

# 1 **3D genomic features across >50 diverse cell types reveal**

## 2 **insights into the genomic architecture of childhood obesity**

3

4 Khanh B. Trang, PhD<sup>1,2</sup>; Matthew C. Pahl, PhD<sup>1,2</sup>; James A. Pippin, BA<sup>1,2</sup>; Chun Su, PhD<sup>1,2</sup>; Sheridan H.

5 Littleton, BS<sup>1,2,3,4</sup>; Prabhat Sharma, PhD<sup>1,5</sup>; Nikhil N. Kulkarni, PhD<sup>1,5</sup>; Louis R. Ghanem, MD, PhD<sup>6</sup>;

6 Natalie A. Terry, MD, PhD<sup>6</sup>; Joan M. O'Brien, MD<sup>7,8</sup>; Yadav Wagley, PhD<sup>9</sup>; Kurt D. Hankenson, DVM, MS,

7 PhD<sup>9</sup>; Ashley Jermusyk, PhD<sup>10</sup>; Jason W. Hoskins, PhD<sup>10</sup>; Laufey T. Amundadottir, PhD<sup>10</sup>; Mai Xu, PhD<sup>10</sup>;

8 Kevin M Brown, PhD<sup>10</sup>; Stewart A. Anderson, PhD<sup>11,12</sup>; Wenli Yang, PhD<sup>13,14</sup>; Paul M. Titchnell, PhD<sup>13,15</sup>;

9 Patrick Seale, PhD<sup>13,14</sup>; Laura Cook, PhD<sup>16,17,18</sup>; Megan K. Levings, PhD<sup>19,20,21</sup>; Babette S. Zemel, PhD<sup>6,22</sup>;

10 Alessandra Chesi, PhD<sup>1,23</sup>; Andrew D. Wells, PhD<sup>1,5,23,24</sup>; Struan F.A. Grant, PhD<sup>1,2,4,13,22,25,26</sup>

- 11
- 12 <sup>1.</sup> Center for Spatial and Functional Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
- 13 <sup>2.</sup> Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
- 14 <sup>3.</sup> Cell and Molecular Biology Graduate Group, Perelman School of Medicine, University of Pennsylvania,
- 15 Philadelphia, PA, USA
- 16 <sup>4.</sup> Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 17 <sup>5.</sup> Department of Pathology, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
- 18 <sup>6.</sup> Division of Gastroenterology, Hepatology, and Nutrition, Children's Hospital of Philadelphia, PA, USA
- 19 <sup>7.</sup> Scheie Eye Institute, Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania,
- 20 Philadelphia, Pennsylvania, PA, USA
- 21 <sup>8.</sup> Penn Medicine Center for Ophthalmic Genetics in Complex Disease
- 22 <sup>9.</sup> Department of Orthopedic Surgery University of Michigan Medical School Ann Arbor, MI, USA
- 23 <sup>10.</sup> Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute,
- 24 Bethesda, MD, USA
- 25 <sup>11.</sup> Department of Child and Adolescent Psychiatry, Children's Hospital of Philadelphia, Philadelphia, PA, USA
- 26 <sup>12.</sup> Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 27 <sup>13.</sup> Institute for Diabetes, Obesity and Metabolism, Perelman School of Medicine, University of Pennsylvania,
- 28 Philadelphia, PA, USA

- 29 <sup>14.</sup> Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania,  
30 Philadelphia, PA, USA
- 31 <sup>15.</sup> Department of Physiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 32 <sup>16.</sup> Department of Microbiology and Immunology, University of Melbourne, Peter Doherty Institute for Infection and  
33 Immunity, Melbourne, VIC, Australia
- 34 <sup>17.</sup> Department of Critical Care, Melbourne Medical School, University of Melbourne, Melbourne, VIC, Australia
- 35 <sup>18.</sup> Division of Infectious Diseases, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
- 36 <sup>19.</sup> Department of Surgery, University of British Columbia, Vancouver, BC, Canada
- 37 <sup>20.</sup> BC Children's Hospital Research Institute, Vancouver, BC, Canada
- 38 <sup>21.</sup> School of Biomedical Engineering, University of British Columbia, Vancouver, BC, Canada
- 39 <sup>22.</sup> Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 40 <sup>23.</sup> Department of Pathology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 41 <sup>24.</sup> Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
- 42 <sup>25.</sup> Division Endocrinology and Diabetes, The Children's Hospital of Philadelphia, Philadelphia, PA, USA
- 43 <sup>26.</sup> Penn Neurodegeneration Genomics Center, Perelman School of Medicine, University of Pennsylvania,  
44 Philadelphia, PA, USA

45  
46

47 **Corresponding author:** Struan F.A. Grant, email: [grants@chop.edu](mailto:grants@chop.edu)

48  
49  
50  
51  
52  
53  
54  
55  
56

57 Word count: 4733

58 **KEY POINTS**

59

60 **Question:** What are the causal variants and corresponding effector genes conferring pediatric  
61 obesity susceptibility in different cellular contexts?

62

63 **Findings:** Our method of assessing 3D genomic data across a range of cell types revealed  
64 heritability enrichment of childhood obesity variants, particularly within pancreatic alpha cells.  
65 The mapping of putative causal variants to cis-regulatory elements revealed candidate effector  
66 genes for cell types spanning metabolic, neural, and immune systems.

67

68 **Meaning:** We gain a systemic view of childhood obesity genomics by leveraging 3D techniques  
69 that implicate regulatory regions harboring causal variants, providing insights into the disease  
70 pathogenesis across different cellular systems.

## 71 **ABSTRACT**

72           The prevalence of childhood obesity is increasing worldwide, along with the associated  
73 common comorbidities of type 2 diabetes and cardiovascular disease in later life. Motivated by  
74 evidence for a strong genetic component, our prior genome-wide association study (GWAS)  
75 efforts for childhood obesity revealed 19 independent signals for the trait; however, the  
76 mechanism of action of these loci remains to be elucidated. To molecularly characterize these  
77 childhood obesity loci we sought to determine the underlying causal variants and the  
78 corresponding effector genes within diverse cellular contexts. Integrating childhood obesity  
79 GWAS summary statistics with our existing 3D genomic datasets for 57 human cell types,  
80 consisting of high-resolution promoter-focused Capture-C/Hi-C, ATAC-seq, and RNA-seq, we  
81 applied stratified LD score regression and calculated the proportion of genome-wide SNP  
82 heritability attributable to cell type-specific features, revealing pancreatic alpha cell enrichment  
83 as the most statistically significant. Subsequent chromatin contact-based fine-mapping was  
84 carried out for genome-wide significant childhood obesity loci and their linkage disequilibrium  
85 proxies to implicate effector genes, yielded the most abundant number of candidate variants  
86 and target genes at the *BDNF*, *ADCY3*, *TMEM18* and *FTO* loci in skeletal muscle myotubes and  
87 the pancreatic beta-cell line, EndoC-BH1. One novel implicated effector gene, *ALKAL2* – an  
88 inflammation-responsive gene in nerve nociceptors – was observed at the key *TMEM18* locus  
89 across multiple immune cell types. Interestingly, this observation was also supported through  
90 colocalization analysis using expression quantitative trait loci (eQTL) derived from the  
91 Genotype-Tissue Expression (GTEx) dataset, supporting an inflammatory and neurologic  
92 component to the pathogenesis of childhood obesity. Our comprehensive appraisal of 3D  
93 genomic datasets generated in a myriad of different cell types provides genomic insights into  
94 pediatric obesity pathogenesis.

## 95 INTRODUCTION

96 The prevalence of obesity has risen significantly worldwide<sup>1</sup>, especially among children and  
97 adolescents<sup>2</sup>. Obesity is associated with chronic diseases, such as diabetes, cardiovascular  
98 diseases, and certain cancers<sup>3-6</sup>, along with mechanical issues including osteoarthritis and sleep  
99 apnea<sup>7</sup>.

100 Modern lifestyle factors, including physical inactivity, excessive caloric intake, and  
101 socioeconomic inequity, along with disrupted sleep and microbiome, represent environmental  
102 risk factors for obesity pathogenesis. However, genetics also play a significant role, with the  
103 estimated heritability ranging from 40% to 70%<sup>8-10</sup>. Studies show that body weight and obesity  
104 remain stable from infancy to adulthood<sup>11-14</sup>, but variation between individuals does exist<sup>15</sup>.  
105 Genome-wide association studies (GWAS) have improved our understanding of the genetic  
106 contribution to childhood obesity<sup>16-21</sup>. However, the functional consequences and molecular  
107 mechanisms of identified genetic variants in such GWAS efforts are yet to be fully elucidated.  
108 Efforts are now being made to predict target effector genes and explore potential drug targets  
109 using various computational and experimental approaches<sup>22-26</sup>, which subsequently warrant  
110 functional follow-up efforts.

111 With our extensive datasets generated on a range of different cell types, by combining 3D  
112 chromatin maps (Hi-C, Capture-C) with matched transcriptome (RNA-seq) and chromatin  
113 accessibility data (ATAC-seq), we investigated heritability patterns of pediatric obesity-  
114 associated variants and their gene-regulatory functions in a cell type-specific manner. This  
115 approach yielded 94 candidate causal variants mapped to their putative effector gene(s) and  
116 corresponding cell type(s) setting. In addition, using methods comparable to our prior efforts in  
117 other disease contexts<sup>27-34</sup>, we also uncovered new variant-to-gene combinations within specific  
118 novel cellular settings, most notably in immune cell types, which further confirmed the  
119 involvement of the immune system in the pathogenesis of obesity in the early stages of life.

## 120 **METHODS**

121 **Data and resource:** Datasets used in prior studies are listed in **eTable 1**. ATAC-seq, RNA-seq,  
122 Hi-C, and Capture-C *library generation* for each cell type is provided in their original published  
123 study and their *pre-processing pipelines and tools* can be found in **eMethods**.

124  
125 **Definition of cis-Regulatory Elements (cREs):** We intersected ATAC-seq open chromatin  
126 regions (OCRs) of each cell type with chromatin conformation capture data determined by Hi-  
127 C/Capture-C of the same cell type, and with promoters (-1,500/+500bp of TSS, which were  
128 referenced by GENCODE v30.

129  
130 **Childhood obesity GWAS summary statistics:** Data on childhood obesity from the EGG  
131 consortium was downloaded from [www.egg-consortium.org](http://www.egg-consortium.org). We used 8,566,179 European  
132 ancestry variants (consisting of 8,613 cases and 12,696 controls in stage I; of 921 cases and  
133 1,930 controls in stage II), representing ~55% of the total 15,504,218 variants observed across  
134 all ancestries in the original study<sup>35</sup>. The sumstats file was reformatted by *munge\_sumstats.py*  
135 to standardize with the weighted variants from HapMap v3 within the LDSC baseline, which  
136 reduced the variants to 1,217,311 (7.8% of total).

137  
138 **Cell type specific partitioned heritability:** We used LDSC (<http://www.github.com/bulik/ldsc>)  
139 v.1.0.1 with *--h2* flag to estimate SNP-based heritability of childhood obesity within 4 defined  
140 sets of input genomic regions: (1) OCRs, (2) OCRs at gene promoters, (3) cREs, and (4) cREs  
141 with an expanded window of  $\pm 500$  bp. The baseline model LD scores, plink filesets, allele  
142 frequencies and variants weights files for the European 1000 genomes project phase 3 in hg38  
143 were downloaded from the provided link  
144 (<https://alkesgroup.broadinstitute.org/LDSCORE/GRCh38/>). The cREs of each cell type were

145 used to create the annotation, which in turn were used to compute annotation-specific LD  
146 scores for each cell types cREs set.

