Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Claussnitzer M, Dankel SN, Kim K-H, et al. *FTO* obesity variant circuitry and adipocyte browning in humans. N Engl J Med 2015;373:895-907. DOI: 10.1056/NEJMoa1502214

Supplementary Appendix

Causal FTO Obesity Variant Circuitry and Adipocyte Browning in Humans

Table of Contents

List of Investigators

Melina Claussnitzer, Simon Nitter Dankel, Kyoung-Han Kim, Gerald Quon, Wouter Meuleman, Christine Haugen, Viktoria Glunk, Isabel Sophia Sousa, Jacqueline L. Beaudry, Vijitha Puviindran, Nezar A. Abdennur, Jannel Liu, Per-Arne Svensson, Yi-Hsiang Hsu, Daniel J. Drucker, Gunnar Mellgren, Chi-Chung Hui, Hans Hauner, Manolis Kellis

Methods

Overview

To dissect the mechanistic basis of the genetic association between non-coding genetic variants in the FTO locus and obesity, we took the following steps:

I. Identify the cell type where obesity-associated variants act in humans: (a) predicted adipocyte progenitors using epigenomic annotations across multiple human tissues and cell types, and (b-d) validated using enhancer assays for risk and non-risk alleles.

II. Identify the genes whose expression is altered by the obesity-associated variants in humans: (a) predicted 8 candidates using topologically-associated domains surrounding the FTO obesity locus, and (b) validated *IRX3* and *IRX5* using eQTL analysis in human adipocyte progenitors from risk and non-risk homozygous participants.

III. Recognize the cellular processes affected by the genetic locus and its targets in humans and mice, (a) predicted using genome-wide expression correlation with *IRX3* and *IRX5*, and (b-d) validated using gene expression and cellular phenotypes measured in risk and non-risk participants.

IV. Demonstrate that *IRX3* and *IRX5* control energy balance phenotypes in humans and mice, using (a) direct manipulation of *IRX3* and *IRX5* in risk and non-risk participants, (b) direct manipulation of *Irx3* and *Irx5* in three mouse cellular models; (c) generation of adipose dominant negative *Irx3* mice followed by organismal and cellular phenotypes.

V. Distinguish the causal nucleotide variant among 89 candidate variants, (a) predicted using evolutionary conservation scores for regulatory motif combinations, and (b-c) validated using single-nucleotide alterations in human enhancer reporter assays and electrophoretic mobility shift assays.

VI. Identify the upstream regulator whose activity is altered by the causal variant, (a-b) predicted ARID5B using regulatory motif matches and high TF expression levels in human adipocyte progenitors and adipose tissue, and (c-d) validated using combined

ARID5B and rs1421085 manipulations in enhancer assays, regulator binding competition experiments via EMSA; and ARID5B manipulations in risk and non-risk participants followed by *IRX3* and *IRX5* expression.

VII. Evaluate variant causality in its genomic context by single-nucleotide editing, (a) using CRISPR/Cas9 in primary human adipocytes from risk and non-risk participants, and evaluating (b) gene expression changes in *IRX3* and *IRX5* after bi-directional editing and (c-d) restoration of adipocyte browning signatures by C-to-T editing in the risk background, including restoration of mitochondrial expression programs, decreased lipid storage markers, and restoration of both basal respiration and stimulated mitochondrial thermogenesis.

We describe each of these steps in detail next, in the order presented in our manuscript.

I: Identify the cell type of action of obesity-associated variants.

Our first goal in dissecting the FTO non-coding association with obesity (Fig. 1A) was to identify the cell type in which the obesity-associated variants may act.

a. Computational prediction of cell type of action using epigenomic annotations.

Rationale. Unlike protein-coding genes, where gene expression levels can be used to predict where a gene is most likely to act, for non-coding regions, we used epigenomic annotations associated with regulatory region activity to predict in an unbiased way the cell type in which the genetic variants may act.

Chromatin state segmentations. We used publicly-available datasets of chromatin state annotations in 127 reference epigenomes (Fig. 1B, S1A) from the Roadmap Epigenomics[1] and ENCODE[2] projects. These include a large number of primary cells and tissue types, enabling an unbiased approach for predicting relevant cell types (Fig. S1B). These are available for download at http://compbio.mit.edu/roadmap/ and can be visualized in the WashU Epigenome Browser.[3]

We used annotations of chromatin states[2] based on combinations of histone modification marks, using a 25-state model predicted by ChromHMM[4], using 12 histone modification marks imputed by ChromImpute[5].

Enhancer annotations. To identify putative distal regulatory elements, we focused on enhancer states (states 9-18), defined by high levels of enhancer-associated marks (H3K4me1 and H3k27ac) and lower levels of promoter-associated marks (H3K4me3, H3K4me2, H3K9ac). To recognize master regulatory loci, we combined consecutive enhancer elements into clusters by joining pairs of elements that were 200bp apart or less (one quarter of the

median length of enhancers), and evaluated total cluster length (Fig. S1C). We included transcribed enhancers (states 9-12) that also show transcription-associated marks (H3K36me3, H4K20me1, H3K79me2), active enhancers (states 13-15) that show relatively-stronger H3K27ac, and weak enhancers (states 16-18) that show relatively weaker H3K27ac.

b. Experimental validation of enhancer activity in a human adipocyte cell line.

Rationale. We next sought to validate experimentally that the region showing epigenomic signatures associated with enhancer function in adipocyte cell types from the Roadmap Epigenomics project[1] (Adipose-Derived Mesenchymal Stem Cells, E025) does indeed function as an enhancer in human adipocytes (Fig. S1D), as predicted by its epigenomic signatures.

Luciferase-based enhancer reporter assays. To examine enhancer activity we performed luciferase reporter gene assays in human SGBS (Simpson–Golabi–Behmel Syndrome) adipocytes. Specifically, we evaluated the activity of candidate regulatory elements by cloning them upstream of a luciferase gene, transferring the resulting reporter construct into cell lines, and measuring expression of the luciferase reporter gene to quantify the activity of the enhancer.

Cloning of reporter constructs. For the promoter construct, we cloned a 752 bp thymidine kinase (TK) promoter upstream of the firefly luciferase gene into the *EcoRV* and *BglII* sites of the pGL4.22 firefly luciferase reporter vector (Promega). We subcloned enhancer regions upstream of the TK promoter into the *KpnI* and *SacI* sites of the pGL4.22-TK vector in forward orientations.

SGBS adipocyte cell culturing. We cultured the human preadipocyte SGBS cell line as previously described[6] in DMEM/Ham's F12 (1:1) medium (supplemented with 10% FCS, 17 µM biotin, 33 µM pantothenic acid and 1% penicillin/streptomycin). To promote adipose differentiation of the SGBS cell line, we grew cells to confluence. For induction of adipocyte differentiation we cultured in serum free MCDB-131/DMEM/Ham's F12 (1:2) medium supplemented with 11 µM biotin, 22 µM pantothenic acid, 1% penicillin/streptomycin, 10 µg/ml human transferrin, 66 nM insulin, 100 nM cortisol, 1nM triiodothyronine, 20 nM dexamethasone, 500 µM 3-isobutyl-1-methyl-xanthine (Serva, Germany) and 2 µM rosiglitazone (Alexis, Germany). We maintained cells at 37°C and 5% CO_2.

Transfection in SGBS adipocytes. We transfected SGBS adipocytes (12-well plate, 8 x 10^4 / well) at day four after the induction of differentiation (80% confluence) with 2 µg of the respective pGL4.22-TK construct and 2 µl Lipofectamine reagent. We co-transfected the firefly luciferase constructs with the ubiquitin promoter-driven Renilla luciferase reporter vector pRL-Ubi to normalize the transfection efficiency. Twenty-four hours after transfection, we washed the cells with PBS and lysed in 1x passive lysis buffer (Promega, Germany) on a rocking platform for 30 min at room temperature.

Measurement of luciferase activity. We measured firefly and Renilla luciferase activity (substrates D-luciferin and Coelenterazine from PJK, Germany) using a Luminoscan Ascent microplate luminometer (Thermo) and a Sirius tube luminometer (Berthold). We calculated the ratios of firefly luciferase expression to Renilla luciferase expression and normalized to the TK promoter control vector, i.e. enhancer activity.

c. Validation of genetic control and narrowing down candidate region using enhancer assays in risk and non-risk haplotypes.

Rationale. Having established that the region underlying the enhancer signatures can function as an enhancer, we next sought to establish genetic control by the FTO obesity-associated variants, by separately testing the risk and the non-risk haplotypes.

Enhancer Tiling Assays. Given the large size of the obesity-associated region, we used tiling assays to narrow down the region of association. We tiled the 47kb *FTO* obesity risk locus in five 10kb segments, and synthesized each as a plasmid vector (Life Technologies). The five 10kb tiles were: segment 1: chr16:53799507-53809507; segment 2: chr16:53809507-53819507; segment 3: chr16:53819507-53829507; segment 4: chr16:53829507-53839507; segment 5: chr16:53839507-53849507. We also tested a 1kb interval centered on the rs1421085 variant (chr16: 53,800,454-53,801,454).

Comparison of risk-allele and protective-allele enhancer constructs. For each 10kb intervals (Fig. 1C), we synthesized both risk (obesity-associated) and non-risk (protective) haplotypes using HapMap individuals information. For the 1kb interval centered on rs1421085 (Fig. S1D), we used site-directed mutagenesis (Quick Change II Site-Directed Mutagenesis Kit, Stratagene) to introduce the C risk allele by T-to-C substitution in the non-risk T background haplotype. We verified the identity of each construct clone by DNA sequencing. We cloned genomic DNA segments upstream of the TK promoter into the *KpnI* and *SacI* sites of the pGL4.22-TK vector in forward orientation. We performed transfection in human SGBS adipocytes as described above (section Ib). We used a paired t-test to compute P-values comparing luciferase expression from risk and non-risk alleles. We transfected human SGBS adipocytes as described above (section Ib).

d. Validation of adipocyte specificity by enhancer assays in diverse cell types.

Rationale. Given the cell type specific nature of epigenomic signatures of enhancer activity, we predicted that the observed enhancer would function in a cell type specific way. In order to validate this prediction, we evaluated the 1kb enhancer reporter assay constructs (Fig. S1D) in multiple cell types using both the risk and the non-risk alleles of rs1421085.

Cell types tested. To examine cell type specific enhancer activity for the risk and the protective allele we performed enhancer assays in multiple cell lines, including human SGBS (Simpson–Golabi–Behmel Syndrome) cells, human Huh7 hepatoma, mouse C2C12 myoblasts, HT22 neuronal cells and human embryonic kidney K562 cells.

Cell culturing. The human Huh7 hepatoma, mouse C2C12 myoblasts, HT22 neuronal cells and human embryonic kidney K562 cells were cultured in DMEM medium (supplemented with penicillin/streptomycin and 10 % FBS). The human preadipocyte SGBS (Simpson–Golabi–Behmel Syndrome) cell line was cultured as described above (section Ib). We maintained all cells at 37°C and 5% CO_2.

Cloning of luciferase reporter gene constructs. We cloned the promoter construct, a 752 bp thymidine kinase (TK) promoter upstream of the firefly luciferase gene as described above (Section Ib). Enhancer regions were cloned upstream of the TK promoter into the *KpnI* and *SacI* sites of the pGL4.22-TK vector in forward orientations.

Transfection in multiple cell types. We transfected Huh7 cells (96-well plate, 1.1 x 10^4 / well) one day after plating with approximately 90% confluence, K562 cells (12-well plate, 8 x 10^4 / well) three days after plating with approximately 90% confluence, SGBS adipocytes (12-well plate, 8 x 10^4 / well) at day eight after the induction of differentiation with approximately 80% confluence and C2C12 cells (12-well plate, 2 x 10^5 / well) at day four after induction of differentiation with approximately 90% confluence. We transfected Huh7 with 0.5 µg of the respective firefly luciferase reporter vector and 1 µl Lipofectamine 2000 transfection reagent (Invitrogen, Germany), differentiated C2C12 myocytes with 1 µg of the respective pGL4.22-TK construct and 2 µl Lipofectamine reagent, and both K562-cells and differentiated SGBS adipocytes with 2 µg of the respective pGL4.22-TK construct and 2 µl Lipofectamine reagent. We co-transfected the firefly luciferase constructs with the ubiquitin promoter-driven Renilla luciferase reporter vector pRL-Ubi to normalize the transfection efficiency. Twenty-four hours after transfection, we washed the cells with PBS and lysed in 1x passive lysis buffer (Promega, Germany) on a rocking platform for 30 min at room temperature.

Luciferase measurements. We measured Firefly and Renilla luciferase activity (substrates D-luciferin and Coelenterazine from PJK) using a Luminoscan Ascent microplate luminometer (Thermo) and a Sirius tube luminometer (Berthold). We calculated the ratios of firefly lucifer-

ase expression to Renilla luciferase expression and normalized to the TK promoter control vector, i.e. enhancer activity. We compared luciferase expression between risk and non-risk alleles using the P-value of a paired t-test.

II: Identifying the target genes of the obesity-associated variants.

As genetic variants can act at potentially large distances, we next sought to identify the genes that are targeted by the regulatory region harboring the *FTO* association with obesity, whose expression may be affected by the risk haplotype.

a. Predict putative target genes using long-range chromosome interactions.

Rationale. To predict putative target genes we first examined topologically-associated domains surrounding the *FTO* locus (Fig. 1D) using chromosome conformation capture data (Hi-C)[7], representing physical interactions between genomically-distal chromosomal regions that are proximal when folded in three dimensions within the nucleus (Fig. 1E), reasoning that genes within topologically-associated domains[8] may also be genetic targets.

Hi-C data processing and visualization. We obtained Hi-C data for human IMR90 myofibroblastic cells[9] study from GEO (accession number: GSE35156), mapped them to human genome assembly hg19, and processed them as described in Ref. [10]. We visualized interaction heatmap patterns using the matplotlib package for Python. We visualized region chr16:52800000-56000000, binned at 40kb resolution (Fig. 1E, top right). We also computed pairwise LD r^2 values in Europeans derived from the 1000 Genomes Phase 1 call set and plotted the mean r^2 values between variants in each pair of 40kb bins (Fig. 1E, bottom left).

Prediction of candidate target genes. We considered all genes in the two topologically associated domains surrounding the FTO genetic locus as potential candidates. This resulted in a total of eight genes, for which we undertook eQTL analysis as we describe next.

b. eQTL analysis of putative target genes to identify targets under genetic control.

Rationale. To identify the genetic target genes whose expression is altered by the obesity-associated variants, we performed targeted expression quantitative trail locus (eQTL) analysis for all potential target genes (Fig. 1F) using primary human adipocytes using patient samples from risk and protective allele carriers.

Primary subcutaneous adipose tissue isolation. We obtained isolated cells from subcutaneous adipose tissue excisions for 38 healthy European persons 20 to 50 years of age and with a normal body-mass index (BMI) (20 to 24 kg/m2). We obtained written informed consent from each participant and approval by the local ethics committee of the Faculty of Medicine

of the Technical University of Munich, Germany, or the Regional Committee for Medical Research Ethics (REK) of Haukeland University Hospital, Bergen, Norway.

Selection of homozygous risk and the non-risk participants. The 38 participants were selected to consist of 18 homozygous carriers of the protective genotype, and 20 homozygous carriers of the risk genotype. No heterozygous participants were selected. We performed genotyping for both the GWAS reported tag SNP and the predicted causal variant rs1421085.

Cell culturing and differentiation. For each participant, we isolated primary mesenchymal adipocyte progenitors as previously described[11] with some modifications. Briefly, after expansion and freezing, we cultured the cells in DMEM/F12 (1:1) medium (supplemented with 10% FCS and 1% penicillin/streptomycin) for 18 h, followed by expansion in DMEM/F12 medium (supplemented with 2.5% FCS, 1% penicillin/streptomycin), 33µM biotin, 17µM pantothenic acid, 132nM insulin, 10ng/ml EGF, and 1ng/ml FGF until confluence. We induced adipogenic differentiation for two days by supplementing with 66nM insulin, 100nM cortisol, 10µg/ml transferrin, 1nM triiodo-L-thyronin (T3), 2µM rosiglitazone, 25nM dexamethasone and 0.5mM IBMX.

