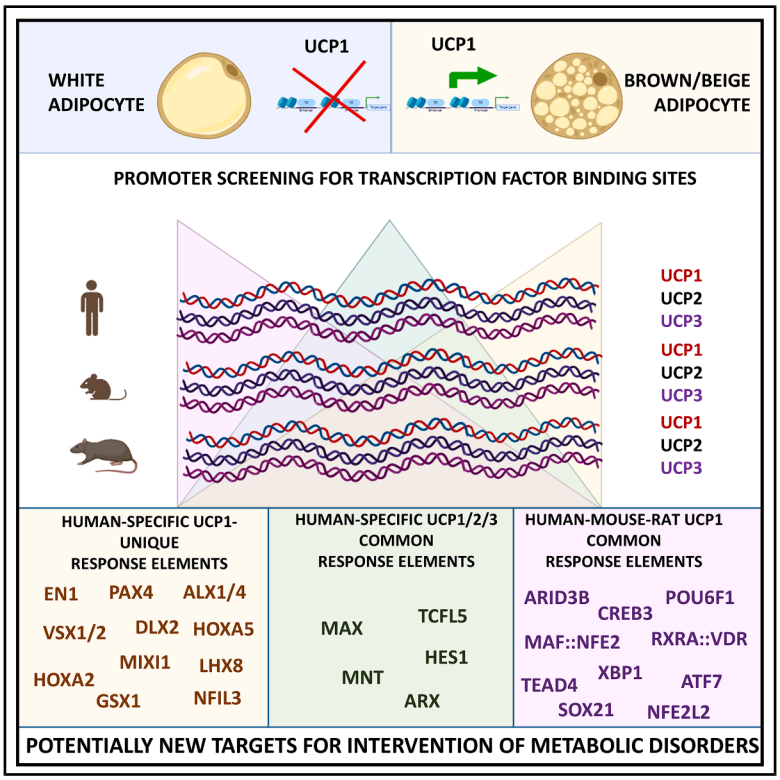
iScience

Article

In silico analysis of *UCP* promoters predicts conserved and human specific regulators of adipocyte thermogenesis

Graphical abstract



Highlights

- Cross-species promoter mapping predicts conserved and human-specific *UCP1* regulators
- Scoring system ranks 579 TFs by binding motif frequency and *UCP1* co-expression
- Transposable elements contribute to human-rodent *UCP1* promoter divergence
- The RE profile of UCP1 in humans is closer to UCP2 than to *UCP1* in mice and rats

Authors

Beáta B. Tóth, Géza Hegedűs, Eszter Virág, Gergely Tamás Pethő, Levente Laczkó, László Fésüs

Correspondence

toth.beata@etk.unideb.hu

In brief

Genomic analysis; Sequence analysis; Sequence homology

Tóth et al., 2025, iScience *28*, 112969
July 18, 2025
https://doi.org/10.1016/j.isci.2025.112969

CellPress

iScience

CellPress
OPEN ACCESS

Article

In silico analysis of *UCP* promoters predicts conserved and human specific regulators of adipocyte thermogenesis

Beáta B. Tóth,[1,6,*] Géza Hegedűs,[1,2] Eszter Virág,[1,3] Gergely Tamás Pethő,[1] Levente Laczkó,[1,4] and László Fésüs[5]

[1]"One Health" Institute, Faculty of Health Science, University of Debrecen, Egyetem Tér 1, 4032 Debrecen, Hungary
[2]Department of Information Technology and Its Applications, Faculty of Information Technology, University of Pannonia, Gasparich Márk Str 18/A, 8900 Zalaegerszeg, Hungary
[3]Research Institute for Medicinal Plants and Herbs Ltd., Budakalász, Hungary
[4]HUN-REN-DE Conservation Biology Research Group, University of Debrecen, Egyetem Tér 1, 4032 Debrecen, Hungary
[5]Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Egyetem Tér 1, 4032 Debrecen, Hungary
[6]Lead contact
*Correspondence: toth.beata@etk.unideb.hu
https://doi.org/10.1016/j.isci.2025.112969

SUMMARY

Adipocyte thermogenesis is a promising therapeutic target to treat metabolic diseases. We predicted transcription factors (TFs) and response elements (REs) regulating the human *UCP1* gene, a key regulator of thermogenesis, by comparing its promoter with those of *UCP2* and *UCP3* in humans, mice, and rats. Using the Eukaryotic Promoter Database, we identified conserved and human-specific TF-REs and developed a scoring system based on RE frequency and TF-*UCP1* co-expression in human neck adipocytes. Additional databases refined the list of potential regulators. Transposable elements and base substitutions partially explained sequence divergence between species, and RE analysis revealed greater similarity between the human *UCP1* and *UCP2* promoters than between the *UCP1* orthologs. While some REs (e.g., EN1, POU6F1, and TEAD4) were *UCP1*-specific, others (e.g., PPARG and RXR) were shared by all *UCP*s. These results reveal conserved and species-specific regulatory patterns that contribute to the understanding of the transcriptional control of thermogenesis in human adipocytes.

INTRODUCTION

Adipocytes play a crucial role in energy balance and metabolism.[1–6] While white adipocytes store energy, brown and beige adipocytes release it as heat through thermogenesis, a process with potential therapeutic applications in the treatment of metabolic disorders.[7–10] Thermogenesis is activated by, among other things, cold exposure, hormonal stimuli, or a high-fat diet, and it is mainly driven by the action of Uncoupling Protein 1 (UCP1) in mitochondria, which uncouples oxidative phosphorylation to generate heat instead of ATP.[11] Skeletal muscle also contributes to non-shivering thermogenesis via SERCA-mediated calcium cycling and sarcolipin (SLN)-regulated ATP consumption, indicating multiple adaptive pathways of heat generation, particularly in species or contexts where brown adipose tissue (BAT)-dependent UCP1 thermogenesis is limited or absent.[12–15] Additional UCP-independent futile cycles, including the mitochondrial creatine/phosphocreatine and triacylglycerol/fatty acid cycles, also contribute to maintaining energy balance. Nevertheless, UCP1 remains an essential target for metabolic modulation.[16–18]

The human genome encodes six UCP isoforms that belong to the solute carrier family 25 (SLC25). UCP1, UCP2, and UCP3 likely originate from a single precursor gene and diverged early in vertebrate evolution; therefore, they remain the focus of our study. *UCP4* and *UCP5* appear to have evolved separately and are more conserved across species, particularly in neuronal tissues.[19–21] UCP6 (KMPC1) transports inorganic anions and dicarboxylates, phylogenetically a separate member of the classical UCP1–UCP5 cluster, reflecting its divergent evolutionary pathway and different functional roles.[22] While UCP1 is the only isoform directly involved in adipocyte thermogenesis, UCP2-UCP5 are known to play a role in the regulation of mitochondrial energy, metabolic efficiency and oxidative stress, however, their exact role is unclear.[23–25] UCP2, which is expressed in various tissues, regulates reactive oxygen species (ROS) and insulin secretion,[26,27] while UCP3, which is mainly found in skeletal muscle, prevents lipid-induced oxidative damage and can support muscle thermogenesis.[28–30]

To better understand the transcriptional regulation of *UCP1*, we mapped and compared 579 transcription factor (TF) response elements (REs) in the promoters of human, mouse, and rat *UCP1* and its closest paralogues *UCP2* and *UCP3*. Taking advantage of functional differences between UCP family members and evolutionary conservation between species, we narrowed our dataset to highlight TFs potentially involved in

species- and gene-specific regulation of human *UCP1* expression. Several TFs have been proposed as direct regulators, including PPARs, RXRs, RARs, THR, CREBPs, NFE2L2, SREBP1, EBF1-2, CEBPA, and the c-Jun:c-Fos complex, while PGC-1α and PRDM16 act as coactivators.[8,31–43] However, the complexity of epigenetic and transcriptional regulation is not yet fully understood.

Using computational analyses, we investigated the TF-DNA binding motifs in the −5kb to +1kb regions of *UCP1*, *UCP2* and *UCP3* in human, mouse and rat using JASPAR Position Weight Matrixes (PWMs) by Eucaryotic Promoter Database (EPD) (EPD: https://epd.epfl.ch; JASPAR: https://jaspar.genereg.net).[44,45] Additional databases (ChiPBase, TFlink, Contra v3, AdipoNet, ISMARA) refined the list of potential regulators. Our ranking system prioritized the TFs based on their frequency in *UCP* promoters and their correlation with the expression of *UCP1* in human neck fat cells, distinguishing between species- and gene-specific regulators. To simplify the nomenclature, the term ''promoter'' is used here for the entire analyzed regulatory sequence region. The results provide a predictive framework for experimental validation and help identify common and human *UCP1*-specific regulators that may not be detected in conventional model organisms. Due to its probabilistic nature, our approach is valuable for target selection in adipocyte browning studies and potential metabolic interventions.

RESULTS

Comparative analysis of the promoter, gene, and protein sequences of UCP1, UCP2, and UCP3 in different species: Structural and evolutionary aspects

To elucidate the DNA composition and the structural features underlying the specific function of three mitochondrial SLC25 proteins, UCP1 (SLC25A7), UCP2 (SLC25A8), and UCP3 (SLC25A9), a thorough comparative analysis was performed, covering the nucleotide sequences of the genes and their promoters in the human, mouse and rat genomes. We also aligned the amino acid sequences of the encoded proteins and analyzed their complex three-dimensional (3D) spatial structure.

DNA sequence similarity analysis shows distinct evolutionary trends between genes and their regulatory regions

A phylogenetic tree reconstructed by integrating the human, mouse and rat *UCP1, UCP2* and *UCP3* data clearly shows that orthologous sequences were grouped, indicating a closer relationship than to their paralogous counterparts when considering amino acid composition, coding (mRNA) and regulatory DNA sequences, with high support values in all cases (Figure 1A). As expected, among orthologous genes, mouse and rat are more tightly clustered compared to human, indicating greater similarity. However, when we look at the relationship between paralogous genes, we see a difference between clustering by promoter sequence and clustering by gene or amino acid sequence. While the reconstruction of the mRNA and amino acid sequences identified *UCP2* and *UCP3* as a separate cluster from *UCP1*; *UCP1* and *UCP2* were grouped based on the promoter sequence. These results are also confirmed by pairwise distance analysis of the coding and promoter sequences of the *UCP* genes (Figure S1A).

Transposable and repetitive elements drive sequence divergence in the human UCP1 promoter

The thermogenic *UCP1* gene evolved in placental mammals around 120–150 million years ago.[46,47] However, Jastroch et al. demonstrated that three members of the core *UCP* family (UCP1/2/3) were present before the divergence of the ray-finned and lobe-finned vertebrate lineages around 420 million years ago.[48] Nevertheless, some were later lost or inactivated in several clades, such as birds, while their regulatory sequences underwent significant evolutionary changes.[49] A comparative analysis of 5000 bp upstream and 1000 bp downstream of the *UCP1* gene revealed substantial sequence divergence between humans and some rodents (e.g., mouse and rat), with two large human-specific regions (1419 bp in the middle part and 658 bp at the end of investigated promoter region) missing in mice and rats (Figures 1B and S1B). These segments likely arose after placental and marsupial divergence but were secondarily lost in specific lineages (Multiz Alignment of 100 Vertebrate Tool in the UCSC Genome Browser (UCSC: https://genome.ucsc.edu/)). The absence was confirmed by a BLAST search in the NCBI database (NCBI: https://blast.ncbi.nlm.nih.gov/). Short palindromic sequences or residues can be detected at the beginning and end of these evolutionary novel DNA fragments in the human *UCP1* promoter (Figures S1C–S1G; red boxes). Eukaryotic Promoter Database analysis ($p < 1e\text{-}5$) shows that the RE of CREB3 at −3498 bp corresponds to the palindromic ''TGACGTCA'' sequence at the end of the 1419 bp long segment (Figure S1D). RepeatMasker analysis identified a 468 bp long terminal repeat (LTR) retrotransposon (MLT1G, ERVL-MaLR family) within the 1419 bp long segment, flanked by homeodomain TF-RE (DLX2, HOXD8, HOXD9, HOXB5, and POU6F1). A BLAST search shows that this 468 bp sequence is integrated into the human genome at multiple sites (Figure S1H). The 658 bp long region contained a short, interspersed core element (SINE, Alu sz6) framed by a long interspersed core element (LINE, L1M5). Several additional repeat elements were detected along the entire length of the promoter studied (Figures 1B and S2). Typically, TF REs' motifs are identified at their start and endpoints, such as MEOX2 at the boundary of the AT-rich repeat element or RHOX11 for L1M5.

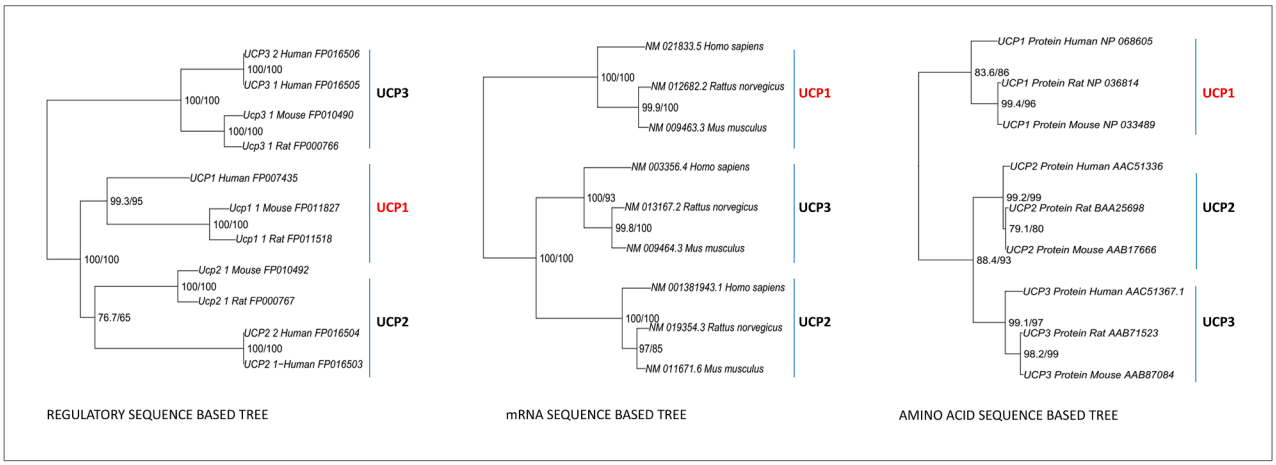
Despite this divergence, conserved regions near +250 bp and −3600 bp exhibited high sequence similarity between species and harbored TF-REs for MZF1, GMEB1 and GLIS2, suggesting their role in transcriptional regulation (Figure 1B, blue boxes, UCSC genome browser, Eucaryotic Promoter Database). These results shed light on the influence of transposable elements on the evolution of regulatory sequences and their potential impact on human *UCP1* gene expression.

Subtle structural differences between orthologous and paralogous UCP1 proteins may have functional implications

The 3D structures of human UCP1, UCP2, and UCP3 proteins resemble those of their mouse and rat counterparts, showing closer relationships between the orthologues compared to the paralogues, as predicted by AlphaFold 2 (Figure S3A) and structural alignment using the RCSB PDB interface (Figure 2A, Methods). Figure S3A highlights their conformational similarities and shows a low-confidence interdomain region (red), suggesting structural flexibility that may influence proton conductance

A

Phylogenetic tree reconstruction of nucleotide and amino acid sequences of UCP orthologs and paralogs



B

Comparative analysis of the human UCP1 promoter region (+1000 to –5000 bp) with orthologous sequences from other species

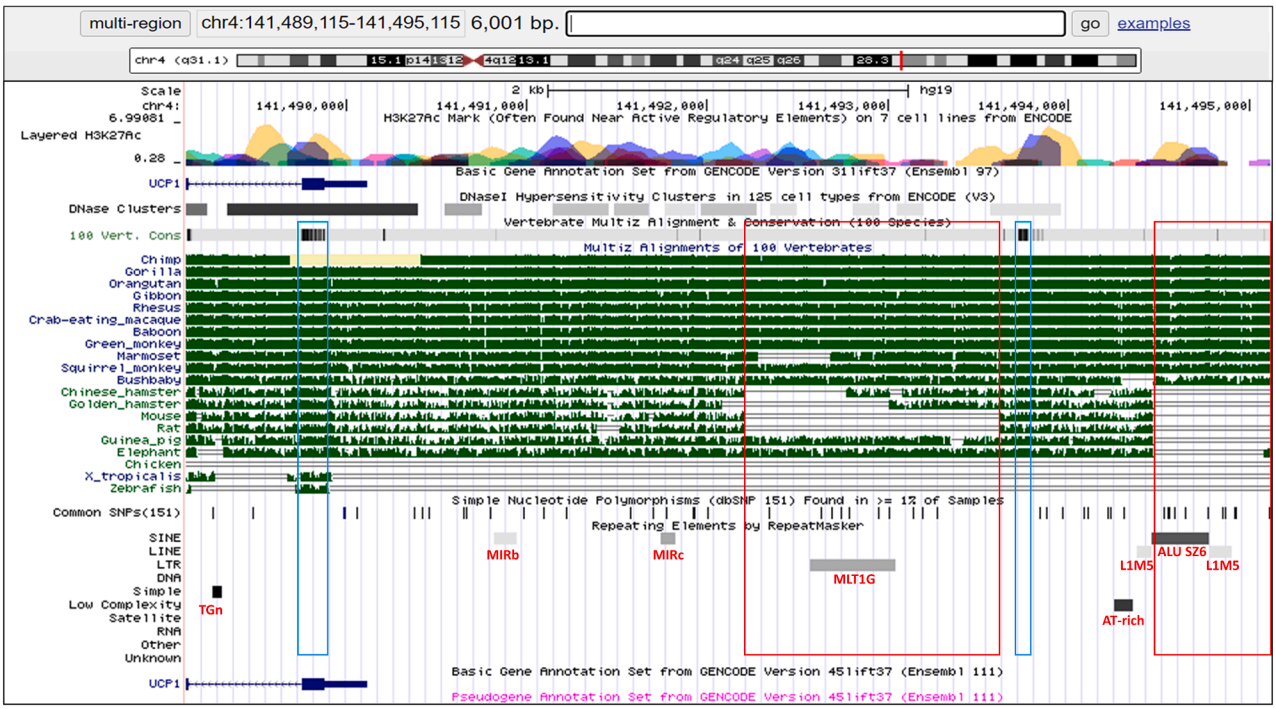


Figure 1. Comparative analysis of the promoter, gene and protein sequences of UCP1, UCP2 and UCP3 in different species

(A) Phylogenetic tree based on the nucleotide sequences of the promoter, mRNA and amino acid sequences of the corresponding proteins of UCP1, UCP2 and UCP3 in human, mouse and rat. The tree is a rooted phylogram with support values (aLRT/rapid bootstrap) on the branches (R package ape v5.7.1).

