ADIPOCYTE
2026, VOL. 15, NO. 1, 2680625
https://doi.org/10.1080/21623945.2026.2680625

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Multi-omics integration reveals convergent extracellular matrix remodelling and lipid metabolic reprogramming as central axes of adipocyte differentiation from mouse embryonic stem cells

M. Al-Sayegh [a,b,c], M. Khalili[a], M. Alzaabi[a], M. Sultana[b], L. Ali[d], M. Ali[d] and M. El-Hadidi[e]

[a]Division of Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates; [b]Center of Genomics and System Biology, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates; [c]College of Medicine and Health, University of Birmingham, Birmingham, UK; [d]Core Technology Platforms Operations, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates; [e]Department of Cancer and Genomic Sciences, College of Medicine and Health, University of Birmingham Dubai, Dubai, United Arab Emirates

ABSTRACT

Adipogenesis from mouse embryonic stem cells (mESCs) offers a tractable model for dissecting early adipocyte commitment, yet the mechanisms coordinating this transition across multiple biological layers remain incompletely understood. Here we present the first simultaneous five-layer multi-omics characterization of mESC-derived adipocyte differentiation, integrating transcriptomics, proteomics, secretomics, lipidomics, and metabolomics from matched adipogenic (Pos) and non-differentiating (Neg) cell populations at day 30. Applying Multi-Omics Factor Analysis (MOFA+), we identified a dominant shared latent axis that perfectly segregated Pos from Neg cells across all five views. Layer-specific functional enrichment converged on two principal biological axes: ECM remodeling — encompassing collagens, laminins, thrombospondins, and lysyl oxidases — and lipid metabolic reprogramming, with phospholipid and glycerolipid metabolic processes dominating the lipidomics/metabolomics layer. Ensemble Machine Learning Feature Ranking (EMFR) identified a secreted factor (Scpep1; importance score 0.90) as the top-ranked discriminatory feature. Network analysis revealed indirect ECM-lipid connectivity mediated by four bridging nodes (Plod1, Thbs2, Plg, Pmp22) through a hub subnetwork of phospholipid-metabolizing enzymes. Targeted qPCR validation of six candidate regulators (Itga5, Igfbp6, Pik3cg, Lpl, Acer3, Sirt1) confirmed RNA-seq concordance. These findings establish convergent ECM remodeling and lipid metabolic reprogramming as central axes of adipocyte identity acquisition from mESCs.

ARTICLE HISTORY
Received 5 March 2026
Revised 22 May 2026
Accepted 23 May 2026

KEYWORDS
Adipogenesis; mouse embryonic stem cells; multi-omics integration; extracellular matrix remodelling; lipidomics; MOFA+; multi-omics factor analysis

Introduction

Adipose tissue is a dynamic endocrine organ that integrates nutrient storage with systemic metabolic regulation, including glucose homoeostasis, inflammation, and insulin sensitivity [1,2]. In obesity and type 2 diabetes mellitus (T2DM), adipose dysfunction contributes directly to insulin resistance, ectopic lipid deposition, and chronic low-grade inflammation [3]. At the cellular level, these pathologies are rooted in alterations in adipocyte differentiation, expansion, and functional maturation.

Adipogenesis is a tightly orchestrated process involving coordinated transcriptional activation, metabolic reprogramming, structural remodelling, and extracellular matrix (ECM) reorganization [4,5]. Classical in vitro systems such as 3T3-L1 cells and mouse embryonic fibroblasts (MEFs) have established central regulatory pathways, including the well-characterized PPARγ and C/EBP transcriptional cascade, which remains the primary driver of adipogenic commitment, as well as hormonal inputs such as insulin and glucocorticoids [6–9]. While PPARγ-driven transcription is indisputably central to adipogenesis, these models largely capture terminal differentiation and may not fully recapitulate early lineage commitment and developmental priming events that precede adipocyte maturation [10].

CONTACT M. El-Hadidi m.el-hadidi@bham.ac.uk Department of Cancer and Genomic Sciences, College of Medicine and Health, University of Birmingham Dubai, Dubai International Academic City, PO Box 341799, Dubai, United Arab Emirates
Supplemental data for this article can be accessed online at https://doi.org/10.1080/21623945.2026.2680625

Mouse embryonic stem cells (mESCs) provide a unique system to interrogate adipogenesis from a pluripotent state through lineage commitment to mature adipocytes [6,11,12]. In this system, differentiating (Pos) cells are induced towards the adipogenic lineage, while non-differentiating (Neg) cells maintained under standard mESC culture serve as reference controls. Consistent with the known progressive suppression of naive pluripotency during extended culture, OCT4 expression is downregulated in both Pos and Neg cells; however, Neg cells do not activate adipogenic transcriptional programmes, as confirmed by the absence of FABP4, PPARG, and CEBPB induction and negligible Oil Red O staining at Day 30 (Supplementary Figure S2; [13]). Neg cells therefore represent a well-defined non-differentiating reference condition rather than a fully pluripotent state, providing a meaningful transcriptional and phenotypic contrast for multi-omics comparison. Recent transcriptomic analyses have demonstrated that mESC-derived adipocytes display canonical features of adipogenesis, including suppression of pluripotency genes and induction of markers such as Pparg, Cebpa, and Fabp4 [13,14]. These models offer the opportunity to examine adipogenesis as a developmental trajectory rather than solely as terminal differentiation.

Despite extensive characterization of transcriptional regulators, adipocyte differentiation is not governed by gene expression alone. Proteomic remodelling enforces cytoskeletal restructuring and ECM dynamics [2,15], lipidomic and metabolomic shifts support triglyceride biosynthesis and membrane expansion [8,9,14], and secreted factors modulate paracrine signalling within the adipose niche [15–17]. The extracellular matrix itself is actively remodelled during adipogenesis and influences cell fate stability and metabolic phenotype [18,19]. Thus, adipogenesis represents a multi-layered biological transition integrating intracellular regulation with extracellular communication.

Single-omics approaches provide only partial insight into this complexity. While transcriptomic studies have identified thousands of differentially expressed genes during adipocyte differentiation [10,13], corresponding changes in protein abundance, lipid composition, metabolite flux, and secreted signalling molecules are not necessarily concordant. Integrative multi-omics strategies allow identification of shared and modality-specific regulatory axes by detecting features whose variation is associated with the biological condition of interest — here, the adipogenic versus non-differentiating state — across multiple molecular layers simultaneously [20]. Statistical frameworks such as Multi-Omics Factor Analysis (MOFA+) enable decomposition of heterogeneous datasets into latent factors that capture coordinated variation across layers [21,22]. These approaches have been successfully applied to complex metabolic disorders, including early stages of T2DM [23], but remain underutilized in developmental adipogenesis models.

Furthermore, lipid metabolic enzymes, sphingolipid regulators, and growth factor–binding proteins have emerged as modulators of metabolic homoeostasis beyond classical transcription factors [24–30]. However, their integrated roles during early adipocyte lineage commitment remain incompletely defined. A systems-level understanding of how transcriptional priming, structural remodelling, lipid metabolism, and extracellular signalling converge during adipogenesis is still lacking.

In this study, we applied an integrated multi-omics framework to dissect adipocyte differentiation from mouse embryonic stem cells. By combining transcriptomic, proteomic, metabolomic, lipidomic, and secretomic profiling at day 30 under differentiating and non-differentiating conditions, we aimed to (1) Identify coordinated regulatory programmes across molecular layers, (2) Decompose shared and modality-specific variance using MOFA+, (3) Prioritize central regulatory hubs through network integration, and (4) validate key candidate genes associated with adipocyte commitment.

Through this integrative systems biology approach, we define convergent regulatory axes linking transcriptional control, extracellular matrix remodelling, lipid metabolism, and secreted signalling during adipogenesis. These findings provide a multi-layered framework for understanding adipocyte differentiation and nominate candidate regulators relevant to metabolic disease pathophysiology.

Materials and methods

Experimental design and sample overview

Mouse embryonic stem cells (mESCs) were differentiated under adipogenic conditions ('Pos') and compared with matched non-differentiating controls ('Neg') at day 30 (D30). All omics layers were generated from the same biological replicates ($n = 3$ per condition), enabling cross-modality integration. This study focuses on terminal differentiation at D30 to capture stable adipocyte phenotypes [13].

Cell culture and adipogenic differentiation

Mouse ESC (mESC:hb9:GFP, strain HBG3) were maintained under ES growth medium consisting of Knockout™ DMEM (Thermo Fisher Scientific), 14% Fetal Bovine Serum (Corning®), 1 × 2-Mercaptoethanol (Thermo Fisher Scientific), 1× L-Glutamine, 1× NEAA, 1× Nucleosides (all Thermo Fisher Scientific), and 10^7 U/mL Leukemia Inhibitory Factor (LIF, Thermo Fisher Scientific) under humidified conditions of 5% CO_2/37°C. Non-differentiating control (Neg) cells were maintained throughout the 30-day period in this LIF-supplemented medium. As shown in Supplementary Figure S2 [13], OCT4 expression is progressively suppressed in both Pos and Neg cells, consistent with the known attenuation of naive pluripotency during extended culture. Importantly, Neg cells do not activate adipogenic transcriptional programmes — FABP4, PPARG, and CEBPB expression remain at baseline — and show negligible Oil Red O staining, confirming their identity as a non-differentiating reference condition throughout this study.

Adipogenic differentiation (Pos condition) was initiated using the same basal medium without LIF. Differentiation proceeded first through embryoid body (EB) formation via the hanging-drop method at 1×10^5 cells/mL for 2 days. EBs were collected and plated onto 0.1% gelatin-coated wells, then treated with 1 µM Retinoic Acid and 12.5 µg/mL Ascorbic Acid in differentiation medium for 3 days (Days 3–6). From Day 7, attached EBs were maintained in differentiation medium supplemented with 12.5 µg/mL Ascorbic Acid, 3 nM Thyroid hormone T3, and 0.5 µg/mL insulin, with medium changes every 2–3 days until Day 25. On Day 25, differentiation medium was replaced with serum-free medium for 5 days. Cells were harvested at Day 30 for all omics profiling. Adipogenic maturation was confirmed by brightfield microscopy and Oil Red O staining at D30 (Supplementary Figure S1 and S2).

For transcriptomic profiling, cells were harvested by TRIzol® extraction (Invitrogen) from confluent 6-well plates (approximately 2×10^6 cells per replicate). For proteomic and secretomic analyses, cells were grown in T75 flasks (approximately 8×10^6 cells per replicate); conditioned media were collected and concentrated using Amicon Ultra-15 centrifugal filters (3 kDa MWCO, Merck Millipore) prior to digestion. For metabolomic and lipidomic profiling, approximately 5×10^6 cells per replicate were harvested by scraping in ice-cold PBS, pelleted by centrifugation (300 × g, 5 min, 4°C), snap-frozen in liquid nitrogen, and stored at −80°C until extraction.

Transcriptomic data processing

Raw RNA-seq count data were obtained from Alzaabi et al. (BioProject PRJNA1126392) [13]. Gene-level counts were merged by Ensembl gene ID using R (v4.3.1). Low-count genes were filtered using DESeq2 (v1.38.3), and normalization was performed using variance-stabilizing transformation (VST). Differential expression between Pos and Neg samples were computed using the Wald test within DESeq2. False discovery rate (FDR) correction was applied using the Benjamini–Hochberg method, and genes with FDR <0.05 were considered significant. The VST-normalized matrix was used for downstream feature ranking, PCA, and MOFA+ integration.

