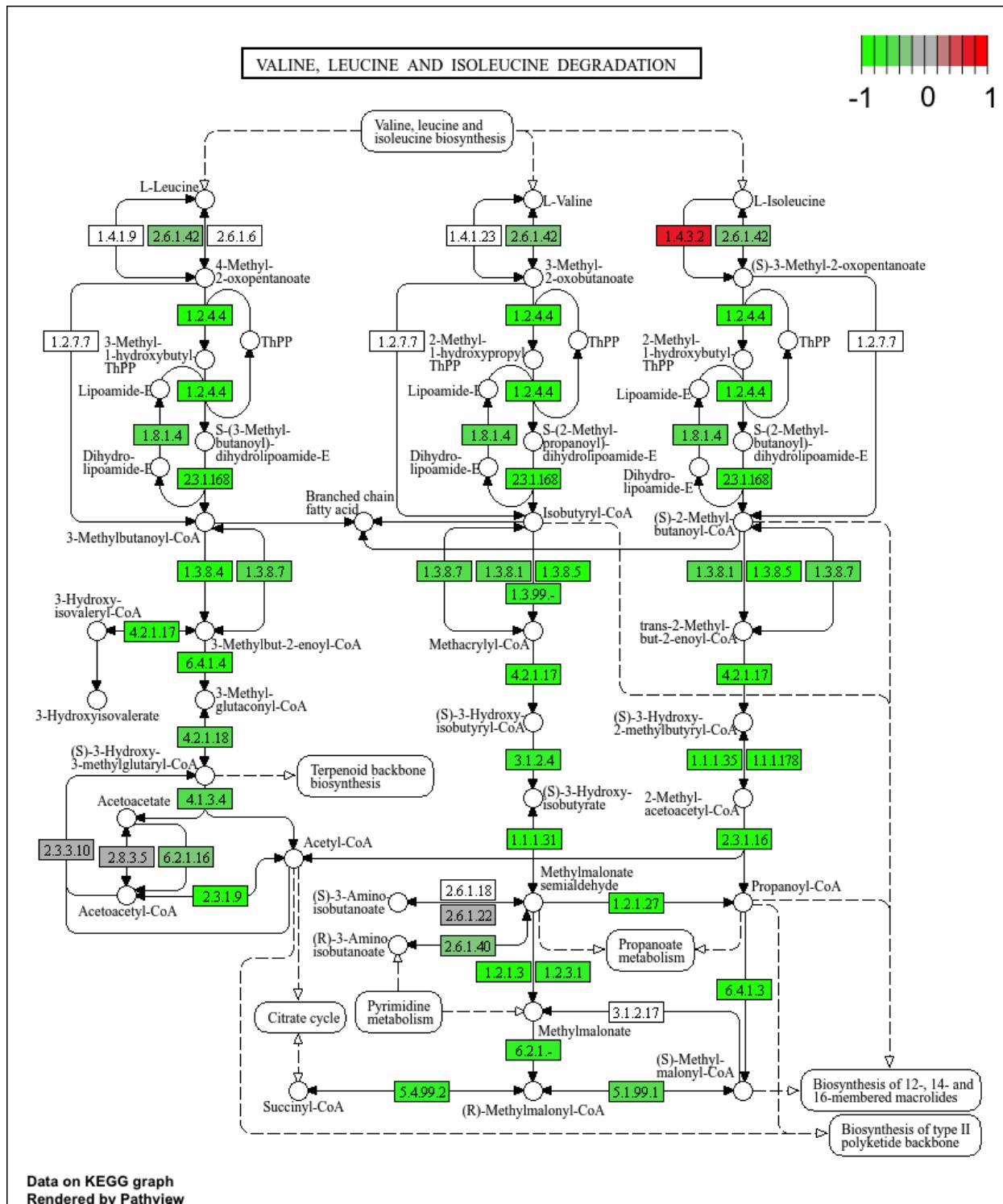


Figure S8: The KEGG pathway for branched-chain amino acid degradation in mice. Each element is colored based on its HDMA loading from adipose tissue normalized scaled to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Almost the entire pathway was strongly negatively loaded indicating that increased expression of genes involved in branched-chain amino acid degradation was associated with reduced MDI.

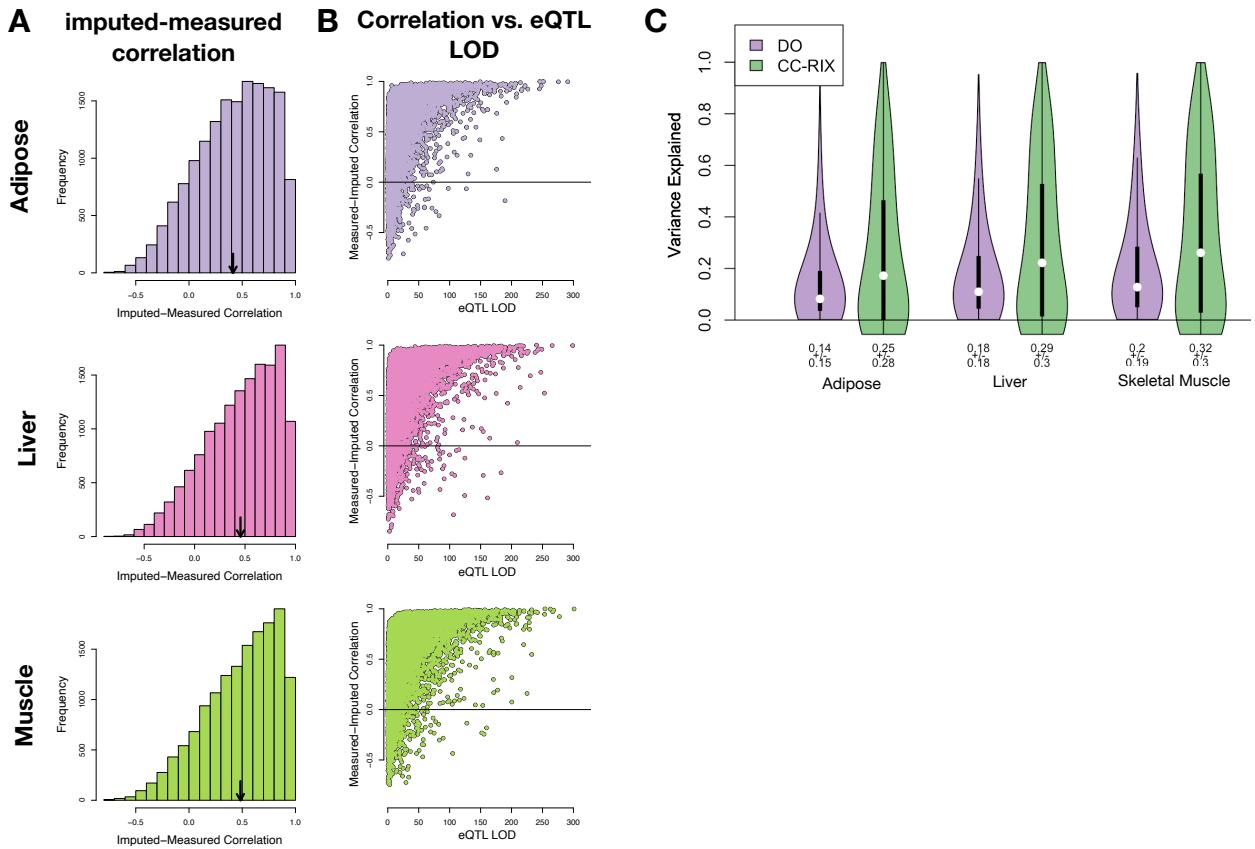


Figure S9: Validation of transcript imputation in the CC-RIX. **A.** Distributions of correlations between imputed and measured transcripts in the CC-RIX. The mean of each distribution is shown by the red line. All distributions were skewed toward positive correlations and had positive means near a Pearson correlation (r) of 0.5. **B.** The relationship between the correlation between measured and imputed expression in the CC-RIX (x-axis) and eQTL LOD score. As expected, imputations are more accurate for transcripts with strong local eQTLs. **C.** Variance Distributions of variance explained by local genotype across all transcripts in the DO and CC-RIX.

