Supplementary Information

BMI-anchored transcriptomics uncovers oxidative and inflammatory poles of obesity heterogeneity

SI Materials, Methods, and Results 3

Repeated subsampling analysis of gene-selection stability 3

Repeated subsampling with a fixed regularization parameter 3

Matched-sparsity sensitivity analysis 3

Gene-level selection frequency 3

Functional mapping of enriched GO terms 4

Adipose tissue histology and immunohistochemistry 8

Independent histological characterization of adipose depots in DU6, DU6P, and FZTDU mice 9

Sensitivity analyses for BMI-range restrictions 9

Gene set curation and enrichment framework for mitochondrial metabolism, lipid biology, and hormone signaling 9

Multivariable logistic regression model for classifying the insulin status of the separate validation dataset 10

Prediction-oriented nested cross-validation analysis 10

Enrichment of primary-model genes among prediction-selected genes 11

Sex-adjusted sensitivity analysis 11

Primary modeling cohort 12

Separate validation cohort 12

SI Figures 13

Figure S1. Measured BMI distribution in human visceral adipose tissue samples. 13

Figure S2. Number of selected genes under repeated subsampling of the primary modeling cohort. 14

Figure S3. Circular representation of relative enrichment contributions of individual functions grouped by biological category in model-selected genes. 15

Figure S4. Independent histological characterization of adipose depots in DU6, DU6P, and FZTDU mice. 16

Figure S5. Robustness of cross-species BMI-anchored transcriptomic scores to alternative BMI-range definitions used for human model construction. 17

Figure S6. Circular representation of relative enrichment contributions of individual functions grouped by biological category in DU6 mice. 18

Figure S7. Circular representation of relative enrichment contributions of individual functions grouped by biological category in DU6P mice. 19

Figure S8. Circular representation of relative enrichment contributions of individual functions grouped by biological category in MHO-like patients. 20

Figure S9. Circular representation of relative enrichment contributions of individual functions grouped by biological category in MUO-like patients. 21

Figure S10. Prediction-oriented evaluation of the BMI-anchored transcriptomic framework. 22

SI Tables 23

Table S1. Functional enrichment of mitochondrial complexes, regulatory machinery, and lipid metabolism pathways in DU6P and DU6 mouse lines. 23

Table S2. Rule-based functional category assignment for GO biological process terms. 24

Table S3. Curated functional gene sets used for DU6 and DU6P up-regulated DEG enrichment analysis. 24

Table S4. List of primary model-selected genes 24

References 25

SI Materials, Methods, and Results

Repeated subsampling analysis of gene-selection stability

To evaluate whether the genes retained by the BMI-anchored transcriptomic model represented a reproducible feature-selection preference rather than a sample-specific sparse solution, we performed repeated subsampling analyses using the primary modeling cohort. The analysis included the same 1,298 participants and 16,042 high-confidence human–mouse orthologous genes used for construction of the primary model. The full-cohort LASSO model, fitted using α = 0.123, retained 458 genes with non-zero coefficients.

Repeated subsampling with a fixed regularization parameter

Participants were randomly sampled without replacement at 40%, 50%, 60%, 70%, 80%, and 90% of the primary modeling cohort, corresponding to 519, 649, 779, 909, 1,038, and 1,168 participants, respectively. For each sampling fraction, 500 repeated subsamples were generated. A LASSO model was fitted within each subsample using the same regularization parameter as the primary model (α = 0.123). For iteration i, the selected feature set S_i was defined as:

$$S_i = \{g : \beta_{ig} \neq 0\}$$

where β_{ig} denotes the LASSO coefficient of gene g in subsampling iteration i. The number of genes selected in the iteration i is defined as $K_i = |S_i|$. Model size was summarized across the 500 iterations using the median and interquartile range (IQR). As expected under a fixed α, progressively smaller subsamples generally produced smaller sparse models (Fig. S2), motivating an additional analysis in which model sparsity was approximately controlled.

Matched-sparsity sensitivity analysis

To determine whether preferential reselection of the primary genes was simply a consequence of differences in the total number of selected genes, we performed a matched-sparsity sensitivity analysis. For each sampling fraction, a single fraction-specific α value was calibrated using repeated subsamples to obtain a median number of selected genes as close as possible to the 458-gene size of the primary model. The calibrated α was subsequently fixed and applied to 500 repeated subsamples at the corresponding sampling fraction. Across the 40–90% sampling fractions, the resulting median model sizes ranged from 455 to 461 genes, closely approximating the 458-gene primary model (Fig. S2). Thus, comparison between the fixed-α and matched-sparsity analyses distinguishes whether observed gene-selection behavior is attributable to changes in overall model size or whether the relative preference for primary model-selected genes persists after approximately controlling for model sparsity.

Gene-level selection frequency

For each gene g, selection frequency was defined as the proportion of the 500 subsampling iterations in which the gene received a non-zero LASSO coefficient:

$$f_g = \frac{1}{500}\sum_{b=1}^{500} I,$$

$$I = \begin{cases} 1, g \in S_b \\ 0, g \notin S_b \end{cases}$$

where I is an indicator function equal to 1 when gene g is selected and 0 otherwise. In other words, f_g is the number of iterations in which gene g was selected divided by 500. Thus, f_g = 1 indicates that a gene was selected in all 500 models, whereas f_g = 0 indicates that it was never selected. Selection-frequency distributions were compared between the 458 genes selected by the primary full-cohort model and the remaining 15,584 candidate genes. Genes that were never selected in any subsampling iteration were retained because their absence from the resampled models is informative regarding relative selection preference.

The primary genes showed markedly higher selection frequencies than the remaining candidate genes. In the fixed-α analysis, their mean selection frequency increased from approximately 0.20 at 40% subsampling to 0.66 at 90% subsampling, whereas the mean frequency among other genes remained approximately 0.01–0.02 (Fig.

1b). A similar separation was observed after matching model sparsity (Fig. 1c). *P*-values were adjusted separately within the fixed-α and matched-sparsity analyses across the six evaluated sampling fractions using the Holm procedure. These results indicate preferential reselection of the primary gene set. However, the low selection frequency appearing in the analyses of smaller subsampling showed that not every primary gene was consistently recovered in every subsample.

Functional mapping of enriched GO terms

To determine whether the enriched biological processes were relevant to obesity, we systematically mapped the enriched GO terms from the GSEA analysis to eleven predefined functional themes: (1) obesity/adiposity, (2) immune/inflammatory, (3) lipid metabolism/catabolism, (4) other metabolism, (5) energy production/consumption, (6) GH–IGF axis, (7) cell cycle/division, (8) cytoskeleton/adhesion, (9) development/signaling, (10) autophagy/protein quality control, and (11) miscellaneous.

Gene Ontology biological process terms were assigned to functional categories using a deterministic, rule-based text classification procedure based on the GO term description. For each enriched GO term, the description was searched against a curated set of case-insensitive regular expression patterns. The classification was performed sequentially according to a fixed category priority order. If a GO term matched patterns from more than one category, the first matching category in the ordered rule set was assigned. The classification used only the GO term description and did not use row-by-row manual reassignment.

The following rules were applied in the order shown below. Pattern matching was case-insensitive. Notably, regular expression patterns were written using standard regex syntax. Parentheses were used to group terms, and the question mark symbol "?" indicated that the preceding group was optional. For example, the pattern:

(positive)?regulation of antigen receptor-mediated signaling

matches both "regulation of antigen receptor-mediated signaling" and "positive regulation of antigen receptor-mediated signaling", because the grouped phrase positive , including the trailing space, is optional.