Proteomic and secretomic profiling

Sample preparation

Following the differentiation methods, proteins from cell lysates (proteome) and conditioned media (secretome) were reduced with 20 mM DTT (Sigma, 70°C, 30 min) and S-alkylated with 50 mM iodoacetamide (IAA, Sigma, 1 h, dark). Remaining IAA was quenched with 20 mM DTT. Proteins were precipitated

with 30% trichloroacetic acid (TCA) at 4°C overnight, centrifuged at 16,000 × g (4°C, 15 min), washed with ice-cold acetone, air-dried, and resuspended in 50 mM ammonium bicarbonate (ABC). Protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Fifty micrograms of protein per sample were digested with MS-grade trypsin/Lys-C mix (1:50 w/w enzyme:protein, Thermo Scientific) for 24 h at 37°C. Digestion was quenched with formic acid, peptides were desalted using ZipTip C18 (Merck Millipore), dried, and resuspended in 15 µL of 0.2% formic acid prior to LC-MS/MS analysis.

LC-MS/MS analysis

Proteomics analysis was performed using a fully automated EASY-nLC 1200 system (Thermo Scientific) coupled to a Q-Exactive High-Field Orbitrap mass spectrometer (Thermo Scientific) equipped with an Easy Spray ion source. Peptides were separated on a PepMap RSLC C18 column (75 µm i.d., 15 cm length, maintained at 40°C). Mobile phases were 0.1% formic acid (solvent A) and 0.1% formic acid in 95% acetonitrile (solvent B). A linear gradient was applied: 0% B for 3 min, 0–35% B over 45 min, 35–60% B over 15 min, followed by a 10-min wash at 90% B and 15-min re-equilibration at 0% B. The spray voltage was set to 1.8 kV, S-lens RF level to 35, and ion transfer tube to 275°C. Full MS scans were acquired at 120,000 resolution (m/z range 350–1800; AGC target 3E6; maximum ion time 50 ms). Data-dependent MS/MS was performed on the top 20 most intense precursors using HCD fragmentation (resolution 15,000; AGC target 1E5; minimum AGC 8.0E3; intensity threshold 4.0E5; maximum ion time 20 ms; isolation width 1.4 m/z; charge states 2–5; dynamic exclusion 30 s; normalized collision energy 32%).

Database searching

Database searching was performed using Proteome Discoverer v2.5 (Thermo Fisher Scientific) against the Mus musculus UniProtKB/Swiss-Prot database (release 2025_03). Search parameters: enzyme, trypsin; fragmentation, HCD; precursor mass tolerance, 10 ppm; fragment mass tolerance, 0.2 Da; fixed modification, carbamidomethylation (C); variable modifications, oxidation (M) and N-terminal acetylation. Peptide and protein identifications were filtered at 1% FDR. Only proteins identified with ≥2 unique peptides and detected in at least 2 out of 3 biological replicates within at least one condition were retained for downstream analysis.

In the context of protein retention, 'reproducible signal' was defined as detection with ≥2 unique peptides in at least 2 out of 3 biological replicates within at least one experimental condition (Pos or Neg). This threshold was applied to minimize stochastic identifications while retaining condition-specific proteins that may be absent in the opposing condition.

Metabolomic and lipidomic profiling

Sample preparation

Following the differentiation methods, cells were washed with phosphate-buffered saline (PBS) and homogenized in ice-cold methanol (1 mL) using a dounce homogenizer. Chloroform (2 mL) was added, the mixture was vortexed for 2 min, and phase separation was induced by adding Milli-Q water (1 mL), followed by vortexing and a 5-min rest. The lower organic phase (lipids) and upper aqueous phase (metabolites) were separately collected, air-dried, and resuspended in solvent A prior to LC-MS analysis.

LC-MS/MS analysis

Untargeted metabolomic and lipidomic profiling was performed using a Dionex U3000 UHPLC system coupled to a Q-Exactive High-Field Orbitrap (Thermo Scientific). Metabolites and lipids were separated on a 2.1 × 50 mm, 1.8 µm C18 column. Mobile phases consisted of 10 mM ammonium acetate in 80% isopropanol/10% acetonitrile (solvent A) and isopropanol (solvent B) at a flow rate of 300 µL/min. The gradient started at 30% B, increased to 35% B over 2 min and to 100% B over 8 min, held at 100% B for 9 min for column washing, followed by 3-min re-equilibration at 30% B. Data were acquired in positive and negative electrospray ionization modes (ESI+/ESI–) with full-scan MS at 60,000 resolution (m/z 100–1500) and data-dependent MS/MS acquisition.

Data processing

Raw MS data were processed using Compound Discoverer v3.3 (Thermo Fisher Scientific), enabling peak detection, alignment, and adduct deconvolution. Metabolite and lipid identification was achieved by matching accurate mass to the BioCyc database and by comparing retention times and MS/MS fragmentation profiles against the mzCloud spectral library. Features with reproducible retention time, accurate mass (m/z), and MS/MS confirmation across replicates were retained. Putative identifications were assigned based on spectral matching confidence tiers.

Data preprocessing and normalization

Each omics layer was processed independently prior to integration [1]: Features with complete missingness were removed [2]. Sporadic missing values were imputed using mean imputation within each group [3]. Feature matrices were standardized with z-score prior to PCA and MOFA+. Given the limited sample size (n = 6 total), preprocessing steps were optimized to minimize noise amplification while preserving condition-specific variation.

Feature ranking via expectation–maximization feature ranking (EMFR)

To prioritize biologically relevant signals, a supervised Expectation–Maximization Feature Ranking (EMFR) approach was applied separately to each omics layer. Features were ranked based on their discriminative power between Pos and Neg samples. The top 100 ranked features per modality were retained for downstream dimensionality reduction and integration. This supervised filtering strategy was implemented to reduce dimensionality and enhance detection of condition-relevant variation in small-sample, high-dimensional datasets. We acknowledge that this approach introduces bias towards condition-associated signals; therefore, latent factor interpretation was performed conservatively. The complete list of EMFR-ranked features for each omics layer, including feature identifiers and ranking scores, is provided in Supplementary Table S2. PCA was performed separately for each omics layer using sklearn.decomposition. PCA (Python v3.11; scikit-learn v1.3.0). The top two principal components were extracted. Variance explained by PC1 and PC2 were recorded. Sample clustering was visualized to assess condition separability.

Multi-omics factor analysis (MOFA+)

Multi-omics integration was performed using MOFA2 (v1.8) [21,22]. The EMFR-selected features were formatted as separate 'views' (features × samples) for Transcriptome, Proteome, Lipidome, Metabolome, and Secretome. Only shared biological replicates (n = 6) were included. Models were trained with up to K = 5 latent factors and maxiter = 1000. The following outputs were extracted sample-level factor scores, feature weights per factor and view, and variance explained per factor and per view. Associations between latent factors and experimental condition were assessed using the Wilcoxon rank-sum test with Benjamini–Hochberg correction.

Feature weights were extracted from the trained MOFA+ model for each factor–view combination. Top-weighted features were defined as those with absolute weight values in the top 10% per factor–view pair; these features were used for downstream GO enrichment analysis and network construction. The complete feature weight matrix for all retained factors across all five views is provided in Supplementary Table S3.

Factor 1, the dominant condition-associated axis, is characterized in the transcriptomic view by upregulation of ECM structural genes (including Col1a1, Col3a1, Thbs2, Lama1) and lipid metabolic enzymes (including Lpl, Fasn, Scd1), and in the proteomic view by ECM components and cytoskeletal remodelling proteins. Factors 2–4 were not condition-dependent and captured modality-specific residual and inter-replicate variation. The complete annotated feature lists per factor are provided in Supplementary Table S3.

Gene ontology (GO) enrichment analysis

GO over-representation analysis was performed on top-weighted features (primarily Factor 1 drivers) for each modality. Protein identifiers were mapped to gene symbols using UniProt ID mapping. Metabolite and

lipid features were mapped to enzymes/genes using LIPEA (KEGG/LIPID MAPS). Enrichment analysis was conducted using g:Profiler (g:GOSt) with BH correction. All terms with FDR < 0.05 were considered significant. Top enriched Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories were visualized as −log10(FDR). The complete list of all statistically significant GO enrichment results across all omics layers and GO categories is provided in Supplementary Table S1.

Network construction and hub identification

Interaction networks were generated using STRING v12.0 (Mus musculus; taxonomy ID 10,090). Only interactions with a combined confidence score ≥0.7 were retained. Networks were constructed using NetworkX (v3.2.1). Nodes were annotated by omics layer, and node size was scaled according to degree of centrality, where node size is proportional to the number of interaction partners (edges) normalized by the maximum possible degree in the network. Additional centrality metrics (betweenness, eigenvector centrality) were calculated. To prioritize core regulators while maintaining balanced representation, hub subnetworks were extracted by selecting the top 20% of nodes ranked by degree centrality within each omics layer. Hub subnetworks were defined as the subgraph induced by nodes with degree centrality values within the top 20% of all nodes in the full condition-associated interaction network, consistent with established network biology frameworks for identifying topologically central regulators [31]. Hub subnetwork visualization was performed in Cytoscape v3.10. The complete node statistics including degree, betweenness, and eigenvector centrality values are provided in Supplementary Table S4.

qPCR validation

Selected hub genes (Lpl, Itga5, Pik3cg, Igfbp6, Plg, Acer3) were validated using quantitative PCR under identical D30_Pos and D30_Neg conditions ($n = 3$ per group). These six candidates were prioritized for validation based on a multi-criteria scoring strategy: (1) high degree centrality in the hub subnetwork (top 10% of all nodes); (2) cross-layer representation, defined as detectable signal in at least two omics layers (e.g. transcriptome and proteome, or proteome and secretome); (3) literature-supported biological relevance to adipogenesis or metabolic disease; and (4) availability of validated qPCR primer sets. Candidates satisfying all four criteria were selected for experimental confirmation. Relative expression levels were calculated using the ΔΔCt method. Statistical comparisons were performed using unpaired two-tailed t-tests. Data are presented as mean ± SEM. Raw VST expression values for validated candidates across all biological replicates are provided in Supplementary Table S5.

Statistical analysis and reproducibility

All statistical analyses were conducted in R (v4.3.1) and Python (v3.11). Multiple testing corrections were performed using the Benjamini–Hochberg method where applicable. All custom scripts, processed matrices, and network statistics are publicly available at: https://github.com/mxa1429/msc-bioinformatics-mohamed-alsayegh. Processed omics data matrices are provided in Supplementary Tables S1–S5. Raw mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE repository. Accession numbers are provided in the Data Availability Statement.