| id | norm_ss | cell_iname | pert_type | raw_ss▲ | fdr_q_nlog10 | set_type | src_set_id |
|-----------|----------------|-------------------|------------------|----------------|---------------------|-----------------|-----------------------------------|
| | | HA1E | TRT_CP | -0.97 | 15.65 | PCL | CP_PROTEIN_SYNTHESIS_INHIBITOR |
| | | PC3 | TRT_SH.CGS | -0.90 | 15.65 | PATHWAY_SET | BIOCARTA_EIF4_PATHWAY |
| | | A375 | TRT_CP | -0.87 | 15.65 | MOA_CLASS | RAF_INHIBITOR |
| | | HCC515 | TRT_CP | -0.84 | 15.65 | PCL | CP_TOPOISOMERASE_INHIBITOR |
| | | HEPG2 | TRT_SH.CGS | -0.82 | 15.65 | PATHWAY_SET | BIOCARTA_BCR_PATHWAY |
| | | PC3 | TRT_CP | -0.77 | 15.65 | MOA_CLASS | MTOR_INHIBITOR |
| | | HCC515 | TRT_CP | -0.76 | 15.65 | PCL | CP_GLUCOCORTICOID_RECECTORAGONIST |
| | | HCC515 | TRT_CP | -0.76 | 15.65 | MOA_CLASS | GLUCOCORTICOID_RECECTORAGONIST |
| | | A375 | TRT_CP | -0.72 | 15.65 | MOA_CLASS | MTOR_INHIBITOR |
| | | -666 | TRT_CP | -0.70 | 15.65 | PCL | CP_PROTEIN_SYNTHESIS_INHIBITOR |
| | | -666 | TRT_CP | -0.68 | 15.65 | PCL | CP_JAK_INHIBITOR |
| | | A549 | TRT_CP | -0.67 | 15.65 | PCL | CP_GLUCOCORTICOID_RECECTORAGONIST |
| | | A549 | TRT_CP | -0.67 | 15.65 | MOA_CLASS | GLUCOCORTICOID_RECECTORAGONIST |
| | | -666 | TRT_CP | -0.57 | 15.65 | PCL | CP_MTOR_INHIBITOR |
| | | -666 | TRT_CP | -0.55 | 15.65 | MOA_CLASS | MTOR_INHIBITOR |
| | | -666 | TRT_CP | -0.55 | 15.65 | PCL | CP_PI3K_INHIBITOR |
| | | -666 | TRT_CP | 0.85 | 15.65 | MOA_CLASS | PKC_ACTIVATOR |

Figure S10: CMAP results using the *adipose* tissue composite transcript as an input. Table includes results from *all cell types* sorted with a $-\log_{10}(q) > 15$. The results are sorted by the correlation of the query to the input with the most negative results at the top.

| id | norm_CS | cell_iname | pert_type | raw_CS▲ | fdr_q_nlog10 | set_type | src_set_id |
|------|---------|------------|------------|---------|--------------|--|------------|
| | | VCAP | TRT_SH.CGS | -0.99 | 15.65 | PATHWAY_SET REACTOME_DOWNSTREAM_TCR_SIGNALING | |
| | | VCAP | TRT_SH.CGS | -0.99 | 15.65 | PATHWAY_SET REACTOME_NOD1_2_SIGNALING_PATHWAY | |
| | | A549 | TRT_SH.CGS | -0.92 | 15.65 | PATHWAY_SET BIOCARTA_TNFR1_PATHWAY | |
| | | VCAP | TRT_SH.CGS | -0.92 | 15.65 | PATHWAY_SET HALLMARK_WNT_BETA_CATENIN_SIGNALING | |
| | | HT29 | TRT_CP | -0.92 | 15.65 | PCL CP_TUBULIN_INHIBITOR | |
| -666 | | | TRT_OE | -0.88 | 15.65 | PCL OE_CELL_CYCLE_INHIBITION | |
| | | VCAP | TRT_SH.CGS | -0.87 | 15.65 | PATHWAY_SET REACTOME_P75_NTR_RECECTOR_MEDIATED_SIGNALLING | |
| | | HT29 | TRT_CP | -0.86 | 15.65 | MOA_CLASS TUBULIN_INHIBITOR | |
| | | MCF7 | TRT_CP | -0.85 | 15.65 | PCL CP_TUBULIN_INHIBITOR | |
| -666 | | | TRT_CP | -0.81 | 15.65 | PCL CP_PROTEASOME_INHIBITOR | |
| -666 | | | TRT_SH.CGS | -0.80 | 15.65 | PATHWAY_SET REACTOME_DOWNREGULATION_OF_ERBB2_ERBB3_SIGNALING | |
| | | HCC515 | TRT_CP | -0.80 | 15.65 | PCL CP_GLUCOCORTICOID_RECECTORAGONIST | |
| | | HCC515 | TRT_CP | -0.80 | 15.65 | MOA_CLASS GLUCOCORTICOID_RECECTORAGONIST | |
| | | A549 | TRT_OE | -0.78 | 15.65 | PATHWAY_SET REACTOME_RAF_MAP_KINASE CASCADE | |
| | | A549 | TRT_OE | -0.78 | 15.65 | PATHWAY_SET PID_RAS_PATHWAY | |
| -666 | | | TRT_SH.CGS | -0.78 | 15.65 | PCL KD_RIBOSOMAL_40S_SUBUNIT | |
| | | A549 | TRT_OE | -0.76 | 15.65 | PATHWAY_SET REACTOME_SIGNALLING_TO_P38_VIA_RIT_AND_RIN | |
| | | A549 | TRT_OE | -0.76 | 15.65 | PATHWAY_SET REACTOME_PROLONGED_ERK_ACTIVATION_EVENTS | |
| | | A549 | TRT_OE | -0.73 | 15.65 | PATHWAY_SET PID_TCR_RAS_PATHWAY | |
| | | HA1E | TRT_OE | -0.73 | 15.65 | PATHWAY_SET REACTOME_SHC RELATED_EVENTS | |
| | | HA1E | TRT_OE | -0.71 | 15.65 | PATHWAY_SET PID_EPHB_FWD_PATHWAY | |
| -666 | | | TRT_CP | -0.70 | 15.65 | MOA_CLASS GLYCOGEN_SYNTHASE_KINASE_INHIBITOR | |
| | | HA1E | TRT_OE | -0.70 | 15.65 | PATHWAY_SET PID_GMCSF_PATHWAY | |
| | | A549 | TRT_OE | -0.69 | 15.65 | PATHWAY_SET REACTOME_SIGNALLING_TO_ERKS | |
| -666 | | | TRT_LIG | -0.69 | 15.65 | PATHWAY_SET PID_ERBB_NETWORK_PATHWAY | |
| -666 | | | TRT_CP | -0.67 | 15.65 | MOA_CLASS PROTEASOME_INHIBITOR | |
| -666 | | | TRT_CP | -0.66 | 15.65 | PCL CP_GLYCOGEN_SYNTHASE_KINASE_INHIBITOR | |
| -666 | | | TRT_CP | 0.73 | 15.65 | MOA_CLASS MTOR_INHIBITOR | |