(B) Comparative analysis of the human *UCP1* promoter region (+1000 bp to −5000 bp) in vertebrates using the UCSC Multiz Alignment of 100 Vertebrates. Red boxes indicate human-specific segments missing in mouse and rat; blue boxes indicate conserved regions between species. Repeat elements are shown with abbreviations (red letters) and positions (gray lines) from the UCSC RepeatMasker.

and ligand binding. Pairwise structural alignments confirm strong homology to human UCP1, with TM scores of 0.97 (mouse/rat UCP1), 0.84 (human UCP2) and 0.77 (human UCP3) (Table S1). Despite the different sequence homology, the UCP proteins share conserved structural features within the SLC25 family, as in the case of the ADP/ATP translocase (TM score 0.8, 21% sequence identity) (Figure 2B). These results support the evolutionary conservation of the UCP structure[50] and raise questions about the functional role of sequence variations in fuzzy interdomain regions.

A

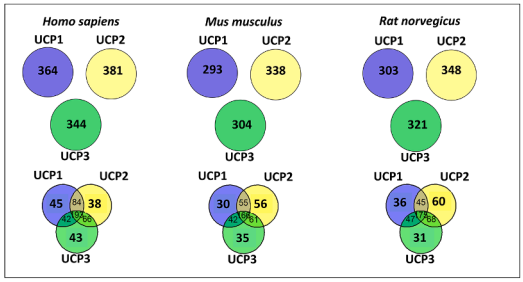
Aligned predicted protein structure of UCP1, UCP2 and UCP3 in human, mouse and rat.

B

H-UCP1 aligned with H-ANT1

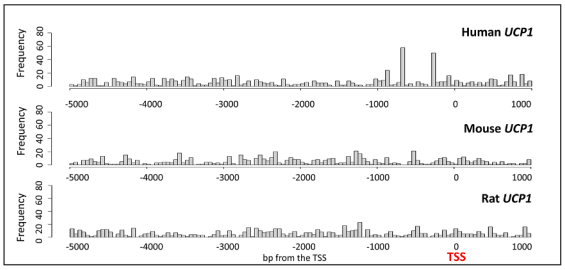
C

Number of TFs with RE in human, mouse and rat *UCP1*, *UCP2* and *UCP3* Promoter



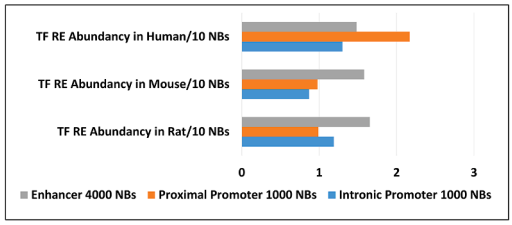
D

TF RE frequency in human, mouse and rat *UCP1* gene promoter



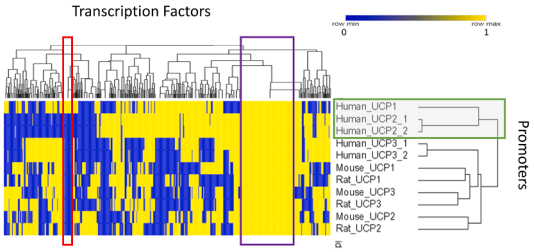
E

TF RE abundancy in *UCP1* gene promoter regions of human, mouse and rat



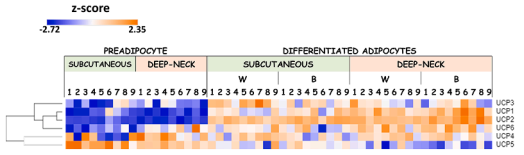
F

Hierarchical cluster analysis of the human, mouse and rat *UCP1, UCP2* and *UCP3* genes based on the TF-REs detected in their promoters.



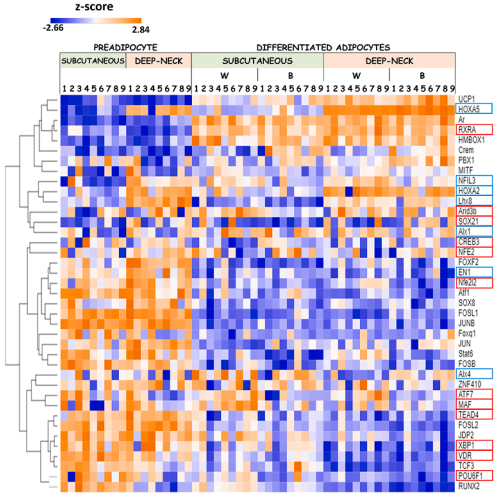
G

Heatmap shows the gene expression profile of the six *UCP* genes in adipocyte samples of cervical origin.



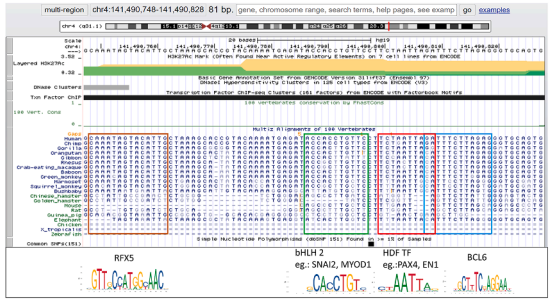
H

Heatmap shows the gene expression profile of 45 TFs with RE in the human *UCP1* promoter in adipocyte samples of cervical origin.



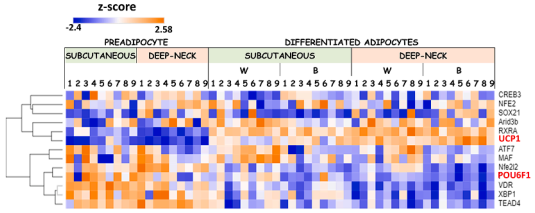
I

Comparison of the HDF1-class TF-RE sequences at the human *UCP1* promoter -690 BP with sequences from other species.



J

Heatmap shows the gene expression profile of 12 TFs with RE in the human, mouse and rat *UCP1* promoter in adipocyte samples of cervical origin.



(legend on next page)

Comparative analysis of transcription factor binding sites/response elements of UCP promoters in humans, mice, and rats

The prevalence of putative TF-REs in the UCP1 promoter shows greater similarity to human UCP2 than to rodent UCP1

We investigated the specific transcriptional regulation of human *UCP1* using the EPD and JASPAR TF PWMs and analyzed the presence of 579 TF motifs.[51] At a significance threshold of $p < 1e\text{-}4$, we identified 293–381 TFs with putative REs in nine genes (Figures 2C and S3B; Table S2). Some TFs had multiple binding sites (e.g., CREB1: 4, SP1: 14 in the analyzed *UCP1* promoter), while certain REs corresponded to multiple TFs (e.g., at −262 bp RE for 19 bHLH TFs, −690 bp RE for 22 homeodomain TFs) (Tables S3A–S3C). A more stringent threshold ($p < 1e\text{-}5$) resulted in 63–123 TFs per species (Figure S3C). The distribution of TF-RE varied: humans had the highest density in the proximal promoter, whereas mice and rats had more in the enhancer region (Figures 2D and 2E; Tables S3B and S3C). Heatmap analysis and hierarchical clustering showed that the TF-RE profile of human *UCP1* is more similar to its paralog *UCP2* than to its orthologues, regardless of whether the analysis is based on binary presence-absence data (Figure 2F) or on a weighted occurrence that considers the number of REs identified in the promoter regions (Figure S4). Similarly, Spearman's rank correlation confirmed that the gene expression of *UCP1* in subcutaneous and deep-necked adipocytes correlated most strongly with *UCP2* (Figure 2G).

Comparative analysis of transcription factor response elements reveals unique regulatory sequences for human UCP1

Our analysis identified 364 TFs (monomers, dimers, trimers and variants) with putative REs in the human *UCP1* promoter. Of these, 319 were also present in the human *UCP2* and/or *UCP3* promoters, while 45 were exclusive to *UCP1* (Figure 2C; Table S2). Analysis of gene expression in subcutaneous and deep neck preadipocytes and mature adipocytes in which browning processes were induced by prolonged rosiglitazone treatment revealed that five TFs (AR, RXRA, HMBOX1, CREM, and HOXA5) correlated positively with *UCP1*, while 17, including JUNB, FOSL1/2, ATF1, VDR, and NFE2L2, showed a negative correlation (Figure 2H, Table S4, STARMethods). Our RNA-seq dataset reflects the combined effect of factors that promote adipocyte browning, including tissue origin, PPARγ activator drug rosiglitazone used in the brown differentiation protocol, and FTO pro-obesity rs1421085 T to C SNP-based adipocyte phenotype (Methods 10). Remarkably, the specificity of RXRA for *UCP1* was only observed when it was dimerized with VDR, and JUNB_var2 was unique when paired with FOSL1 or FOSL2. Of these 45 UCP1-specific TF-REs, 13 (ALX1, ALX4, DLX2, EN1, GSX1, HOXA2, LHX8, MIXL1, NFIL3, PAX4, VSX1, VSX2, and HOXA5) were absent from the mouse and rat *UCP1*/2/3 promoters, with all but NFIL3 belonging to the homeodomain family (HDF) (Figure 2F, blue box). For many of them, analysis of the EPD at strict confidence ($p < 1e\text{-}5$) supports their role in species-specific thermogenic regulation (Tables S3C and S5A), though only EN1, HOXA5 and NFIL3 show high expression in neck-derived fat cells. A palindromic sequence (5′-TCTAAT TAGA-3′) at −690 bp (EPD database; UCSC Genome Browser GRCh 37/hg19: 141.490.801–141.490.810), which acts as RE for 9 of these homeodomain TFs, was present with minor variations in primates, indicating a newly formed regulatory element in adipocyte thermogenesis (Figure 1B; Multiz Vertebrate Sequence Alignment in UCSC). DLX2, HOXA2 and HOXA5, other human *UCP1*-specific homeodomain TFs with RE at different positions (Table S5A), show significant expression in cervical adipocytes, but only *HOXA5* expression is correlated with *UCP1* (positively). In addition, NFIL3, a basic leucine zipper TF, had a potential binding site at −881 bp, albeit with a probability of $p < 1e\text{-}4$, but is highly expressed in mature adipocytes. Conversely, 93 TF-REs were conserved in all UCP promoters of the three species, suggesting common regulatory mechanisms (Figure 2F Purple box, Table S2). These include known TFs such as PPARA, PPARG, RARE, TFAP2, SP, THR, and SREBF1, indicating evolutionary conservation regulating *UCP1* gene expression.[52]

Figure 2. Comparative 3D structure and promoter-TF binding landscape of UCP1, UCP2 and UCP3 in humans, mice and rats

(A) 3D structures of human, mouse and rat UCP1, UCP2 and UCP3 proteins predicted by AlphaFold2, aligned and visualized from three perspectives (intermembrane, matrix and membrane-embedded sides) (RCSB PDB).

(B) Structural alignment of human UCP1 (brown) with human ANT1 (blue) showing conservation within the mitochondrial SLC25 carrier family (AlphaFold2 and RCSB PDB).

(C) Venn diagrams showing the number of TFs with predicted REs in the −5000 bp to +1000 bp promoter region of *UCP1, UCP2* and *UCP3* in human, mouse and rat (EPD, $p < 1e\text{–}4$).

(D) Distribution of TF-REs across the human, mouse and rat *UCP1* promoters in 50 bp bins (−5000 bp to +1000 bp).(Excel).

(E) Relative abundance of TF-REs per 10 bp in three promoter segments: distal enhancer (−5000 to −1001 bp), proximal/core promoter (−1000 to TSS) and intronic promoter (TSS to +1000 bp) from human, mouse and rat. (Excel).

(F) Hierarchical clustering of all *UCP* gene promoters based on shared TF-REs. Red box: 13 TFs unique to the human *UCP1* promoter; purple box: 93 TFs common to all *UCP*s; green box: TF-RE clusters most similar to human *UCP1* (Morpheus interface).

(G) Heatmap showing gene expression profiles of all *UCP* genes in human cervical adipocytes. Clustering is based on Spearman rank correlation (Morpheus interface).

(H) Relative expression profile of 45 TFs with predicted REs in the human *UCP1* promoter in the cervical adipocyte samples. Blue boxes: TFs absent in the mouse and rat *UCP1/2/3* promoters; red boxes: TFs also present in the mouse and rat *UCP1* promoters. Very weakly expressed TFs were excluded (Morpheus interface).

(I) Multiz alignment of the human *UCP1* promoter shows a species-specific HDF1 TF RE at −690 bp. Nearby REs for RFX5 (brown), MYOD1 (green) and BCL6 (blue) are also shown (UCSC Genome Browser and EPD).

(J) Heatmap of 12 TFs with common REs in human, mouse and rat *UCP1* promoters showing gene expression in cervical adipocytes. *UCP1* is shown as reference. Hierarchical clustering shows TFs with *UCP1*-like expression (Morpheus interface). Adipocytes were differentiated using a white (W) or brown (B) protocol.

Identification of transcription factors that potentially bind to the unique palindromic nucleotide sequence of the human UCP1 regulatory region at −690 bp

To predict TFs that bind to the unique palindromic sequence "5′-TCTAATTAGA-3′" at −690 bp in the human *UCP1* regulatory region, we analyzed TFs from the HDF using several databases. The EPD (p < 1e-4) identified several homeodomain TFs, some of which were specific for human *UCP1* (Table S5A), while others lacked this specificity (Tables S2, S3, and S5B). To refine this selection, we used ConTra v3 (stringency score = 0.95, similarity matrix = 0.85; ConTra v3: http://bioit2.irc.ugent.be/contra/v3/#/step/1) to confirm the predicted binding motifs and their evolutionary conservation (Figure S5). Gene expression data from neck-derived adipocytes helped prioritize TFs that showed significant positive or negative correlation with *UCP1* (Figure S6; Tables S4, S5A, and S5B). Additional criteria, including ChIP-Seq data (ChIPBase and TFlink), a more stringent analysis of the promoter database (p < 1e-5) and PWM from JASPAR, refined the list of potential TFs (Tables S5A and S5B). AdipoNET categorized these TFs based on their association with white, brown or intermediate (linker) fat phenotypes to predict their functional role in the adipocyte browning process.[53] An automated PubMed search identified the co-occurrence of these TFs with *UCP1* in abstracts to determine what we know about these TFs in relation to UCP1 (Figure S7 and Table S6). Based on the data collected, the most likely candidates for binding at −690 bp were *EN1*, *PRRX2, HOXA5*, *HOXD8*, *LHX8*, *PAX4*, and *VSX2*. While *EN1, PRRX2* and *HOXD8* showed high-to-medium adipocyte expression with a strong negative correlation to *UCP1*, *HOXA5* showed a positive correlation, indicating the promotion of the browning process. *PAX4* and *VSX2*, although not appearing in our transcriptome studies,[54] were also identified in a more stringent promoter analysis (p < 1e-5), supporting their possible regulatory role. In addition, *PAX4* showed high motif activity in adipose tissue when analyzed using the ISMARA platform (Figure S8, the red boxes highlight the adipose tissue data). These results suggest several homeodomain TFs that may be involved in the regulation of *UCP1* and whose role needs to be further investigated.[55]

Investigation of potential regulatory hubs at human UCP1-specific homeodomain and NFIL3 transcription factor response elements

Homeodomain TFs play a key role in embryonic development and may regulate tissue-specific *UCP1* expression during adipocyte maturation by modulating chromatin accessibility and forming cooperative complexes.[56–59] We examined TF-REs within ±40 bp of human-specific homeodomain (Table S5A) and NFIL3 REs to identify potential regulatory interactions.