Results

Multi-layer profiling reveals robust condition-specific molecular remodelling during mESC adipogenesis

To characterize terminal adipogenic differentiation, we profiled mouse embryonic stem cell–derived cultures at day 30 (D30) under adipogenic (Pos) and non-differentiating (Neg) conditions across five molecular layers: transcriptome, proteome, lipidome, metabolome, and secretome (Figure 1). All datasets were generated from matched biological replicates ($n = 3$ per condition), enabling cross-layer integration. Principal component analysis (PCA) performed independently on the top 100 EMFR-ranked features per modality

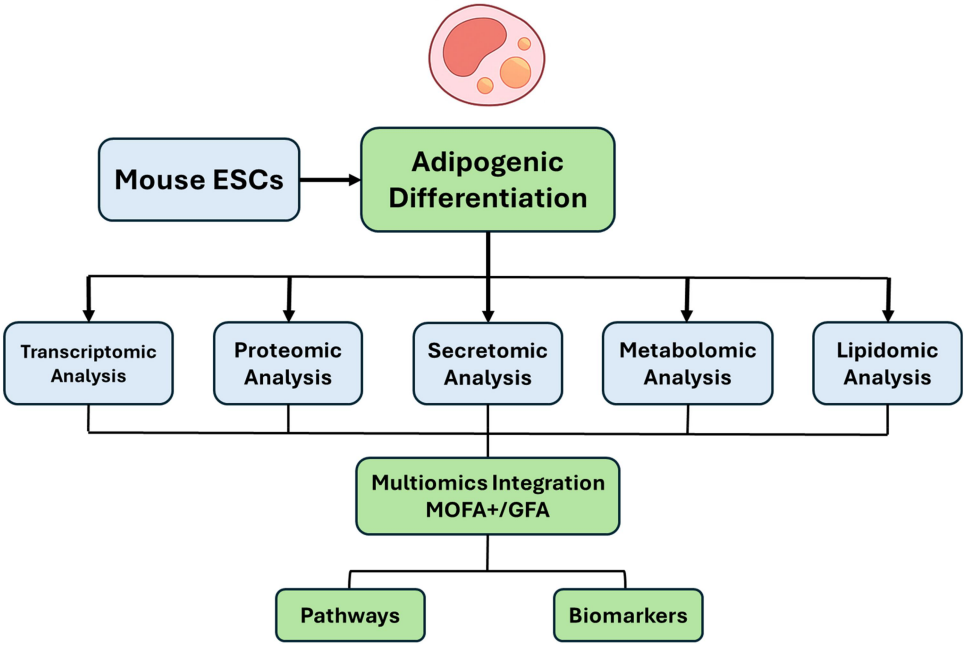

Figure 1. Study design and integrative multi-omics analysis workflow. Schematic overview of the experimental and computational framework used to investigate adipocyte differentiation from mouse embryonic stem cells (mESCs): cells were cultured under adipogenic (Pos) and non-differentiating (Neg) conditions and profiled at day 30 across five molecular layers: transcriptomics, proteomics, secretomics, metabolomics, and lipidomics. Layer-specific datasets were independently processed and subsequently integrated using Multi-Omics Factor Analysis (MOFA+) to identify shared and modality-specific sources of variation. Downstream analyses included functional enrichment, interaction network construction, hub prioritization, and targeted qPCR validation of candidate regulatory genes.

demonstrated clear segregation of Pos and Neg samples across all layers (Figure 2). Transcriptomic data exhibited the strongest separation (PC1 = 88.9%), indicating dominant transcriptional remodelling. Proteomic and lipidomic datasets also showed robust discrimination along PC1, whereas secretomic separation, while present, explained comparatively less variance. These findings confirm that adipogenic differentiation at D30 induces coordinated and detectable molecular shifts across intracellular and extracellular compartments.

MOFA+ identifies a dominant shared latent axis associated with adipogenic differentiation

To determine whether condition-specific variation is shared across omics layers or modality-restricted, we applied Multi-Omics Factor Analysis (MOFA+) to the integrated dataset (Figure 3). Four latent factors were retained based on model convergence and variance structure.

Variance decomposition revealed that Factor 1 accounted for the largest proportion of variance across transcriptomic and proteomic layers, with substantial contributions from lipidomic and metabolomic views (Figure 3(A)). In contrast, secretomic variance was distributed across Factors 1, 2, and 3 (Figure 3 (A)), with Factor 1 contributing a smaller proportion of total secretomic variance relative to its dominant contribution to transcriptomic and proteomic layers. This distribution pattern indicates partial independence of extracellular remodelling from the dominant intracellular differentiation axis. Overall variance explained was highest in transcriptomics and proteomics, followed by lipidomics and metabolomics (Figure 3(B)). Association testing demonstrated that Factor 1 was strongly associated with experimental condition (Wilcoxon $p = 2.9 \times 10^{-5}$; BH-adjusted $p = 2.9 \times 10^{-5}$), whereas Factors 2–4 were not condition-dependent (Figure 4(A)). Projection of samples in factor space demonstrated complete segregation of Pos and Neg samples along Factor 1 (Figure 4(B); all sample identifiers are fully displayed in the revised figure). Increasing the number of latent factors improved total explained variance but did not alter the

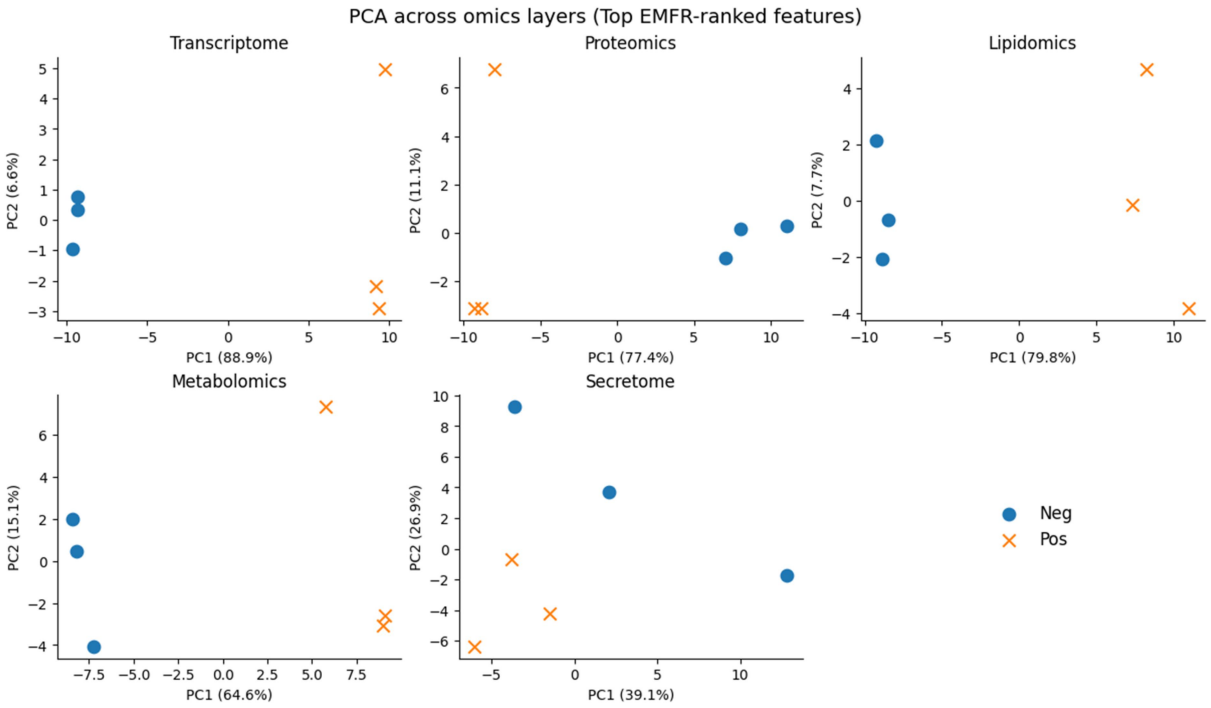

Figure 2. Principal component analysis (PCA) demonstrates condition-specific separation across individual omics layers: PCA was performed independently for each omics modality using the top 100 features ranked by Expectation–Maximization Feature Ranking (EMFR). Plots show Transcriptome, Proteomics, Lipidomics, Metabolomics, and Secretome datasets from differentiated (Pos, orange crosses) and non-differentiated (Neg, blue circles) samples ($n = 3$ per group). Across all layers, samples segregate along PC1, indicating robust condition-associated molecular variation. The percentage of variance explained by PC1 and PC2 is indicated on each axis. Transcriptomic data exhibits the strongest condition-specific separation (PC1 = 88.9%), followed by proteomics (PC1 = 77.4%), lipidomics (PC1 = 79.8%), and metabolomics (PC1 = 64.6%). The secretome exhibited comparatively lower PC1-driven separation (PC1 = 39.1%), consistent with the partial independence of extracellular remodelling from the dominant intracellular differentiation axis identified by MOFA+ Factor 1.

dominance of Factor 1 (Figure 4(C)). Importantly, quality control indicated partial correlation between Factors 1–2 and total detected features in some modalities, consistent with known normalization sensitivity in small-sample integrative frameworks. However, the magnitude of condition separation and cross-layer concordance strongly support Factor 1 as a biologically meaningful differentiation axis rather than a technical artefact.

Together, these results define a shared multi-omics differentiation programme linking transcriptional activation, proteomic remodelling, and lipid metabolic reprogramming.

Functional enrichment highlights extracellular matrix remodelling and lipid metabolic reprogramming

To interpret the biological processes driving the dominant differentiation axis, we performed Gene Ontology enrichment analysis on top-weighted features per modality (Figure 5). The complete list of all statistically significant enriched terms (FDR <0.05) across all omics layers and GO categories (Biological Process, Cellular Component, Molecular Function) is provided in Supplementary Table S1.

Across transcriptomic datasets, the top enriched Biological Process terms included cell migration, anatomical structure development, cell motility, system development, and blood vessel morphogenesis (Figure 5). In the proteomic dataset, enriched Biological Process terms included response to wounding, wound healing, supramolecular fibre organization, tissue development, and actomyosin structure organization. Cellular Component categories were dominated by collagen-containing extracellular matrix, external encapsulating structure, and extracellular region. Molecular Function terms highlighted extracellular matrix structural constituent, growth factor binding, and protein–protein interaction activity.

