Intestinal epigenomic alterations are associated with a dysregulated nutrient absorption phenotype in obesity

Dilhana S. Badurdeen[1], Zhen Li[1], Jeong-Heon Lee[2], Tao Ma[3], Aditya Vijay Bhagwate[3], Rachel Latanich[4], Arjit Dogiparthi[4], Tamas Ordog[5], Olga Kovbasnjuk[6], Vivek Kumbhari[1], Jennifer Foulke-Abel[4]*

[1]Division of Gastroenterology and Hepatology, Mayo Clinic in Florida, Jacksonville, FL 32224, USA

[2]Division of Experimental Pathology and Laboratory Medicine, Department of Laboratory Medicine, and Pathology, and Center for Cell Signaling in Gastroenterology (C-SiG), Mayo Clinic, Rochester, MN 55905, USA

[3]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN 55905, USA

[4]Division of Gastroenterology and Hepatology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[5]Department of Physiology and Biomedical Engineering, Division of Gastroenterology and Heptaology, and Center for Cell Signaling in Gastroenterology (C-SiG), Mayo Clinic, Rochester, MN 55905, USA

[6]National Institute for General Medical Sciences, Bethesda, MD 20892, USA

* **Corresponding author:** Jennifer Foulke-Abel, PhD, Johns Hopkins University School of Medicine, Division of Gastroenterology & Hepatology, 720 Rutland Ave, Ross 929, Baltimore, MD 21205. e-mail: jfoulke@jhmi.edu

ABBREVIATIONS

ATAC (assay for transposase-accessible chromatin), AUC (area under the curve), BMI (body mass index), DEG (differentially expressed gene), FDR (false discovery rate), FPKM (fragments per kilobase of transcript per million mapped reads), GOBP (Gene Ontology Biological Process), ISC (intestinal stem cell), qRT-PCR (quantitative real-time polymerase chain reaction), ROC (receiver-operator curves), TF (transcription factor), TSS (transcriptional start site)

ABSTRACT

Obesity is an epidemic with myriad health effects, but little is understood regarding individual obese phenotypes and how they may respond to therapy. Epigenetic changes associated with obesity have been detected in blood, liver, pancreas, and adipose tissues. Previous work found that dietary glucose hyperabsorption occurs in some obese subjects, but detailed transcriptional or epigenomic features of the intestine associated with this phenotype are unknown. This study evaluated differentially expressed genes and relative chromatin accessibility in intestinal organoids established from donors classified as lean, obese, or obese hyperabsorptive by body mass index and glucose transport assays. Transcriptomic analysis indicated that obese hyperabsorptive subjects have significantly upregulated dietary nutrient absorption proteins and downregulated type I interferon targets. Chromatin accessibility and transcription factor footprinting suggested that enhanced binding of HNF4G promotes the obese hyperabsorption phenotype. Quantitative PCR assessment in a larger subject cohort suggested that intestinal epithelial expression of CUBN, GIP, and SLC2A5 have high correlation with hyperabsorption. The obese hyperabsorption phenotype is characterized by transcriptional changes that support increased nutrient uptake and may be driven by differentially accessible chromatin. Recognizing unique intestinal phenotypes in obesity provides new perspective in considering therapeutic targets and options to manage the disease.

Keywords: metabolic disease; intestinal organoids; ATAC-sequencing; nutrient absorption

INTRODUCTION

Obesity is a growing global public health crisis, with over 1.9 billion adults being overweight, and 650 million obese.[1] In 2016, the annual direct healthcare cost of obesity in the U.S. was $480 billion.[2] The World Health Organization recognizes obesity as a risk factor for type 2 diabetes, coronary heart disease, asthma, stroke, and multiple cancers. Obesity-associated diseases place an additional burden on the healthcare system, with medical costs estimated to be 30% higher for obese individuals (BMI ≥30) compared to those that are lean (BMI ≤25). Among the available therapies, bariatric surgery is the most effective intervention for significant and sustained weight loss.[3] However, surgical outcomes are

heterogeneous with 15-35% of patients experiencing weight loss failure (<50% excess weight loss at 12-18 months) after roux-en-Y gastric bypass (RYGB).[4,5] Numerous hypotheses have fallen short of fully identifying and characterizing molecular changes catalyzed by bariatric procedures.[6] Individualized weight loss success prediction based on phenotypic classification would be important not only to guide selection of an effective intervention, but also to minimize potential risks and complications from ineffective procedures.

Among the many phenotypes associated with obesity, accelerated intestinal glucose absorption observed in nondiabetic obese subjects[7] is a striking example of an altered physiological response that likely exacerbates the disease yet lacks full mechanistic understanding. Intestinal organoids/enteroids are a unique experimental tool that stably retain intestinal region- and donor-specific genotype and phenotypes.[8] In our prior study using patient-derived jejunal organoids from lean (control group, G1) and obese subjects,[9] we found that a subset of obese donors exhibited significantly increased epithelial glucose absorption capacity (obese hyperabsorption group, G2) relative to G1 lean controls or other obese donors (G3). This observation suggested that intestinal metabolic phenotypes may be genetically encoded and differ among obese subjects. It further intimated that genetic regulation of dietary sugar hyperabsorption may contribute to metabolic disease in G2 subjects. Considering the relatively recent global surge in obesity development and the paucity of evidence for specific protein-coding gene mutations to drive this outcome, we hypothesized that potential stable epigenetic modifications in G2 individuals may account for the hyperabsorption phenotype. In this study, we analyzed the transcriptomes of 3 representative donors from each group (G1, G2, and G3) to identify differentially expressed genes characteristic of each group. We also conducted an assay for transposase-accessible chromatin (ATAC) sequencing to determine differential chromatin accessibility, and then carried out ATAC footprinting and motif analysis to detect putative transcription factors differentially occupying chromatin among the three groups. Transcript levels distinctive for differing metabolic profiles were additionally validated in an expanded cohort of lean and obese subject organoids. Collectively, our data provides important insights into the dysregulated transcriptional state and altered open chromatin landscape of intestinal epithelia in obese glucose hyperabsorptive subjects.

82

83 **RESULTS**

84 **Differential enrichment of metabolism-associated transcripts in obese donors exhibiting glucose**
85 **hyperabsorption**

86 Donor organoid lines were previously characterized as lean controls (G1, BMI ≤25, n=3), obese with
87 elevated intestinal glucose absorption relative to lean controls (G2, BMI ≥35, n=3), or obese with intestinal
88 glucose absorption similar to lean controls (G3, BMI ≥35, n=3).[9] Bulk RNA-seq was conducted to
89 identify differentially expressed genes (DEG, absolute $\log_2$ fold change>2, p<0.05) between the three
90 groups (**Fig 1A** and **Supplementary Table S1**). Principal component analysis demonstrated that the
91 overall transcriptional profiles among the three groups have considerable overlap, and that one individual
92 in each group appeared to segregate slightly from the data set as a whole (**Fig 1B**). Comparison of G2
93 to either G1 or G3 each revealed 456 DEGs, many of which function in dietary nutrient absorption,
94 metabolism, and regulation (**Fig 1C, D**). In contrast, G1 and G3 were distinguished by only 27 DEGs,
95 suggesting highly similar transcriptomes (**Fig 1E**). The enriched gene ontology biological processes
96 (GOBP) in G2 included glucose transport, cholesterol transport, and triglyceride homeostasis (Cluster
97 #1, **Fig 1F, G**). Downregulated GOBP in G2 epithelia included type I interferon signaling associated with
98 regulation of innate immune responses (Cluster #2, **Fig 1F, G**).

99 ***Sugar transport and metabolism.*** Significantly increased transcripts in G2 included the major
100 facilitators of intestinal glucose and fructose absorption: SLC2A2, SLC2A5, SLC5A1, and SLC5A4 (**Fig**
101 **1H**), which confirmed and expanded the findings of prior work.[9] As seen previously, levels of
102 gluconeogenic enzymes glucose-6-phosphatase (G6PC) and phosphoenolpyruvate carboxykinase
103 (PCK1) were also elevated. Ketohexokinase (KHK) and aldolase B (ALDOB) were also increased in G2,
104 suggesting that the epithelia is primed to catabolize an influx of dietary fructose entering through GLUT5.
105 Elevated transcripts for the brush border enzyme sucrase-isomaltase (SI) may enhance digestion of
106 sucrose and maltose into simple sugars for intestinal uptake, and dipeptidyl peptidase IV (DPP4) would
107 have potential to activate gut hormones involved in metabolic regulation.