Figure S11: CMAP results using the *pancreatic islet* tissue composite transcript as an input. Table includes results from *all cell types* sorted with a $-\log_{10}(q) > 15$. The results are sorted by the correlation of the query to the input with the most negative results at the top.

| id | norm_ss | cell_iname | pert_type | raw_ss ▲ | fdr_q_nlog10 | set_type | src_set_id |
|-----------|----------------|-------------------|------------------|-----------------|---------------------|-----------------|---|
| | | ASC | TRT_CP | -0.94 | 0.79 | PCL | CP_PARP_INHIBITOR |
| | | ASC | TRT_CP | -0.94 | 0.79 | MOA_CLASS | PROTEIN_TYROSINE_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.84 | 0.45 | MOA_CLASS | BTK_INHIBITOR |
| | | ASC | TRT_CP | -0.81 | 0.39 | MOA_CLASS | LEUCINE_RICH_REPEAT_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.81 | 0.79 | PCL | CP_HSP_INHIBITOR |
| | | ASC | TRT_CP | -0.80 | 0.93 | PCL | CP_EGFR_INHIBITOR |
| | | ASC | TRT_CP | -0.79 | 0.32 | MOA_CLASS | T-TYPE_CALCIUM_CHANNEL_BLOCKER |
| | | ASC | TRT_CP | -0.79 | 1.09 | PCL | CP_MTOR_INHIBITOR |
| | | ASC | TRT_CP | -0.76 | 0.97 | PCL | CP_PI3K_INHIBITOR |
| | | ASC | TRT_CP | -0.75 | 0.20 | MOA_CLASS | HISTONE_DEMETHYLASE_INHIBITOR |
| | | ASC | TRT_CP | -0.74 | 0.42 | PCL | CP_IKK_INHIBITOR |
| | | ASC | TRT_CP | -0.74 | 0.83 | PCL | CP_AURORA_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.74 | 0.17 | PCL | CP_LEUCINE_RICH_REPEAT_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.72 | 0.36 | PCL | CP_BROMODOMAIN_INHIBITOR |
| | | ASC | TRT_CP | -0.71 | 1.09 | MOA_CLASS | TYROSINE_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.70 | 0.82 | PCL | CP_PROTEIN_SYNTHESIS_INHIBITOR |
| | | ASC | TRT_CP | -0.67 | 0.69 | PCL | CP_SRC_INHIBITOR |
| | | ASC | TRT_CP | -0.67 | 0.81 | MOA_CLASS | AURORA_KINASE_INHIBITOR |
| | | ASC | TRT_CP | -0.65 | 0.89 | MOA_CLASS | FLT3_INHIBITOR |
| | | ASC | TRT_CP | -0.62 | 0.40 | MOA_CLASS | FGFR_INHIBITOR |
| | | ASC | TRT_CP | -0.59 | 0.66 | MOA_CLASS | MEK_INHIBITOR |
| | | ASC | TRT_CP | -0.59 | 0.13 | MOA_CLASS | SYK_INHIBITOR |
| | | ASC | TRT_CP | -0.58 | 0.01 | PCL | CP_PKC_INHIBITOR |
| | | ASC | TRT_CP | -0.58 | 0.65 | PCL | CP_HDAC_INHIBITOR |
| | | ASC | TRT_CP | -0.58 | 0.65 | PCL | CP_ATPASE_INHIBITOR |
| | | ASC | TRT_CP | -0.53 | 0.09 | PCL | CP_FLT3_INHIBITOR |
| | | ASC | TRT_CP | -0.53 | 0.42 | PCL | CP_P38_MAPK_INHIBITOR |
| | | ASC | TRT_CP | -0.53 | 0.22 | MOA_CLASS | IKK_INHIBITOR |
| | | ASC | TRT_CP | -0.52 | 0.58 | PCL | CP_VEGFR_INHIBITOR |
| | | ASC | TRT_CP | -0.51 | -0.00 | PCL | CP_T-TYPE_CALCIUM_CHANNEL_BLOCKER |

Figure S12: CMAP results using the *adipose* tissue composite transcript as an input. Table includes the top 30 results derived *only from normal adipocytes* (ASC) regardless of significance. The results are sorted by the correlation of the query to the input with the most negative results at the top.

| id | norm_CS | cell_iname | pert_type | raw_CS | fdr_q_nlog10 | set_type | src_set_id |
|----|---------|------------|-----------|--------|--------------|-------------|---------------------------------------|
| | | YAPC | TRT_CP | -1.00 | 0.67 | MOA_CLASS | ABL_KINASE_INHIBITOR |
| | | YAPC | TRT_CP | -0.99 | 0.66 | PCL | CP_CDK_INHIBITOR |
| | | YAPC | TRT_CP | -0.97 | 1.41 | PCL | CP_TOPOISOMERASE_INHIBITOR |
| | | YAPC | TRT_CP | -0.95 | 0.70 | MOA_CLASS | THYMIDYLATE_SYNTHASE_INHIBITOR |
| | | YAPC | TRT_CP | -0.95 | 0.62 | MOA_CLASS | ADRENERGIC_INHIBITOR |
| | | YAPC | TRT_CP | -0.94 | 0.50 | MOA_CLASS | BENZODIAZEPINE_RECECTOR_ANTAGONIST |
| | | YAPC | TRT_CP | -0.89 | 0.63 | PCL | CP_RIBONUCLEOTIDE_REDUCTASE_INHIBITOR |
| | | YAPC | TRT_CP | -0.88 | 0.52 | MOA_CLASS | VASOPRESSIN_RECECTOR_ANTAGONIST |
| | | YAPC | TRT_CP | -0.85 | 0.63 | MOA_CLASS | ANGIOTENSIN_RECECTOR_ANTAGONIST |
| | | YAPC | TRT_CP | -0.85 | 0.33 | PCL | CP_CANNABINOID_RECECTORAGONIST |
| | | YAPC | TRT_CP | -0.84 | 0.30 | PCL | CP_RETINOID_RECECTORAGONIST |
| | | YAPC | TRT_CP | -0.83 | 1.19 | MOA_CLASS | NFKB_PATHWAY_INHIBITOR |
| | | YAPC | TRT_CP | -0.83 | 0.54 | MOA_CLASS | DNA_ALKYLATING_DRUG |
| | | YAPC | TRT_CP | -0.80 | 0.50 | MOA_CLASS | CHOLESTEROL_INHIBITOR |
| | | YAPC | TRT_CP | -0.79 | 0.15 | MOA_CLASS | SULFONYLUREA |
| | | YAPC | TRT_CP | -0.78 | 0.52 | MOA_CLASS | HIV_INTEGRASE_INHIBITOR |
| | | YAPC | TRT_CP | -0.78 | 0.13 | MOA_CLASS | LEUKOTRIENE_INHIBITOR |
| | | YAPC | TRT_CP | -0.78 | 0.45 | PCL | CP_PPAR_RECECTORAGONIST |
| | | YAPC | TRT_CP | -0.78 | 0.54 | MOA_CLASS | INSULIN_SENSITIZER |
| | | YAPC | TRT_CP | -0.77 | 0.51 | MOA_CLASS | ESTROGEN_RECECTOR_ANTAGONIST |
| | | YAPC | TRT_CP | -0.77 | 0.76 | MOA_CLASS | DNA_SYNTHESIS_INHIBITOR |
| | | YAPC | TRT_XPR | -0.77 | 0.67 | PATHWAY_SET | BIOCARTA_PARKIN_PATHWAY |
| | | YAPC | TRT_CP | -0.77 | 0.51 | PCL | CP_VEGFR_INHIBITOR |
| | | YAPC | TRT_CP | -0.75 | 0.39 | MOA_CLASS | RNA_SYNTHESIS_INHIBITOR |
| | | YAPC | TRT_CP | -0.72 | 0.60 | MOA_CLASS | BCR-ABL_KINASE_INHIBITOR |
| | | YAPC | TRT_XPR | -0.71 | 0.66 | PATHWAY_SET | BIOCARTA_EIF_PATHWAY |
| | | YAPC | TRT_XPR | -0.69 | 0.54 | PATHWAY_SET | PID_CIRCADIAN_PATHWAY |
| | | YAPC | TRT_CP | -0.68 | 0.77 | MOA_CLASS | TOPOISOMERASE_INHIBITOR |
| | | YAPC | TRT_XPR | -0.64 | 0.49 | PATHWAY_SET | BIOCARTA_CBL_PATHWAY |
| | | YAPC | TRT_CP | -0.64 | 0.53 | MOA_CLASS | TUBULIN_INHIBITOR |