1. obesity/adiposity

GO terms were assigned to **obesity/adiposity** if their descriptions matched one of the following patterns:

cellular response to (hexose|glucose|carbohydrate)
glucose (metabolic process|starvation)
extracellular matrix disassembly
cytokinesis
negative regulation of cytokine production involved in inflammatory response

2. immune/inflammatory

GO terms were assigned to i**mmune/inflammatory** if their descriptions matched immune activation, inflammatory response, leukocyte, cytokine, antigen-processing, or host-defense-related patterns:

Interleukin
interleukin-\d+
tumor necrosis factor
inflammatory response
neuroinflammatory
chemokine
(leukocyte|granulocyte|neutrophil|macrophage) chemotaxis
leukocyte
lymphocyte
macrophage
mast cell
myeloid leukocyte
toll-like receptor
pattern recognition receptor
nf-kappa
antigen processing and presentation

(positive)?regulation of antigen receptor-mediated signaling
(positive)?regulation of T cell receptor signaling
negative regulation of B cell apoptotic process
regulation of B cell apoptotic process
B cell apoptotic process
symbiont-mediated
CCR7
Pyroptotic
defense response to bacterium

3. lipid metabolism/catabolism

GO terms were assigned to l**ipid metabolism/catabolism** if their descriptions matched lipid, fatty-acid, sterol, cholesterol, phospholipid, sphingolipid, or catecholamine/dopamine metabolic processes:

lipid
fatty acid
fatty acid beta-oxidation
beta-oxidation
acyl-CoA
phospholipid
glycerophospholipid
glycolipid
glycosphingolipid
sphingolipid
sterol
cholesterol
monocarboxylic acid biosynthetic
catecholamine metabolic
dopamine metabolic

4. energy production/consumption

GO terms were assigned to **energy production/consumption** if their descriptions matched mitochondrial, respiratory-chain, oxidative phosphorylation, electron-transport, ATP synthesis, oxygen transport, pyruvate-to-acetyl-CoA, or GTP-related processes:

GTP metabolic
small GTPase-mediated signal transduction
mitochondrial depolarization
mitochondrial respiratory chain complex
respiratory electron transport
oxidative phosphorylation
electron transport chain
mitochondrial electron transport
ATP synthesis coupled electron transport
proton motive force-driven
oxygen transport
cytochrome complex assembly
cytochrome c from mitochondria
pyruvate decarboxylation to acetyl-CoA

5. GH–IGF axis

GO terms were assigned to **GH–IGF axis** if their descriptions matched hormone receptor signaling, steroid hormone response, MAPK/JNK activity, receptor tyrosine kinase signaling, Wnt signaling, or growth-factor-related pathways:

determination of left/right symmetry
steroid hormone receptor signaling
estrogen receptor signaling

nuclear receptor-mediated steroid hormone signaling
cellular response to steroid hormone stimulus
MAP kinase activity
JUN kinase activity
ERBB
Wnt signaling
fibroblast growth factor
transforming growth factor beta
epidermal growth factor
vascular endothelial growth factor
platelet-derived growth factor

6. cell cycle/division

GO terms were assigned to **cell cycle/division** if their descriptions matched cell-cycle, mitotic, meiotic, chromosomal segregation, chromatid, spindle, nuclear division, checkpoint, nonhomologous end joining, rRNA maturation, or related division-associated terms:

cell cycle
mitotic
mitosis
meiosis
chromosome (segregation|separation|organization)
chromosome segregation
chromatid
protein localization to (condensed)?chromosome
spindle
nuclear division
G1/S
G2/M
metaphase
anaphase
checkpoint
alternative nonhomologous end joining
single-strand annealing
nucleolar large rRNA
LSU-rRNA

7. cytoskeleton/adhesion

GO terms were assigned to **cytoskeleton/adhesion** if their descriptions matched cytoskeletal organization, microtubule, actin, adhesion, adherens junction, desmosome, cortical cytoskeleton, or cell projection-related processes:

cytoskeleton
microtubule
actin
cell-matrix adhesion
cell-substrate adhesion
adherens junction
desmosome
cortical cytoskeleton
cell projection

8. development/signaling

GO terms were assigned to **development/signaling** if their descriptions matched angiogenesis, Notch signaling, neuronal or synaptic processes, neurotrophic signaling, axon development, brain-region development, epiboly, muscle development, oligodendrocyte development, or Rap protein signaling:

sprouting angiogenesis

Notch receptor
neuron
neurotrophic
axon
synapse
presynapse
olfactory bulb
forebrain
epiboly
cardiac muscle tissue
ventricular cardiac muscle tissue
striated muscle tissue development
muscle tissue (development|morphogenesis)
oligodendrocyte development
Rap protein signal transduction

9. autophagy/protein quality control

GO terms were assigned to **autophagy/protein quality control** if their descriptions matched autophagy, autophagosome, macroautophagy, lytic vacuole, vacuole organization, lysosome organization, ubiquitin/proteasome pathways, protein quality control, misfolded protein response, unfolded protein response, or protein localization to lysosome:

protein polymerization
autophagy
macroautophagy
autophagosome
lytic vacuole organization
vacuole organization
lysosome organization
ubiquitin
ubiquitination
ubiquitin-protein transferase
ubiquitin protein ligase
proteasome
protein quality control
misfolded protein
unfolded protein response
protein localization to lysosome

10. other metabolism

GO terms were assigned to **other metabolism** if their descriptions matched non-lipid metabolic processes, including amino-acid, nucleotide, ribonucleotide, purine, RNA, rRNA, mRNA, carboxylic acid, organic acid, phosphate, phosphorus, thioester, ion homeostasis, acetyl-CoA, pyruvate, ATP, TCA-cycle-related, NAD-related, glycoprotein, reactive oxygen species, phosphatidylinositol phosphate, skeletal muscle contraction, or osteoclast differentiation terms:

polyamine
spermidine metabolic
ribonucleotide
deoxyribonucleoside
purine
tRNA pseudouridine synthesis
ribose phosphate
RNA (biosynthetic|metabolic)
rRNA metabolic
mRNA metabolic

nucleic acid biosynthetic
nucleotide
nucleoside phosphate metabolic
nucleotide-sugar
tRNA aminoacylation
amino acid (activation|transport|metabolic|biosynthetic)
proteinogenic amino acid
cellular response to amino acid
oxoacid
carboxylic acid
organic acid biosynthetic
monocarboxylic acid (metabolic|catabolic)
dicarboxylic acid
phosphate metabolic
phosphorus metabolic
secondary alcohol
thioester
intracellular sodium ion homeostasis
regulation of calcium ion transport$
acetyl-CoA (metabolic|biosynthetic)
pyruvate metabolic
ATP metabolic
tricarboxylic acid metabolic
NAD+ metabolic
glycoprotein metabolic
positive regulation of reactive oxygen species metabolic process
phosphatidylinositol phosphate biosynthetic
skeletal muscle contraction
osteoclast differentiation

11. miscellaneous

GO terms that did not match any of the above category-specific rules were assigned to **miscellaneous**.

For each GO term, the *z*-score of the enrichment score was used to calculate its relative contribution within each dataset. The term-level proportion was calculated as:

$$Proportion_i = \frac{z_i}{Z} \times 100$$

where z_i is the z-score of term i, Z is the sum of the z-score values of all GO terms in the corresponding dataset. Category-level enrichment contribution was calculated by summing the z-score values of all GO terms assigned to the same functional category and dividing by the Z. In addition, the one-to-one mapping of GO terms to functional categories is provided in Table S2.

Adipose tissue histology and immunohistochemistry

Paraffin-embedded white adipose tissue (WAT) was sectioned into 10 µm slices, dewaxed, dehydrated in an alcohol series, and stained with hematoxylin (Carl Roth, T865.1) and eosin (Carl Roth, 3137.1).

For immunohistochemical detection of F4/80 as a macrophage marker a rat monoclonal antibody anti-mouse F4/80 was used (Cat.-No: 123101, Biolegend, 1:400 dilution). For immunohistochemical detection of macrophages, tissue sections were deparaffinized, and quenched with 3% hydrogen peroxide for 10 min. Slides were heat retrieved with citrate buffer at pH 6.1 (Catalog S1699, DakoCytomation) for 4 min at 123°C, then cooled for 15 min, followed by a running tap water rinse. Sections were then incubated with the anti-F4/80 antibody at room temperature for 60 min and rinsed with TBS. F4/80-stained sections were incubated for 30 min in Rat Probe and then Rat Polymer-HRP (Rat-on-Mouse HRP Kit, Cat-No: RT517H, Biocare Medical/Zytomed systems). Slides were rinsed with TBS and developed with DAB-Plus (DakoCytomation), counterstained in

modified Mayer's hematoxylin and blued in 0.3% ammonia water, followed by a tap water rinse[1].

Brightfield images were taken with a Keyence BZ-X (Keyence Corporation of America) at 20× magnification (for adipocyte size determination: FZTDU n = 9 epigonadal, n = 7 inguinal, DU6 n = 12 both inguinal and epigonadal, DU6P n = 6 both inguinal and epigonadal; for evaluating F4/80 staining/crown-like structures: FZTDU n = 6 for epigonadal and n = 4 inguinal, DU6 n = 12 for both epigonadal and inguinal, DU6P n = 4 epigonadal and n = 1 inguinal). For adipocyte size measurements, three images per sample were analyzed; for crown-like structures, at least two images per sample were analyzed. Adipocyte size was measured using the ImageJ[2] plugin Adiposoft[3]. Crown-like structures were counted manually and normalized to adipocyte number as determined by Adiposoft analysis. Statistically significant differences among three groups were tested with one-way Brown-Forsythe and Welch ANOVA followed by Dunnett T3 correction for multiple comparisons and differences between two groups were tested with Mann–Whitney non-parametric test and visualized using GraphPad Prism 10.