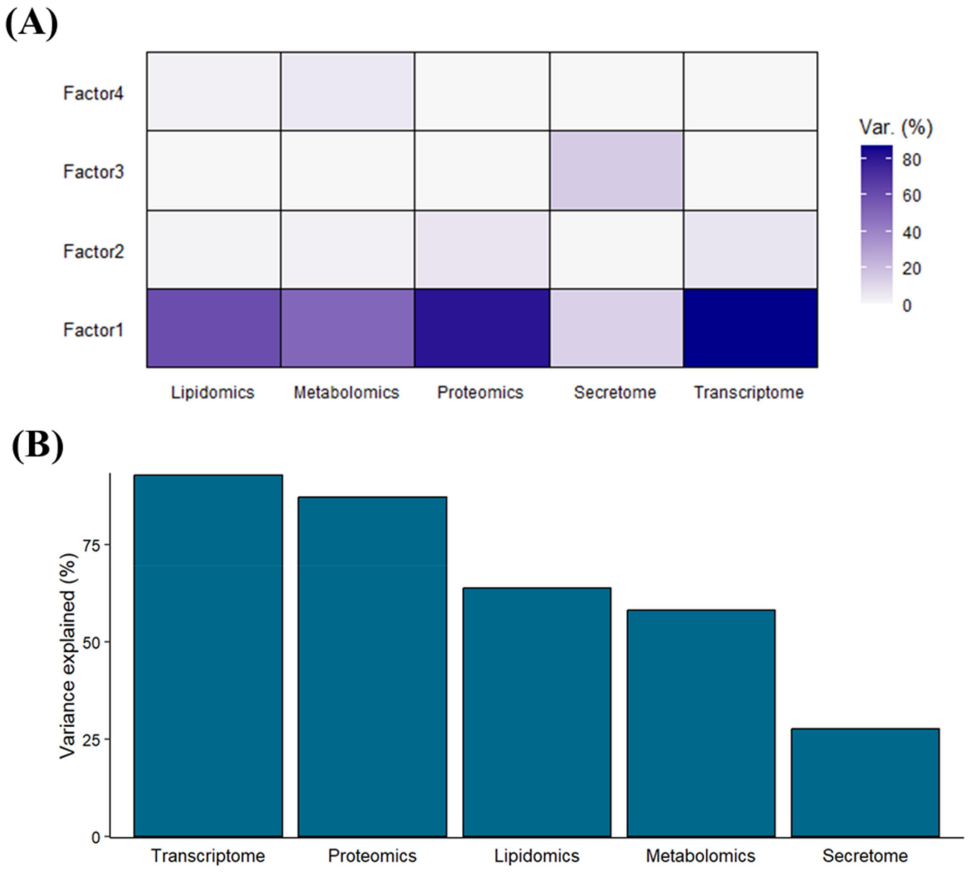

Figure 3. MOFA+ variance decomposition reveals a dominant shared differentiation axis across molecular layers: (A) Heatmap showing the percentage of variance explained by each latent factor (Factors 1–4) across the five omics views (Lipidomics, Metabolomics, Proteomics, Secretome, Transcriptome). Color intensity reflects variance explained (%), ranging from 0% (white) to >80% (dark blue). Factor 1 accounts for the dominant proportion of variance in the Transcriptome and Proteomics views, with substantial contributions to Lipidomics and Metabolomics. In contrast, secretomic variance is distributed across Factors 1 and 3, indicating partial independence of extracellular remodeling from the dominant intracellular differentiation axis. (B) Total variance explained per omics layer summed across all retained latent factors. Transcriptomics (~92%) and proteomics (~87%) exhibit the highest total explained variance, followed by lipidomics (~65%) and metabolomics (~60%), with the secretome showing the lowest total explained variance (~28%), consistent with its greater modality-specific variation.

Secretomic enrichment strongly reinforced extracellular matrix localization and external encapsulating structure categories, indicating active remodelling of the extracellular microenvironment. In contrast, lipidomic and metabolomic layers were dominated by lipid metabolic processes, including lipid metabolic process, phospholipid and glycerolipid metabolism, organophosphate metabolism, and phospholipase and lipase activity.

These results reveal two coordinated but distinct biological themes [1] structural and extracellular remodelling, and [2] lipid metabolic reprogramming. This dual remodelling signature suggests that adipogenic differentiation involves not only intracellular transcriptional shifts but also extracellular matrix restructuring coupled to lipid biosynthetic activation.

Hierarchical clustering reveals coordinated cross-layer molecular restructuring

Unsupervised hierarchical clustering of z-scored features confirmed coordinated remodelling across modalities (Figure 6). In all four visualized layers (transcriptome, proteome, lipid/metabolome, secretome), Pos samples clustered distinctly from Neg samples. As shown in Figure 6, the functional category distribution of upregulated and downregulated features per omics layer is shown in Figure 6(B), demonstrating that ECM/adhesion features dominate upregulated proteins in the proteome (67%) and secretome (53%), while lipid

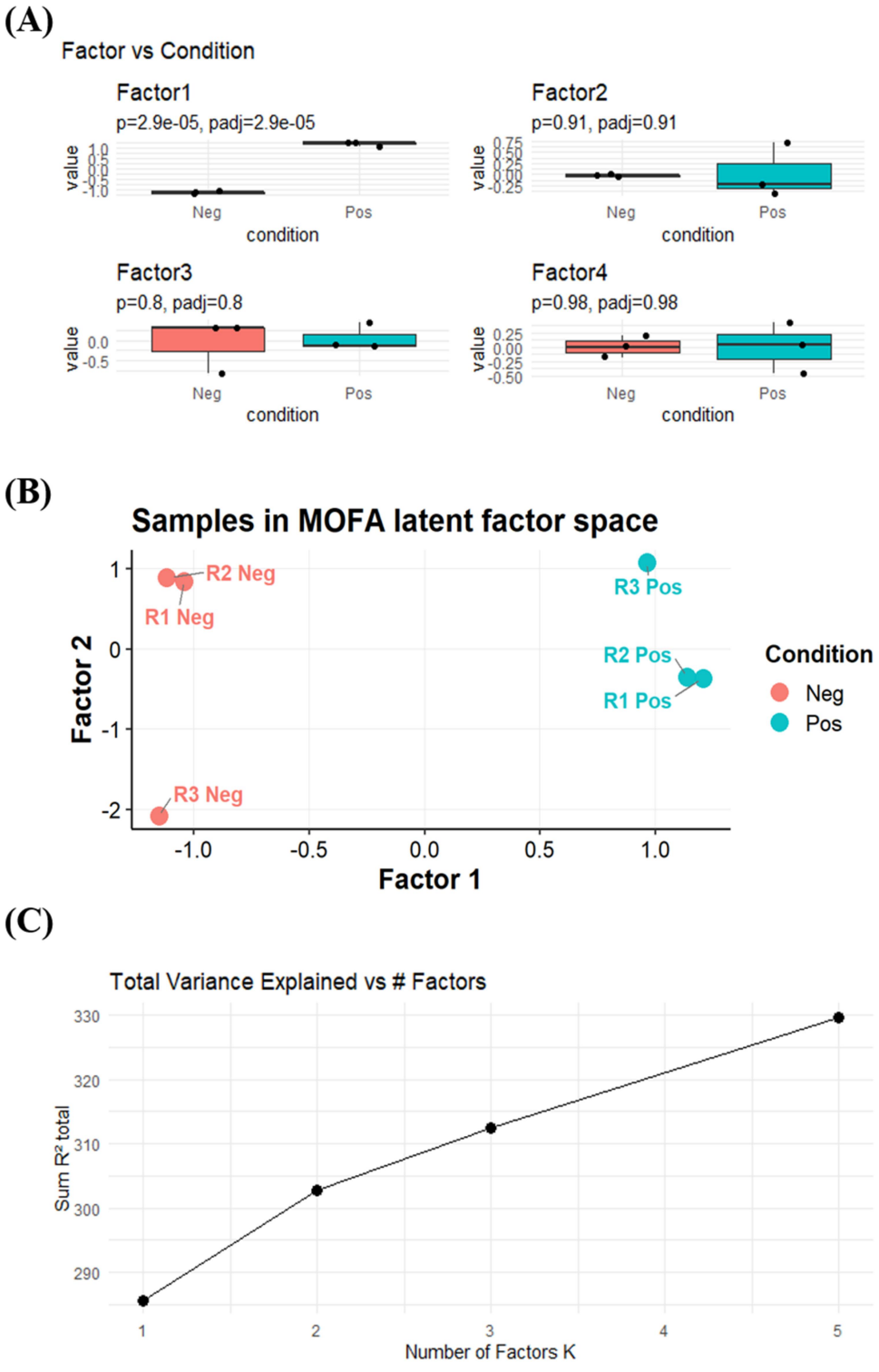

Figure 4. MOFA+ factor scores identify a dominant condition-associated latent factor across all omics layers: (A) Boxplots showing the association of MOFA+ latent factors (factors 1–4) with experimental condition (Neg, red; Pos, teal). Statistical significance was assessed using the Wilcoxon rank-sum test with Benjamini–Hochberg correction. Factor 1 shows strong and significant separation between conditions ($p = 2.9 \times 10^{-5}$; padj $= 2.9 \times 10^{-5}$), whereas factors 2–4 are not condition-dependent ($p = 0.91$, 0.80, and 0.98, respectively). Points represent individual samples ($n = 3$ per group). (B) Projection of all six biological replicates in MOFA+ latent factor space (Factor 1 vs Factor 2), demonstrating complete segregation of Pos (teal) and Neg (red) samples along Factor 1. Sample labels (R1–R3 per condition) are fully displayed. Inter-replicate variation along Factor 2 within the Neg condition, particularly for R3 Neg, reflects expected biological heterogeneity rather than technical artifact, as Factor 2 was not condition-associated (padj = 0.91). (C) Total variance explained (Sum R^2 total) as a function of the number of retained latent factors (K = 1–5), demonstrating incremental gains in explained variance with additional factors while confirming the dominant contribution of Factor 1 at K = 1.

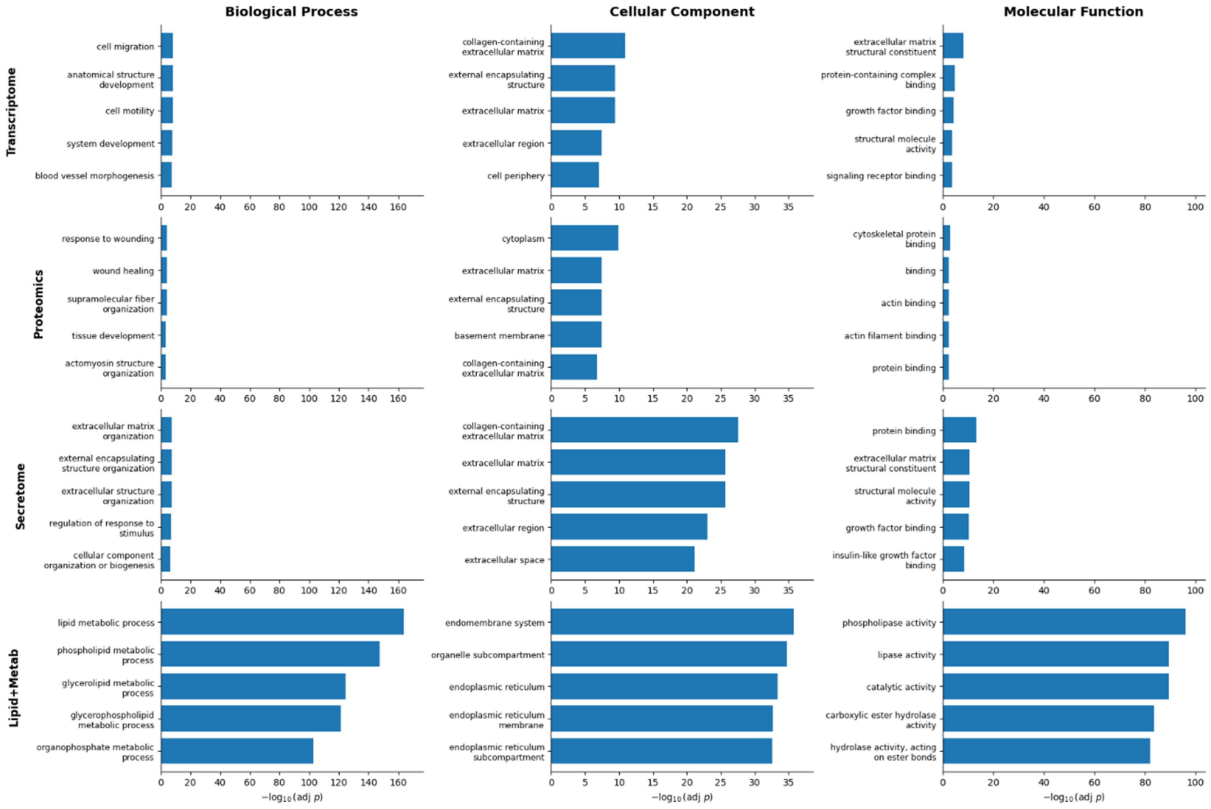

Figure 5. Gene Ontology enrichment analysis across omics layers highlights extracellular matrix remodelling and lipid metabolic reprogramming during adipogenic differentiation: GO over-representation analysis was performed on condition-associated top-weighted features from each omics layer. The five top enriched terms ranked by adjusted *p*-value (BH correction; FDR < 0.05) are shown for Biological Process (left), Cellular Component (middle), and Molecular Function (right) categories across Transcriptome, Proteomics, Secretome, and Lipidome & Metabolome layers. Bar lengths represent $-\log_{10}$ (adjusted *p*-value). Transcriptomic and proteomic layers are enriched for ECM organization, tissue development, and structural molecule activity. The secretome shows strong enrichment for extracellular matrix and external encapsulating structure categories, indicating active remodelling of the extracellular microenvironment by differentiating adipocytes. Lipid and metabolomic layers are dominated by lipid metabolic processes including phospholipid, glycerolipid, and organophosphate metabolism, as well as phospholipase and lipase activities. Displayed terms represent the highest-ranked enrichment results by statistical significance; the complete list of all significant GO terms across all layers and categories is provided in Supplementary Table S1.

metabolism accounts for 100% of upregulated and 92% of downregulated features in the lipidome/metabolome.