108 ***Lipid transport and metabolism.*** Dietary fatty acids can be either passively absorbed or transit via

membrane-bound proteins. Transporters situated at the apical plasma membrane including fatty acid translocase (CD36) and apolipoprotein M (APOM) were significantly upregulated in G2 (**Fig 1H**), suggesting enhanced uptake ability. Additional fatty acid carriers included secreted apolipoproteins APOA1, APOA4, APOB, APOC2, and APOC3 that participate in lipid transport and act as enzyme cofactors for lipid metabolism. Other amplified transcripts of note included receptors for lipoproteins (CUBN and SCARB1), a channel for cholesterol-containing micelle import (NPC1L1), and SOAT2, which catalyzes cholesterol esterification necessary for lipoprotein assembly. The rate limiting enzyme in mitochondrial fatty acid oxidation (CPT1A) was relatively unchanged, suggesting that enterocytes are primarily facilitating fatty acid transport rather than executing catabolism.

Bile acid absorption and signaling. Bile acids are normally introduced in the duodenum to assist with solubilizing dietary fats and are subsequently reabsorbed in the distal ileum. Suprisingly, G2 demonstrated transcriptional enrichment in the jejunum for proteins involved in bile acid reabsorption that are usually most prominent in distal ileum (**Fig 1H**). These included the major intracellular carrier for bile acids, fatty acid binding protein 6 (FABP6), and the basolateral bile acid exporter subunits SLC51A and SLC51B, as well as the bile acid responsive transcription factor FXR (NR1H4) and MALRD1, a cofactor known to augment intestinal FGF19 secretion.[10] FGF19 transcripts were similar among the three groups, which would be expected due to the absence of exogenous bile acids in the organoid culture media to induce expression.

Intestinal hormones. Peptide hormones were another class of signaling molecule with differential expression that distinguished G2 from G1 and G3. G2 subjects were found to overexpress guanylin (GUCA2A), an intestinal peptide hormone produced by enterocytes and Paneth cells with pleiotropic effects on epithelial transport via induction of cGMP synthesis (**Fig 1H**). Upregulated transcripts related to enteroendocrine cell (EEC)-specific hormones, including glucose-dependent insulinotropic polypeptide (GIP), motilin (MLN), somatostatin (SST), and the prohormone chromogranin A (CHGA) were also detected. Given the relative rarity of EECs in organoids grown under conditions that do not optimize EEC development,[11] the intrinsic emergence of these hormone transcripts is further evidence of differential transcriptional controls in the G2 state.

136 Collectively, these results indicated that the obese glucose hyperabsorption phenotype (G2) is
137 characterized by upregulated transcription of nutrient transporters and metabolic regulators that may
138 contribute to dysregulated energy homeostasis in a subset of obese subjects.

139 **Downregulation of epithelial innate immune response transcripts in G2 and G3**

140 Many of the downregulated targets distinctive in G2 included interferon-stimulated genes (ISGs) and
141 downstream molecules linked to the type I interferon (IFN) signaling pathway (**Fig 1I**). These included
142 antiviral components such as interferon induced transmembrane proteins IFITM1, IFITM2, and IFITM3
143 that restrict viral entry, and OAS2 and OASL that act as sensors for double-stranded RNA. Innate defense
144 molecules including antimicrobial proteins REG1A, REG1B, REG3A, SLPI, and SEMG1 were also
145 significantly decreased relative to G1. Receptors with immune-associated functions including IL18R1,
146 LIFR, OSMR, CD40, and LTK were also reduced. Interestingly, G3 shared downregulation of many innate
147 immune-related genes that were observed in G2, implying that innate immune repression among obese
148 subjects may be more predictable than dietary nutrient absorption phenotype.

149 **Differential open chromatin regions in G2 suggest potential epigenetic modifications**

150 ATAC-seq was performed to investigate the potential for an altered chromatin landscape that would
151 account for the marked differential gene expression in G2 relative to G1 or G3 (**Supplementary Table**
152 **S2**). Comparison of G2 to G1 yielded 2,051 differential sites (**Fig 2A**), and G2 vs G3 indicated 3,549
153 differential sites (**Fig 2B**). G1 and G3 differed by a more modest 491 sites (**Fig 2C**). The categorical
154 locations of differential open chromatin were similarly distributed in each of the three paired group
155 comparisons. Notably, there were distinct open chromatin regions proximal to transcriptional start sites
156 (TSS) in G2 that were not equivalent in G1 nor G3 (**Fig 2D**). Many of these sites coincided with increased
157 levels of downstream gene transcripts, suggesting an increase in enhancer binding activity.

158 When integrated with transcriptional data, pathways uniquely enriched in G2 were centered on
159 digestion and lipid transport/metabolism, an example of which is APOA4 (**Fig 2E**). In contrast, innate
160 immune-related transcripts appeared to be less prominent due to more restricted chromatin near the
161 promoter region of genes such as interferon-inducible antiviral enzyme 2'-5'-oligoadenylate synthetase 2
162 (OAS2) (**Fig 2F**).

163 **ATAC footprinting indicates differential binding of transcription factors (TFs) associated with**
164 **metabolic and innate immune regulation in G2**

165 ATAC footprinting in G2 compared to G1 (**Fig 3A**) or to G3 (**Fig 3B**) yielded significant differential
166 binding predictions that were in general agreement with transcriptional analyses and were highly similar
167 between the two comparators. Comparison of G1 and G3 did not render any significant differential binding
168 results, which further reflects their transcriptional similarity. G2 footprints suggested increased binding of
169 nuclear receptor ligand-activated TFs known to regulate digestive and metabolic processes, including
170 hepatocyte nuclear factor 4 paralogs (HNF4A, HNF4G) and numerous retinoic acid X receptor (RXR)
171 heterodimers. Lower scoring but significant binding prediction also implicated peroxisome proliferator
172 activated receptors (PPARA, PPARD, PPARG) which function as key regulators of lipid metabolism.

173 Integrated binding motif analysis ranked HNF4G as the top factor involved in G2 upregulated gene
174 transcription (p = 1e-40) (**Fig 3C**). HNF4G transcripts were also slightly upregulated ($\log_2$FC = 1.15) in
175 G2, suggesting that it may exert dominance in transcriptional regulation in the G2 state. Further, RNA-
176 seq FPKM values for HNF4A and HNF4G were approximately equivalent in G1 and G3, but HNF4G
177 FPKM was double that of HNF4A in G2. A combination of increased open chromatin sites conducive to
178 HNF4 paralog binding and increased HNF4G levels suggest that the G2 state may be driven at least in
179 part by HNF4G activity.

180 Decreased footprinting was observed for Jun proto-oncogene (JUN), Fos proto-oncogene (FOS), and
181 activating transcription factor (ATF) family proteins that constitute the activator protein-1 (AP-1) TF
182 complex. AP-1 complexes regulate diverse cell responses related to inflammation, proliferation,
183 differentiation, and apoptosis. Motif analysis ranked FOSL1 sites as least accessible in G2 compared to
184 G1 (p = 1e-245) or to G3 (p = 1e-427). Although there were no significant differential footprints detected
185 in G1 versus G3, motif analysis did indicate that FOSL2::JUN sites were less accessible in G3 (p = 1e-
186 20), hinting that some AP-1 mediated processes may be attenuated in obese subjects with normative
187 dietary nutrient absorption profiles. This was corroborated in the RNA-seq analysis showing that some of
188 the innate immune-associated transcripts notably downregulated in G2 were also decreased in G3
189 relative to G1 (**Fig 1I**).

190 **Analysis of additional subjects corroborates upregulated genes characteristic of G2**

191 To follow up on the general findings from the nine selected donors analyzed in RNA-seq and ATAC-
192 seq experiments, we conducted quantitative real-time PCR (qRT-PCR) analysis of cDNA isolated from
193 additional organoids representing G1 and G2 subjects to validate transcriptional trends using a curated
194 panel of markers. The expanded analysis did not identify additional G3 donors outside of the original
195 three analyzed by RNA-seq. As expected, donors that were phenotyped as G2 in glucose transport
196 assays exhibited markedly increased transcript levels for select genes (presented quantitatively as ΔC_T
197 values) relative to G1 donors that were generally consistent with the RNA-seq findings (**Fig 4A**). When
198 comparing individual genes among larger groups of G1 and G2 donors, some transcripts appeared to
199 have high predictive value (GIP, CUBN, and SLC2A5), while others were less robust (GUCA2A, APOA4)
200 (**Fig 4B**). This implies that the G2 phenotype could be characterized through transcriptomic evaluation of
201 intestinal epithelia in lieu of time- and labor-intensive functional organoid-based assays, and that with
202 further cohort expansion and validation, the development of a qRT-PCR-based diagnostic panel may be
203 feasible.