Independent histological characterization of adipose depots in DU6, DU6P, and FZTDU mice

To validate our findings at the transcriptional level, we performed histological analyses on epigonadal and inguinal WAT (Supplementary Fig. S4a,b). Adipocyte size measurement revealed that DU6 mice exhibited pronounced hypertrophy across both depots; epigonadal adipocytes were 2.3-fold larger than FZTDU controls ($p < 0.0001$) and 3.8-fold larger than those of DU6P ($p < 0.0001$). This hypertrophic phenotype was more evident in the inguinal depot, where DU6 cells were 4.9-fold larger than FZTDU ($p = 0.0010$) and 6.3-fold larger than DU6P ($p = 0.0007$). In contrast, DU6P mice maintained a lean cellular profile, with epigonadal adipocytes being significantly smaller than those of the FZTDU control group ($p = 0.0123$). Notably, in the inguinal depots of both FZTDU and DU6P mice, we identified clusters of multilocular cells resembling beige adipocytes (indicated by arrows in Supplementary Fig. S4b), which were absent in DU6 tissue, suggesting that high-fat selection may impair the browning potential of subcutaneous fat. Inflammatory infiltration was assessed via F4/80 immunohistochemical staining to visualize crown-like structures (CLS), which represent macrophages surrounding dying adipocytes (Supplementary Fig. S4c,d). In the epigonadal depot, no statistically significant differences in CLS number were observed between the three mouse lines. However, in the inguinal depot, DU6 mice displayed a significant increase in CLS compared to FZTDU, with the median number being 3.8-fold higher ($p = 0.0022$). This increase reflects a localized pro-inflammatory response associated with adipocyte size expansion seen in the DU6 line. Due to a lack of available tissue, only a single sample was analyzed for the DU6P inguinal group, preventing its inclusion in the formal statistical comparisons for this specific site.

Sensitivity analyses for BMI-range restrictions

To assess the robustness of the BMI-range restrictions used for primary model construction, the BMI-anchored transcriptomic model was refitted using alternative cohort definitions while retaining the same preprocessing and Spearman's ρ-guided model-selection procedure. Specifically, we evaluated models trained using participants with (i) BMI < 60 and (ii) BMI ≥ 30. Cross-species robustness was assessed by applying each refitted model to the three mouse lines and comparing the resulting BMI-anchored transcriptomic scores and pairwise effect sizes with those obtained using the primary 30 ≤ BMI < 60 model.

Gene set curation and enrichment framework for mitochondrial metabolism, lipid biology, and hormone signaling

To systematically evaluate the biological processes represented in the differentially expressed genes (DEGs) between DU6 and DU6P, we curated functional gene categories relevant to mitochondrial metabolism, lipid biology, and hormone signaling. These categories were not adopted wholesale from a single pathway or publication. Instead, each category was assembled by integrating established pathway and annotation resources, including KEGG[4], Reactome[5], Gene Ontology[6], MSigDB Hallmark collections[7], MitoCarta3.0[8], and MGI nomenclature[9], followed by literature-guided manual refinement to ensure that the final gene sets reflected the biological processes being tested. For example, oxidative phosphorylation categories were curated primarily from mitochondrial and respiratory-chain annotations, whereas lipid synthesis/storage and mitochondrial translation

categories were further refined using studies and reviews focused on triacylglycerol metabolism, lipid droplet biology, sterol handling, and mammalian mitochondrial protein synthesis. Gene symbols were harmonized to official Mus musculus nomenclature using MGI to ensure consistency across datasets and pathway sources. The definitions of each category are described below, and the complete gene membership used for the enrichment analyses is provided in Supplementary Table S3.

Respiratory chain complexes (Complexes I–V). Subunit genes of the five oxidative phosphorylation complexes were obtained from the MitoCarta3.0 database and cross-referenced with KEGG OXPHOS pathway annotations. Complex I (NADH dehydrogenase) contained 44 nuclear-encoded subunits, Complex II (succinate dehydrogenase) 4 subunits, Complex III (ubiquinol–cytochrome c reductase) 11 subunits, Complex IV (cytochrome c oxidase) 18 structural subunits and 17 assembly factors, and Complex V (ATP synthase) 24 subunits.

Mitochondrial translation machinery. Nuclear-encoded mitochondrial ribosomal proteins (MRPs; 79 genes) and translational activators (11 genes including *Mss51, Lrpprc, Ptcd1*) were obtained from MitoCarta3.0 and recent mitochondrial translation studies[10-12].

Mitochondrial regulatory machinery. Genes regulating organelle biogenesis, calcium homeostasis, protein import, and transcriptional control (52 genes) were defined by merging mitochondrial biogenesis regulators (*Ppargc1b, Sirt1*), import machinery (TIM/SAM components such as *Dnajc19, Samm50*), calcium uniporter regulators (*Mcur1*), and mitochondrial translational activators.

Fatty-acid oxidation (FAO) axis. Genes encoding β-oxidation enzymes, electron transfer flavoproteins (ETF), and peroxisomal fatty-acid metabolism proteins (41 genes) were collated from KEGG FAO pathway, MitoCarta3.0, and the PEX gene family. This unified "FAO axis" captured total fatty-acid breakdown capacity, including mitochondrial FAO (*Acadvl, Cpt2, Echs1*), ETF coupling (*Etfb*), and peroxisomal lipid metabolism (*Pex11g, Pex26*).

Lipid synthesis and storage. Genes involved in de novo fatty-acid synthesis, triglyceride biosynthesis, lipid droplet formation, and sterol/cholesterol handling (38 genes) were defined using Reactome lipid metabolism pathways and prior lipogenesis-focused studies[13-17]. Examples include cholesterol synthesis (*Hmgcr*), mono-/diacylglycerol acyltransferases (*Mogat2, Dagla, Daglb*), lipoprotein handling (*Apobr, Ldlrap1*), and sterol transfer proteins (*Stard5*).

Growth hormone (GH) signaling axis. A "core" GH axis gene set (18 genes) was defined based on canonical GH–JAK–STAT signaling, including receptors (*Gh, Ghr*), downstream effectors (*Stat5a/b, Jak2, Socs2/3, Cish*), and insulin-like growth factor genes (*Igf1, Igfals, Igfbp1–6*). An "extended" GH axis (28 genes) was additionally defined to capture downstream metabolic mediators (*Igf2, Igf2r, Igf1r, Ppargc1a, Ppargc1b, Cebpa/b, Stat3, Socs1, Irs1*).

Finally, up-regulated DEGs identified between DU6 and DU6P were intersected with each curated category using official gene symbols. For each category, enrichment was tested using a two-tailed Fisher's exact test, with the full set of 16,687 annotated mouse genes as background. For each comparison, we report the observed number of DEGs in the category, expected overlap by chance, odds ratio (OR), and *p*-value. This approach enabled quantification of overrepresentation for both broad mitochondrial pathways (OXPHOS, FAO) and focused functional modules (lipid synthesis, GH axis).

Multivariable logistic regression model for classifying the insulin status of the separate validation dataset

In the multivariable logistic regression model that included LOWESS-derived quartile z score, measured BMI, age, and gender, quartile z score showed an odds ratio of 0.645 ($p = 0.141$), whereas measured BMI showed an odds ratio of 1.074 ($p = 0.113$). Age remained significantly associated with insulin resistance status, with an odds ratio of 1.111 per year ($p = 9.60 \times 10^{-4}$), while gender was not significant (OR = 0.712, $p = 0.576$). The overall model showed modest explanatory performance, with an AIC of 88.872.

Prediction-oriented nested cross-validation analysis

To characterize the predictive behavior of the BMI-anchored transcriptomic framework under an explicitly prediction-oriented setting, we performed repeated outer-split analyses using 40%, 50%, 60%, 70%, 80%, and 90% of the primary modeling cohort as the outer training set. For each training fraction, a random subset

containing the corresponding proportion of participants was drawn without replacement from the primary modeling cohort, with all remaining participants constituting the outer hold-out testing set. This procedure was repeated independently for 100 iterations at each training fraction.