Analyzing the −690 bp region with the EPD (p < 1e-4), we identified BCL6 (−702 bp), RFX5 (−651/650 bp) and bHLH2 class TF-REs (−679 bp) at this position (Figure 2I). *BCL6* and *RFX5* showed high expression and a strong positive correlation with *UCP1* in the neck adipocytes (Figure S9; Table S4), while *ID4*, *SNAI2*, and *TCF3/4/12* from the bHLH2 class showed high expression but a negative correlation. Analysis of the interactions of PAX4 with ISMARA revealed possible cooperations with BCL6, TCF4, ID4, SNAI2, and RFX5, supporting the formation of a regulatory complex of these players at this site (Figure S10). Examining the vicinity of *UCP1*-specific REs (Table S5A; Figure S5, Contra v3) HIC1, which has RE close to the RE of HOXA5 (−4785 bp), was highly expressed in pre-adipocytes but was significantly reduced in mature adipocytes and showed a negative correlation with *UCP1* (Figure S11A; Tables S3 and S4). TEADs (RE at −2496 bp, close to DLX2 RE at −2478 bp) expression showed variable correlations with *UCP1*; *TEAD1/3/4* were negatively correlated, while *TEAD2* was positively correlated (Figure S11B), suggesting context-dependent regulation. HOXA2-associated REs around at −3959 bp (NFKB2 at −3976 bp, ZNF410 at −3947 bp and ZNF283 at −3948 bp) could also form a regulatory complex (Figure S11C). *NFKB2* gene expression showed a negative trend with *UCP1*, while *ZNF410* and *ZNF283* showed no clear correlation. These REs were also detected at p < 1e-5, underlining their possible regulatory role. Finally, we identified the NFIL3 close REL RE at −902 bp (Table S5A; Figures S11D and S11E), which has low to moderate expression regardless of adipocyte differentiation status and is conserved in primates but is absent in the *UCP1* promoter in mice and rats (Figure S11F).

Common REs in orthologous UCP1 promoters differ in their genomic location

A comparison of the regulatory elements of the *UCP1* promoter in humans, mice and rats revealed that several TF-REs—ARID3B, ATF7, CREB3, MAF:NFE2, NFE2L2, POU6F1, RXRA:VDR, SOX21, TEAD4, and XBP1—which are present in the human *UCP1* promoter but absent in the human *UCP2* and *UCP3* promoters, were also present in the mouse *UCP1* promoter. Many of these TFs, such as CREB/ATF, RXR, NFE2 and NFE2L2, are known regulators of *UCP1* expression.[33,60–62] Gene expression analyses (Figures 2H and 2J), consistent with AdipoNET pathway predictions, indicate that POU6F1, TEAD4, XBP1, and NFE2L2 predominantly promote the white adipocyte phenotype and have a strong negative correlation with *UCP1* (Table S7A). *RXRA* is the only TF in this group that correlates positively with *UCP1* (Table S4), suggesting a possible promoting role in browning processes. AdipoNET also predicts that RXRA, CREB3 and XBP1 may regulate the balance between white and brown adipocytes. Remarkably, POU6F1, the only homeodomain TF in this group, is associated exclusively with *UCP1* and absent from the *UCP2* and *UCP3* promoters in the organisms studied, suggesting a gene-specific role in *UCP1* regulation and thermogenesis. However, a comparative analysis of *UCP1* regulatory sequences across species (UCSC: https://genome.ucsc.edu/), shows that these REs are conserved in primates but absent in mouse consensus sequences. Instead, they appear at different genomic positions in mice and rats, indicating a lack of positional conservation (Figures S12A and S12B; Table S2).

Human-specific bHLH transcription factors REs in the proximal promoter of UCP1-2-3

In a comparative analysis, five regulatory TF-REs (HES1, MAX, MNT, TCFL5, and ARX) were identified exclusively in the human *UCP1*, *UCP2*, and *UCP3* promoters, which were absent in the mouse and rat *UCP* promoters (Table S2). Most of these TFs are expressed in human adipocytes, with higher expression in

progenitor cells and a significant negative correlation with *UCP1*, except *HES1*, whose expression is more associated with subcutaneous adipocytes (Table S4; Figure S13A). ARX, a member of the homeobox TF family, is not expressed in our experiments with cervical fat cells. The other four TFs share a common RE at −263 bp in the proximal promoter of *UCP1*, which is typically recognized by TFs of the bHLH family. The EPD also identifies other potential bHLH TFs at this sequence, such as ARNT:HIF1A (Tables S3A and S3B; Figures S13B and S13D). Further experiments are needed to clarify the context-dependent regulatory mechanisms that may influence which TFs bind to this RE.

Limited overlap between predicted and experimentally validated TF binding sites in the human UCP1 promoter

We compared the list of putative TF-REs predicted by the EPD (−5000 to +1000 bp from the TSS, p < 1e-4) with experimentally validated TFs from the ChiPBase and TFlink databases for the human *UCP1* promoter. EPD predicted 364 TF-REs, while ChiPBase and TFlink identified 170 and 131 binding TFs, respectively (Figure S14A). Only a small fraction of the experimental data agreed with EPD: 23 TFs (14%) with ChiPBase and 41 TFs (31%) with TFlink. Remarkably, only 11 TFs co-occurred in all three databases (GATA2, HIF1A, JUN, KLF4, MXI1, MYC, MYCN, NFKB1, RBPJ, TEAD4, and USF1), indicating their probable role in regulating *UCP1* expression. Analysis of gene expression in adipocytes from the neck showed that the expression of most of these TFs correlated with *UCP1*, and predominantly negatively, except *USF1*, which showed a positive correlation (Figure S14B; Table S4).[54] USF1 REs were present in all *UCP* promoters examined, suggesting a general role in *UCP* regulation rather than specific involvement in *UCP1* and browning (Table S2).[63] The AdipoNET predictions also suggest that most of these TFs are associated with the white adipocyte phenotype, although JUN and NFKB1 can be involved in both white and brown adipocyte formation (Table S7B). REs with a more stringent threshold (p < 1e-5) were identified for HIF1A, JUN, RBPJ and TEAD4, supporting their possible regulatory role in *UCP1* expression. HIF1A, MYC, MYCN and MXI1 share a common binding site at −263 bp of the proximal promoter, a region suggested above to be specific for the regulation of human *UCP* genes (Figures S13A–S13D), as some TFs with binding motifs here have no REs in mouse and rat *UCP* genes. The proximity of the RE of SP2 and EGR1 at −263 bp (Table S3B) also suggests the possibility of the formation of a regulatory complex at this position, which is confirmed by the ConTra v3 data (Figures S13B and S13D).

Transcription factor ranking indicates *UCP1*-specific and evolutionarily conserved *UCP* gene expression regulators

We ranked the analyzed 579 TFs to assess their role in regulating the human *UCP1* gene based on a scoring system comprising two key parameters: (1) the frequency of TF-RE motifs in the analyzed *UCP* promoters and (2) the correlation coefficient of TF gene expression with *UCP1* in human adipocytes (Table S4). Three independent scores were assigned to each TF based on the *UCP* gene promoters used to calculate the frequency of the binding motif: (1) the frequency of TF-RE motifs in the human *UCP1* promoter (Table S8A), (2) the frequency of TF-RE motifs in the human, mouse, and rat *UCP1* promoters (to estimate their conservation (Table S8B), and (3) the prevalence of TF-RE in the promoters of *UCP1*, *UCP2* and *UCP3* in all species studied (Table S8C). This approach allowed a comparative ranking of the TFs, with the TFs with the highest score at the top, assessing both evolutionary conservation and functional significance. Figure 3 shows the heat maps based on the 60 highest TF scores from the three alternative score calculations (Figures 3A–3C). In the *UCP* promoters analyzed, commonly identifiable REs such as ZNF263, EWSR1-FLI and SP1/3 are at the top of the ranking, while the highest scoring human *UCP1*-specific TFs include HIF1A:ARNT, NRF1, TBX19, THR, ESSRA, NR1H3:RXRA.

TFs previously associated with thermogenesis generally scored high, highlighting the validity of our scoring system in assessing the role of TFs in regulating *UCP1* gene expression. THRB, THRA, JUN and CREB1 showed some human *UCP1* specificity, while RXRA was dominant in all *UCP1* promoters analyzed. Others, including PPARG, CREB1, SREBF1 and RARA, appeared to be general regulators of *UCP* genes. Interestingly, ZIC1, CEBPA and CEBPB were not found to be significant regulators of human *UCP1* (Figure 4A). Among the TFs that showed the highest specificity for human *UCP1* in our *UCP* promoter analyses (Table S5A), EN1, HOXA5, and LHX8 stood out (Figure 4B). In addition, POU6F1, PRRX2, HOXD8 and LBX2, all belonging to the HDF, were also identified as potential *UCP1* gene-expression regulators at −690 bp; however, based on the EPD, their binding sequences did not show exclusive specificity for human *UCP1* (Figure 4C). The heatmap shows that, except for POU6F1, the REs of these TFs were dominant in the human *UCP1* promoter, suggesting their possible gene-specificity in regulating gene expression. Conversely, TFs such as POU6F1, TEAD4, NFE2L2 and XBP1, which ranked high based on RE frequency in human *UCP1*, showed a similar ranking when conservation in different species was considered (second columns), suggesting their ancient roles in *UCP1* regulation (Figure 4D). While TFs predicted to be *UCP1* regulators based on EPD and chip-seq data generally perform well, RBPJ, HIF1A, and TEAD4 stand out (Figure 4E). Our results establish a systematic ranking of 579 TFs in influencing *UCP1* expression and provide insights into their potential regulatory role in human adipocyte thermogenesis.

DISCUSSION

Evolutionary dynamics of UCP1 genes and their promoters differ in humans and rodents

Computational insights into the transcriptional regulation of human UCP1

In our analysis, *in silico* approaches were used to investigate the regulatory landscape of *UCP* genes. While the coding sequences of *UCP1* have been extensively studied in different species, much less is known about the evolution of its regulatory elements.[48,49,64–71] A previous work by Gaudry and Campbell (2017) identified conserved motifs such as the TATA box, RARE-3 and PPAR in mammals and confirmed their fundamental role in the regulation of *UCP* genes.[49,72] Our results are

A

Top 60 TFs based on TF RE frequency in h UCP1 promoter	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
ZNF263	1	1	1
EWSR1-FLI1	2	2	2
IRF1	3	10	11
IRF3	4	13	44
SP1	5	9	3
ZNF740	6	12	18
SP3	7	3	4
RBPJ	8	4	19
PRDM1	9	8	45
RREB1	10	5	8
ARNT::HIF1A	11	79	111
NRF1	12	43	75
THRB	13	93	116
TBX19	14	102	132
NR4A2	15	44	50
THRA	16	104	123
EGR3	17	34	51
CREB1	18	27	57
Nr1h3::RXRA	19	57	97
MZF1	20	39	25
Esrra	21	29	39
KLF16	22	6	6
PPARG	23	18	34
Sox6	24	19	47
EGR1	25	42	69
STAT3	26	64	52
GLIS2	27	30	37
KLF5	28	11	5
Hic1	29	66	54
ZNF384	30	48	60
PBX3	31	38	13
NR2C2	32	7	14
Arid5a	33	97	136
RFX5	34	98	92
TCF4	35	22	10
POU6F1	36	58	190
ZNF354C	37	52	53
Nfe2l2	38	46	114
TCF3	39	33	16
HIF1A	40	108	149
ID4	41	40	28
TEAD4	42	70	169
KLF9	43	15	24
JUN	44	83	139
NR2F2	45	60	78
RXRG	46	14	43
Nr2f6	47	26	35
MAFG::NFE2L1	48	47	62
SREBF1	49	59	23
E2F8	50	78	129
TEAD1	51	80	148
Prrx2	52	156	277
Stat5a	53	21	42
ZBTB7B	54	84	117
MXI1	55	119	65
Bach1::Mafk	56	65	79
SP4	57	31	17
Arid3a	58	62	76
PAX5	59	45	31
Stat5b	60	23	46

B

Top 60 TFs based on TF RE frequency in h/m/r UCP1 promoter	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
ZNF263	1	1	1
EWSR1-FLI1	2	2	2
SP3	7	3	4
RBPJ	8	4	19
RREB1	10	5	8
KLF16	22	6	6
NR2C2	32	7	14
PRDM1	9	8	45
SP1	5	9	3
IRF1	3	10	11
KLF5	28	11	5
ZNF740	6	12	18
IRF3	4	13	44
RXRG	46	14	43
KLF9	43	15	24
EBF1	140	16	36
Klf12	94	17	27
PPARG	23	18	34
Sox6	24	19	47
RXRA	68	20	63
Stat5a	53	21	42
TCF4	35	22	10
Stat5b	60	23	46
NFIX	62	24	33
ZEB1	87	25	12
Nr2f6	47	26	35
CREB1	18	27	57
CTCFL	82	28	9
Esrra	21	29	39
GLIS2	27	30	37
SP4	57	31	17
PLAG1	108	32	21
TCF3	39	33	16
EGR3	17	34	51
ESR2	80	35	22
REST	98	36	29
RUNX1	126	37	55
PBX3	31	38	13
MZF1	20	39	25
ID4	41	40	28
TGIF1	114	41	30
EGR1	25	42	69
NRF1	12	43	75
NR4A2	15	44	50
PAX5	59	45	31
Nfe2l2	38	46	114
MAFG::NFE2L1	48	47	62
ZNF384	30	48	60
PPARA::RXRA	71	49	58
VDR	132	50	59
NR1H2::RXRA	77	51	66
ZNF354C	37	52	53
ZBTB7A	121	53	32
NFIA	395	54	67
GLI2	373	55	61
CREB3L1	67	56	126
Nr1h3::RXRA	19	57	97
POU6F1	36	58	190
SREBF1	49	59	23
NR2F2	45	60	78

C

Top 60 TFs based on TF RE frequency in h/m/r UCP1/2/3 promoter	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
ZNF263	1	1	1
EWSR1-FLI1	2	2	2
SP1	5	9	3
SP3	7	3	4
KLF5	28	11	5
KLF16	22	6	6
Zfx	169	63	7
RREB1	10	5	8
CTCFL	82	28	9
TCF4	35	22	10
IRF1	3	10	11
ZEB1	87	25	12
PBX3	31	38	13
NR2C2	32	7	14
Tcf12	147	110	15
TCF3	39	33	16
SP4	57	31	17
ZNF740	6	12	18
RBPJ	8	4	19
Rarg	353	221	20
PLAG1	108	32	21
ESR2	80	35	22
SREBF1	49	59	23
KLF9	43	15	24
MZF1	20	39	25
SNAI2	88	71	26
Klf12	94	17	27
ID4	41	40	28
REST	98	36	29
TGIF1	114	41	30
PAX5	59	45	31
ZBTB7A	121	53	32
NFIX	62	24	33
PPARG	23	18	34
Nr2f6	47	26	35
EBF1	140	16	36
GLIS2	27	30	37
TFAP2A	117	81	38
Esrra	21	29	39
NHLH1	420	210	40
Klf1	101	61	41
Stat5a	53	21	42
RXRG	46	14	43
IRF3	4	13	44
PRDM1	9	8	45
Stat5b	60	23	46
Sox6	24	19	47
GLIS3	65	72	48
CTCF	191	82	49
NR4A2	15	44	50
EGR3	17	34	51
STAT3	26	64	52
ZNF354C	37	52	53
Hic1	29	66	54
RUNX1	126	37	55
TFDP1	81	77	56
CREB1	18	27	57
PPARA::RXRA	71	49	58
VDR	132	50	59
ZNF384	30	48	60

Figure 3. Ranking of transcription factors based on RE frequency and co-expression with *UCP1*

The tables show the top 60 TFs of 579 candidates from the EPD database, ranked according to a composite score combining (i) the correlation of TF and *UCP1* gene expression in human adipocytes and (ii) the frequency of their predicted binding sites (REs) in *UCP* promoters.

(A) Top 60 TFs based on RE frequency in the human *UCP1* promoter.

(B) Top 60 TFs based on RE frequency in the human, mouse and rat *UCP1* promoters.

(C) The top 60 TFs based on RE frequency in the *UCP1, UCP2,* and *UCP3* promoters of all three species.

consistent with these findings and confirm that these ancient elements are widely distributed in *UCP* promoters of humans, mice, and rats. However, their ubiquity throughout the *UCP* gene family makes them less suitable as selective targets for *UCP1*-specific modulation in metabolic disorders. Importantly, our analysis has also uncovered human *UCP1*-specific elements not found in rodents, which may help to develop more targeted therapeutic interventions in humans.