147

148 **Genetic loci included in variant-to-genes mapping:** 19 sentinel signals that achieved  
149 genome-wide significance in the trans-ancestral meta-analysis study<sup>35</sup> were leveraged for our  
150 analyses. Proxies for each sentinel SNP were queried using TopLD<sup>36</sup> and LDlinkR tool<sup>37</sup> with  
151 the GRCh38 Genome assembly, 1000 Genomes phase 3 v5 variant set, European population,  
152 and LD threshold of  $r^2 > 0.8$ , which resulted in 771 proxies, including the 21 SNPs from the 99%  
153 credible set of the original study (**eTable 2**).

154

155 **GWAS-eQTL colocalization:** The summary statistics for the European ancestry subset from  
156 the EGG consortium GWAS for childhood obesity was used. Common variants ( $MAF \geq 0.01$ )  
157 from the 1000 Genomes Project v3 samples were used as a reference panel. We used non-  
158 overlapped genomic windows of  $\pm 250,000$  bases extended in both directions from the median  
159 genomic position of each of 19 sentinel loci as input. We used ColocQuiaL<sup>38</sup> to test genome-  
160 wide colocalization of all possible variants included in each inputted window against GTEx v.8  
161 eQTLs associations for all 49 tissues available from <https://www.gtexportal.org/home/datasets>.  
162 A conditional posterior probability of colocalization of 0.8 or greater was imposed.

163

164

## 165 RESULTS

### 166 ***Enrichment assessment of childhood obesity variants across cell types***

167 To explore the enrichment of childhood obesity GWAS variants across cell types, we carried  
168 out Partitioned Linkage Disequilibrium Score Regression (LDSR)<sup>39</sup> on all ATAC-seq-defined  
169 OCRs for each cell type. We assessed cell-type specific enrichment of GWAS signals in four  
170 main categories of genomic regions (**Fig. 1A**): (1) Total OCRs: open chromatin regions defined  
171 by ATAC-seq; (2) Promoter OCRs: the subset of OCRs overlapping a gene promoter; (3) cREs:  
172 the subset of OCRs that form chromatin loops (as determined by Hi-C/Promoter Capture-C) with  
173 a gene promoter, and are therefore considered putative enhancers or suppressors regulating  
174 gene expression; (4) cREs  $\pm$  500bases: extended cREs by 500 bases in both directions. The  
175 rationale behind this approach is that different GWAS variants can influence phenotypes by  
176 regulating gene expression in a cell-type specific manner through various regulatory  
177 mechanisms. For example, they may alter enhancer function (cREs category) or affect the  
178 binding of a transcription factor at a gene's promoter (Promoter OCRs category).

179 We observed that 41 of 57 cell types – including 22 metabolic, 21 immune, 7 neural cell  
180 types and 7 independent cell lines (**eTable 1**) – showed at least a degree of directional  
181 enrichment with the total set of OCRs (**Fig. 1B – Total OCRs**). However, only four cell types –  
182 two pancreatic alpha and two pancreatic beta cell-based datasets – had statistically significant  
183 enrichments ( $P < 0.05$ ). These enrichments were less pronounced when focusing on promoter  
184 OCRs only (**Fig. 1B – Promoter OCRs**). To further limit the LD enrichment assessment to just  
185 those OCRs that can putatively regulate gene expression via chromatin contacts with gene  
186 promoters, we used the putative cREs<sup>27,29</sup>. This reduced the number of cell types showing at  
187 least nominal enrichment (31 of 57), enlarged the dispersion of enrichment ranges across  
188 different cell types, increased the 95% confidence intervals (CI) of enrichments, and hence  
189 increased the  $P$ -value of the resulting regression score. cREs from pancreatic alpha cells



190 derived from single-cell ATAC-seq were the only dataset that remained statically significant  
191 (**Fig. 1B – putative cREs**).

192 The original reported LDSR method analyzed enrichment in the 500bp flanking regions of  
193 their regulatory categories<sup>39</sup>. However, when we expanded our analysis to the  $\pm 500$ bp window  
194 for our cREs, albeit incorporating more weighted variants into the enrichment (represented by  
195 larger dots in **Fig. 1B – cREs  $\pm 500$  bases**), this resulted in a decrease in the number of cell  
196 types yielding at least nominal enrichment (26 cell types), the enrichment range across cell  
197 types, the 95% CI, and level of significance. The pancreatic alpha cell observation also dropped  
198 below the bar for significance with this expanded window definition.

199

### 200 ***Consistency and diversity of childhood obesity proxy variants mapped to cREs***

201 Despite the enrichments above only being limited to just a small number of cell types, it is  
202 likely that individual loci have differing levels of contributions in various cellular contexts and  
203 could not be detected at the genome wide assessment scale. As such we elected to further  
204 explore the candidate effector genes that are directly affected by cREs harboring childhood  
205 obesity-associated variants by systematically mapping the genomic positions of the LD proxies  
206 onto each cell type's cREs. Most proxies fall within chromatin contact regions (**blue area in**  
207 **Venn diagram Fig. 2A**) or OCRs (**yellow area**) or open chromatin contact regions (**red area**),  
208 or completely outside (**white area**) any defined region. Only 94 proxies fall within our defined  
209 cREs (**overlapped area with dotted green border in Fig. 2A**), they clustered at 13 original loci  
210 (**eTable 3**). **eFig. 1A** outlines the number of signals at each locus included or excluded based  
211 on the criteria we defined for our regions of interest. The *TMEM18* locus yielded the most  
212 variants through cREs mapping, with 46 proxies for the two lead independent variants,  
213 rs7579427 and rs62104180. The second most abundant locus was *ADCY3*, with 21 proxies for  
214 lead variant rs4077678 (**Fig. 2B**). The higher number of variants at one locus did not correlate

215 with implicating more genes or cell types through mapping. The mapping frequency of various  
216 variants within a specific locus exhibited substantial differences.

217 Inspecting individual variants regardless of their locus, we found that 28 of 94 proxies  
218 appeared in cREs across multiple cell types, with another 66 observed in just one cell type (**Fig.**  
219 **2D**). 45 variants of these 66 just contacted one gene promoter, such as at the *GPR1* and  
220 *TFAP2B* loci (**eFig. 2**).

221 Overall, the number of cell types in which a variant was observed in open chromatin  
222 correlated with the number of genes contacted via chromatin loops (**eFig. 3A**). However, we  
223 also observed that some variants found in cREs in multiple cell types were more selective with  
224 respect to their candidate effector genes (**eFig. 3B-red arrow**), or conversely, more selective  
225 across given cell types but implicated multiple genes (**eFig. 3B-blue arrow**). **eFig. 4** outlines  
226 our observations at the *TMEM18* locus – an example locus involved in both scenarios.

227  
228 ***Implicated genes cluster at loci strongly associated with childhood obesity consistently***  
229 ***across multiple cell types***

230 Mapping the variants across all the cell types resulted in a total of 111 implicated childhood  
231 obesity candidate effector genes (**Table 1**). Among these, 45 genes were specific to just one  
232 cell type (**eFig. 5A**), including 13 in myotubes and 7 in natural killer cells. Conversely and  
233 notably, *BDNF* appeared across 42 different cell types. Across the metabolic, neural, and  
234 immune systems and seven other cell lines, there were 9 genes consistently implicated in all  
235 four categories (top panel **Fig. 3** – red stars, **eFig. 5B**: “all”), while 5 genes were consistently  
236 implicated in metabolic, neural, and immune systems (top panel **Fig. 3** – blue stars, **eFig. 5B**:  
237 “all\_main”). Two genes, *ADCY3* and *BDNF*, had variants both at their promoters and contacted  
238 variants in cREs via chromatin loops (**eFig. 6**).

239 At the *TMEM18* locus on chr 2p25.3, a highly significant human obesity locus that has long  
240 been associated with both adult and childhood obesity, we observed differing degrees of

241 evidence for 16 genes, but noted that rs6548240, rs35796073, and rs35142762 consistently  
242 contacted the *SH3YL1*, *ACP1*, and *ALKAL2* promoters across multiple cell types (**Fig. 2D-third**  
243 **and fourth column**).

244 At the chr 2p23 locus, *ADCY3* yielded the most contacts (i.e. many proxies contacting the  
245 same gene via chromatin loops), suggesting this locus acts as a regulatory hub. However, we  
246 observed a similar composition in cell types for four other genes: *DNAJC27*, *DNAJC27-AS1*  
247 (both previously implicated in obesity and/or diabetes traits<sup>40</sup>), *AC013267.1*, and *SNORD14*  
248 (*RF00016*). *ITSN2*, *NCOA1*, and *EFR3B* were three genes within this locus that were only  
249 implicated in immune cell types. *NCOA1* encodes a prominent meta-inflammation factor<sup>41</sup>  
250 known to reduce adipogenesis and shift the energy balance between white and brown fat, and  
251 its absence known to induce obesity<sup>42</sup>.

252 *CALCR* was the most frequently implicated gene at its locus, supported by 20 cell types  
253 across all systems. While within the *BDNF* locus, *METTL15* and *KIF18A* – two non-cell-type-  
254 specific genes - plus some lncRNA genes, were contacted by childhood obesity-associated  
255 proxies within the same multiple cell types as *BDNF*, again suggesting the presence of a  
256 regulatory hub.

257 At the *FAIM2* locus on chr 12q13.12, we observed known genes associated with obesity,  
258 eating patterns, and diabetes-related traits, including *ASIC1*, *AQP2*, *AQP5*, *AQP6*, *RACGAP1*,  
259 and *AC025154.2* (*AQP5-AS1*) along with *FAIM2* (**Table 1**). These genes were harbored within  
260 cREs of astrocytes, neural progenitors, hypothalamic neurons, and multiple metabolic cell types.  
261 Plasmacytoid and CD1c+ conventional dendritic cells were the only two immune cell types that  
262 harbored such proxies within their cREs, implicating *ASIC1*, *PRPF40B*, *RPL35AP28*, *TMBIM6*,  
263 and *LSM6P2* at the *FAIM2* locus.

264 The independent *ADCY9* and *FTO* loci are both located on chromosome 16. Genes at the  
265 *ADCY9* locus were only implicated in a subset of immune cell types. Interestingly, genes at the  
266 *FTO* locus were only implicated in Hi-C datasets (as opposed to Capture C), including 6

267 metabolic cell types and astrocytes. Most genes at the *FTO* locus were implicated in skeletal  
268 myotubes, differentiated osteoblasts, and astrocytes, namely *FTO* and *IRX3*; while *IRX5*,  
269 *CRNDE*, and *AC106738.1* were also implicated in adipocytes and hepatocytes.

270

### 271 ***The most implicated cell types by two sets of analyses***

272 EndoC-BH1 and myotubes are the two cell types in which we implicated the most effector  
273 genes, with 38 and 42, respectively – **Fig. 3** side panel. This phenomenon is likely proportional  
274 in the case of myotubes, given the large number of cREs identified by overlapped Hi-C contact  
275 data and ATAC-seq open regions (**Fig. 1A**), but not for EndoC-BH1. Albeit harboring an  
276 average number of cREs compared to other cell types, EndoC-BH1 cells were consistently  
277 among the top-ranked heritability estimates for the childhood obesity variants resulting from the  
278 EGG consortium GWAS (**Fig. 1**) and harbored a significant number of implicated genes by the  
279 mapping of proxies. Interestingly, the pancreatic alpha cell type – shown above to be the most  
280 significant for heritability estimate by LDSC – revealed only 6 implicated genes contacted by the  
281 defined proxies, namely *BDNF* and five lncRNA genes.