Within each outer training subset, the LASSO regularization parameter α was selected by inner 5-fold cross-validation. Participants in the outer training subset were randomly divided into five folds, and the same fold assignments were used to evaluate all candidate α values within that outer iteration. For each candidate α, each participant in the outer training subset received exactly one out-of-fold prediction from a model fitted using the other four inner folds, ensuring that the participant being predicted had not contributed to model fitting. These inner out-of-fold predictions were pooled across the five folds, and Spearman's correlation between measured and predicted BMI was calculated over the complete outer training subset.

The α yielding the highest pooled inner-CV Spearman's ρ was selected. If two α values produced numerically indistinguishable optimization scores, the larger α was selected to favor the sparser solution. The candidate grid comprised logarithmically spaced α values from 0.02 to 1.00, with the primary-model α of 0.123 additionally included in the grid. After α selection, the final LASSO model was refitted using all participants in the corresponding outer training subset and was then evaluated once in the complementary outer hold-out subset. The outer hold-out participants were therefore not used for α selection or LASSO model fitting.

Prediction performance was summarized across the 100 outer iterations at each training fraction using the median and interquartile range (IQR) of outer hold-out Spearman's ρ. Despite explicitly selecting α to optimize inner out-of-fold BMI prediction, median hold-out Spearman's ρ remained modest across the evaluated training fractions, ranging from 0.148 to 0.212 (Fig. S10a).

Enrichment of primary-model genes among prediction-selected genes

For outer iteration i, let k_i denote the total number of genes with non-zero coefficients in the prediction-oriented model, and let x_i denote the number of these genes that were also present in the primary 458-gene set. The proportion of primary-model genes among prediction-selected genes in iteration i was defined as:

$$r_i = \frac{x_i}{k_i}$$

For each outer training fraction, the distribution of r_i across the 100 iterations was summarized by its median and IQR. This statistic gives equal weight to each fitted prediction-oriented model regardless of its number of selected genes.

Because model sparsity varied across iterations, we additionally calculated the pooled proportion of primary-model genes among all prediction-selected gene events:

$$R_{pooled} = \frac{\sum_{i=1}^{100} x_i}{\sum_{i=1}^{100} k_i}$$

Unlike the median of the iteration-level proportions, R_{pooled} weights each iteration according to the number of genes selected by that model. Both summaries are therefore shown in Fig. S10b: bars represent the pooled proportion, whereas the overlaid points and error bars represent the median and IQR of the iteration-level proportions.

The primary gene set comprised 458 of the 16,042 candidate genes, corresponding to a background proportion of 2.86%. To descriptively quantify enrichment above this background, a 2 × 2 contingency table was constructed separately for each training fraction by aggregating gene-selection events across the 100 iterations. Gene-selection events were classified according to whether the gene belonged to the primary 458-gene set and whether it was selected by the prediction-oriented model. Enrichment was summarized using an odds ratio and a two-sided Fisher's exact test. The pooled enrichment increased with training-set size, with odds ratios ranging from 18.7 at 40% training to 149.1 at 90% training; all two-sided Fisher's exact p-values were $< 10^{-1500}$ (Fig. S10b). Because repeated gene-selection events are not statistically independent, transcriptomic features are correlated, and the primary gene set was derived from the same modeling cohort, these p-values were interpreted as descriptive measures of the magnitude of enrichment rather than as independent replication evidence.

Sex-adjusted sensitivity analysis

Primary modeling cohort

To assess whether the female-predominant composition of the primary modeling cohort materially influenced the BMI-anchored transcriptomic model, we performed a full-cohort sensitivity analysis in which sex was explicitly included as an unpenalized covariate. The analysis retained the same 1,298 participants and 16,042 high-confidence human–mouse orthologous genes used for primary model construction. Sex was included as an unpenalized covariate (male = 1 and female = 0) using a partial-regression approach. BMI and each gene-expression feature were residualized with respect to an intercept and sex, and LASSO regression was then fitted to the residualized data using the same regularization parameter as the primary model ($\alpha = 0.123$). This allowed the effect of sex adjustment to be evaluated while keeping the degree of regularization unchanged. The resulting gene coefficients were compared with those from the primary unadjusted model. Model concordance was assessed using Spearman's correlation across the 458 genes selected in the primary model, and the number of primary genes that remained selected after sex adjustment was recorded.

Explicit adjustment for sex in the full primary cohort had minimal impact on the BMI-anchored transcriptomic model. Using the original regularization parameter ($\alpha = 0.123$), the sex-adjusted model selected 462 genes, of which 434 overlapped the 458 genes in the primary model (94.8% recovery; Jaccard index = 0.893). Coefficients across the complete primary gene set remained highly concordant with those of the unadjusted model (Spearman's $\rho = 0.987$), and coefficient directions were fully concordant among genes selected in both models. Retuning the regularization parameter after sex adjustment yielded a nearly identical α of 0.124 and retained 433 of the 458 primary genes (94.5%).

Separate validation cohort

On the other hand, to assess whether sex imbalance influenced the insulin-resistance-related transcriptomic signal observed in the separate validation cohort, we analyzed the PCA scores obtained by projecting validation samples onto the PCA space defined in the primary modeling cohort. The PCA model was not refitted in the validation cohort; the training-derived mean vector and PCA loadings were retained. The association between insulin-resistance status and PC1 was then evaluated using linear regression with PC1 as the dependent variable and insulin-resistance status and sex as independent variables. The corresponding unadjusted model containing insulin-resistance status alone was evaluated for comparison. As a complementary analysis, insulin-resistance status was modeled using logistic regression with standardized PC1 and sex as predictors, with the PC1 effect expressed as an odds ratio per 1-standard-deviation increase in PC1. An exploratory insulin-resistance-status-by-sex interaction was additionally examined but was interpreted descriptively because of the limited number of male participants.

In linear regression analyses, the coefficient for insulin-resistance status (β) represents the estimated mean difference in projected PC1 score between insulin-resistant and insulin-sensitive participants, with insulin-sensitive participants as the reference group. In the unadjusted linear regression, insulin-resistant participants had higher PC1 scores than insulin-sensitive participants ($\beta = 1.270$, 95% CI 0.152–2.388, p-value = 0.027). After adjustment for sex, the estimated difference remained essentially unchanged ($\beta = 1.266$, 95% CI 0.160–2.371, p-value = 0.025). Consistently, in a complementary logistic regression model, PC1 remained associated with insulin resistance after adjustment for sex (OR = 1.87 per 1-SD increase in PC1, 95% CI 1.06–3.31, p-value = 0.031).

SI Figures

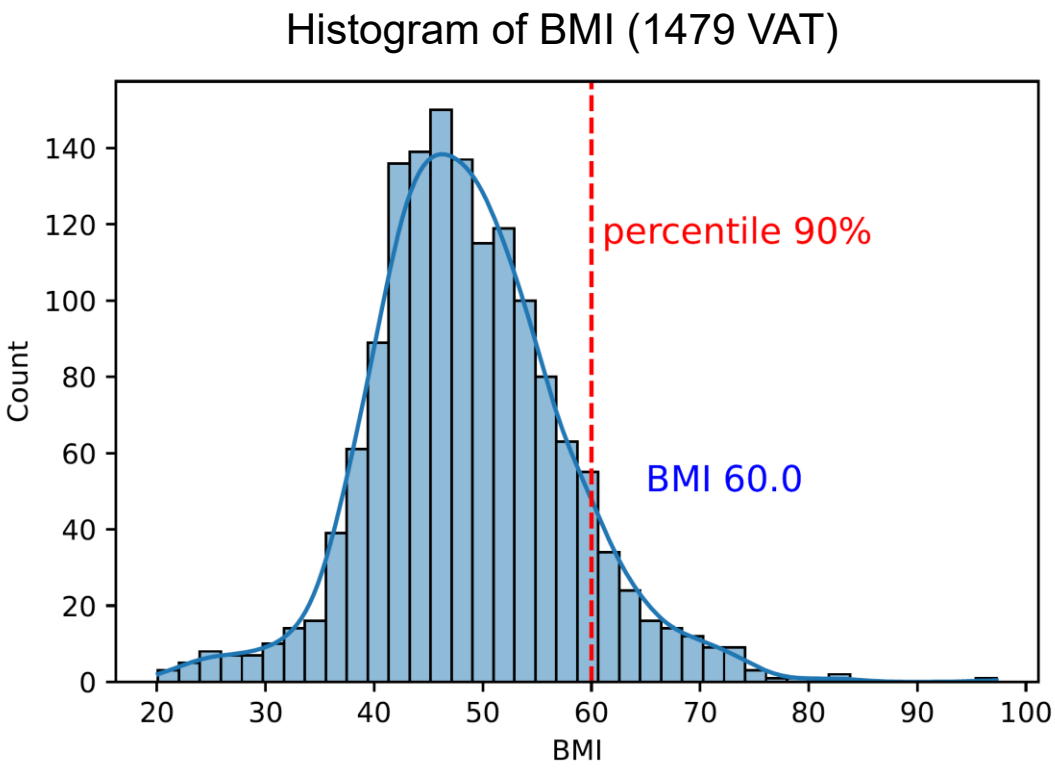


Figure S1. Measured BMI distribution in human visceral adipose tissue samples.