Clusters enriched in Pos samples included extracellular matrix components, lipid metabolic enzymes, and growth factor–associated proteins. Conversely, Neg-enriched clusters contained features associated with alternative cellular programmes, including cytoskeletal organization and non-adipogenic structural signatures. These findings reinforce the presence of a coherent differentiation programme spanning intracellular metabolic rewiring and extracellular structural adaptation.

Network integration identifies lipid–ECM interaction hubs across molecular layers

To move beyond enrichment lists and identify central regulators, we constructed an integrated interaction network using STRING-derived high-confidence interactions (combined score ≥0.7). The top 20% highest-degree nodes per modality were extracted to define a balanced hub subnetwork (Figure 7).

The resulting network revealed a dense lipid metabolic module centred on phospholipases (Pla2g family), ceramide synthases (Cers), acyltransferases (Mboat family), and lipid-processing enzymes (Plpp, Lpcat, Dgk). The ECM component cluster — comprising collagens (Col1a1, Col3a1, Col5a2), laminins (Lama1, Lamb1), and

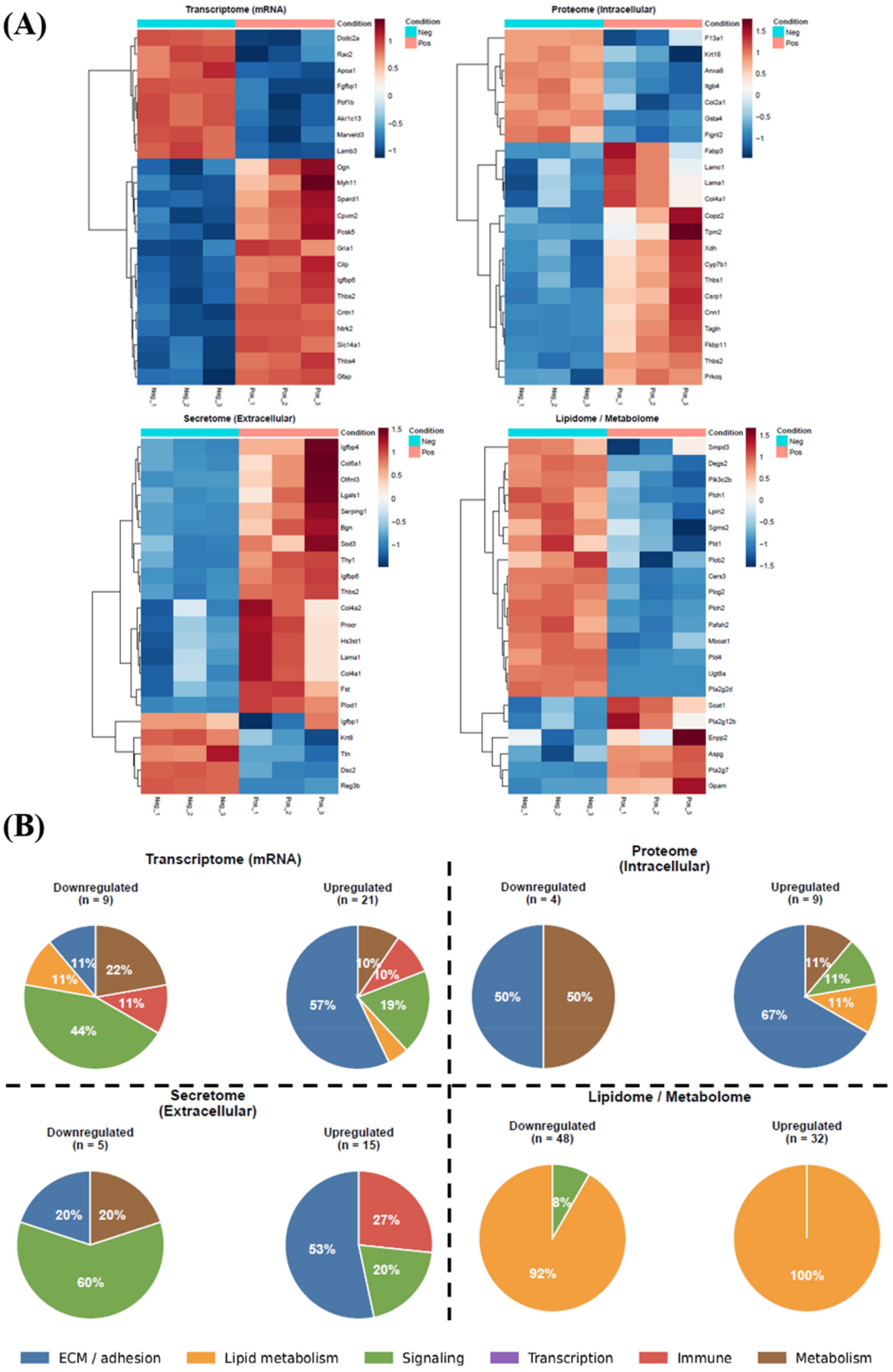

Figure 6. Hierarchical clustering reveals coordinated molecular remodelling across omics layers during adipogenic differentiation: (A) Heatmaps showing scaled expression and abundance profiles of the top 22 condition-associated features per omics layer, selected by EMFR ranking score, across Transcriptome (mRNA), Proteome (Intracellular), Secretome (Extracellular), and Lipidome/Metabolome datasets. Rows represent individual molecular features and columns represent biological replicates (Neg, teal; Pos, red; $n = 3$ per group). Data were z-score scaled per feature to highlight relative changes across conditions. Hierarchical clustering was performed using Euclidean distance and complete linkage. Across all four

Thbs2 — was connected to the broader network primarily through bridging nodes including Plod1, Thbs2, Plg, and Pmp22, which share interaction partners with lipid metabolic and signalling nodes. While direct high-confidence edges between lipid enzymatic hubs and ECM structural proteins were limited in the STRING network, indirect connectivity through shared bridging partners suggests functional coupling between these two modules. Cross-layer links additionally involved secreted signalling mediators, including Plg and growth factor–binding proteins, extending network connectivity across omics layers. A smaller disconnected subnetwork (bottom left, Figure 7) includes cytoskeletal and growth factor-associated candidates (Tpm1, Tpm2, Myl3, Igfbp6), reflecting their partial connectivity within the broader network at the applied confidence threshold.

Notably, hubs were not restricted to classical adipogenic transcription factors. Instead, lipid metabolic enzymes and extracellular matrix regulators formed the structural backbone of the network, suggesting that metabolic–structural coupling is central to adipocyte commitment.

Targeted qPCR validation confirms integrative multi-omics prioritization

To validate candidates emerging from integrative and network analyses, we performed qPCR on six prioritized hub genes: Lpl, Itga5, Pik3cg, Igfbp6, Plg, and Acer3 (Figure 8). Independent qPCR measurements demonstrated consistent directional changes between Pos and Neg conditions across all candidates. Lpl and Igfbp6 were significantly upregulated in differentiated cells, while Plg, Acer3, Itga5, and Pik3cg showed significant downregulation, and all candidates displayed directional concordance between RNA-seq VST and qPCR measurements, with variable magnitude of change. Statistical testing confirmed significance for selected targets ($*p < 0.05$; $**p < 0.01$). These results confirm mRNA-level concordance between RNA-seq and independent measurements and support the robustness of the EMFR + MOFA+ + network-based prioritization strategy as a hypothesis-generating framework for regulatory candidate nomination. RNA-seq VST expression values and protein-level abundance data (where detected in the proteomic dataset) for all six candidates are provided in Supplementary Table S5.

Integrated model of adipogenic differentiation from pluripotency

Collectively, these results define adipogenic differentiation from mESCs as a coordinated multi-layer transition characterized by (1) A dominant shared latent differentiation axis (MOFA Factor 1), (2) convergent extracellular matrix remodelling and lipid metabolic activation, (3) network centrality of lipid enzymes and ECM regulators, emerging alongside, rather than replacing, canonical PPARγ-driven transcriptional regulators; and (4) mRNA-level validation of prioritized cross-layer hub candidates. This integrative framework demonstrates that adipogenesis is not merely transcriptional activation of PPARγ-driven programmes, but a systems-level restructuring linking intracellular lipid biosynthesis to extracellular architectural remodelling and secreted signalling dynamics.

---

layers, differentiated (Pos) samples cluster distinctly from non-differentiated (Neg) samples, with coordinated upregulation of ECM-related and lipid metabolic features and reciprocal downregulation of alternative cellular programs. The complete ranked feature lists for all layers are provided in Supplementary Table S2. (B) Pie charts showing the functional category distribution of upregulated and downregulated condition-associated features per omics layer, annotated by six functional categories: ECM/adhesion (blue), Lipid metabolism (orange), Signaling (green), Transcription (purple), Immune (red), and Metabolism (brown). In the Transcriptome, lipid metabolism dominated upregulated features (57%), while ECM/adhesion represented the largest downregulated category (44%). In the Proteome, ECM/adhesion dominated upregulated features (67%). In the Secretome, ECM/adhesion dominated upregulated features (53%), while lipid metabolism and signaling accounted for the downregulated fraction. In the Lipidome/Metabolome, lipid metabolism accounted for 100% of upregulated and 92% of downregulated features, reflecting the layer-specific nature of lipid metabolic reprogramming during adipogenic differentiation. These patterns reinforce the two principal biological themes identified by GO enrichment analysis: structural ECM remodelling and broad lipid metabolic reprogramming.