282

### 283 ***Pathway analysis***

284 Of the 111 implicated genes in total, PubMed query revealed functional studies for 66  
285 genes. The remaining were principally lncRNA and miRNA genes with currently undefined  
286 functions (**Table 1**). To investigate how our implicated genes could confer obesity risk, we  
287 performed several pathway analyses keeping them either separated for each cell type or  
288 pooling into the respective metabolic, neural, or immune system sets. **eFig. 7** shows simple  
289 Gene Ontology (GO) biological process terms enrichment results.

290 Leveraging the availability of our expression data generated via RNA-seq (available for 46 of  
291 57 cell types), we performed pathway analysis. Given that our gene sets from the variant-to-  
292 gene process was stringently mapped, the sparse enrichment from normal direct analyses is not

293 ideal for exploring obesity genetic etiology. Thus, we incorporated two methods from the  
294 *pathfindR* package<sup>43</sup> and our customized *SPIA* (details in **eMethods**). The result of 60 enriched  
295 KEGG terms is shown in **eFig. 8 (eTable 4)**, with 13 genes in 14 cell types for *pathfindR* and 39  
296 enriched KEGG terms shown in **eFig. 9 (eTable 5)**, with 10 genes in 42 cell types for  
297 customized *SPIA*. There were 20 overlapping pathways between the two approaches (yellow  
298 rows in **eTable 4&5**) including many signaling pathways such as the GnRH (hsa04912), cAMP  
299 (hsa04024), HIF-1 (hsa04066), Glucagon (hsa04922), Relaxin (hsa04926), Apelin (hsa04371),  
300 and Phospholipase D (hsa04072) signaling pathways. They were all driven by one or more of  
301 these 5 genes: *ADCY3*, *ADCY9*, *CREBBP*, *MMP2*, and *NCOA1*. Interestingly, we observed the  
302 involvement of natural killer cells in nearly all the enriched KEGG terms from *pathfindR* due to  
303 the high expression of the two adenylyl cyclase encoded genes, *ADCY3* and *ADCY9*, along with  
304 *CREBBP*. The *SPIA* approach disregarded the aquaporin genes (given they appear so  
305 frequently in so many pathways that involve cellular channels) but highlighted the central role of  
306 *BDNF* which single-handedly drove four signaling pathways: the Ras, Neurotrophin, PI3K-Akt,  
307 and MAPK signaling pathways. This also revealed the role of *TRAP1* in neurodegeneration.

308 These two approaches did not discount the role of *FAIM2* and *CALCR*. However, their  
309 absence was mainly due to the content of the current KEGG database. On the other hand,  
310 these approaches accentuated the role of the *MMP2* gene at the *FTO* locus in skeletal  
311 myotubes, given its consistency within the GnRH signaling pathway (**eFig. 10**), which is in line  
312 with previous studies linking its expression with obesity<sup>44-46</sup>.

313

### 314 ***Supportive evidence by colocalization of target effector genes with eQTLs***

315 The GTEx consortium has characterized thousands of eQTLs, albeit in heterogeneous bulk  
316 tissues<sup>47</sup>. To assess how many observed gene-SNP pairs agreed with our physical variant-to-  
317 gene mapping approach in our multiple separate cellular settings, we performed colocalization  
318 analysis using ColocQuiaL<sup>38</sup>.

319 282 genes were reported to be associated with the variants within 13 loci from our variant-  
320 to-genes analysis. We found 114 colocalizations for ten of our loci that had high conditioned  
321 posterior probabilities ( $\text{cond.PP.H4.abf} \geq 0.8$ ), involving 44 genes and 41 tissues among the  
322 eQTLs. We extracted the posterior probabilities for each SNP within each colocalization and  
323 selected the 95% credible set as the likely causal variants (complete list in **eTable 6**). Despite  
324 sensitivity differences and varying cellular settings, when compared with our variant-to-gene  
325 mapping results, colocalization analysis yielded consistent identification for 21 pairs of SNP-  
326 gene interactions when considering the analyses across all our cell types, composed of 20  
327 SNPs and 7 genes. Details of these SNP-gene pairs are shown in **Fig. 4A** and **B**.

328 Of these 20 SNPs, 15 were at the *ADCY3* locus, in LD with sentinel variant rs4077678, and  
329 all implicated *ADCY3* as the effector gene in 29 cell types – 15 metabolic, 6 immune, 4 neural  
330 cell types and 4 independent cell lines (**Fig. 4C**). Indeed, missense mutations have been  
331 previously reported for this gene in the context of obesity<sup>48,49</sup> while another member of this gene  
332 family, *ADCY5*, has also been extensively implicated in metabolic traits<sup>50</sup>.

333  
334 ***Predicting transcription factors (TFs) binding disruption at implicated genes contributing***  
335 ***to obesity risk***

336 TFs regulate gene expression by binding to DNA motifs at enhancers and silencers, where  
337 any disruption by a SNP can potentially cause dysregulation of a target gene. Thus, we used  
338 *motifbreakR* (R package) to predict such possible events at the loci identified by our variant-to-  
339 gene mapping. Each variant was predicted to disrupt the binding of several different TFs, thus  
340 requiring further literature cross-examination to select the most probable effects. For example,  
341 rs7132908 (consistently contacting *FAIM2* in 25 cell types) was predicted to disrupt the binding  
342 of 12 different transcription factors. Among them, SREBF1 (**eFig. 11A**) was the only TF that  
343 concurred with evidence that it regulates *AQP2* and *FAIM2* at the same enhancer<sup>51</sup>. The full  
344 prediction list can be found in **eTable 7**.

345 To narrow down the list of putative TF binding sites at each variant position, we leveraged  
346 the ATAC-seq footprint analysis using the RGT suite<sup>52</sup>. The final set of Motif-Predicted Binding  
347 Sites (MPBS) within each cell type ATAC-seq footprints was used to overlap with the genomic  
348 locations of the OCRs, and then overlapped with our obesity variants, resulting in annotated 29  
349 variants. Mosaic plot in **eFig. 11B** shows the number and proportions of variants predicted by  
350 *motifbreakR* and/or overlapped with MPBS. Insignificant *P*-value from Fisher's exact test  
351 indicated the independence of the two analyses. Only seven variants were found within the  
352 cREs for the same TF motifs predicted to be disrupted by *motifbreakR* (**eFig. 11C**). **eFig. 12**  
353 outlines the seven variants that *motifbreakR* and ATAC-seq footprint analysis agreed on the TF  
354 bindings they might disrupt.

## 355 DISCUSSION

356 Given the challenge of uncovering the underlying molecular mechanisms driving such a  
357 multifactorial disease as obesity, our approach leveraging GWAS summary statistics, RNA-seq,  
358 ATAC-seq, and promoter Capture C / Hi-C offers new insights. This is particularly true as it is  
359 becoming increasingly evident that multiple effector genes can operate in a temporal fashion at  
360 a given locus depending on cell state, including at the *FTO* locus<sup>53</sup>. Our approach offers an  
361 opportunity to implicate relevant cis-regulatory regions across different cell types contributing to  
362 the genetic etiology of the disease. By assigning GWAS signals to candidate causal variants  
363 and corresponding putative effector genes via open chromatin and chromatin contact  
364 information, we enhanced the fine-mapping process with an experimental genomic perspective  
365 to yield new insights into the biological pathways influencing childhood obesity.

366 LD score regression is a valuable method that estimates the relationship between linkage  
367 disequilibrium score and the summary statistics of GWAS SNPs to quantify the separate  
368 contributions of polygenic effects and various confounding factors that produce SNP-based  
369 heritability of disease. The general positive heritability enrichment across our open chromatin  
370 features spanning multiple cell types (**Fig. 1B.a**) reinforces the notion that obesity etiology  
371 involves many systems in our body.

372 While obesity has long been known to be a risk factor for pancreatitis and pancreatic cancer,  
373 the significant enrichment of pancreatic alpha and beta cell related 3D genomic features for  
374 childhood obesity GWAS signals demonstrates the bidirectional relationship between obesity  
375 and the pancreas; indeed, it is well established that insulin has obesogenic properties.  
376 Moreover, the comorbidity of obesity and diabetes (either causal or a result of the overlap  
377 between SNPs associated with these two diseases) is tangible. When focusing on genetic  
378 annotation of the cREs only, the association with obesity became more diverse across cell  
379 types, especially in metabolic cells. Interestingly, the lack of enrichment (only 8 of 57 cell types  
380 yielded no degree of enrichment) of obesity SNPs heritability in open gene promoters (**Fig.**



381 **1B.b)** reveals that cRE regions harboring obesity SNPs are more involved in gene regulation  
382 than disruption, and therefore potentially contributing more weight to the manifestation of the  
383 disease.

384 Of course, we should factor in the effective sample sizes of the GWAS efforts that are wide-  
385 ranging (2,000-24,000 – given that the N for each variant is different within a single dataset,  
386 thus contributing to the weights and *P*-value of each SNP when the algorithm calculates the  
387 genome-wide heritability), which could result in noise and negative enrichment observed in the  
388 analysis – a methodology limitation of partial linkage regression that has been extensively  
389 discussed in the field<sup>54</sup>. Thus, it is crucial to interpret the enrichment (or lack thereof) of disease  
390 variants in a certain cellular setting with an *ad hoc* biological context.

391 From mapping the common proxies of 19 independent sentinel SNPs that were genome-  
392 wide significantly associated with childhood obesity to putative effector genes through chromatin  
393 contacting cREs, one striking finding was the several potential “hubs” of putatively core effector  
394 genes, whose occurrence spread across three human physiological systems. With the data  
395 available from so many cell types, our approach connected new candidate causal variants to  
396 known obesity-related genes and new implications of cell modality for previously known  
397 associations.

398 A potential application of this association could be to fine-tune the effect of a drug toward  
399 controlling appetite. An example of bringing new aspects to the old is for the signal within the  
400 *FTO* locus that contacted *IRX3* and *IRX5*: previous studies have suggested these obesogenic  
401 effects operate in adipocytes<sup>55</sup>, brain<sup>56</sup>, or pancreas<sup>57</sup>; here we confirmed this association in  
402 adipocytes and uncover the presence of distal chromatin contacts in myotubes for the first time.

403 Besides the above-mentioned genes with known associations with obesity, we discovered  
404 newly implicated genes. For example, the *LRR1Q3* gene at the *TNNI3K* locus had its open  
405 promoter contacted by two SNPs, rs1040070 and rs10493544, in NTERA2 cells only. The  
406 published studies<sup>58,59</sup> that associated *LRR1Q3* with major depressive disorder and opioid usage

407 acknowledged the overlapping promoter of this gene, albeit in the opposite direction, with a run-  
408 through transcript of *FPGT-TNNI3K* – previously shown to be associated with BMI in  
409 European<sup>60</sup> and Korean populations<sup>61</sup>.

410 It is apparent that not all the implicated genes we report would contribute equally to the  
411 susceptibility of obesity pathogenesis. Each locus comprises genes whose functions are  
412 obviously related to obesity or similar traits like BMI, fat weight, etc., while other genes are not  
413 so directly obvious in their relation to these traits.

414 It is encouraging that for implicated genes within these multi-cell-type loci across different  
415 physiological systems we could find previous associations to the corresponding cell types or  
416 systems. Examples are the two aforementioned genes at the *TMEM18* locus (*SH3YL1* and  
417 *ACP1*)<sup>62-66</sup> with the broad spectrum of their functions, *HEPACAM2* implicated in the NCIH716  
418 cell line at the *CALCR* locus<sup>67,68</sup>, and *LRR1Q3* in the NTERA2 cell line at the *TNNI3K* locus<sup>69</sup>.

