

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

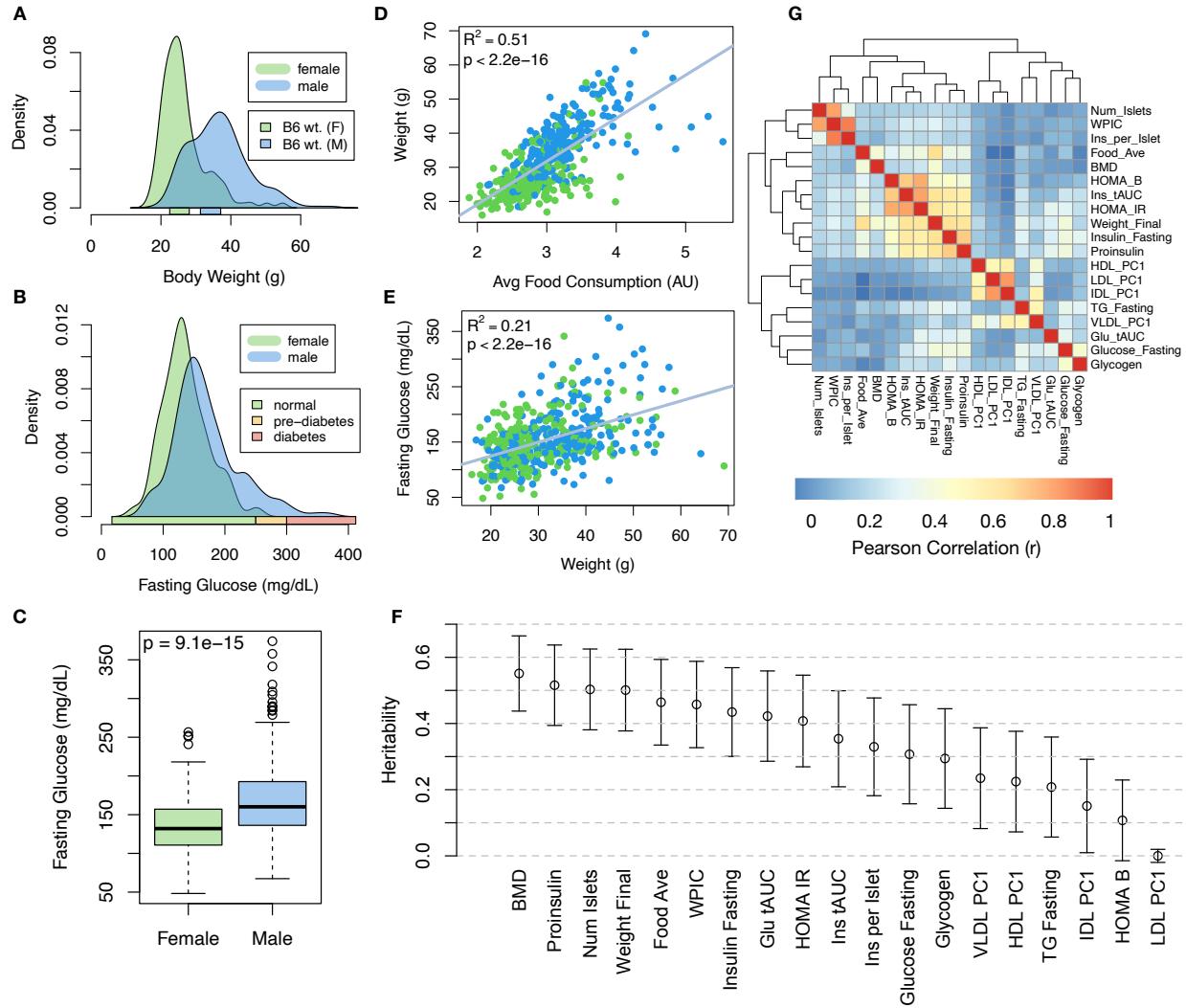


Figure 1: Clinical overview. **A.** Distributions of final body weight in the diversity outbred mice. Sex is indicated by color. The average B6 male and female adult weights at 24 weeks of age are indicated by blue and green bars on the x-axis. **B.** The distribution of final fasting glucose across 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 only 20-30% of total heritability,

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

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

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

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

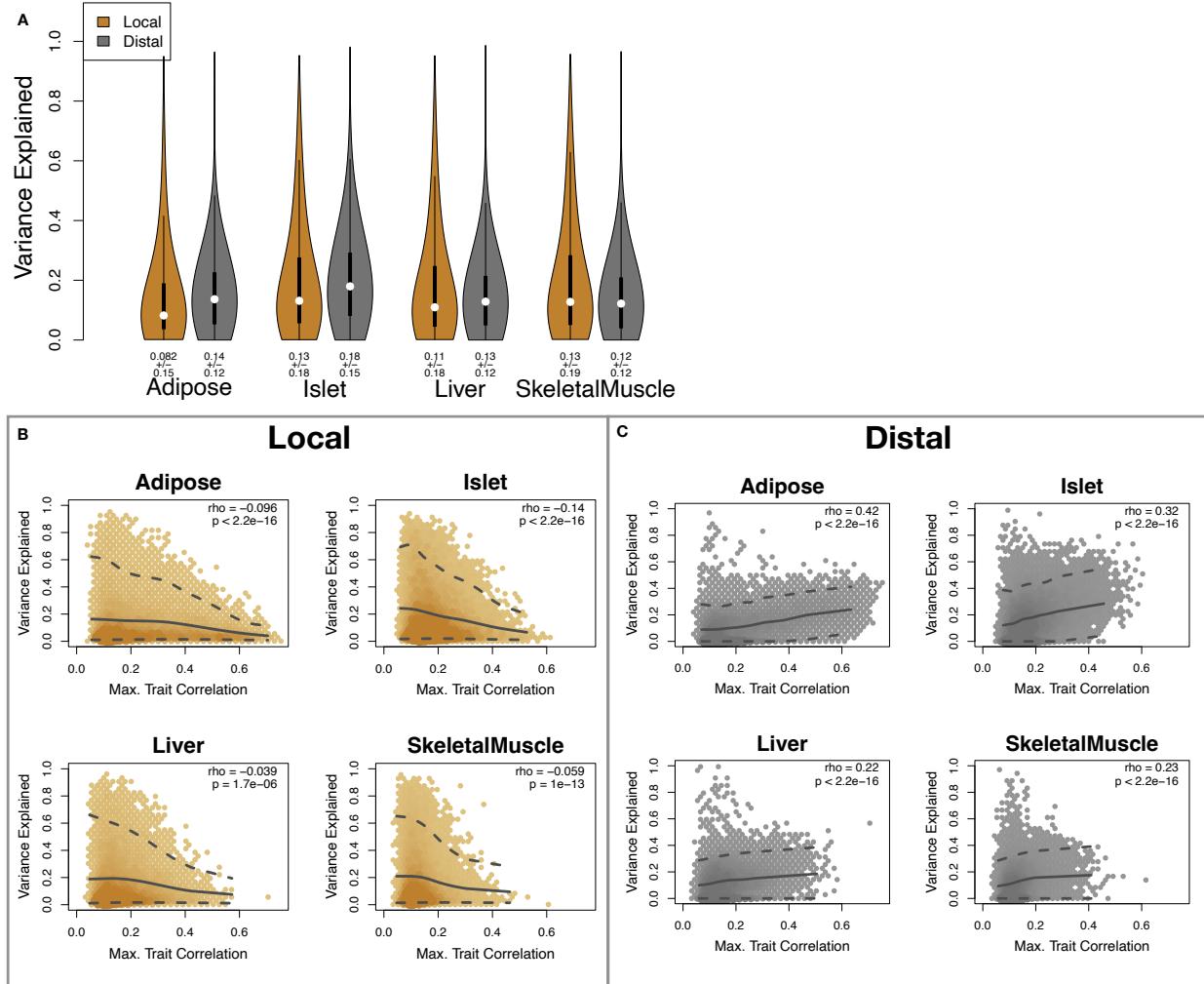


Figure 2: Transcript heritability and trait relevance. **A.** Distributions of distal and local heritability of transcripts across the four tissues. Overall local and distal factors contribute equally to transcript heritability. Arrows indicate the median of each distribution. The relationship between **(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 correlated with trait relevance, while distal heritability was positively correlated with trait relevance. Pearson (r) and p values for each correlation are shown in the upper right of each panel.

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

157 one-dimensional scores—a composite genome score (G_C), a composite transcriptome score (T_C), and a com-
 158 posite phenome score (P_C)—and ~~uses~~used the univariate theory of mediation to constrain these projections
 159 to satisfy the hypotheses of perfect mediation, namely that upon controlling for the transcriptomic score, the
 160 genome score is uncorrelated to the phenome score. A complete mathematical derivation and implementation
 161 details for HDMA are available in [Supp. the Methods](#).

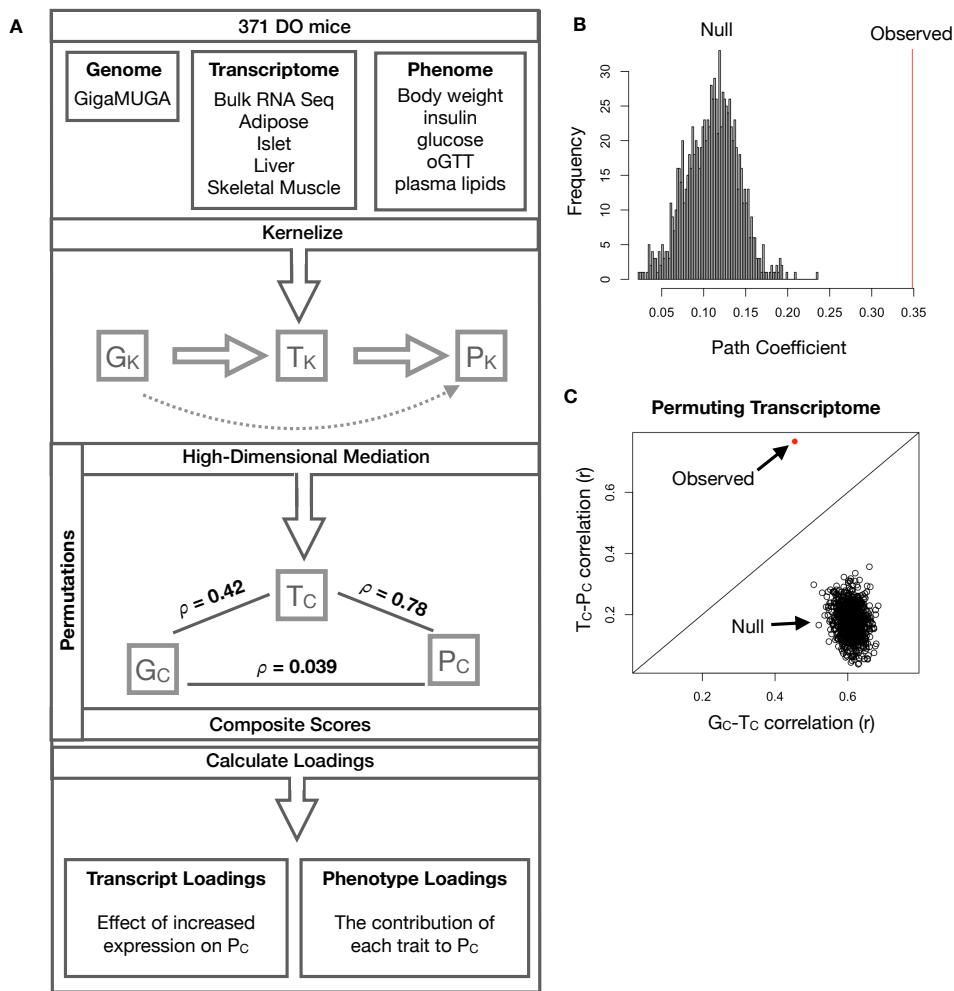


Figure 3: High-dimensional mediation. **A.** Workflow indicating major steps of high-dimensional mediation. The genotype, transcriptome, and phenotype matrices were independently normalized and converted to kernel matrices representing the pairwise relationships between individuals for each data modality (K_G = genome kernel, K_T = transcriptome kernel; K_P = phenome 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 , 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 methods. **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).