Sample preparation and qPCR gene expression measurements. We extracted total RNA with TRIzol (Invitrogen). We synthesized copy DNA (cDNA) with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) or M-MLV reverse transcriptase (Invitrogen) with oligo(dT). We performed qPCR using SYBR Green with 60ºC annealing temperature. We calculated relative gene expression by the delta delta Ct method. We normalized target gene expression to expression of *HPRT*[12] or *TBP*.

Whole-adipose tissue qPCR. We repeated the analysis for IRX3 and IRX5 using whole-subcutaneous adipose tissue from the same participants (Fig. S2A), in order to recognize whether the effect found in preadipocytes is also visible at the whole-tissue level for the same individuals. The presence of a strong eQTL signal at the preadipocyte level and absence of an eQTL signal at the whole-tissue level for the same set of individuals indicates that the genetic locus likely functions in a cell type specific way, and that the majority of cells represented in whole-adipose tissue (including mature adipocytes, connective tissue, blood vessels, accessory cells such as monocytes and macrophages, T cells) do not show expression differences associated with the obesity-associated variants for IRX3 and IRX5.

III: Recognize cellular processes affected by obesity-associated variants in humans.

Having identified two master regulators (IRX3 and IRX5) as genetic targets of the obesity-associated variants in preadipocytes, we used correlation with their expression patterns to predict the cellular processes affected by the obesity-associated variants, which we validate by molecular and cellular phenotyping in risk and non-risk individuals.

a. Predict target cellular processes using expression correlation across participants.

Rationale. To predict the cellular processes affected by obesity-associated variants, we used an unbiased approach based on genome-wide correlations in gene expression patterns. We measured expression in adipose tissue from a separate cohort of 10 non-genotyped participants, and identified genes showing strong positive or strong negative correlation with IRX3 and IRX5 in their expression patterns.

Participants and genome-wide expression analysis. We collected perirenal white adipose tissue containing brown adipocytes from a separate cohort of 10 healthy kidney donors that were not previously genotyped and for which obesity status was not a selection criterion. We obtained the tissue laparoscopically under the incision plane created in the renal fascia (Gerotas fascia) proximal to the renal vein. We extracted total RNA with RNeasy Lipid Tissue Kit (Qiagen). We measured gene expression using Affymetrix Gene 1.0 ST microarrays (Affymetrix, Santa Clara, CA), as previously described.[13]

Gene ontology enrichment in positively- and negatively-correlated genes. We ranked all genes by their absolute correlation with IRX3 and also with IRX5 (Pearson coefficient), and measured overrepresentation of gene function categories for the 500 most-correlated genes (Supplementary Table 1). We selected the two Gene Ontology terms (Fig. 2A) most enriched in the positively-correlated set with *IRX5* (FXR/RXR activation) and most negatively-correlated set with both *IRX3* and *IRX5* (mitochondrial function), to predict roles of the obesity-associated variants in switching from energy dissipation (mitochondrial activity) to energy storage (lipid storage). Overrepresented categories were determined with the PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (pantherdb.org), based on binomial statistics with Bonferroni-correction for multiple testing.

b. Provide expression support for roles in adipocyte browning and thermogenesis.

Rationale. Given the correlation of IRX3 and IRX5 gene expression with energy balance genes, we reasoned that they may show expression differences in human white and brown fat, possibly correlated with thermogenesis regulator UCP1. Specifically, given their positive correlation with lipid storage and negative correlation with mitochondrial activity, we expected

IRX3 and IRX5 would show higher expression in white fat tissue, lower expression in brown fat tissue and negative correlation with thermogenesis regulator UCP1.

Evaluate differential IRX3 and IRX5 expression between brown and white fat tissue. To evaluate differential IRX3 and IRX5 expression levels brown and white fat, we analyzed 9 brown (perithyroid) and 9 paired subcutaneous white adipose tissue biopsies from healthy participants using Affymetrix microarrays (IRX3 probe set 229638_at, IRX5 probe set 210239_at) (Fig. S2B).

Evaluate expression correlation of IRX3 and IRX5 with thermogenesis regulators UCP1 and PGC1A. Using Affymetrix microarray data we also analyzed IRX3 and IRX5 co-expressed genes UCP1 and PGC1a in perirenal white adipose tissue containing brown adipocytes obtained from healthy kidney donors (n=10) (Fig. S2C).

c. Validate genetic control of expression for energy balance genes in adipocytes.

Rationale. To establish that the pathways positively- and negatively-correlated with IRX3 and IRX5 across participants are indeed under the genetic control of the obesity-associated variants at the FTO locus, we used qPCR to evaluate the expression of key marker genes of mitochondrial function and lipid storage in risk and non-risk homozygous participants.

Targeted qPCR of energy balance regulators in risk and non-risk participants. We performed isolation, culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb). We extracted RNA using TRIzol (Invitrogen) from primary human adipose progenitors at day 2 of differentiation from 18 homozygous non-risk and 20 homozygous risk allele carriers. We synthesized cDNA with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). We performed qPCR using SYBR Green with 60ºC annealing temperature and calculated relative gene expression by the delta delta Ct method. We normalized target gene expression to expression of *HPRT*[12] or *TBP* (human).

Trans-eQTL analysis of targeted genes. For each targeted gene, we calculated a Mann–Whitney U test P-value to detect statistically-significant differences in the expression of risk and non-risk participants. We found significant differences ($P<0.05$) for all targeted genes (Fig. S2D-E), with risk participants showing significantly increased expression of lipid storage genes, and significantly increased expression of mitochondrial function and lipid catabolism genes. This indicates that these distally-located genes are trans-acting expression quantitative locus (trans-eQTL) targets of the obesity-associated genetic variants, indicating genetic control of energy balance genes by the *FTO* locus in human adipocytes.

To verify whether genes in the predicted energy balance pathways are under genetic control by the obesity-associated variants, we evaluated the regulation of UCP1 gene expression under thermogenic stimulatory conditions (Fig. S2F). For stimulation experiments, we stimu-

lated cells with isoproterenol (1 µmol/l) for 12 hours or exposed to 30 °C for 6 hours before mRNA was harvested.[14]

d. Validate genetic control of browning-associated cellular phenotypes in humans.

Rationale. The reduced participants at risk showed reduced expression mitochondrial, browning and respiration genes, and higher expression for lipid storage markers, indicating a shift from energy dissipation to storage, we next sought to test if those expression differences also reflect in cellular signatures of obesity.

Triglyceride accumulation in primary human adipocytes. To assess the extent of triglyceride accumulation in adipocytes from risk compared to protective allele participants we measured cell size in isolated differentiated adipocytes (Fig. 2B). For preparation of mature adipocytes and cell size determination, we minced subcutaneous adipose tissue from rs1421085 CC risk (n=16) and TT non-risk (n=26) allele carriers and digested in Krebs-Ringer-Phosphate buffer (KRP; 154 mm NaCl, 100 mm NaH_2PO_4, 154 mm KCl, 154 mm $MgSO_4$, 110 mm $CaCl_2$, pH 7.4) containing 100 U/ml collagenase and 4% BSA for 60 min at 37 C in a shaking water bath. After this step, we removed the undigested tissue by filtration through a nylon mesh with a pore size of 250 µm (VWR, Darmstadt, Germany). We washed the floating adipocytes three times with KRP containing 0.1% BSA and used to assess the mean cell diameter based on the measurement of 100 cells from each fraction under the microscope. We did genotyping by MassARRAY (Sequenom), Omni express (Illumina) or Sanger Sequencing.

Mitochondrial biosynthesis in primary human adipocytes. To examine an association of the rs1421085 variant with mitochondrial mass, as a surrogate measure of cellular respiratory capacity, we compared mitochondrial DNA (mtDNA) content in primary human adipose cell cultures of CC risk and TT non-risk allele carriers (n=8 per group) (Fig. 2C). We quantified the relative amount of mtDNA by comparison of the mitochondrial MTCO2 gene with a nuclear target (18S) by qPCR. mtDNA content given as ratio between mtDNA (MTCO2) and 18S. We performed isolation, culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb).

Oxygen Consumption Rate (OCR) in primary human adipocytes. To directly assess effects of the rs1421085 on mitochondrial function, we performed oxygen consumption rate (OCR) measurements under basal and isoproterenol-stimulated conditions (Fig. S2G). We seeded genotyped adipocyte progenitor cells isolated from BMI-matched persons (21 rs1421085 CC risk and 21 TT non-risk) at 70% confluence and induced to differentiate (6 x 10^4 cells / well) the next day. On day 2 of differentiation, we exchanged the medium by XF Assay Medium (1g/L glucose, 2mM L-glutamine, 2mM Na pyruvate) and 2% fatty acid-free BSA before 2 h incubation at 37ºC without CO_2. We treated cells with isoproterenol (1 µmol/l) or DMSO (ba-

sal) for 4 hours. We measured total OCR using the Seahorse XF24 (Seahorse Bioscience) and subtracted the minimum OCR level after rotenone/antimycin A treatment (5 μM) from the initial level without treatment, according to the manufacturer's protocol. Furthermore, to more directly assess effects on mitochondrial thermogenesis, we calculated uncoupled respiration (proton leak) by subtracting the minimum OCR level after rotenone/antimycin from the minimum level after oligomycin treatment (5 μM). We ran the protocol with 3 min mixing and 2 min measuring. We performed isolation culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb). The OCR method was re-used in Section IIId and VIc.

Conditional oxygen consumption rate analyses in primary human adipocytes. To establish causality we used *ARID5B* knock-down in primary human adipocytes of risk and non-risk participants and measured regulation of oxygen consumption rate conditional on the presence of ARID5B by qPCR (Fig. S6K). We performed siRNA-mediated ARID5B knock-down with ON-TARGET-plus SMARTpool siRNA (Dharmacon) and HiPerFect (Qiagen) and overexpression with ARID5B cDNA derived from SGBS total cDNA, which we inserted into the doxycycline-inducible Tet-on system (Tet-On® Advanced Inducible Gene Expression System, BD Biosciences, Clontech, San Diego, CA, USA). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis of *IRX3* and *IRX5* as described above (Section IIb).

Lipolysis rate in primary human adipocytes. Since increased release of glycerol and free fatty acids from adipose cells through lipolysis may indicate reduced lipid storage (an anti-obesity phenotype), we also measured glycerol in the medium after an 18-h incubation comparing primary human adipose cell cultures of risk (n=20) and non-risk (n=18) allele carriers (Fig. S6L). We measured glycerol spectrophotometrically using a glycerol 3-phosphate oxidase trinder kit (Sigma). For stimulated lipolysis measurements, we added 1 μmol/l isoproterenol (Sigma) for 1 hour. We performed isolation, culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb).

IV: Demonstrate causal roles for IRX3 and IRX5 in adipocyte thermogenesis and obesity in humans and mice.

Having established that the obesity-associated genetic variants in the FTO locus are also associated with molecular and cellular differences in mitochondrial function and lipid storage, we next sought to establish that the altered expression of target genes *IRX3* and *IRX5* mediates these phenotypes, by direct manipulation of *IRX3* and *IRX5* in humans and mice, followed by cellular and organismal phenotyping.

a. Validation of molecular and cellular phenotypes of thermogenesis in humans.

Rationale. To examine causal roles for *IRX3* and *IRX5* on energy balance regulation in human adipocytes, we quantified the effect of their manipulation on downstream target gene expression and thermogenesis in primary human adipocytes from both risk and non-risk participants.

IRX3 and ***IRX5*** perturbations. To establish epistasis we used *IRX3* and *IRX5* knock-down and overexpression in primary human adipocytes of homozygous risk (n=8) and homozygous non-risk (n=10) participants (Fig. 2D, Fig. S3A-C). siRNA-mediated *IRX3* and *IRX5* knock-down was performed with ON-TARGET-plus SMARTpool siRNA (Dharmacon) or siRNA from Life Technologies for IRX3/5 and HiPerFect (Qiagen). For overexpression of IRX3 and IRX5 we inserted cDNAs derived from human SGBS total cDNA into the doxycycline-inducible Tet-on system (Tet-On® Advanced Inducible Gene Expression System, BD Biosciences, Clontech, San Diego, CA, USA). We performed isolation, culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb).

qPCR validation of causal roles of ***IRX3*** and ***IRX5*** in humans. Having established the genetic control of mitochondrial function and lipid accumulation of the *FTO* association with obesity, we sought to validate the causal role of IRX3 and IRX5 in target gene regulation (Fig. 3A-C). We carried out siRNA mediated IRX3 and IRX5 knockdown and IRX3 and IRX5 overexpression in our patient samples (n=18 homozygous non-risk, n=20 homozygous risk) and measured regulation of downstream targets by qPCR. We performed isolation, culturing, and differentiation of primary human adipose progenitor cells as described above (section IIb).

Oxygen Consumption Rate (OCR) after IRX3 and IRX5 perturbations. To show that IRX3 and IRX5 levels recapitulate the impact of *FTO* genetic variant on thermogenesis and adipocyte browning, we performed OCR measurements (Fig. 2D). We first carried out siRNA mediated IRX3 and IRX5 knockdown and IRX3 and IRX5 overexpression in our patient samples and measured both basal and isoproterenol-stimulated respiratory rate of primary human adipocytes from risk and non-risk participants (n=8 homozygous non-risk, n=10 homozygous risk). We treated cells with isoproterenol (1 µmol/l) or DMSO (basal) for 4 hours. We measured total OCR using the Seahorse XF24 (Seahorse Bioscience) and subtracted the minimum OCR level after rotenone/antimycin A treatment (5 µM) from the initial level without treatment. Isolation, culturing, and differentiation of primary human adipose progenitor cells and OCR assays is described in more detail in Section IIb and IIIc, respectively.

b. Anti-obesity effects of Irx3 adipose dominant negative in mice.

Having established the molecular and cellular phenotypic roles of IRX3 and IRX5, we sought to examine the effects of adipose repression of Irx3 at the organism level (Fig. S4). We used adipose dominant negative Irx3 mice (aP2-Irx3DN).

Generation of mice. All animal studies were approved by Animal Care Committee of the Toronto Centre of Phenogenomics and conformed to the standards of the Canadian Council on Animal Care. We used *Rosa26*$^{EnR\text{-}Irx3}$ conditional transgenic mice (Irx3EnR)[8], aP2-Cre mice[15], and Irx3/5 double knockout mice as described previously.[16] For generating adipose aP2-Irx3DN, we crossed aP2-Cre male mice with *Rosa26*$^{EnR\text{-}Irx3}$ homozygous female mice. We used *Rosa26*$^{EnR\text{-}Irx3}$ heterozygous mice as a control group to compare with Adipo-Irx3DN (aP2-Cre; *Rosa26*$^{EnR\text{-}Irx3}$) mice. We genotyped mice by PCR assay using following sets of primers. The primer set for *EnR-Irx3* transgene consists of ROSA-FP (5'-AAAGTCGCTCTGAGTTGTTAT-3'), ROSA-WT-RP (5'-GGAGCGGGAGAAATGGATATG-3') and ROSA-MUT-RP (5'-GCGAAGAGTTTGTCCTCAACC-3'). The expected size of the specific PCR products are 600 base pair (bp) for wild-type and 350 bp for EnR-Irx3 transgene. The primer set for *Cre* gene are forward (5'-ATCCGAAAAGAAAACGTTGA-3') and reverse (5'-ATCCAGGTTACGGATATAGT-3'). The expected size of the PCR product is 600 bp. We used only male mice in this study. We maintained mice on 12-hour light/dark cycles and provided with food (Harlan #2918) and water *ad libitum*. For diet-induced obesity studies, we subjected 8 week old male mice to 45% high-fat diet (Research Diets) for 10 weeks. We measured body weight every week from 3 to 18 weeks of age (Fig. **S4A**-F,**J**-K,**P**).

Histological analysis. After assessment of energy metabolism, we sacrificed mice to harvest tissues for organ weight measurement, gene expression analysis and histological analysis (Fig. S4G). For histological analysis, we fixed tissues in 4% paraformaldehyde and embedded in paraffin. We stained 5 μm sections with hematoxylin and eosin (H&E).