Species-specific differences in UCP1 regulatory sequences

Using platforms such as EPD, ConTra v3, Multiz Alignment of 100 Species and the UCSC genome browser, we identified

A

Browning genes	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
THRB	13	93	116
THRA	16	104	123
CREB1	18	27	57
PPARG	23	18	34
NFE2L2	38	46	114
JUN	44	83	139
SREBF1	49	59	23
RXRA	68	20	63
PPARA::RXRA	71	49	58
RARA	74	101	86
AR	122	159	210
FOXO1	123	91	147
EBF1	140	16	36
CREB3	290	252	322
CEBPA	352	220	71
CEBPD	387	474	554
CEBPB	392	253	98
ZIC1	444	76	83
CEBPG	472	516	570

B

Human UCP1 specific TFs	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
EN1	89	194	345
HOXA5	187	300	458
LHX8	204	318	465
GSX1	251	387	525
MIXL1	256	398	526
PAX4	257	402	527
VSX1	261	412	528
VSX2	262	413	529
ALX1	299	426	539
DLX2	303	432	551
HOXA2	413	487	558
ALX4	458	512	567
NFIL3	577	578	581

C

Homeodomain Family TFs	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
POU6F1	36	58	190
PRRX2	52	156	277
HOXD8	75	174	168
EN1	89	194	345
LBX2	105	213	326
HOXA5	187	300	458
PAX4	257	402	527
VSX2	262	413	529
PHOX2A	319	404	477
PROP1	333	354	436
PRRX1	378	471	411
DLX4	401	263	272
NOTO	428	495	562
VAX2	429	496	459

D

Human/Mouse/Rat UCP1 specific TFs	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
POU6F1	36	58	190
NFE2L2	38	46	114
TEAD4	42	70	169
XBP1	164	122	195
ATF7	213	232	318
ARID3B	241	310	406
Maf::NFE2	276	247	378
CREB3	290	252	322
RXRA::VDR	335	356	404
SOX21	477	488	540

E

EPD and ChipSec Identified TFs	Rank based on h UCP1	Rank based on h/m/r UCP1	Rank based on h/m/r UCP1/2/3
RBPJ	8	4	19
HIF1A	40	108	149
TEAD4	42	70	169
JUN	44	83	139
MXI1	55	119	65
NFKB1	76	75	87
MYC	90	120	77
USF1	143	106	73
KLF4	162	74	68
GATA2	171	149	204
MYCN	209	228	249

Figure 4. Ranking of transcription factors based on predicted regulatory potential in *UCP* gene expression

The heatmaps summarize the ranking of TFs identified in our study as potential regulators of *UCP* gene expression, focusing on human-specific or *UCP1*-associated regulatory roles. The ranking is based on a composite score calculated from the following factors: (i) the correlation coefficient between gene

(legend continued on next page)

and compared the binding motifs of TFs in nine *UCP* promoters in different species. We observed both conserved and species-specific patterns.[73,74] Notably, *UCP1* in humans showed greater regulatory similarity to *UCP2* in humans than to *UCP1* in rodents, whereas *UCP1* in mice and rats showed more similarities. These results suggest that *UCP* regulation is under different evolutionary pressures in primates and rodents.

Large sequence regions within the human distal *UCP1* enhancer are not present in rodents (Figure 1B), likely due to lineage-specific deletions and TE insertion events. Interestingly, some TF binding sites in humans overlap with conserved promoter regions in rodents, but base changes have created new, human-specific binding motifs. This emphasizes the importance of large sequential and base-level changes in shaping species-specific regulation. In addition, there is emerging evidence for differences in 3′UTR *UCP1* mRNA processing between rodents and humans that affect protein concentration and thermogenic potential,[75] suggesting that extrapolations from rodent models should be made cautiously.

The role of repetitive elements in the human UCP1 promoter

We have identified numerous repetitive elements in the human *UCP1* promoter, including endogenous retrovirus sequences such as MLT1G or Alu repeats, which are not found in rodents (Figures 1B and S2).[76] Many of these elements are flanked by TF-binding motifs, and some contain palindromic sequences that frequently serve as protein binding sites.[77,78] This suggests that the insertion-coordination sequences of some repetitive elements may have been repurposed as functional regulatory modules during evolution.[79–81] For example, repeated insertions of MLT1G or L1M5 elements into the human genome may influence nearby gene expression and coordinate regulatory programs relevant to the formation of thermogenic fat cells,[82,83] as is the case with MER20 elements in the regulation of progesterone-responsive genes. LTR elements such as MLT1G can also act as alternative promoters or influence chromatin structure.[84]

Interestingly, we found a human-specific palindromic 5′-TCT AATTAGA-3′ homeodomain RE motif at −690 bp in the *UCP1* promoter that resembles a viral transposable element identified in Mamestra configurata nucleopolyhedrovirus variants, although it is not classified as such by RepeatMasker.[85] This emphasizes the potential of ancient cryptic insertions to influence human gene regulation.[86–89] There is increasing evidence that transposable elements and tandem repeats can influence gene expression and even modulate transcription via interactions with microRNAs (e.g., SINE).[90]

Comparative analyzes of the regulatory elements in the promoters of *UCP1*, *UCP2*, and *UCP3* in different species

Homeodomain family class I transcription factors and human-specific sequences

Our comparative analysis of the *UCP1*, *UCP2* and *UCP3* promoters in humans, mice and rats highlights the potential role of homeodomain TFs in the regulation of *UCP1*, particularly in humans. These TFs, which are known to control embryonic development, may also be utilized in adipose tissue to influence thermogenic capacity.[91] While their role in adipocyte development and plasticity is increasingly recognized,[56,57,92–94] the direct regulation of *UCP1* by these TFs is still poorly understood. Homeodomain TFs can influence gene expression via chromatin remodeling and often function as dimers[95] to increase binding specificity.[58,59,96,97]

Among the homeodomain TFs targeting human-specific REs, EN1 stands out. It is highly expressed in human cervical preadipocytes and negatively correlates with *UCP1* expression.[54] In mice, EN1 promotes brown adipogenesis but is repressed upon cold exposure, suggesting a role in early cell lineage commitment rather than direct control of *UCP1*.[98–100] Our previous data showed upregulation of EN1 in adipocytes from individuals with the obesity-associated FTO rs1421085 risk allele (unpublished), supporting its possible species-specific regulatory function.[54]

Another candidate, HOXA5, is upregulated in thermogenic deep-neck adipocytes in humans and is known to promote fat browning in rodents, likely via anti-inflammatory pathways (Figure 2H, Table S4).[100–102] However, its expression in white adipocytes derived from induced pluripotent stem cells suggests a context-dependent role.[103] PAX4, which is typically associated with pancreatic beta cell development[104,105] and diabetes,[106] is also expressed in adipocyte progenitor cells[105] and may influence the balance between white and brown adipocyte formation. PRRX2 and VSX1/2, although less studied in thermogenic fat, show differential expression between different adipocyte types and may also contribute to thermogenic programming.[107]

Evolutionary conservation and divergence in UCP1 promoter elements

Many regulatory motifs found in the human *UCP1* promoter are conserved in primates, but not in the investigated rodents. Interestingly, some TF binding sites —such as POU6F1— occur within promoter segments unique to humans, although they are common *UCP1* regulatory elements in the species studied. This suggests evolutionary reorganization or independent acquisition of the regulatory elements.

expression of each TF and *UCP1* in human adipocytes, and (ii) the frequency of predicted TF response elements (REs) in the analyzed promoter regions. Three rankings are shown: the first is based on RE frequency in the human *UCP1* promoter, the second contains the RE frequency in the human, mouse and rat *UCP1* promoters, and the third extends the analysis to the *UCP1, UCP2* and *UCP3* promoters in these three species. For reference, TFs with experimental support (ChIP-seq data) or previously reported roles in adipocyte browning are also listed.

(A) TFs previously reported to regulate adipocyte browning.

(B) Predicted human *UCP1*-specific TFs.

(C) TFs of the Homeodomain family whose gene expression correlates (positively or negatively) with *UCP1* and whose REs are located in human-specific *UCP1* promoter segments (EPD data).

(D) Predicted TFs with *UCP1*-specific regulatory potential.

(E) TFs with predicted REs in the human *UCP1* promoter (EPD) that also show binding evidence in the same promoter region in the ChIPBase and TFlink databases. h, human; m, mouse; r, rat.

Although POU6F1 function is not traditionally associated with adipose tissue, it is involved in cell fate determination and tissue remodeling.[108,109] Its strong negative correlation with *UCP1* expression in our data suggests a possible role as a suppressor of thermogenesis in mature human adipocytes (Table S4).

Transcription factor ranking predicts conserved and evolutionarily novel UCP1 regulators

Our TF ranking analysis confirmed many known regulators of *UCP1*, including PPARG, RXRA, CREB1, SREBF1, THR, and RARA—factors that are critical for brown and beige adipocyte differentiation. Their high scores across species reinforce the conserved nature of key thermogenic regulatory pathways.

Remarkably, our analysis also revealed TFs that were not previously directly involved in the human-specific regulation of *UCP1*. These include HIF1A:ARNT, NRF1, ESRRA (ERRα), and NR1H3:RXRA (LXRα:RXRα). HIF1A generally responds to hypoxia and is associated with glucose metabolism and adipocyte inflammation.[110,111] NRF1 and ESSRA play a central role in mitochondrial biogenesis, which is well compatible with the mitochondrial dependence of thermogenesis.[112] LXRα:RXRα regulates lipid metabolism[113] and may act as a link between lipid signaling and *UCP1* induction.[114]

Highly ranked *UCP1*-specific TFs such as TEAD4, NFE2L2 (NRF2), and XBP1 are associated with mechanotransduction,[115,116] oxidative stress[117] and ER stress,[118] respectively, suggesting that *UCP1* regulation is linked to broader cellular stress and environmental adaptation pathways, consistent with the complex regulation of thermogenesis in response to physiological and metabolic cues.[119]

Interestingly, some TFs previously thought to be important regulators of browning—such as CEBPA, CEBPB and ZIC1— were not highly associated with human *UCP1*, possibly reflecting species-specific roles or functions restricted to upstream regulatory cascades rather than direct gene activation.[120,121]

TFs such as RBPJ (a key player in Notch signaling), HIF1A, and TEAD4 were strongly supported by both the presence of motifs and ChIP-seq data, making them strong candidates for future experimental validation. Notch signaling is known to affect the balance between white and beige lipids, highlighting the importance of RBPJ.[122]

Regulatory hotspots in the human *UCP1* promoter

Our predictions using the EPD revealed areas with a high density of TF-REs within the human *UCP1* promoter (Figure 5). These RE clusters likely function as enhanceosomes—cooperative assemblies of TFs and cofactors that integrate multiple signals to fine-tune gene expression. Such complex regulatory structures are typical of multicellular organisms.[123] Several of these clusters consist of *UCP1*-specific homeodomain TF binding sites on or near repetitive elements, indicating a possible interplay between genomic structural variation and regulatory complexity. Another notable region is located within the proximal promoter, near a bHLH family TF-RE at approximately −262/−263 bp (Figures S13B–S13D). The neighboring binding sites for EGR1 and SP2 suggest a possible regulatory interaction with HIF1A. Similar TF cooperation has been observed in non-small cell lung cancer, where HIF1A interacts with EGR1 and SP1 to regulate erythropoietin receptor (EPOR) expression under hypoxic conditions.[124] Given the established role of HIF1A in adipose tissue inflammation and energetic remodeling, it may modulate transcription of *UCP1* by binding to its promoter.[53,110,111] EGR1 is also known for its role in metabolic regulation and influences insulin sensitivity via the PI3K/Akt and MAPK signaling pathways, supporting its contribution to *UCP1*-specific regulatory complexes.[60,125]

Studies in mice show that slight variations in the composition of *UCP1*-enhanceosomes can significantly affect gene expression.[126] Therefore, a precise knowledge of the regulatory architecture is crucial to safely and efficiently modulate *UCP1* expression for therapeutic purposes.

Among the co-regulators, the p300/CBP-associated factor, a histone acetyltransferase, was found to be a significant human-specific cofactor in the *UCP1* promoter (Figure 5). This suggests that epigenetic modulation may play an important role in the species-specific regulation of *UCP1*. In contrast, PGC1A, a known co-activator of PPARG in many species, predominantly associates with TFs common to *UCP* genes, suggesting a broader, less selective regulatory role.

Overall, these results emphasize that understanding the composition, dynamics and potential plasticity of *UCP1* regulatory modules is critical for the development of therapeutic strategies to activate *UCP1* expression. Future research should investigate the conditional interchangeability of these factors under different physiological and pathological conditions.

Limitations of the study

Our comparative analysis revealed limited agreement between the computationally predicted TF-REs and the sites identified by ChIP-Seq. This highlights a fundamental challenge in accurately defining functional regulatory elements, particularly in complex mammalian promoters.

A central problem lies in the biophysics of TF-DNA interactions: While TFs often bind nonspecifically to the DNA backbone to facilitate targeting by linear diffusion, high-affinity binding requires more selective base-specific interactions.[77] This nonspecific background binding may challenge the interpretation of ChIP-Seq, which is further complicated by contextual variability – binding profiles may differ between tissues, developmental stages and experimental conditions.

Computational predictions based on position weight matrices provide a useful framework, but are also subject to significant limitations. Short TF-binding motifs often occur randomly throughout the genome, making distinguishing between functional from false matches difficult. Even when combined with databases such as the EPD, PWM-based methods may miss regulatory sequences because, for example, (1) not all TFs have well-characterized PWMs, (2) binding tolerance allows TFs to recognize a wide range of sequences beyond the consensus motif, and (3) contextual factors —such as cooperative binding with other TFs and cofactors— are not taken into account. In addition, epigenetic factors such as chromatin accessibility, DNA methylation and histone modifications play a crucial role in determining whether a predicted site will be utilized at a given time point *in vivo*. Therefore, sequence-based predictions alone are not sufficient to determine functional significance. To overcome these limitations, integrative approaches are required. Techniques such as ATAC-Seq, DNase-Seq, and histone ChIP-Seq can provide insights into chromatin

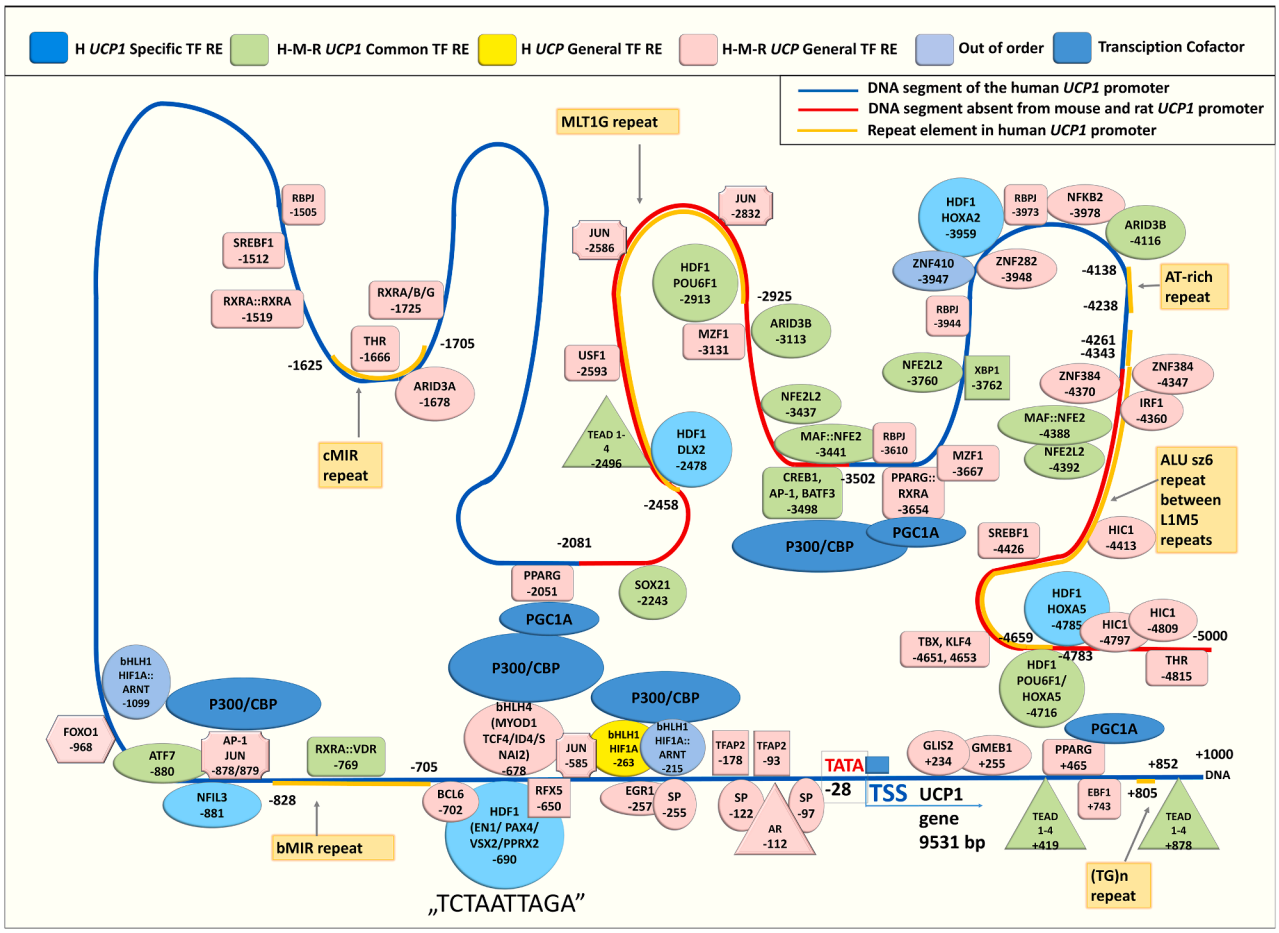


Figure 5. Overview of the TF-REs in the human *UCP1* promoter and their gene or species specificity

This figure provides a comprehensive overview of the TF-REs in the human *UCP1* promoter identified and filtered in our study. The highlighted TF-REs are gene or species-specific, while some previously reported TF-REs without such specificity are indicated for reference. Red lines indicate human promoter segments not present in the mouse and rat *UCP1* promoters. Yellow lines represent repetitive elements identified with the RepeatMasker database. Base pair positions are indicated relative to the transcription start site (TSS). A complete list of transcription factor abbreviations can be found in the "List of abbreviations" section at the end of the article. H, Human; M, Mouse; R, Rat; HDF, Homeodomain family; bHLH, basic-helix-loop-helix family (PowerPoint).

context and regulatory accessibility, while functional assays (e.g., reporter gene assays, CRISPR interference, or targeted mutagenesis) are essential for validating TF-REs and determining their biological role. Ultimately, experimental validation remains the gold standard and should be combined with bioinformatic predictions to refine our understanding of *UCP1* regulation.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and data should be directed to and will be fulfilled by the Lead Contact, Dr. Beáta B. Tóth (toth.beata@etk.unideb.hu).