204

205 **DISCUSSION**

206 Intestinal organoids derived from individuals with metabolic syndrome indicate that some subjects
207 have distinct transcriptional profiles consistent with dysregulated energy balance. The intestinal stem
208 cells of G2 subjects appeared to maintain genetic memory of their in vivo state with entrained upregulation
209 of fatty acid uptake, dietary sugar transport and metabolism, and specific gut hormones that was
210 independent of their immediate diet, given that all of the organoids were provided with the same
211 media/nutrition throughout the experimental process. It should be noted that organoid media could be
212 introducing bias in some of the transcriptional outputs relative to native tissue due to the specific growth
213 factors necessary for culture sustenance. However, the transcriptional differences observed were intrinsic
214 to the epithelia and not dependent on other tissue components (mesenchyme, immune, muscle,
215 endothelia, lymphatics, neurons, or microbes). The transcriptional signatures appear stable in organoid
216 culture, as the cell lines tested were from a range of passages between 10-38 and have been

217 cryopreserved and rethawed multiple times.

218 Diet may play a major role in catalyzing the intestinal adaptation toward a hyperabsorptive phenotype.
219 Epithelia has the ability to undergo a rapid transcriptional shift toward increased fatty acid absorption
220 capacity in response to dietary nutrient influx, even observed within days of transitioning mice from
221 standard chow to a high fat diet.[12] Numerous studies have examined the influence of diet composition
222 on intestinal stem cell (ISC) behavior in animal models, finding that obesogenic high sugar and/or high
223 fat diets alter epithelial proliferation and lineage differentiation trajectories, and confirming that ISCs can
224 sense and respond to nutrients.[13-15] Studies evaluating the effect of diet on ISCs could also explain
225 the altered gut hormones evident in G2 subjects. Increased levels of GIP, SST, and MLN in G2 subjects
226 are consistent with observations that dietary exposures alter development of enteroendocrine cell
227 subtypes in mice.[12,13] However, the duration and magnitude of dietary imbalance that may lead to
228 durable epigenetic modifications is not well defined. This study provides insight not only toward the
229 contribution of intestinal epithelia as a facilitator of excess nutrient absorption, but also suggests that
230 acquisition of hard-wired transcriptional changes in these functions may undermine efforts to restrict
231 calorie intake.

232 TF footprinting and DNA binding motif analysis implicated HNF4 as the top-scoring regulator of the
233 G2 hyperabsorptive phenotype, and the role of HNF4 in intestinal absorptive/metabolic function is
234 strongly supported by prior animal studies. HNF4A and HNF4G function as highly redundant paralogs in
235 the intestinal epithelium, with HNF4G being predominantly intestine-restricted.[16] HNF4 has become
236 known for orchestrating critical programming of intestinal enterocyte maturation, brush border function,
237 and fatty acid oxidation.[17,18] Increased occupancy of PPARD and retinoic acid X receptor α (RXRA)
238 binding sites were also predicted in G2, consistent with known roles in sensing and responding to diet.
239 Despite evidence of more active HNF4, PPARD, and RXRA, the transcripts for rate-limiting enzymes
240 involved in fatty acid oxidation (FAO) such as CPT1A were not upregulated. This could be a consequence
241 of the limited fatty acid content in standard organoid medium, so it would be interesting to assess whether
242 FAO pathways are indeed primed and would be more rapidly mobilized in G2 organoid monolayers than
243 their G1 counterparts if provided lipid-enriched media.

Downregulated transcripts in G2 were more difficult to functionally interpret as they group loosely with various facets of innate immune responses. Footprinting and motif analysis suggest restricted binding of AP-1 complex members which have pleiotropic cellular functions that include inflammation and immunity. While intestinal epithelia is not first on the list of major immunity determinants, it does function as the interface with microbial and environmental constituents by coordinating barrier function, luminal sampling by professional immune cells, and secretion of defense molecules. A number of the downregulated genes in G2 are interferon inducible and suggest blunting of type I IFN signaling. A correlation between obesity and repressed type I IFN/antiviral responses has been noted in previous studies using human peripheral blood mononuclear cells[19] and diet-induced obese mice.[20,21] Thought to arise as a consequence of chronic inflammation in metabolic disease, a dampened ability to mount IFN responses during obesity may also be epigenetically controlled since the effect persisted in organoid culture and chromatin accessibility was restricted at associated TF binding sites. Interestingly, some degree of overlap in the downregulated G2 and G3 transcripts relative to G1 hint that the innate immune changes are not necessarily linked to the hyperabsorptive phenotype.

Since some obese subjects (G3) clearly lacked evidence of an intestinal absorptive and metabolic dysregulation profile, genetic imprinting cannot be interpreted as an absolute causal or correlative factor in obesity. Profiles of additional donors are necessary to define the expected proportion of overweight or obese patients with altered intestinal absorption characteristics, and to understand how this phenotype impacts the severity of their metabolic disease, comorbidities, and potential response to therapeutic intervention. However, this study and prior work[9] highlight the importance of personalized approaches to characterizing metabolic disease phenotypes. It is not surprising that the intestinal epithelium undergoes profound imprinted alterations in the obesogenic state given that prior studies have detected differential DNA methylation patterns from other tissues (blood, liver, pancreas, adipose) during obesity.[22] Future studies to understand drivers of the epigenetic changes and identify strategies to limit hyperabsorption in these patients could transform clinical approaches to obesity management.

METHODS

271 **Human organoid cultures**

272 Organoid cultures were established from adult jejunal tissue obtained during upper endoscopy
273 procedures or tissue discarded following bariatric surgeries as described previously.[9] Informed consent
274 from study participants was obtained under protocols approved by the Institutional Review Boards at
275 Johns Hopkins University (NA_0038329), Baylor College of Medicine (H-31793 and H31910), and Mayo
276 Clinic Florida (22-001407). All experiments were performed in accordance with review board guidelines
277 and regulations. Organoids classified as Group 1 were established from lean donors (BMI ≤ 25 kg/m2)
278 and exhibited low intestinal epithelial glucose absorption. Organoids in Group 2 were from obese donors
279 (BMI ≥ 35 kg/m2) and showed elevated intestinal epithelial glucose absorption relative to lean donors.
280 Organoids in Group 3 were from obese donors (BMI ≥ 35 kg/m2) with intestinal epithelial glucose
281 absorption profiles that were similar to lean donors. BMI classification was based on World Health
282 Organization guidelines. Information regarding patient characteristics (age, sex, BMI, absorption
283 phenotype) is provided in **Table 1**. Organoids were cultured in Expansion Media (EM), as previously
284 described (Advanced Dulbecco's Modified Eagle Medium/Ham's F-12, 10 mM HEPES, 0.2 mM
285 GlutaMAX, 50% v/v Wnt3A-conditioned media, 15% v/v R-spondin-1-conditioned media, 10% v/v of
286 Noggin-conditioned media, 50 ng/ml human epidermal growth factor (EGF), 500 nM A83-01, 10 μM SB
287 202190, 1× B27 supplement, 1 mM N-acetylcysteine).[23,24] At the time of culture initation and each
288 weekly culture passage, EM was supplemented with 10 μM CHIR99021 and 10 μM Y-27632 for the first
289 48 h. Growth factor conditioned media was obtained from cell lines expressing Wnt3A (murine L-Wnt3A,
290 CRL-2647, American Type Culture Collection), R-spondin-1 (human HA-Rspo1-Fc/HEK293, kindly
291 provided by Dr. Calvin Kuo, Stanford University, USA), and Noggin (murine Noggin-Fc/HEK293).[25] EM
292 was refreshed every 2-3 d, and organoids were passaged every 7 days. Passage numbers used in this
293 study ranged from 10 to 81.