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

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

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

192 Each composite score is a weighted combination of the measured variables. The magnitude and sign of the
193 weights, called loadings, correspond to the relative importance and directionality of each variable in the
194 composite score. The loadings of each measured trait onto P_C indicate how much each contributed to the
195 composite phenotype. Body weight contributed the most (Fig. 4), followed by homeostatic insulin resistance
196 (HOMA_IR) and fasting plasma insulin levels (Insulin_Fasting). We can thus interpret P_C as an index
197 of metabolic disease (Fig. 4B). Individuals with high values of P_C have a higher metabolic disease index
198 (MDI) and greater metabolic disease, including higher body weight and higher insulin resistance. We refer to
199 P_C as the MDI going forward. Traits contributing the least to the MDI were measures of cholesterol and
200 pancreas composition. Thus, when we interpret the transcriptomic signature identified by HDMA, we are
201 explaining primarily the putative transcriptional mediation of body weight and insulin resistance, as opposed
202 to cholesterol measurements.

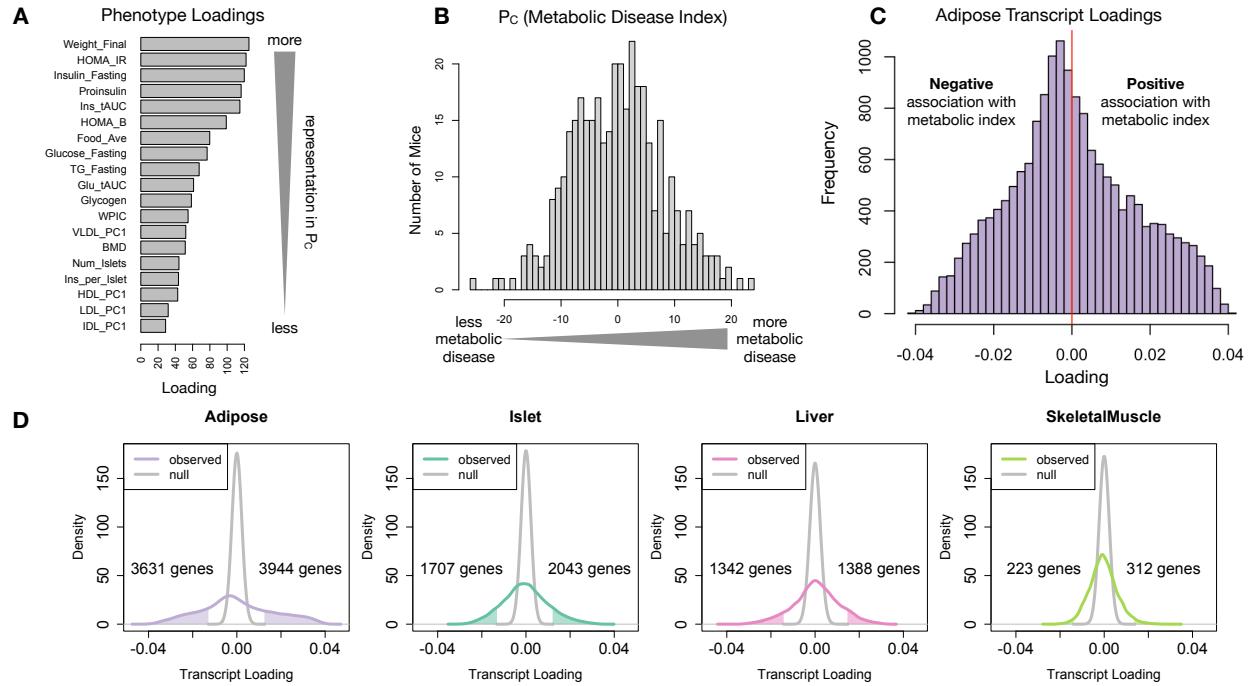


Figure 4: Interpretation of loadings. **A.** Loadings across traits. Body weight and insulin resistance contributed the most to the composite trait. **B.** Phenotype scores across individuals. Individuals with large positive phenotype scores had higher body weight and insulin resistance than average. Individuals with large negative phenotype scores 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 Distributions 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. **E.** Dot plots showing correlations between the composite genome (G_C), composite transcriptome (T_C), and the composite phenotype (P_C). The G_C and T_C are positively correlated, the G_T and G_P are positively correlated, but the G_C and G_P are uncorrelated.

- 203 **High-loading transcripts had low local heritability, high distal heritability, and were linked**
 204 **mechanistically to obesity**
- 205 We interpreted large loadings onto transcripts as indicating strong mediation of the effect of genetics on the
 206 MDI. Large positive loadings indicate that higher expression was associated with a higher MDI (i.e. higher
 207 risk of obesity and metabolic disease on the HFHS diet) (Fig. 4C). Conversely, large negative loadings
 208 indicate that high expression of these transcripts was associated with a lower MDI (i.e. lower risk of obesity
 209 and metabolic disease on the HFHS diet) (Fig. 4C). We used gene set enrichment analysis (GSEA)^{39;40} to
 210 look for biological processes and pathways that were enriched at the top and bottom of this list (Methods).
- 211 In adipose tissue, both GO processes and KEGG pathway enrichments pointed to an axis of inflammation and

212 metabolism (Figs. S3 and S4). GO terms and KEGG pathways associated with inflammation were positively
213 associated with the MDI, indicating that increased expression in inflammatory pathways was associated
214 with a higher burden of disease. It is well established that adipose tissue in obese individuals is inflamed
215 and infiltrated by macrophages^{41–45}, and the results here suggest that this may be a dominant heritable
216 component of metabolic disease.

217 The strongest negative enrichments in adipose tissue were related to mitochondrial activity in general, and
218 thermogenesis in particular (Figs. S3 and S3). Genes in the KEGG oxidative phosphorylation pathway were
219 almost universally negatively loaded in adipose tissue, suggesting that increased expression of these genes was
220 associated with reduced MDI (Supp. Fig. S5). Consistent with this observation, it has been shown previously
221 that mouse strains with greater thermogenic potential are also less susceptible to obesity on an obesigenic
222 diet⁴⁶.

223 Transcripts associated with the citric acid cycle as well as the catabolism of the branched-chain amino acids
224 (valine, leucine, and isoleucine) were strongly enriched with negative loadings in adipose tissue (Supp. Figs.
225 S3, S6 and S7). Expression of genes in both pathways (for which there is some overlap) has been previously
226 associated with insulin sensitivity^{12;47;48}, suggesting that heritable variation in regulation of these pathways
227 may influence risk of insulin resistance.

228 Looking at the 10 most positively and negatively loaded transcripts from each tissue, it is apparent that
229 transcripts in the adipose tissue had the largest loadings, both positive and negative (Fig. 5A bar plot). This
230 suggests that much of the effect of genetics on body weight and insulin resistance is mediated through gene
231 expression in adipose tissue. This finding does not speak to the relative importance of tissues not included
232 in this study, such as brain, in which transcriptional variation may mediate a large portion of the genetic
233 effect on obesity. The strongest loadings in liver and pancreas were comparable, and those in skeletal muscle
234 were the weakest (Fig. 5A), suggesting that less of the genetic effects were mediated through transcription
235 in skeletal muscle. As expected, heritability analysis showed that transcripts with the largest loadings had
236 higher distal heritability than local heritability (Fig. 5A heat map and box plot). This pattern contrasts
237 with transcripts nominated by TWAS We also performed TWAS in this population by imputing transcript
238 levels for each gene based on local genotype only and correlating the imputed transcript levels with each
239 trait. In contrast to HDMA, the TWAS procedure tended to nominate transcripts with lower loadings (Fig.
240 5B), which tended to have lower loadings, higher local heritability and lower distal heritability. Transcripts
241 Finally, we focused on transcripts with the highest local heritability in each tissue (Fig. 5C)had the lowest
242 loadings. This procedure selected transcripts with low loadings on average, consistent with our findings
243 above (Fig. 2B).

²⁴⁴ We performed a literature search for the genes in each of these groups along with the terms “diabetes”,
²⁴⁵ “obesity”, and the name of the expressing tissue to determine whether any of these genes had previous
²⁴⁶ associations with metabolic disease in the literature (Methods). Multiple genes in each group had been
²⁴⁷ previously associated with obesity and diabetes (Fig. 5 bolded gene names). Genes with high loadings were
²⁴⁸ most highly enriched for previous literature support. They were 2.42.1 times more likely than TWAS hits
²⁴⁹ and 3.8 times more likely than genes with high local heritability to be previously associated with obesity or
²⁵⁰ diabetes.

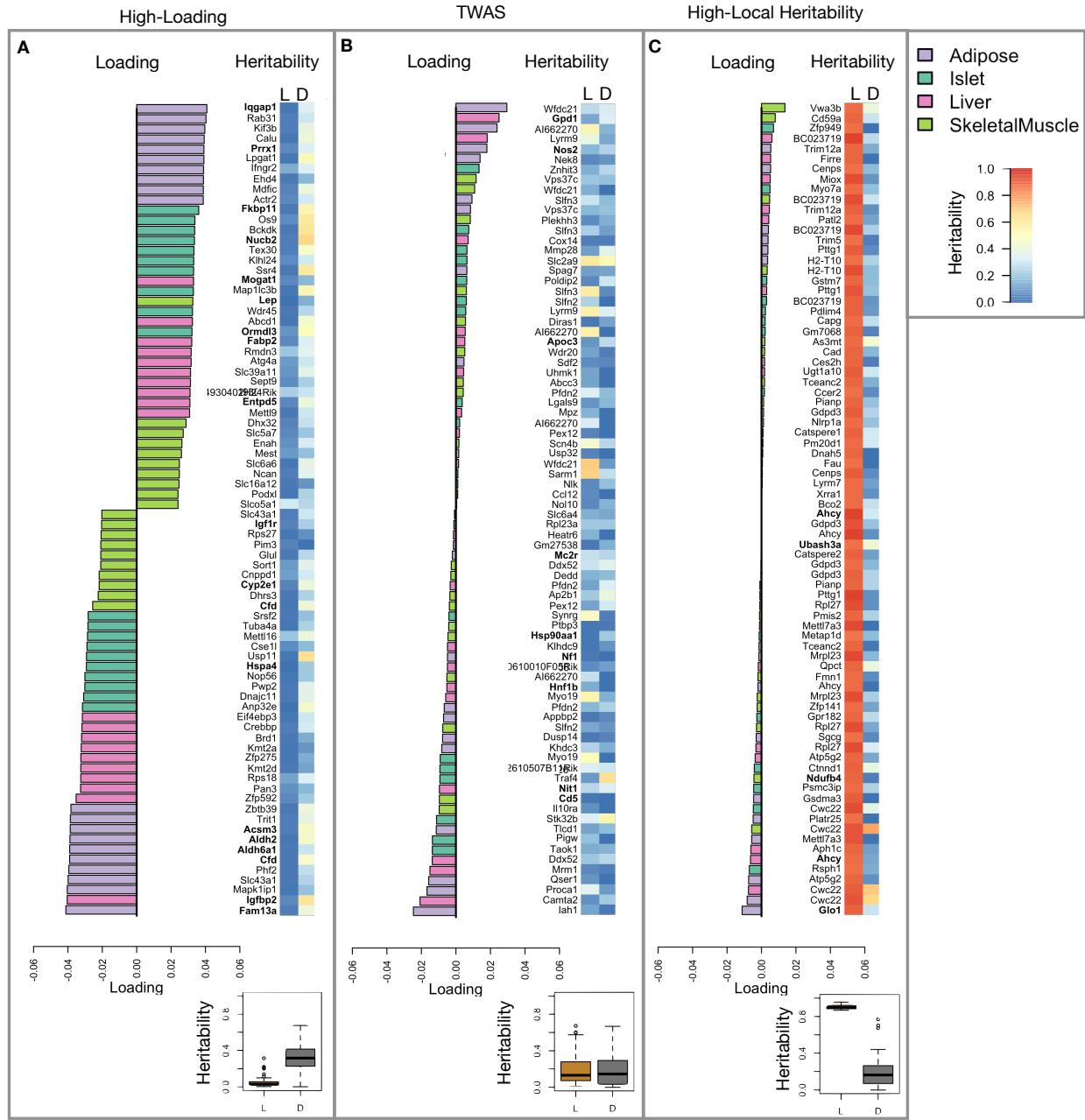


Figure 5: Transcripts with high loadings have high distal heritability and literature support. Each panel has a bar plot showing the loadings of transcripts selected by different criteria. Bar color indicates the tissue of origin. The heat map shows the local (L - left) and distal (D - right) heritability 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 $p < 2.2^{-16}$)