Materials availability statement

This study did not generate new materials.

Data and code availability

All data generated or analyzed during this study are available in the supplementary files or from publicly accessible databases as cited in the manuscript.

The used RNA-seq data have been deposited to Sequence Read Archive (SRA) database (SRA: https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA607438. The new codes developed for this study for automated data retrival from EPD database is available in github (Github: https://github.com/DEpt-metagenom/TFRE-in-UCP1-promoter). The main and supplementary figures and tables are available in the Mendeley database: Mendeley Data: https://data.mendeley.com/preview/z5b6stz3zx?a=828a4650-5239-487a-b17f-43c667b02454. Specific data queries and scripts used for analysis are available from the lead contact upon reasonable request.

ACKNOWLEDGMENTS

We would like to thank Professors Jan Nedergaard and Barbara Canon for their insightful feedback and valuable suggestions to improve the clarity of the manuscript. On behalf of the project "Bioinformatic analysis of third generation sequencing data" ("Harmadik generációs szekvenálási adatok bioinformatikai elemzése"), we would like to thank for the opportunity to use the HUN-REN Cloud: https://science-cloud.hu/, which helped us to achieve the results published in this paper. During this research B.B.T., L.L., E.V., and G.T.P. was funded by the European Union and the European Regional

Development Fund (GINOP-2.3.4-15-2016-00002, A felsőoktatás és az ipar együttműködése az egészségiparban"), G.H. was supported by the European Union and the European Regional Development Fund (GINOP-2.3.4-15-2020-00008, Komplex Egészségipari Multidiszciplináris Kompetencia Központ kialakítása a Debreceni Egyetemen új innovatív termékek és technológiák fejlesztése érdekében"), and we are thankful for this. B.B.T. was also supported by the University of Debrecen Scientific Research Bridging Fund (DETKA24). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, B.B.T., methodology, B.B.T., G.H., E.V., G.T.P., and L.L.; software, B.B.T., G.H., E.V., G.T.P., and L.L.; formal analysis, B.B.T. and G.T.P.; visualization, B.B.T. and L.L.; investigation, B.B.T.; data curation, B.B.T.; writing – original draft preparation, B.B.T.; writing – review and editing, B.B.T., G.T.P., L.L., and L.F.; project administration, L.L.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Reconstruction of the phylogenetic trees
  - 3D protein structure prediction, alignment and pairwise structural comparison
  - BLAST search to identify specific nucleotide sequences in the human, mouse and rat genome
  - Comparison of orthologous sequences to estimate evolutionary divergence
  - Computational identification of putative transcription factor response elements in gene regulatory sequences using the EPD interface
  - Comparative analysis of the putative TFs in the regulatory sequences of human, mouse and rat *UCP1*, *UCP2* and *UCP3* genes
  - Chromosome position of the investigated *UCP1*, *UCP2* and *UCP3* promoters and genes in the human genome
  - Investigation of TF-RE clusters in the human *UCP1* promoter and their conservation in different species
  - Heatmap visualization and hierarchical cluster analyses of the TF-RE profile of the *UCP* genes and the transcriptomic data of human adipocytes
  - Experimental setup and differentiation protocol of human adipocytes
  - Comparison of the identified putative TF-REs in the regulatory sequence of *UCP1* with the corresponding data found in the chip-seq-based databases
  - Identification of repeat elements in the genome
  - Literature mining using a custom script
  - Identification of TF motif activity in different tissues and exploration of TF interaction partners using the ISMARA online platform
  - Ranking of the transcription factors
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.112969.

Received: April 10, 2024
Revised: October 6, 2024
Accepted: June 18, 2025
Published: June 21, 2025

REFERENCES

1. Tayanloo-Beik, A., Nikkhah, A., Alaei, S., Goodarzi, P., Rezaei-Tavirani, M., Mafi, A.R., Larijani, B., Shouroki, F.F., and Arjmand, B. (2023). Brown adipose tissue and alzheimer's disease. Metab. Brain Dis. *38*, 91–107. https://doi.org/10.1007/s11011-022-01097-z.
2. Seki, T., Yang, Y., Sun, X., Lim, S., Xie, S., Guo, Z., Xiong, W., Kuroda, M., Sakaue, H., Hosaka, K., et al. (2022). Brown-fat-mediated tumour suppression by cold-altered global metabolism. Nature *608*, 421–428. https://doi.org/10.1038/s41586-022-05030-3.
3. Wibmer, A.G., Becher, T., Eljalby, M., Crane, A., Andrieu, P.C., Jiang, C. S., Vaughan, R., Schöder, H., and Cohen, P. (2021). Brown adipose tissue is associated with healthier body fat distribution and metabolic benefits independent of regional adiposity. Cell Rep. Med. *2*, 100332. https://doi.org/10.1016/j.xcrm.2021.100332.
4. Becher, T., Palanisamy, S., Kramer, D.J., Eljalby, M., Marx, S.J., Wibmer, A.G., Butler, S.D., Jiang, C.S., Vaughan, R., Schöder, H., et al. (2021). Brown adipose tissue is associated with cardiometabolic health. Nat. Med. *27*, 58–65. https://doi.org/10.1038/s41591-020-1126-7.
5. Razzoli, M., and Bartolomucci, A. (2016). The Dichotomous Effect of Chronic Stress on Obesity. Trends Endocrinol. Metab. *27*, 504–515. https://doi.org/10.1016/j.tem.2016.04.007.
6. Nedergaard, J., Bengtsson, T., and Cannon, B. (2007). Unexpected evidence for active brown adipose tissue in adult humans. Am. J. Physiol. Endocrinol. Metab. *293*, E444–E452. https://doi.org/10.1152/ajpendo.00691.2006.
7. Feldmann, H.M., Golozoubova, V., Cannon, B., and Nedergaard, J. (2009). UCP1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality. Cell Metab. *9*, 203–209. https://doi.org/10.1016/j.cmet.2008.12.014.
8. Ricquier, D. (2011). Uncoupling Protein 1 of Brown Adipocytes, the Only Uncoupler: A Historical Perspective. Front. Endocrinol. *2*, 85. https://doi.org/10.3389/fendo.2011.00085.
9. Nedergaard, J., Golozoubova, V., Matthias, A., Asadi, A., Jacobsson, A., and Cannon, B. (2001). UCP1: the only protein able to mediate adaptive non-shivering thermogenesis and metabolic inefficiency. Biochim. Biophys. Acta *1504*, 82–106. https://doi.org/10.1016/S0005-2728(00)00247-4.
10. Enerbäck, S., Jacobsson, A., Simpson, E.M., Guerra, C., Yamashita, H., Harper, M.E., and Kozak, L.P. (1997). Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese. Nature *387*, 90–94. https://doi.org/10.1038/387090a0.
11. Hoang, T., Smith, M.D., and Jelokhani-Niaraki, M. (2013). Expression, folding, and proton transport activity of human uncoupling protein-1 (UCP1) in lipid membranes: evidence for associated functional forms. J. Biol. Chem. *288*, 36244–36258. https://doi.org/10.1074/jbc.M113.509935.
12. Price, E.R., Sirsat, T.S., Sirsat, S.K.G., and Dzialowski, E.M. (2019). Sarcoplasmic reticulum Ca2+-ATPase (SERCA) activity during the transition to endothermy in an altricial bird. J. Exp. Biol. *222*, jeb201111. https://doi.org/10.1242/jeb.201111.
13. Ikeda, K., Kang, Q., Yoneshiro, T., Camporez, J.P., Maki, H., Homma, M., Shinoda, K., Chen, Y., Lu, X., Maretich, P., et al. (2017). UCP1-independent signaling involving SERCA2b-mediated calcium cycling regulates beige fat thermogenesis and systemic glucose homeostasis. Nat. Med. *23*, 1454–1465. https://doi.org/10.1038/nm.4429.
14. Bal, N.C., Maurya, S.K., Sopariwala, D.H., Sahoo, S.K., Gupta, S.C., Shaikh, S.A., Pant, M., Rowland, L.A., Bombardier, E., Goonasekera, S.

A., et al. (2012). Sarcolipin is a newly identified regulator of muscle-based thermogenesis in mammals. Nat. Med. *18*, 1575–1579. https://doi.org/10.1038/nm.2897.

15. Tattersall, G.J., Sinclair, B.J., Withers, P.C., Fields, P.A., Seebacher, F., Cooper, C.E., and Maloney, S.K. (2012). Coping with thermal challenges: physiological adaptations to environmental temperatures. Compr. Physiol. *2*, 2151–2202. https://doi.org/10.1002/cphy.c110055.
16. Kazak, L., and Spiegelman, B.M. (2020). Mechanism of futile creatine cycling in thermogenesis. Am. J. Physiol. Endocrinol. Metab. *319*, E947–E949. https://doi.org/10.1152/ajpendo.00444.2020.
17. Cohen, P., and Kajimura, S. (2021). The cellular and functional complexity of thermogenic fat. Nat. Rev. Mol. Cell Biol. *22*, 393–409. https://doi.org/10.1038/s41580-021-00350-0.
18. Nedergaard, J., and Cannon, B. (2014). The Browning of White Adipose Tissue: Some Burning Issues. Cell Metab. *20*, 396–407. https://doi.org/10.1016/j.cmet.2014.07.005.
19. Hanák, P., and Jezek, P. (2001). Mitochondrial uncoupling proteins and phylogenesis–UCP4 as the ancestral uncoupling protein. FEBS Lett. *495*, 137–141. https://doi.org/10.1016/s0014-5793(01)02338-9.
20. Sokolova, I.M., and Sokolov, E.P. (2005). Evolution of mitochondrial uncoupling proteins: novel invertebrate UCP homologues suggest early evolutionary divergence of the UCP family. FEBS Lett. *579*, 313–317. https://doi.org/10.1016/j.febslet.2004.11.103.
21. Hughes, J., and Criscuolo, F. (2008). Evolutionary history of the UCP gene family: gene duplication and selection. BMC Evol. Biol. *8*, 306. https://doi.org/10.1186/1471-2148-8-306.
22. Gorgoglione, R., Porcelli, V., Santoro, A., Daddabbo, L., Vozza, A., Monné, M., Di Noia, M.A., Palmieri, L., Fiermonte, G., and Palmieri, F. (2019). The human uncoupling proteins 5 and 6 (UCP5/SLC25A14 and UCP6/SLC25A30) transport sulfur oxyanions, phosphate and dicarboxylates. Biochim. Biophys. Acta. Bioenerg. *1860*, 724–733. https://doi.org/10.1016/j.bbabio.2019.07.010.
23. Nedergaard, J., and Cannon, B. (2003). The 'Novel''Uncoupling' Proteins UCP2 and UCP3: What Do They Really do? Pros and Cons for Suggested Functions. Exp. Physiol. *88*, 65–84. https://doi.org/10.1113/eph8802502.
24. Rousset, S., Alves-Guerra, M.-C., Mozo, J., Miroux, B., Cassard-Doulcier, A.-M., Bouillaud, F., and Ricquier, D. (2004). The Biology of Mitochondrial Uncoupling Proteins. Diabetes *53*, S130–S135. https://doi.org/10.2337/diabetes.53.2007.S130.
25. Nicholls, D.G. (2021). Mitochondrial proton leaks and uncoupling proteins. Biochim. Biophys. Acta. Bioenerg. *1862*, 148428. https://doi.org/10.1016/j.bbabio.2021.148428.
26. Hirschenson, J., Melgar-Bermudez, E., and Mailloux, R.J. (2022). The Uncoupling Proteins: A Systematic Review on the Mechanism Used in the Prevention of Oxidative Stress. Antioxidants *11*, 322. https://doi.org/10.3390/antiox11020322.
27. Cadenas, S. (2018). Mitochondrial uncoupling, ROS generation and cardioprotection. Biochim. Biophys. Acta. Bioenerg. *1859*, 940–950. https://doi.org/10.1016/j.bbabio.2018.05.019.
28. Himms-Hagen, J., and Harper, M.-E. (2001). Physiological Role of UCP3 May Be Export of Fatty Acids from Mitochondria When Fatty Acid Oxidation Predominates: An Hypothesis. Exp. Biol. Med. *226*, 78–84. https://doi.org/10.1177/153537020122600204.
29. Schrauwen, P., Hinderling, V., Hesselink, M.K.C., Schaart, G., Kornips, E., Saris, W.H.M., Westerterp-Plantenga, M., and Langhans, W. (2002). Etomoxir-induced increase in UCP3 supports a role of uncoupling protein 3 as a mitochondrial fatty acid anion exporter. FASEB J. *16*, 1688–1690. https://doi.org/10.1096/fj.02-0275fje.
30. Tsuboyama-Kasaoka, N., and Ezaki, O. (2001). Mitochondrial uncoupling protein 3 (UCP3) in skeletal muscle. Front. Biosci. *6*, D570–D574. https://doi.org/10.2741/Kasaoka.
31. Rosen, E.D., Walkey, C.J., Puigserver, P., and Spiegelman, B.M. (2000). Transcriptional regulation of adipogenesis. Genes Dev. *14*, 1293–1307.
32. del Mar Gonzalez-Barroso, M., Pecqueur, C., Gelly, C., Sanchis, D., Alves-Guerra, M.-C., Bouillaud, F., Ricquier, D., and Cassard-Doulcier, A.-M. (2000). Transcriptional Activation of the Human ucp1 Gene in a Rodent Cell Line. J. Biol. Chem. *275*, 31722–31732. https://doi.org/10.1074/jbc.M001678200.
33. Rim, J.S., and Kozak, L.P. (2002). Regulatory motifs for CREB-binding protein and Nfe2l2 transcription factors in the upstream enhancer of the mitochondrial uncoupling protein 1 gene. J. Biol. Chem. *277*, 34589–34600. https://doi.org/10.1074/jbc.M108866200.
34. Seale, P., Kajimura, S., Yang, W., Chin, S., Rohas, L.M., Uldry, M., Tavernier, G., Langin, D., and Spiegelman, B.M. (2007). Transcriptional Control of Brown Fat Determination by PRDM16. Cell Metab. *6*, 38–54. https://doi.org/10.1016/j.cmet.2007.06.001.
35. Timmons, J.A., Wennmalm, K., Larsson, O., Walden, T.B., Lassmann, T., Petrovic, N., Hamilton, D.L., Gimeno, R.E., Wahlestedt, C., Baar, K., et al. (2007). Myogenic gene expression signature establishes that brown and white adipocytes originate from distinct cell lineages. Proc. Natl. Acad. Sci. USA *104*, 4401–4406. https://doi.org/10.1073/pnas.0610615104.
36. Villarroya, F., Peyrou, M., and Giralt, M. (2017). Transcriptional regulation of the uncoupling protein-1 gene. Biochimie *134*, 86–92. https://doi.org/10.1016/j.biochi.2016.09.017.
37. ZHANG, X., WANG, C., and LIU, M. (2023). 77-OR: Targeting JunB Induces a Shift from Low- to High-Thermogenic Adipocytes and Ameliorates Diet-Induced Insulin Resistance. Diabetes *72(Supplement_1)*, 77–OR. https://doi.org/10.2337/db23-77-OR.
38. Li, Y., Schwalie, P.C., Bast-Habersbrunner, A., Mocek, S., Russeil, J., Fromme, T., Deplancke, B., and Klingenspor, M. (2019). Systems-Genetics-Based Inference of a Core Regulatory Network Underlying White Fat Browning. Cell Rep. *29*, 4099–4113.e5. https://doi.org/10.1016/j.celrep.2019.11.053.
39. Shapira, S.N., and Seale, P. (2019). Transcriptional Control of Brown and Beige Fat Development and Function. Obesity *27*, 13–21. https://doi.org/10.1002/oby.22334.
40. Ambele, M.A., and Pepper, M.S. (2017). Identification of transcription factors potentially involved in human adipogenesis in vitro. Mol. Genet. Genomic Med. *5*, 210–222. https://doi.org/10.1002/mgg3.269.
41. Inagaki, T., Sakai, J., and Kajimura, S. (2016). Transcriptional and epigenetic control of brown and beige adipose cell fate and function. Nat. Rev. Mol. Cell Biol. *17*, 480–495. https://doi.org/10.1038/nrm.2016.62.
42. Harms, M.J., Lim, H.-W., Ho, Y., Shapira, S.N., Ishibashi, J., Rajakumari, S., Steger, D.J., Lazar, M.A., Won, K.-J., and Seale, P. (2015). PRDM16 binds MED1 and controls chromatin architecture to determine a brown fat transcriptional program. Genes Dev. *29*, 298–307. https://doi.org/10.1101/gad.252734.114.
43. Bonet, M.L., Oliver, P., and Palou, A. (2013). Pharmacological and nutritional agents promoting browning of white adipose tissue. Biochim. Biophys. Acta *1831*, 969–985. https://doi.org/10.1016/j.bbalip.2012.12.002.
44. Perier, R.C. (2000). The Eukaryotic Promoter Database (EPD). Nucleic Acids Res. *28*, 302–303. https://doi.org/10.1093/nar/28.1.302.
45. Sandelin, A., Alkema, W., Engström, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res. *32*, 91D–D94. https://doi.org/10.1093/nar/gkh012.
46. Gaudry, M.J., Jastroch, M., Treberg, J.R., Hofreiter, M., Paijmans, J.L.A., Starrett, J., Wales, N., Signore, A.V., Springer, M.S., et al. (2017). Inactivation of thermogenic UCP1 as a historical contingency in multiple placental mammal clades. Sci. Adv. *3*, e1602878.
47. Keipert, S., Gaudry, M.J., Kutschke, M., Keuper, M., Dela Rosa, M.A.S., Cheng, Y., Monroy Kuhn, J.M., Laterveer, R., Cotrim, C.A., Giere, P., et al. (2024). Two-stage evolution of mammalian adipose tissue thermogenesis. Science *384*, 1111–1117. https://doi.org/10.1126/science.adg1947.