Figure S13: CMAP results using the *pancreatic islet* composite transcript as an input. Table includes the top 30 results derived *only from YAPC cells*, which are derived from pancreatic carcinoma cells. Results are shown regardless of significance and are sorted by the correlation of the query to the input with the most negative results at the top.

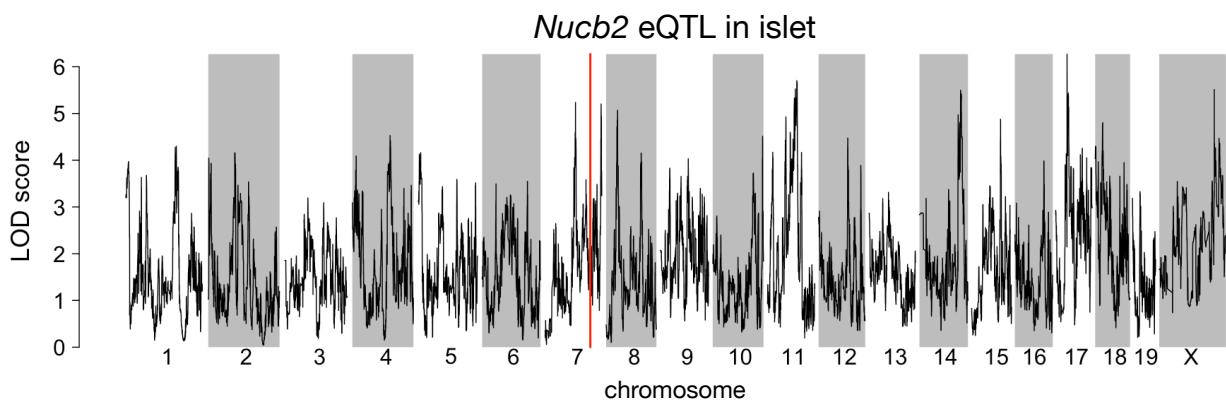


Figure S14: Regulation of *Nucb2* expression in islet. *Nucb2* is encoded on mouse chromosome 7 at 116.5 Mb (red line). In islets the heritability of *Nucb2* expression levels is 69% heritable. This LOD score trace shows that there is no local eQTLs at the position of the gene, nor any strong distal eQTLs anywhere else in the genome.

483 **References**

- 484 [1] M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds,
485 R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutyavin, S. Stehling-Sun, A. K.
486 Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld,
487 A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, and J. A.
488 Stamatoyannopoulos. Systematic localization of common disease-associated variation in regulatory DNA.
489 *Science*, 337(6099):1190–1195, Sep 2012.
- 490 [2] K. K. Farh, A. Marson, J. Zhu, M. Kleinewietfeld, W. J. Housley, S. Beik, N. Shores, H. Whitton, R. J.
491 Ryan, A. A. Shishkin, M. Hatan, M. J. Carrasco-Alfonso, D. Mayer, C. J. Luckey, N. A. Patsopoulos,
492 P. L. De Jager, V. K. Kuchroo, C. B. Epstein, M. J. Daly, D. A. Hafler, and B. E. Bernstein. Genetic
493 and epigenetic fine mapping of causal autoimmune disease variants. *Nature*, 518(7539):337–343, Feb
494 2015.
- 495 [3] E. Pennisi. The Biology of Genomes. Disease risk links to gene regulation. *Science*, 332(6033):1031, May
496 2011.
- 497 [4] L. A. Hindorff, P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta, F. S. Collins, and T. A. Manolio.
498 Potential etiologic and functional implications of genome-wide association loci for human diseases and
499 traits. *Proc Natl Acad Sci*, 106(23):9362–9367, Jun 2009.
- 500 [5] J. K. Pickrell. Joint analysis of functional genomic data and genome-wide association studies of 18
501 human traits. *Am J Hum Genet*, 94(4):559–573, Apr 2014.
- 502 [6] D. Welter, J. MacArthur, J. Morales, T. Burdett, P. Hall, H. Junkins, A. Klemm, P. Flieck, T. Manolio,
503 L. Hindorff, and H. Parkinson. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations.
504 *Nucleic Acids Res*, 42(Database issue):D1001–1006, Jan 2014.
- 505 [7] Y. I. Li, B. van de Geijn, A. Raj, D. A. Knowles, A. A. Petti, D. Golan, Y. Gilad, and J. K. Pritchard.
506 RNA splicing is a primary link between genetic variation and disease. *Science*, 352(6285):600–604, Apr
507 2016.
- 508 [8] D. Zhou, Y. Jiang, X. Zhong, N. J. Cox, C. Liu, and E. R. Gamazon. A unified framework for joint-tissue
509 transcriptome-wide association and Mendelian randomization analysis. *Nat Genet*, 52(11):1239–1246,
510 Nov 2020.
- 511 [9] E. R. Gamazon, H. E. Wheeler, K. P. Shah, S. V. Mozaffari, K. Aquino-Michaels, R. J. Carroll, A. E.