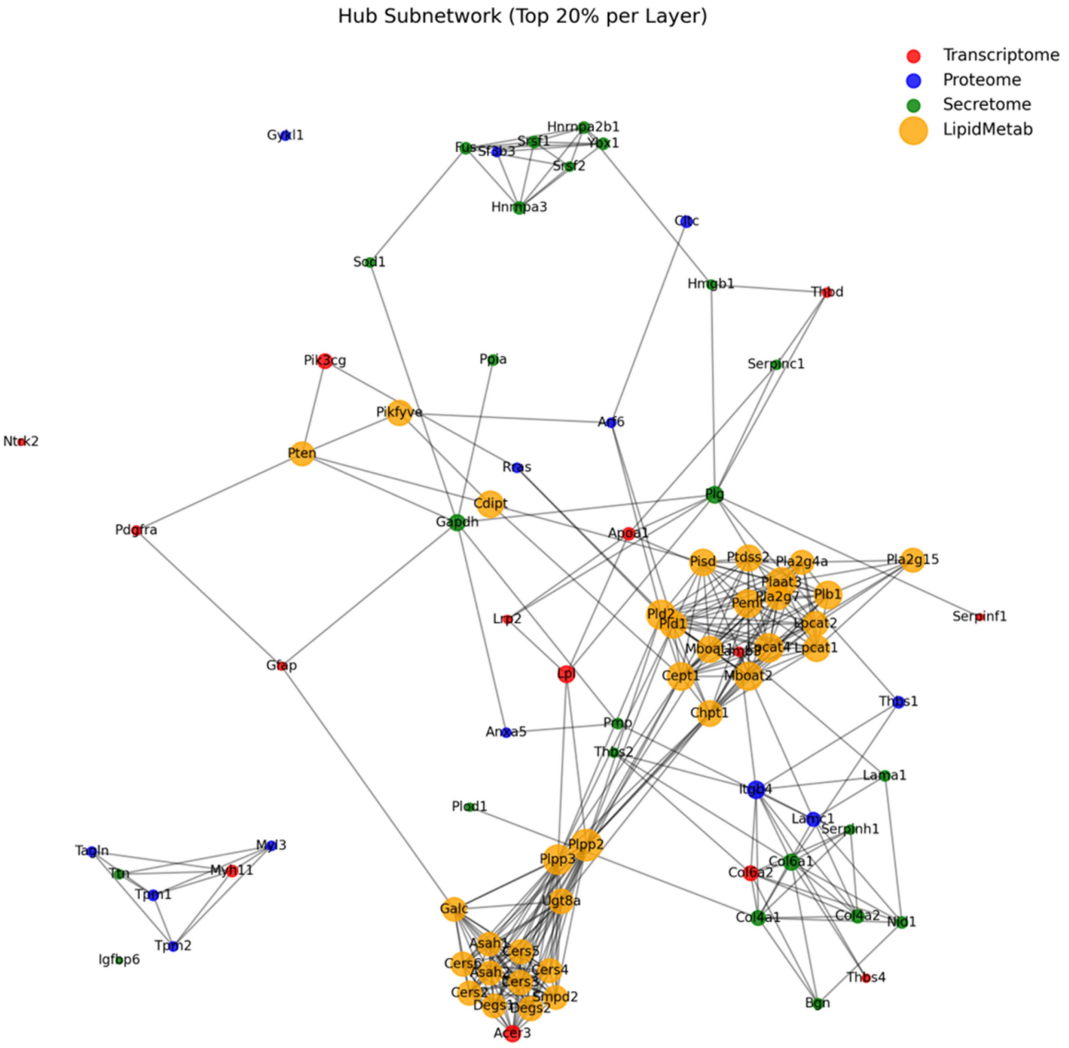

Figure 7. Integrated multi-omics hub subnetwork highlighting cross-layer regulatory connectivity during adipogenic differentiation: Network representation of condition-associated hub features constructed from high-confidence STRING interactions (combined score ≥0.7; Mus musculus). Nodes represent molecular features from four omics layers: Transcriptome (red), Proteome (blue), Secretome (green), and Lipidome/Metabolome (orange). Node size is proportional to degree centrality, where larger nodes indicate a greater number of high-confidence interaction partners within the full condition-associated network, scaled to the maximum possible degree. Edges represent known or predicted protein–protein interactions derived from the STRING database. The subnetwork displays the top 20% highest-degree nodes per omics layer. The network reveals a dense lipid metabolic module (orange) centred on phospholipases (Pla2g family), ceramide synthases (Cers, Degs, Smpd), acyltransferases (Mboat family), and lipid-processing enzymes (Plpp2, Plpp3, Lpcat1, Lpcat2, Ugt8a). A structurally distinct ECM component cluster (right) comprises collagens (Col4a1, Col4a2, Col6a1, Col6a2), laminins (Lama1, Lamb2), and associated proteins (Itgb4, Serpinh1, Bgn, Thos4). Connectivity between the ECM cluster and the lipid metabolic module occurs primarily through bridging nodes including Plod1, Thbs2, Plg, and Pmp22, which share interaction partners with both lipid metabolic and ECM structural nodes, rather than through direct high-confidence edges between lipid enzymes and ECM structural proteins. A smaller disconnected subnetwork (bottom left) includes cytoskeletal and growth factor-associated candidates (Tpm1, Tpm2, Myl3, Igfbp6), reflecting their partial connectivity within the broader network at the applied confidence threshold. Cross-layer hub nodes including Plg, Plod1, and Thbs2 bridge multiple omics layers, supporting their nomination as multi-modal regulatory candidates.

Discussion

Adipose tissue is a metabolically active endocrine organ whose dysfunction underpins the pathogenesis of obesity, type 2 diabetes mellitus (T2DM), and insulin resistance [1,2]. At the cellular level, these pathologies are rooted in disturbances of adipocyte differentiation, expansion, and functional maturation. Despite decades of molecular dissection, the full regulatory complexity of adipogenesis remains incompletely

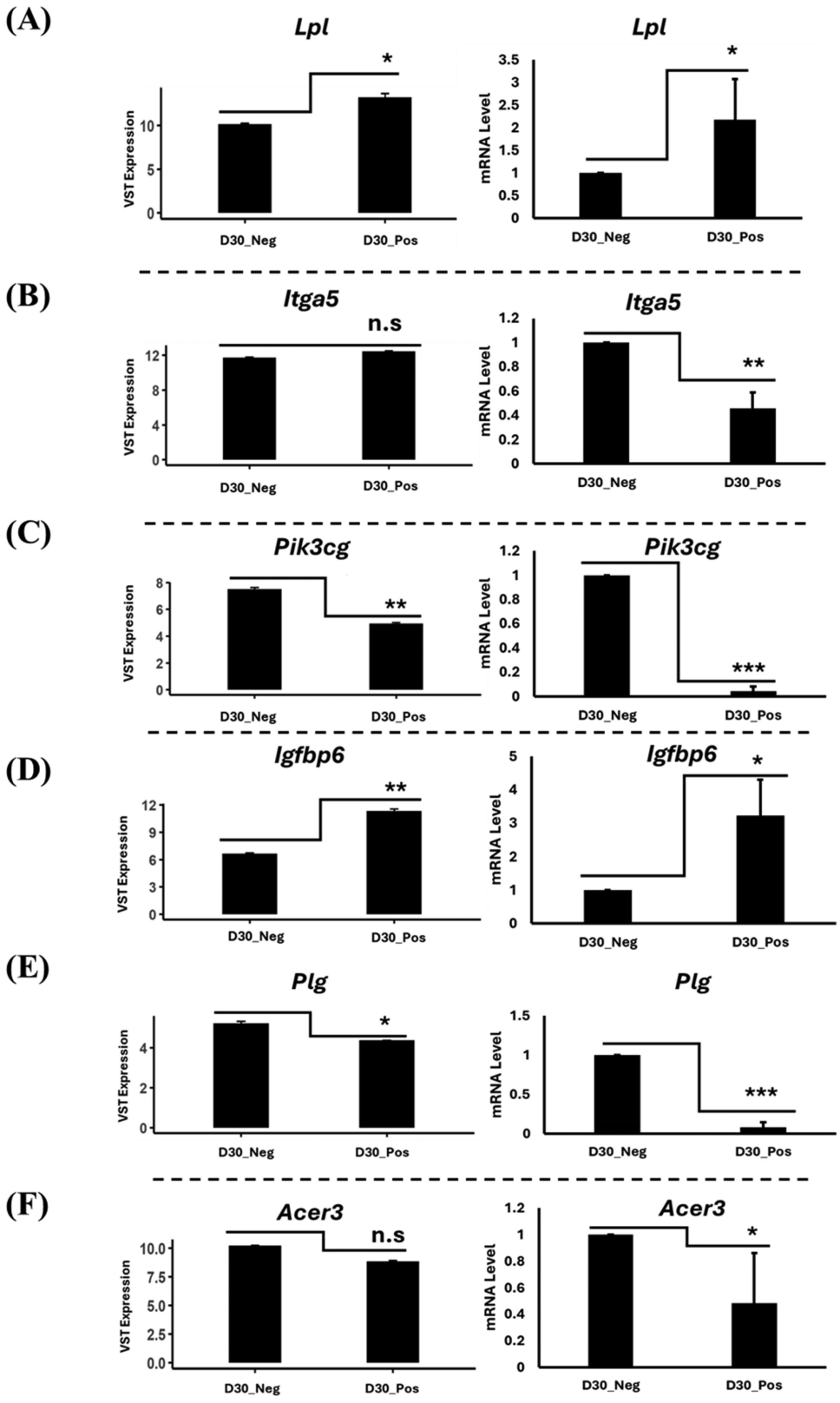

Figure 8. Targeted qPCR validation confirms mRNA-level concordance of integrative multi-omics prioritization: Comparison of RNA-seq variance-stabilized transformed (VST) expression values (left panels) and independent quantitative PCR (qPCR) relative mRNA levels (right panels) for six prioritized hub candidates at day 30 under non-differentiating (D30_Neg) and adipogenic (D30_Pos) conditions. Candidates were selected based on degree centrality, cross-layer representation, literature relevance, and primer availability (see Methods). Bar plots represent mean ± SEM ($n = 3$ per group). Statistical significance was assessed using unpaired two-tailed Student's t-test (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; n.s., not significant). (A) Lpl shows significant upregulation in D30_Pos by both VST and qPCR. (B) Itga5 shows non-significant change by VST but

defined, in part because single-omics approaches capture only one dimension of a fundamentally multi-layered biological transition [5,8]. In this study, we applied an integrative multi-omics framework to mESC-derived adipogenesis, combining transcriptomic, proteomic, secretomic, metabolomic, and lipidomic profiling at day 30 under differentiating and non-differentiating conditions. Through MOFA+-based factor decomposition [21,22], functional enrichment, interaction network analysis, and qPCR validation, we identified a dominant, cross-layer differentiation axis converging on two principal biological themes: extracellular matrix (ECM) remodelling and lipid metabolic reprogramming. Critically, network hub analysis revealed that lipid metabolic enzymes and ECM regulators constitute high-degree hubs alongside canonical PPARγ-driven transcriptional regulators, underscoring the multi-layered nature of the adipogenic regulatory landscape. These findings reframe adipogenesis as a systems-level restructuring event and nominate candidate regulators of direct relevance to metabolic disease.

A dominant shared multi-omics differentiation axis

The centrepiece of our integrative analysis is the identification of MOFA+ Factor 1 as a dominant latent axis strongly associated with adipogenic condition across all five molecular layers (Wilcoxon $p = 2.9 \times 10^{-5}$; BH-adjusted $p = 2.9 \times 10^{-5}$). Factor 1 explained the largest proportion of variance in transcriptomic and proteomic data, with substantial contributions from lipidomic and metabolomic views, while secretomic variance was more distributed across additional factors, suggesting partial independence of extracellular remodelling from the dominant intracellular axis. The ability of a single latent factor to capture coordinated condition-associated variation across five distinct molecular modalities is consistent with a hierarchically organized differentiation programme in which transcriptional activation propagates downstream to enforce proteomic and lipid metabolic reconfiguration [5,8,9].

PCA performed independently on EMFR-ranked features confirmed robust condition-specific separation across all layers, with transcriptomics exhibiting the strongest discrimination (PC1 = 88.9%), consistent with the established primacy of transcriptional control in orchestrating PPARγ and C/EBP-driven adipogenic cascades [9,10]. The comparatively lower but present secretomic separation is consistent with the dynamic and context-sensitive nature of adipokine and ECM factor secretion, which is modulated by developmental stage, cell density, and paracrine signalling [15,16].