419 Chronic inflammation is an essential characteristic of obesity pathogenesis. Adipose tissue-  
420 resident immune cells have been observed, leading to an increased focus in recent years on  
421 their potential contribution to metabolic dysfunction. On the other hand, neurological or  
422 psychological conditions, such as stress, induce the secretion of both glucocorticoids (increase  
423 motivation for food) and insulin (promotes food intake and obesity). Pleasure feeding then  
424 reduces activity in the stress-response network, reinforcing the feeding habit. It has been shown  
425 that voluntary behaviors, stimulated by external or internal stressors or pleasurable feelings,  
426 memories, and habits, can override the basic homeostatic controls of energy balance<sup>70</sup>. The  
427 potential link between the immune system and metabolic disease, and moreover, through the  
428 neural system, was tangible in our findings.

429 Two of the three SNPs which ranked the third most consistent in our variant-to-gene  
430 mapping (**Fig. 2C**) – rs35796073 and rs35142762 – contacted the *ALKAL2* promoter (supported  
431 by GTEx evidence to colocalize with *ALKAL2* expression). The anaplastic lymphoma kinase  
432 (encoded by *ALK* gene) is a receptor tyrosine kinase, belongs to the insulin receptor family, and

433 has been reported to promote nerve cell growth and differentiation<sup>71,72</sup>. Despite *ALKAL2* (ALK  
434 and LTK ligand 2) being studied principally in the context of immunity, a recent study using the  
435 EGCUT biobank GWAS identified *ALK* as a candidate thinness gene and genetic deletion  
436 showed that its expression in hypothalamic neurons acts as a negative regulator in controlling  
437 energy expenditure via sympathetic control of adipose tissue lipolysis<sup>73</sup>. *ALKAL2* – encoding a  
438 high-affinity agonist of *ALK/LTK* receptors – which has been reported to enhance expression in  
439 response to inflammatory pain in nociceptors<sup>74,75</sup> - has been recently implicated as a novel  
440 candidate gene for childhood BMI by transcriptome-wide association study<sup>76</sup>, and achieved  
441 genome-wide significance in a GWAS study contrasting persistent healthy thinness with severe  
442 early-onset obesity using the STILTS and SCOOP cohorts<sup>77</sup>. The finding that overexpression of  
443 *ALKAL2* could potentiate neuroblastoma progression in the absence of ALK mutation<sup>78</sup> echoes  
444 the relationship between *ADCY3* and *MC4R*<sup>79</sup>, where a peripheral gene, *ADCY3*, can  
445 regulate/impair the function of a core gene, i.e. *MC4R*, within the energy-regulating  
446 melanocortin signaling pathway<sup>80</sup>.

447 Our approach implicates putative target genes based on a mechanism of regulation for  
448 these variants to alter gene expression – through regulator TF(s) that bind to these contact  
449 sites. A potential limitation of the predictions from *motifbreakR* and matching TF motifs to ATAC-  
450 seq footprint by the RGT toolkit is that they were both based on the position probability matrixes  
451 of Jaspar and Hocomoco, which come from public motif databases. The ATAC-seq footprint  
452 analysis also carries sequence bias that can lead to false positive discovery. Thus, our attempt  
453 to call such regulators by predicting TF binding disruption can only serve as nominations – but  
454 warrant further functional follow up.

455 Another limitation of this work is the diversity in data quality among different samples, since  
456 different datasets were sampled and collected at different time points, from different patients,  
457 using different protocols, with libraries sequenced at different depths and qualities, and initially  
458 preprocessed with different pipelines and parameters. Thus, it is crucial to keep in mind that the

459 discrepancy in data points might have resulted from variations in data quality. Importantly, any  
460 association discovered must be validated functionally before effector genes of the genetic  
461 variants can be leveraged to develop new therapies. Their putative function(s) must be  
462 characterized, together with the mechanism whereby the given variant's alleles differentially  
463 affect the expression of the targeted genes. The next step is to explore how the target genes  
464 affect the trait of interest more directly.

465 Our results have provided a set of leads for future exploratory experiments in specific  
466 cellular settings in order to further expand our knowledge of childhood obesity genomics and  
467 hence equip us with more effective means to overcome the burden of this systematic disease.

468

## 469 **CONCLUSION**

470 Our approach of combining RNA-seq, ATAC-seq, and promoter Capture C/Hi-C datasets  
471 with GWAS summary statistics offers a systemic view of the multi-cellular nature of childhood  
472 obesity, shedding light on potential regulatory regions and effector genes. By leveraging  
473 physical properties, such as open chromatin status and chromatin contacts, we enhanced the  
474 fine-mapping process and gained new insights into the biological pathways influencing the  
475 disease. Although further functional validation is required, our findings provide valuable leads  
476 together with their cellular contexts for future research and the development of more effective  
477 strategies to address the burden of childhood obesity.

478 **ACKNOWLEDGMENTS**

479 This work was supported by National Institutes of Health awards R01 HD056465, R01

480 DK122586 and UM1 DK126194, and the Daniel B. Burke Endowed Chair for Diabetes

481 Research.

482 Given the use of de-identified datasets and biospecimens was not considered human subjects

483 research, ethical oversight was waived by the Institutional Review Board of the Children's

484 Hospital of Philadelphia

485 **TABLE AND FIGURE LEGENDS**

486 **Figure 1: Partitioned Linkage Disequilibrium Score Regression analysis for open**  
487 **chromatin regions of all cell types.**

488 A. The schematic shows the different types of regions defined in our study and 3 different  
489 ways overlapping chromatin contact regions – OCRs – gene promoters define cREs.

490 B. Heritability enrichment by LDSC analysis for each cell type

491 a) Bar-plot shows the total number of OCRs identified by ATAC-seq for each cell type  
492 on bulk cells - blue, or on single cells – red; and the portion of OCRs that fall within  
493 cREs identified by incorporating Hi-C (green) or by Capture-C (orange).

494 b-e) 4 panels of dot-plots show heritability enrichment by LDSC analysis for each cell  
495 type, with standard error whiskers. Dots' colors correspond to  $-\log_{10}(p\text{-values})$ ,  
496 dots with white asterisks are significant  $p\text{-values} < 0.05$ , and dots' sizes  
497 corresponding to the proportion of SNP contribute to heritability. Dash line at 1, i.e.,  
498 no enrichment.

499 b) Analysis done on whole OCRs set of each cell type (whiskers colors match  
500 with bulk/single cell from bar-plot a);

501 c) On only OCRs that overlapped with promoters (whiskers' colors match with  
502 bulk/single cell from bar-plot a);

503 d) On the putative cREs of each cell type (whiskers' colors match with Hi-  
504 C/Capture-C from bar-plot a);

505 e) On the same cREs as (c) panel with their genomic positions expanded  $\pm 500$   
506 bases on both sides (whiskers' colors match with Hi-C/Capture-C from bar-  
507 plot a)

508

509 **Figure 2: Mapping 771 proxies to the open chromatin regions of each cell type**

510 A. Venn diagram shows how 771 proxies mapped to the OCRs:

511 • Blue area: 758 proxies were located within contact regions of at least one cell  
512 type regardless of chromatin state;

513 • Red area: 417 proxies were located within contact regions marked as open by  
514 overlapping with OCR;

515 • Yellow area: If we only considered open chromatin regions, 178 proxies were  
516 included;

- 517           • Dotted green bordered area: To focus on just those variants residing within open  
518 chromatin and contacting promoter regions in any cell type, we overlapped the  
519 genomic positions of these proxies with each cell type's cRE set, yielding 90  
520 variants (3 from the 99% credible set) directly contacting open gene promoters  
521 **(eTable 3)**, with 10 of which located within a promoter of one gene but contacting  
522 another different gene promoter. There were an additional 4 variants located  
523 within gene promoters but in chromatin contact with promoter(s) of nearby  
524 transcript(s) of the same gene (correspond to 3 cREs illustrations in Figure 1A).
- 525           • White area: proxies that fall into neither defined region of interest.
- 526 B. Bar-plot shows number of proxies, cell types and target genes mapped at each locus.
- 527 C. The upSet plot shows the degree of overlap across cell types of the variants; ranked  
528 from the most common variant (red) – rs61888800 from *BDNF* locus, a well-known 5'  
529 untranslated region variant of this gene that is associated with anti-depression and  
530 therapeutic response<sup>81,82</sup> – appeared in 39 cell types, to the group of variants (grey)  
531 which appeared in only one cell type.

532

533 **Figure 3: Profiles of 111 implicated genes by 94 proxies through cREs of each cell**  
534 **type**

535 Main panel: Bubble plot show corresponding expression level (size) and number of variants  
536 (color) target each implicated gene of each cell type. Squares represent genes with  
537 variants at their promoters. Circles represent genes with variants contacted through  
538 chromatin loops. Some genes were implicated by both types, these “double implications”  
539 are represented as diamond shapes, and were identified across several cell types: two  
540 cell types (plasmacytoid dendritic cells and pre-differentiated adipocytes) for *ADCY3*  
541 gene, and five for *BDNF* (human embryonic stem cells - hESC, differentiated human  
542 fetal osteoblast cells - hFOB\_Diff, neural progenitor cells derived from induced  
543 pluripotent stem cells - NPC\_iPSC, PANC-1, and NCIH716 cell lines)

544 Genes with expression undetected in our arrays are shown as triangles.

545 Top panel: bar-plot shows numbers of cell types each gene was implicated within, color-  
546 coded by which systems the cell types belong to.

547 Right panel: bar-plot shows numbers of genes implicated by the variants with each cell type.

548

549 **Figure 4: Colocalization of target effector genes with eQTLs**

550 A. Venn diagram shows the overlaps between sets of genes yielded by ColocQuiaL and the  
551 variant-to-gene mapping process.

552 B. Circos plot of the 10 loci demonstrates the differences in the ranges of associations between  
553 the two approaches, with long-ranged chromatin contacts between obesity variants and  
554 target genes displayed as orange links and short-range eQTLs colocalizations as green  
555 links.

556 Two SNPs – rs35796073, and rs35142762 within the *TMEM18* locus, in linkage  
557 disequilibrium with rs7579427 – were estimated with high probability (cond.PP.H4=0.78) of  
558 colocalizing with the expression of *ALKAL2* gene in subcutaneous adipose tissue. These  
559 pairs of SNP-gene were also identified by our variant-to-gene mapping approach in natural  
560 killer cells, plasmacytoid dendritic cells, unstimulated PBMC naïve CD4 T cells and  
561 astrocytes.

562 The rs7132908 variant at the *FAIM2* locus colocalized with the expression of *AQP6* in  
563 thyroid tissue and with *ASIC1* in prostate tissue, not only with high cond.PP.H4 but also with  
564 high individual SNP causal probability (SNP.PP.H4 > 0.95). rs7132908 was the second most  
565 consistent observation in our variant-to-gene mapping, namely across 25 different cell types  
566 (**Figure 2B**) and all three systems plus the other independent cell lines. The pair of  
567 rs7132908-contacting-*AQP6* was observed in 15 different cell types - 8 metabolic and 4  
568 neural cell types, and 3 independent cell lines. The pair of rs7132908-contacting-*ASIC1* was  
569 observed in 11 different cell types - 8 metabolic and 2 neural cell types, and plasmacytoid  
570 dendritic cells.

571 The other eQTL signals that overlapped with our variant-to-gene mapping results were:  
572 *BDNF* at the *METTL15* locus with its promoter physically contacted by rs11030197 in 4 cell  
573 types and its expression significantly colocalized (cond.PP.H4=0.82) in tibial artery; *ADCY9*  
574 at its locus with its promoter physically contacted by rs2531995 in natural killer cells and its  
575 expression significantly colocalized in skin tissue (“Skin\_Not\_Sun\_Exposed\_Suprapubic”,  
576 cond.PP.H4=0.97). And *ADCY3* in the C panel.