Metabolic phenotyping. To examine body composition by measuring fat and lean mass as well as free water, we subjected 18 week old mice to a EchoMRI device (Echo Medical Systems). We also utilized an indirect calorimetry (Oxymax System, Columbus Instruments) to measure energy metabolism of the mice. Each mouse was added to the airtight chamber and provided with about 20g of chow and water bottle. From measurement over periods of 24 hours, we recoded oxygen consumption (VO_2), carbon dioxide production (VCO_2), respiratory exchange ratio (RER = VCO_2/VO_2) from each chamber every 15 minutes (Fig. S4I). We simultaneously measured the locomotor

activity of each mouse using the break counts of infrared beams (Fig. S4L). We calculated energy expenditure by multiplying oxygen consumption (VO_2) by the calorific value (CV = 3.815 + 1.232 × respiratory exchange ratio) and normalized by lean mass measured by a EchoMRI (Fig. S4J). After 24 hours of measurement, we weighted the amount of food left in the cage to calculate food intake (Fig. S4K).

Thermoneutrality VO_2 measurements. For a mouse, room temperature (22°C) is a cold stress, thus requires increased heat generation to maintain body temperature, compared to mice in the thermoneutral zone (30°C). Thus, we reasoned that smaller body size in AdipoIrx3DN mice could lead to the development of a browning phenotype due to higher heat loss. To exclude this possibility and measure energy expenditure without thermal stress, we measured energy expenditure mice adapted to thermoneutrality (Fig. S4H). We housed mice at 30 °C for 8 days with free access to food and water. After acclimation under thermoneutral condition, we subjected mice to anesthesia (sodium pentoabarbital, 66mg/kg i.p.) to exclude potential influence from physical activity. Then we measured O_2 consumption (VO_2) using the indirect calorimeter and normalized with lean mass.

Targeted qPCR analysis of energy balance regulators in adipose dominant negative Irx3 mice versus littermate controls. We extracted RNA using TRIzol (Invitrogen) from different adipose tissue depots of the aP2-Irx3DN mice and controls, including perigonadal, inguinal and brown adipose tissue and measured mRNA levels from regulators of mitochondrial function, thermogenesis, lipid catabolism, lipid stores, and adipocyte differentiation (Fig. S4M-N). We synthesized cDNA with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). We performed qPCR using SYBR Green with 60ºC annealing temperature and calculated relative gene expression by the delta delta Ct method. We normalized target gene expression to expression of *Actb* (mouse).

qPCR analysis of EnR in different tissues. To evaluate expression of Adipo-IRX3DN adipose-specific dominant negative Irx3-EnR transgene in all measured adipose tissues, and robust expression of Hypo-IRX3DN hypothalamus-specific dominant negative Irx3-EnrR transgene in Hypothalamus, but no significant expression in the wrong tissues, we performed qPCR analysis of EnR mRNA levels in the respective tissue from adipose dominant negative aP2-Irx3DN mice and hypothalamus dominant Ins2-Irx3DN mice[8] (Fig. S4Q).

Adipocyte size measurements in adipose dominant negative Irx3 mice versus littermate controls. We measured the size of an adipocyte as cross-sectional surface area using Image J by manual tracing of ingunial adipocytes on three randomly selected H&E stained sections

per mouse (Fig. S4O). We performed histological analysis as described above (section IVb) We analyzed histograms of adipocyte cross-sectional area using Origin 8.0.

c. Validation of energy balance phenotypes using Irx3 and Irx5 manipulations in mouse adipocytes.

To confirm tissue-autonomous effects of *Irx3* and *Irx5* on energy balance, we evaluated the effects of *Irx3* and *Irx5* knockout and overexpression on lipid accumulation and/or UCP1 activation in three cellular mouse models (Fig. S5). First, we carried out a double knock-out of *Irx3* and *Irx5*, extracted mouse embryonic fibroblasts (MEFs), and *in vitro* differentiated them into adipocytes. Second, we over expressed Irx3 and Irx5 in mouse 3T3-L1 pre-adipocytes and differentiated them *in vitro.* Third, to examine a causal role of IRX3 and IRX5 in mitochondrial thermogenesis repression we used ME3 cells as a model of beige adipocytes. We over expressed IRX3 and IRX5 in ME3 cells and performed UCP1 reporter studies.

Mouse embryonic fibroblasts (MEFs) from Irx3/Irx5 knockout mice. We isolated MEFs from E14.5 embryos of wild-type and *Irx3/5* double knockout mice, as described previously[17]. We cultured MEFs in DMEM supplemented with 10% fetal bovine serum (FBS). Two days after cells become 100% confluent, we induced 3T3-L1 cells or MEFs were induced to differentiate to adipocytes by administrating DMEM containing 10% FBS, 500 μM 3-isobutyl-1-methylxanthine, 1 μM dexamethasone, and 10 μg/ml insulin. Two days after induction, cells were switched to the maintenance medium containing 10% FBS and 10 μg/ml insulin and cultured for 6-8 more days (up to 8-10 days after differentiation induction). Then cells were stained with Oil Red O to examine lipid accumulation (Fig. S5A) or were harvested for gene expression analysis (Fig. S5B).

Over-expression of Irx3 and Irx3 in 3T3-L1 pre-adipocytes. We cultured 3T3-L1 preadipocytes in DMEM supplemented with 10% fetal bovine serum (FBS). We achieved overexpression of *Irx3* and *Irx5* in 3T3-L1 cells by either piggyBAC transposon system[18] or infection with adenovirus carrying *Irx3* or *Irx5*. For piggyBAC transpson-mediated overexpression, we subcloned cDNAs encoding *Irx3* and *Irx5* into the PB-TAC vector. Then we transfected a mixture of the PB-TAC, PB-RB (i.e. the reverse tetracycline transactivator) and PBase (i.e. PB transposase) plasmids at a ratio of 5:1:1 into undifferentiated 3T3-L1 cells. One day after transfection, we started drug selection with puromycin and G418 (Geneticin) started and continued until the end of the experiment. When cell were 80% confluent, we added 1 μg/mL doxycycline to culture media in order to induce *Irx3* and *Irx5* overexpression and maintained throughout the experiment. For adenovirus-mediated overexpression, we infected 3T3-L1 cells with adenovirus at multiplicity of infection (MOI) of 50, when cell reached confluence. After overnight infection, we washed cells and added with fresh culture media. Two days

after cells become 100% confluent, we induced 3T3-L1 cells to differentiate to adipocytes by administrating DMEM containing 10% FBS, 500 µM 3-isobutyl-1-methylxanthine, 1 µM dexamethasone, and 10 µg/ml insulin. Two days after induction, we switched cells to the maintenance medium containing 10% FBS and 10 µg/ml insulin and cultured for 6-8 more days (up to 8-10 days after differentiation induction). Then we stained cells with Oil Red O to examine lipid accumulation (Fig. S5C) or were harvested for gene expression analysis (Fig. S5D).

ME3 cells over expressing IRX3 and IRX5. We derived ME3 cells from *RB*$^{-/-}$ mouse embryos, as a model of beige adipocytes.[19,20] We cultured cells in AmnioMax C100 basal medium (Life Technologies) containing 7.5% AmnioMax C100 Supplement (Life Technologies), 7.5% foetal bovine serum (FBS) (Life Technologies), 2 mM L-glutamine (Lonza) and Penicillin-Streptomycin (Sigma-Aldrich). On day 0, we seeded 50,000 cells in each well of a 24-well plate. We transfected the cells on day 1 using Superfect Transfection Reagent (Qiagen) and Opti-Mem Reduced Serum Medium (Life Technologies), with 0.5 µg pGL3 Basic UCP1 reporter plasmid and three different concentrations (0.1 µg, 0.5 µg and 1.0 µg) of pCMV6-XL5 IRX5 (SC128090, Origene) or pCMV6-AC IRX3 (SC319551, Origene). We adjusted the total concentration of plasmid using the empty vector pCMV6-XL4 (pCMV6-XL4, Origene). We induced UCP1 expression with 100 nM isoproterenol (Sigma-Aldrich) and 1 µM 9-cis-retinoic acid (Sigma-Aldrich). Controls were given an equivalent volume of DMSO. All experiments in triplicates. We lysed the cells with a buffer containing 10% glycerol, 25 mM TAE (pH 7.8), 1% Triton X_{100}, 1 mM EDTA and 2 mM DTT. We measured luciferase activity using FLUOstar Optima (BMG Labtech) according to manufacturer's protocol, with 50 µl ATP Substrate (BioThema) and Luciferin Substrate (BioThema) added to each well (Fig. S5E).

V: Identification and validation of the causal variant rs1421085.

Given that the resolution of GWAS is limited by the haplotype structure of the human genome (a common variant in the population is usually strongly genetically linked to dozens of neighboring variants), the probability that the reported marker SNP of the FTO locus is indeed the causal SNP driving the phenotypic association with obesity is minuscule (the FTO obesity locus harbors 89 common variants in high linkage disequilibrium). We sought to identify the causal variant underlying the FTO association with obesity.

a. Computational prediction of causal nucleotide.

To predict the causal variant that confers the increased risk of obesity in the population, we used several conservation scores based on a phylogenetic framework developed in our lab, i.e. Phylogenetic Module Complexity Analysis (PMCA)[21], which analyses evolutionary conservation of *cis*-regulatory modules across related species. Consistent with the chromatin state annotations across 127 human cell types and our tiling experiments, the highest Phylogenetic Module Complexity Analysis (PMCA) scores occur in the first 10kb interval of the locus. The conservation analysis predicted variant rs1421085 at position chr16:53,800,954 (hg19) as a conserved motif module with the highest-score of any of the non-coding variants

PMCA functional conservation analysis. To predict the causal variant underlying the FTO association with obesity, we used the PMCA method described in Ref.[21] with several modifications. Briefly, 972 position weight matrices grouped in 192 motif matrix families[22] (Genomatix, Munich, Germany) were used to scan a variant-flanking regions of the human reference genome (variant at mid-position) and its orthologous regions for cross-species functionally conserved motif modules (a module is defined as a set of binding site motifs, whose order and distance range is conserved across species).[21] (Fig. 3A, Fig. S6A). This method counts instances of conserved single motifs, motifs within conserved modules and conserved modules within a variant flanking region and computes the enrichment for any of those features using randomizations of orthologous sets (120bp centered on the variant) (Fig. 3B, Fig. S6B). Randomizations were done by local shuffling and an experimentally validated scoring scheme was applied as descibed in Ref. [21], with the following modification: background probabilities were computed using a set of 10,000 shuffled orthologous sequence sets.

b. Experimental validation that predicted nucleotide alters enhancer activity.

To validate that the single-nucleotide T-to-C rs1421085 alteration leads to a robust and reliable increase in enhancer activity we used enhancer luciferase assays in human SGBS adipocytes for elements of varying sizes (Fig. S6C) and cloned in different orientations (Fig. S6D) in the reporter gene construct.

Allelic enhancer activity assays for elements of varying sizes. To assay the enhancer context of rs1421085 allelic activity, we synthesized fragments with different sizes, including a 10kb, 1000bp and 100bp fragment encompassing rs1421085 at mid-position (segment 1_10kb: chr16:53799507-53809507; segment 1_1kb: chr16:53800454-53801454; segment 1_100bp: chr16:53800904-53801004) (Fig. S6C). We synthesized 10kb and 1000bp regions as plasmid vectors (Life Technologies) and 100bp as double-stranded oligonucleotides (MWG, Germany). We annealed complementary oligonucleotides and purified on a 12% polyacrylamide gel. For each of these fragments, we generated constructs carrying both alternate alleles were generated.

Cloning of constructs with elements of varying sizes. We subcloned genomic DNA segments with three different sizes (100bp, 1kb, 10kb) upstream of the TK promoter into the *KpnI* and *SacI* sites of the pGL4.22-TK vector in forward orientation as described above (section Ib). We performed site-directed mutagenesis (Quick Change II Site-Directed Mutagenesis Kit, Stratagene) to introduce the T-to-C substitution in the 10kb and 1000bp fragments. We verified the identity of each construct clone by DNA sequencing. We performed transfection in human SGBS adipocytes and enhancer activity measurements as described above (Section Ib).

Allelic enhancer activity assays with elements in different orientations. To confirm enhancer function of the rs1421085 surrounding regulatory region for the C risk allele, we further used reporter gene construct that harbor our identified enhancer element in the forward or reverse orientation and upstream or downstream orientation of the reporter gene (Fig. S6D).

Cloning of constructs in different orientations to the luciferase gene. We subcloned genomic DNA segments either upstream of the TK promoter into the *KpnI* and *SacI* sites of the pGL4.22-TK vector or downstream of the luciferase gene into the *BamHI* site of the pGL4.22-TK vector in both forward and reverse orientations. We used site-directed mutagenesis (Quick Change II Site-Directed Mutagenesis Kit, Stratagene) to introduce the T-to-C substitution in the non-risk background. We verified the identity of each construct clone by DNA sequencing. We performed transfection in human SGBS adipocytes as described above (Section Ib).

c. Experimental validation that predicted nucleotide alters transcription factor binding.

To test if the risk allelic enhancer activity and increase in IRX3 and IRX5 expression is provoked by differential transcription factor binding, we performed Electrophoretic Mobility Shift Assays (EMSA) using adipocytic nuclear extract.

Electrophoretic mobility shift assays (EMSA). To test if the variant rs1421085 shows allele-specific binding transcriptional regulators we performed electrophoretic mobility shift assays (EMSA) using the rs1421085 surrounding sequence in the two allelic forms (Fig. S6E). We performed EMSA with Cy5-labelled oligonucleotide probes. rs1421085-flanking region oligonucleotides were commercially synthesized containing either the risk or the protective allele (MWG). We annealed Cy5-labelled forward strands with non-labelled reverse strands, and we separated the double-stranded probes from single-stranded oligonucleotides on a 12% polyacrylamide gel. We visualized complete separation by DNA shading. We tested the efficiency of the labeling by a dot plot, which confirmed that all of the primers were labeled similarly. We prepared nuclear protein extracts from SGBS human adipose cells with adapted protocols based on the method described in Ref. [21]. We recovered the supernatant and

stored at -80°C. We conducted DNA-protein binding reactions in 50 mM Tris-HCl, 250 mM NaCl, 5 mM MgCl2, 2.5 mM EDTA, 2.5 mM DTT, 20% v/v glycerol and the appropriate concentrations of poly (dI-dC). For DNA-protein interactions, we incubated 3-5 µg of nuclear protein extract from the respective cell line for 10 min on ice, and added Cy-5-labelled genotype-specific DNA probe for another 20 min.

VI: Prediction and validation of the upstream regulator ARID5B.

Hypothesizing that the rs1421085 T-toC substitution alters a transcription factor binding site motif, we used our conservation algorithms to predict the binding regulator and we performed expression and enhancer assays to validate the causal role of the binding regulator.

a. Computational prediction of candidate regulators based on motif matches.

To establish the binding regulator we analyzed the sequence surrounding rs1421085 and searched for evolutionary conserved transcription factor binding site (TFBS) motif modules across mammalian species and determined disturbed TFBS motifs as described above (section IVa).

b. Expression-based prediction of likely causal regulator in humans.

Given that the T-to-C substitution at rs1421085 disrupts three AT-rich regulatory motifs for NKX6-2, LHX6, and the ARID (AT-rich interaction domain) motif family, suggesting them as putative upstream regulators we used gene expression analysis in whole human adipose tissue biopsies and isolated human adipose progenitor cells (Fig. S6F-G). We then tested diverse members of the ARID family of transcription factors for the expression patterns.

Microarray analysis in whole adipose tissue and isolated primary human adipocyte progenitor cells. We measured global gene expression in whole abdominal subcutaneous adipose tissue from 13 lean and 17 obese persons using Illumina HumanRef-8 v.3 BeadChip microarrays, as described previously.[23] (Fig. S6F). For isolated human adipose progenitor cells from human abdominal subcutaneous adipose tissue (12 lean and 12 obese persons), we fractionated about 700-800 mg of tissue by collagenase immediately after biopsy collection, as described previously.[24] (Fig. S6F). We analyzed the human adipose progenitor cells by Illumina HumanHT-12 v.3 BeadChip microarrays. We quantile normalized signal intensities.