The partial independence of the secretomic layer from Factor 1 carries important biological implications beyond technical observation. Extracellular remodelling during adipogenesis involves not only transcriptionally regulated ECM gene expression but also post-translational processing, vesicle-mediated secretion, and unconventional secretory routes that may be governed by regulatory circuits distinct from the dominant intracellular transcriptional programme. Three mechanisms may account for this independence: (1) temporal asynchrony, whereby secretory programmes are initiated or resolved on a different timescale than the intracellular changes captured at the D30 snapshot; (2) distinct post-transcriptional regulatory circuits, including protein processing and trafficking machinery not captured by standard transcriptomic or proteomic profiling; and (3) cellular heterogeneity within the differentiating culture, where subpopulations at different commitment stages may contribute differentially to the conditioned media secretome. Dedicated multi-time-point secretomic profiling combined with single-cell approaches will be required to fully resolve these possibilities.

Feature pre-selection via EMFR introduces a supervised bias towards condition-discriminative signals. This approach was deliberate and necessary given the high-dimensional, small-sample structure of our dataset ($n = 3$ per condition), but latent factor interpretation should be considered conservative and focused on condition-relevant variation. Future studies with larger biological replicates and unsupervised feature

significant downregulation by qPCR, reflecting post-transcriptional regulation or sensitivity differences between platforms. (C) Pik3cg shows significant downregulation in both VST and qPCR. (D) Igfbp6 shows significant upregulation in both VST and qPCR. (E) Plg shows significant change in both VST and qPCR. (F) Acer3 shows non-significant change by VST but significant downregulation by qPCR, consistent with detection sensitivity differences between bulk RNA-seq and targeted qPCR. Directional concordance between RNA-seq and qPCR across candidates supports the robustness of the EMFR + MOFA + + network-based prioritization strategy as a hypothesis-generating framework.

selection will be required to fully characterize the scope of multi-omics remodelling during mESC adipogenesis.

ECM remodelling: an orchestrated structural program during adipogenesis

Across transcriptomic, proteomic, and secretomic layers, Gene Ontology enrichment analyses converged consistently upon extracellular matrix organization as the most robustly enriched biological process. Enriched terms including collagen-containing extracellular matrix, external encapsulating structure, ECM structural constituent, and growth factor binding were identified across multiple modalities, reflecting coordinated molecular remodelling at both intracellular and secretory levels. These observations are consonant with the established biology of adipocyte differentiation, in which ECM composition actively governs the structural and signalling environment required for successful lipid droplet expansion and morphological maturation [18,19].

The preadipocyte-to-adipocyte transition requires dissolution of a stiff, fibrillar ECM that restricts cell rounding, and its replacement by a compliant matrix compatible with adipocyte morphology and function [18]. Collagens, laminins, fibronectins, and their associated remodelling enzymes are central to this process [19]. Our network analysis demonstrates that ECM components constitute high-degree hubs connected to the broader network primarily through bridging nodes such as Plod1, Thbs2, and Plg, rather than through direct high-confidence interactions with lipid enzymatic hubs. This indirect connectivity nonetheless suggests functional coupling between structural ECM remodelling and lipid metabolic reprogramming, mediated through shared signalling intermediaries. The prominent enrichment of ECM-related terms in the secretome specifically indicates that mESC-derived adipocytes actively remodel their extracellular microenvironment through secretory activity consistent with in vivo observations of adipocyte-driven niche remodelling in adipose tissue [18,19].

In the context of metabolic disease, pathological ECM remodelling — particularly collagen cross-linking and adipose fibrosis — impairs adipocyte insulin sensitivity, reduces metabolic flexibility, and drives chronic low-grade inflammation in obese adipose tissue [2,3]. The identification of ECM remodelling as a central axis of developmental adipogenesis thus provides a mechanistic foundation for understanding how disruption of this structural programme during early commitment may predispose adipocytes to pathological phenotypes. The mESC model employed here [6,11,12] is particularly well-suited to interrogate these early remodelling events, given its ability to recapitulate adipogenesis from a stem cell state through lineage commitment, capturing regulatory stages that precede terminal differentiation in classical models such as 3T3-L1 cells and MEFs [7,10,32,33].

Lipid metabolic reprogramming: a multi-enzyme regulatory network beyond triglyceride storage

Lipidomic and metabolomic enrichment analyses revealed coordinated upregulation of phospholipid metabolism, glycerolipid biosynthesis, organophosphate metabolism, and phospholipase and lipase activities during adipogenic differentiation. These findings demonstrate that lipid metabolic reprogramming during adipogenesis encompasses substantially more than the induction of triglyceride biosynthetic pathways classically attributed to PPARγ activation [9]. The breadth of lipid metabolic enrichment suggests that membrane phospholipid remodelling and production of bioactive lipid species are integral components of the adipogenic programme [8,9].

The interaction network revealed a dense lipid metabolic module centred on phospholipases (Pla2g family), ceramide synthases (Cers), acyltransferases (Mboat family), and lipid-processing enzymes including phospholipid phosphatases (Plpp) and lysophosphatidylcholine acyltransferases (Lpcat). Their emergence as high-degree network hubs suggests that lipid enzymatic activity constitutes an important regulatory layer in adipocyte commitment, consistent with recognized roles of lipid phosphate phosphatases and sphingolipid-metabolizing enzymes in modulating adipose tissue function [30].

Among validated candidates, lipoprotein lipase (Lpl) was significantly upregulated in differentiated cells, corroborating its identity as a canonical PPARγ target gene and mediator of fatty acid uptake from circulating lipoproteins [25], confirming biological fidelity of the D30 differentiation system. Acid ceramidase 3 (Acer3) was significantly downregulated, implicating regulated suppression of ceramide hydrolysis during

adipogenic commitment. Ceramides are bioactive sphingolipids with established roles in insulin resistance, lipotoxicity, and inflammation, and their accumulation in obese adipose tissue contributes to impaired insulin receptor signalling [28]. The downregulation of Acer3 raises the hypothesis that ceramide accumulation may be functionally permissive during early adipogenic commitment, though this requires direct functional validation. Acer3 represents a compelling candidate for future loss-of-function studies in adipogenic differentiation models.

Validated hub candidates: integrin signalling, growth factor modulation, and PI3K connectivity

Three additional validated candidates — Itga5, Igfbp6, and Pik3cg — illustrate the cross-layer regulatory breadth of the integrative pipeline. Itga5 (integrin alpha-5) mediates cell-fibronectin interactions and mechanosensing pathways linking ECM physical properties to cytoskeletal and transcriptional responses [34]. During adipogenesis, integrin expression shifts permit the morphological transition from fibroblastic preadipocyte to rounded adipocyte through cytoskeletal depolymerization and matrix decoupling [18,34]. Morandi et al. demonstrated that ITGA5 knockdown promotes adipogenic commitment in human adipose-derived stem cells [34], consistent with the significant downregulation of Itga5 observed in Pos cells by qPCR in our mESC system (Figure 8(B)), supporting a conserved role for integrin-ECM crosstalk in adipocyte commitment across model systems.

Igfbp6 was identified as a differentially expressed secreted hub with roles in IGF-II bioavailability modulation, immunity, fibrosis, and cellular senescence [29]. Its secretion by differentiating adipocytes suggests a paracrine function in modulating IGF signalling within the adipogenic niche, with translational relevance given dysregulation of IGFBP family members in obese and insulin-resistant adipose tissue [29]. Pik3cg displayed directionally concordant but non-significant expression changes; its emergence as a network hub despite borderline differential expression suggests that network topology captures regulatory relevance not fully reflected in fold-change metrics alone [26,27]. Functional perturbation studies for all three candidates are warranted.

Plasminogen as a novel regulator of the adipogenic extracellular microenvironment

Plasminogen (Plg) was significantly downregulated in differentiated Pos cells relative to Neg controls (Figure 8(E)), despite emerging as a high-connectivity bridging hub between proteomic and secretomic network layers. Plasminogen is the zymogen of plasmin, a serine protease whose activity extends beyond fibrinolysis to MMP activation, ECM-sequestered growth factor release, and integrin-mediated cell adhesion regulation — all directly relevant to ECM remodelling during adipogenesis [18]. The downregulation of Plg in differentiating cells is consistent with the known suppression of fibrinolytic activity during adipocyte maturation, where reduced plasminogen availability may limit excessive ECM degradation and stabilize the nascent adipocyte matrix. This interpretation is also consistent with the well-documented elevation of plasminogen activator inhibitor-1 (PAI-1) in mature adipose tissue, where the fibrinolytic system is actively restrained [2,18]. The high network centrality of Plg as a bridging node despite its downregulation at the mRNA level highlights the importance of incorporating network topology alongside expression magnitude when prioritizing regulatory candidates; protein-level and secretory dynamics of Plg during adipogenesis warrant further investigation.

Integrating findings into a systems model of adipogenesis

The mESC differentiation model employed here provides access to early regulatory events not captured by classical preadipocyte lines; SIRT1-mediated regulation of retinoic acid receptor expression via NCOR1 acetylation has been identified as a key developmental mechanism in this system [35], highlighting the unique biology of stem cell-derived adipogenesis relative to terminal differentiation models. Collectively, our data support a model in which adipocyte commitment from pluripotency is orchestrated through the coordinated action of at least three interacting systems: (1) transcriptional and proteomic remodelling activating ECM structural programmes; (2) lipid metabolic enzymatic networks governing phospholipid,

sphingolipid, and neutral lipid composition; and (3) secretory remodelling of the extracellular microenvironment through ECM components, growth factor modulators, and fibrinolytic mediators.

The lipid–ECM coupling identified in our hub network is of conceptual significance. Phospholipid remodelling influences membrane curvature, vesicle trafficking, and lipid droplet surface dynamics, while ECM mechanical properties regulate integrin signalling pathways that feedback on lipid metabolic enzyme activity through PI3K-Akt and MAPK cascades [26,27,34]. The indirect connectivity between lipid metabolic enzymes and ECM regulators, identified through shared bridging nodes, suggests that these programmes are functionally coupled rather than simply parallel — with important implications for understanding how metabolic and structural disruptions in obese adipose tissue may co-evolve and perpetuate each other [1–3].

Compared to earlier integrative omics studies of adipogenesis largely limited to paired transcriptomic-proteomic or transcriptomic-metabolomic analyses, our five-layer integration provides substantially higher resolution of the cross-modal regulatory landscape [15,21]. The application of MOFA+ to this system, previously deployed primarily in cancer and early T2DM contexts [23], demonstrates the generalizability of this framework for interrogating complex developmental transitions in endocrine tissues.

Limitations and future directions

Several important limitations must be acknowledged. The restricted sample size ($n = 3$ per condition) limits statistical power, precludes robust detection of inter-replicate heterogeneity, and necessitates conservative interpretation of multi-omics outputs. Replication in larger independent cohorts ($n \geq 6$ per condition) will be necessary to confirm the robustness of identified regulatory axes and validate nominated hub candidates. The supervised EMFR feature selection strategy introduces ascertainment bias towards condition-discriminate features; unsupervised integration approaches applied to the full feature space will be important for validation. The single D30 time-point does not resolve temporal dynamics; longitudinal profiling across adipogenic progression would substantially enrich mechanistic interpretation, particularly for the secretomic layer, which exhibited the most dynamic behaviour across factors. While qPCR validation confirms directional concordance between RNA-seq and independent mRNA measurements for six candidates, protein-level and functional validation through gain- and loss-of-function perturbation experiments will be required to establish causal roles for nominated hubs in adipogenesis. Additionally, although mESC-derived adipocytes replicate canonical differentiation features [6,11–13], their full equivalence to primary adipocytes and their direct relevance to human adipose biology requires further benchmarking [33].