577 C. ColocQuiaL estimated that these SNPs highly colocalize with the expression of *ADCY3* in  
578 11 different tissues, where the overlapping with the 16 cell types is represented, color-coded  
579 by the proxies rs numbers.

580

581 **Table 1 – PubMed-query known functions for 111 genes implicated by obesity variants.**

Locus	Implicated genes	Obesity or related traits	Different traits
-------	------------------	---------------------------	------------------



<b>TNNI3K</b>	LRRIQ3	(NA)	Associated with opioid usage <a href="#">[PMID:34728798]</a> and MDD <a href="#">[PMID: 31748543]</a>
	FPGT	Predict BMI in Korean pop. <a href="#">[PMID: 28674662]</a>	(NA)
	FPGT-TNNI3K		Associated with MDD <a href="#">[PMID: 31748543]</a>
	LRRC53	Associated with high BMI increased risk heart attack <a href="#">[PMID: 32471361]</a>	(NA)
	ASTN1	Identified as obesity QTL in rat <a href="#">[PMID: 35729251]</a>	Associated with neurodevelopmental traits <a href="#">[PMID: 24381304]</a> and variety of cancers <a href="#">[PMID: 32945491]</a>
	BRINP2	(NA)	Associated with neurodevelopmental traits <a href="#">[PMID: 34267256]</a>
<b>SEC16B</b>	AL122019.1	(NA)	
	AL162431.1		
<b>TMEM18</b>	FAM110C	(NA)	Overexpression induces microtubule aberrancies <a href="#">[PMID: 17499476]</a> , involved in cell spreading and migration <a href="#">[PMID: 19698782]</a>
	SH3YL1	Associated with BMI in type 2 diabetes nephropathy <a href="#">[PMID: 33223406]</a>	Influence on T cell activation <a href="#">[PMID: 31427643]</a> , involved in different cancer types <a href="#">[PMID: 26305679,24508479]</a>
	ACP1	Associated with early-onset obesity <a href="#">[PMID: 24129437]</a> , correlated with cardiovascular risks <a href="#">[PMID: 19570551]</a> , drive adipocyte differentiation via control of pdgfra signaling <a href="#">[PMID: 33615467]</a>	Associated with bipolar disorder <a href="#">[PMID: 31830721]</a>
	ALKAL2	Associated with childhood BMI <a href="#">[PMID: 33627773]</a>	Enhance expression in response to inflammatory pain in nociceptors <a href="#">[PMID: 35608912, 35610945]</a>
	MYT1L	Associated with early-onset obesity <a href="#">[PMID: 24129437]</a>	(NA)
	AC079779.1	(NA)	
	AC079779.2		
	AC079779.3		
AC079779.4			
LINC01865			
AC105393.2			
AC105393.1			

	LINC01874			
	LINC01875			
	AC093326.1			
	AC141930.2			
	ITSN2	(NA)	Regulate T-cells function <a href="#">[PMID: 32618424]</a> and help the interaction with B-cells <a href="#">[PMID: 29337666]</a>	
<b>ADCY3</b>	NCOA1	Meta-inflammation gene <a href="#">[PMID: 25647480]</a> , reduce adipogenesis, shift the energy balance between white and brown fat <a href="#">[PMID: 31133421]</a>	(NA)	
	ADCY3	Regulate/impair MC4R within energy-regulating melanocortin signaling pathway <a href="#">[PMID: 29311635,32955435]</a>		
	DNAJC27-AS1	Linked to obesity, diabetes traits <a href="#">[PMID: 30131766]</a>		
	DNAJC27	Linked to obesity, diabetes traits <a href="#">[PMID: 30131766]</a>		
	EFR3B	Associated with T1D <a href="#">[PMID: 21980299]</a> , down-regulated in rare obesity-related disorder <a href="#">[PMID: 25705109]</a>		
	WDR43	(NA)		Associated with breast cancer <a href="#">[PMID: 27117709]</a>
	AC013267.1	(NA)		
RF00016	(NA)			
<b>GPR1</b>	GPR1	Increase expression in obese phenotype <a href="#">[PMID: 34174278]</a>	(NA)	
<b>TFAP2B</b>	TFAP2D		Involve in embryogenesis <a href="#">[PMID: 12711551]</a>	
<b>CALCR</b>	HEPACAM2	(NA)	Associated with colorectal cancer <a href="#">[PMID: 29659199, 29973580]</a>	
	VPS50		Involve in neurodevelopmental disorders and defects <a href="#">[PMID: 30828385, 34037727]</a>	
	MIR653		Involve in different types of cancer <a href="#">[PMID: 35777307]</a>	
	MIR489		Promote adipogenesis in mice <a href="#">[PMID: 34004251]</a>	(NA)
	CALCR		Associated with BMI and control of food-intake <a href="#">[PMID: 34462445, 34210852, 31955990, 29522093]</a>	
	TFPI2		(NA)	

	BET1	Involved in triacylglycerol metabolism [ <a href="#">PMID: 24423365</a> ]	Associated with muscular dystrophy [ <a href="#">PMID: 34310943, 34779586</a> ]
	AC003092.1	(NA)	Association with glioblastoma [ <a href="#">PMID: 33815468, 30442884</a> ]
	AC002076.1	(NA)	
<b>BDNF</b>	LIN7C	Associated in T2D [ <a href="#">PMID: 20215397</a> ], obesity [ <a href="#">PMID: 23044507</a> ]	Associated with psychopathology [ <a href="#">PMID: 23044507</a> ]
	BDNF-AS	Regulate <i>BDNF</i> and <i>LIN7C</i> expression [ <a href="#">PMID: 22960213, 22446693</a> ]	(NA)
	BDNF	Regulate eating behavior and energy balance [ <a href="#">PMID: 34556834</a> ]	(NA)
	MIR610	(NA)	Involve in different types of cancer [ <a href="#">PMID: 34408418, 29228616, 26885452</a> ]
	KIF18A		Involve in different types of cancer [ <a href="#">PMID: 35591854, 35286090</a> ]
	METTL15	Associated with childhood obesity [ <a href="#">PMID: 31504550</a> ]	(NA)
	AC090124.1	(NA)	Reported to differentially prognostic of pancreatic cancer [ <a href="#">PMID: 34307375</a> ]
	ARL14EP		Involve in WAGR syndrome [ <a href="#">PMID: 36011342, 31511512</a> ]
	DCDC1		Involvement with eyes anomalies [ <a href="#">PMID: 34773354, 34703991</a> ]
	THEM7P		Associated with mechanisms underlying inguinal hernia [ <a href="#">PMID: 34392144</a> ]
	AL035078.2		(NA)
	ELP4		
	LINC00678		
	AC023206.1		
	RN7SKP158		
	AC104978.1		
MIR8068			
AC013714.1			
AC100773.1			
AC090833.1			
AC090791.1			
AC110056.1			
AL035078.2			

<b>FAIM2</b>	PRPF40B	(NA)	Splicing regulator involved in T-cell development [ <a href="#">PMID: 31088860, 34323272</a> ]	
	TMBIM6	Deficiency leads to obesity by increasing Ca <sup>2+</sup> -dependent insulin secretion [ <a href="#">PMID: 32394396</a> ]	Immune cell function and survival [ <a href="#">PMID: 26470731</a> ]	
	BCDIN3D	Associated with obesity, T2D [ <a href="#">PMID: 20215397</a> ]	(NA)	
	FAIM2	Associated with childhood obesity [ <a href="#">PMID: 31504550</a> ]		
	AQP2	Associated with obesity, diabetes [ <a href="#">PMID: 33367818</a> ]		
	AQP5	Associated with non-obese diabetes [ <a href="#">PMID: 25635992, 22320885</a> ]		Responsible for transporting water, involve in Sjogren's syndrome [ <a href="#">PMID: 25635992, 31557796</a> ]
	AQP6	Down-regulated in retina in diabetes [ <a href="#">PMID: 21851171</a> ]	Associated with renal diseases [ <a href="#">PMID: 30654539</a> ]	(NA)
	RACGAP1	Involve in diabetes nephropathy [ <a href="#">PMID: 35222021</a> ]		
	ASIC1	Inhibition increase food intake and decrease energy expenditure [ <a href="#">PMID: 35894166</a> ]		
	LSM6P2	(NA)	(NA)	
	RPL35AP28			
	LINC02396			
	LINC02395			
	AC025154.1			
AC025154.2				
<b>ADCY9</b>	SLX4	(NA)	Associated with blood pressure [ <a href="#">PMID: 30671673</a> ]	
	DNASE1	Associated with obesity hypertension [ <a href="#">PMID: 33351325</a> ]	(NA)	
	TRAP1	Involve in global metabolic network, deletion reduce obesity incidence [ <a href="#">PMID: 25088416</a> ]		
	CREBBP	Associated with high adiposity and low cardiometabolic risk [ <a href="#">PMID: 33619380</a> ]		
	ADCY9	Associated with BMI, obesity [ <a href="#">PMID: 33619380, 23563607</a> ]		
	SRL	(NA)	Involve in cardiac dysfunction [ <a href="#">PMID: 22119571</a> ]	
	LINC01569		Associated with cancer and endometriosis [ <a href="#">PMID: 35341703, 34422671</a> ]	
	TFAP4		Associated with BMI, birth weight,	(NA)

		maternal glycemic <a href="#">[PMID: 35708509]</a>		
	AC012676.1	(NA)	Involve in hepatocellular carcinoma <a href="#">[PMID: 35210216]</a>	
	AC009171.2	(NA)		
<b>FTO</b>	FTO	Most extensively studied obesity locus <a href="#">[PMID: 34556834]</a>	(NA)	
	IRX3	Obesogenic effects in adipocytes <a href="#">[PMID: 26760096]</a> , brain <a href="#">[PMID: 24646999]</a> , pancreas[93]		
	IRX5			
		AC018553.1	(NA)	Associated with melanoma <a href="#">[PMID: 35611195]</a>
		CRNDE	Regulator of angiogenesis in obesity-induced diabetes <a href="#">[PMID: 31863035]</a>	(NA)
		MMP2	Involve in obesity-relate angiogenesis <a href="#">[PMID: 35919566]</a>	
		CAPNS2	(NA)	Associated with thyroid-related traits <a href="#">[PMID: 23408906]</a>
		AMFR	Involve in hepatic lipid metabolism <a href="#">[PMID: 33591966]</a>	(NA)
		CETP	Involve in monogenic hyperalphalipoproteinemia <a href="#">[PMID: 34878751]</a>	
		RPGRIP1L	Hypomorphism of this ciliary gene linked to morbid obesity <a href="#">[PMID: 27064284, 30597647, 29657248]</a>	Required for hypothalamic arcuate neuron development <a href="#">[PMID: 30728336]</a>
		LINC02169	(NA)	Associated with occupational exposure to gases/fumes and mineral dust <a href="#">[PMID: 31152171]</a>
		AC007491.1	(NA)	
		AC018553.2		
	LINC02140			
	AC106738.1			
	AC106738.2			
	MTND5P34			
	AC007336.1			
<b>MC4R</b>	AC090771.1	(NA)		