Sample preparation for microarray analyses. We extracted total RNA with RNeasy Lipid Tissue Kit (Qiagen).

qPCR-based correlation analysis of IRX3/IRX5 with ARID5B. We analyzed the correlation of gene expression of *IRX3* and *IRX5*, with ARID5B in isolated human adipose progenitors cells

from 18 participants carrying the TT homozygous non-risk allele haplotype and 20 patients carrying the CC risk haplotype (Fig. S6G).

Sample preparation for qPCR analysis. We extracted total RNA with TRIzol (Invitrogen). We synthesized copy DNA (cDNA) with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) as described above (section IIb). We performed qPCR using SYBR Green with 60ºC annealing temperature. We calculated relative gene expression by the delta delta Ct method. We normalized target gene expression to expression of *HPRT*[12] or *TBP* (human). All genotyping was done for both the GWAS reported tag SNP and the identified variant rs1421085.

c. Experimental validation of the predicted regulator ARID5B in human cells.

To validate the repressor function of ARID5B specifically for the rs1421085 non-risk allele we performed cis/trans conditional enhancer assays and cis/trans endogenous gene expression analyses in primary human adipocytes from patients and controls (Fig. S6H-J, regulatory model in Fig. S6M).

Conditional gene expression analyses in primary human adipocytes. To establish causality we used *ARID5B* knock-down (Fig. S6H) and overexpression (Fig. S6J) in primary human adipocytes of risk and non-risk participants and measured regulation of *IRX3* and *IRX5* and regulation of downstream targets conditional on the presence of ARID5B by qPCR. We performed siRNA-mediated ARID5B knock-down with ON-TARGET-plus SMARTpool siRNA (Dharmacon) and HiPerFect (Qiagen) and overexpression with ARID5B cDNA derived from SGBS total cDNA, which we inserted into the doxycycline-inducible Tet-on system (Tet-On® Advanced Inducible Gene Expression System, BD Biosciences, Clontech, San Diego, CA, USA). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis of *IRX3* and *IRX5* as described above (Section IIb).

Conditional enhancer activity assays. To test for a causal role of ARID5B in regulating the enhancer at the *FTO* obesity locus, we performed ARID5B knock-down enhancer studies. siRNA-mediated ARID5B knock-down was performed with ON-TARGET-plus SMARTpool siRNA (Dharmacon) or siRNA from Life Technologies for IRX3/5 and HiPerFect (Qiagen) (Fig. S6I). We performed enhancer activity assays as described above (Section Ib).

d. Experimental validation of regulator binding

To further validate that ARID5B is the binding repressor we performed competition experiments where the labeled DNA-protein complex is competed away by increasing concentrations of unlabeled oligonucleotides containing the ARID5B consensus sequence (Fig. S6E).

EMSA competition experiments. We performed EMSA assay as described above (Section IVc). For competition experiments 11-, 33- and 100-fold molar excess of unlabeled probe as competitor was included with the reaction prior to addition of Cy5-labeled DNA probes (Fig. S6E). Binding reactions were incubated for 20 min at 4°C. The DNA-protein complexes were resolved on a nondenaturation 5.3% polyacrylamide gel in 0.5x Tris/borate/EDTA buffer. Fluorescence was visualized with a Typhoon TRIO+ imager (GE Healthcare).

VII: Validation of variant causality by genome editing.

Our regulatory model is that the rs1421085 single-nucleotide T-to-C alteration underlies the *FTO* association with obesity, by disrupting an evolutionarily-conserved motif for the ARID5B repressor, resulting in de-repression of a mesenchymal adipocyte progenitor super-enhancer in the first intron of the *FTO* gene, which activates downstream targets *IRX3* and *IRX5* up to 1.1Mb away and represses mitochondrial function, thermogenesis and adipocyte browning. This overexpression of *IRX3* and *IRX5* results in a cell-autonomous energetic shift from energy dissipation by mitochondrial thermogenesis in inducible beige adipocytes, to energy accumulation and lipid storage in white adipocytes, providing a mechanistic basis for the increased fat stores and body weight gain at the organismal level in risk allele carriers (mechanistic model in Fig. 4D). We sought to demonstrate causality of the single-nucleotide alteration in its endogenous context and used CRISPR/Cas9 genome editing in primary human patient samples (Fig. 3C-D, Fig. 4A-C).

a. Genome editing methodology.

CRISPR/Cas9 genome editing. To establish causality of the single-nucleotide alteration in its endogenous context for obesity cellular phenotypes, we performed CRISPR/Cas9 genome editing in primary human adipose cells to alter the rs1421085 single-nucleotide polymorphism (SNP) from TT non-risk allele to CC risk allele and vice versa We purchased the hCas9 plasmid and the guide RNAs (gRNA) cloning vector from Addgene. To change the T to the C allele and the C to the T allele, respectively, we performed Site-directed Mutagenesis using the Q5® Site-Directed Mutagenesis Kit (New England Biolabs). We used the mutagenesis primer forward 5'-TAAGGCATGAcATTGATTAAGTGTC-3' and reverse 5'-GGACCTGAACTGCTACTATAAAATC-3'. To exclude off-target effects we carried out reversion after initial editing with a separate guide RNA. We designed the guide RNAs (gRNAs) using the CRISPR design online tool from the Zhang lab (http://crispr.mit.edu/). We used the following guide RNAs: 5' CACCGCTGGAAGGAACGCGTTTGTT-3', 5'-CACCGGGACAGTGCGTAGACTAAAC-3'. For transfection, we co-transfected GFP and Cas9 with sgRNAs, the homology vector, and pMACS plasmid 4.1 (Milteny Biotec) in human adipose progenitor cells using the Amaxa-Nucleofector device (program U-033) (Lonza). We

sorted cells using the MACSelectTM Transfected Cell Selection cell sorting kit (Miltenyi Biotec). We cultured sorted cells for 3–5 days and clones propagated from a single cell were picked out. We confirmed nucleotide exchange by DNA sequencing.

b. Bi-directional editing of risk and non-risk adipocytes.

We used bi-directional editing of the rs1421085 variant using CRISPR/Cas9 genome editing from both the protective and the risk allele. We started from isolated adipose-derived mesenchymal cells from one patient carrying the non-risk haplotype (Fig. 3C). We also carried out the converse editing event, starting with a CC homozygous risk participant (Fig. 3D). Genomic DNA was amplified from the respective rs1421085 CC risk allele carrier and TT non-risk allele carrier. Genome editing was performed as described above (Section VIIa).

c. Establishment of causality for repression and temporal context dependency.

To confirm that ARID5B is necessary for the repression of *IRX3* and *IRX5*, specifically for the C-to-T rescued allele we performed conditional analysis of single-nucleotide editing and knock-down of the upstream regulator ARID5B (mechanistic model in Fig. 4D).

Conditional gene expression analyses in CRISPR/Cas9 edited primary human adipocytes. To establish repressor causality we carried out *ARID5B* knock-down in our CRISPR/Cas9 edited patient cells and measured regulation of *IRX3* and *IRX5* conditional on the presence of ARID5B by qPCR (Fig. 3D). We performed C-toT editing by CRISPR/Cas9 as described above (Section VIa). siRNA-mediated ARID5B knock-down was performed with ON-TARGET-plus SMARTpool siRNA (Dharmacon) and HiPerFect (Qiagen). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis as described (Section IIb).

Temporal dynamics of rs1421085 allelic activity on IRX3 and IRX3 expression in CRISPR/Cas9 edited cells from a risk allele carrier. We performed C-toT editing by CRISPR/Cas9 as described above (Section VIa). *IRX3* and *IRX5* gene expression was measured at day 0, 2, 4, 6 of differentiation (Fig. 4A). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis as described above (Section IIb).

d. Establishment of causality for cellular phenotypes of mitochondrial function and thermogenesis.

To confirm that the C-to-T genome editing in primary human adipose progenitor cells from a homozygous risk allele carrier reverses obesity cellular signatures, we performed gene expression analyses of mitochondrial function and thermogenic marker genes and performed oxygen consumption rate measurements as a surrogate for thermogenic activity of cells (mechanistic model in Fig. 4D).

Gene expression analyses of mitochondrial function, thermogenic marker genes and lipid metabolism regulators in CRISPR/Cas9 edited cells from a risk allele carrier. We performed gene expression analyses using qPCR in 5 clonal expansions from CRISPR/Cas9 edited cells (Fig. 4B, Fig. S7A). Genome editing was performed as described above (Section VIa). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis as described above (Section IIb).

Oxygen Consumption Rate Measurements in CRISPR/Cas9 edited cells from a risk allele carrier. We performed basal and stimulated OCR measurements in 5 clonal expansions from CRISPR/Cas9 edited cells. For OCR, we treated cells with isoproterenol (1 µmol/l) or DMSO (basal) for 4 hours. We measured total OCR using the Seahorse XF24 (Seahorse Bioscience) and subtracted the minimum OCR level after rotenone/antimycin A treatment (5 µM) from the initial level without treatment (Fig. S7B). Furthermore, to more directly assess effects on mitochondrial thermogenesis, we calculated uncoupled respiration (proton leak) by subtracting the minimum OCR level after rotenone/antimycin from the minimum level after oligomycin treatment (5 µM) (Fig. 4C). OCR is described in more detail in Section IIIc. We ran the protocol with 3 min mixing and 2 min measuring. CRISPR/Cas9 genome editing was performed as described above (Section VIa).

Conditional oxygen consumption rate analyses in CRISPR/Cas9 edited primary human adipocytes. To establish repressor causality we carried out *ARID5B* knock-down in our CRISPR/Cas9 edited patient cells and measured regulation of basal and stimulated oxygen consumption rate conditional on the presence of ARID5B by qPCR (**Fig. S7C**). We performed C-toT editing by CRISPR/Cas9 as described above (Section VIa). siRNA-mediated ARID5B knock-down was performed with ON-TARGET-plus SMARTpool siRNA (Dharmacon) and HiPerFect (Qiagen). We performed sample preparation of primary human adipose cells, cDNA preparation and qPCR analysis as described above (Section IIb).

Accession codes. Microarray data have been deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress/) or GEO (http://www.ncbi.nlm.nih.gov/GEO). The data for abdominal subcutaneous adipose tissue (13 lean and 17 obese) is available with ArrayExpress accession E-TABM-862. The perirenal adipose tissue dataset (10 lean persons) is available with GEO accession (GSE59325).

Author Contributions

[M.C.] designed the study, [M.C., S.N.D., K-H. K., C.H., V.G., I.S.S., J.B., V.P., J.L., P.A.S.,] gathered the data, [M.C., S.N.D., G.Q., W.M., N.A.A., Y.H., D.J.D., G.M., C.C.H., H.H., M.K.] analyzed the data, [M.C. and M.K.] vouch for the data and the analysis, [M.C. and M.K.]

wrote the paper with input from [S.N.D.], [M.C. and M.K.] decided to publish the paper. The authors declare no competing financial interests.

Acknowledgements

We thank Jan-Inge Bjune and Elisabeth Hofmair for technical assistance, and the Norwegian Genomics Consortium (NGC), Bergen, Norway for technical support, Iacovos P Michael and Andras Nagy for providing reagents and technical support for PiggyBAC transposon system. We thank Prof. Jan W Eriksson for obtaining perirenal adipose tissue biopsies. We thank Dr. Thomas Skurk for obtaining whole adipose tissue samples. Grants were provided by the National Institutes of Health National Human Genome Research Institute NHGRI through RC1HG005334 and R01HG004037 (M.K.); Else Kröner-Fresenius Foundation, Bad Homburg v.d.H, Germany; Virtual Institute “Molecular basis of glucose regulation and type 2 diabetes” received from the Helmholtz Zentrum München, München-Neuherberg, Germany; Clinical Cooperation Group Nutrigenomics and Type 2 Diabetes of the Helmholtz Zentrum München, German Research Center for Environmental Research, 85764 Neuherberg, and Technische Universität München, 85350 Freising-Weihenstephan; German Federal Ministry of Education and Research to the German Centre for Diabetes Research (DZD e.V.); Samarbeidsorganet Helse Vest RHF, Norway, and the Swedish federal government under the LUA/ALF agreement. M.C. was supported by a fellowship from the German Federal Ministry of Education and Research to the German Centre for Diabetes Research (DZD e.V.); C-C.H. and DJD were supported by a grant from the Canadian Institute of Health Research. K.-H.K. was supported by fellowships from the Heart and Stroke Foundation of Canada and the Hospital for Sick Children.

Statistical Analysis Plan

Statistical analyses were performed using two-tailed Student’s t-test or ANOVA for comparing the means of two or multiple groups, respectively. Nonparametric testing (U-Mann–Whitney test) was used where appropriate, that is, when normal distribution of sample sets was not evident.