251 **Tissue-specific transcriptional programs were associated with metabolic traits**

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

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

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

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

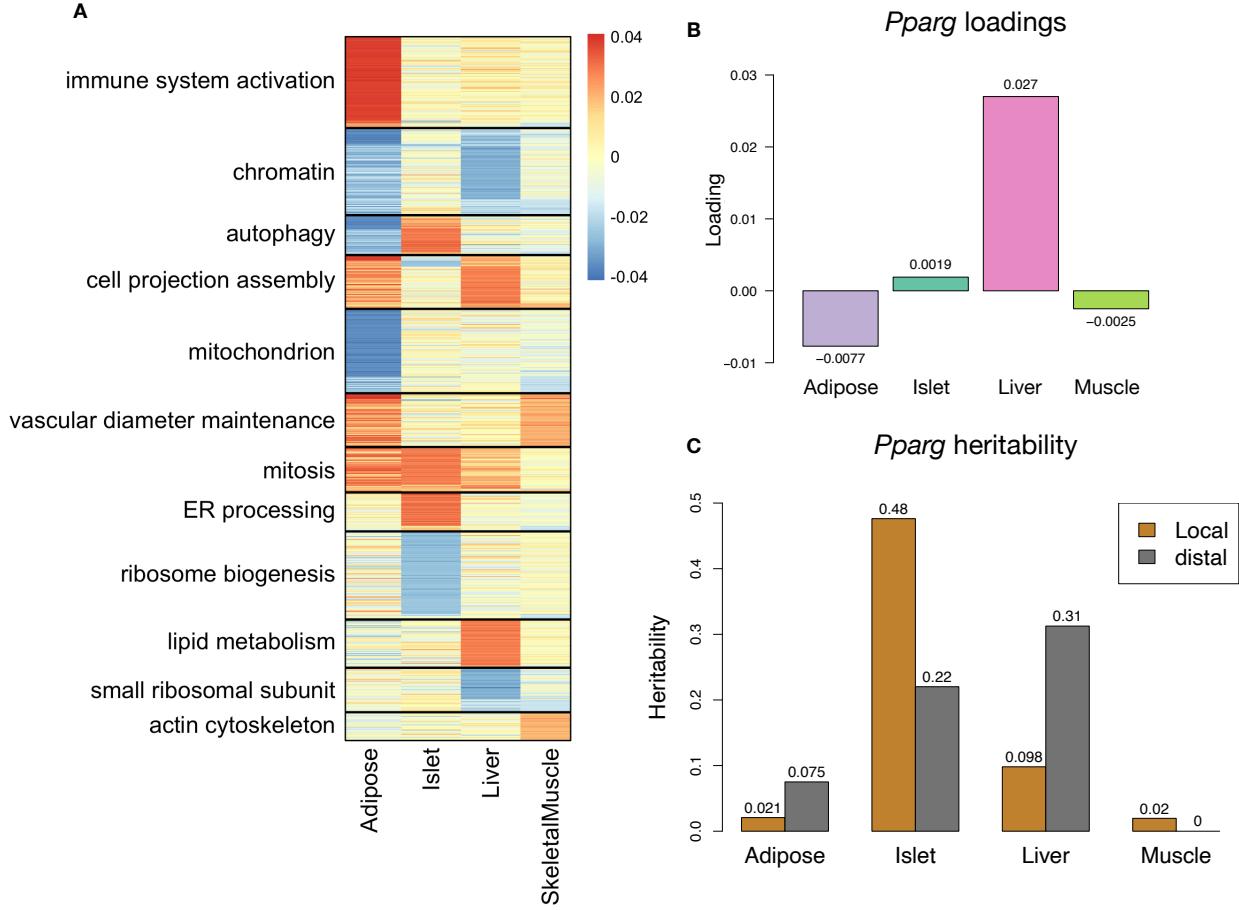


Figure 6: Tissue-specific transcriptional programs were associated with obesity and insulin resistance. **A** Heat map showing the loadings of all transcripts with loadings greater than 2.5 standard deviations from the mean in any tissue. The heat map was clustered using k medoid clustering. Functional 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.

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

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

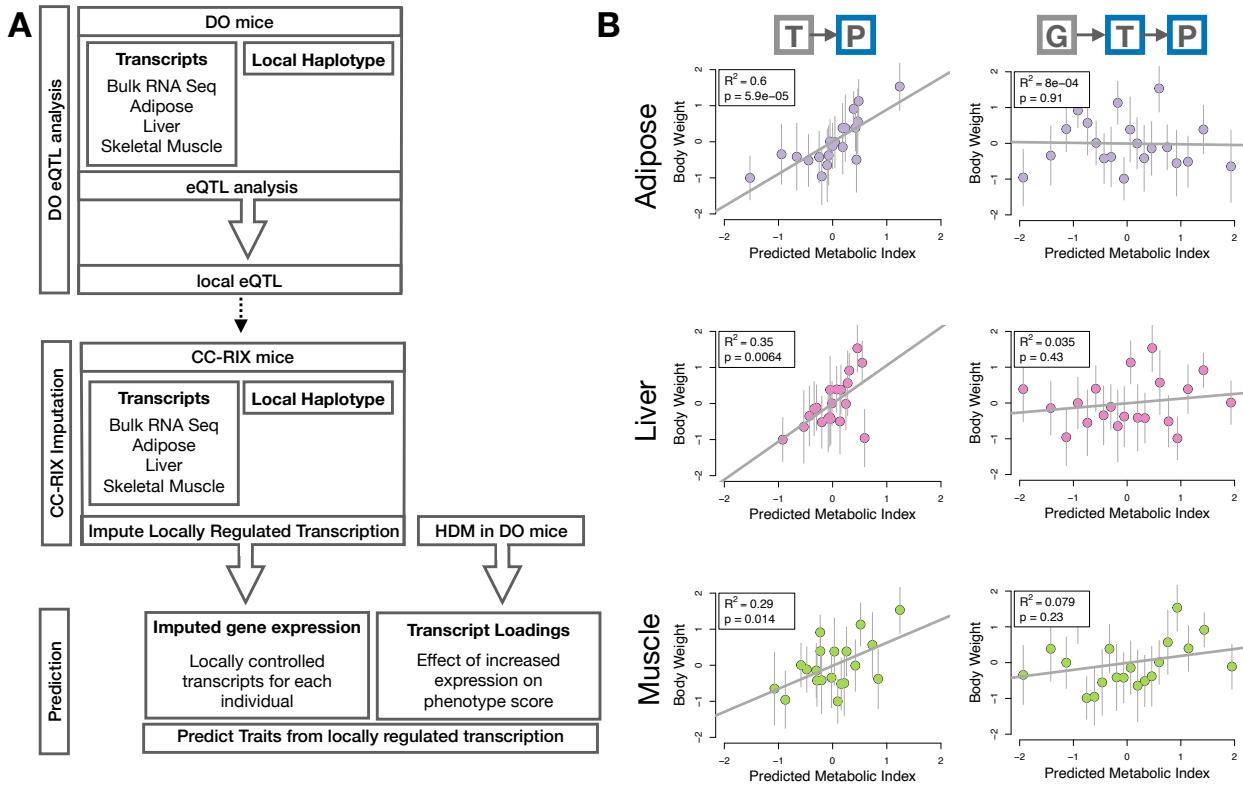


Figure 7: Transcription, but not local genotype, predicts phenotype in the CC-RIX. **A.** Workflow showing procedure for translating HDMA results to an independent population of mice. **B.** Relationships between the predicted metabolic disease index (MDI) and measured body weight. The left column shows 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.

292 expression in the CC-RIX using local genotype and were able to estimate variation in gene transcription
 293 robustly (Supp. Fig. S8). However, these imputed values failed to predict body weight in the CC-RIX when
 294 weighted with the loadings from HDMA. (Fig. 7B right column). This result suggests that local regulation of
 295 gene expression is not the primary factor driving heritability of complex traits. It is also consistent with our
 296 findings in the DO population that distal heritability was a major driver of trait-relevant gene expression and
 297 that high-loading transcripts had comparatively high distal and low local heritability.

298 **Distally heritable transcriptomic signatures reflected variation in composition of adipose tissue
 299 and islets**

300 The interpretation of global genetic influences on gene expression and phenotype is potentially more challenging
 301 than the interpretation and translation of local genetic influences, as genetic effects cannot be localized to
 302 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.