Future work should extend this multi-omics framework to human pluripotent stem cell models of adipogenesis, include time-resolved profiling to capture the dynamic evolution of each molecular layer, and apply causal network modelling approaches [36] and machine learning-based feature interpretation [37–39] to refine regulatory hierarchy inference. Single-cell multi-omics technologies, combined with spatial transcriptomics of developing adipose tissue, would further resolve the cellular heterogeneity within differentiating mESC cultures and contextualize our population-level findings within the diversity of adipocyte progenitor states.

Conclusions

This study establishes that adipocyte differentiation from mouse embryonic stem cells is a systems-level biological transition characterized by a dominant, cross-layer multi-omics differentiation axis integrating transcriptional activation, proteomic remodelling, lipid metabolic reprogramming, and partially independent extracellular secretory remodelling [5,8,21,22]. Network hub analysis identifies lipid metabolic enzymes and ECM regulators — including Lpl, Acer3, Plg (a network hub candidate warranting further protein-level investigation), Itga5, Igfbp6, and Pik3cg — as central nodes of the adipogenic regulatory network, emerging alongside canonical PPARγ-driven transcriptional regulators, providing a more complete, multi-dimensional view of the adipogenic regulatory network [9,25,28,29,34]. ECM remodelling and lipid metabolic reprogramming emerge as two coordinated but mechanistically distinct biological programmes that converge to define the mature adipocyte phenotype, with direct relevance to the pathological ECM fibrosis and sphingolipid dysregulation observed in obese and insulin-resistant adipose tissue [1–3,18,19]. This integrative framework

provides a rich molecular resource for the adipogenesis community and establishes a foundation for future functional and translational investigations aimed at identifying novel therapeutic targets in metabolic disease.

Acknowledgments

Authors would like to acknowledge University of Birmingham for sponsoring this publication.

Author contributions

CRediT: **M. Al-Sayegh:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Resources, Writing – original draft, Writing – review & editing; **M. Khalili:** Methodology; **M. Alzaabi:** Methodology; **M. Sultana:** Methodology; **L. Ali:** Methodology; **M. Ali:** Methodology; **M. El-Hadidi:** Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the New York University Abu Dhabi cooperation internal grant under [76-71202-ADHPG-AD231-63210].

ORCID

M. Al-Sayegh http://orcid.org/0000-0003-3865-1688

Data availability statement

Proteome, lipidome, metabolome, and secretome data related to the study is presented in supplementary material. Annotation of data is presented in supplementary material. Transcriptome data is available with accession number PRJNA1126392 on NCBI. Computational codes and scripts are available at https://github.com/mxa1429/msc-bioinformatics-mohamed-alsayegh. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD078584 and 10.6019/PXD078584.

References

[1] Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. J Clin Endocrinol Metab. 2004;89(6):2548–2556. doi: 10.1210/jc.2004-0395
[2] Rosen ED, Spiegelman BM. What we talk about when we talk about fat. Cell. 2014;156(1–2):20–44. doi: 10.1016/j.cell.2013.12.012
[3] Hirosumi J, Tuncman G, Chang L, et al. A central role for JNK in obesity and insulin resistance. Nature. 2002;420(6913):333–336. doi: 10.1038/nature01137
[4] Okunogbe A, Nugent R, Spencer G, et al. Economic impacts of overweight and obesity: current and future estimates for 161 countries. BMJ Glob Health. 2022;7(9). doi: 10.1136/bmjgh-2022-009773
[5] Rosen ED, Spiegelman BM. Molecular regulation of adipogenesis. Annu Rev Cell Dev Biol. 2000;16(1):145–171. doi: 10.1146/annurev.cellbio.16.1.145
[6] Dani C, Smith AG, Dessolin S, et al. Differentiation of embryonic stem cells into adipocytes in vitro. J Cell Sci. 1997;110(Pt 11):1279–1285.
[7] Al-Sayegh M, Ali H, Jamal MH, et al. Mouse embryonic fibroblast adipogenic potential: a comprehensive transcriptome analysis. Adipocyte. 2021;10(1):1–20. doi: 10.1080/21623945.2020.1859789
[8] Cristancho AG, Lazar MA. Forming functional fat: a growing understanding of adipocyte differentiation. Nat Rev Mol Cell Biol. 2011;12(11):722–734. doi: 10.1038/nrm3198
[9] Tontonoz P, Spiegelman BM. Fat and beyond: the diverse biology of PPARgamma. Annu Rev Biochem. 2008;77(1):289–312. doi: 10.1146/annurev.biochem.77.061307.091829
[10] Billon N, Kolde R, Reimand J, et al. Comprehensive transcriptome analysis of mouse embryonic stem cell adipogenesis unravels new processes of adipocyte development. Genome Biol. 2010;11(8):R80. doi: 10.1186/gb-2010-11-8-r80

[11] Cuaranta-Monroy I, Simandi Z, Kolostyak Z, et al. Highly efficient differentiation of embryonic stem cells into adipocytes by ascorbic acid. Stem Cell Res. 2014;13(1):88–97. doi: 10.1016/j.scr.2014.04.015
[12] Cuaranta-Monroy I, Simandi Z, Nagy L. Differentiation of adipocytes in monolayer from mouse embryonic stem cells. Methods Mol Biol. 2016;1341:407–415.
[13] Alzaabi M, Khalili M, Sultana M, et al. Transcriptional dynamics and key regulators of adipogenesis in mouse embryonic stem cells: insights from robust rank aggregation analysis. Int J Mol Sci. 2024;25(17):9154. doi: 10.3390/ijms25179154
[14] Berry DC, Jiang Y, Graff JM. Emerging roles of adipose progenitor cells in tissue development, homeostasis, expansion and thermogenesis. Trends Endocrinol Metab. 2016;27(8):574–585. doi: 10.1016/j.tem.2016.05.001
[15] Lehr S, Hartwig S, Sell H. Adipokines: a treasure trove for the discovery of biomarkers for metabolic disorders. Proteomics Clin Appl. 2012;6(1–2):91–101. doi: 10.1002/prca.201100052
[16] Thomou T, Mori MA, Dreyfuss JM, et al. Adipose-derived circulating miRNAs regulate gene expression in other tissues. Nature. 2017;542(7642):450–455. doi: 10.1038/nature21365
[17] Deng ZB, Poliakov A, Hardy RW, et al. Adipose tissue exosome-like vesicles mediate activation of macrophage-induced insulin resistance. Diabetes. 2009;58(11):2498–2505. doi: 10.2337/db09-0216
[18] Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. Nat Rev Mol Cell Biol. 2014;15(12):786–801. doi: 10.1038/nrm3904
[19] Jaaskelainen I, Petaisto T, Mirzarazi Dahagi E, et al. Collagens regulating adipose tissue formation and functions. Biomedicines. 2023;11(5). doi: 10.3390/biomedicines11051412
[20] Misra BB, Langefeld CD, Olivier M, et al. Integrated omics: tools, advances, and future approaches. J Mol Endocrinol. 2018;62(1):R21–R45. doi: 10.1530/JME-18-0055
[21] Argelaguet R, Arnol D, Bredikhin D, et al. Mofa+: a statistical framework for comprehensive integration of multi-modal single-cell data. Genome Biol. 2020;21(1):111. doi: 10.1186/s13059-020-02015-1
[22] Argelaguet R, Velten B, Arnol D, et al. Multi-omics factor analysis-a framework for unsupervised integration of multi-omics data sets. Mol Syst Biol. 2018;14(6):e8124. doi: 10.15252/msb.20178124
[23] Emam MTA, Antunes M, Soudy A, et al. Comparative evaluation of multiomics integration tools for the study of prediabetes: insights into the earliest stages of type 2 diabetes mellitus. Netw Model Anal Health Inf Bioinforma. 2024;13(8). doi: 10.1007/s13721-024-00442-9
[24] Sethi JK, Vidal-Puig AJ. Thematic review series: adipocyte biology. Adipose tissue function and plasticity orchestrate nutritional adaptation. J Lipid Res. 2007;48(6):1253–1262. doi: 10.1194/jlr.R700005-JLR200
[25] Goldberg IJ, Eckel RH, Abumrad NA. Regulation of fatty acid uptake into tissues: lipoprotein lipase- and CD36-mediated pathways. J Lipid Res. 2009;50 Suppl(Suppl):S86–90.
[26] Zhang Y, Yang JH. Activation of the PI3K/Akt pathway by oxidative stress mediates high glucose-induced increase of adipogenic differentiation in primary rat osteoblasts. J Cell Biochem. 2013;114(11):2595–2602. doi: 10.1002/jcb.24607
[27] Hopkins BD, Goncalves MD, Cantley LC. Insulin-PI3K signalling: an evolutionarily insulated metabolic driver of cancer. Nat Rev Endocrinol. 2020;16(5):276–283. doi: 10.1038/s41574-020-0329-9
[28] Parveen F, Bender D, Law SH, et al. Role of ceramidases in sphingolipid metabolism and human diseases. Cells. 2019;8(12):1573. doi: 10.3390/cells8121573
[29] Liso A, Venuto S, Coda ARD, et al. Igfbp-6: at the crossroads of immunity, tissue repair and fibrosis. Int J Mol Sci. 2022;23(8):4358. doi: 10.3390/ijms23084358
[30] Tang X, Brindley DN. Lipid phosphate phosphatases and cancer. Biomolecules. 2020;10(9):1263. doi: 10.3390/biom10091263
[31] Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. 2004;5(2):101–113. doi: 10.1038/nrg1272
[32] Al-Sayegh MA, Mahmood SR, Khair SBA, et al. Beta-actin contributes to open chromatin for activation of the adipogenic pioneer factor CEBPA during transcriptional reprograming. Mol Biol Cell. 2020;31(23):2511–2521.
[33] Bahmad HF, Daouk R, Azar J, et al. Modeling adipogenesis: current and future perspective. Cells. 2020;9(10). doi: 10.3390/cells9102326
[34] Morandi EM, Verstappen R, Zwierzina ME, et al. Itgav and Itga5 diversely regulate proliferation and adipogenic differentiation of human adipose derived stem cells. Sci Rep. 2016;6(1):28889. doi: 10.1038/srep28889
[35] Jung YJ, Park W, Noh JM, et al. SIRT1 induces the adipogenic differentiation of mouse embryonic stem cells by regulating RA-induced RAR expression via NCOR1 acetylation. STEM Cell Res. 2020;44:101771. doi: 10.1016/j.scr.2020.101771
[36] Liu A, Trairatphisan P, Gjerga E, et al. From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. NPJ Syst Biol Appl. 2019;5(1):40. doi: 10.1038/s41540-019-0118-z
[37] Eraslan G, Avsec Z, Gagneur J, et al. Deep learning: new computational modelling techniques for genomics. Nat Rev Genet. 2019;20(7):389–403. doi: 10.1038/s41576-019-0122-6
[38] Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet. 2015;16(6):321–332. doi: 10.1038/nrg3920
[39] Chen H, Lundberg SM, Lee SI. Explaining a series of models by propagating Shapley values. Nat Commun. 2022;13(1):4512. doi: 10.1038/s41467-022-31384-3