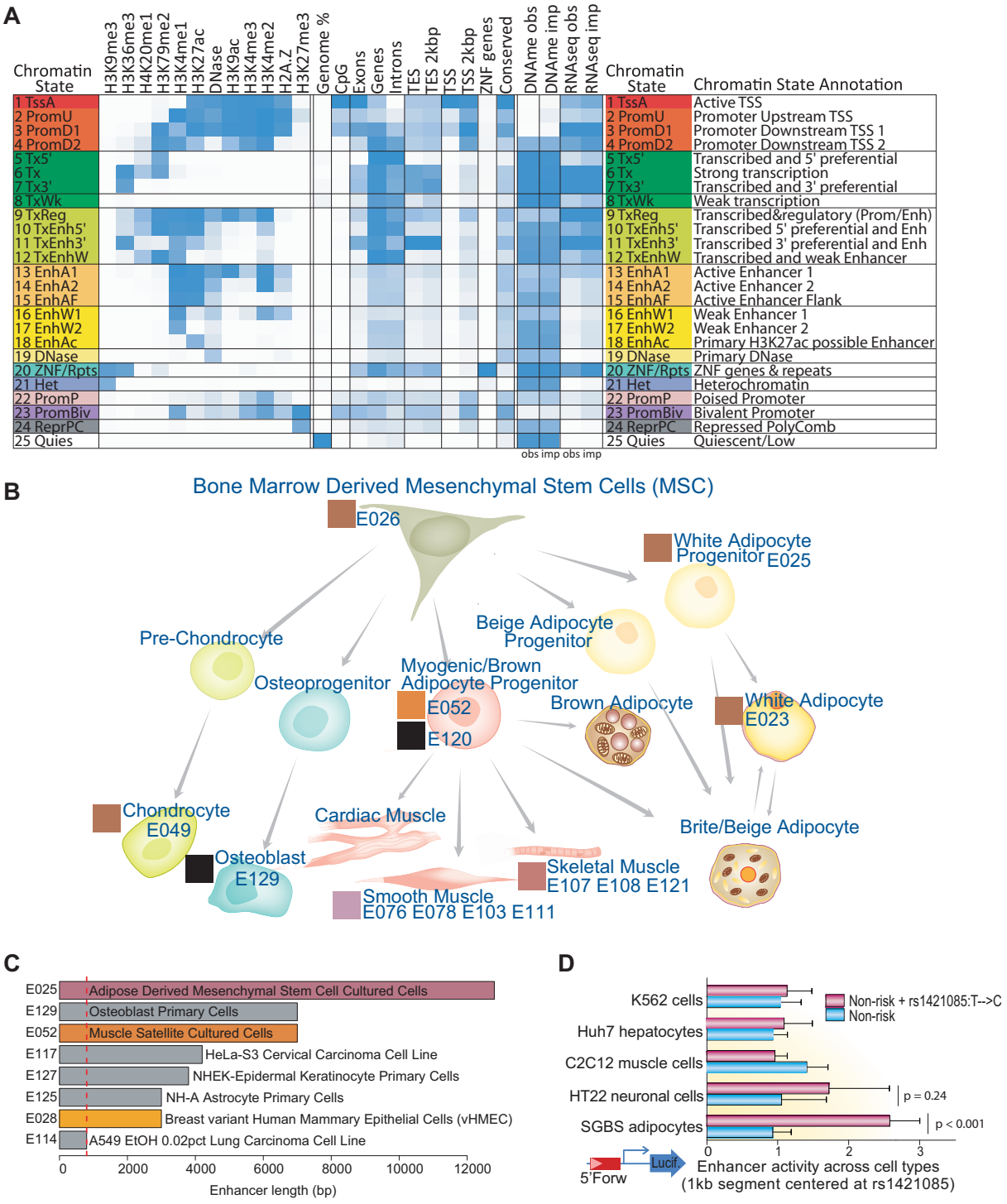
Supplementary Appendix References

1. Roadmap_Epigenomics_Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. Nature 2015;518:317-30.
2. ENCODE_Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.
3. Zhou X, Maricque B, Xie M, et al. The Human Epigenome Browser at Washington University. Nat Methods 2011;8:989-90.
4. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods 2012;9:215-6.
5. Ernst J, Kellis M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. Nat Biotechnol 2015.
6. Fischer-Posovszky P, Newell FS, Wabitsch M, Tornqvist HE. Human SGBS cells - a unique tool for studies of human fat cell biology. Obes Facts 2008;1:184-9.
7. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 2009;326:289-93.
8. Smemo S, Tena JJ, Kim KH, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature 2014;507:371-5.
9. Dixon JR, Selvaraj S, Yue F, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 2012;485:376-80.
10. Imakaev M, Fudenberg G, McCord RP, et al. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods 2012;9:999-1003.
11. Veum VL, Dankel SN, Gjerde J, et al. The nuclear receptors NUR77, NURR1 and NOR1 in obesity and during fat loss. Int J Obes (Lond) 2012;36:1195-202.
12. de Kok JB, Roelofs RW, Giesendorf BA, et al. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab Invest 2005;85:154-9.
13. Svensson PA, Lindberg K, Hoffmann JM, et al. Characterization of brown adipose tissue in the human perirenal depot. Obesity (Silver Spring) 2014.
14. Ye L, Wu J, Cohen P, et al. Fat cells directly sense temperature to activate thermogenesis. Proc Natl Acad Sci U S A 2013;110:12480-5.
15. Shan T, Liu W, Kuang S. Fatty acid binding protein 4 expression marks a population of adipocyte progenitors in white and brown adipose tissues. FASEB J 2013;27:277-87.
16. Li D, Sakuma R, Vakili NA, et al. Formation of proximal and anterior limb skeleton requires early function of Irx3 and Irx5 and is negatively regulated by Shh signaling. Dev Cell 2014;29:233-40.
17. Li ZJ, Nieuwenhuis E, Nien W, et al. Kif7 regulates Gli2 through Sufu-dependent and -independent functions during skin development and tumorigenesis. Development 2012;139:4152-61.
18. Woltjen K, Michael IP, Mohseni P, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 2009;458:766-70.
19. Lukas J, Bartkova J, Rohde M, Strauss M, Bartek J. Cyclin D1 is dispensable for G1 control in retinoblastoma gene-deficient cells independently of cdk4 activity. Mol Cell Biol 1995;15:2600-11.
20. Hansen JB, Jorgensen C, Petersen RK, et al. Retinoblastoma protein functions as a molecular switch determining white versus brown adipocyte differentiation. Proc Natl Acad Sci U S A 2004;101:4112-7.
21. Claussnitzer M, Dankel SN, Klocke B, et al. Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. Cell 2014;156:343-58.
22. Cartharius K, Frech K, Grote K, et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics 2005;21:2933-42.

23. Dankel SN, Fadnes DJ, Stavrum AK, et al. Switch from stress response to homeobox transcription factors in adipose tissue after profound fat loss. PLoS One 2010;5:e11033.
24. Methlie P, Dankel S, Myhra T, et al. Changes in adipose glucocorticoid metabolism before and after bariatric surgery assessed by direct hormone measurements. Obesity (Silver Spring) 2013;21:2495-503.
25. Spalding KL, Arner E, Westermark PO, et al. Dynamics of fat cell turnover in humans. Nature 2008;453:783-7.

Supplementary Figures and Tables.

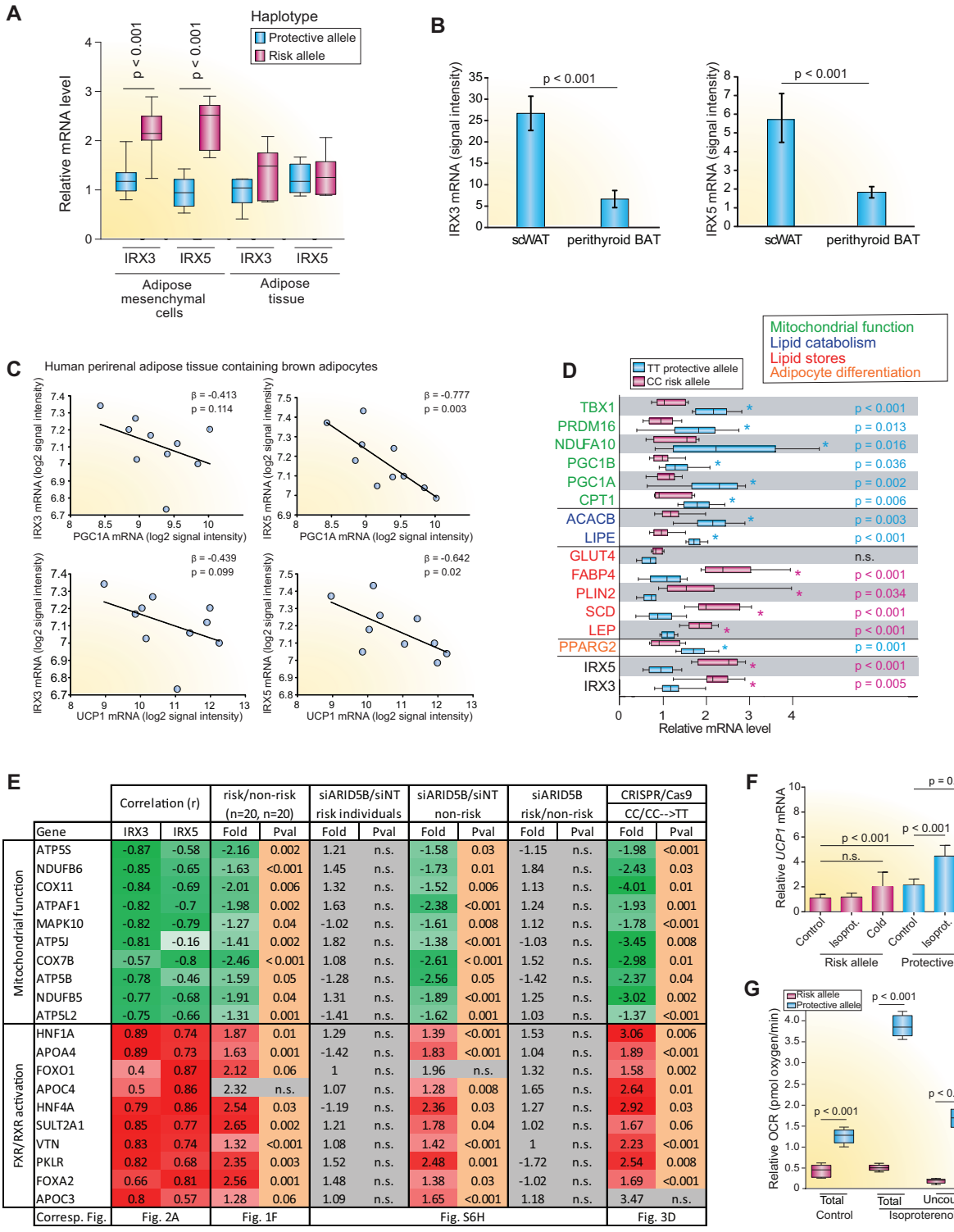
Figure S1. Cell type specificity and epigenomic annotations



Panel A shows 25-state chromatin state model using 12 imputed epigenomic marks learned jointly across 127 reference epigenomes. For each state (rows) are shown: the state abbreviation; the frequency with which each mark is expected in each state (emission parameters);

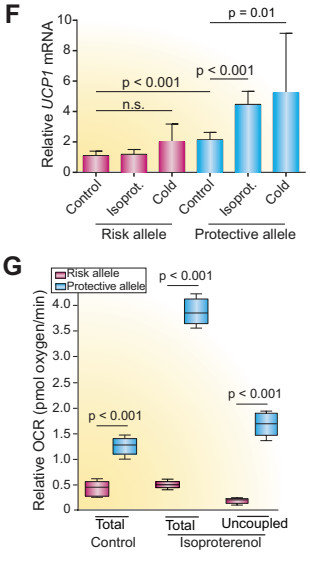
the genome coverage and functional enrichments for diverse annotations and conserved elements; the median observed and imputed DNA methylation and RNA-Seq signal; and a short description. Panel B shows lineage relationships of mesenchymal stem cells (MSCs) and MSC-derived lineages, including brown, white, and brite/beige adipocytes involved in lipid storage and energy dissipation. Epigenome identity (EID) numbers correspond to Roadmap Epigenomics[1] numbering of reference epigenomes. E114-E129 correspond to ENCODE project[2] reference epigenomes. Cell graphics modified with permission by Bio-Techne. Panel C shows that tissues and cell types with the longest enhancer lengths in the *FTO* locus include multiple mesenchymal-derived lineages, e.g., adipose-derived mesenchymal cells (12.8kb, E025), muscle progenitor cells (7kb, E052), and osteoblasts (1kb, E129). The super-enhancer active in adipose-derived mesenchymal cells is 15.2-fold longer than the median enhancer length (800bp, dashed red line). Panel D displays the result of allelic enhancer assays for Segment 1 across five cell types, including both BMI-relevant cell types and control cell types, measured by relative luciferase expression. Non-risk haplotype (blue) and risk rs1421085 allele (red) introduced as a single nucleotide alteration on the non-risk haplotype were tested for segment 1, by transfection in SGBS adipocytes, HT22 neuronal cells, C2C12 muscle cells, Huh7 hepatocytes and K562 lymphoma cells. Enhancer activity was measured by relative luciferase expression and normalized to the mean ratio from pGL4.22-TK promoter construct. Risk allelic enhancer activity was specifically observed in human SGBS adipocytes (n=7). P-values: t-test, Error bars: SD.

Figure S2. Allele-dependent association of IRX3 and IRX5 with lipid storage and mitochondrial function genes in human adipose cells



E

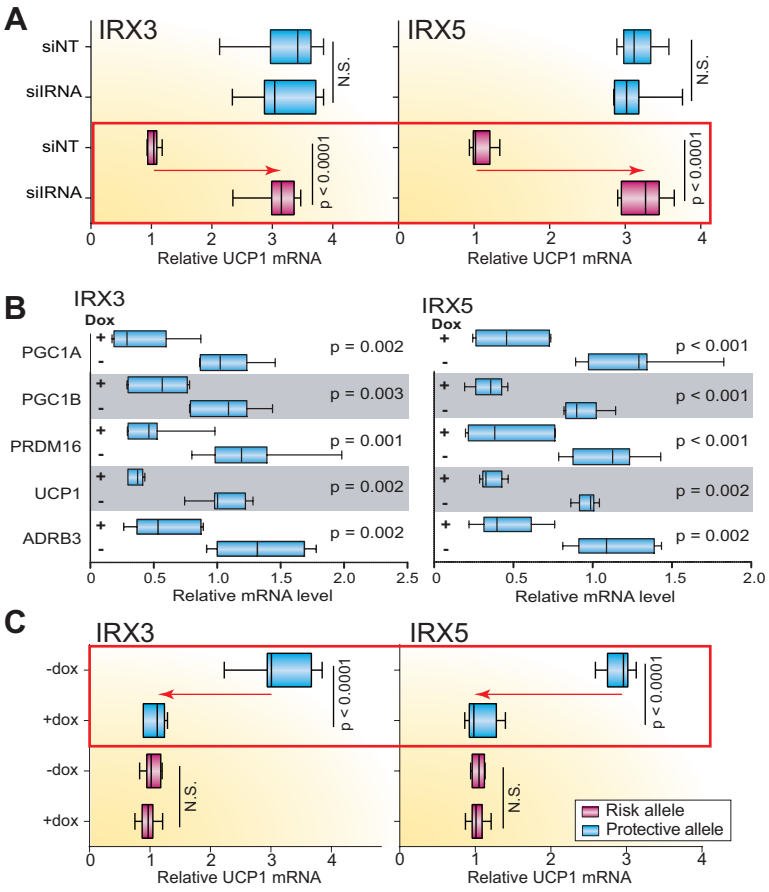
		Correlation (r)		risk/non-risk (n=20, n=20)		siARID5B/siNT risk individuals		siARID5B/siNT non-risk		siARID5B risk/non-risk		CRISPR/Cas9 CC/CC-->TT	
	Gene	IRX3	IRX5	Fold	Pval	Fold	Pval	Fold	Pval	Fold	Pval	Fold	Pval
Mitochondrial function	ATP5S	-0.87	-0.58	-2.16	0.002	1.21	n.s.	-1.58	0.03	-1.15	n.s.	-1.98	<0.001
	NDUFB6	-0.85	-0.65	-1.63	<0.001	1.45	n.s.	-1.73	0.01	1.84	n.s.	-2.43	0.03
	COX11	-0.84	-0.69	-2.01	0.006	1.32	n.s.	-1.52	0.006	1.13	n.s.	-4.01	0.01
	ATPAF1	-0.82	-0.7	-1.98	0.002	1.63	n.s.	-2.38	<0.001	1.24	n.s.	-1.93	0.001
	MAPK10	-0.82	-0.79	-1.27	0.04	-1.02	n.s.	-1.61	0.008	1.12	n.s.	-1.78	<0.001
	ATP5J	-0.81	-0.16	-1.41	0.002	1.82	n.s.	-1.38	<0.001	-1.03	n.s.	-3.45	0.008
	COX7B	-0.57	-0.8	-2.46	< 0.001	1.08	n.s.	-2.61	< 0.001	1.52	n.s.	-2.98	0.01
	ATP5B	-0.78	-0.46	-1.59	0.05	-1.28	n.s.	-2.56	0.05	-1.42	n.s.	-2.37	0.04
	NDUFB5	-0.77	-0.68	-1.91	0.04	1.31	n.s.	-1.89	<0.001	1.25	n.s.	-3.02	0.002
	ATP5L2	-0.75	-0.66	-1.31	0.001	-1.41	n.s.	-1.62	0.001	1.03	n.s.	-1.37	<0.001
FXR/RXR activation	HNF1A	0.89	0.74	1.87	0.01	1.29	n.s.	1.39	<0.001	1.53	n.s.	3.06	0.006
	APOA4	0.89	0.73	1.63	0.001	-1.42	n.s.	1.83	<0.001	1.04	n.s.	1.89	<0.001
	FOXO1	0.4	0.87	2.12	0.06	1	n.s.	1.96	n.s.	1.32	n.s.	1.58	0.002
	APOC4	0.5	0.86	2.32	n.s.	1.07	n.s.	1.28	0.008	1.65	n.s.	2.64	0.01
	HNF4A	0.79	0.86	2.54	0.03	-1.19	n.s.	2.36	0.03	1.27	n.s.	2.92	0.03
	SULT2A1	0.85	0.77	2.65	0.002	1.21	n.s.	1.78	0.04	1.02	n.s.	1.67	0.06
	VTN	0.83	0.74	1.32	<0.001	1.08	n.s.	1.42	<0.001	1	n.s.	2.23	<0.001
	PKLR	0.82	0.68	2.35	0.003	1.52	n.s.	2.48	0.001	-1.72	n.s.	2.54	0.008
	FOXA2	0.66	0.81	2.56	0.001	1.48	n.s.	1.38	0.03	-1.02	n.s.	1.69	<0.001
	APOC3	0.8	0.57	1.28	0.06	1.09	n.s.	1.65	<0.001	1.18	n.s.	3.47	n.s.
	Corresp. Fig.	Fig. 2A		Fig. 1F		Fig. S6H						Fig. 3D	



Panel A shows risk (n=13) and non-risk (n=13) homozygous participant mRNA expression for *IRX3* and *IRX5* measured using qPCR (normalized to *HPRT*) in adipose mesenchymal cells (left, as in Fig. 1F) cultured, induced to differentiate, and harvested after 2 days, and for whole adipose tissue (right), demonstrating lack of eQTLs in whole-adipose tissue and the