- 512 Eyler, J. C. Denny, D. L. Nicolae, N. J. Cox, and H. K. Im. A gene-based association method for
513 mapping traits using reference transcriptome data. *Nat Genet*, 47(9):1091–1098, Sep 2015.
- 514 [10] Z. Zhu, F. Zhang, H. Hu, A. Bakshi, M. R. Robinson, J. E. Powell, G. W. Montgomery, M. E. Goddard,
515 N. R. Wray, P. M. Visscher, and J. Yang. Integration of summary data from GWAS and eQTL studies
516 predicts complex trait gene targets. *Nat Genet*, 48(5):481–487, May 2016.
- 517 [11] A. Gusev, A. Ko, H. Shi, G. Bhatia, W. Chung, B. W. Penninx, R. Jansen, E. J. de Geus, D. I. Boomsma,
518 F. A. Wright, P. F. Sullivan, E. Nikkola, M. Alvarez, M. Civelek, A. J. Lusis, T. ki, E. Raitoharju,
519 M. nen, I. ä, O. T. Raitakari, J. Kuusisto, M. Laakso, A. L. Price, P. Pajukanta, and B. Pasaniuc.
520 Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*, 48(3):245–252,
521 Mar 2016.
- 522 [12] M. P. Keller, D. M. Gatti, K. L. Schueler, M. E. Rabaglia, D. S. Stapleton, P. Simecek, M. Vincent,
523 S. Allen, A. T. Broman, R. Bacher, C. Kendzierski, K. W. Broman, B. S. Yandell, G. A. Churchill, and
524 A. D. Attie. Genetic Drivers of Pancreatic Islet Function. *Genetics*, 209(1):335–356, May 2018.
- 525 [13] W. L. Crouse, G. R. Keele, M. S. Gastonguay, G. A. Churchill, and W. Valdar. A Bayesian model
526 selection approach to mediation analysis. *PLoS Genet*, 18(5):e1010184, May 2022.
- 527 [14] J. M. Chick, S. C. Munger, P. Simecek, E. L. Huttlin, K. Choi, D. M. Gatti, N. Raghupathy, K. L. Svenson,
528 G. A. Churchill, and S. P. Gygi. Defining the consequences of genetic variation on a proteome-wide scale.
529 *Nature*, 534(7608):500–505, Jun 2016.
- 530 [15] H. E. Wheeler, S. Ploch, A. N. Barbeira, R. Bonazzola, A. Andaleon, A. Fotuhi Siahpirani, A. Saha,
531 A. Battle, S. Roy, and H. K. Im. Imputed gene associations identify replicable trans-acting genes enriched
532 in transcription pathways and complex traits. *Genet Epidemiol*, 43(6):596–608, Sep 2019.
- 533 [16] B. D. Umans, A. Battle, and Y. Gilad. Where Are the Disease-Associated eQTLs? *Trends Genet*,
534 37(2):109–124, Feb 2021.
- 535 [17] N. J. Connally, S. Nazeen, D. Lee, H. Shi, J. Stamatoyannopoulos, S. Chun, C. Cotsapas, C. A. Cassa,
536 and S. R. Sunyaev. The missing link between genetic association and regulatory function. *Elife*, 11, Dec
537 2022.
- 538 [18] H. Mostafavi, J. P. Spence, S. Naqvi, and J. K. Pritchard. Systematic differences in discovery of genetic
539 effects on gene expression and complex traits. *Nat Genet*, 55(11):1866–1875, Nov 2023.
- 540 [19] D. W. Yao, L. J. O'Connor, A. L. Price, and A. Gusev. Quantifying genetic effects on disease mediated
541 by assayed gene expression levels. *Nat Genet*, 52(6):626–633, Jun 2020.

- 542 [20] X. Liu, J. A. Mefford, A. Dahl, Y. He, M. Subramaniam, A. Battle, A. L. Price, and N. Zaitlen. GBAT:
543 a gene-based association test for robust detection of trans-gene regulation. *Genome Biol*, 21(1):211, Aug
544 2020.
- 545 [21] H. J. Westra, M. J. Peters, T. Esko, H. Yaghoobkar, C. Schurmann, J. Kettunen, M. W. Christiansen,
546 B. P. Fairfax, K. Schramm, J. E. Powell, A. Zhernakova, D. V. Zhernakova, J. H. Veldink, L. H. Van den
547 Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P. A. C. ' Hoen,
548 E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A. B. Singleton, D. G. Hernandez,
549 M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Iker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey,
550 S. A. Gharib, D. A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder,
551 M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R. C. Jansen, P. M. Visscher, J. C. Knight,
552 B. M. Psaty, S. Ripatti, A. Teumer, T. M. Frayling, A. Metspalu, J. B. J. van Meurs, and L. Franke.
553 Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*,
554 45(10):1238–1243, Oct 2013.
- 555 [22] Y. Gilad, S. A. Rifkin, and J. K. Pritchard. Revealing the architecture of gene regulation: the promise
556 of eQTL studies. *Trends Genet*, 24(8):408–415, Aug 2008.
- 557 [23] Aparna Nathan, Samira Asgari, Kazuyoshi Ishigaki, Cristian Valencia, Tiffany Amariuta, Yang Luo,
558 Jessica I Beynor, Yuriy Baglaenko, Sara Suliman, Alkes L Price, et al. Single-cell eqtl models reveal
559 dynamic t cell state dependence of disease loci. *Nature*, 606(7912):120–128, 2022.
- 560 [24] E. Sollis, A. Mosaku, A. Abid, A. Buniello, M. Cerezo, L. Gil, T. Groza, O. §, P. Hall, J. Hayhurst,
561 A. Ibrahim, Y. Ji, S. John, E. Lewis, J. A. L. MacArthur, A. McMahon, D. Osumi-Sutherland,
562 K. Panoutsopoulou, Z. Pendlington, S. Ramachandran, R. Stefancsik, J. Stewart, P. Whetzel, R. Wilson,
563 L. Hindorff, F. Cunningham, S. A. Lambert, M. Inouye, H. Parkinson, and L. W. Harris. The NHGRI-EBI
564 GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res*, 51(D1):D977–D985, Jan
565 2023.
- 566 [25] R. J. F. Loos and G. S. H. Yeo. The genetics of obesity: from discovery to biology. *Nat Rev Genet*,
567 23(2):120–133, Feb 2022.
- 568 [26] R. K. Singh, P. Kumar, and K. Mahalingam. Molecular genetics of human obesity: A comprehensive
569 review. *C R Biol*, 340(2):87–108, Feb 2017.
- 570 [27] P. Arner. Obesity—a genetic disease of adipose tissue? *Br J Nutr*, 83 Suppl 1:9–16, Mar 2000.
- 571 [28] Mark P Keller, Mary E Rabaglia, Kathryn L Schueler, Donnie S Stapleton, Daniel M Gatti, Matthew

- 572 Vincent, Kelly A Mitok, Ziyue Wang, Takanao Ishimura, Shane P Simonett, et al. Gene loci associated
573 with insulin secretion in islets from nondiabetic mice. *The Journal of Clinical Investigation*, 129(10):4419–
574 4432, 2019.
- 575 [29] G. A. Churchill, D. M. Gatti, S. C. Munger, and K. L. Svenson. The Diversity Outbred mouse population.
576 *Mamm Genome*, 23(9-10):713–718, Oct 2012.
- 577 [30] Elissa J Chesler, Darla R Miller, Lisa R Branstetter, Leslie D Galloway, Barbara L Jackson, Vivek M
578 Philip, Brynn H Voy, Cymbeline T Culiat, David W Threadgill, Robert W Williams, et al. The
579 collaborative cross at oak ridge national laboratory: developing a powerful resource for systems genetics.
580 *Mammalian Genome*, 19:382–389, 2008.
- 581 [31] Michael C Saul, Vivek M Philip, Laura G Reinholdt, and Elissa J Chesler. High-diversity mouse
582 populations for complex traits. *Trends in Genetics*, 35(7):501–514, 2019.
- 583 [32] D. W. Threadgill, D. R. Miller, G. A. Churchill, and F. P. de Villena. The collaborative cross: a
584 recombinant inbred mouse population for the systems genetic era. *ILAR J*, 52(1):24–31, 2011.
- 585 [33] S. M. Clee and A. D. Attie. The genetic landscape of type 2 diabetes in mice. *Endocr Rev*, 28(1):48–83,
586 Feb 2007.
- 587 [34] K. W. Broman, D. M. Gatti, P. Simecek, N. A. Furlotte, P. Prins, Š. Sen, B. S. Yandell, and G. A.
588 Churchill. R/qt12: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and
589 Multiparent Populations. *Genetics*, 211(2):495–502, Feb 2019.
- 590 [35] Klaasjan G Ouwens, Rick Jansen, Michel G Nivard, Jenny van Dongen, Maia J Frieser, Jouke-Jan
591 Hottenga, Wibowo Arindrarto, Annique Claringbould, Maarten van Iterson, Hailiang Mei, et al. A
592 characterization of cis-and trans-heritability of rna-seq-based gene expression. *European Journal of
593 Human Genetics*, 28(2):253–263, 2020.
- 594 [36] Alkes L Price, Agnar Helgason, Gudmar Thorleifsson, Steven A McCarroll, Augustine Kong, and Kari
595 Stefansson. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related
596 or unrelated individuals. *PLoS genetics*, 7(2):e1001317, 2011.
- 597 [37] Julien Bryois, Alfonso Buil, David M Evans, John P Kemp, Stephen B Montgomery, Donald F Conrad,
598 Karen M Ho, Susan Ring, Matthew Hurles, Panos Deloukas, et al. Cis and trans effects of human
599 genomic variants on gene expression. *PLoS genetics*, 10(7):e1004461, 2014.
- 600 [38] M. Helmer, S. Warrington, A. R. Mohammadi-Nejad, J. L. Ji, A. Howell, B. Rosand, A. Anticevic,