582

583

584 **REFERENCES**

585

- 586 1. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975  
587 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million  
588 children, adolescents, and adults. *Lancet*. Dec 16 2017;390(10113):2627-2642.  
589 doi:10.1016/s0140-6736(17)32129-3
- 590 2. National Health and Nutrition Examination Survey 2017–March 2020 Prepandemic Data  
591 Files Development of Files and Prevalence Estimates for Selected Health Outcomes,  
592 <http://dx.doi.org/10.15620/cdc.106273> (2021). <https://stacks.cdc.gov/view/cdc/106273>
- 593 3. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *New England*  
594 *Journal of Medicine*. 2017;377(1):13-27. doi:10.1056/NEJMoa1614362
- 595 4. Singh GM, Danaei G, Farzadfar F, et al. The Age-Specific Quantitative Effects of  
596 Metabolic Risk Factors on Cardiovascular Diseases and Diabetes: A Pooled Analysis. *PLOS ONE*.  
597 2013;8(7):e65174. doi:10.1371/journal.pone.0065174
- 598 5. The Emerging Risk Factors C. Separate and combined associations of body-mass index  
599 and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective  
600 studies. *The Lancet*. 2011/03/26/ 2011;377(9771):1085-1095.  
601 doi:[https://doi.org/10.1016/S0140-6736\(11\)60105-0](https://doi.org/10.1016/S0140-6736(11)60105-0)
- 602 6. Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K. Body Fatness  
603 and Cancer — Viewpoint of the IARC Working Group. *New England Journal of Medicine*.  
604 2016;375(8):794-798. doi:10.1056/NEJMs1606602
- 605 7. Fontaine KR, Barofsky I. Obesity and health-related quality of life. *Obes Rev*. Aug  
606 2001;2(3):173-82. doi:10.1046/j.1467-789x.2001.00032.x
- 607 8. Loos RJF, Yeo GSH. The genetics of obesity: from discovery to biology. *Nature Reviews*  
608 *Genetics*. 2022/02/01 2022;23(2):120-133. doi:10.1038/s41576-021-00414-z
- 609 9. Maes HH, Neale MC, Eaves LJ. Genetic and environmental factors in relative body  
610 weight and human adiposity. *Behavior genetics*. Jul 1997;27(4):325-51.
- 611 10. Elks CE, Den Hoed M, Zhao JH, et al. Variability in the heritability of body mass index: a  
612 systematic review and meta-regression. *Frontiers in endocrinology*. 2012;3:29.
- 613 11. Demerath EW, Choh AC, Czerwinski SA, et al. Genetic and environmental influences on  
614 infant weight and weight change: the Fels Longitudinal Study. *Am J Hum Biol*. Sep-Oct  
615 2007;19(5):692-702. doi:10.1002/ajhb.20660
- 616 12. Dubois L, Girard M, Girard A, Tremblay R, Boivin M, Pérusse D. Genetic and  
617 environmental influences on body size in early childhood: a twin birth-cohort study. *Twin Res*  
618 *Hum Genet*. Jun 2007;10(3):479-85. doi:10.1375/twin.10.3.479
- 619 13. Wardle J, Carnell S, Haworth CM, Plomin R. Evidence for a strong genetic influence on  
620 childhood adiposity despite the force of the obesogenic environment. *Am J Clin Nutr*. Feb  
621 2008;87(2):398-404. doi:10.1093/ajcn/87.2.398
- 622 14. Bouchard C. Childhood obesity: are genetic differences involved? *The American Journal*  
623 *of Clinical Nutrition*. 2009;89(5):1494S-1501S. doi:10.3945/ajcn.2009.27113C
- 624 15. Littleton SH, Berkowitz RI, Grant SFA. Genetic Determinants of Childhood Obesity. *Mol*  
625 *Diagn Ther*. Dec 2020;24(6):653-663. doi:10.1007/s40291-020-00496-1

- 626 16. Vogelesang S, Bradfield JP, Ahluwalia TS, et al. Novel loci for childhood body mass index  
627 and shared heritability with adult cardiometabolic traits. *PLoS Genetics*. 2020;16(10):e1008718.  
628 doi:10.1371/journal.pgen.1008718
- 629 17. Yaghootkar H, Zhang Y, Spracklen CN, et al. Genetic Studies of Leptin Concentrations  
630 Implicate Leptin in the Regulation of Early Adiposity. *Diabetes*. Dec 2020;69(12):2806-2818.  
631 doi:10.2337/db20-0070
- 632 18. Couto Alves A, De Silva NMG, Karhunen V, et al. GWAS on longitudinal growth traits  
633 reveals different genetic factors influencing infant, child, and adult BMI. *Sci Adv*. Sep  
634 2019;5(9):eaaw3095. doi:10.1126/sciadv.aaw3095
- 635 19. Fu J, Wang Y, Li G, et al. Childhood sleep duration modifies the polygenic risk for obesity  
636 in youth through leptin pathway: the Beijing Child and Adolescent Metabolic Syndrome cohort  
637 study. *Int J Obes (Lond)*. Aug 2019;43(8):1556-1567. doi:10.1038/s41366-019-0405-1
- 638 20. Turcot V, Lu Y, Highland HM, et al. Protein-altering variants associated with body mass  
639 index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet*. Jan  
640 2018;50(1):26-41. doi:10.1038/s41588-017-0011-x
- 641 21. Littleton SH, Grant SFA. Strategies to identify causal common genetic variants and  
642 corresponding effector genes for paediatric obesity. *Pediatr Obes*. Dec 2022;17(12):e12968.  
643 doi:10.1111/ijpo.12968
- 644 22. Yu F, Cato LD, Weng C, et al. Variant to function mapping at single-cell resolution  
645 through network propagation. *Nature Biotechnology*. 2022/06/06 2022;doi:10.1038/s41587-  
646 022-01341-y
- 647 23. Avsec Ž, Agarwal V, Visentin D, et al. Effective gene expression prediction from sequence  
648 by integrating long-range interactions. *Nature Methods*. 2021/10/01 2021;18(10):1196-1203.  
649 doi:10.1038/s41592-021-01252-x
- 650 24. Gazal S, Weissbrod O, Hormozdiari F, et al. Combining SNP-to-gene linking strategies to  
651 identify disease genes and assess disease omnigenicity. *Nature Genetics*. 2022/06/01  
652 2022;54(6):827-836. doi:10.1038/s41588-022-01087-y
- 653 25. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-  
654 based sequence model. *Nature Methods*. 2015/10/01 2015;12(10):931-934.  
655 doi:10.1038/nmeth.3547
- 656 26. Nasser J, Bergman DT, Fulco CP, et al. Genome-wide enhancer maps link risk variants to  
657 disease genes. *Nature*. 2021/05/01 2021;593(7858):238-243. doi:10.1038/s41586-021-03446-x
- 658 27. Chesi A, Wagley Y, Johnson ME, et al. Genome-scale Capture C promoter interactions  
659 implicate effector genes at GWAS loci for bone mineral density. OriginalPaper. *Nature*  
660 *Communications*. 2019-03-19 2019;10(1):1-11. doi:doi:10.1038/s41467-019-09302-x
- 661 28. Su C, Johnson ME, Torres A, et al. Mapping effector genes at lupus GWAS loci using  
662 promoter Capture-C in follicular helper T cells. *Nature Communications*. 2020/07/03  
663 2020;11(1):3294. doi:10.1038/s41467-020-17089-5
- 664 29. Pahl MC, Doege CA, Hodge KM, et al. Cis-regulatory architecture of human ESC-derived  
665 hypothalamic neuron differentiation aids in variant-to-gene mapping of relevant complex traits.  
666 OriginalPaper. *Nature Communications*. 2021-11-19 2021;12(1):1-12. doi:doi:10.1038/s41467-  
667 021-27001-4

- 668 30. Cousminer DL, Wagley Y, Pippin JA, et al. Genome-wide association study implicates  
669 novel loci and reveals candidate effector genes for longitudinal pediatric bone accrual. *Genome*  
670 *Biol.* Jan 4 2021;22(1):1. doi:10.1186/s13059-020-02207-9
- 671 31. Vujkovic M, Ramdas S, Lorenz KM, et al. A multiancestry genome-wide association study  
672 of unexplained chronic ALT elevation as a proxy for nonalcoholic fatty liver disease with  
673 histological and radiological validation. *Nat Genet.* Jun 2022;54(6):761-771.  
674 doi:10.1038/s41588-022-01078-z
- 675 32. Pahl MC, Le Coz C, Su C, et al. Implicating effector genes at COVID-19 GWAS loci using  
676 promoter-focused Capture-C in disease-relevant immune cell types. *Genome Biology.*  
677 2022/06/03 2022;23(1):125. doi:10.1186/s13059-022-02691-1
- 678 33. Su C, Gao L, May CL, et al. 3D chromatin maps of the human pancreas reveal lineage-  
679 specific regulatory architecture of T2D risk. *Cell Metabolism.* 6 September 2022  
680 2022;34(9)1409. doi:<https://doi.org/10.1016/j.cmet.2022.08.014>
- 681 34. Palermo J, Chesi A, Zimmerman A, et al. Variant-to-gene mapping followed by cross-  
682 species genetic screening identifies GPI-anchor biosynthesis as a regulator of sleep. *Sci Adv.* Jan  
683 6 2023;9(1):eabq0844. doi:10.1126/sciadv.abq0844
- 684 35. Bradfield JP, Vogelesang S, Felix JF, et al. A trans-ancestral meta-analysis of genome-  
685 wide association studies reveals loci associated with childhood obesity. *Human Molecular*  
686 *Genetics.* 2019;28(19):3327-3338. doi:10.1093/hmg/ddz161
- 687 36. Huang L, Rosen JD, Sun Q, et al. TOP-LD: A tool to explore linkage disequilibrium with  
688 TOPMed whole-genome sequence data. *The American Journal of Human Genetics.* 2022/06/02/  
689 2022;109(6):1175-1181. doi:<https://doi.org/10.1016/j.ajhg.2022.04.006>
- 690 37. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R Package for Rapidly Calculating  
691 Linkage Disequilibrium Statistics in Diverse Populations. *Front Genet.* 2020;11:157.  
692 doi:10.3389/fgene.2020.00157
- 693 38. Chen BY, Bone WP, Lorenz K, Levin M, Ritchie MD, Voight BF. ColocQuiaL: a QTL-GWAS  
694 colocalization pipeline. *Bioinformatics.* 2022;38(18):4409-4411.  
695 doi:10.1093/bioinformatics/btac512
- 696 39. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional  
697 annotation using genome-wide association summary statistics. OriginalPaper. *Nature Genetics.*  
698 2015-09-28 2015;47(11):1228-1235. doi:doi:10.1038/ng.3404
- 699 40. Cherian PT, Al-Khairi I, Sriraman D, et al. Increased Circulation and Adipose Tissue Levels  
700 of DNAJC27/RBJ in Obesity and Type 2-Diabetes. *Front Endocrinol (Lausanne).* 2018;9:423.  
701 doi:10.3389/fendo.2018.00423
- 702 41. Rollins DA, Coppo M, Rogatsky I. Minireview: nuclear receptor coregulators of the p160  
703 family: insights into inflammation and metabolism. *Mol Endocrinol.* Apr 2015;29(4):502-17.  
704 doi:10.1210/me.2015-1005
- 705 42. Mohsen G A-M, Abu-Taweel GM, Rajagopal R, et al. Betulinic acid lowers lipid  
706 accumulation in adipocytes through enhanced NCoA1-PPAR $\gamma$  interaction. *Journal of Infection*  
707 *and Public Health.* 2019/09/01/ 2019;12(5):726-732.  
708 doi:<https://doi.org/10.1016/j.jiph.2019.05.011>
- 709 43. Ulgen E, Ozisik O, Sezerman OU. pathfindR: An R Package for Comprehensive  
710 Identification of Enriched Pathways in Omics Data Through Active Subnetworks. *Methods.*  
711 *Frontiers in Genetics.* 2019-September-25 2019;10doi:10.3389/fgene.2019.00858