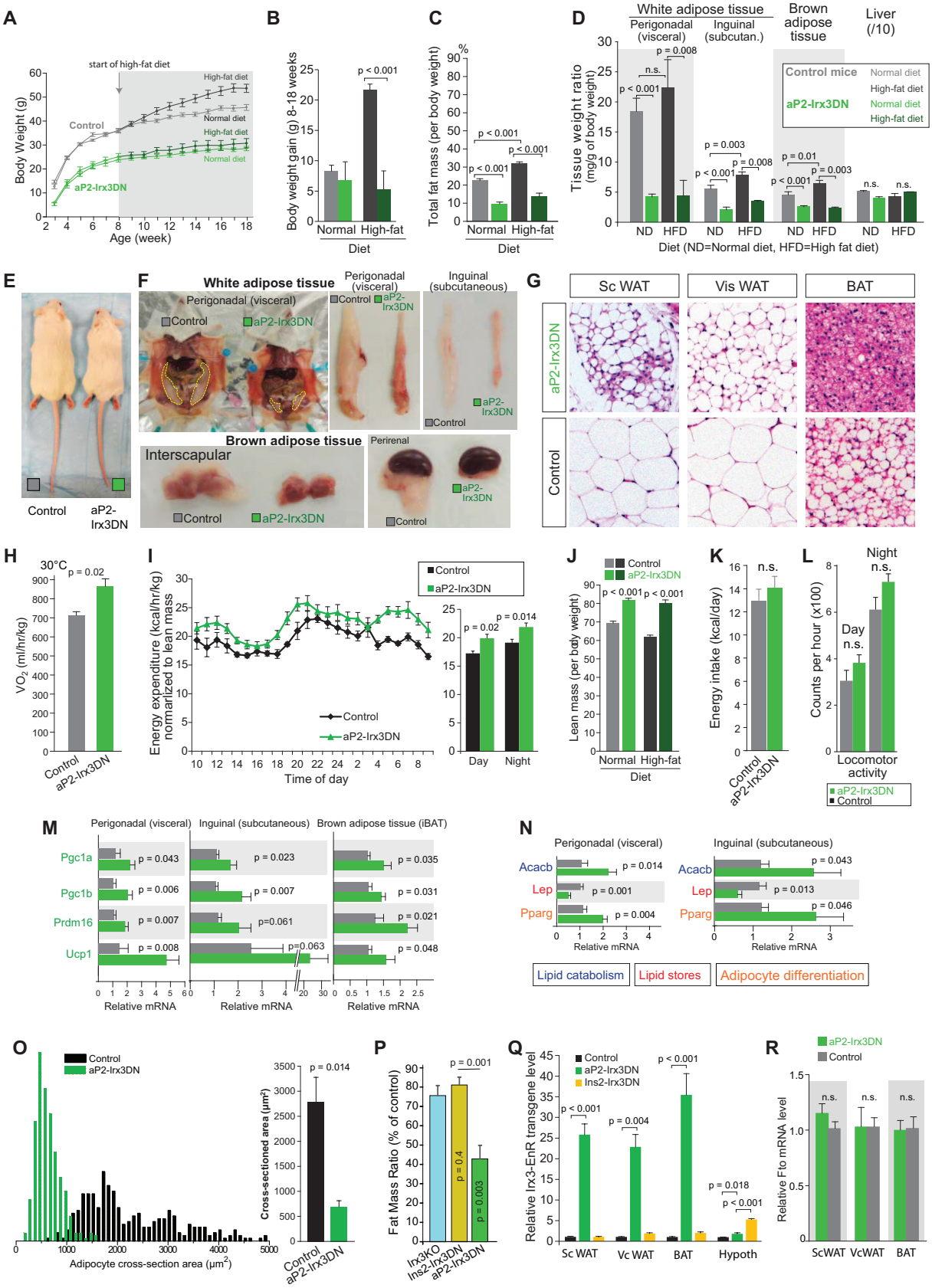
importance of adipocyte-specific measurements. Panel B shows reduced expression of *IRX3* and *IRX5* in perithyroid brown adipose tissue (BAT) vs. subcutaneous white adipose tissue (scWAT) for healthy persons, measured by Affymetrix microarrays (n=9) (*IRX3* probe set 229638_at, *IRX5* probe set 210239_at), consistent with repression of thermogenesis in white WAT. P-values: t-test, Error bars: SEM. Panel C shows negative correlation between *IRX3* (Affymetrix microarray probe set 8001449) and *IRX5* (probe set 7995668) mRNA levels and both mitochondrial biogenesis marker *PGC1A* (probe set 8099633), and uncoupling protein *UCP1* (probe set 8102904) in human perirenal white subcutaneous adipose tissue containing brown adipocytes, obtained from healthy kidney donors (n=10). Panel D shows reduced risk-allele expression of gene markers for mitochondrial function (green), catabolic lipid metabolism (blue), and adipocyte differentiation (orange), and increased risk-allele expression of gene markers for anabolic lipid metabolism (red), measured using qPCR in primary adipose cells isolated from participants homozygous for the non-risk (n=18) and risk allele (n=20). Risk allele carriers show reduced adipocyte differentiation marker *PPARG2* (orange), indicating reduced adipocyte turnover.[25] *IRX3* and *IRX5* (black) are also shown as reference. Panel E shows the fold change (Fold) and significance (Pval) of expression changes for the 20 genes that are most positively (n=10) and negatively (n=10) correlated with inter-individual expression of IRX3 and IRX5 (Fig. 2A) for three settings: the risk allele-dependent gene expression in isolated adipocyte progenitors (as shown in Fig. 1F for IRX3 and IRX5); upon ARID5B knock-down (as shown in Fig. S6H for IRX3 and IRX5); and after T-to-C allele editing by CRISRP/Cas9 (as shown in Fig. 4A for IRX3 and IRX5). Correlation coefficients are Spearman's rho (n = 10 perirenal white adipose tissue containing brown adipocytes from Fig. 2A). p-values were calculated by Mann Whitney U test. Panel F shows reduced UCP1 gene expression from risk allele patients (n=8) compared to non-risk allele participants (n=8) for baseline, cold exposure (incubation at 30°C for 6 hours) and isoproterenol-mediated β3-adrenergic receptor activation (1 µmol/l, 12 hours) indicative of thermogenic response in adipocytes. Panel G shows reduced basal and isoproterenol-stimulated (coupled and uncoupled) respiration (oxygen consumption rate, OCR) in isolated adipocyte progenitor cells from risk (n = 26) vs. non-risk (n=20) homozygous participants after 2 days after differentiation, revealing disrupted mitochondrial function and thermogenesis response.

Figure S3. Regulatory roles of ***IRX3*** and ***IRX5*** in adipocyte mitochondrial thermogenesis in humans.



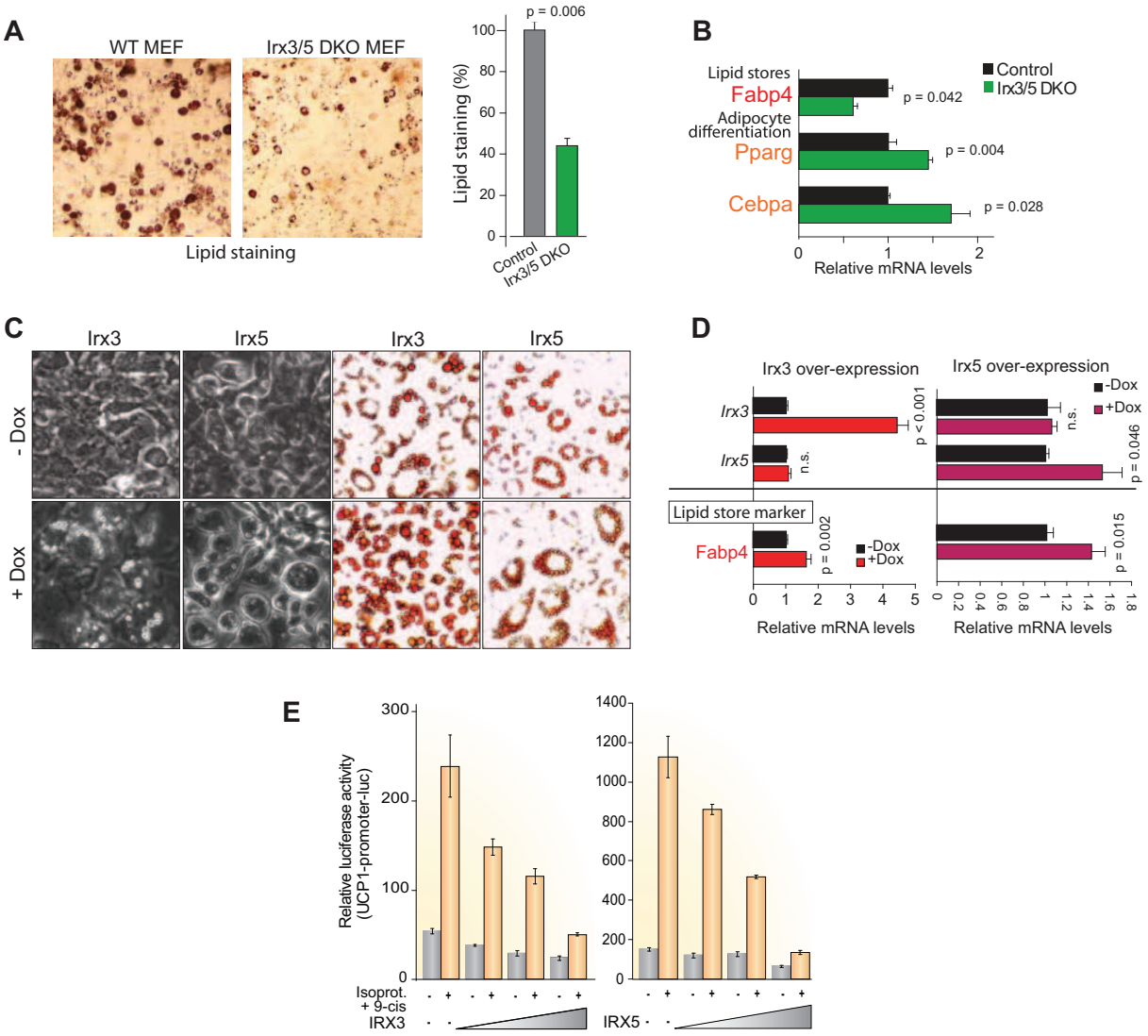
Panel A shows rescue of UCP1 gene expression levels upon knock-down of IRX3 or IRX5 in risk-allele carriers (n=10), restoring protective-allele UCP1 expression levels, but no change in UCP1 expression for IRX3 or IRX5 knock-down in protective-allele participants (n=8). The cells were induced to differentiate concomitant with siRNA treatment and collected after two days. Panel B shows reduced expression of mitochondrial activity marker genes upon doxycycline-induced overexpression of *IRX3* and *IRX5* in adipocyte progenitors specifically in non-risk (n=8) vs. risk (n=10) participants. Panel C shows reduced UCP1 expression upon IRX3 or IRX5 overexpression in differentiating primary human adipocyte progenitor cells from protective-allele (n=8) participants, resulting in risk allele participant UCP1 levels, but no change in UCP1 expression upon IRX3 or IRX5 overexpression in risk-allele participants (n = 10). The cells were induced to differentiate concomitant with Doxycycline and collected after two days. Target gene expression measured by qPCR was normalized to *HPRT* mRNA.

Figure S4. Anti-obesity phenotypes in Adipo-Irx3DN transgenic mice.



All panels show phenotypic consequences of a adipocyte-specific repression of Irx3 by expression a dominant-negative form of Irx3 in adipose cells (green; aP2-Cre;Irx3EnR), compared to littermate controls (grey; EnR). Panels A and B show reduced body weight for Adipo-Irx3DN mice, and reduced body weight gain on a high-fat diet (n=11 normal diet, n=10 high-fat) vs. controls (n=12 normal diet, n=6 high-fat diet). Panel C shows reduced total fat mass as a percentage of total body mass in Adipo-Irx3DN mice vs. controls (white and brown adipose tissue combined). Panel D shows reduced adipose tissue weight for major fat stores for Adipo-Irx3DN mice vs. controls, but no difference in liver. Panel E shows reduced body size of aP2-Irx3DN transgenic mice expressing a dominant negative form of *Irx3* in adipose cells (aP2-Cre;*Irx3*EnR). Panel F shows reduced adipose tissue mass in Adipo-Irx3DN vs. controls measured at 18 weeks of age after a high-fat diet introduced at 8 weeks. Panel G shows 4-fold reduced adipocyte size in white and brown fat sections upon high-fat diet for Adipo-Irx3DN mice vs. control. Panel H shows increased energy expenditure for aP2-Irx3DN mice vs. controls after 8 days at thermoneutrality (30ºC), measured by indirect calorimetry (n = 3 per group). Panel I shows increased oxygen consumption in Adipo-Irx3DN vs. controls (normalized to lean mass), a marker of increased energy expenditure. Difference is seen both during the day and during the night, indicating that it is not due to increased locomotor activity. Panel J shows increased fraction of lean mass per body weight in Adipo-IRX3DN mice vs. controls. Panel K shows unchanged energy intake in Adipo-Irx3DN mice vs. controls. Panel L shows unchanged locomotor activity in Adipo-Irx3DN mice vs. controls. Panel M shows increased expression of mitochondrial function genes in adipose tissues of Adipo-Irx3DN mice vs. controls. Panel N shows increased expression of lipid catabolism marker *Acacb* in Aidpo-Irx3DN mice vs. controls, decreased expression of lipid storage marker *Lep*, and increased expression of adipocyte differentiation marker *Pparg*, measured by qPCR with normalization to β-actin (*Actb*) mRNA. (n = 9-11). Panel O show distribution and and average of cross-sectioned area of subcutaneous white adipocytes, revealing reduced white adipocyte size for Adipo-IRX3DN compared to control mice. Panel P shows 57% reduced fat-mass ratio vs. control for aP2-Irx3DN mice vs. 25% reduction for *Irx3* knock-out mice, and 19% for Ins2-Irx3DN. Panel Q confirms robust expression of Adipo-IRX3DN adipose-specific dominant negative Irx3-EnR transgene in all measured adipose tissues, and robust expression of Hypo-IRX3DN hypothalamus-specific dominant negative Irx3-EnrR transgene in Hypothalamus, but no significant expression in the wrong tissues, indicating that the phenotypes we observe for Adipo-Irx3DN are not due to ectopic expression in hypothalamus. Panel R shows unchanged endogenous *Fto* mRNA level in Adipo-Irx3DN vs. controls in adipose tissues, indicating that Fto itself had no role in the anti-obesity phenotype.
Sc=subcutaneous/inguinal. Vc=Visceral. WAT=White Adipose Tissue. BAT=interscapular brown adipose tissue.