- 601 S. N. Sotiropoulos, and J. D. Murray. On the stability of canonical correlation analysis and partial least
602 squares with application to brain-behavior associations. *Commun Biol*, 7(1):217, Feb 2024.
- 603 [39] Gennady Korotkevich, Vladimir Sukhov, and Alexey Sergushichev. Fast gene set enrichment analysis.
604 *bioRxiv*, 2019.
- 605 [40] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich,
606 S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: a
607 knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*,
608 102(43):15545–15550, Oct 2005.
- 609 [41] S. Subramanian and A. Chait. The effect of dietary cholesterol on macrophage accumulation in adipose
610 tissue: implications for systemic inflammation and atherosclerosis. *Curr Opin Lipidol*, 20(1):39–44, Feb
611 2009.
- 612 [42] I. Akoumianakis, N. Akawi, and C. Antoniades. Exploring the Crosstalk between Adipose Tissue and
613 the Cardiovascular System. *Korean Circ J*, 47(5):670–685, Sep 2017.
- 614 [43] I. S. Stafeev, A. V. Vorotnikov, E. I. Ratner, M. Y. Menshikov, and Y. V. Parfyonova. Latent Inflammation
615 and Insulin Resistance in Adipose Tissue. *Int J Endocrinol*, 2017:5076732, 2017.
- 616 [44] I. P. Fischer, M. Irmler, C. W. Meyer, S. J. Sachs, F. Neff, M. de Angelis, J. Beckers, M. H. p, S. M.
617 Hofmann, and S. Ussar. A history of obesity leaves an inflammatory fingerprint in liver and adipose
618 tissue. *Int J Obes (Lond)*, 42(3):507–517, Mar 2018.
- 619 [45] S. Chung, H. Cuffe, S. M. Marshall, A. L. McDaniel, J. H. Ha, K. Kavanagh, C. Hong, P. Tontonoz,
620 R. E. Temel, and J. S. Parks. Dietary cholesterol promotes adipocyte hypertrophy and adipose tissue
621 inflammation in visceral, but not in subcutaneous, fat in monkeys. *Arterioscler Thromb Vasc Biol*,
622 34(9):1880–1887, Sep 2014.
- 623 [46] V. Kus, T. Prazak, P. Brauner, M. Hensler, O. Kuda, P. Flachs, P. Janovska, D. Medrikova, M. Rossmeisl,
624 Z. Jilkova, B. Stefl, E. Pastalkova, Z. Drahota, J. Houstek, and J. Kopecky. Induction of muscle
625 thermogenesis by high-fat diet in mice: association with obesity-resistance. *Am J Physiol Endocrinol
Metab*, 295(2):E356–367, Aug 2008.
- 627 [47] C. B. Newgard. Interplay between lipids and branched-chain amino acids in development of insulin
628 resistance. *Cell Metab*, 15(5):606–614, May 2012.
- 629 [48] D. D. Sears, G. Hsiao, A. Hsiao, J. G. Yu, C. H. Courtney, J. M. Ofrecio, J. Chapman, and S. Subramaniam.

- 630 Mechanisms of human insulin resistance and thiazolidinedione-mediated insulin sensitization. *Proc Natl*
631 *Acad Sci U S A*, 106(44):18745–18750, Nov 2009.
- 632 [49] R. Stienstra, C. Duval, M. ller, and S. Kersten. PPARs, Obesity, and Inflammation. *PPAR Res*,
633 2007:95974, 2007.
- 634 [50] O. Gavrilova, M. Haluzik, K. Matsusue, J. J. Cutson, L. Johnson, K. R. Dietz, C. J. Nicol, C. Vinson,
635 F. J. Gonzalez, and M. L. Reitman. Liver peroxisome proliferator-activated receptor gamma contributes
636 to hepatic steatosis, triglyceride clearance, and regulation of body fat mass. *J Biol Chem*, 278(36):34268–
637 34276, Sep 2003.
- 638 [51] K. Matsusue, M. Haluzik, G. Lambert, S. H. Yim, O. Gavrilova, J. M. Ward, B. Brewer, M. L. Reitman,
639 and F. J. Gonzalez. Liver-specific disruption of PPARgamma in leptin-deficient mice improves fatty
640 liver but aggravates diabetic phenotypes. *J Clin Invest*, 111(5):737–747, Mar 2003.
- 641 [52] D. Patsouris, J. K. Reddy, M. ller, and S. Kersten. Peroxisome proliferator-activated receptor alpha
642 mediates the effects of high-fat diet on hepatic gene expression. *Endocrinology*, 147(3):1508–1516, Mar
643 2006.
- 644 [53] S. E. Schadinger, N. L. Bucher, B. M. Schreiber, and S. R. Farmer. PPARgamma2 regulates lipogenesis
645 and lipid accumulation in steatotic hepatocytes. *Am J Physiol Endocrinol Metab*, 288(6):E1195–1205,
646 Jun 2005.
- 647 [54] W. Motomura, M. Inoue, T. Ohtake, N. Takahashi, M. Nagamine, S. Tanno, Y. Kohgo, and T. Okumura.
648 Up-regulation of ADRP in fatty liver in human and liver steatosis in mice fed with high fat diet. *Biochem*
649 *Biophys Res Commun*, 340(4):1111–1118, Feb 2006.
- 650 [55] A. Srivastava, A. P. Morgan, M. L. Najarian, V. K. Sarsani, J. S. Sigmon, J. R. Shorter, A. Kashfeen,
651 R. C. McMullan, L. H. Williams, P. guez, M. T. Ferris, P. Sullivan, P. Hock, D. R. Miller, T. A. Bell,
652 L. McMillan, G. A. Churchill, and F. P. de Villena. Genomes of the Mouse Collaborative Cross. *Genetics*,
653 206(2):537–556, Jun 2017.
- 654 [56] A. Roberts, F. Pardo-Manuel de Villena, W. Wang, L. McMillan, and D. W. Threadgill. The poly-
655 morphism architecture of mouse genetic resources elucidated using genome-wide resequencing data:
656 implications for QTL discovery and systems genetics. *Mamm Genome*, 18(6-7):473–481, Jul 2007.
- 657 [57] G. A. Churchill, D. C. Airey, H. Allayee, J. M. Angel, A. D. Attie, J. Beatty, W. D. Beavis, J. K.
658 Belknap, B. Bennett, W. Berrettini, A. Bleich, M. Bogue, K. W. Broman, K. J. Buck, E. Buckler,
659 M. Burmeister, E. J. Chesler, J. M. Cheverud, S. Clapcote, M. N. Cook, R. D. Cox, J. C. Crabbe,