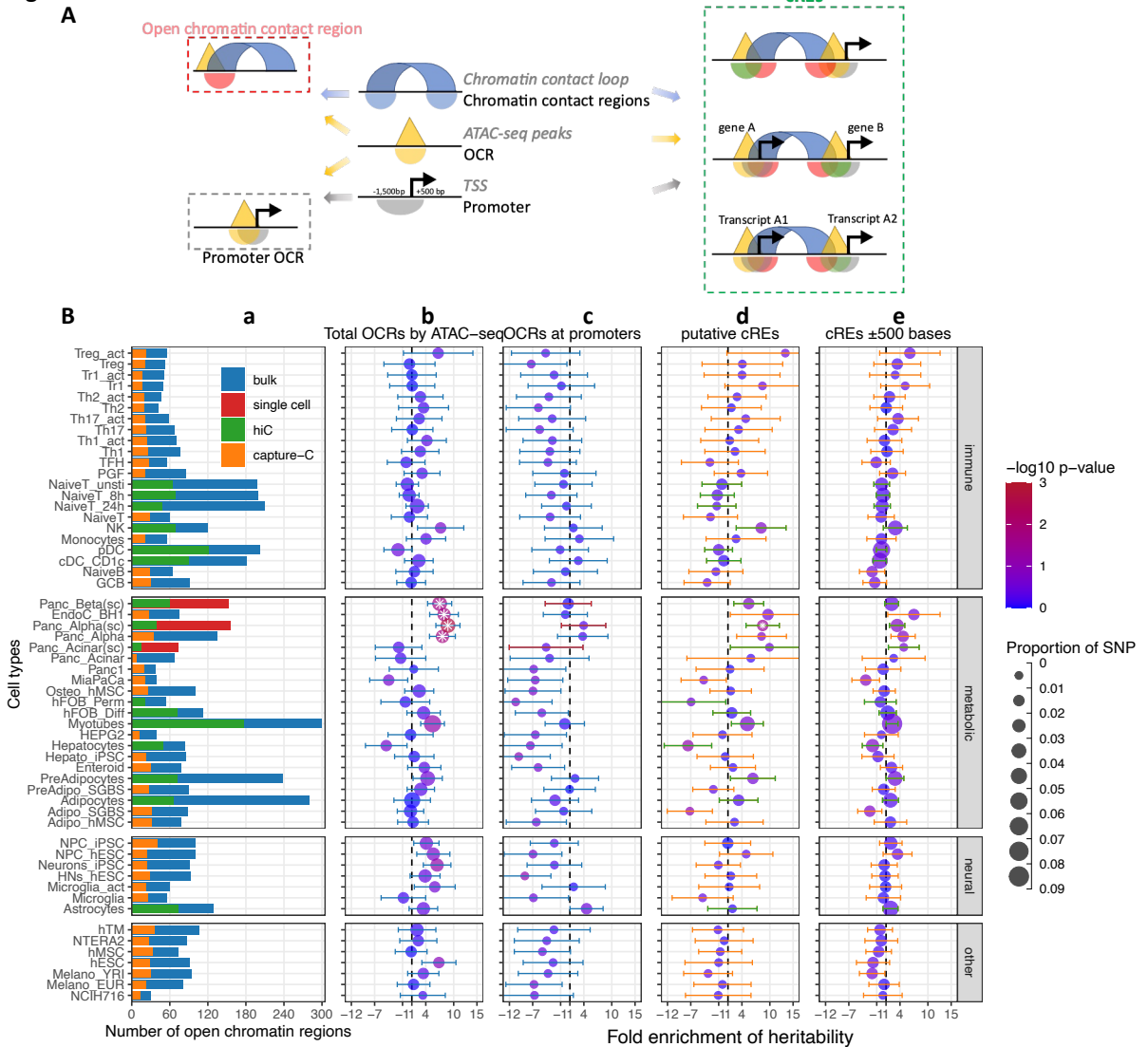


- 712 44. Derosa G, Ferrari I, D'Angelo A, et al. Matrix metalloproteinase-2 and -9 levels in obese  
713 patients. *Endothelium*. Jul-Aug 2008;15(4):219-24. doi:10.1080/10623320802228815
- 714 45. Sezgin SBA, Bayoglu B, Ersoz F, et al. Downregulation of MMP-2 and MMP-9 genes in  
715 obesity patients and their relation with obesity-related phenotypes. *Turkish Journal of*  
716 *Biochemistry*. 2022;47(4):425-433. doi:10.1515/tjb-2021-0124
- 717 46. Nonino CB, Noronha NY, de Araújo Ferreira-Julio M, et al. Differential Expression of  
718 MMP2 and TIMP2 in Peripheral Blood Mononuclear Cells After Roux-en-Y Gastric Bypass. Brief  
719 Research Report. *Frontiers in Nutrition*. 2021-October-13 2021;8doi:10.3389/fnut.2021.628759
- 720 47. Mohammadi Pejman 5 6 Park YoSon 11 Parsana Princy 12 Segrè Ayellet V. 1 Strober  
721 Benjamin J. 9 Zappala Zachary 7 8 GCLaAFBAACSEDJRHYJB, P. 19 Volpi Simona 19  
722 NpmAAGPKSLARLNCMHMRASJ, 16 PSLBMEBPA, 137 NCFNCR. Genetic effects on gene  
723 expression across human tissues. *Nature*. 2017;550(7675):204-213.
- 724 48. Gararup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk  
725 of obesity and type 2 diabetes. *Nature genetics*. 2018;50(2):172-174. doi:10.1038/s41588-017-  
726 0022-7
- 727 49. Stergiakouli E, Gaillard R, Tavaré JM, et al. Genome-wide association study of height-  
728 adjusted BMI in childhood identifies functional variant in ADCY3. *Obesity (Silver Spring)*. Oct  
729 2014;22(10):2252-9. doi:10.1002/oby.20840
- 730 50. Sinnott-Armstrong N, Sousa IS, Laber S, et al. A regulatory variant at 3q21.1 confers an  
731 increased pleiotropic risk for hyperglycemia and altered bone mineral density. *Cell Metabolism*.  
732 2021/03/02/ 2021;33(3):615-628.e13. doi:<https://doi.org/10.1016/j.cmet.2021.01.001>
- 733 51. Kikuchi H, Jung HJ, Raghuram V, et al. Bayesian identification of candidate transcription  
734 factors for the regulation of Aqp2 gene expression. *Am J Physiol Renal Physiol*. Sep 1  
735 2021;321(3):F389-f401. doi:10.1152/ajprenal.00204.2021
- 736 52. Li Z, Schulz MH, Look T, Begemann M, Zenke M, Costa IG. Identification of transcription  
737 factor binding sites using ATAC-seq. *Genome Biology*. 2019/02/26 2019;20(1):45.  
738 doi:10.1186/s13059-019-1642-2
- 739 53. Sobreira DR, Joslin AC, Zhang Q, et al. Extensive pleiotropism and allelic heterogeneity  
740 mediate metabolic effects of IRX3 and IRX5. *Science (New York, NY)*. 2021;372(6546):1085-  
741 1091. doi:10.1126/science.abf1008
- 742 54. Steinsaltz D, Dahl A, Wachter KW. On Negative Heritability and Negative Estimates of  
743 Heritability. *Genetics*. 2020;215(2):343-357. doi:10.1534/genetics.120.303161
- 744 55. Claussnitzer M, Hui CC, Kellis M. FTO Obesity Variant and Adipocyte Browning in  
745 Humans. *N Engl J Med*. Jan 14 2016;374(2):192-3. doi:10.1056/NEJMc1513316
- 746 56. Smemo S, Tena JJ, Kim K-H, et al. Obesity-associated variants within FTO form long-  
747 range functional connections with IRX3. *Nature*. 2014/03/01 2014;507(7492):371-375.  
748 doi:10.1038/nature13138
- 749 57. Ragvin A, Moro E, Fredman D, et al. Long-range gene regulation links genomic type 2  
750 diabetes and obesity risk regions to HHEX, SOX4, and IRX3. *Proc Natl Acad Sci U S A*. Jan 12  
751 2010;107(2):775-80. doi:10.1073/pnas.0911591107
- 752 58. Johnston KJA, Adams MJ, Nicholl BI, et al. Identification of novel common variants  
753 associated with chronic pain using conditional false discovery rate analysis with major  
754 depressive disorder and assessment of pleiotropic effects of LRFN5. *Transl Psychiatry*. Nov 20  
755 2019;9(1):310. doi:10.1038/s41398-019-0613-4