294 **Organoid monolayers**

295 Organoid monolayers were prepared as previously described.[26] Briefly, isolated organoid fragments
296 resuspended in EM were plated on 24-well plate permeable tissue culture inserts (Corning 3470; 0.4 μm
297 pore, 0.33 cm^2 surface area) coated with collagen IV (10 μg/cm^2, Sigma C5533). To monitor progress

298 toward confluence, transepithelial electrical resistance (TER) was measured using an EVOM2
299 voltohmmeter (World Precision Instruments). Monolayers judged to be confluent by TER measurement
300 and inspection by light microscopy were transferred to differentiation media (DM; Expansion Media
301 lacking Wnt3A, R-spondin-1, A83-01, and SB 202190). Monolayers were used for experiments 5 days
302 after transfer to DM. Differentiated organoid monolayers were digested with Accutase (MilliporeSigma)
303 at 25 °C for 20 min or until cells exhibited rounding and dissociation by light microscopy. The cell
304 suspension was transferred to a low-binding microcentrifuge tube (USA Scientific) and washed with
305 Sorting Buffer (PBS, 25 mM HEPES pH 7.4, 0.5% bovine serum albumin, 5 mM EDTA). Cells were
306 pelleted at 500 xg, 4 °C, 5 min. Buffer was aspirated and cell pellets were snap frozen in liquid nitrogen
307 for RNA-seq, or resuspended at 75,000 cells/100 µL Anti-Freezing Buffer (Advanced DMEM/F-12, 10%
308 fetal calf serum, 10% DMSO) for ATAC-seq, and immediately stored at -80 °C until processed for the
309 respective sequencing applications.

310 **RNA sequencing**

311 Total RNA was isolated from cells and purified by DNase treatment using the RNeasy Micro kit (Qiagen).
312 RNA-seq libraries were prepared by Illumina TruSeq RNA Library Preparation Kit v2 and sequenced on
313 the Illumina HiSeq 4000 platform in the Mayo Clinic Medical Genomics Core. Mayo RNA-sequencing
314 (MAP-RSeq) analysis pipeline was used to perform gene expression analysis. FASTQ readings (paired-
315 end; 100 bp) were aligned to hs38 with TopHat2, generating raw gene and exon counts using
316 featureCount, and sample quality was determined by RSeQC. EdgeRv3.8.6 was used for differential
317 expression analysis and to compare groups. Genes with an absolute $\log_2$ fold change ($\log_2$FC) >1.5 and
318 false discovery rate (FDR) < 0.05 were considered to be significantly differentially expressed.

319 **ATAC sequencing**

320 Approximately 50,000 nuclei were utilized for ATAC-seq. The transposition reaction was carried out in a
321 thermomixer at 37° C for 30 minutes and oscillated at 1,000 rpm. The size of the library DNA was
322 determined by a Fragment Analyzer from the amplified and purified library. The ATAC-seq library was
323 amplified by PCR. By targeting known highly accessible and closed chromatin sites with real-time PCR,
324 the enrichment of each library's accessible regions was determined and expressed as a fold difference.

BWA was used to map paired-end reads to the hs38 genome. Picard SortSam was used to convert SAM files into BAM files and sort by chromosomal coordinates. Picard MarkDuplicates was used to remove PCR duplicates. Processed BAM files were used to call the peak by MACS2. BAM files and peak files obtained from MACS2 was processed using DiffBind to identify differentially accessible genomic loci. Using an FDR threshold of 0.05, Homer was used to annotate the differentially accessible loci to the closest transcription start site (TSS). GREAT was used to conduct gene ontology on the genome regions determined in the ATAC-seq analysis with default settings. Motif analysis was completed by the findMotifsGenome.pl command in Homer. Transcription factor footprinting analysis was conducted using the TOBIAS computational framework.[27]

Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted from differentiated enteroid monolayers using the PureLink RNA Mini Kit (Invitrogen) and reverse transcribed to complementary DNA using SuperScript VILO Master Mix (Invitrogen). Reactions for qRT-PCR were carried out in triplicate using 5 ng cDNA, gene-targeted oligonucleotide pairs (sequences in **Table 2**), and Power SYBR Green Master Mix (Invitrogen) on a QuantStudio 12K Flex instrument (Applied Biosystems). Each gene target was normalized to both 18S ribosomal RNA (RN18S) and GAPDH, and relative ΔC_T values were calculated using the average ΔC_T obtained from all the lean (G1) samples.

DATA AVAILABILITY

Data is provided within the manuscript and in supplementary tables. The complete RNA-seq and ATAC-seq datasets generated and analyzed during the current study are available online in the Gene Expression Omnibus (accession number pending).

REFERENCES

1 Congdon, P. & Amugsi, D. Editorial: The obesity epidemic: Causes, context, prevention. *Front Public Health* **10**, 1030180 (2022). https://doi.org/10.3389/fpubh.2022.1030180

2 Waters, H. & Graf, M. America's Obesity Crisis: The Health and Economic Costs of Excess Weight.

1-30 (Milken Institute, 2020).

3 Maciejewski, M. L. *et al.* Bariatric Surgery and Long-term Durability of Weight Loss. *JAMA Surgery* **151**, 1046-1055 (2016). https://doi.org/10.1001/jamasurg.2016.2317

4 Christou, N. V., Look, D. & MacLean, L. D. Weight Gain After Short- and Long-Limb Gastric Bypass in Patients Followed for Longer Than 10 Years. *Annals of Surgery* **244**, 734-740 (2006). https://doi.org/10.1097/01.sla.0000217592.04061.d5

5 Higa, K., Ho, T., Tercero, F., Yunus, T. & Boone, K. B. Laparoscopic Roux-en-Y gastric bypass: 10-year follow-up. *Surg Obes Relat Dis* **7**, 516-525 (2011). https://doi.org/10.1016/j.soard.2010.10.019

6 Kohli, R., Stefater, M. A. & Inge, T. H. Molecular insights from bariatric surgery. *Rev Endocr Metab Disord* **12**, 211-217 (2011). https://doi.org/10.1007/s11154-011-9172-6

7 Nguyen, N. Q. *et al.* Accelerated intestinal glucose absorption in morbidly obese humans: relationship to glucose transporters, incretin hormones, and glycemia. *J Clin Endocrinol Metab* **100**, 968-976 (2015). https://doi.org/10.1210/jc.2014-3144

8 Günther, C., Winner, B., Neurath, M. F. & Stappenbeck, T. S. Organoids in gastrointestinal diseases: from experimental models to clinical translation. *Gut* **71**, 1892 (2022). https://doi.org/10.1136/gutjnl-2021-326560

9 Hasan, N. M. *et al.* Intestinal stem cell-derived enteroids from morbidly obese patients preserve obesity-related phenotypes: Elevated glucose absorption and gluconeogenesis. *Mol Metab* **44**, 101129 (2021). https://doi.org/10.1016/j.molmet.2020.101129

10 Wang, L. X., Frey, M. R. & Kohli, R. The Role of FGF19 and MALRD1 in Enterohepatic Bile Acid Signaling. *Front Endocrinol (Lausanne)* **12**, 799648 (2021). https://doi.org/10.3389/fendo.2021.799648

11 Zeve, D. *et al.* Robust differentiation of human enteroendocrine cells from intestinal stem cells. *Nature Communications* **13**, 261 (2022). https://doi.org/10.1038/s41467-021-27901-5

12 Enriquez, J. R. *et al.* A dietary change to a high-fat diet initiates a rapid adaptation of the intestine. *Cell Reports* **41**, 111641 (2022). https://doi.org/https://doi.org/10.1016/j.celrep.2022.111641