- 660 W. E. Crusio, A. Darvasi, C. F. Deschepper, R. W. Doerge, C. R. Farber, J. Forejt, D. Gaile, S. J.
661 Garlow, H. Geiger, H. Gershenfeld, T. Gordon, J. Gu, W. Gu, G. de Haan, N. L. Hayes, C. Heller,
662 H. Himmelbauer, R. Hitzemann, K. Hunter, H. C. Hsu, F. A. Iraqi, B. Ivandic, H. J. Jacob, R. C. Jansen,
663 K. J. Jepsen, D. K. Johnson, T. E. Johnson, G. Kempermann, C. Kendzierski, M. Kotb, R. F. Kooy,
664 B. Llamas, F. Lammert, J. M. Lassalle, P. R. Lowenstein, L. Lu, A. Lusis, K. F. Manly, R. Marcucio,
665 D. Matthews, J. F. Medrano, D. R. Miller, G. Mittelman, B. A. Mock, J. S. Mogil, X. Montagutelli,
666 G. Morahan, D. G. Morris, R. Mott, J. H. Nadeau, H. Nagase, R. S. Nowakowski, B. F. O'Hara, A. V.
667 Osadchuk, G. P. Page, B. Paigen, K. Paigen, A. A. Palmer, H. J. Pan, L. Peltonen-Palotie, J. Peirce,
668 D. Pomp, M. Pravenec, D. R. Prows, Z. Qi, R. H. Reeves, J. Roder, G. D. Rosen, E. E. Schadt, L. C.
669 Schalkwyk, Z. Seltzer, K. Shimomura, S. Shou, M. J. ä, L. D. Siracusa, H. W. Snoek, J. L. Spearow,
670 K. Svenson, L. M. Tarantino, D. Threadgill, L. A. Toth, W. Valdar, F. P. de Villena, C. Warden,
671 S. Whatley, R. W. Williams, T. Wiltshire, N. Yi, D. Zhang, M. Zhang, and F. Zou. The Collaborative
672 Cross, a community resource for the genetic analysis of complex traits. *Nat Genet*, 36(11):1133–1137,
673 Nov 2004.
- 674 [58] J. Y. Huh, Y. J. Park, M. Ham, and J. B. Kim. Crosstalk between adipocytes and immune cells in
675 adipose tissue inflammation and metabolic dysregulation in obesity. *Mol Cells*, 37(5):365–371, May 2014.
- 676 [59] J. Lamb, E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet,
677 A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty,
678 P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, and T. R. Golub. The Connectivity Map: using
679 gene-expression signatures to connect small molecules, genes, and disease. *Science*, 313(5795):1929–1935,
680 Sep 2006.
- 681 [60] Aravind Subramanian, Rajiv Narayan, Steven M Corsello, David D Peck, Ted E Natoli, Xiaodong Lu,
682 Joshua Gould, John F Davis, Andrew A Tubelli, Jacob K Asiedu, et al. A next generation connectivity
683 map: L1000 platform and the first 1,000,000 profiles. *Cell*, 171(6):1437–1452, 2017.
- 684 [61] S. Amin, A. Lux, and F. O'Callaghan. The journey of metformin from glycaemic control to mTOR
685 inhibition and the suppression of tumour growth. *Br J Clin Pharmacol*, 85(1):37–46, Jan 2019.
- 686 [62] A. Kezic, L. Popovic, and K. Lalic. mTOR Inhibitor Therapy and Metabolic Consequences: Where Do
687 We Stand? *Oxid Med Cell Longev*, 2018:2640342, 2018.
- 688 [63] A. D. Barlow, M. L. Nicholson, and T. P. Herbert. -cells and a review of the underlying molecular
689 mechanisms. *Diabetes*, 62(8):2674–2682, Aug 2013.

- 690 [64] Y. Gu, J. Lindner, A. Kumar, W. Yuan, and M. A. Magnuson. Rictor/mTORC2 is essential for
691 maintaining a balance between beta-cell proliferation and cell size. *Diabetes*, 60(3):827–837, Mar 2011.
- 692 [65] E. B. Geer, J. Islam, and C. Buettner. Mechanisms of glucocorticoid-induced insulin resistance: focus on
693 adipose tissue function and lipid metabolism. *Endocrinol Metab Clin North Am*, 43(1):75–102, Mar 2014.
- 694 [66] J. X. Li and C. L. Cummins. Fresh insights into glucocorticoid-induced diabetes mellitus and new
695 therapeutic directions. *Nat Rev Endocrinol*, 18(9):540–557, Sep 2022.
- 696 [67] R. A. Lee, C. A. Harris, and J. C. Wang. Glucocorticoid Receptor and Adipocyte Biology. *Nucl Receptor Res*, 5, 2018.
- 698 [68] S. Viengchareun, P. Penfornis, M. C. Zennaro, and M. s. Mineralocorticoid and glucocorticoid receptors
699 inhibit UCP expression and function in brown adipocytes. *Am J Physiol Endocrinol Metab*, 280(4):E640–
700 649, Apr 2001.
- 701 [69] J. Liu, X. Kong, L. Wang, H. Qi, W. Di, X. Zhang, L. Wu, X. Chen, J. Yu, J. Zha, S. Lv, A. Zhang,
702 P. Cheng, M. Hu, Y. Li, J. Bi, Y. Li, F. Hu, Y. Zhong, Y. Xu, and G. Ding. -HSD1 in regulating brown
703 adipocyte function. *J Mol Endocrinol*, 50(1):103–113, Feb 2013.
- 704 [70] L. E. Ramage, M. Akyol, A. M. Fletcher, J. Forsythe, M. Nixon, R. N. Carter, E. J. van Beek,
705 N. M. Morton, B. R. Walker, and R. H. Stimson. Glucocorticoids Acutely Increase Brown Adipose
706 Tissue Activity in Humans, Revealing Species-Specific Differences in UCP-1 Regulation. *Cell Metab*,
707 24(1):130–141, Jul 2016.
- 708 [71] J. L. Barclay, H. Agada, C. Jang, M. Ward, N. Wetzig, and K. K. Ho. Effects of glucocorticoids on
709 human brown adipocytes. *J Endocrinol*, 224(2):139–147, Feb 2015.
- 710 [72] M. ó, R. Gupte, W. L. Kraus, P. Pacher, and P. Bai. PARPs in lipid metabolism and related diseases.
711 *Prog Lipid Res*, 84:101117, Nov 2021.
- 712 [73] P. Bai, C. ó, H. Oudart, A. nszki, Y. Cen, C. Thomas, H. Yamamoto, A. Huber, B. Kiss, R. H.
713 Houtkooper, K. Schoonjans, V. Schreiber, A. A. Sauve, J. Menissier-de Murcia, and J. Auwerx. PARP-1
714 inhibition increases mitochondrial metabolism through SIRT1 activation. *Cell Metab*, 13(4):461–468,
715 Apr 2011.
- 716 [74] A. Chiarugi and M. A. Moskowitz. Cell biology. PARP-1—a perpetrator of apoptotic cell death? *Science*,
717 297(5579):200–201, Jul 2002.
- 718 [75] M. Althubiti, R. Almainani, S. Y. Eid, M. Elzubaier, B. Refaat, S. Idris, T. A. Alqurashi, and M. Z.