Histogram of measured BMI values from 1,479 human visceral adipose tissue samples. The 90th percentile (red dashed line) corresponds to measured BMI ≈ 60.0. Sample counts by measured BMI range: < 30 (n = 31), 30 – 40 (n = 159), 40 – 50 (n = 669), 50 – 60 (n = 470), ≥ 60 (n = 150).

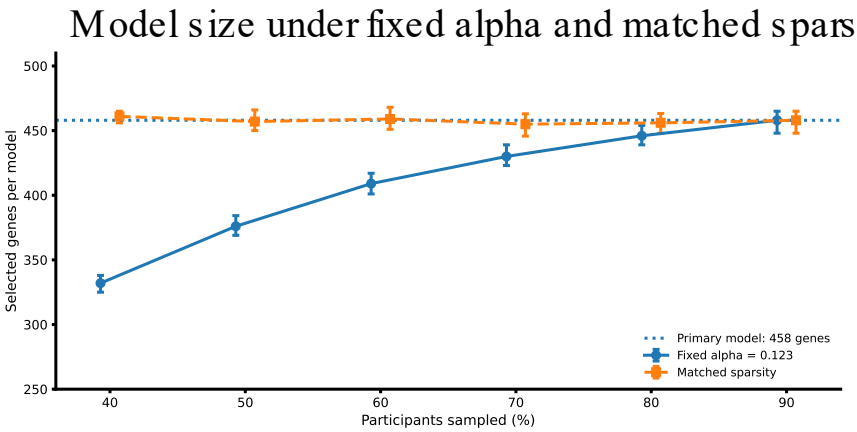


Figure S2. Number of selected genes under repeated subsampling of the primary modeling cohort.

Number of genes selected per model across 500 repeated subsampling iterations using 40–90% of the primary modeling cohort. Points indicate the median and error bars the interquartile range. Models were fitted either using the primary-model regularization parameter (fixed $\alpha = 0.123$) or using a sampling-fraction-specific α calibrated to approximate the 458-gene sparsity of the primary model.

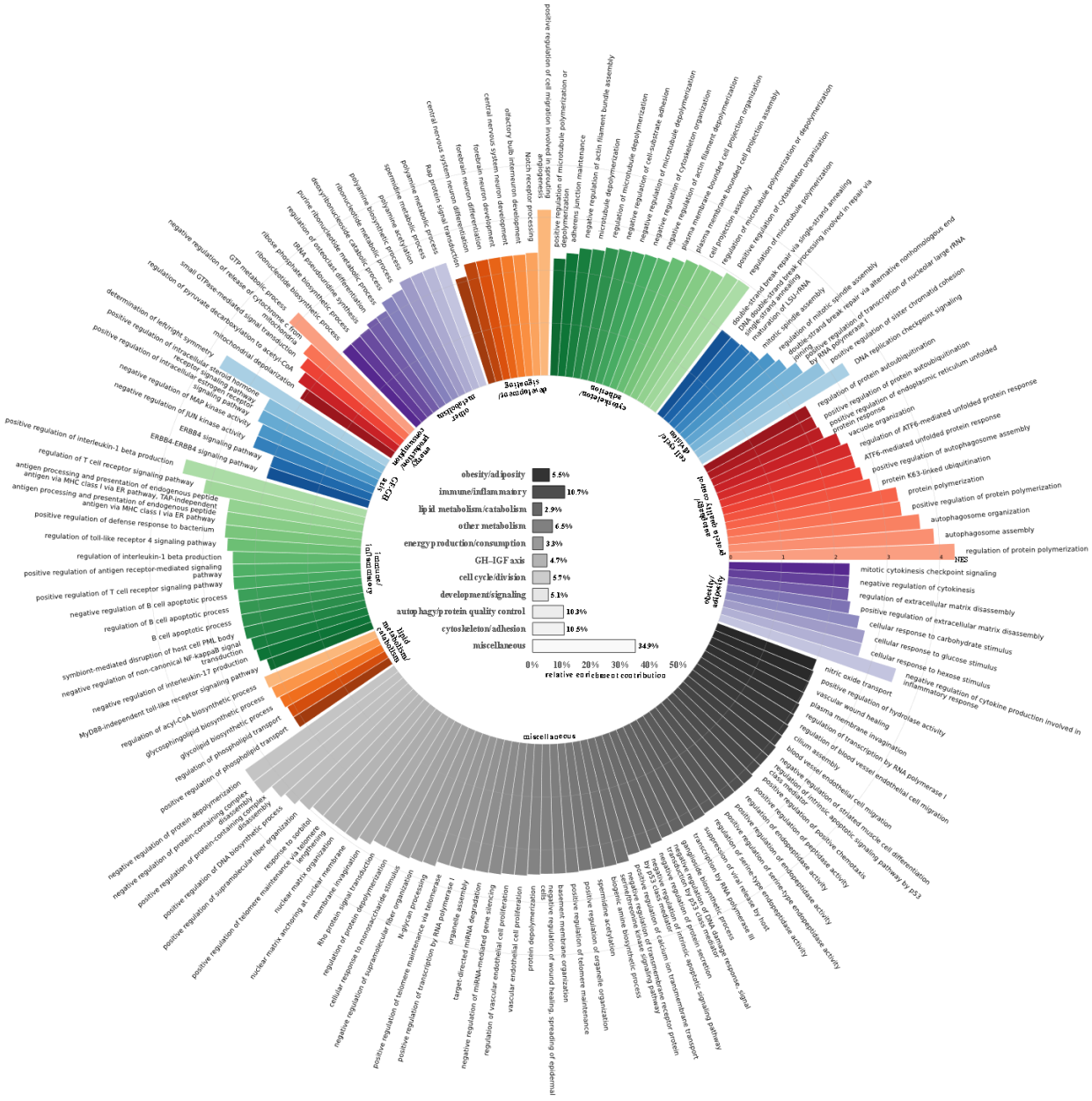


Figure S3. Circular representation of relative enrichment contributions of individual functions grouped by biological category in model-selected genes.

Circular bar plot depicting the normalized enrichment scores (NES) of individual functions identified from Gene Set Enrichment Analysis (GSEA) of model-selected genes. Each bar represents a single biological function, with bar length corresponding to its NES value. Functions are grouped into higher-order biological categories, labeled in the inner circle. The inner bar plot indicates the percentage contribution of each category to the total enrichment signal. This percentage was calculated as the sum of NES values for all functions within a given category divided by the summed NES values across all plotted categories and multiplied by 100. Thus, the percentages represent the relative enrichment contribution of each biological category. This visualization highlights the integrated transcriptional architecture of obesity, demonstrating coordinated contributions from metabolic, immune, structural, and regulatory processes.