13 Aliluev, A. *et al.* Diet-induced alteration of intestinal stem cell function underlies obesity and

379 prediabetes in mice. *Nature Metabolism* **3**, 1202-1216 (2021). https://doi.org/10.1038/s42255-021-
380 00458-9
381 14 Lickwar, C. R. *et al.* Transcriptional Integration of Distinct Microbial and Nutritional Signals by the
382 Small Intestinal Epithelium. *Cell Mol Gastroenterol Hepatol* **14**, 465-493 (2022).
383 https://doi.org/10.1016/j.jcmgh.2022.04.013
384 15 Alonso, S. & Yilmaz, Ö. H. Nutritional Regulation of Intestinal Stem Cells. *Annual Review of*
385 *Nutrition* **38**, 273-301 (2018). https://doi.org/10.1146/annurev-nutr-082117-051644
386 16 Chen, L. *et al.* A reinforcing HNF4–SMAD4 feed-forward module stabilizes enterocyte identity.
387 *Nature Genetics* **51**, 777-785 (2019). https://doi.org/10.1038/s41588-019-0384-0
388 17 Chen, L. *et al.* HNF4 Regulates Fatty Acid Oxidation and Is Required for Renewal of Intestinal Stem
389 Cells in Mice. *Gastroenterology* **158**, 985-999.e989 (2020).
390 https://doi.org/10.1053/j.gastro.2019.11.031
391 18 Chen, L. *et al.* HNF4 factors control chromatin accessibility and are redundantly required for
392 maturation of the fetal intestine. *Development* **146** (2019). https://doi.org/10.1242/dev.179432
393 19 Teran-Cabanillas, E., Montalvo-Corral, M., Caire-Juvera, G., Moya-Camarena, S. Y. & Hernández,
394 J. Decreased interferon-α and interferon-β production in obesity and expression of suppressor of
395 cytokine signaling. *Nutrition* **29**, 207-212 (2013).
396 https://doi.org/https://doi.org/10.1016/j.nut.2012.04.019
397 20 Smith, A. G., Sheridan, P. A., Harp, J. B. & Beck, M. A. Diet-Induced Obese Mice Have Increased
398 Mortality and Altered Immune Responses When Infected with Influenza Virus12. *The Journal of*
399 *Nutrition* **137**, 1236-1243 (2007). https://doi.org/https://doi.org/10.1093/jn/137.5.1236
400 21 Gaur, P. *et al.* Defective interferon amplification and impaired host responses against influenza virus
401 in obese mice. *Obesity (Silver Spring)* **29**, 1272-1278 (2021). https://doi.org/10.1002/oby.23196
402 22 Multhaup, M. L. *et al.* Mouse-human experimental epigenetic analysis unmasks dietary targets and
403 genetic liability for diabetic phenotypes. *Cell Metab* **21**, 138-149 (2015).
404 https://doi.org/10.1016/j.cmet.2014.12.014
405 23 Foulke-Abel, J. *et al.* Human Enteroids as a Model of Upper Small Intestinal Ion Transport

406 Physiology and Pathophysiology. *Gastroenterology* **150**, 638-649.e638 (2016).
407 https://doi.org/10.1053/j.gastro.2015.11.047
408 24 Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma,
409 adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762-1772 (2011).
410 https://doi.org/10.1053/j.gastro.2011.07.050
411 25 Heijmans, J. *et al.* ER Stress Causes Rapid Loss of Intestinal Epithelial Stemness through
412 Activation of the Unfolded Protein Response. *Cell Reports* **3**, 1128-1139 (2013).
413 https://doi.org/https://doi.org/10.1016/j.celrep.2013.02.031
414 26 In, J. *et al.* Enterohemorrhagic Escherichia coli reduce mucus and intermicrovillar bridges in human
415 stem cell-derived colonoids. *Cell Mol Gastroenterol Hepatol* **2**, 48-62.e43 (2016).
416 https://doi.org/10.1016/j.jcmgh.2015.10.001
417 27 Bentsen, M. *et al.* ATAC-seq footprinting unravels kinetics of transcription factor binding during
418 zygotic genome activation. *Nat Commun* **11**, 4267 (2020). https://doi.org/10.1038/s41467-020-
419 18035-1
420

421 **ACKNOWLEDGEMENTS**

422 The authors thank Dr. Mary K. Estes (Baylor College of Medicine) for sharing four jejunal organoid lines
423 (J2, J4, J11, and J12) used in this study. Tissue culture facilities were provided by the Johns Hopkins
424 Digestive Diseases Basic & Translational Research Core Center (National Institutes of Health P30
425 DK089502). RNA-seq was performed by the Genome Analysis Core, Mayo Clinic Rochester. ATAC-seq
426 was conducted by the Epigenomics Development Laboratory, Center for Cell Signaling in
427 Gastroenterology, Mayo Clinic Rochester (National Institutes of Health P30 DK084567). Sequence
428 assembly pipelines, differential gene expression, and ATAC footprinting were carried out by the
429 Bioinformatics Core, Mayo Clinic Rochester. Quantitative PCR was conducted on instrumentation
430 provided by the Johns Hopkins Genetic Resources Core Facility.
431

432 **FUNDING**

433 Mayo Clinic internal funds (DSB, VK)

434 Johns Hopkins Division of Gastroenterology & Hepatology internal funds (JFA)

435 National Institutes of Health K01 DK113043 (JFA)

436

437 **AUTHOR CONTRIBUTIONS**

438 Project conceptualization and design: DSB, ZL, OK, VK, JFA

439 Data acquisition: DSB, ZL, JHL, TM, AVB, RL, AD, JFA

440 Analysis of results: DSB, ZL, TM, AVB, TO, OK, JFA

441 Initial manuscript drafting: DSB, ZL, OK, JFA.

442 Final manuscript editing and approval: All authors

443

444 **ADDITIONAL INFORMATION**

445 **Competing Interests**

446 Dr. Kovbasnjuk is an employee of the National Institutes of Health, USA. The content of this manuscript
447 is solely the responsibility of the author and does not necessarily represent the official views of the
448 National Institutes of Health, USA. All other authors declare no competing interests.

449

450 **FIGURE LEGENDS**

451 **Figure 1. Differential gene expression among the organoid donor groups.** (**A**) Experimental design
452 and donor phenotypes of groups 1, 2, and 3. (**B**) Principal component analysis (PCA) of the nine donors
453 following RNA-seq analysis. Colors distinguish donors in each of the three groups. (**C**) Volcano plot
454 depicting significantly upregulated (red) or downregulated (blue) transcripts ($\log_2$FC > |2|, $p < 0.05$) in
455 paired comparisons of G2/G1, (**D**) G2/G3, and (**E**) G1/G3. (**F**) Heatmap of relative expression for
456 significant DEGs in G2 compared to G1 or G3. Cluster 1 contains DEG that were significantly upregulated
457 in G2, and Cluster 2 contains DEG that were significantly downregulated in G2. Z scores for each gene
458 were computed from $\log_2$(FPKM+1) contained in the respective clusters. (**G**) Gene ontology (GO) terms
459 representing major biological processes characteristic of Clusters 1 and 2. (**H**) Heatmap ($\log_2$FC) for

460 selected transcripts distinctively upregulated in G2 that are associated with dietary sugar
461 transport/metabolism, fat absorption/transport, bile acid absorption, and gut hormones. (**I**) Heatmap of
462 selected transcripts distinctively downregulated in G2 that loosely group as factors involved in innate
463 immune functions.
464
465 **Figure 2. Differential open chromatin regions among the organoid donor groups.** (**A**) Differential
466 open chromatin regions ($\log_2 FC > |2|$, $p < 0.05$) identified from ATAC-seq in G1 vs G2, (**B**) G2 vs G3, and
467 (**C**) G1 vs G3. (**D**) Heatmaps of differential open chromatin proximal to transcription start sites (TSS,
468 within 2 kb upstream or downstream) for each donor. (**E**) Summed ATAC-seq peak tracks for each of the
469 three groups representing the upstream and protein coding region for APOA4, which had increased
470 accessibility in G2. (**F**) Summed peak tracks for OAS2, which had decreased accessibility in G2 relative
471 to the other groups.
472
473 **Figure 3. Differential transcription factor binding events inferred from ATAC footprinting.** (**A**)
474 Volcano plot of differential binding scores in G1 vs G2 and (**B**) G2 vs G3. There were no significant
475 differential binding events in G1 vs G3. (**C**) Top three most and least accessible TF motifs in G2 when
476 compared to G1, ranked by *p*-value.
477
478 **Figure 4. Validation of differentially expressed intestinal epithelial biomarkers in obese**
479 **hyperabsorptive subjects.** (**A**) Selected transcripts measured by quantitative real-time PCR (qRT-PCR)
480 and expressed as the difference in cycle number required for detection of exponential target amplification
481 ($-\Delta C_T$) relative to the average cycle number ($C_{T,\ avg}$) required to detect each target in the lean cohort.
482 Increased expression, red; decreased expression, blue. N = 10 lean (G1) and 10 obese hyperabsorptive
483 (G2) donors that were unique from those analyzed previously by next-generation sequencing. (**B**)
484 Receiver-operator curves (ROC) for each biomarker analyzed by qRT-PCR. Each graph contains the
485 respective area under the curve (AUC) and *p*-value.
486

Table 1. Organoid donor characteristics. Deidentified subject designator (organoid line), group classification based on glucose transport phenotype, demographic information (age, sex), body mass index (BMI), and passage number for each organoid line studied.