- 719 El-Readi. BTK targeting suppresses inflammatory genes and ameliorates insulin resistance. *Eur Cytokine Netw*, 31(4):168–179, Dec 2020.
- 720
- 721 [76] C. Skrabs, W. F. Pickl, T. Perkmann, U. ger, and A. Gessl. Rapid decline in insulin antibodies and
722 glutamic acid decarboxylase autoantibodies with ibrutinib therapy of chronic lymphocytic leukaemia. *J
723 Clin Pharm Ther*, 43(1):145–149, Feb 2018.
- 724 [77] Hans Hacker and Michael Karin. Regulation and function of ikk and ikk-related kinases. *Science's
725 STKE*, 2006(357):re13–re13, 2006.
- 726 [78] E. A. Oral, S. M. Reilly, A. V. Gomez, R. Meral, L. Butz, N. Ajluni, T. L. Chenevert, E. Korytnaya,
727 A. H. Neidert, R. Hench, D. Rus, J. F. Horowitz, B. Poirier, P. Zhao, K. Lehmann, M. Jain, R. Yu,
728 C. Liddle, M. Ahmadian, M. Downes, R. M. Evans, and A. R. Saltiel. and TBK1 Improves Glucose
729 Control in a Subset of Patients with Type 2 Diabetes. *Cell Metab*, 26(1):157–170, Jul 2017.
- 730 [79] M. C. Arkan, A. L. Hevener, F. R. Greten, S. Maeda, Z. W. Li, J. M. Long, A. Wynshaw-Boris, G. Poli,
731 J. Olefsky, and M. Karin. IKK-beta links inflammation to obesity-induced insulin resistance. *Nat Med*,
732 11(2):191–198, Feb 2005.
- 733 [80] M. Clark, C. J. Kroger, Q. Ke, and R. M. Tisch. The Role of T Cell Receptor Signaling in the
734 Development of Type 1 Diabetes. *Front Immunol*, 11:615371, 2020.
- 735 [81] P. ndell, A. C. Carlsson, A. Larsson, O. Melander, T. Wessman, J. v, and T. Ruge. TNFR1 is associated
736 with short-term mortality in patients with diabetes and acute dyspnea seeking care at the emergency
737 department. *Acta Diabetol*, 57(10):1145–1150, Oct 2020.
- 738 [82] A. M. Keestra-Gounder, M. X. Byndloss, N. Seyffert, B. M. Young, A. vez Arroyo, A. Y. Tsai, S. A.
739 Cevallos, M. G. Winter, O. H. Pham, C. R. Tiffany, M. F. de Jong, T. Kerrinnes, R. Ravindran, P. A.
740 Luciw, S. J. McSorley, A. J. umler, and R. M. Tsolis. NOD1 and NOD2 signalling links ER stress with
741 inflammation. *Nature*, 532(7599):394–397, Apr 2016.
- 742 [83] A. M. Keestra-Gounder and R. M. Tsolis. NOD1 and NOD2: Beyond Peptidoglycan Sensing. *Trends
743 Immunol*, 38(10):758–767, Oct 2017.
- 744 [84] J. Montane, L. Cadavez, and A. Novials. Stress and the inflammatory process: a major cause of
745 pancreatic cell death in type 2 diabetes. *Diabetes Metab Syndr Obes*, 7:25–34, 2014.
- 746 [85] B. B. Kahn and T. E. McGraw. , and type 2 diabetes. *N Engl J Med*, 363(27):2667–2669, Dec 2010.

- 747 [86] S. Del Prato and N. Pulizzi. The place of sulfonylureas in the therapy for type 2 diabetes mellitus.
748 *Metabolism*, 55(5 Suppl 1):S20–27, May 2006.
- 749 [87] X. Liu, Y. I. Li, and J. K. Pritchard. Trans Effects on Gene Expression Can Drive Omnipotent Inheritance.
750 *Cell*, 177(4):1022–1034, May 2019.
- 751 [88] U. Vosa, A. Claringbould, H. J. Westra, M. J. Bonder, P. Deelen, B. Zeng, H. Kirsten, A. Saha,
752 R. Kreuzhuber, S. Yazar, H. Brugge, R. Oelen, D. H. de Vries, M. G. P. van der Wijst, S. Kasela,
753 N. Pervjakova, I. Alves, M. J. é, M. Agbessi, M. W. Christiansen, R. Jansen, I. ä, L. Tong, A. Teumer,
754 K. Schramm, G. Hemani, J. Verlouw, H. Yaghootkar, R. nmez Flitman, A. Brown, V. Kukushkina,
755 A. Kalnayenlis, S. eger, E. Porcu, J. Kronberg, J. Kettunen, B. Lee, F. Zhang, T. Qi, J. A. Hernandez,
756 W. Arindrarto, F. Beutner, J. Dmitrieva, M. Elansary, B. P. Fairfax, M. Georges, B. T. Heijmans, A. W.
757 Hewitt, M. nen, Y. Kim, J. C. Knight, P. Kovacs, K. Krohn, S. Li, M. Loeffler, U. M. Marigorta, H. Mei,
758 Y. Momozawa, M. ller Nurasyid, M. Nauck, M. G. Nivard, B. W. J. H. Penninx, J. K. Pritchard, O. T.
759 Raitakari, O. Rotzschke, E. P. Slagboom, C. D. A. Stehouwer, M. Stumvoll, P. Sullivan, P. A. C. 't Hoen,
760 J. Thiery, A. njes, J. van Dongen, M. van Iterson, J. H. Veldink, U. lker, R. Warmerdam, C. Wijmenga,
761 M. Swertz, A. Andiappan, G. W. Montgomery, S. Ripatti, M. Perola, Z. Katalik, E. Dermitzakis,
762 S. Bergmann, T. Frayling, J. van Meurs, H. Prokisch, H. Ahsan, B. L. Pierce, T. ki, D. I. Boomsma, B. M.
763 Psaty, S. A. Gharib, P. Awadalla, L. Milani, W. H. Ouwehand, K. Downes, O. Stegle, A. Battle, P. M.
764 Visscher, J. Yang, M. Scholz, J. Powell, G. Gibson, T. Esko, L. Franke, P. A. C. 't Hoen, J. van Meurs,
765 J. van Dongen, M. van Iterson, M. A. Swertz, and M. Jan Bonder. Large-scale cis- and trans-eQTL
766 analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat
767 Genet*, 53(9):1300–1310, Sep 2021.
- 768 [89] B. Hallgrímsson, R. M. Green, D. C. Katz, J. L. Fish, F. P. Bernier, C. C. Roseman, N. M. Young,
769 J. M. Cheverud, and R. S. Marcucio. The developmental-genetics of canalization. *Semin Cell Dev Biol*,
770 88:67–79, Apr 2019.
- 771 [90] M. L. Siegal and A. Bergman. Waddington's canalization revisited: developmental stability and evolution.
772 *Proc Natl Acad Sci U S A*, 99(16):10528–10532, Aug 2002.
- 773 [91] A. B. Paaby and G. Gibson. Cryptic Genetic Variation in Evolutionary Developmental Genetics. *Biology
774 (Basel)*, 5(2), Jun 2016.
- 775 [92] E. A. Boyle, Y. I. Li, and J. K. Pritchard. An Expanded View of Complex Traits: From Polygenic to
776 Omnipotent. *Cell*, 169(7):1177–1186, Jun 2017.

- 777 [93] Naomi R Wray, Cisca Wijmenga, Patrick F Sullivan, Jian Yang, and Peter M Visscher. Common disease
778 is more complex than implied by the core gene omnigenic model. *Cell*, 173(7):1573–1580, 2018.
- 779 [94] H. Shimizu and A. Osaki. Nesfatin/Nucleobindin-2 (NUCB2) and Glucose Homeostasis. *Curr Hypertens
780 Rev*, pages Nesfatin/Nucleobindin-2 (NUCB2) and Glucose Homeostasis., Jul 2014.
- 781 [95] M. Nakata and T. Yada. Role of NUCB2/nesfatin-1 in glucose control: diverse functions in islets,
782 adipocytes and brain. *Curr Pharm Des*, 19(39):6960–6965, 2013.
- 783 [96] M. Riva, M. D. Nitert, U. Voss, R. Sathanoori, A. Lindqvist, C. Ling, and N. Wierup. Nesfatin-1
784 stimulates glucagon and insulin secretion and beta cell NUCB2 is reduced in human type 2 diabetic
785 subjects. *Cell Tissue Res*, 346(3):393–405, Dec 2011.