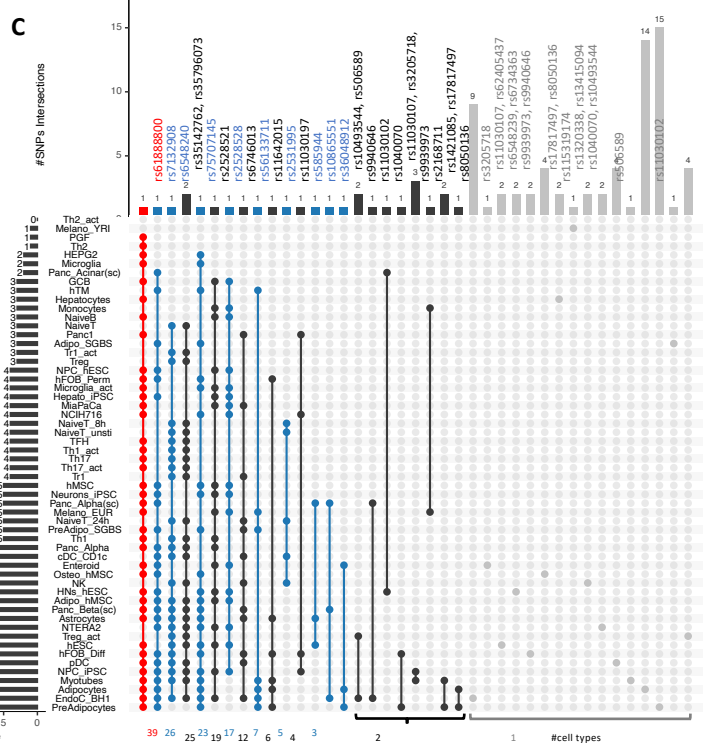
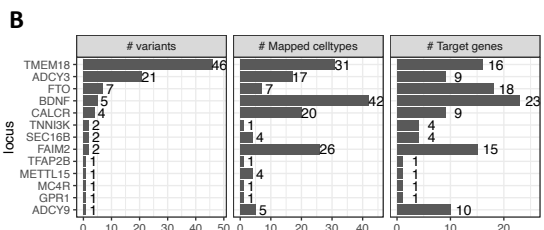
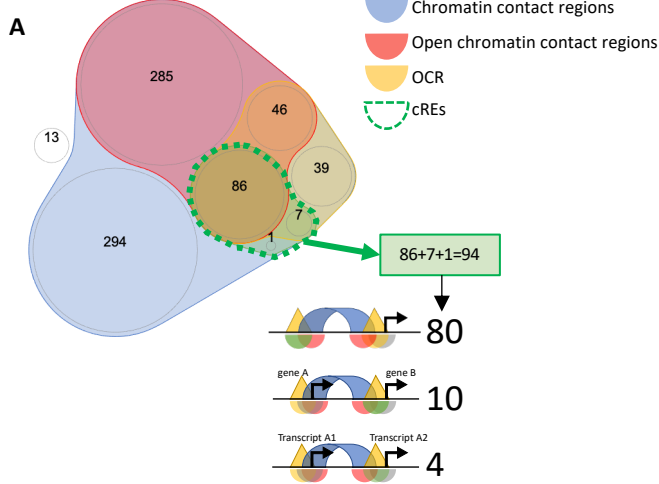
- 756 59. Sanchez-Roige S, Fontanillas P, Jennings MV, et al. Genome-wide association study of  
757 problematic opioid prescription use in 132,113 23andMe research participants of European  
758 ancestry. *Mol Psychiatry*. Nov 2021;26(11):6209-6217. doi:10.1038/s41380-021-01335-3
- 759 60. Graff M, Ngwa JS, Workalemahu T, et al. Genome-wide analysis of BMI in adolescents  
760 and young adults reveals additional insight into the effects of genetic loci over the life course.  
761 *Hum Mol Genet*. Sep 1 2013;22(17):3597-607. doi:10.1093/hmg/ddt205
- 762 61. Lee JS, Cheong HS, Shin HD. BMI prediction within a Korean population. *PeerJ*.  
763 2017;5:e3510. doi:10.7717/peerj.3510
- 764 62. Fernandes SJ, Morikawa H, Ewing E, et al. Non-parametric combination analysis of  
765 multiple data types enables detection of novel regulatory mechanisms in T cells of multiple  
766 sclerosis patients. *Sci Rep*. Aug 19 2019;9(1):11996. doi:10.1038/s41598-019-48493-7
- 767 63. Blessing AM, Ganesan S, Rajapakshe K, et al. Identification of a Novel Coregulator,  
768 SH3YL1, That Interacts With the Androgen Receptor N-Terminus. *Mol Endocrinol*. Oct  
769 2015;29(10):1426-39. doi:10.1210/me.2015-1079
- 770 64. Kobayashi M, Harada K, Negishi M, Katoh H. Dock4 forms a complex with SH3YL1 and  
771 regulates cancer cell migration. *Cell Signal*. May 2014;26(5):1082-8.  
772 doi:10.1016/j.cellsig.2014.01.027
- 773 65. Choi GS, Min HS, Cha JJ, et al. SH3YL1 protein as a novel biomarker for diabetic  
774 nephropathy in type 2 diabetes mellitus. *Nutr Metab Cardiovasc Dis*. Feb 8 2021;31(2):498-505.  
775 doi:10.1016/j.numecd.2020.09.024
- 776 66. Gaynor SC, Monson ET, Gaine ME, et al. Male-specific association of the 2p25 region  
777 with suicide attempt in bipolar disorder. *J Psychiatr Res*. Feb 2020;121:151-158.  
778 doi:10.1016/j.jpsychires.2019.11.009
- 779 67. Wu Z, Liu Z, Ge W, et al. Analysis of potential genes and pathways associated with the  
780 colorectal normal mucosa-adenoma-carcinoma sequence. *Cancer Med*. Jun 2018;7(6):2555-  
781 2566. doi:10.1002/cam4.1484
- 782 68. Huang Z, Yang Q, Huang Z. Identification of Critical Genes and Five Prognostic  
783 Biomarkers Associated with Colorectal Cancer. *Med Sci Monit*. Jul 5 2018;24:4625-4633.  
784 doi:10.12659/msm.907224
- 785 69. Pleasure SJ, Lee VM. NTera 2 cells: a human cell line which displays characteristics  
786 expected of a human committed neuronal progenitor cell. *J Neurosci Res*. Aug 15  
787 1993;35(6):585-602. doi:10.1002/jnr.490350603
- 788 70. Dallman MF. Stress-induced obesity and the emotional nervous system. *Trends*  
789 *Endocrinol Metab*. Mar 2010;21(3):159-65. doi:10.1016/j.tem.2009.10.004
- 790 71. Iwahara T, Fujimoto J, Wen D, et al. Molecular characterization of ALK, a receptor  
791 tyrosine kinase expressed specifically in the nervous system. *Oncogene*. Jan 30 1997;14(4):439-  
792 49. doi:10.1038/sj.onc.1200849
- 793 72. Motegi A, Fujimoto J, Kotani M, Sakuraba H, Yamamoto T. ALK receptor tyrosine kinase  
794 promotes cell growth and neurite outgrowth. *J Cell Sci*. Jul 1 2004;117(Pt 15):3319-29.  
795 doi:10.1242/jcs.01183
- 796 73. Orthofer M, Valsesia A, Mägi R, et al. Identification of ALK in Thinness. *Cell*. 2020/06/11/  
797 2020;181(6):1246-1262.e22. doi:<https://doi.org/10.1016/j.cell.2020.04.034>

- 798 74. Defaye M, Iftinca MC, Gadotti VM, et al. The neuronal tyrosine kinase receptor ligand  
799 ALKAL2 mediates persistent pain. *The Journal of Clinical Investigation*. 06/15/  
800 2022;132(12)doi:10.1172/JCI154317
- 801 75. Sun W, Yang S, Wu S, et al. Transcriptome analysis reveals dysregulation of  
802 inflammatory and neuronal function in dorsal root ganglion of paclitaxel-induced peripheral  
803 neuropathy rats. *Mol Pain*. May 24 2022;17448069221106167.  
804 doi:10.1177/17448069221106167
- 805 76. Yao S, Wu H, Ding J-M, et al. Transcriptome-wide association study identifies multiple  
806 genes associated with childhood body mass index. *International Journal of Obesity*. 2021/05/01  
807 2021;45(5):1105-1113. doi:10.1038/s41366-021-00780-y
- 808 77. Riveros-McKay FA-O, Mistry V, Bounds RA-O, et al. Genetic architecture of human  
809 thinness compared to severe obesity. 2019;(1553-7404 (Electronic))
- 810 78. Borenäs M, Umapathy G, Lai W-Y, et al. ALK ligand ALKAL2 potentiates MYCN-driven  
811 neuroblastoma in the absence of ALK mutation. *The EMBO Journal*. 2021;40(3):e105784.  
812 doi:<https://doi.org/10.15252/emj.2020105784>
- 813 79. Siljee JE, Wang Y, Bernard AA, et al. Subcellular localization of MC4R with ADCY3 at  
814 neuronal primary cilia underlies a common pathway for genetic predisposition to obesity. *Nat*  
815 *Genet*. Feb 2018;50(2):180-185. doi:10.1038/s41588-017-0020-9
- 816 80. Timshel PN, Thompson JJ, Pers TH. Genetic mapping of etiologic brain cell types for  
817 obesity. *eLife*. 2020/09/21 2020;9:e55851. doi:10.7554/eLife.55851
- 818 81. Licinio J, Dong C Fau - Wong M-L, Wong ML. Novel sequence variations in the brain-  
819 derived neurotrophic factor gene and association with major depression and antidepressant  
820 treatment response. (1538-3636 (Electronic))
- 821 82. Colle R, Deflesselle E, Martin S, et al. BDNF/TRKB/P75NTR polymorphisms and their  
822 consequences on antidepressant efficacy in depressed patients. (1744-8042 (Electronic))  
823

**Figure 1**



**Figure 2**



**Figure 3**

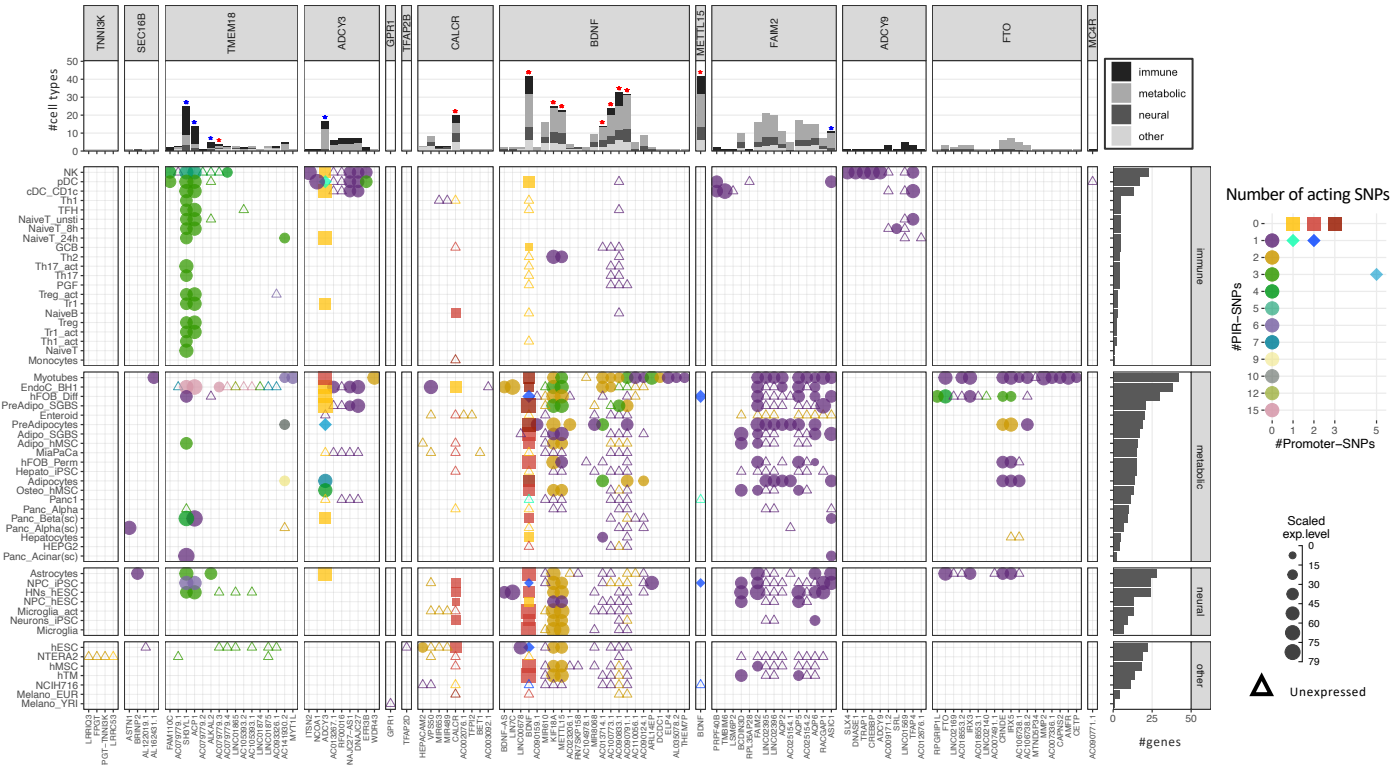
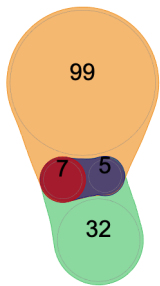


Figure 4

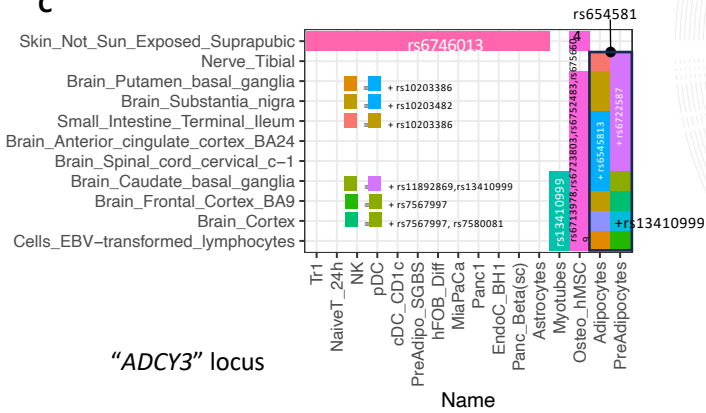
A



Number of putative target genes identified by:

- ColocQuial
- Variant-to-gene mapping
- Overlapped but linked different effector variant
- Overlapped by same effector variant

C



B