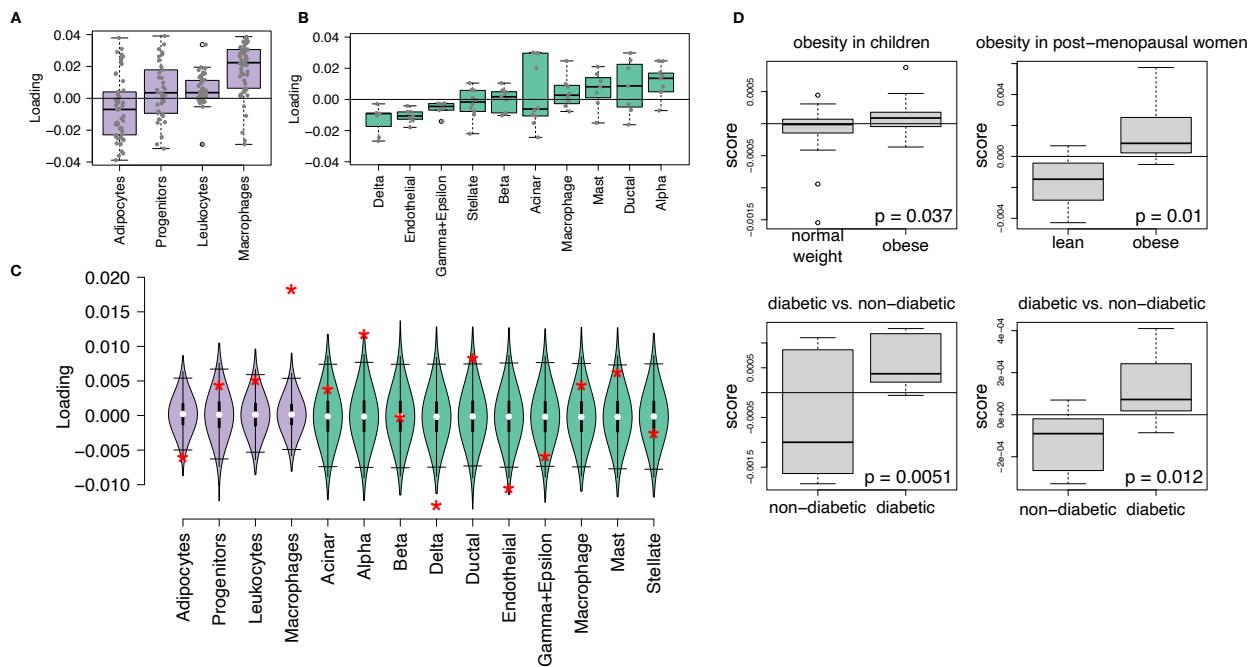


Figure 8: HDMA results translate to humans. **A.** Distribution of loadings for cell-type-specific transcripts in adipose tissue. **B.** Distribution of loadings for cell-type-specific transcripts in pancreatic islets (green). **C.** Null distributions for the mean loading of randomly selected transcripts 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 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

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

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

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

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

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

338 Another **potential** application of the transcript loading landscape is in ranking potential drug candidates for the
339 treatment of metabolic disease. Although high-loading transcripts may be good candidates for understanding
340 specific biology related to obesity, the transcriptome overall is highly interconnected and redundant. The
341 ConnectivityMap (CMAP) database [59–59;60](#) developed by the Broad Institute allows querying thousands of
342 compounds that reverse or enhance the extreme ends of transcriptomic signatures in multiple different cell
343 types. By identifying drugs that reverse pathogenic transcriptomic signatures, we can potentially identify
344 compounds ~~that~~th~~Fat~~ have favorable effects on gene expression. To test this hypothesis, we queried the
345 CMAP database through the CLUE online query tool (<https://clue.io/query/>, version 1.1.1.43) (Methods).
346 We identified top anti-correlated hits across all cell types (Supp. Figs S9 and S10). To get more tissue-specific

347 results, we also looked at top results in cell types that most closely resembled our tissues. We looked at
348 results in adipocytes (ASC) as well as pancreatic tumor cells (YAPC) regardless of *p* value (Supp. Figs S11
349 and S12).

350 Looking across all cell types, the notable top hits from the adipose tissue loadings included mTOR inhibitors
351 and glucocorticoid agonists (Supp. Fig. S9). It is thought that metformin, which is commonly used to
352 improve glycemic control, acts, at least in part, by inhibiting mTOR signaling^{61:62}. However, long-term use
353 of other mTOR inhibitors, such as rapamycin, are known to cause insulin resistance and β -cell toxicity⁶²⁻⁶⁴.
354 Glucocorticoids are used to reduce inflammation, which was a prominent signature in the adipose tissue,
355 but these drugs also promote hyperglycemia and diabetes^{65:66}. Acute treatment with glucocorticoids has
356 further been shown to reduce thermogenesis in rodent adipocytes⁶⁷⁻⁶⁹, but increase thermogenesis in human
357 adipocytes^{70:71}. Thus, the pathways identified by CMAP across all cell types were highly related to the
358 transcript loading profiles, but the relationship was not a simple reversal.

359 The top hit for the adipose composite transcript in CMAP adipocytes was a PARP inhibitor (Supp. Fig.
360 S11). PARPs play a role in lipid metabolism and are involved in the development of obesity and diabetes⁷².
361 PARP1 inhibition increases mitochondrial biogenesis⁷³. Inhibition of PARP1 activity can further prevent
362 necrosis in favor of the less inflammatory apoptosis⁷⁴, thereby potentially reducing inflammation in stressed
363 adipocytes. Other notable hits among the top 20 were BTK inhibitors, which have been observed to suppress
364 inflammation and improve insulin resistance⁷⁵. The CMAP database identified both known diabetes drugs
365 (e.g. sulfonylureas), as well as to reduce insulin antibodies in type I diabetes⁷⁶. IkappaB kinase (IKK) is an
366 enzyme complex involved in regulating cellular responses to inflammation⁷⁷. Inhibitors of IKK have been
367 shown to improve glucose control in type II diabetes^{78:79}.

368 Among the top most significant hits for the transcript loadings from pancreatic islets (Supp. Fig. S10),
369 was suppression of T cell receptor signaling, which is drugs that target pathways known to be involved in
370 Type 1 diabetes⁸⁰, as well as TNFR1, which has been associated with mortality in diabetes patients⁸¹.
371 Suppression of NOD1/2 signaling was also among the top hits. NOD1 and 2 sense ER stress^{82:83}, which
372 is associated with β -cell death in type 1 and type 2 diabetes⁸⁴. This cell death process is dependent on
373 NOD1/2 signaling⁸², although the specifics have not yet been worked out.

374 We also looked specifically at hits in pancreatic tumor cells (YAPC) regardless of significance level to get
375 a transcriptional response more specific to the pancreas (Supp. Fig. S12). Hits in this list included widely
376 used diabetes drugs, such as sulfonylureas, PPAR receptor agonists, and insulin sensitizers. Rosiglitazone is
377 a PPAR- γ agonist and was one of the most prescribed drugs for type 2 diabetes before its use was reduced

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

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

390 Discussion

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

401 Supplemental Discussion

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

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

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

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

428 An extension of the omnigenic model proposed that most heritability of complex traits is driven by weak
429 distal eQTLs that are potentially below the detection threshold in studies with feasible sample sizes⁸⁷. This
430 is consistent with what we observed here. For example, *Nucb2*, had a high loading in islets and was also
431 strongly distally regulated (66% distal heritability) (Fig. 5). This gene is expressed in pancreatic β cells and
432 is involved in insulin and glucagon release^{94–96}. Although its transcription was highly heritable in islets, that
433 regulation was distributed across the genome, with no clear distal eQTL (Supp. Fig. S13). Thus, although
434 distal regulation of some genes may be strong, this regulation is likely to be highly complex and not easily
435 localized.

436 Individual high-loading transcripts also demonstrated biologically interpretable, tissue-specific patterns. We
437 highlighted *Pparg*, which is known to be protective in adipose tissue⁴⁹ where it was negatively loaded, and
438 harmful in the liver^{50–54}, where it was positively loaded. Such granular patterns may be useful in generating

439 hypotheses for further testing, and prioritizing genes as therapeutic targets. The tissue-specific nature of
440 the loadings also may provide clues to tissue-specific effects, or side effects, of targeting particular genes
441 system-wide.

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

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

458 Data and Code Availability

459 **DO mice:** Genotypes, phenotypes, and pancreatic islet gene expression data were previously published¹².
460 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
461 - GSE266569; DO skeletal muscle - GSE266567. Expression data with calculated eQTLs are available at
462 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)

466 **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

469 Figshare https://figshare.com/articles/dataset/Data_and_code_for_High-Dimensional_Mediation_Anal
470 ysis_HDMA_in_diversity_outbred_mice/27066979 DOI: 10.6084/m9.figshare.27066979

471 **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
472 _mice/27066979 DOI: 10.6084/m9.figshare.27066979
473

474 **Acknowledgements**

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

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

483 **Supplemental Figures**

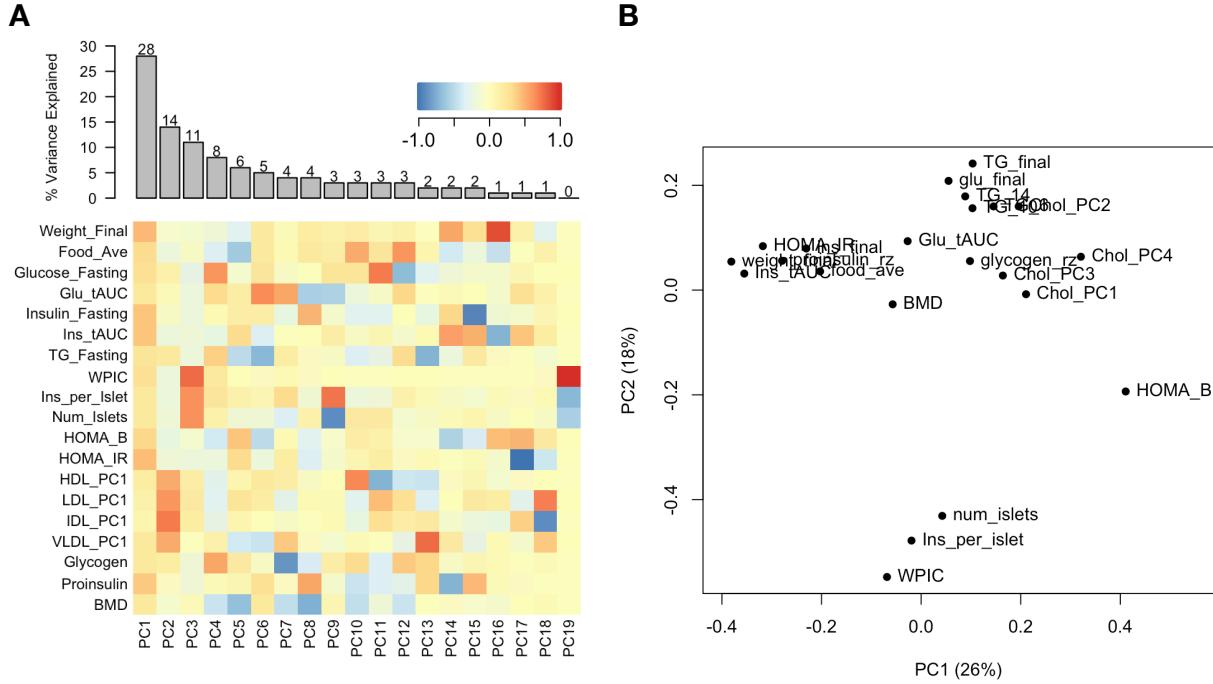


Figure S1: Trait matrix decomposition. **A** The heat map shows the loadings of each trait onto each principal component of the trait matrix. The bars at the top show the percent variance explained for each principal component. **B** Traits plotted by the first and second principal components of the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