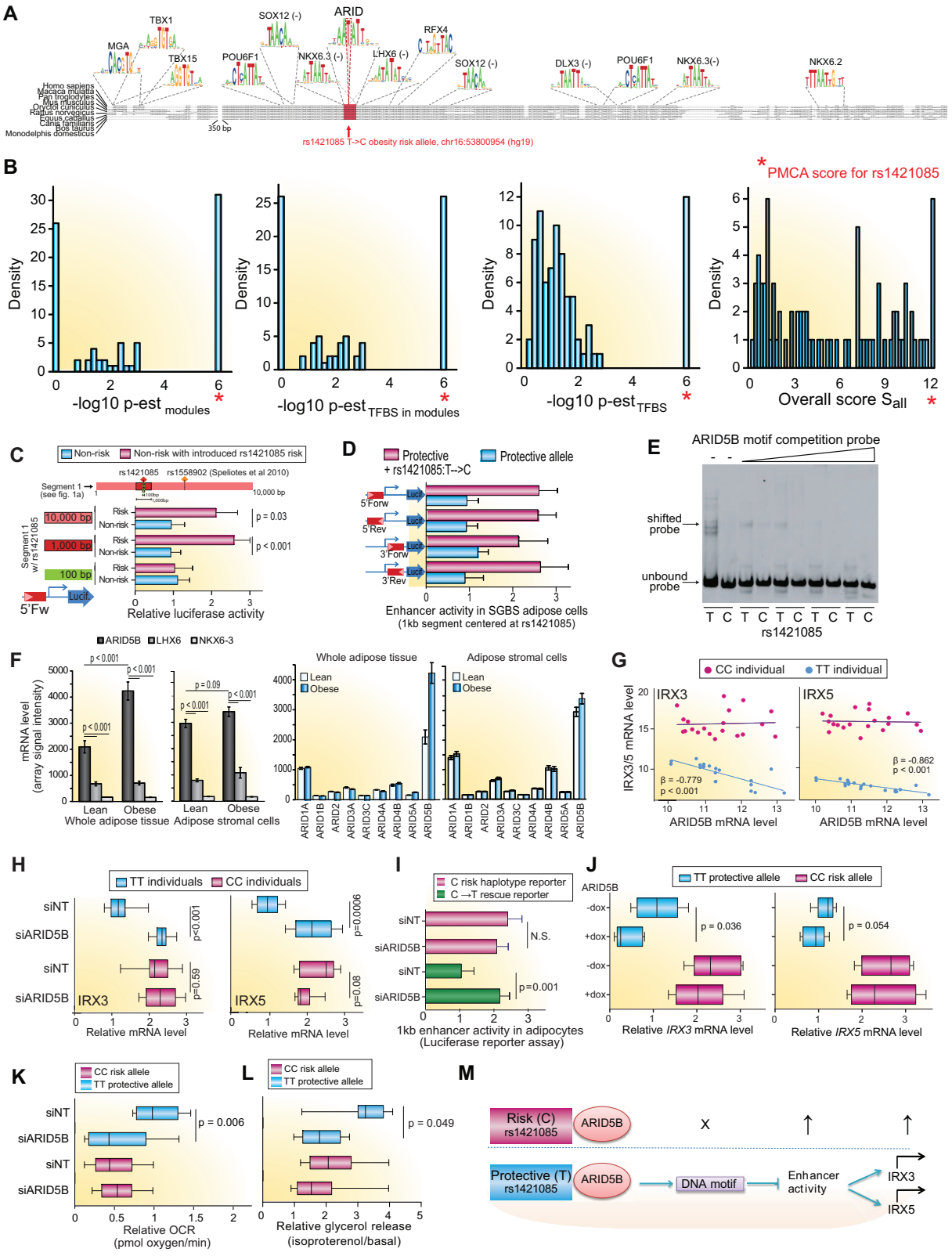
Figure S5. Cell-autonomous regulation of obesity phenotypes by Irx3 and Irx5 in mouse adipocytes.



Panel A shows decreased lipid accumulation in cultured mouse embryonic fibroblasts (MEFs) from *Irx3/5* double knock-out (DKO) mice, measured by Oil Red O lipid staining after 8 days of *in vitro* adipocyte differentiation, indicating cell-autonomous regulation of obesity-related phenotypes. Panel B shows reduced expression of the lipid carrier gene aP2/Fabp4, a marker of lipid stores, but increased expression of adipocyte differentiation markers *Pparg* and *Cepbalpha* in *Irx3/5* DKO MEFs. Panels C-D show that overexpression of *Irx3* or *Irx5* in 3T3-L1 cells leads to increased lipid accumulation visualized by Oil-Red-O lipid staining (Panel C) and increased mRNA expression of *Fabp4* (Panel D). The images are representative of phase-contrast microscopic views of the 3T3-L1 adipocytes on day 4 post-differentiation, and Oil-Red-O lipid staining on day 8 post-differentiation. Irx3 overexpression did not influence *Irx5* levels, and Irx5 overexpression did not influence *Irx3* levels, indicating that the two act in parallel without cross-regulation or feedback. Panel E shows that *IRX3*

and *IRX5* overexpression leads to reduced *UCP1* expression, measured in a luciferase assay using a reporter construct containing the human *UCP1* promoter fused to the luciferase gene. The reporter was co-transfected in ME3 beige preadipocytes with plasmid encoding human *IRX3* or *IRX5* (0.1, 0.5 or 1.0 µg). Repression is seen for both basal (grey, no isoproterenol and 9-cis retinoic acid) and stimulated (blue, isoproterenol and 9-cis retinoic acid), with a significantly more pronounced repression upon stimulation, consistent with reduced thermogenesis. n = 3. Error bars: SEM.

Figure S6. Regulatory circuitry of ARID5B, rs1421085, IRX3 and IRX5 in human adipocytes.

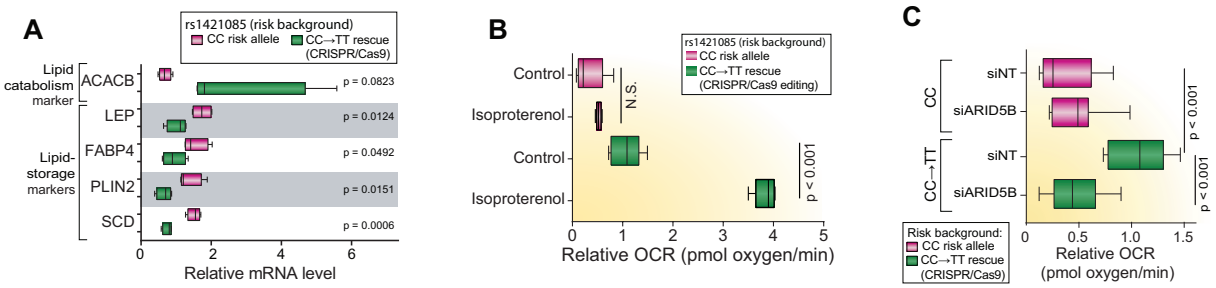


Panel A shows a multiple sequence alignment of the 1kb region surrounding the evolutionarily-conserved motif module disrupted by the predicted causal variant rs1421085. Panel B

shows the Phylogenetic Complexity Module Analysis (PMCA) scores for all single nucleotide variants (SNVs) in the *FTO* obesity locus, revealing rs1421085 as the top-scoring SNV for all tested parameters. TFBS=transcription factor binding sites. Panel C shows increased enhancer activity upon transfection into human SGBS preadipocytes (n=9) for 10kb and 1kb segments centered at rs1421085 upon introduction of the rs1421085 risk variant (red) on the non-risk haplotype (blue), but no change for 100bp. Panel D shows orientation- and position-independent gain of enhancer activity for the 1kb region centered on rs1421085, upon T-to-C single-nucleotide editing of the non-risk allele in a luciferase enhancer reporter assay fused to a basal pGL4.22-TK promoter in SGBS adipocytes (n = 9). Panel E shows increased binding for the non-risk T allele (intact ARID5B mtoif) vs. the risk C allele (disrupted ARID5B motif) using electrophoretic mobility shift assays (EMSA) using human adipocytic nuclear extract. Binding in the T allele is competed away by increasing amounts of canonical ARID5B motif (AATAT) unlabeled competing probe. The C allele shows no binding (no shifted probe band) even without unlabeled probe competition. Panel F shows whole tissue and adipose stromal cell expression for *ARID5B*, *LHX6*, *NKX6-3* and other ARID family members based on Illumina microarrays for lean (n=13) and obese (n=17) participants. Panel G shows inverse correlation between *ARID5B* and *IRX3*/*IRX5* expression using *HPRT*-normalized qPCR in adipocyte progenitors specifically from non-risk participants (n=18), but no correlation for risk participants (n=20). Panel H shows increased *IRX3* and *IRX5* levels upon siRNA knock-down of ARID5B in primary adipocytes progenitor cells non-risk participants (n=18), bringing expression to risk levels, but no change for risk participants (n=20). Panel I shows loss of enhancer activity for the 1kb region surrounding rs1421085 upon C-to-T single-nucleotide editing rescue of the C risk allele (restoring the disrupted ARID5B motif). Enhancer activity measures luciferase expression relative to a basal promoter in SGBS adipocytes (n=9). Knock-down of the ARID5B repressor has no effect on enhancer activity in the risk haplotype (lacking the motif), but it leads to loss of repression in the edited construct. This *cis*-/*trans*- conditional analysis indicates that enhancer repression occurs only when both the intact motif and the intact factor are present (green, siNT), and absence of either the repressor (siARID5B) or the repressor motif (red) leads to loss of repression, indicating causal role for the ARID5B repressor and the rs1421085 nucleotide variant. Panel J shows endogenous decrease of *IRX3* and *IRX5* mRNA levels upon ARID5B overexpression in homozygous non-risk (n=18) primary human adipocyte progenitor cells, indicating moderately increased repression. No change is seen in risk allele (n=20) carriers (disrupted ARID5B motif). Panel K shows that knock-down of the ARID5B repressor leads to reduced basal oxygen consumption rate (OCR) in primary human adipose cells from non-risk (n=18) participants (intact ARID5B motif), resulting in risk-allele levels. ARID5B knock-down has no effect on risk allele (n=20) participants (disrupted ARID5B motif). This *cis-/trans*- conditional analysis indicates

causality of the ARID5B upstream repressor acting through the non-risk haplotype. Panel L shows that knock-down of the ARID5B repressor leads to reduced lipid catabolism (catecholamine-stimulated lipolysis rate) estimated by glycerol concentration in the culture medium after isoproterenol stimulation vs. unstimulated (basal) in primary human adipose cells from non-risk (n=18) participants. No change is seen in risk participants (n=20). This *cis-/trans-* conditional analysis indicates causality of the ARID5B upstream repressor acting through the non-risk haplotype. Panel M summarizes our regulatory model: in the non-risk allele (blue), the ARID5B repressor binds an intact DNA motif and represses enhancer activity. In the risk allele (red), and T-to-C single-nucleotide variant at rs1421085 disrupts the ARID5B motif, leading to loss of DNA binding, activation of the pre-adipocyte enhancer, and activation of the downstream target genes *IRX3* and *IRX5*.

Figure S7. Phenotypic effects of rs1421085 circuitry on oxygen consumption and glycerol release.



Panel A shows increased endogenous expression of lipid catabolism marker *ACACB* and reduced expression of lipid storage markers upon CRISPR/Cas9 single-nucleotide C-to-T genome editing of the rs1421085 risk allele, indicating a shift from energy storage to energy dissipation. Panel B shows increased basal (control) and stimulated (isoproterenol) respiration (oxygen consumption rate) in differentiating primary human adipocyte progenitors upon CRISPR/Cas9 T-to-C editing of rs1421085 rescuing the risk allele (n=5 clonal expansions), indicating that the single-nucleotide change is sufficient to rescue isoproterenol-stimulated increase in cellular energy expenditure indicative of uncoupled respiration and thermogenesis. Panel C shows increased respiration (oxygen consumption rate, OCR) upon C-to-T single-nucleotide rescue of rs1421085 in primary adipocytes from risk-allele participants by CRISPR/Cas9 genome editing of the endogenous locus, indicating single-nucleotide variant causality. Knock-down of ARID5B has no effect on risk-allele carriers (red), but results in decreased respiration in the edited-allele adipocytes (green), indicating causality of the ARID5B repressor acting through the rs1421085 non-risk allele. n = 10 clonal expansions were induced to differentiate into adipocytes concomitant with siRNA transfection and analyzed after two days.

Table S1. Enrichment in functional categories for genes co-expressed with IRX3 and IRX5 (Ingenuity Toxicity Lists)

Categories with enrichment p-value < 0.05 are shown. Gene expression was measured by Affymetrix microarrays in human perirenal white adipose tissue containing brown adipocytes (10 healthy persons).

NEGATIVE CORRELATION WITH IRX3	p-value	Correlated genes (Pearson's r > 0.5 or < -0.5)
Mitochondrial Dysfunction	1.47e-12	MAP2K4, NDUFA9, UQCR11, NDUFB5, ACO2, NDUFB8, MT-CO2, VDAC2, NDUFB10, PDHA1, NDUFS1, NDUFA5, MAOB, NDUFA10, UQCRFS1, GPX4, ATP5F1, NDUFA8, ATP5J, COX7C, SDHC, COX7A1, PRDX3, ATP5B, TXN2, NDUFA6, UQCRC2, MAPK10, NDUFA12, MT-ND3, COX7B, ATP5H, COX7A2L, ATP5S, NDUFA2, MT-CYB, NDUFAB1, ATP5J2, ATPAF1, MT-CO3, NDUFB6, PARK2, OGDH, COX4I1, NDUFS4, AIFM1, MT-CO1, NDUFAF1, UCP2, ATP5O, MAPK8, MAPK9, LRRK2, VDAC3, NDUFS3, UQCRB, MT-ND5, MT-ND1, COX11, ATP5L2, NDUFV2, UQCR10, CYC1, CAT, SDHD, COX15
TR/RXR Activation	0.003	AKT2, PIK3CA, AKR1C1/AKR1C2, UCP2, MED1, NCOA6, THRA, ME1, PIK3R4, THRSP, F10, NCOA2, PDE3B, PIK3C3, NCOA1, NCOA4, ACACA, STRBP, PIK3CB, TBL1XR1, THRB, RXRB, ATM, PPARGC1A
RAR Activation	0.004	MAP2K4, PRKACB, NSD1, PIK3CA, ADH1C, PDPK1, SMAD5, MNAT1, PRKAG1, TRIM24, ALDH1A1, KAT2B, CSNK2A1, GTF2H5, SMAD4, RXRB, PRKD3, CITED2, PNPLA4, GTF2H3, SMAD2, AKT2, GNAS, MED1, NRIP2, MAPK8, MAPK9, SNW1, PARP1, CSNK2A2, MAPK14, PRKAR2B, SMARCA2, PRKCD, ERCC3, MAPK10, NCOA1, PRKACA, PIK3CB, HLTF, PPARGC1A, RBP4
Biogenesis of Mitochondria	0.008	PTCD2, PRDX3, GNAS, TFAM, CAV2, CEBPA, ADRB3, PPARGC1A
Fatty Acid Metabolism	0.009	ADH1C, ACAA2, HADHB, ALDH2, ADH1A, ALDH1A1, ECI2, ALDH3A2, ACAD8, ACADM, HSD17B4, HADHA, ALDH5A1, ALDH7A1, ECHS1, ACAA1, ACOX1, ECH1, ALDH9A1, ACADS, ACADL, AUH, ACAT1, CPT2, ADH1B, IVD, ADHFE1, HADH, ACSL1
POSITIVE CORRELATION WITH IRX3		

Primary Glomerulonephritis Biomarker Panel (Human)	3.7e-05	TYMS, ELF3, EGR1, HBEGF, IER3, SERPINE1, SAMD4A, MAFF, MCL1
Cardiac Necrosis/Cell Death	4.1e-05	SOCS3, LIF, SPRR1A, XDH, IL6, NOS3, GHRH, NTN1, mir-154, BBC3, SOD2, PIM1, JUND, FOSL1, CAV3, ABCB8, MCL1, RRAD, LCN2, TRIM54, STAT3, CSF3, THBD, INHBA, BACH1, DAXX, MPO, APLN, PPP1R15A, E2F1, THBS2, ZYX, MAP2K3, MDK, CYR61, CDK2, WNT1, mir-21, PVRL2, EPO, MYH6, LEP, PTK2B, HSPB8, MYLK3, TNFAIP3, SIRT7, HIF1A, CAMK2N2, E2F3, MAPK11, mir-34, FAS, RNASE3, IRAK1, NRG1, FSTL1, TIMP1, UCN, MANF, SST, NAMPT, MAPKAPK2, MAP2K1, GHRL, PLAT, IFNG, ADAMTS13, PPIF, CXCR4, MT1A, BIRC5, IL17A, STAR, GCG, IVNS1ABP, GNAI2, APOA1, CDKN1A, BRCA2, CYP2J2, NPPA, CXADR, MYBPC3, CALCA
Hepatic Cholestasis	4.46e-05	GCGR, LIF, IL6, SLC22A6, CYP8B1, ATP8B1, IL1RL2, NR0B2, RARA, PPRC1, ABCC1, OSM, NFKBIB, HNF4A, IL17F, IL1RAP, FGF19, TNFRSF11B, MAP3K14, CXCL8, PRKCQ, SLCO1A2, NFKB2, MAPK12, PRKACG, IL36RN, TGFB3, ESR1, RELA, IL1A, IL12A, IL17C, NFKBIE, IL1F10, SLC22A7, HNF1A, PRKCZ, IRAK1, IL17D, NGFR, SLC10A2, CYP7A1, PRKCE, LBP, TNFRSF1B, ABCB1, IFNG, TNFRSF1A, RELB, IL1R1, IL17A, GCG, TRAF2, SLCO3A1, LTA, ADCY1, ABCC3, IL11
Cardiac Hypertrophy	1.33e-04	CHGA, CTGF, TNNT2, CACNA1H, HRAS, IL6, NOS3, ATP2A2, EEF1D, CABIN1, FHL2, SOD2, GAL, TCAP, GAA, PNKD, TXN, SERPINE1, AHR, CACNA1C, HBEGF, MYH7, HMGA1, STAT3, IER3, CSF3, DES, MNT, AKAP13, TRIM63, IRX4, MAP2K3, MAPK7, AGTR2, mir-21, EPO, MB, GUCY2C, HSPB8, MYLK3, SIRT7, DUSP5, UCN, FGF23, PRKCE, VDR, LTK, SCNN1B, MMP1, HTR2A, PLAT, RGS2, PTGIR, GRB2, XIRP1, mir-30, MYOCD, CDKN1A, ADORA2A, INHA, LIF, SMAD3, FKBP1B, PIM1, OSM, CAV3, MIF, ATF3, RRAD, IL6R, TRIM54, INHBA, LMCD1, ANKRD1, ADRA2A, DUSP1, APLN, mir-210, SMTN, ABCC4, S100A6, ANGPT2, MYH6, LEP, PDGFA, DACT1, TNFAIP3, MKL1, HIF1A, NFATC1, KCNJ11, HTR2C, SHC1, TIMP1, GATA6, MAP2K1, RASSF1, VAV2, EPAS1, UCN2, TRIM55, PPARD, NFATC4, BIRC5, GNAI2, RCAN1, NFATC2, ADRA2C, CYP2J2, NPPA, RCAN2, MYBPC3, IL11