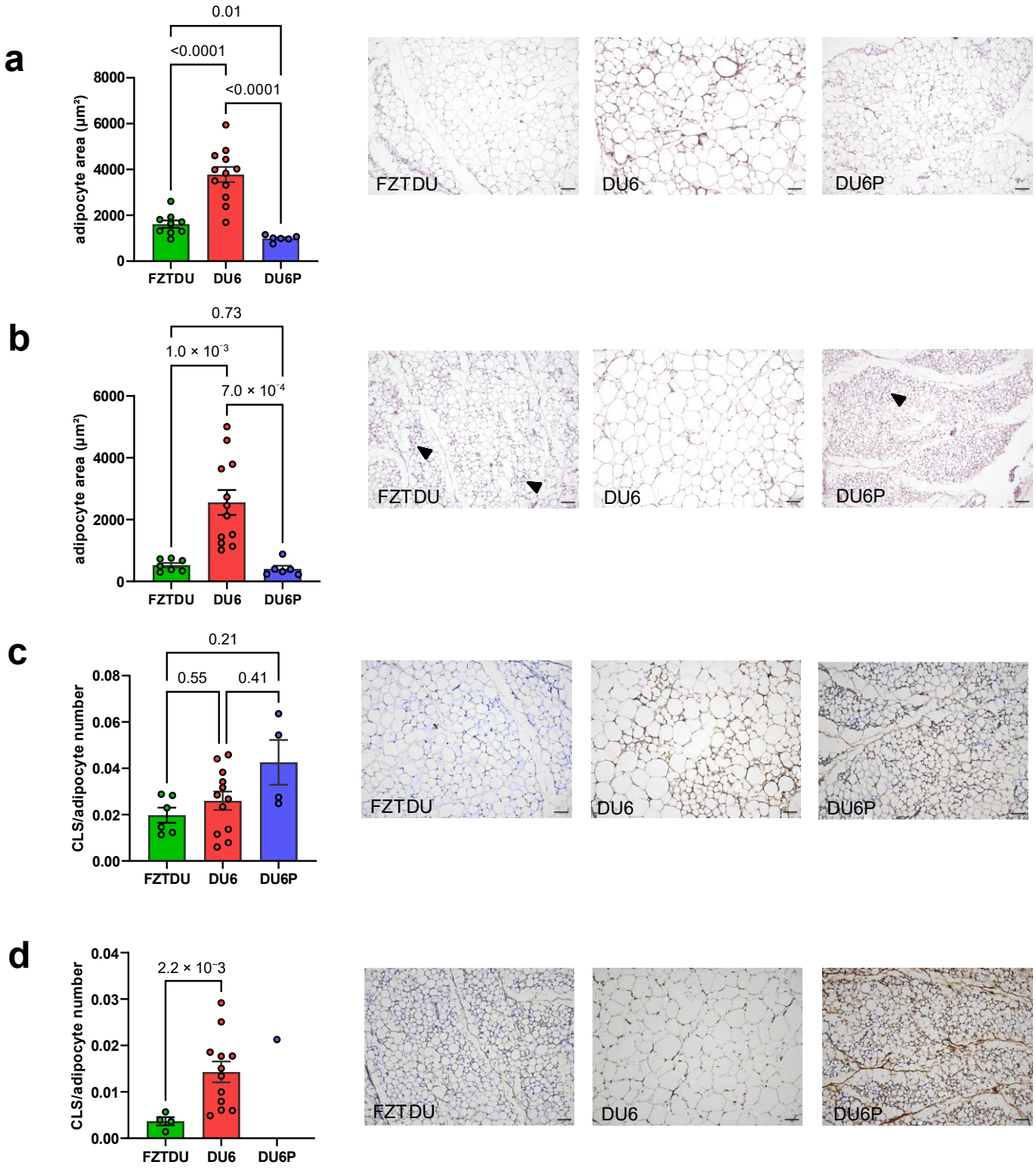


Figure S4. Independent histological characterization of adipose depots in DU6, DU6P, and FZTDU mice.

(a and b) Representative hematoxylin and eosin staining and adipocyte size quantification in epigonadal (a) and inguinal (b) white adipose tissue. (c and d) Representative F4/80 immunohistochemical staining and crown-like structure quantification in epigonadal (c) and inguinal (d) white adipose tissue. Analyses were performed in an independent cohort. ns, not significant. *P*-values below 0.0001 are presented as $p < 0.0001$ because exact values are not provided in the GraphPad Prism output.

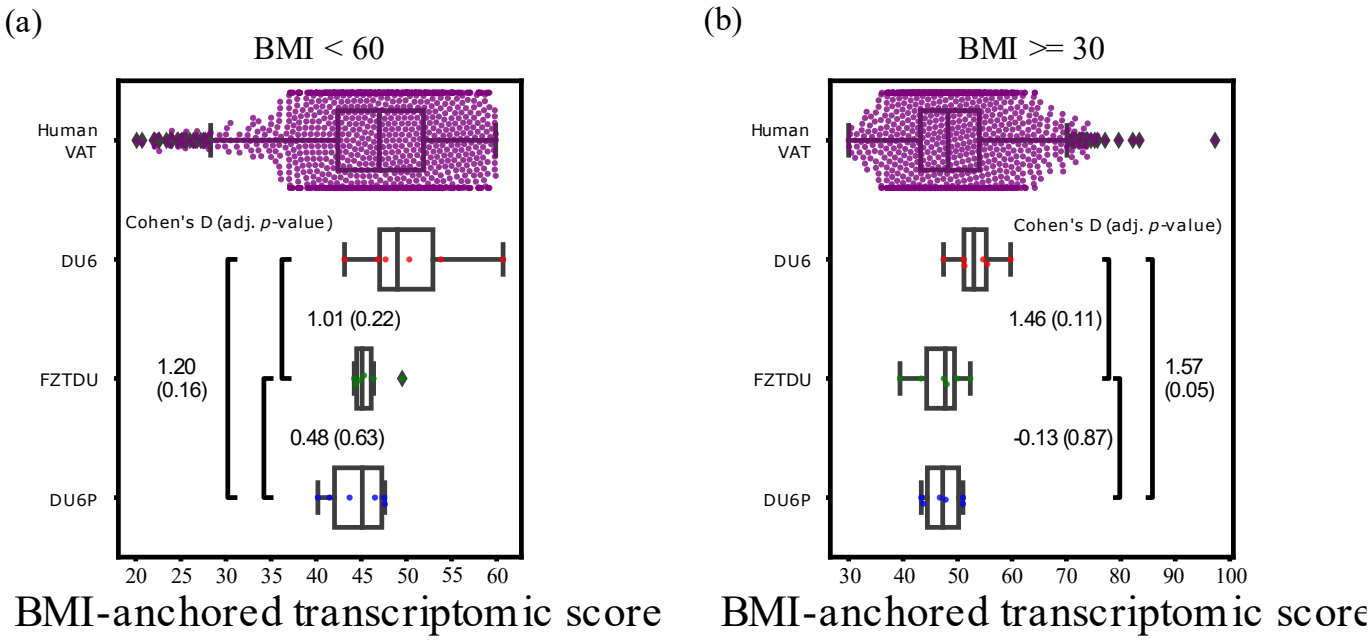


Figure S5. Robustness of cross-species BMI-anchored transcriptomic scores to alternative BMI-range definitions used for human model construction.

The BMI-anchored transcriptomic model was reconstructed using the same modeling procedure as the primary analysis but with two alternative human VAT cohort definitions: (a) BMI <60, thereby additionally including participants with BMI <30, and (b) BMI ≥30, thereby removing the upper-BMI restriction and including the previously excluded highest-BMI participants. Human BMI distributions used for model construction are shown at the top of each panel. Each refitted model was subsequently applied to epigonadal adipose transcriptomes from DU6, FZTDU, and DU6P mice (n = 6 per line) to derive BMI-anchored transcriptomic scores. Individual mice are shown as colored points with boxplots summarizing the group distributions. Brackets indicate pairwise Cohen's d effect sizes, with Holm–Bonferroni-adjusted *p*-values from Wilcoxon rank-sum tests shown in parentheses. *P*-values were adjusted separately for the three pairwise mouse-line comparisons within each BMI-range model. The characteristic elevation of the DU6 transcriptomic score was retained under both alternative cohort definitions.



Figure S6. Circular representation of relative enrichment contributions of individual functions grouped by biological category in DU6 mice.