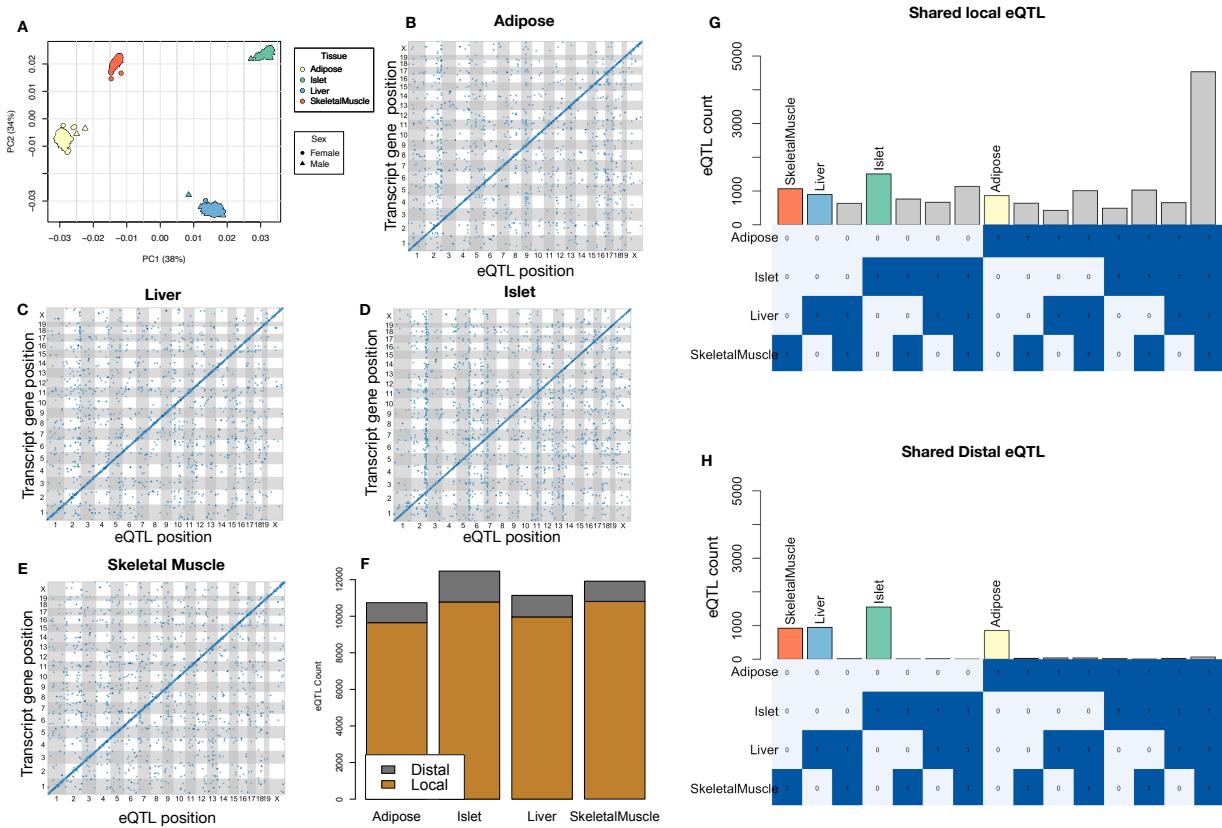


Figure S2: Overview of eQTL analysis in DO mice. **A.** RNA seq samples from the four different tissues clustered by tissue. **B.-E.** eQTL maps are shown for each tissue. The *x*-axis shows the position of the mapped eQTL, and the *y*-axis shows the physical position of the gene encoding each mapped 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.

KEGG pathway enrichments by GSEA

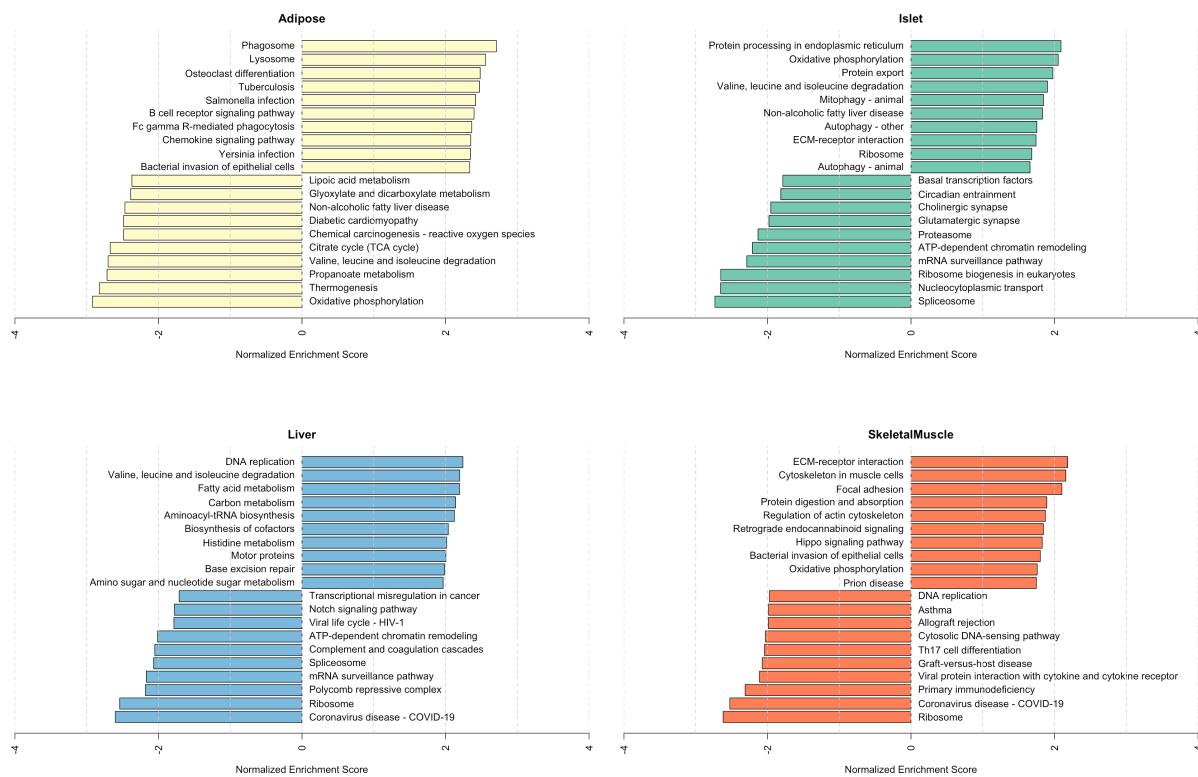


Figure S3: Bar plots showing normalized enrichment scores (NES) for KEGG pathways as determined by fast gene score enrichment analysis (fgsea). Only the top 10 positive and top 10 negative scores are shown. Colors indicate tissue. The name beside each bar shows the name of each enriched KEGG pathway.

Top GO term enrichments by GSEA

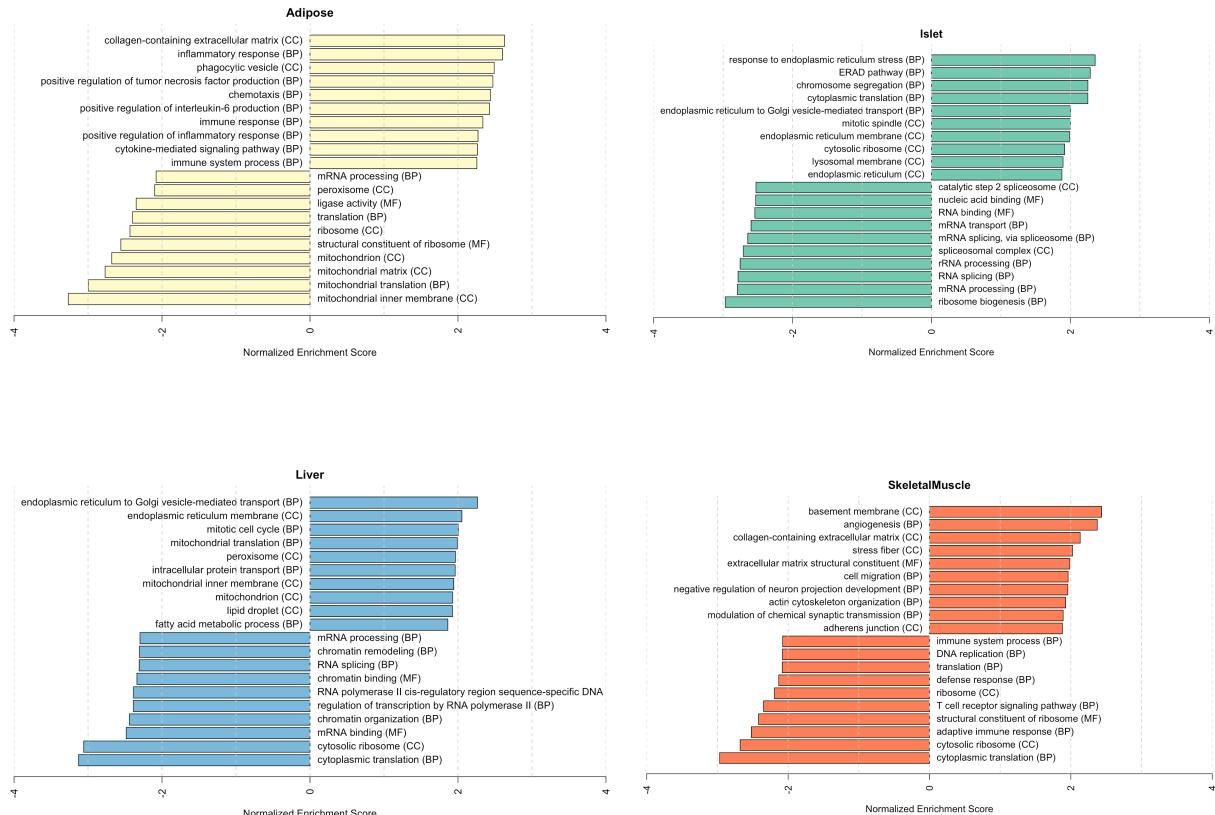


Figure S4: Bar plots showing normalized enrichment scores (NES) for GO terms as determined by fast gene score enrichment analysis (fgsea). Only the top 10 positive and top 10 negative scores are shown. Colors indicate tissue. The name beside each bar shows the name of each enriched GO term. The letters in parentheses indicate whether the term is from the biological process ontology (BP), the molecular function ontology (MF), or the cellular compartment ontology (CC).

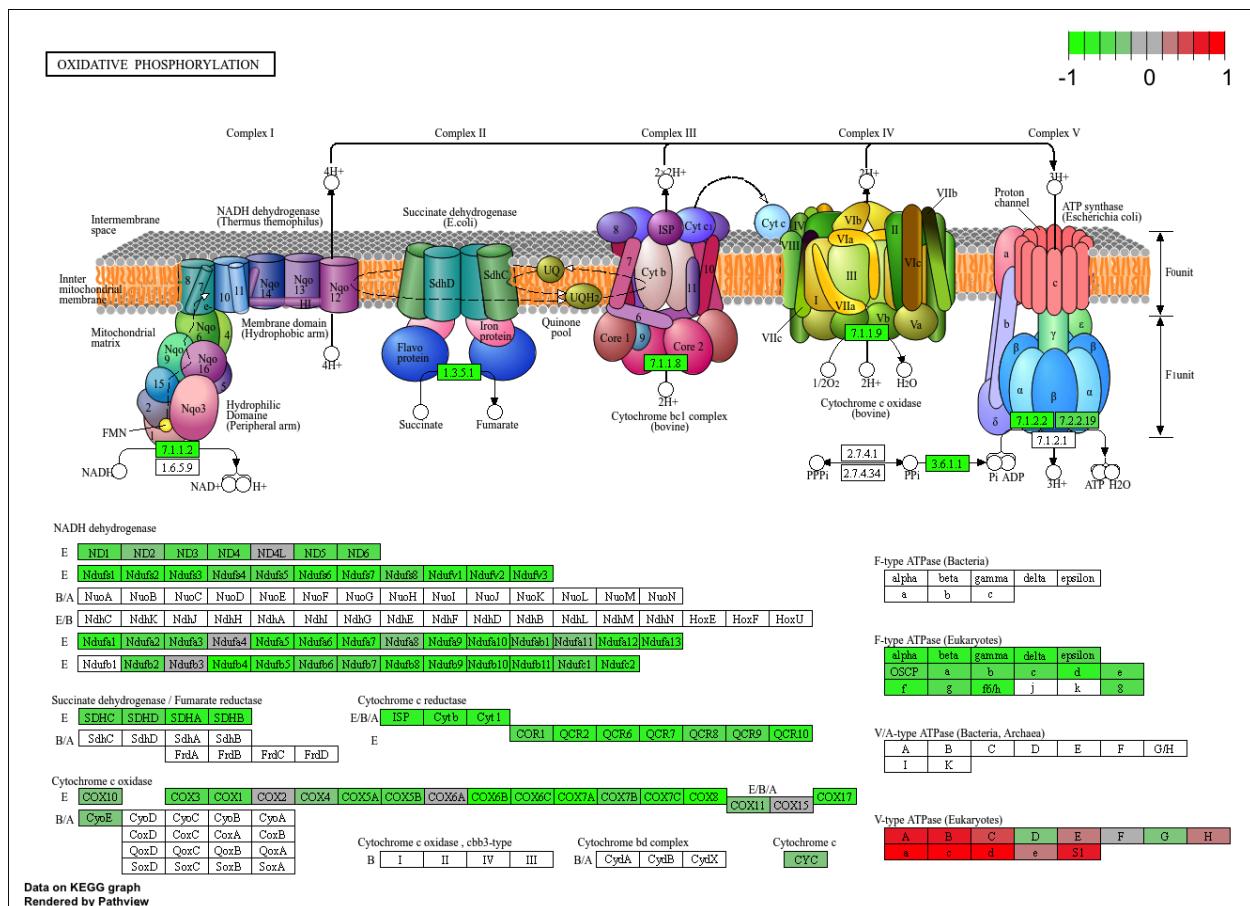


Figure S5: The KEGG pathway for oxidative phosphorylation in mice. Each element is colored based on its HDMA loading from adipose tissue normalized to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Almost the entire pathway was strongly negatively loaded indicating that increased expression of genes involved in oxidative phosphorylation was associated with reduced MDI.