Liver Necrosis/Cell Death	6.81e-04	mir-146, SOCS1, SOCS3, INHA, CD40LG, LIF, SMAD3, HRAS, IL6, CCND1, PTGER1, MYC, F11, CXCL3, SOD2, BBC3, NR0B2, CPB2, OSM, JUND, IGFBP1, HNF4A, SERPINE1, NFKBIB, AHR, MCL1, HRH2, SELE, IL6R, HBEGF, STAT3, IER3, INHBA, MNT, DAXX, SELP, E2F1, FGA, ESR1, EPO, SIGIRR, RELA, GADD45B, SLC26A1, UGCG, FKBP1A, SMPD1, IRF3, USP2, FAS, TACR1, SHC1, FGF4, FLNA, TIMP1, PROS1, NGFR, CIT, NR1I3, PTPN1, CYP7A1, SST, TNFRSF1B, MAP2K1, RHBDD3, SLC20A1, MYCN, IFNG, IL10, TNFRSF1A, IL15, XBP1, PLK1, BAX, IL17A, BUB1, PLG, LDLR, ABL2, CDKN1A, MAP3K8, TNFSF14
Cardiac Fibrosis	8.39e-04	SMAD3, TNNT2, CACNA1H, HRAS, NOS3, ATP2A2, F11, HTR1B, SOD2, PDCD1, NDUFS6, PNKD, SERPINE1, CAV3, AHR, PTX3, ATF3, SNAI1, HBEGF, TNNI3, LMNA, CACNA1C, STAT3, CSF3, BUB1B, DES, DIO3, ADRA2A, THBS2, mir-21, AGTR2, PVRL2, EPO, MYH6, ANGPT2, TBXA2R, MYLK3, HSPB8, TNFAIP3, SIRT7, MAPK11, mir-34, TIMP1, VDR, RASSF1, PLAT, VAV2, IFNG, PTGIR, GRB2, EGR1, XIRP1, let-7, PLG, APOA1, mir-30, ADRA2C, MYBPC3
Increases Liver Damage	9.43e-04	LTB, SMPD1, IL6, F2, FAS, PRKCZ, PTGER1, NR0B2, TICAM1, PDCD1, PTPN1, CYP7A1, OSM, LBP, SERPINE1, TNFRSF1B, ALDH3A1, FGR, PLAT, IFNG, MIF, TNFRSF1A, EGR1, IL6R, XBP1, STAT3, IL1R1, GH1, CSF3, IL17A, PLG, FOXA2, LTA, CDKN1A, CD44, MAP2K3, GSTP1
Increases Liver Hyperplasia/Hyperproliferation	0.002	S100A6, LEP, LIF, HMMR, NCAPG, LTB, HRAS, TLX1, CXCL5, IL6, CCND1, TGFBR2, MYC, mir-26, NR1I3, OSM, GHRL, AHR, FGF19, CXCR4, PCSK9, RRM2, AGO2, STAT3, FOXM1, BIRC5, IGFBP2, HPN, BUB1, LTA, E2F1, CENPA, mir-192, TGFA, mir-17, IL11
Increases Renal Nephritis	0.003	IFNG, CD19, CD40LG, LIF, LEP, TNFRSF1A, IL10, IL21R, COL4A3, IL17A, MPO, CSF1, PIM1, LTA, GP1BA, ICOS, TNFSF14
Liver Proliferation	0.005	mir-146, SOCS3, CTGF, XDH, SMAD3, HRAS, LTB, IL6, CCND1, F2, MYC, TGFBR2, CXCL3, RARA, CPB2, SLC7A5, OSM, OSMR, JUND, IGFBP1, HNF4A, AHR, FGF19, IL4R, IL6R, HBEGF, INHBA, E2F1, MDK, CDK2, mir-21, WNT1, EPO, LEP, SMPD1, HNF1A, FAS, TNFRSF12A, TIMP1, THBS1, PROS1, NGFR, NR1I3, CIT, GDF2, SULF2, TNFRSF1B, SLC20A1, IFNG, S1PR2, EDNRB, PPARD, CXCR4, TNFRSF1A, NPPC, XBP1, FOXM1, PLG, WNT3A, LY9, RETN, LTA,

		CDKN1A, CSHL1, TGFA
Increases Liver Hepatitis	0.008	IFNG, TNFRSF18, LEP, TNFRSF1A, IL10, LTB, IL6, STAT3, IL1R1, FAS, IL17A, CCL2, NR0B2, LTA, SERPINE1, TNFRSF1B, CYR61, ESR1, PDX1
Hepatic Fibrosis	0.010	IGFBP4, IL1A, CD40LG, ICAM1, LEP, PDGFA, SMAD3, COL4A3, VTN, IL6, TGFBR2, TNC, CCL2, ECE2, THBS1, TIMP1, IGFBP1, MMP1, AHR, CXCL8, IFNG, SDC1, IL10, BGN, IGFBP2, ACTA2, CSF1, THBS2, TGFA, TGFB3, mir-21
Hepatic Stellate Cell Activation	0.011	CXCL8, RELA, IFNG, IL10, PDGFA, TGFBR3, NFKB2, IL6, TGFBR2, CCL2, TIMP1, TGFA, TGFB3, KLF2
Genes associated with Chronic Allograft Nephropathy (Human)	0.024	TNC, VCAM1, ICAM1, CCL2, PDGFA, THBS1, TIMP1, MATN1, CCL19
Acute Renal Failure Panel (Rat)	0.034	GAS2, VCAM1, ATF3, RRAD, SLC9A3, EGR1, LCN2, FKBP1A, IL6, PRODH2, TNFRSF12A, MYC, ACTA2, CDKN1A, SPHK1, CD44, TAGLN, FOSL1, IGFBP1, COL18A1
Positive Acute Phase Response Proteins	0.044	C4A/C4B, SOD2, APCS, C9, CRP, LBP, SERPINE1, FGA, HMOX2, SERPINF2, FGG
FXR/RXR Activation	0.045	IL1A, APOA4, PKLR, APOF, FOXA1, APOA2, VTN, C9, AMBP, APOC2, IL1F10, SLC22A7, HNF1A, CYP8B1, C4A/C4B, SLC27A5, NR0B2, SLC51A, RARA, SLC10A2, CYP7A1, HNF4A, FGF19, SDC1, FETUB, AHSG, MAPK12, SULT2A1, SERPINF2, APOA1, ORM1, FOXA2, IL36RN, FGA, FOXA3, APOC3
Genes Downregulated in Response to Chronic Renal Failure (Rat)	0.045	CYP1A1, SLC22A12, SLC17A1, SLC22A7, SLC22A6

LXR/RXR Activation	0.045	RELA, IL1A, LPA, APOA4, APOF, APOA2, VTN, C9, AMBP, ABCG1, APOC2, IRF3, IL6, IL1F10, C4A/C4B, CCL2, IL1RL2, NGFR, CYP7A1, LBP, TNFRSF1B, IL1RAP, TNFRSF11B, TNFRSF1A, RELB, AHSG, NFKB2, IL1R1, SERPINF2, LDLR, APOA1, ORM1, IL36RN, FGA, APOC3
Decreases Depolarization of Mitochondria and Mitochondrial Membrane	0.049	ABCC1, CDKN1A, PLIN3, MAP2K1, ATP2A2, BIRC5, FAS, MCL1
NEGATIVE CORRELATION WITH IRX5		
Mitochondrial Dysfunction	0.008	COX7B, NDUFA9, NDUFB5, ATP5S, MT-CO2, ATP5E, NDUFA2, VDAC2, CASP9, GPD2, MT-CYB, NDUFAB1, ATP5J2, NDUFB6, ATPAF1, MT-CO3, PARK2, COX4I1, AIFM1, MT-CO1, SDHA, MAPK8, MAPK9, SDHC, VDAC3, GPX7, UQCRB, MT-ND1, MT-ND5, COX11, ATP5L2, NDUFV2, CAT, MAPK10, NDUFA12, SDHD, MT-ND3, UQCRQ
Aryl Hydrocarbon Receptor Signaling	0.014	TRIP11, POLA1, NQO2, RBL1, RARG, PTGES3, ARNT, SP1, ALDH3A2, GSTM4, ALDH6A1, ALDH5A1, GSTK1, ATM, MGST1, RBL2, MED1, GSTA4, GSTM3, CYP2C18, CDK6, MAPK8, ALDH9A1, RXRG, ALDH1L2, CYP3A4, NFIA, CYP3A43, NFIB, DHFR, DCT, CDKN1B, ESR1, MGST3, MCM7
TR/RXR Activation	0.022	PIK3C2B, AKT2, AKR1C1/AKR1C2, MED1, UCP1, ME1, THRSP, RXRG, PDE3B, SREBF1, PIK3C3, NCOA1, NCOA4, ACACA, STRBP, PIK3CB, TBL1XR1, THRB, PPARGC1A, ATM
POSITIVE CORRELATION WITH IRX5		
FXR/RXR Activation	0.002	G6PC2, IL1A, APOA4, SLC10A1, PKLR, APOH, FOXA1, APOA2, APOC4, VTN, C9, APOC2, IL1F10, HNF1A, SLC27A5, NR0B2, SLC51A, CYP19A1, RARA, SLC10A2, CYP7A1, GC, HNF4A, FGF19, TTR, SDC1, FETUB, IL36A, MAPK12, SULT2A1, A1BG, SERPINF2, APOA1, FABP6, FOXO1, ORM1, FOXA2, IL36RN, G6PC, SLC51B, FGA, FOXA3, APOC3

Hepatic Cholestasis	0.004	SLC22A6, IL1R2, IL1RL2, NR0B2, ADCY5, PPRC1, RARA, OSM, HNF4A, NFKBIB, IL1RAP, FGF19, CXCL8, IL36A, MAPK12, IL36RN, PRKACG, ESR1, RELA, IL1A, IL12A, SLC10A1, IL17C, IL1RL1, IL1F10, HNF1A, IL25, IL17B, IRAK1, NFKBIA, NGFR, SLC10A2, IL17D, CYP7A1, PRKAR1B, PRKCE, CHUK, ADCY8, TNFRSF1B, ADCY2, IL3, RELB, IL1R1, IRAK3, PRKCG, GCG, TRAF2, FABP6, IL12B, LTA, ADCY1, IL11
Vasopressin-induced Genes in Inner Medullary Renal Collecting Duct Cells (Rat)	0.017	AQP3, WNK4, WNK1, AQP4, AQP2
Increases Liver Hepatitis	0.023	TNFRSF18, LEP, LTB, STAT3, IL1R1, IL17RA, TBX21, IL25, NR0B2, IL12B, CNR2, LTA, BID, SERPINE1, TNFRSF1B, CYR61, ESR1, PDX1
Cytochrome P450 Panel - Substrate is a Vitamin (Human)	0.025	CYP24A1, CYP26C1, CYP26A1, CYP27B1
LXR/RXR Activation	0.027	RELA, IL1A, LPA, APOA4, APOH, ABCG4, IL1RL1, APOA2, APOC4, VTN, C9, APOC2, IRF3, IL1F10, IL1R2, IL1RL2, NGFR, CYP7A1, NOS2, TNFRSF1B, GC, IL1RAP, TTR, RELB, IL36A, APOA5, IL1R1, A1BG, SERPINF2, APOA1, ORM1, IL36RN, NCOR2, FGA, MMP9, APOC3, CCL7
Cardiac Hypertrophy	0.036	CHGA, TCF15, TNNC1, TNNT2, CACNA1H, CORIN, ATP2A2, SOD2, GAL, mir-15, TXN, SERPINE1, MYH7, STAT3, HMGA1, CSF3, DES, MYL9, AKAP13, IRX4, HPRT1, AGTR2, EPO, CA2, MB, GUCY2C, IL1RL1, HSPB8, MYLK3, SIRT7, mir-1, F2RL1, UCN, FGF23, PRKCE, VDR, LTK, SCNN1B, MMP1, HTR2A, ENDOG, ALOX15, SLC4A1, FLT1, MYOD1, XIRP1, mir-30, SLC6A4, ADORA2A, CSRP3, INHA, HAND1, FKBP1B, OSM, CAV3, ELN, MIF, HDAC4, RRAD, IL6R, RAPGEF3, INHBA, BMP10, LMCD1, ANKRD1, APLN, mir-210, SMTN, MMP7, LEP, TERT, MKL1, NFATC1, PGF, KCNJ11, SHC1, NFKBIA, CYP19A1, FOXO3, NOS2, ACTC1, RASSF1, TIMP3, EPAS1, TRIM55, CKM, PPARD, NFATC4, BIRC5, HOPX, NKX2-5, mir-208, ADRA2C, CYP2J2, RCAN2, COL9A2, MMP9, GATA4, MYBPC3, IL11
Protection from Hypoxia-induced Renal Ischemic Injury (Rat)	0.037	VEGFA, EPO, SLC2A1

Cardiac Fibrosis	0.040	TNNT2, CACNA1H, ATP2A2, HTR1B, SOD2, PDCD1, CNR2, SERPINE1, CAV3, PTX3, SNAI1, TNNI3, STAT3, CSF3, DES, TLR2, AGTR2, PVRL2, EPO, GPX1, TBXA2R, HSPB8, MYLK3, PPP1CB, SIRT7, MAPK11, mir-34, DTNA, mir-1, CYP11B2, VDR, NOS2, ACTC1, RASSF1, TIMP3, SLC4A1, LIMS2, XIRP1, let-7, F3, PLG, HOPX, mir-30, APOA1, CAPNS1, mir-208, SLC6A4, ADRA2C, CSRP3, MMP9, MYBPC3
Cardiac Necrosis/Cell Death	0.043	XDH, CTSG, GHRH, VEGFA, mir-154, CASR, SOD2, BAG1, ADCY5, CNR2, FOSL1, JUND, mir-320, CAV3, ABCB8, RRAD, LCN2, STAT3, CSF3, THBD, mir-144, BMP10, INHBA, BACH1, TLR2, BCL2L1, MPO, APLN, PPP1R15A, E2F1, ZYX, CYR61, WNT1, PVRL2, EPO, LEP, TERT, MYLK3, HSPB8, SIRT7, CAMK2N2, E2F3, MAPK11, RNASE3, mir-34, IRAK1, NRG1, FSTL1, mir-1, UCN, MANF, FOXO3, SST, CACNB2, MAPKAPK2, NOS2, ACTC1, GHRL, ADAMTS13, MT1A, BIRC5, STAR, GCG, NKX2-5, APOA1, WISP1, CYP2J2, CXADR, GATA4, MYBPC3, CALCA