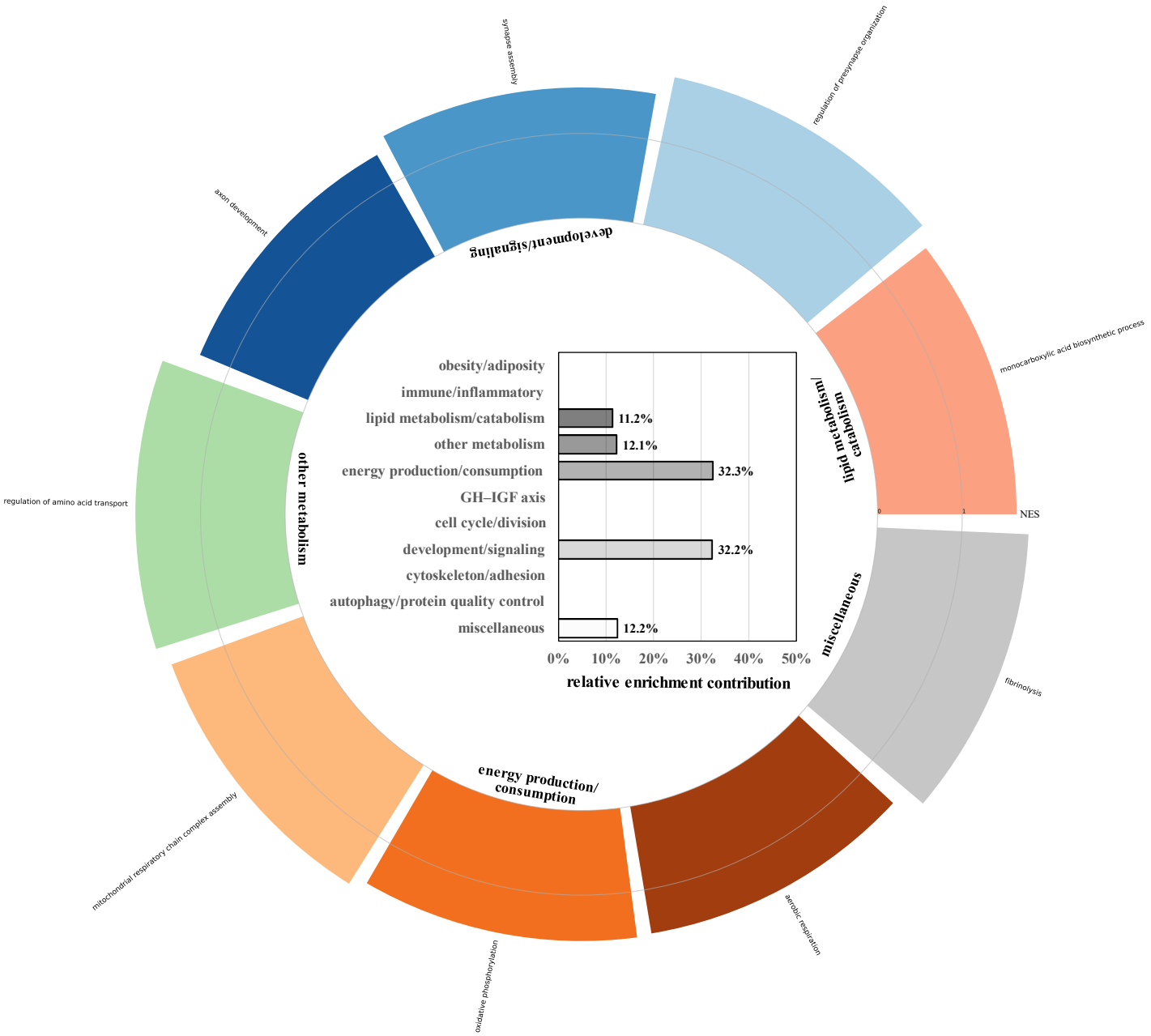


Figure S7. Circular representation of relative enrichment contributions of individual functions grouped by biological category in DU6P mice.

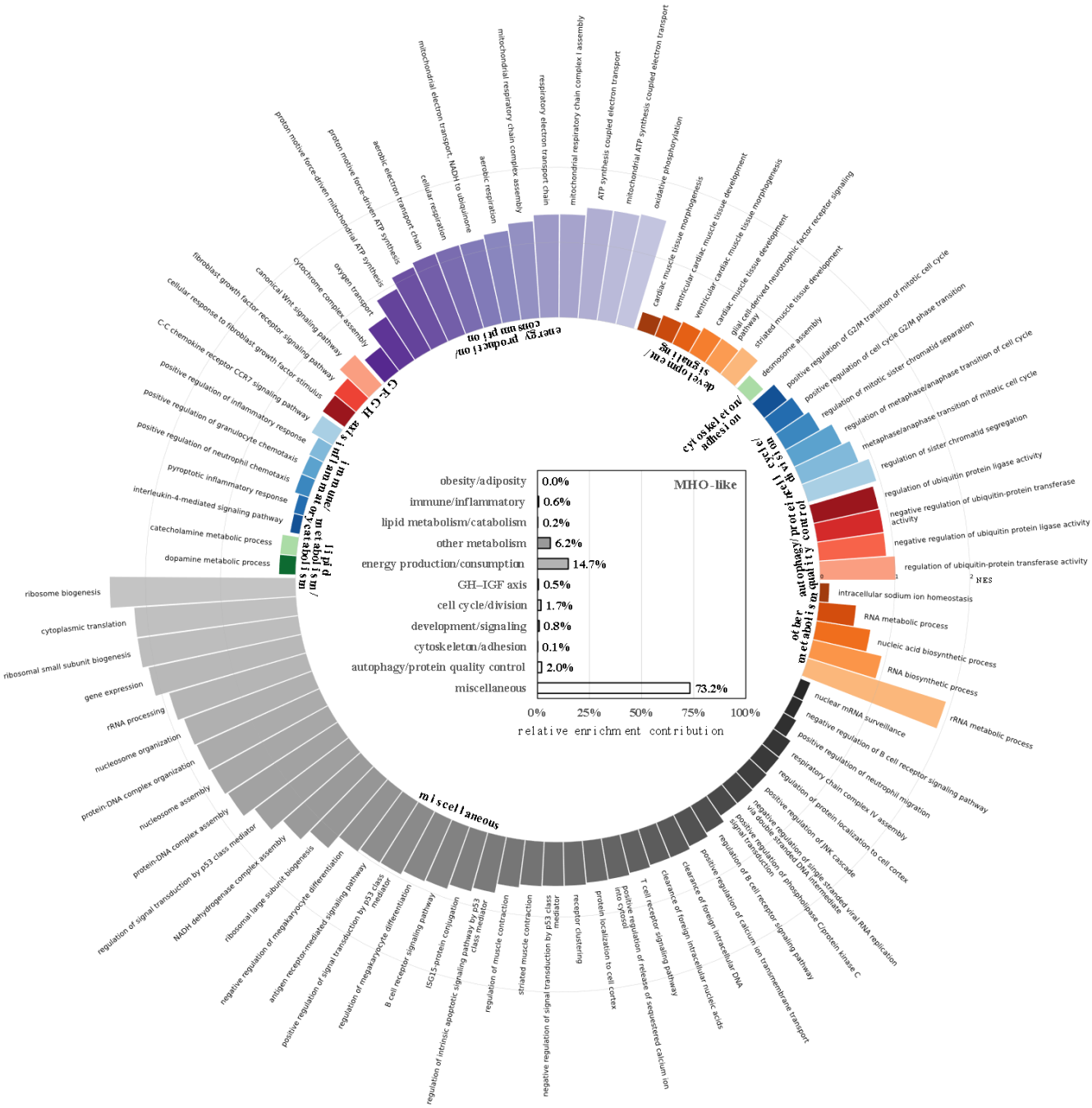


Figure S8. Circular representation of relative enrichment contributions of individual functions grouped by biological category in MHO-like patients.



Figure S9. Circular representation of relative enrichment contributions of individual functions grouped by biological category in MUO-like patients.

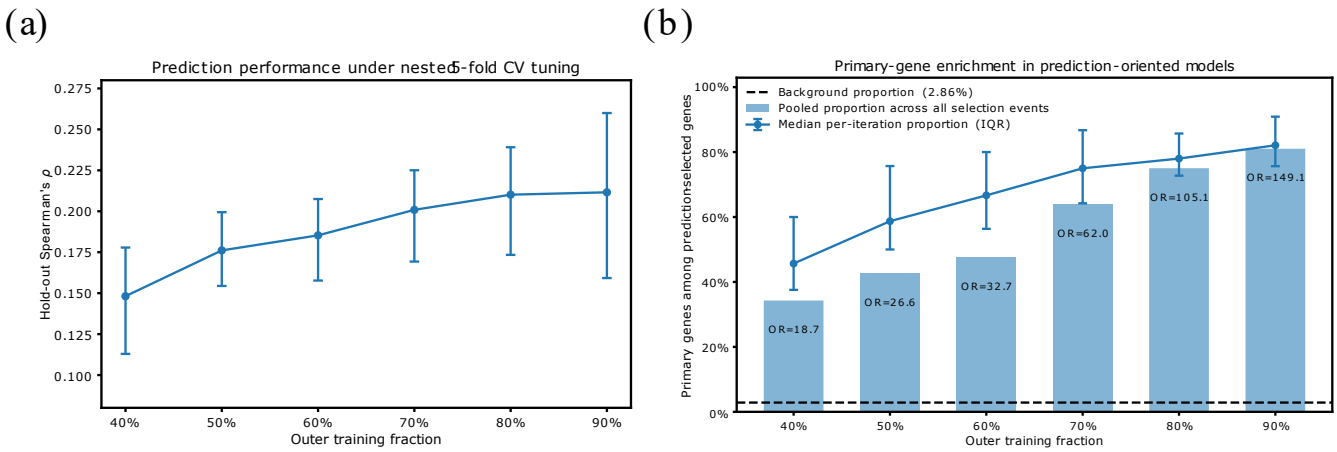


Figure S10. Prediction-oriented evaluation of the BMI-anchored transcriptomic framework.