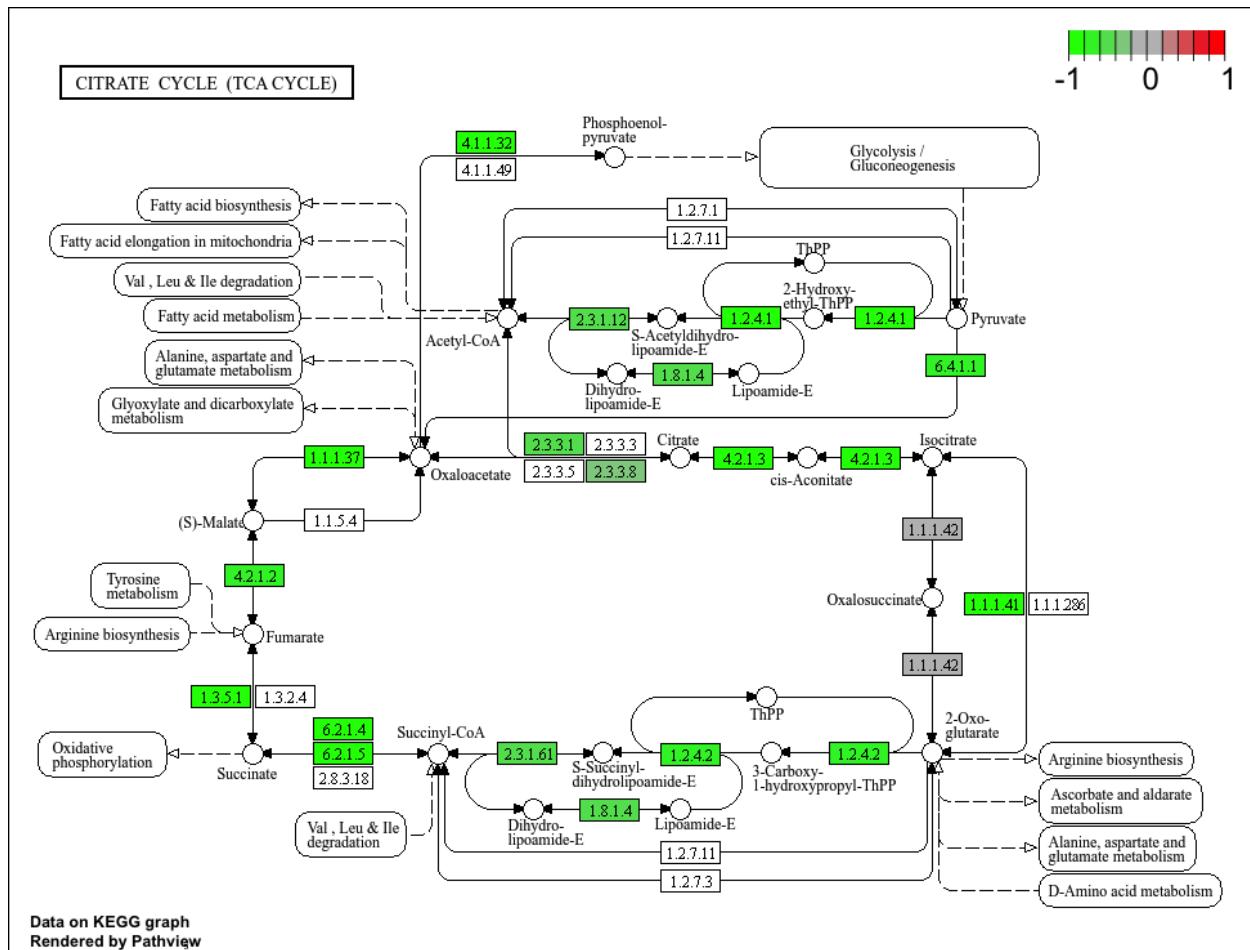


Figure S6: The KEGG pathway for the TCA (citric acid) cycle in mice. Each element is colored based on its HDMA loading from adipose tissue normalized to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Many genes in the cycle were strongly negatively loaded indicating that increased expression of genes involved in branched-chain amino acid degradation was associated with reduced MDI.

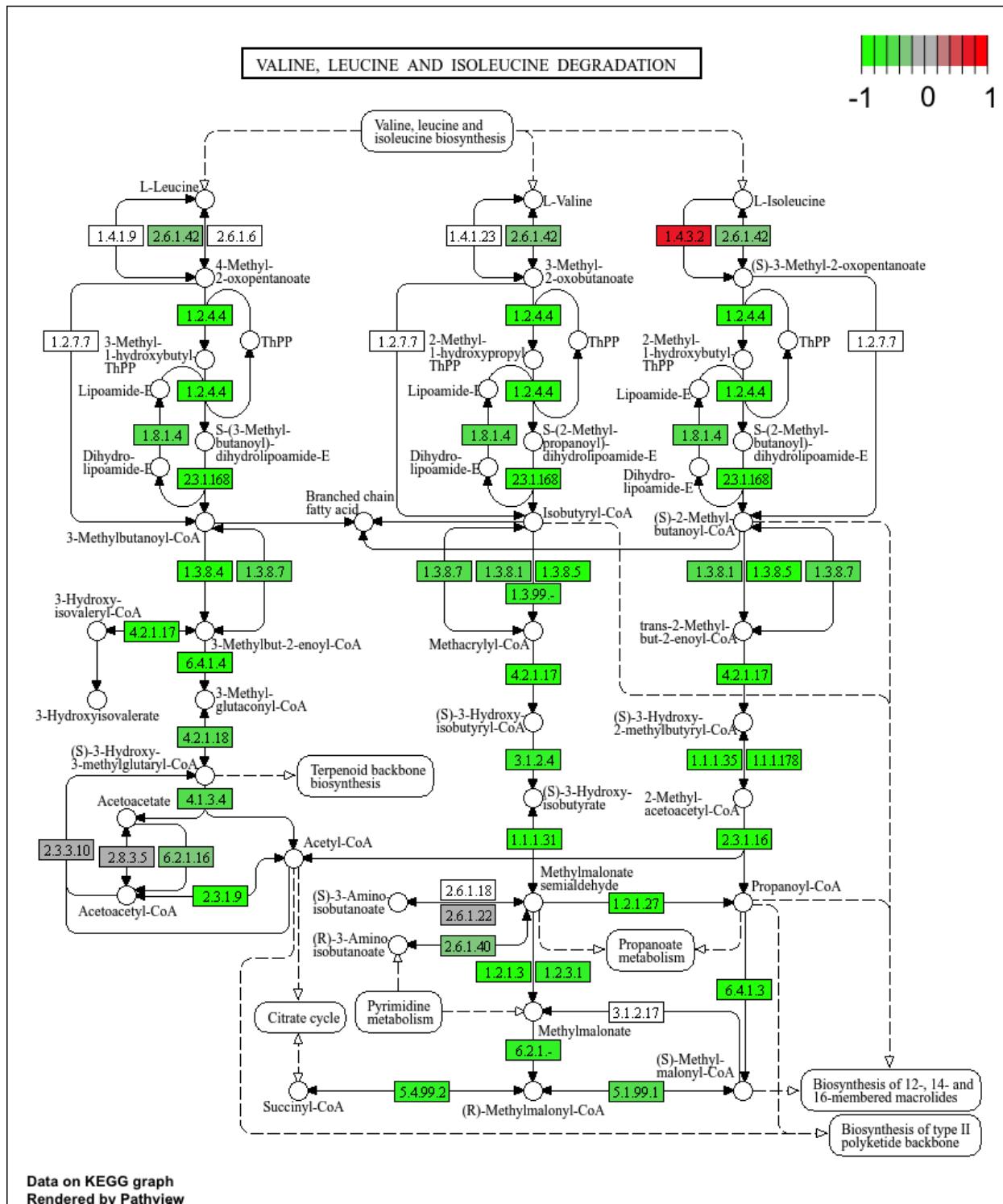


Figure S7: The KEGG pathway for branched-chain amino acid degradation in mice. Each element is colored based on its HDMA loading from adipose tissue normalized to run from -1 to 1. Genes highlighted in green had negative loadings, and those highlighted in red had positive loadings. Almost the entire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