Group	Organoid Line	Originating Institution	Absorption Phenotype	Age	Sex	BMI	Organoid passage used in experiments
G1	J1	Johns Hopkins	low	n.a.	F	≤25	P38
G1	J11	Baylor	low	n.a.	F	23	P29
G1	J23	Johns Hopkins	low	46	F	21	P11
G2	J2	Baylor	high	52	F	≥35	P81
G2	J4	Baylor	high	n.a.	F	≥35	P33
G2	J21	Johns Hopkins	high	42	F	43	P10
G3	J12	Baylor	low	n.a.	M	≥35	P28
G3	J17	Johns Hopkins	low	49	F	53	P15
G3	J20	Johns Hopkins	low	48	F	51	P13
G1	1	Johns Hopkins	low	81	F	22.8	P16
G1	23	Johns Hopkins	low	57	M	23	P9
G1	24	Johns Hopkins	low	63	F	28	P7
G1	25	Johns Hopkins	low	63	M	28.2	P8
G1	26	Johns Hopkins	low	65	F	24	P6
G1	27	Johns Hopkins	low	73	F	23.8	P4
G1	28	Johns Hopkins	low	67	F	27.1	P5
G1	0020	Mayo	low	20	M	23.1	P12
G1	0021	Mayo	low	21	F	22.1	P10
G1	0022	Mayo	low	37	M	19.5	P10
G2	0003	Mayo	high	51	F	32.7	P7
G2	0004	Mayo	high	58	M	36.3	P5
G2	0005	Mayo	high	73	F	36.3	P8
G2	0009	Mayo	high	73	F	36.7	P6
G2	0012	Mayo	high	55	F	34.3	P4
G2	0013	Mayo	high	32	F	42.7	P6
G2	0014	Mayo	high	47	F	36.9	P7
G2	0015	Mayo	high	46	F	43.5	P4
G2	0018	Mayo	high	71	F	38.3	P3
G2	0031	Mayo	high	55	F	32.1	P4

Table 2. qRT-PCR oligo sequences.

Gene Symbol	Accession Number	Forward primer (5'-3')	Reverse primer (5'-3')
18S rRNA	NR_003286	GTAACCCGTTGAACCCCATT	CCATCCAATCGGTAGTAGCG
GAPDH	NM_002046	AGGGCTGCTTTTAACTCTGGT	CCCCACTTGATTTTGGAGGGA
APOA4	NM_000482	CAGTGTGGCAAGAAACTCCT	GTAGTCCCACATCACCGTG
CD36	NM_001001548	TGTCCTGGCTGTGTTTGG	TGTCTGTAAACTTCTGTGCCTG

CUBN	NM_001081	TCTCAGCAATCCAAATCAGGG	ACCACAAAGTCTCCACATCAG
G6PC	NM_000151	AGGGAAAGATAAAGCCGACC	AGCAAGGTAGATTCGTGACAG
GIP	NM_004123	GCCAGTCAAGCTAATAGGAAGG	AGAGGTTTGTCTGATCCAGC
GUCA2A	NM_033553	ATGAATGCCTTCCTGCTCTC	CCTGGAGGTCTTTGAGCTTC
KHK	NM_006488	ACTCCAATGGCAACCGTAC	ACCTTCTCAAAGTCTGTAGCAG
SLC2A2	NM_000340	TGCTGTCTCTGTATTCCTTGTG	TGCTCACATAACTCATCCAAGAG
SLC2A5	NM_003039	CCGTGTCCATGTTTCCATTTG	ATCCCATTAAGATCGCAGGC
SLC5A11	NM_052944	CTGTACTTCCTCTTTGTCCTGG	GTCCACTTCCAACATTGCTG

492

Figure 1.

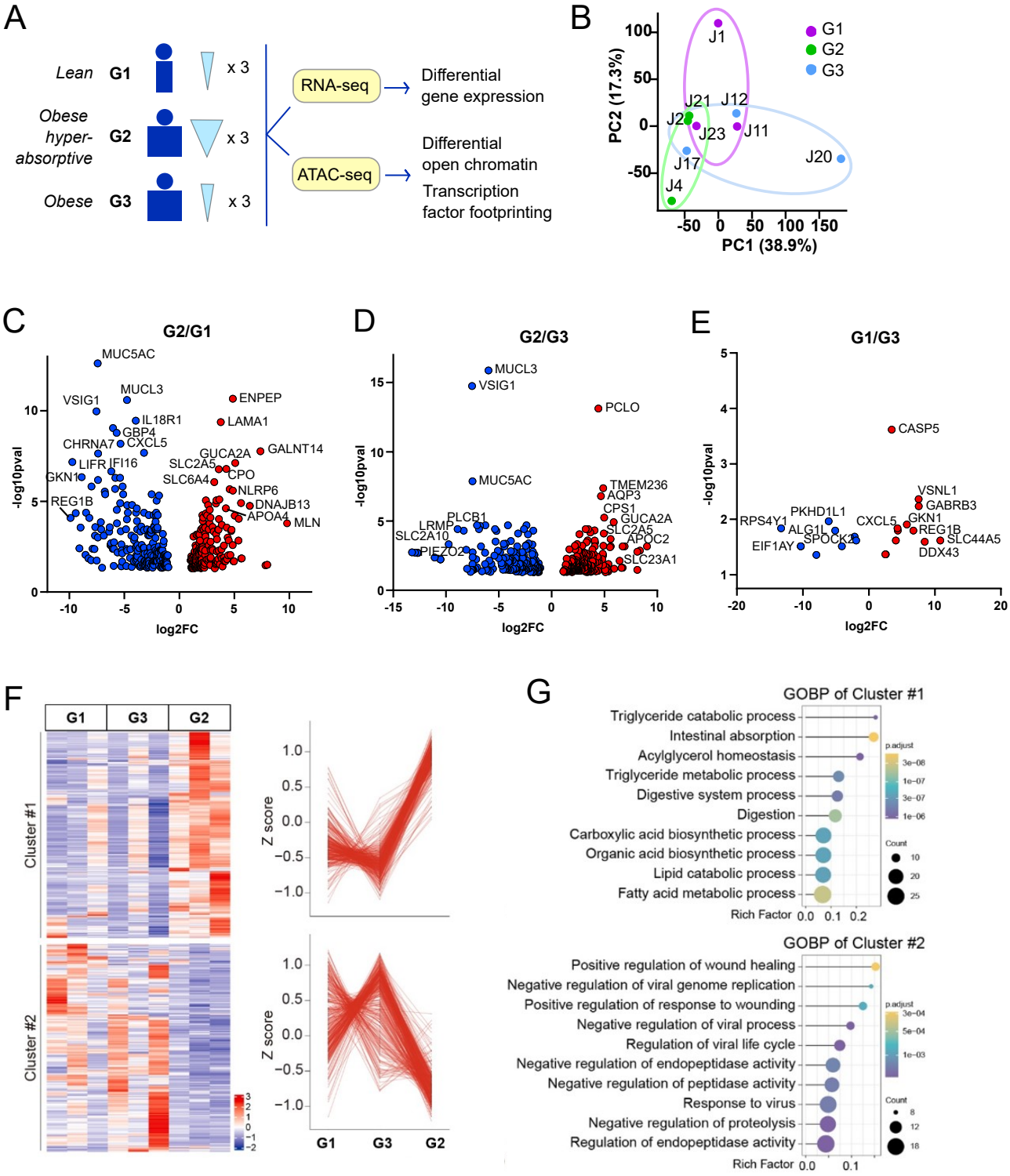
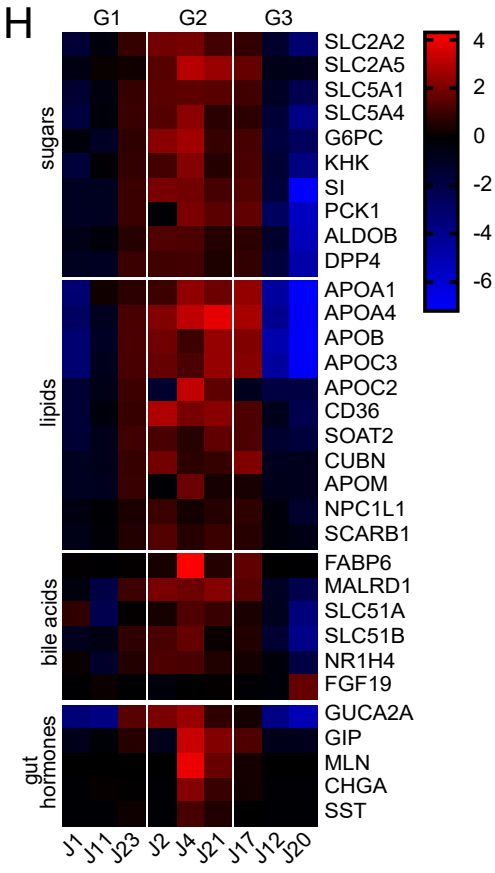