(a) Prediction performance under nested 5-fold cross-validation tuning. Points and error bars indicate the median and interquartile range of outer hold-out Spearman's ρ across 100 repeated outer-split iterations. Within each outer training set, α was selected by inner 5-fold cross-validation using pooled out-of-fold Spearman correlation as the tuning criterion, after which the model was refitted on the complete outer training set and evaluated on the untouched outer hold-out set. Median hold-out Spearman's ρ ranged from 0.148 to 0.212 across training fractions, indicating modest out-of-sample BMI prediction despite explicit optimization for predictive performance. (b) Primary-gene enrichment in prediction-oriented models. Bars show the pooled proportion of prediction-selected genes belonging to the primary 458-gene set across all gene-selection events from 100 repeated outer-split iterations at each training fraction. The overlaid line and error bars show the median and interquartile range of the corresponding proportion calculated separately within each outer iteration. The dashed horizontal line indicates the background proportion of primary genes among all 16,042 candidate genes (458/16,042 = 2.86%). Odds ratios from two-sided Fisher's exact tests based on aggregated selected and non-selected gene events are shown above the bars. Odds ratios increased from 18.7 to 149.1 across the 40–90% training fractions, with all two-sided Fisher's exact p-values $< 10^{-1500}$.

SI Tables

Table S1. Functional enrichment of mitochondrial complexes, regulatory machinery, and lipid metabolism pathways in DU6P and DU6 mouse lines.

Category	Odds Ratio	p-value	Observed Genes
DU6P			
Complex I (NADH dehydrogenase; structural)	28.4	3.75×10^{-12}	*Ndufa1, Ndufa3, Ndufa7, Ndufa9, Ndufb1, Ndufb7, Ndufb10, Ndufs1, Ndufs3, Ndufv1, Ndufv3*
Complex II (Succinate dehydrogenase)	27.1	4.8×10^{-2}	*Sdhb*
Complex III (Ubiquinol–cytochrome c reductase)	69.0	1.13×10^{-7}	*Cyc1,Uqcrfs1, Uqcrq, Uqcr11, Uqcrh*
Complex IV – Assembly factors	10.9	1.8×10^{-2}	*Surf1, Coa8*
Complex IV – Structural subunits	31.8	2.0×10^{-6}	*Cox5a, Cox7a2, Cox7c, Cox8a, mt-Co3*
Complex V (ATP synthase)	27.7	3.5×10^{-7}	*Atp5f1d, Atp5mc1, Atp5mc3, Atp5mk, Atp5pf, Atp5po*
Mitochondrial ribosomal proteins (MRPs)	5.57	2.8×10^{-3}	*Mrpl14, Mrpl47, Mrpl48, Mrps28, Mrps31*
Mitochondrial translational activators	30.7	2.8×10^{-4}	*Lrpprc, Mss51, Ptcd1*
Mitochondrial regulatory machinery (merged: translational activators + biogenesis regulators + import machinery + MCU regulators)	15.3	2.1×10^{-7}	*Mss51, Lrpprc, Ptcd1, Ppargc1b, Sirt1, Dnajc19, Samm50, Mcur1*
FAO axis (total fatty-acid breakdown) (FAO + ETF + PEX)	20.3	2.9×10^{-8}	*Acadvl, Cpt2, Crot, Echs1, Etfb, Pex11g, Pex26, Stard7*
Lipid synthesis & storage	9.68	1.2×10^{-3}	*Acacb, Elovl2, Scd1, Stard7*
GH signaling extended	13.7	3.5×10^{-4}	*Irs1, Ppargc1b, Socs1, Stat5a*
DU6			
Lipid synthesis & storage	76.5	3.8×10^{-11}	*Hmgcr, Mogat2, Dagla, Daglb, Apobr, Ldlrap1, Stard5*
Complex IV – Structural subunits	17.8	5.9×10^{-2}	*Cox7a1*
GH signaling extended	11.2	9.0×10^{-2}	*Igf2r*

Fisher's exact tests were performed separately for DU6-upregulated and DU6P-upregulated DEGs using 16,687 genes as the background. Only categories with at least one observed upregulated DEG are shown in the table.

Table S2. Rule-based functional category assignment for GO biological process terms.

Each unique GO biological process description included in the analysis was assigned to one of eleven functional categories using the rule-based classification scheme described in the Supplementary Methods. Table S2 provides the one-to-one mapping between GO term description and assigned functional category.

Table S3. Curated functional gene sets used for DU6 and DU6P up-regulated DEG enrichment analysis.

Table S3 lists the full gene membership and category size for each functional category used in the Fisher's exact tests of DU6-upregulated and DU6P-upregulated DEGs, including oxidative phosphorylation complexes, mitochondrial translation and regulatory modules, fatty-acid oxidation, lipid synthesis and storage, and growth hormone signaling categories.

Table S4. List of primary model-selected genes

Table S4 lists the information of genes selected by the primary model.

References

1 Rankin, A. L. *et al.* IL-33 induces IL-13-dependent cutaneous fibrosis. *J Immunol* **184**, 1526-1535, doi:10.4049/jimmunol.0903306 (2010).

2 Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics International* **11**, 36-42 (2004).

3 Galarraga, M. *et al.* Adiposoft: automated software for the analysis of white adipose tissue cellularity in histological sections. *J Lipid Res* **53**, 2791-2796, doi:10.1194/jlr.D023788 (2012).

4 Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* **51**, D587-D592, doi:10.1093/nar/gkac963 (2023).

5 Milacic, M. *et al.* The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res* **52**, D672-D678, doi:10.1093/nar/gkad1025 (2024).

6 Gene Ontology, C. The Gene Ontology knowledgebase in 2026. *Nucleic Acids Res* **54**, D1779-D1792, doi:10.1093/nar/gkaf1292 (2026).

7 Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425, doi:10.1016/j.cels.2015.12.004 (2015).

8 Rath, S. *et al.* MitoCarta3.0: an updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Res* **49**, D1541-D1547, doi:10.1093/nar/gkaa1011 (2021).

9 Baldarelli, R. M. *et al.* Mouse Genome Informatics: an integrated knowledgebase system for the laboratory mouse. *Genetics* **227**, doi:10.1093/genetics/iyae031 (2024).

10 Mai, N., Chrzanowska-Lightowlers, Z. M. & Lightowlers, R. N. The process of mammalian mitochondrial protein synthesis. *Cell Tissue Res* **367**, 5-20, doi:10.1007/s00441-016-2456-0 (2017).

11 D'Souza, A. R. & Minczuk, M. Mitochondrial transcription and translation: overview. *Essays Biochem* **62**, 309-320, doi:10.1042/EBC20170102 (2018).

12 Lightowlers, R. N., Rozanska, A. & Chrzanowska-Lightowlers, Z. M. Mitochondrial protein synthesis: figuring the fundamentals, complexities and complications, of mammalian mitochondrial translation. *FEBS Lett* **588**, 2496-2503, doi:10.1016/j.febslet.2014.05.054 (2014).

13 Coleman, R. A. & Mashek, D. G. Mammalian triacylglycerol metabolism: synthesis, lipolysis, and signaling. *Chem Rev* **111**, 6359-6386, doi:10.1021/cr100404w (2011).

14 Yen, C. L., Stone, S. J., Koliwad, S., Harris, C. & Farese, R. V., Jr. Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *J Lipid Res* **49**, 2283-2301, doi:10.1194/jlr.R800018-JLR200 (2008).

15 Walther, T. C. & Farese, R. V., Jr. Lipid droplets and cellular lipid metabolism. *Annu Rev Biochem* **81**, 687-714, doi:10.1146/annurev-biochem-061009-102430 (2012).

16 Farese, R. V. & Walther, T. C. Essential Biology of Lipid Droplets. *Annu Rev Biochem* **94**, 447-477, doi:10.1146/annurev-biochem-091724-013733 (2025).

17 Alpy, F. & Tomasetto, C. Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. *J Cell Sci* **118**, 2791-2801, doi:10.1242/jcs.02485 (2005).