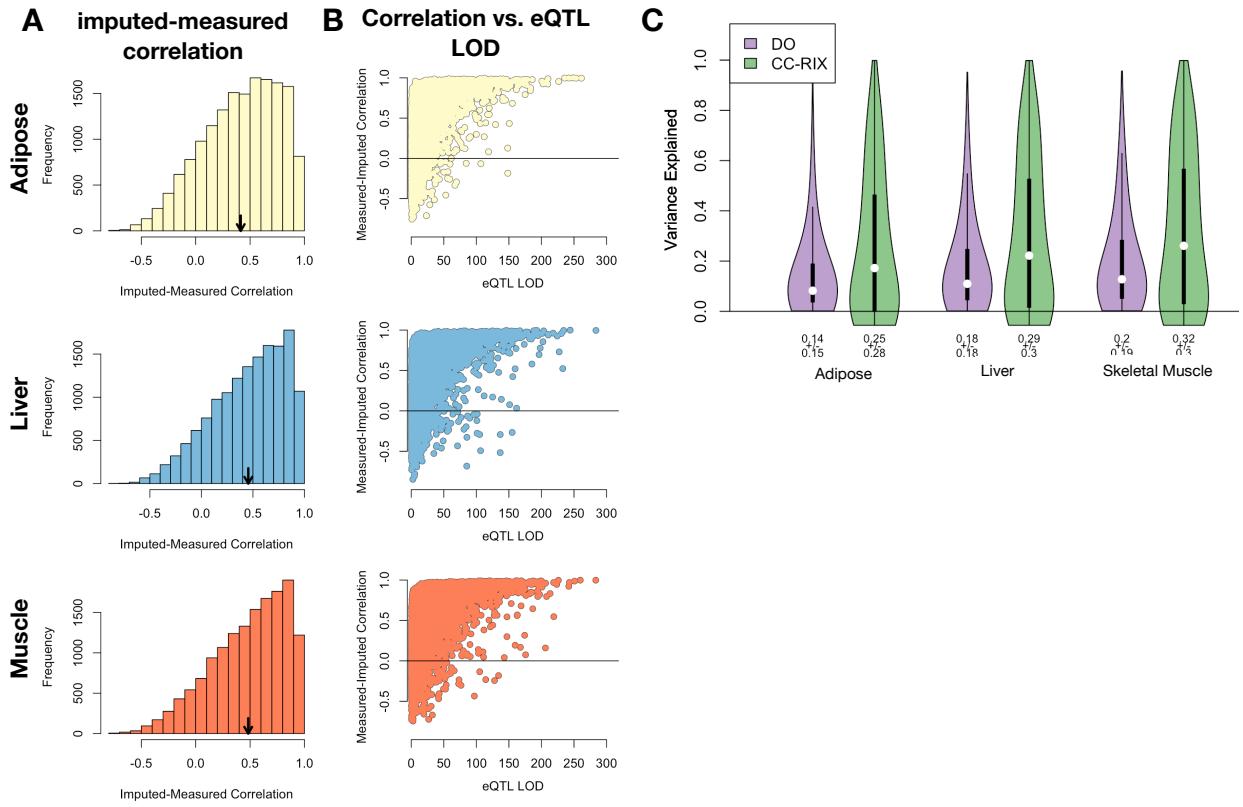


Figure S8: Validation of transcript imputation in the CC-RIX. **A.** Distributions of correlations between imputed and measured transcripts in the CC-RIX. The mean of each distribution is shown by the red line. All distributions were skewed toward positive correlations and had positive means near a Pearson 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 explained by local genotype 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 S9: 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 S10: 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 S11: 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 S12: 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.

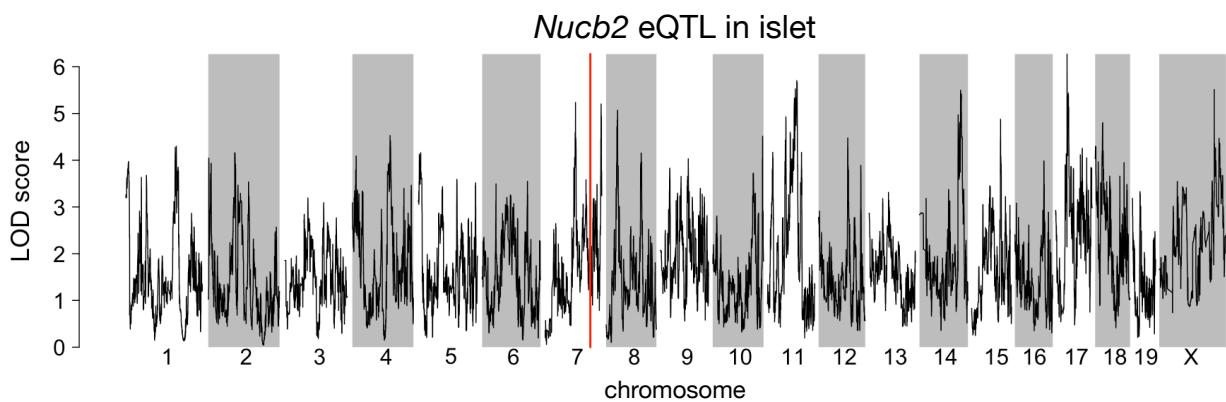


Figure S13: Regulation of *Nucb2* expression in islet. *Nucb2* is encoded on mouse chromosome 7 at 116.5 Mb (red line). In islets the heritability of *Nucb2* expression levels is 69% heritable. This LOD score trace shows that there is no local eQTLs at the position of the gene, nor any strong distal eQTL anywhere else in the genome.

484 **References**

- 485 [1] M. T. Maurano, R. Humbert, E. Rynes, R. E. Thurman, E. Haugen, H. Wang, A. P. Reynolds,
486 R. Sandstrom, H. Qu, J. Brody, A. Shafer, F. Neri, K. Lee, T. Kutyavin, S. Stehling-Sun, A. K.
487 Johnson, T. K. Canfield, E. Giste, M. Diegel, D. Bates, R. S. Hansen, S. Neph, P. J. Sabo, S. Heimfeld,
488 A. Raubitschek, S. Ziegler, C. Cotsapas, N. Sotoodehnia, I. Glass, S. R. Sunyaev, R. Kaul, and J. A.
489 Stamatoyannopoulos. Systematic localization of common disease-associated variation in regulatory DNA.
490 *Science*, 337(6099):1190–1195, Sep 2012.
- 491 [2] K. K. Farh, A. Marson, J. Zhu, M. Kleinewietfeld, W. J. Housley, S. Beik, N. Shores, H. Whitton, R. J.
492 Ryan, A. A. Shishkin, M. Hatan, M. J. Carrasco-Alfonso, D. Mayer, C. J. Luckey, N. A. Patsopoulos,
493 P. L. De Jager, V. K. Kuchroo, C. B. Epstein, M. J. Daly, D. A. Hafler, and B. E. Bernstein. Genetic
494 and epigenetic fine mapping of causal autoimmune disease variants. *Nature*, 518(7539):337–343, Feb
495 2015.
- 496 [3] E. Pennisi. The Biology of Genomes. Disease risk links to gene regulation. *Science*, 332(6033):1031, May
497 2011.
- 498 [4] L. A. Hindorff, P. Sethupathy, H. A. Junkins, E. M. Ramos, J. P. Mehta, F. S. Collins, and T. A. Manolio.
499 Potential etiologic and functional implications of genome-wide association loci for human diseases and
500 traits. *Proc Natl Acad Sci*, 106(23):9362–9367, Jun 2009.
- 501 [5] J. K. Pickrell. Joint analysis of functional genomic data and genome-wide association studies of 18
502 human traits. *Am J Hum Genet*, 94(4):559–573, Apr 2014.
- 503 [6] D. Welter, J. MacArthur, J. Morales, T. Burdett, P. Hall, H. Junkins, A. Klemm, P. Flieck, T. Manolio,
504 L. Hindorff, and H. Parkinson. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations.
505 *Nucleic Acids Res*, 42(Database issue):D1001–1006, Jan 2014.
- 506 [7] Y. I. Li, B. van de Geijn, A. Raj, D. A. Knowles, A. A. Petti, D. Golan, Y. Gilad, and J. K. Pritchard.
507 RNA splicing is a primary link between genetic variation and disease. *Science*, 352(6285):600–604, Apr
508 2016.
- 509 [8] D. Zhou, Y. Jiang, X. Zhong, N. J. Cox, C. Liu, and E. R. Gamazon. A unified framework for joint-tissue
510 transcriptome-wide association and Mendelian randomization analysis. *Nat Genet*, 52(11):1239–1246,
511 Nov 2020.
- 512 [9] E. R. Gamazon, H. E. Wheeler, K. P. Shah, S. V. Mozaffari, K. Aquino-Michaels, R. J. Carroll, A. E.

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

- 543 [20] X. Liu, J. A. Mefford, A. Dahl, Y. He, M. Subramaniam, A. Battle, A. L. Price, and N. Zaitlen. GBAT:
544 a gene-based association test for robust detection of trans-gene regulation. *Genome Biol*, 21(1):211, Aug
545 2020.
- 546 [21] H. J. Westra, M. J. Peters, T. Esko, H. Yaghoobkar, C. Schurmann, J. Kettunen, M. W. Christiansen,
547 B. P. Fairfax, K. Schramm, J. E. Powell, A. Zhernakova, D. V. Zhernakova, J. H. Veldink, L. H. Van den
548 Berg, J. Karjalainen, S. Withoff, A. G. Uitterlinden, A. Hofman, F. Rivadeneira, P. A. C. ' Hoen,
549 E. Reinmaa, K. Fischer, M. Nelis, L. Milani, D. Melzer, L. Ferrucci, A. B. Singleton, D. G. Hernandez,
550 M. A. Nalls, G. Homuth, M. Nauck, D. Radke, U. Iker, M. Perola, V. Salomaa, J. Brody, A. Suchy-Dicey,
551 S. A. Gharib, D. A. Enquobahrie, T. Lumley, G. W. Montgomery, S. Makino, H. Prokisch, C. Herder,
552 M. Roden, H. Grallert, T. Meitinger, K. Strauch, Y. Li, R. C. Jansen, P. M. Visscher, J. C. Knight,
553 B. M. Psaty, S. Ripatti, A. Teumer, T. M. Frayling, A. Metspalu, J. B. J. van Meurs, and L. Franke.
554 Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet*,
555 45(10):1238–1243, Oct 2013.
- 556 [22] Y. Gilad, S. A. Rifkin, and J. K. Pritchard. Revealing the architecture of gene regulation: the promise
557 of eQTL studies. *Trends Genet*, 24(8):408–415, Aug 2008.
- 558 [23] Aparna Nathan, Samira Asgari, Kazuyoshi Ishigaki, Cristian Valencia, Tiffany Amariuta, Yang Luo,
559 Jessica I Beynor, Yuriy Baglaenko, Sara Suliman, Alkes L Price, et al. Single-cell eqtl models reveal
560 dynamic t cell state dependence of disease loci. *Nature*, 606(7912):120–128, 2022.
- 561 [24] E. Sollis, A. Mosaku, A. Abid, A. Buniello, M. Cerezo, L. Gil, T. Groza, O. §, P. Hall, J. Hayhurst,
562 A. Ibrahim, Y. Ji, S. John, E. Lewis, J. A. L. MacArthur, A. McMahon, D. Osumi-Sutherland,
563 K. Panoutsopoulou, Z. Pendlington, S. Ramachandran, R. Stefancsik, J. Stewart, P. Whetzel, R. Wilson,
564 L. Hindorff, F. Cunningham, S. A. Lambert, M. Inouye, H. Parkinson, and L. W. Harris. The NHGRI-EBI
565 GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res*, 51(D1):D977–D985, Jan
566 2023.
- 567 [25] R. J. F. Loos and G. S. H. Yeo. The genetics of obesity: from discovery to biology. *Nat Rev Genet*,
568 23(2):120–133, Feb 2022.
- 569 [26] R. K. Singh, P. Kumar, and K. Mahalingam. Molecular genetics of human obesity: A comprehensive
570 review. *C R Biol*, 340(2):87–108, Feb 2017.
- 571 [27] P. Arner. Obesity—a genetic disease of adipose tissue? *Br J Nutr*, 83 Suppl 1:9–16, Mar 2000.
- 572 [28] Mark P Keller, Mary E Rabaglia, Kathryn L Schueler, Donnie S Stapleton, Daniel M Gatti, Matthew