48. Jastroch, M., Wuertz, S., Kloas, W., and Klingenspor, M. (2005). Uncoupling protein 1 in fish uncovers an ancient evolutionary history of mammalian nonshivering thermogenesis. Physiol. Genomics *22*, 150–156. https://doi.org/10.1152/physiolgenomics.00070.2005.

49. McGaugh, S., and Schwartz, T.S. (2017). Here and there, but not everywhere: repeated loss of *uncoupling protein 1* in amniotes. Biol. Lett. *13*, 20160749. https://doi.org/10.1098/rsbl.2016.0749.

50. Crichton, P.G., Lee, Y., and Kunji, E.R.S. (2017). The molecular features of uncoupling protein 1 support a conventional mitochondrial carrier-like mechanism. Biochimie *134*, 35–50. https://doi.org/10.1016/j.biochi.2016.12.016.

51. Dreos, R., Ambrosini, G., Groux, R., Cavin Périer, R., and Bucher, P. (2017). The eukaryotic promoter database in its 30th year: focus on non-vertebrate organisms. Nucleic Acids Res. *45*, D51–D55. https://doi.org/10.1093/nar/gkw1069.

52. Gaudry, M.J., and Campbell, K.L. (2017). Evolution of UCP1 Transcriptional Regulatory Elements Across the Mammalian Phylogeny. Front. Physiol. *8*, 670. https://doi.org/10.3389/fphys.2017.00670.

53. Tóth, B., Barta, Z., Barta, Á.B., and Fésüs, L. (2021). Regulatory modules of human thermogenic adipocytes: functional genomics of large cohort and Meta-analysis derived marker-genes. BMC Genom. *22*, 886. https://doi.org/10.1186/s12864-021-08126-8.

54. Tóth, B.B., Arianti, R., Shaw, A., Vámos, A., Veréb, Z., Póliska, S., Győry, F., Bacso, Z., Fésüs, L., and Kristóf, E. (2020). FTO Intronic SNP Strongly Influences Human Neck Adipocyte Browning Determined by Tissue and PPARγ Specific Regulation: A Transcriptome Analysis. Cells *9*, 987. https://doi.org/10.3390/cells9040987.

55. Balwierz, P.J., Pachkov, M., Arnold, P., Gruber, A.J., Zavolan, M., and van Nimwegen, E. (2014). ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. Genome Res. *24*, 869–884. https://doi.org/10.1101/gr.169508.113.

56. Gehring, W.J., Affolter, M., and Bürglin, T. (1994). HOMEODOMAIN PROTEINS. Annu. Rev. Biochem. *63*, 487–526. https://doi.org/10.1146/annurev.bi.63.070194.002415.

57. Banerjee-Basu, S., and Baxevanis, A.D. (2001). Molecular evolution of the homeodomain family of transcription factors. Nucleic Acids Res. *29*, 3258–3269.

58. Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., and Mann, R.S. (2011). Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. Cell *147*, 1270–1282. https://doi.org/10.1016/j.cell.2011.10.053.

59. Cain, B., Webb, J., Yuan, Z., Cheung, D., Lim, H.-W., Kovall, R.A., Weirauch, M.T., and Gebelein, B. (2023). Prediction of cooperative homeodomain DNA binding sites from high-throughput-SELEX data. Nucleic Acids Res. *51*, 6055–6072. https://doi.org/10.1093/nar/gkad318.

60. Cao, W., Medvedev, A.V., Daniel, K.W., and Collins, S. (2001). β-Adrenergic Activation of p38 MAP Kinase in Adipocytes. J. Biol. Chem. *276*, 27077–27082. https://doi.org/10.1074/jbc.M101049200.

61. Alvarez, R., de Andrés, J., Yubero, P., Viñas, O., Mampel, T., Iglesias, R., Giralt, M., and Villarroya, F. (1995). A Novel Regulatory Pathway of Brown Fat Thermogenesis. J. Biol. Chem. *270*, 5666–5673. https://doi.org/10.1074/jbc.270.10.5666.

62. Liu, Y., Maekawa, T., Yoshida, K., Muratani, M., Chatton, B., and Ishii, S. (2019). The Transcription Factor ATF7 Controls Adipocyte Differentiation and Thermogenic Gene Programming. iScience *13*, 98–112. https://doi.org/10.1016/j.isci.2019.02.013.

63. Laurila, P.-P., Soronen, J., Kooijman, S., Forsström, S., Boon, M.R., Surakka, I., Kaiharju, E., Coomans, C.P., Van Den Berg, S.A.A., Autio, A., et al. (2016). USF1 deficiency activates brown adipose tissue and improves cardiometabolic health. Sci. Transl. Med. *8*, 323ra13. https://doi.org/10.1126/scitranslmed.aad0015.

64. Cassard, A.M., Bouillaud, F., Mattei, M.G., Hentz, E., Raimbault, S., Thomas, M., and Ricquier, D. (1990). Human uncoupling protein gene: Structure, comparison with rat gene, and assignment to the long arm of chromosome 4. J. Cell. Biochem. *43*, 255–264. https://doi.org/10.1002/jcb.240430306.

65. Klingenspor, M., Fromme, T., Hughes, D.A., Manzke, L., Polymeropoulos, E., Riemann, T., Trzcionka, M., Hirschberg, V., and Jastroch, M. (2008). An ancient look at UCP1. Biochim. Biophys. Acta *1777*, 637–641. https://doi.org/10.1016/j.bbabio.2008.03.006.

66. Hughes, D.A., Jastroch, M., Stoneking, M., and Klingenspor, M. (2009). Molecular evolution of UCP1 and the evolutionary history of mammalian non-shivering thermogenesis. BMC Evol. Biol. *9*, 4. https://doi.org/10.1186/1471-2148-9-4.

67. Jastroch, M., Withers, K.W., Taudien, S., Frappell, P.B., Helwig, M., Fromme, T., Hirschberg, V., Heldmaier, G., McAllan, B.M., Firth, B.T., et al. (2008). Marsupial uncoupling protein 1 sheds light on the evolution of mammalian nonshivering thermogenesis. Physiol. Genomics *32*, 161–169. https://doi.org/10.1152/physiolgenomics.00183.2007.

68. Jastroch, M., Oelkrug, R., and Keipert, S. (2018). Insights into brown adipose tissue evolution and function from non-model organisms. J. Exp. Biol. *221*, jeb169425. https://doi.org/10.1242/jeb.169425.

69. Berg, F., Gustafson, U., and Andersson, L. (2006). The uncoupling protein 1 gene (UCP1) is disrupted in the pig lineage: A genetic explanation for poor thermoregulation in piglets. PLoS Genet. *2*, e129–e1181. https://doi.org/10.1371/journal.pgen.0020129.

70. Hou, L., Shi, J., Cao, L., Xu, G., Hu, C., and Wang, C. (2017). Pig has no uncoupling protein 1. Biochem. Biophys. Res. Commun. *487*, 795–800. https://doi.org/10.1016/j.bbrc.2017.04.118.

71. Gaudry, M.J., Jastroch, M., Treberg, J.R., Hofreiter, M., Paijmans, J.L.A., Starrett, J., Wales, N., Signore, A.V., Springer, M.S., and Campbell, K.L. (2017). Inactivation of thermogenic UCP1 as a historical contingency in multiple placental mammal clades. Sci. Adv. *3*, e1602878. https://doi.org/10.1126/sciadv.1602878.

72. Gaudry, M.J., and Campbell, K.L. (2017). Evolution of UCP1 Transcriptional Regulatory Elements Across the Mammalian Phylogeny. Front. Physiol. *8*, 670. https://doi.org/10.3389/fphys.2017.00670.

73. Collins, S., Cao, W., and Robidoux, J. (2004). Learning new tricks from old dogs: β-adrenergic receptors teach new lessons on firing up adipose tissue metabolism. Mol. Endocrinol. *18*, 2123–2131. https://doi.org/10.1210/me.2004-0193.

74. Saito, S., Saito, C.T., and Shingai, R. (2008). Adaptive evolution of the uncoupling protein 1 gene contributed to the acquisition of novel nonshivering thermogenesis in ancestral eutherian mammals. Gene *408*, 37–44. https://doi.org/10.1016/j.gene.2007.10.018.

75. Lu, W.H., Chang, Y.M., and Huang, Y.S. (2021). Alternative Polyadenylation and Differential Regulation of Ucp1: Implications for Brown Adipose Tissue Thermogenesis across Species. Front. Pediatr. *8*, 612279. https://doi.org/10.3389/fped.2020.612279.

76. Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. Trends Genet. *16*, 418–420. https://doi.org/10.1016/s0168-9525(00)02093-x.

77. Pingoud, A., and Jeltsch, A. (2001). Structure and function of type II restriction endonucleases. Nucleic Acids Res. *29*, 3705–3727. https://doi.org/10.1093/nar/29.18.3705.

78. Linheiro, R.S., and Bergman, C.M. (2008). Testing the palindromic target site model for DNA transposon insertion using the Drosophila melanogaster P-element. Nucleic Acids Res. *36*, 6199–6208. https://doi.org/10.1093/nar/gkn563.

79. McClintock, B. (1950). The Origin and Behavior of Mutable Loci in Maize. Proc. Natl. Acad. Sci. USA *36*, 344–355.

80. Britten, R.J., and Davidson, E.H. (1969). Gene Regulation for Higher Cells: A Theory. Science *165*, 349–357. https://doi.org/10.1126/science.165.3891.349.

81. Thompson, P.J., Macfarlan, T.S., and Lorincz, M.C. (2016). Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. Mol. Cell *62*, 766–776. https://doi.org/10.1016/j.molcel.2016.03.029.

82. Cowley, M., and Oakey, R.J. (2013). Transposable elements re-wire and fine-tune the transcriptome. PLoS Genet. *9*, e1003234. https://doi.org/10.1371/journal.pgen.1003234.

83. Gifford, W.D., Pfaff, S.L., and Macfarlan, T.S. (2013). Transposable elements as genetic regulatory substrates in early development. Trends Cell Biol. *23*, 218–226. https://doi.org/10.1016/j.tcb.2013.01.001.

84. Lynch, V.J., Leclerc, R.D., May, G., and Wagner, G.P. (2011). Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat. Genet. *43*, 1154–1159. https://doi.org/10.1038/ng.917.

85. Li, L., Li, Q., Willis, L.G., Erlandson, M., Theilmann, D.A., and Donly, C. (2005). Complete comparative genomic analysis of two field isolates of Mamestra configurata nucleopolyhedrovirus-A. J. Gen. Virol. *86*, 91–105. https://doi.org/10.1099/vir.0.80488-0.

86. Horton, C.A., Alexandari, A.M., Hayes, M.G.B., Marklund, E., Schaepe, J. M., Aditham, A.K., Shah, N., Suzuki, P.H., Shrikumar, A., Afek, A., et al. (2023). Short tandem repeats bind transcription factors to tune eukaryotic gene expression. Science *381*, eadd1250. https://doi.org/10.1126/science.add1250.

87. Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science *351*, 1083–1087. https://doi.org/10.1126/science.aad5497.

88. Emera, D., and Wagner, G.P. (2012). Transformation of a transposon into a derived prolactin promoter with function during human pregnancy. Proc. Natl. Acad. Sci. USA *109*, 11246–11251. https://doi.org/10.1073/pnas.1118566109.

89. Kunarso, G., Chia, N.-Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.-S., Ng, H.-H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. *42*, 631–634. https://doi.org/10.1038/ng.600.

90. Scarpato, M., Angelini, C., Cocca, E., Pallotta, M.M., Morescalchi, M.A., and Capriglione, T. (2015). Short interspersed DNA elements and miRNAs: a novel hidden gene regulation layer in zebrafish? Chromosome Res. *23*, 533–544. https://doi.org/10.1007/s10577-015-9484-6.

91. Oriente, F., Perruolo, G., Cimmino, I., Cabaro, S., Liotti, A., Longo, M., Miele, C., Formisano, P., and Beguinot, F. (2018). Prep1, A Homeodomain Transcription Factor Involved in Glucose and Lipid Metabolism. Front. Endocrinol. *9*, 346. https://doi.org/10.3389/fendo.2018.00346.

92. Hilton, C., Karpe, F., and Pinnick, K.E. (2015). Role of developmental transcription factors in white, brown and beige adipose tissues. Biochim. Biophys. Acta *1851*, 686–696. https://doi.org/10.1016/j.bbalip.2015.02.003.

93. Waldén, T.B., Hansen, I.R., Timmons, J.A., Cannon, B., and Nedergaard, J. (2012). Recruited vs. nonrecruited molecular signatures of brown, "brite," and white adipose tissues. Am. J. Physiol. Endocrinol. Metab. *302*, E19–E31. https://doi.org/10.1152/ajpendo.00249.2011.

94. Cantile, M., Procino, A., D'armiento, M., Cindolo, L., and Cillo, C. (2003). *HOX* gene network is involved in the transcriptional regulation of in vivo human adipogenesis. J. Cell. Physiol. *194*, 225–236. https://doi.org/10.1002/jcp.10210.

95. Park, M.J., Kim, H.Y., Kim, K., and Cheong, J. (2009). Homeodomain transcription factor CDX1 is required for the transcriptional induction of PPARgamma in intestinal cell differentiation. FEBS Lett. *583*, 29–35. https://doi.org/10.1016/j.febslet.2008.11.030.

96. Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-Binding Specificities of Human Transcription Factors. Cell *152*, 327–339. https://doi.org/10.1016/j.cell.2012.12.009.

97. Dror, I., Zhou, T., Mandel-Gutfreund, Y., and Rohs, R. (2014). Covariation between homeodomain transcription factors and the shape of their DNA binding sites. Nucleic Acids Res. *42*, 430–441. https://doi.org/10.1093/nar/gkt862.

98. Zhang, G., Sun, Q., and Liu, C. (2016). Influencing Factors of Thermogenic Adipose Tissue Activity. Front. Physiol. *7*, 29. https://doi.org/10.3389/fphys.2016.00029.

99. Singh, S., Rajput, Y.S., Barui, A.K., Sharma, R., and Grover, S. (2015). Expression of developmental genes in brown fat cells grown in vitro is linked with lipid accumulation. In Vitro Cell. Dev. Biol. Anim. *51*, 1003–1011. https://doi.org/10.1007/s11626-015-9930-y.

100. Singh, S., Rajput, Y.S., Barui, A.K., Sharma, R., and Datta, T.K. (2016). Fat accumulation in differentiated brown adipocytes is linked with expression of Hox genes. Gene Expr. Patterns *20*, 99–105. https://doi.org/10.1016/j.gep.2016.01.002.

101. Cao, W., Huang, H., Xia, T., Liu, C., Muhammad, S., and Sun, C. (2018). Homeobox a5 Promotes White Adipose Tissue Browning Through Inhibition of the Tenascin C/Toll-Like Receptor 4/Nuclear Factor Kappa B Inflammatory Signaling in Mice. Front. Immunol. *9*, 647. https://doi.org/10.3389/fimmu.2018.00647.

102. Kato, H., Ario, T., Kishida, T., Tadano, M., Osawa, S., Maeda, Y., Takakura, H., and Izawa, T. (2021). Homeobox A5 and C10 genes modulate adaptation of brown adipose tissue during exercise training in juvenile rats. Exp. Physiol. *106*, 463–474. https://doi.org/10.1113/EP089114.