A
Lean G1
Obese hyper-absorptive G2
Obese G3
x 3
x 3
x 3
RNA-seq
ATAC-seq
Differential gene expression
Differential open chromatin
Transcription factor footprinting
B
G1
G2
G3
PC2 (17.3%)
PC1 (38.9%)
J1
J21
J12
J2
J23
J11
J17
J20
J4
C
G2/G1
-log10pval
log2FC
MUC5AC
MUCL3
VSIG1
ENPEP
IL18R1
LAMA1
GBP4
CHRNA7
CXCL5
GALNT14
LIFR
IFI16
SLC2A5
GUCA2A
GKN1
SLC6A4
CPO
NLRP6
REG1B
DNAJB13
APOA4
MLN
D
G2/G3
-log10pval
log2FC
MUCL3
VSIG1
PCLO
MUC5AC
TMEM236
AQP3
CPS1
PLCB1
GUCA2A
LRMP
SLC2A5
SLC2A10
APOC2
PIEZO2
SLC23A1
E
G1/G3
-log10pval
log2FC
CASP5
VSNL1
GABRB3
PKHD1L1
RPS4Y1
CXCL5
GKN1
ALG1L
REG1B
SPOCK2
SLC44A5
EIF1AY
DDX43
F
G1
G3
G2
Cluster #1
Cluster #2
Z score
G1
G3
G2
G
GOBP of Cluster #1
Triglyceride catabolic process
Intestinal absorption
Acylglycerol homeostasis
Triglyceride metabolic process
Digestive system process
Digestion
Carboxylic acid biosynthetic process
Organic acid biosynthetic process
Lipid catabolic process
Fatty acid metabolic process
Rich Factor
p.adjust
Count
GOBP of Cluster #2
Positive regulation of wound healing
Negative regulation of viral genome replication
Positive regulation of response to wounding
Negative regulation of viral process
Regulation of viral life cycle
Negative regulation of endopeptidase activity
Negative regulation of peptidase activity
Response to virus
Negative regulation of proteolysis
Regulation of endopeptidase activity
Rich Factor
p.adjust
Count

Figure 1 (continued)



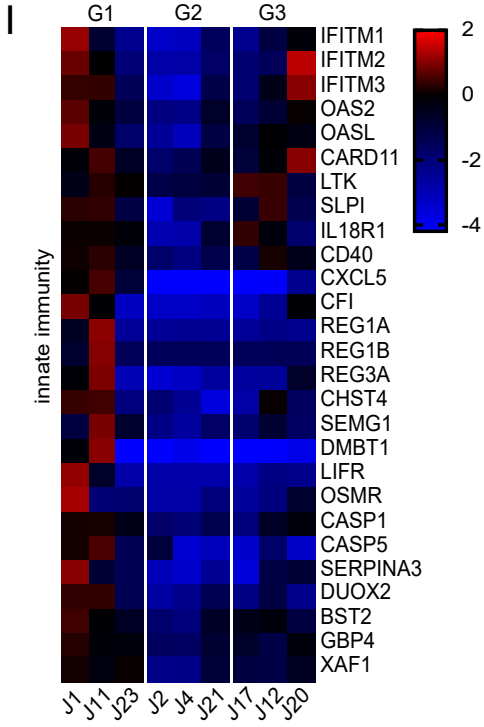
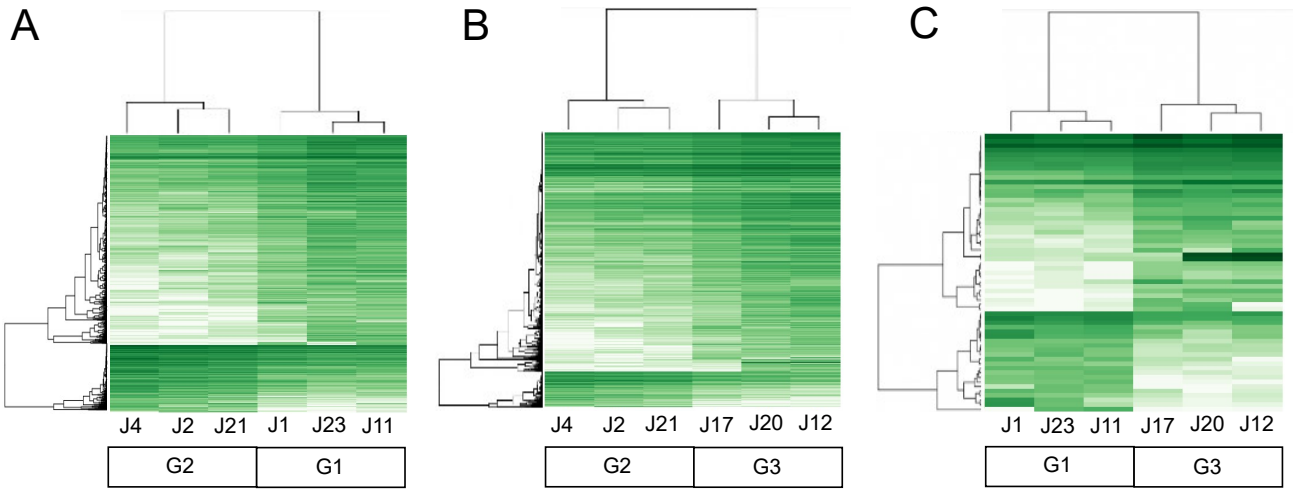
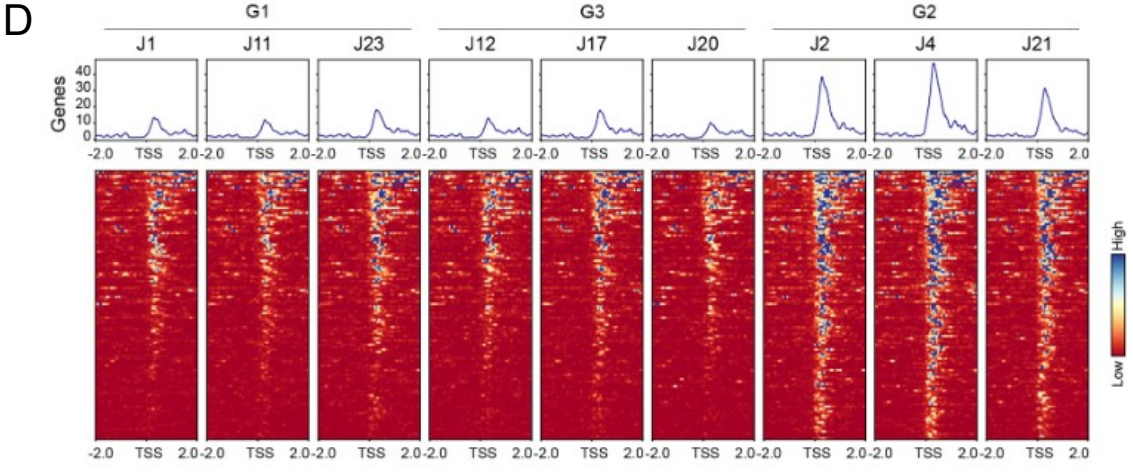
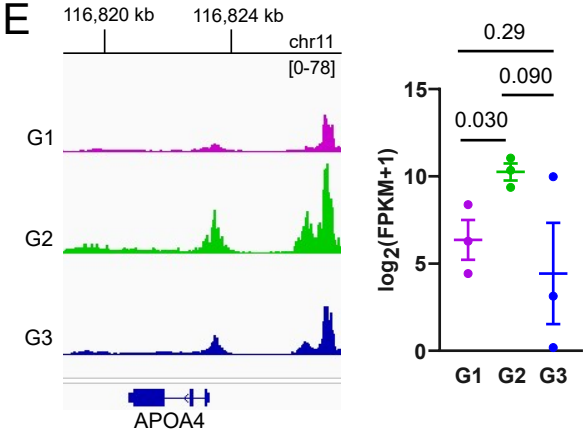


Figure 2.