- 573 Vincent, Kelly A Mitok, Ziyue Wang, Takanao Ishimura, Shane P Simonett, et al. Gene loci associated
574 with insulin secretion in islets from nondiabetic mice. *The Journal of Clinical Investigation*, 129(10):4419–
575 4432, 2019.
- 576 [29] G. A. Churchill, D. M. Gatti, S. C. Munger, and K. L. Svenson. The Diversity Outbred mouse population.
577 *Mamm Genome*, 23(9-10):713–718, Oct 2012.
- 578 [30] Elissa J Chesler, Darla R Miller, Lisa R Branstetter, Leslie D Galloway, Barbara L Jackson, Vivek M
579 Philip, Brynn H Voy, Cymbeline T Culiat, David W Threadgill, Robert W Williams, et al. The
580 collaborative cross at oak ridge national laboratory: developing a powerful resource for systems genetics.
581 *Mammalian Genome*, 19:382–389, 2008.
- 582 [31] Michael C Saul, Vivek M Philip, Laura G Reinholdt, and Elissa J Chesler. High-diversity mouse
583 populations for complex traits. *Trends in Genetics*, 35(7):501–514, 2019.
- 584 [32] D. W. Threadgill, D. R. Miller, G. A. Churchill, and F. P. de Villena. The collaborative cross: a
585 recombinant inbred mouse population for the systems genetic era. *ILAR J*, 52(1):24–31, 2011.
- 586 [33] S. M. Clee and A. D. Attie. The genetic landscape of type 2 diabetes in mice. *Endocr Rev*, 28(1):48–83,
587 Feb 2007.
- 588 [34] K. W. Broman, D. M. Gatti, P. Simecek, N. A. Furlotte, P. Prins, Š. Sen, B. S. Yandell, and G. A.
589 Churchill. R/qt12: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and
590 Multiparent Populations. *Genetics*, 211(2):495–502, Feb 2019.
- 591 [35] Klaasjan G Ouwens, Rick Jansen, Michel G Nivard, Jenny van Dongen, Maia J Frieser, Jouke-Jan
592 Hottenga, Wibowo Arindrarto, Annique Claringbould, Maarten van Iterson, Hailiang Mei, et al. A
593 characterization of cis-and trans-heritability of rna-seq-based gene expression. *European Journal of
594 Human Genetics*, 28(2):253–263, 2020.
- 595 [36] Alkes L Price, Agnar Helgason, Gudmar Thorleifsson, Steven A McCarroll, Augustine Kong, and Kari
596 Stefansson. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related
597 or unrelated individuals. *PLoS genetics*, 7(2):e1001317, 2011.
- 598 [37] Julien Bryois, Alfonso Buil, David M Evans, John P Kemp, Stephen B Montgomery, Donald F Conrad,
599 Karen M Ho, Susan Ring, Matthew Hurles, Panos Deloukas, et al. Cis and trans effects of human
600 genomic variants on gene expression. *PLoS genetics*, 10(7):e1004461, 2014.
- 601 [38] M. Helmer, S. Warrington, A. R. Mohammadi-Nejad, J. L. Ji, A. Howell, B. Rosand, A. Anticevic,

- 602 S. N. Sotiropoulos, and J. D. Murray. On the stability of canonical correlation analysis and partial least
603 squares with application to brain-behavior associations. *Commun Biol*, 7(1):217, Feb 2024.
- 604 [39] Gennady Korotkevich, Vladimir Sukhov, and Alexey Sergushichev. Fast gene set enrichment analysis.
605 *bioRxiv*, 2019.
- 606 [40] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich,
607 S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: a
608 knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*,
609 102(43):15545–15550, Oct 2005.
- 610 [41] S. Subramanian and A. Chait. The effect of dietary cholesterol on macrophage accumulation in adipose
611 tissue: implications for systemic inflammation and atherosclerosis. *Curr Opin Lipidol*, 20(1):39–44, Feb
612 2009.
- 613 [42] I. Akoumianakis, N. Akawi, and C. Antoniades. Exploring the Crosstalk between Adipose Tissue and
614 the Cardiovascular System. *Korean Circ J*, 47(5):670–685, Sep 2017.
- 615 [43] I. S. Stafeev, A. V. Vorotnikov, E. I. Ratner, M. Y. Menshikov, and Y. V. Parfyonova. Latent Inflammation
616 and Insulin Resistance in Adipose Tissue. *Int J Endocrinol*, 2017:5076732, 2017.
- 617 [44] I. P. Fischer, M. Irmler, C. W. Meyer, S. J. Sachs, F. Neff, M. de Angelis, J. Beckers, M. H. p, S. M.
618 Hofmann, and S. Ussar. A history of obesity leaves an inflammatory fingerprint in liver and adipose
619 tissue. *Int J Obes (Lond)*, 42(3):507–517, Mar 2018.
- 620 [45] S. Chung, H. Cuffe, S. M. Marshall, A. L. McDaniel, J. H. Ha, K. Kavanagh, C. Hong, P. Tontonoz,
621 R. E. Temel, and J. S. Parks. Dietary cholesterol promotes adipocyte hypertrophy and adipose tissue
622 inflammation in visceral, but not in subcutaneous, fat in monkeys. *Arterioscler Thromb Vasc Biol*,
623 34(9):1880–1887, Sep 2014.
- 624 [46] V. Kus, T. Prazak, P. Brauner, M. Hensler, O. Kuda, P. Flachs, P. Janovska, D. Medrikova, M. Rossmeisl,
625 Z. Jilkova, B. Stefl, E. Pastalkova, Z. Drahota, J. Houstek, and J. Kopecky. Induction of muscle
626 thermogenesis by high-fat diet in mice: association with obesity-resistance. *Am J Physiol Endocrinol
Metab*, 295(2):E356–367, Aug 2008.
- 628 [47] C. B. Newgard. Interplay between lipids and branched-chain amino acids in development of insulin
629 resistance. *Cell Metab*, 15(5):606–614, May 2012.
- 630 [48] D. D. Sears, G. Hsiao, A. Hsiao, J. G. Yu, C. H. Courtney, J. M. Ofrecio, J. Chapman, and S. Subramaniam.

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

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

- 691 [64] Y. Gu, J. Lindner, A. Kumar, W. Yuan, and M. A. Magnuson. Rictor/mTORC2 is essential for
692 maintaining a balance between beta-cell proliferation and cell size. *Diabetes*, 60(3):827–837, Mar 2011.
- 693 [65] E. B. Geer, J. Islam, and C. Buettner. Mechanisms of glucocorticoid-induced insulin resistance: focus on
694 adipose tissue function and lipid metabolism. *Endocrinol Metab Clin North Am*, 43(1):75–102, Mar 2014.
- 695 [66] J. X. Li and C. L. Cummins. Fresh insights into glucocorticoid-induced diabetes mellitus and new
696 therapeutic directions. *Nat Rev Endocrinol*, 18(9):540–557, Sep 2022.
- 697 [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
700 inhibit UCP expression and function in brown adipocytes. *Am J Physiol Endocrinol Metab*, 280(4):E640–
701 649, Apr 2001.
- 702 [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,
703 P. Cheng, M. Hu, Y. Li, J. Bi, Y. Li, F. Hu, Y. Zhong, Y. Xu, and G. Ding. -HSD1 in regulating brown
704 adipocyte function. *J Mol Endocrinol*, 50(1):103–113, Feb 2013.
- 705 [70] L. E. Ramage, M. Akyol, A. M. Fletcher, J. Forsythe, M. Nixon, R. N. Carter, E. J. van Beek,
706 N. M. Morton, B. R. Walker, and R. H. Stimson. Glucocorticoids Acutely Increase Brown Adipose
707 Tissue Activity in Humans, Revealing Species-Specific Differences in UCP-1 Regulation. *Cell Metab*,
708 24(1):130–141, Jul 2016.
- 709 [71] J. L. Barclay, H. Agada, C. Jang, M. Ward, N. Wetzig, and K. K. Ho. Effects of glucocorticoids on
710 human brown adipocytes. *J Endocrinol*, 224(2):139–147, Feb 2015.
- 711 [72] M. ó, R. Gupte, W. L. Kraus, P. Pacher, and P. Bai. PARPs in lipid metabolism and related diseases.
712 *Prog Lipid Res*, 84:101117, Nov 2021.
- 713 [73] P. Bai, C. ó, H. Oudart, A. nszki, Y. Cen, C. Thomas, H. Yamamoto, A. Huber, B. Kiss, R. H.
714 Houtkooper, K. Schoonjans, V. Schreiber, A. A. Sauve, J. Menissier-de Murcia, and J. Auwerx. PARP-1
715 inhibition increases mitochondrial metabolism through SIRT1 activation. *Cell Metab*, 13(4):461–468,
716 Apr 2011.
- 717 [74] A. Chiarugi and M. A. Moskowitz. Cell biology. PARP-1—a perpetrator of apoptotic cell death? *Science*,
718 297(5579):200–201, Jul 2002.
- 719 [75] M. Althubiti, R. Almainani, S. Y. Eid, M. Elzubaier, B. Refaat, S. Idris, T. A. Alqurashi, and M. Z.

- 720 El-Readi. BTK targeting suppresses inflammatory genes and ameliorates insulin resistance. *Eur Cytokine Netw*, 31(4):168–179, Dec 2020.
- 721
- 722 [76] C. Skrabs, W. F. Pickl, T. Perkmann, U. ger, and A. Gessl. Rapid decline in insulin antibodies and
723 glutamic acid decarboxylase autoantibodies with ibrutinib therapy of chronic lymphocytic leukaemia. *J
724 Clin Pharm Ther*, 43(1):145–149, Feb 2018.
- 725 [77] Hans Hacker and Michael Karin. Regulation and function of ikk and ikk-related kinases. *Science's
726 STKE*, 2006(357):re13–re13, 2006.
- 727 [78] E. A. Oral, S. M. Reilly, A. V. Gomez, R. Meral, L. Butz, N. Ajluni, T. L. Chenevert, E. Korytnaya,
728 A. H. Neidert, R. Hench, D. Rus, J. F. Horowitz, B. Poirier, P. Zhao, K. Lehmann, M. Jain, R. Yu,
729 C. Liddle, M. Ahmadian, M. Downes, R. M. Evans, and A. R. Saltiel. and TBK1 Improves Glucose
730 Control in a Subset of Patients with Type 2 Diabetes. *Cell Metab*, 26(1):157–170, Jul 2017.
- 731 [79] M. C. Arkan, A. L. Hevener, F. R. Greten, S. Maeda, Z. W. Li, J. M. Long, A. Wynshaw-Boris, G. Poli,
732 J. Olefsky, and M. Karin. IKK-beta links inflammation to obesity-induced insulin resistance. *Nat Med*,
733 11(2):191–198, Feb 2005.
- 734 [80] M. Clark, C. J. Kroger, Q. Ke, and R. M. Tisch. The Role of T Cell Receptor Signaling in the
735 Development of Type 1 Diabetes. *Front Immunol*, 11:615371, 2020.
- 736 [81] P. ndell, A. C. Carlsson, A. Larsson, O. Melander, T. Wessman, J. v, and T. Ruge. TNFR1 is associated
737 with short-term mortality in patients with diabetes and acute dyspnea seeking care at the emergency
738 department. *Acta Diabetol*, 57(10):1145–1150, Oct 2020.
- 739 [82] A. M. Keestra-Gounder, M. X. Byndloss, N. Seyffert, B. M. Young, A. vez Arroyo, A. Y. Tsai, S. A.
740 Cevallos, M. G. Winter, O. H. Pham, C. R. Tiffany, M. F. de Jong, T. Kerrinnes, R. Ravindran, P. A.
741 Luciw, S. J. McSorley, A. J. umler, and R. M. Tsolis. NOD1 and NOD2 signalling links ER stress with
742 inflammation. *Nature*, 532(7599):394–397, Apr 2016.
- 743 [83] A. M. Keestra-Gounder and R. M. Tsolis. NOD1 and NOD2: Beyond Peptidoglycan Sensing. *Trends
744 Immunol*, 38(10):758–767, Oct 2017.
- 745 [84] J. Montane, L. Cadavez, and A. Novials. Stress and the inflammatory process: a major cause of
746 pancreatic cell death in type 2 diabetes. *Diabetes Metab Syndr Obes*, 7:25–34, 2014.
- 747 [85] B. B. Kahn and T. E. McGraw. , and type 2 diabetes. *N Engl J Med*, 363(27):2667–2669, Dec 2010.

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

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