103. Mohsen-Kanson, T., Hafner, A.-L., Wdziekonski, B., Takashima, Y., Villageois, P., Carrière, A., Svensson, M., Bagnis, C., Chignon-Sicard, B., Svensson, P.-A., et al. (2014). Differentiation of Human Induced Pluripotent Stem Cells into Brown and White Adipocytes: Role of Pax3. Stem Cell. *32*, 1459–1467. https://doi.org/10.1002/stem.1607.

104. Brun, T., He, K.H.H., Lupi, R., Boehm, B., Wojtusciszyn, A., Sauter, N., Donath, M., Marchetti, P., Maedler, K., and Gauthier, B.R. (2007). The diabetes-linked transcription factor Pax4 is expressed in human pancreatic islets and is activated by mitogens and GLP-1. Hum. Mol. Genet. *17*, 478–489. https://doi.org/10.1093/hmg/ddm325.

105. Xu, L., Xu, C., Zhou, S., Liu, X., Wang, J., Liu, X., Qian, S., Xin, Y., Gao, Y., Zhu, Y., and Tang, X. (2017). PAX4 promotes PDX1-induced differentiation of mesenchymal stem cells into insulin-secreting cells. Am. J. Transl. Res. *9*, 874–886.

106. Lau, H.H., Krentz, N.A.J., Abaitua, F., Perez-Alcantara, M., Chan, J.-W., Ajeian, J., Ghosh, S., Lee, Y., Yang, J., Thaman, S., et al. (2023). PAX4 loss of function increases diabetes risk by altering human pancreatic endocrine cell development. Nat. Commun. *14*, 6119. https://doi.org/10.1038/s41467-023-41860-z.

107. Gupta, A., Efthymiou, V., Kodani, S.D., Shamsi, F., Patti, M.E., Tseng, Y.-H., and Streets, A. (2023). Mapping the transcriptional landscape of human white and brown adipogenesis using single-nuclei RNA-seq. Mol. Metab. *74*, 101746. https://doi.org/10.1016/j.molmet.2023.101746.

108. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872. https://doi.org/10.1016/j.cell.2007.11.019.

109. Xiao, W., Geng, W., Zhou, M., Xu, J., Wang, S., Huang, Q., Sun, Y., Li, Y., Yang, G., and Jin, Y. (2022). POU6F1 cooperates with RORA to suppress the proliferation of lung adenocarcinoma by downregulation HIF1A signaling pathway. Cell Death Dis. *13*, 427. https://doi.org/10.1038/s41419-022-04857-y.

110. Basse, A.L., Isidor, M.S., Winther, S., Skjoldborg, N.B., Murholm, M., Andersen, E.S., Pedersen, S.B., Wolfrum, C., Quistorff, B., and Hansen, J.B. (2017). Regulation of glycolysis in brown adipocytes by HIF-1α. Sci. Rep. *7*, 4052. https://doi.org/10.1038/s41598-017-04246-y.

111. Halberg, N., Khan, T., Trujillo, M.E., Wernstedt-Asterholm, I., Attie, A.D., Sherwani, S., Wang, Z.V., Landskroner-Eiger, S., Dineen, S., Magalang, U.J., et al. (2009). Hypoxia-Inducible Factor 1α Induces Fibrosis and

Insulin Resistance in White Adipose Tissue. Mol. Cell Biol. *29*, 4467–4483. https://doi.org/10.1128/mcb.00192-09.

112. Scarpulla, R.C. (2008). Transcriptional paradigms in mammalian mitochondrial biogenesis and function. Physiol. Rev. *88*, 611–638. https://doi.org/10.1152/physrev.00025.2007.
113. Kalaany, N.Y., and Mangelsdorf, D.J. (2006). LXRS AND FXR: The Yin and Yang of Cholesterol and Fat Metabolism. Annu. Rev. Physiol. *68*, 159–191. https://doi.org/10.1146/annurev.physiol.68.033104.152158.
114. Fedorenko, A., Lishko, P.V., and Kirichok, Y. (2012). Mechanism of fatty-acid-dependent UCP1 uncoupling in brown fat mitochondria. Cell *151*, 400–413. https://doi.org/10.1016/j.cell.2012.09.010.
115. Tharp, K.M., Kang, M.S., Timblin, G.A., Dempersmier, J., Dempsey, G.E., Zushin, P.-J.H., Benavides, J., Choi, C., Li, C.X., Jha, A.K., et al. (2018). Actomyosin-Mediated Tension Orchestrates Uncoupled Respiration in Adipose Tissues. Cell Metab. *27*, 602–615.e4. https://doi.org/10.1016/j.cmet.2018.02.005.
116. Cobbaut, M., Karagil, S., Bruno, L., Diaz de la Loza, M.D.C., Mackenzie, F.E., Stolinski, M., and Elbediwy, A. (2020). Dysfunctional Mechanotransduction through the YAP/TAZ/Hippo Pathway as a Feature of Chronic Disease. Cells *9*, 151. https://doi.org/10.3390/cells9010151.
117. Nguyen, T., Nioi, P., and Pickett, C.B. (2009). The Nrf2-Antioxidant Response Element Signaling Pathway and Its Activation by Oxidative Stress. J. Biol. Chem. *284*, 13291–13295. https://doi.org/10.1074/jbc.R900010200.
118. Yoshida, H., Nadanaka, S., Sato, R., and Mori, K. (2006). XBP1 is Critical to Protect Cells from Endoplasmic Reticulum Stress: Evidence from Site-2 Protease-deficient Chinese Hamster Ovary Cells. Cell Struct. Funct. *31*, 117–125. https://doi.org/10.1247/csf.06016.
119. Ziętak, M., Chabowska-Kita, A., and Kozak, L.P. (2017). Brown fat thermogenesis: Stability of developmental programming and transient effects of temperature and gut microbiota in adults. Biochimie *134*, 93–98. https://doi.org/10.1016/j.biochi.2016.12.006.
120. Farmer, S.R. (2006). Transcriptional control of adipocyte formation. Cell Metab. *4*, 263–273. https://doi.org/10.1016/j.cmet.2006.07.001.
121. Giralt, M., and Villarroya, F. (2013). White, Brown, Beige/Brite: Different Adipose Cells for Different Functions? Endocrinology *154*, 2992–3000. https://doi.org/10.1210/en.2013-1403.
122. Bi, P., Shan, T., Liu, W., Yue, F., Yang, X., Liang, X.-R., Wang, J., Li, J., Carlesso, N., Liu, X., and Kuang, S. (2014). Notch signaling regulates adipose browning and energy metabolism. Nat. Med. *20*, 911–918. https://doi.org/10.1038/nm.3615.
123. Wunderlich, Z., and Mirny, L.A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. Trends Genet. *25*, 434–440. https://doi.org/10.1016/j.tig.2009.08.003.
124. Su, T., Liu, P., Ti, X., Wu, S., Xue, X., Wang, Z., Dioum, E., and Zhang, Q. (2019). HIF1α, EGR1 and SP1 co-regulate the erythropoietin receptor expression under hypoxia: An essential role in the growth of non-small cell lung cancer cells. Cell Commun. Signal. *17*, 152. https://doi.org/10.1186/s12964-019-0458-8.
125. Zhang, J., Zhang, Y., Sun, T., Guo, F., Huang, S., Chandalia, M., Abate, N., Fan, D., Xin, H.B., Chen, Y.E., and Fu, M. (2013). Dietary obesity-induced Egr-1 in adipocytes facilitates energy storage via suppression of FOXC2. Sci. Rep. *3*, 1476. https://doi.org/10.1038/srep01476.
126. Xue, B., Coulter, A., Rim, J.S., Koza, R.A., and Kozak, L.P. (2005). Transcriptional Synergy and the Regulation of Ucp1 during Brown Adipocyte Induction in White Fat Depots. Mol. Cell Biol. *25*, 8311–8322. https://doi.org/10.1128/MCB.25.18.8311-8322.2005.
127. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.
128. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
129. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
130. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics *27*, 592–593. https://doi.org/10.1093/bioinformatics/btq706.
131. Paradis, E. (2010). pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics *26*, 419–420. https://doi.org/10.1093/bioinformatics/btp696.
132. Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. *16*, 111–120. https://doi.org/10.1007/BF01731581.
133. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
134. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242. https://doi.org/10.1093/nar/28.1.235.
135. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402. https://doi.org/10.1093/nar/25.17.3389.
136. Weirauch, M.T., Cote, A., Norel, R., Annala, M., Zhao, Y., Riley, T.R., Saez-Rodriguez, J., Cokelaer, T., Vedenko, A., Talukder, S., et al. (2013). Evaluation of methods for modeling transcription factor sequence specificity. Nat. Biotechnol. *31*, 126–134. https://doi.org/10.1038/nbt.2486.
137. Benita, Y., Kikuchi, H., Smith, A.D., Zhang, M.Q., Chung, D.C., and Xavier, R.J. (2009). An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res. *37*, 4587–4602. https://doi.org/10.1093/nar/gkp425.
138. Kreft, Ł., Soete, A., Hulpiau, P., Botzki, A., Saeys, Y., and De Bleser, P. (2017). ConTra v3: a tool to identify transcription factor binding sites across species, update 2017. Nucleic Acids Res. *45*, W490–W494. https://doi.org/10.1093/nar/gkx376.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Eukaryotic Promoter Database (EPD)	EPDnew	[https://epd.epfl.ch] Accessed Jun 2023
Ensembl Genome Browser	Ensembl	[https://www.ensembl.org] Release 111, Accessed Jan 2024
JASPAR	JASPAR 2024, MA database	[https://jaspar.genereg.net] Accessed Jun 2023
ChipBase	National Institute of Genetics, Japan	https://chip-atlas.org
TFLink	TFLink Database	[https://tflink.net] Accessed Nov 2023
UCSC Genome Browser hg19	UCSC	[https://genome.ucsc.edu] Accessed Nov 2023
R	The R Foundation	Version 4.3.1
Pheatmap	Pheatmap	https://cran.r-project.org/web/packages/pheatmap/index.html
Morpheus	Broad Institute	[https://software.broadinstitute.org/morpheus] Accessed Jan 2024
ISMARA	ISMARA	[https://ismara.unibas.ch/] Accessed Jan 2024
contra v3	Contra v3	[https://bioinfo.lifl.fr/contrav3/] Accessed Jan 2024
Alpha Fold	Alpha Fold 2	https://alphafold.ebi.ac.uk/ Accessed Jan 2024
RCSB	RCSB Protein Data Bank	https://www.rcsb.org
BLAST	NCBI	https://blast.ncbi.nlm.nih.gov/
Multiz Alignments of 100 Vertebrates	UCSC Genome Browser	[https://genome.ucsc.edu] Accessed Nov 2023
Deposited data		
Custom PowerShell and R Scripts	This paper	https://github.com/DEpt-metagenom/TFRE-in-UCP1-promoter; https://data.mendeley.com/preview/z5b6stz3zx?a=828a4650-5239-487a-b17f-43c667b02454
Sequence read archive	NCBI	Accession number PRJNA607438

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study does not involve any experimental models or human participants.

METHOD DETAILS

Reconstruction of the phylogenetic trees

To generate a combined human, mouse and rat *UCP1, UCP2*, and *UCP3* amino acid, coding sequence and promoter sequence phylogenetic tree, the amino acid and the mRNA data set was downloaded from the National Library of Medicine (NCBI: https://www.ncbi.nlm.nih.gov), whereas the sequences of gene promoters were downloaded from the EPD database (EPD: https://epd.epfl.ch/). Accession numbers of the promoters' FASTA sequences: human *UCP1* FP007435, *UCP2_1* FP016503, *UCP2_2* FP016504, *UCP3_1* FP016505, *UCP3_2* FP016506; mouse *UCP1* FP011827, *UCP2* FP010492, *UCP3* FP010490; rat *UCP1* FP011518, *UCP2* FP000767, *UCP3* FP000766; Accession numbers of the mRNAs' FASTA sequences: human UCP1 NM_021833.5, UCP2 NM_001381943.1, UCP3 NM_003356.4; mouse UCP1 NM_009463.3, UCP2 NM_011671.6, UCP3 NM_009464.3; rat UCP1 NM_012682.2, UCP2 NM_019354.3, UCP3 NM_013167.2; Accession numbers of the amino-acids' FASTA sequences: human UCP1 NP_068605, UCP2 AAC51336, UCP3 AAC51367.1; mouse UCP1 NP_033489, UCP2 AAB17666, UCP3 AAB87084; rat UCP1

NP_036814, UCP2 BAA25698, UCP3 AAB71523. The resulting amino acid, exon and promoter sequences were aligned using MAFFT 7.505,[127] and a maximum likelihood tree was constructed using IQtree2 2.0.7[128] with automatic model selection ('-m MFP') turned on and using 1 000 aLRT and 1 000 rapid bootstrap replications to test the robustness of the analysis. The alignment of the mRNA sequences had a length of 3,398 bp, of which 1536 were variable and 878 informative, whereas the promoters had 7,956 variables and 5,412 informative sites on an alignment length of 12585 bp. Of the 313 aligned amino acid positions, we identified 172 polymorphic and 150 informative characters. Phylogenetic trees were visualized with packages ape 5.7.1[129] and phangorn 2.11.1 using R 4.2.2 (R Foundation for Statistical Computing, n.d.).[130] Pairwise genetic distances of *UCP* promoters and exons were calculated with the R package pegas 1.2[131] using the "K80" model of nucleotide substitution (Kimura's 2-parameters distance metric[132]). Pairwise distances were visualized as a neighbour-joining phylogram with pegas, as a UPGMA dendrogram with phangorn, and as a heatmap with pheatmap 1.0.12 (Pheatmap: https://cran.r-project.org/web/packages/pheatmap/index.html).

3D protein structure prediction, alignment and pairwise structural comparison

The 3D structures of human, mouse and rat UCP1, UCP2 and UCP3 proteins were downloaded using the AlphaFold 2 protein structure prediction interface (AlphaFold 2: https://alphafold.ebi.ac.uk/) developed by DeepMind.[133] Colors indicate model confidence, AlphaFold produces a per-residue confidence score (pLDDT) between 0 and 100. Some regions below 50 pLDDT may be unstructured in isolation. Dark blue: Very high (pLDDT >90); Light blue: Confident (90 > pLDDT >70); Yellow: Low (70 > pLDDT >50); Red: Very low (pLDDT <50).

For alignment of the predicted protein structures we used the RCSB PBD (RCSB PBD: https://www.rcsb.org; Berman et al.[134]) 3D protein structure comparison interface. For protein comparisons, pairwise comparisons were performed with human UCP1. The metrics used to describe the degree of pairwise structure alignment and structural similarity are as follows:

RMSD (root-mean-square deviation) is computed between aligned pairs of the backbone C-alpha atoms in superposed structures in Å resolution. The lower the RMSD, the better the structure alignment between the structures. This is the most commonly reported metric when comparing two structures, but it is sensitive to local structure deviation. If a few residues in a loop are not aligned, the RMSD value is large, even though the rest of the structure is well aligned.

TM-score (template modeling score) is a measure of topological similarity between the template and model structures. The TM-score ranges between 0 and 1, where 1 indicates a perfect match and 0 is no match between the two structures. Scores <0.2 usually indicate that the proteins are unrelated while those >0.5 generally have the same protein fold (e.g., classified by SCOP/CATH).

Identity (sequence identity percentage) is the percentage of paired residues in the alignment that are identical in sequence.

Equivalent Residues is the number of residue pairs that are structurally equivalent in the alignment.

Sequence Length is the total number of polymeric residues in the deposited sequence for a given chain.

Modeled Residues is the number of residues with coordinates that were used for structure alignment.

BLAST search to identify specific nucleotide sequences in the human, mouse and rat genome

The BLAST interface (BLAST: https://blast.ncbi.nlm.nih.gov/) was used to identify sequences similar to a given nucleotide sequence in the human, mouse and rat genomes.[135] The human RefSeqGene, rat mRatBN7.2 and mouse GRCm39 reference genome sequences were used in the searches. The graphical summary of the results in Figure S1H shows the percentage of homology in different colors, while the position indicates the sequence region of the analyzed DNA segment in which the homologous segment is located, showing the hits in separate rows.

Comparison of orthologous sequences to estimate evolutionary divergence

We used the phylogenetic tree based on the species' genome provided by the Multiz Alignment of 100 Species interface integrated into the UCSC genome browser. We then manually checked the presence of the 1419 bp sequence missing in the mouse and rat genomes in the corresponding sequences of the other species. In this way, we estimated the evolutionary divergence based on the available genomic data and the presence of the sequence in the different species.