A
J4 J2 J21 J1 J23 J11
G2 G1
B
J4 J2 J21 J17 J20 J12
G2 G3
C
J1 J23 J11 J17 J20 J12
G1 G3


D
G1
G3
G2
J1 J11 J23 J12 J17 J20 J2 J4 J21
Genes
40 30 20 10 0
-2.0 TSS 2.0
High
Low


E
116,820 kb
116,824 kb
chr11
[0-78]
G1
G2
G3
APOA4
0.29
0.090
0.030
15
10
5
0
log2(FPKM+1)
G1 G2 G3

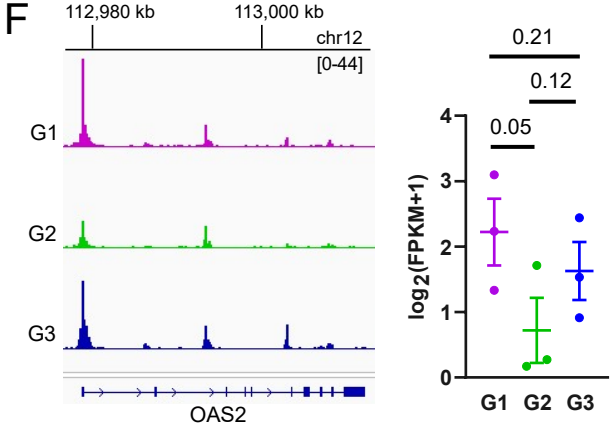
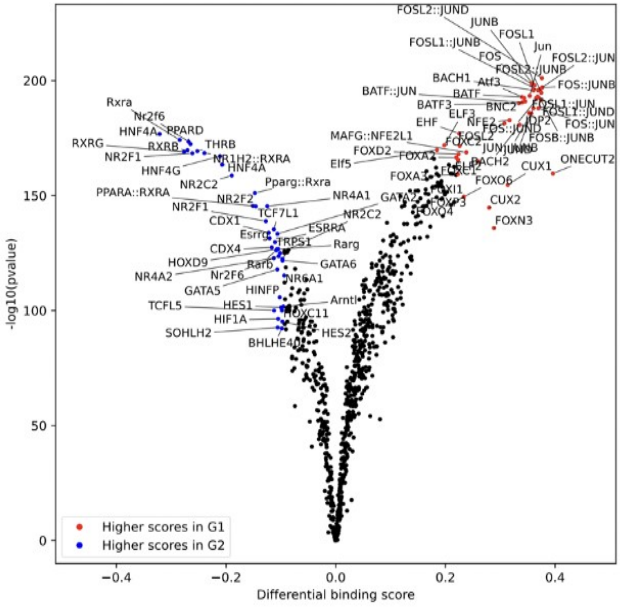
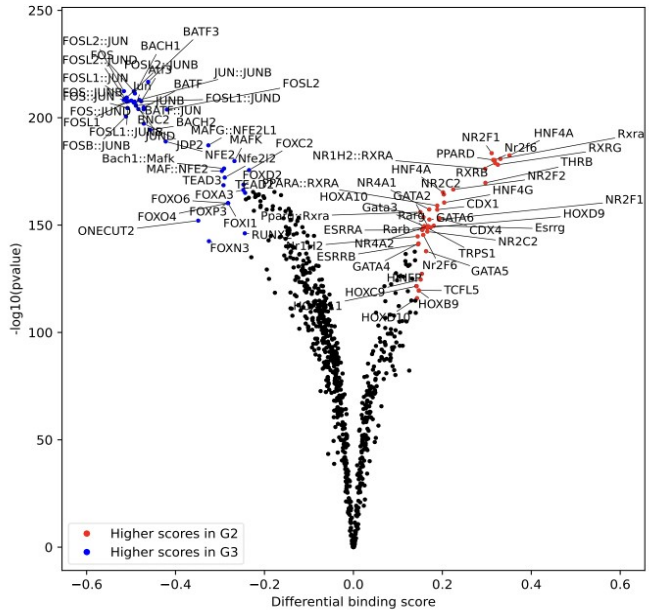

F
112,980 kb
113,000 kb
chr12
[0-44]
G1
G2
G3
OAS2
0.21
0.12
0.05
4
3
2
1
0
log2(FPKM+1)
G1 G2 G3

Figure 3.

A



B



C

Enriched/more accessible in G2

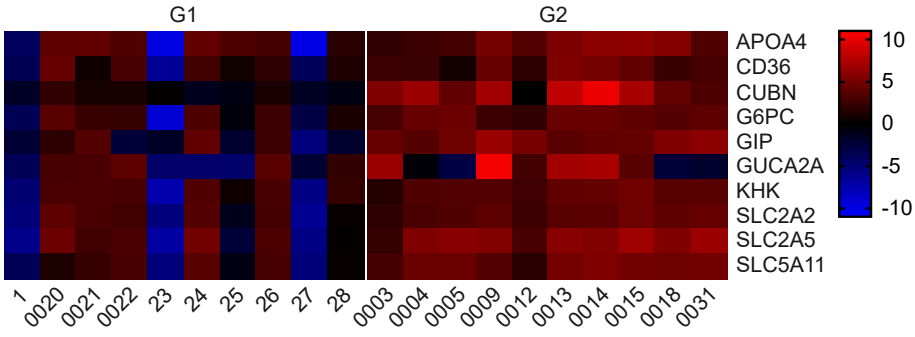
G2 motif	P value
HNF4G	1e-40
CDX2	1e-24
DLX5	1e-14

Less accessible in G2

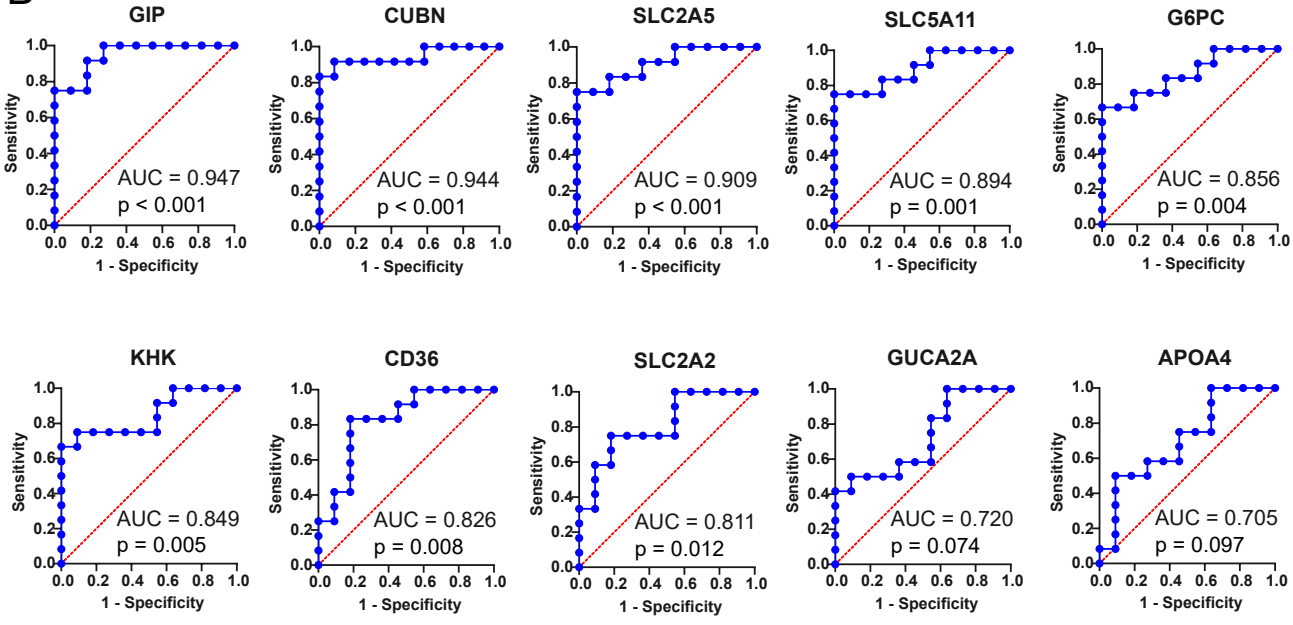
G1 motif	P value
FOSL1	1e-245
ELF3	1e-108
FOXM1	1e-84

Figure 4.

A


G1
G2
APOA4
CD36
CUBN
G6PC
GIP
GUCA2A
KHK
SLC2A2
SLC2A5
SLC5A11
10
5
0
-5
-10
1
0020
0021
0022
23
24
25
26
27
28
0003
0004
0005
0009
0012
0013
0014
0015
0018
0031

B


GIP
AUC = 0.947
p < 0.001
CUBN
AUC = 0.944
p < 0.001
SLC2A5
AUC = 0.909
p < 0.001
SLC5A11
AUC = 0.894
p = 0.001
G6PC
AUC = 0.856
p = 0.004
KHK
AUC = 0.849
p = 0.005
CD36
AUC = 0.826
p = 0.008
SLC2A2
AUC = 0.811
p = 0.012
GUCA2A
AUC = 0.720
p = 0.074
APOA4
AUC = 0.705
p = 0.097
Sensitivity
1 - Specificity