Computational identification of putative transcription factor response elements in gene regulatory sequences using the EPD interface

The identification of putative TF-REs within the regulatory sequences of the genes under study was performed using the Eukaryotic Promoter Database (EPD) interface (EPD: https://epd.epfl.ch).[44] EPD's approach to identify TF-REs relies on the JASPAR Transcription Motifs Library (specifically JASPAR CORE 2018 vertebrates),[45] which includes a diverse set of TF motifs, including generic promoter motifs such as the TATA box, Initiator, GC box and CCAAT box. This library contains a large body of information represented as a position weight matrix (PWM) for 579 TF-binding sequences, including dimers, trimers and variants (based on the PWM of the Jaspar database, it is also possible to search in the EPD interface for the RE of some dimerised or tetrameric TFs described in the literature. Although we do not provide direct evidence for protein-protein interactions or dimerisation in this study, we report the presence of REs corresponding to known dimers. Many approaches have been developed to model and learn a protein's DNA-binding specificity; however, comparative analyses indicate that simple models based on mononucleotide PWMs trained by the best methods perform similarly to more complex models.[136]

To determine whether a given binding sequence is present within the regulatory sequence of the target gene, the EPD interface uses a probability-based algorithm with four different *p*-value thresholds (*p*-value <0.01, 0.001, 0.0001 and 0.00001). In this study, we focused on the two most stringent *p*-value thresholds ($p < 0.0001$ and $p < 0.00001$), as empirical evidence from the literature suggests that more permissive probabilities can lead to high false positive rates.[137] However, our exploratory studies showed that for some genes with experimentally proven regulatory TFs, the software accurately identified the binding sequence only at $p < 0.0001$, as exemplified by the occurrence of the HIF1A TF-RE for genes such as *VEGFA* and *GLUT1*. Furthermore, considering that a specific RE with an average length of 12 bp can be found randomly per 4000 nucleotides,[123] the identified occurrence has limited significance. Rather, it suggests that the TFs for which we did not detect the RE are less likely to be involved in the regulation of gene expression. Therefore, we primarily presented our results using the less stringent criterion of p < 1e-4. To mitigate the problem of false positives, we applied additional screening methods to filter out functionally less relevant REs and highlight potential species- and gene-specific TF-REs. Although a stricter threshold (p < 1e-5) would reduce the number of false-positive results, it would also increase the number of false-negative results, and important results could be overlooked. False-positive results can be controlled by collecting more detailed information and further filtering. In choosing the significance level (p < 1e-4), we were guided by the possibility of comparing the results of similar screenings of paralogous and orthologous *UCP* promoters and filtering out false-positive results in the analysis. However, for reasons of comparison and transparency, we also present results with a stricter threshold (p < 1e-5) in the tables.

In this study, we analyzed regulatory sequences 5 kb upstream and 1 kb downstream of the TSS for the investigated genes (e.g.,: human *UCP1*: GRCh37/hg19; Chr. 4: 141,489,115–141,495,114; TSS: Chr. 4:141,489,115). The relevant base pair boundaries of the upstream regulatory DNA segment may vary from gene to gene, therefore, the broadest DNA region allowed by the EPD interface was analyzed. Data retrieval was automated with a custom script we made publicly available on GitHub and Mendeley to increase transparency and reproducibility.

Comparative analysis of the putative TFs in the regulatory sequences of human, mouse and rat *UCP1*, *UCP2* and *UCP3* genes

The data analysis and screening were performed in an interactive R environment. First, data were extracted from the EPD interface (EPD: https://epd.epfl.ch) as individual data files (text files) for each target gene. EPD utilises the JASPAR database, which contains PWMs of TF-REs, to investigate their presence within specific DNA sequences (JASPAR: https://jaspar.genereg.net). The EPD database contains two promoter sequences for human *UCP2* and *UCP3*, as these genes have alternative transcription start sites, with the TSSs of the two variants differing by a few 100 bp. We did not want to favor either variant, so both were included in the analysis. Subsequently, the data extracted for each gene were summarised in a unified database containing information on the occurrence of the 579 TF-binding sequences in the promoter regions of three genes: *UCP1*, *UCP2* and *UCP3* in humans, mice and rats, respectively. The resulting matrix was then used to identify different clusters of TF-REs based on specific criteria. These criteria included the identification of TF-REs specific to human *UCP1* (not present in human *UCP2* or *UCP3*, but possibly present in mouse and rat *UCP1*, *UCP2* or *UCP3*), unique to human *UCP1* (exclusively present in human *UCP1*, not present in the promoters of the other genes analyzed), specific for *UCP1* (present in human, mouse and rat *UCP1*, not present in human, mouse and rat *UCP2* and *UCP3*), common for human, mouse and rat *UCP1* (but not present in human *UCP2* and *UCP3*, but possibly present in mouse and rat *UCP2* and *UCP3*) and common for human *UCP*s (present in human *UCP1*, *UCP2* and *UCP3*, but not in the other analyzed *UCP* promoters in mouse and rat).

Chromosome position of the investigated *UCP1*, *UCP2* and *UCP3* promoters and genes in the human genome

UCP1 promoter in human: GRCh 37/hg19 Chr 4:141,489,115–141,495,114 size: 6000 bases; *UCP1* gene in human: GRCh 37/hg19 Chr 4:141,480,585–141,490,115 size: 9531 bases, minus strand; *UCP2* promoter in human: GRCh 37/hg19 Chr11:73,693,248 - 73,699,247 size: 6000 bases; *UCP2* gene in human: GRCh 37/hg19 Chr11:73,685,717 - 73,694,247 size: 8531 bases; *UCP3* promoter in human: GRCh 37/hg19 Chr11:73,719,130 - 73,725,129 size: 6000 bases; *UCP3* gene in human: GRCh 37/hg19 chr11:73,711,322 - 73,720,130; size: 8809 bases.

Investigation of TF-RE clusters in the human *UCP1* promoter and their conservation in different species

TF-REs were searched in the vicinity (±40 bp) of the identified putative human-specific and conserved TF-REs to investigate possible regulatory clusters using EPD and Contra v3 (Contra v3: http://bioit2.irc.ugent.be/contra/v3/; Kreft et al.[138]). The identified TF-RE sequences in human *UCP1* were aligned to 100 vertebrate species (hg19 Multiz Alignments of 100 Vertebrates) to assess evolutionary conservation, and the results were checked against NCBI and UCSC databases (UCSC: http://genome.ucsc.edu/).

Heatmap visualization and hierarchical cluster analyses of the TF-RE profile of the *UCP* genes and the transcriptomic data of human adipocytes

Hierarchical cluster analyses and heatmap visualization were performed with the Morpheus web tool (Morpheus: https://software.broadinstitute.org/morpheus/) using Spearman rank correlation of rows and columns and complete linkage based on the presence or absence of the TF-RE in the given gene upstream regulatory region. The same tools were used to visualise the relative gene-expression profile of human neck-derived adipocytes.[54] RNAseq data is deposited in the Sequence Read Archive (SRA) database

(SRA: https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA607438. Heatmaps depict the z-scores calculated from the filtered and normalized transcriptomic data of isolated and maintained preadipocytes from subcutaneous and deep neck adipose stromal cells and lipid-laden mature adipocytes differentiated *in vitro* with white and brown protocols described in Tóth et al.[54] The order of samples in the cluster analysis (see heatmap) was determined to emphasise differential gene expression due to tissue origin and differentiation protocol, while hierarchical cluster analysis (Pearson correlation) was used to examine the similarity of gene expression data (rows).

Experimental setup and differentiation protocol of human adipocytes

Our previous study investigated three factors (anatomical origin, PPARγ stimulation by rosiglitazone and alleles of the FTO locus rs1421085 (three of each FTO rs1421085 genotype: T/T-risk-free donor 1–3, T/C-heterozygous donor 4–6, and C/C-obesity-risk donor 7–9) to determine how they influence browning of adipocytes during differentiation and asked which signaling pathways and processes they trigger for increased browning.[54] To this end, we investigated and compared global gene expression patterns by RNA sequencing of hASC-derived white and brown (in response to sustained PPARγ stimulation by rosiglitazone) differentiated adipocytes. Human adipose stromal cells (hASCs) were isolated from paired DN and SC adipose tissue samples from nine donors, three of each FTO rs1421085 genotype: T/T-risk-free, T/C-heterozygous and C/C-obese-risk. Isolated hASCs were grown to confluence and then differentiated with white and brown differentiation cocktail on 6-well plates at 37°C and 5% CO2. In brief, cells were grown in DMEM-F12-HAM (Sigma-Aldrich cat#D8437) medium supplemented with 10% FBS (Gibco cat#10270106); 10% EGM2 (Lonza #CC3162), 1% biotin (Sigma-Aldrich cat#B4639), 1% pantothenic acid (Sigma-Aldrich cat#P5155), 1% streptomycin-penicillin (Sigma-Aldrich cat#P4333): white differentiations were induced in the first 4 days by hormone cocktails in serum- and additive-free DMEM-F12-HAM (Sigma-Aldrich cat#D8437) medium with biotin, pantothenic acid, apo-transferrin 10ug/ml (Sigma-Aldrich cat#T2252), insulin 20nM (Sigma-Aldrich cat#I9278), T3 3nM (Sigma-Aldrich cat#T6397), dexamethasone 25nM (Sigma-Aldrich cat#D1756), hydrocortisone (Sigma-Aldrich cat#H0888) rosiglitazone 2 μM (Cayman Chemicals cat#71740) and IBMX 250 μM (Sigma-Aldrich cat#I5879). Later, rosiglitazone, dexamethasone and IBMX were omitted from the media for 10 days. To compose the brown media for thermogenic induction, a serum- and additive-free DMEM-F12-HAM (Sigma-Aldrich cat#D8437) medium containing biotin, pantothenic acid, apo-transferrin 10ug/ml (Sigma-Aldrich cat#T2252), insulin 850nM (Sigma-Aldrich cat#I9278), T3 3nM (Sigma-Aldrich cat#T6397) and rosiglitazone 500nM (Cayman Chemicals cat#71740). The insulin concentration was 42.5x higher than that of the white differentiation cocktail.

Comparison of the identified putative TF-REs in the regulatory sequence of *UCP1* with the corresponding data found in the chip-seq-based databases

The list of putative TF-REs identified by the EPD interface (EPD: https://epd.epfl.ch) at $p < 1e\text{-}4$ confidence level in human *UCP1* was compared with similar data from public ChipSeq databases. ChipSeq databases collect experimentally validated DNA binding data for TFs to the promoter of individual genes. Two ChipSeq databases, ChiPBase (ChiPBase: https://rnasysu.com) and TFlink (TFlink: https://tflink.net/) were used to identify proteins/TFs binding in the regulatory DNA sequence of the human *UCP1* gene range between −5000 bp and +1000 bp.

Identification of repeat elements in the genome

The RepeatMasker interface (RepeatMasker: https://www.repeatmasker.org/) integrated into the UCSC genome browser (UCSC: https://genome.ucsc.edu/) was used to identify, categorize and determine the exact position of repeat elements in the human *UCP1* promoter (GRCh 37/hg19 by NCBI Gene, Chr4: 141.489.115–141.495.114).

Literature mining using a custom script

We programmed a Python script which uses the Entrez E-utilities (Entrez E-utilities: A General Introduction to the E-utilities - Entrez Programming Utilities Help - NCBI Bookshelf) ESearch and EFetch to query the PubMed database and retrieve relevant abstracts. The script summarizes the search results in a convenient, human-readable tabular format which is manually examined by the researcher.

We compiled a list of abbreviations of TFs of interest (e.g., HOXA5, JUN, TEAD1, etc.) based on the JASPAR database; Altogether the list of TFs we examined comprised 579 items.

In addition, we added the UCP keyword to the search.

For each TF in the list, the script iteratively executes searches for each pair consisting of that TF and a keyword, e.g., HOXA5 and UCP1, HIF1A and UCP1 etc. These pairs are concatenated with a '+' character and passed to ESearch, e.g., *'HOXA5+UCP1'*. When searching for a complex of several TFs rather than a single TF, e.g., EWSR1 and FLI1, the constituent TFs in the complex were also combined with a '+' character in the search, e.g., *'EWSR1+FLI1+UCP1'*.

When searching for TF abbreviations that are identical or similar to a regular English word, in particular Ar (for which the database returns hits containing the verb form 'are'), CLOCK (English noun), REST (English noun and verb), MAX (short for maximum and a name) and JUN (very frequent abbreviation of the month June), we found that simply searching for these abbreviations returns a very large number of abstracts which contain the English word rather than the TF in question. Thus, we had to add a further search term to disambiguate the abbreviation and narrow down the search to abstracts which contain a reference to the TF that we were

interested in. We found that searching for *"AR androgen receptor", "CLOCK circadian", "REST RE1 silencing", "MAX MYC associated x",* and *"JUN proto-oncogene",* respectively, constrained the search effectively so that the results contained few irrelevant hits.

All combinations of TFs and the keyword were searched, i.e., in total 579 searches were executed. For each such combined TF + keyword search expression, the ESearch tool returns a list of PubMed IDs of articles the abstract of which contains both the TF and the keyword. For each TF, the number of abstracts containing the keyword is summarized in a table by the script. In addition, the full text of all abstracts is downloaded and saved along with the bibliographical data of the corresponding articles, as well as the URL of the full text of the article, where available, up to a pre-set maximum number, e.g., 50 abstracts.

Finally, the downloaded abstracts were manually reviewed. We considered the use of advanced natural language processing techniques, in particular large language models, to partially automate the review of the abstracts, but this was judged unnecessary for this specific set of results.

Identification of TF motif activity in different tissues and exploration of TF interaction partners using the ISMARA online platform

We obtained RNA-seq expression profiles from 16 human cell types using the Illumina Body Map 2 dataset (GEO accession GSE30611), accessible within the Integrated System for Motif Activity Response Analysis (ISMARA) online platform (ISMARA: https://ismara.unibas.ch/supp/dataset1_IBM_v2/ismara_report/).[55] We employed ISMARA to analyze the motif activity of putative species-specific and species-common homeodomain family TFs across different human tissues for the identification of TF motif activity. By leveraging the ISMARA platform, we investigated the motif activity patterns of these TFs, providing insights into their potential roles in tissue-specific gene regulation. Furthermore, we explored potential TFs and their first-level TF interaction partners using ISMARA to uncover regulatory networks and complex formations, such as enhanceosomes.

Ranking of the transcription factors

We ranked the 579 TFs from the EPD database using a scoring system designed to evaluate their role in UCP gene expression. Scores were calculated based on (i) the correlation coefficient between the expression of a given TF and UCP1 in human adipocytes (Table S4, data from our previous experiments, $n = 54$,[54]) and (ii) the frequency of the TF binding sequence/response element motif in the promoter regions analyzed. Based on the UCP gene promoters used for the binding motif frequency calculation, three independent scores were assigned to each TF: (i) the frequency of TF-RE motifs in the human UCP1 promoter, (ii) the frequency of TF-RE motifs in the human, mouse and rat UCP1 promoters (to estimate their conservation) (iii) and the prevalence of TF-RE in the promoters of UCP1, UCP2 and UCP3 in all analyzed species.

For TFs that were not expressed in the samples obtained from human neck adipocytes, we used a correlation coefficient of 0.1 instead of zero, as their expression in other tissue types or experimental conditions cannot be excluded. This adjustment improves the general applicability of our ranking system. In the case of heterodimers or trimers, we considered the correlation coefficients of the individual interacting proteins. The lower value was used unless it was equal to zero.

Following similar principles, TFs for which the Eukaryotic Promoter Database did not identify a binding sequence in the promoter under study were assigned a frequency value of 0.1. This recognizes the possibility that our current knowledge of response element base composition, plasticity, and cooperative binding ability is incomplete and that future discoveries may refine these predictions. The final TF ranking score was calculated by multiplying the absolute value of the correlation coefficient by the frequency data (Tables S6A–S6C).

Although the EPD database contains 579 transcription factor response elements (REs) derived from the JASPAR database, a total of 581 REs were included in our analysis. This discrepancy arises from the fact that EPD does not distinguish between THRA and THRB (thyroid hormone receptor subtypes) or between the individual factors STAT5A and STAT5B, as it only tests their dimerizsed form (STAT5A:STAT5B). In contrast, we treated THRA and THRB as well as STAT5A and STAT5B as separate transcription factors in our gene expression-based correlation analysis and therefore included them individually in the ranking.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of conserved motifs and TF-RE predictions was based on the frequency of occurrence of motifs. TF-REs were identified in the promoter regions of the human, mouse and rat *UCP1, UCP2* and *UCP3* genes using the Eukaryotic Promoter Database (EPD) interface, with significance thresholds of $p < 1e{-}4$ or $p < 1e{-}5$, depending on the analysis. EPD data for each TF were automatically collected using a custom PowerShell script and compiled by R in a matrix format for subsequent analysis.

The Venn package in R was used to compare TF-RE profiles across species and genes and to generate Venn diagrams showing the overlap of TFs across datasets. TF-RE frequencies were visualized using histograms (generated with the Hist package in R) and bar charts (in Excel), with positions normalized relative to the TSS.

Hierarchical clustering of gene expression profiles was performed with Morpheus (Morpheus: https://software.broadinstitute.org/morpheus) using Spearman rank similarity matrices and average linkage. Heatmaps were used to identify TFs and *UCP* homologs with similar expression patterns to *UCP1*.

The structural similarity of UCP proteins between species was assessed using structures predicted by AlphaFold2 and aligned with the RCSB PDB interface tools. Structural comparisons were assessed by pairwise structural alignments as described in the methodology section.

TF ranking was based on composite scores that integrated (i) the correlation coefficient between the TF and *UCP1* gene expression and (ii) the frequency of predicted TF REs in promoter regions across species and genes.

ADDITIONAL RESOURCES

No additional resources were used in this study.