Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi
Afyon Kocatepe University – Journal of Science and Engineering
https://dergipark.org.tr/tr/pub/akufemubid

Araştırma Makalesi / Research Article
e-ISSN: 2149-3367 DOI: https://doi.org/10.35414/akufemubid.1850855
CITE: Afyon Kocatepe University Journal of Science and Engineering. 26(4) (2026) 803-813

Comparative Bioinformatics Analysis of Transcriptomic Signatures and Regulatory Networks in BM-MSCs and UCB-MSCs: Molecular Insights for Regenerative Medicine

*Makale Bilgisi / Article Info
Alındı/Received: 28.12.2025
Kabul/Accepted: 28.06.2026
Yayımlandı/Published: 03.08.2026

Kİ-MKH'ler ve GKK-MKH'lerdeki Transkriptomik İmzaların ve Düzenleyici Ağların Karşılaştırmalı Biyoinformatik Analizi: Rejeneratif Tıp İçin Moleküler Çıkarımlar

Ahmet KARAKUŞ*

Bartın University, Faculty of Science, Department of Biotechnology, Bartın, Türkiye

Abstract

Mesenchymal stem cells (MSCs) play a crucial role in regenerative medicine due to their multifaceted potential and immunomodulatory effects. To investigate the molecular heterogeneity between human BM-MSCs and UCB-MSCs, the GSE6029 dataset was analyzed through an integrative bioinformatics pipeline, including differential expression, pathway enrichment, PPI networks, and transcription factor/miRNA regulatory mapping. The analysis identified 739 differentially expressed genes (DEGs); 374 of these were upregulated in UCB-MSCs and 365 in BM-MSCs. UCB-MSCs exhibited a distinct transcriptomic profile in terms of RNA processing, proteasome, and spliceosome pathways, controlled by key genes such as *MYC* and *HSPA4*. In contrast, BM-MSCs showed significant enrichment in PI3K-Akt signaling, ECM-receptor interactions, and constitutive development. Central to this enrichment were the *COL4A1* and *MET* genes. Post-transcriptional mapping identified miR-20b-5p as a key regulator in UCB-MSCs, while miR-15a-5p and miR-26a-5p were dominant in the BM-MSC regulatory network. These findings suggest that UCB-MSCs exhibit molecular phenotypes associated with rapid proliferation and immunomodulation, while BM-MSCs appear more aligned with structural integration and orthopedic repair. This study proposes a preliminary molecular framework that may inform MSC source selection, suggesting that source-specific molecular characteristics could be leveraged to refine therapeutic strategies according to pathological requirements.

Keywords: *Mesenchymal stem cells; Transcriptomics ;Bioinformatic analysis; Hub genes; miRNA-target network; Regenerative medicine*

Öz

Mezenkimal kök hücreler (MKH'ler), çok yönlü potansiyelleri ve immünomodülatör etkileri nedeniyle rejeneratif tıpta çok önemli bir role sahiptir. İnsan kemik iliği mezenkimal kök hücreleri (Kİ-MKH) ve göbek kordon kanı mezenkimal kök hücreleri (GKK-MKH) arasındaki moleküler heterojenliği araştırmak için, GSE6029 veri seti, diferansiyel ekspresyon, yol zenginleştirme, protein-protein etkileşim ağları ve transkripsiyon faktörü/miRNA düzenleyici haritalama dahil olmak üzere bütünleyici bir biyoinformatik işlem hattı aracılığıyla analiz edildi. Analiz sonucunda 739 farklı şekilde ifade edilen gen (DEG) belirlendi; bunlardan 374'ü GKK-MKH'lerde, 365'i ise Kİ-MKH'lerde yukarı doğru düzenlenmişti. GKK-MKH'ler, *MYC* ve *HSPA4* gibi anahtar genler tarafından kontrol edilen, RNA işleme, proteazom ve splaysom yolları açısından zengin bir transkriptomik profil sergiledi. Bunun aksine, Kİ-MKH'ler PI3K-Akt yolağında, ECM-reseptör etkileşimlerinde ve yapısal gelişimde önemli ölçüde zenginleşme gösterdi; bu zenginleşmenin merkezinde *COL4A1* ve *MET* genleri yer alıyordu. Transkripsiyon sonrası haritalama, GKK-MKH'lerde miR-20b-5p'yi kilit bir düzenleyici olarak tanımlarken, Kİ-MKH'lerde miR-15a-5p ve miR-26a-5p baskın konumdaydı. Bu bulgular, göbek kordonu kanı kaynaklı mezenkimal kök hücrelerin (GKK-MKH'ler) hızlı çoğalma ve immünomodülasyonla ilişkili moleküler fenotipler sergilediğini, kemik iliği kaynaklı mezenkimal kök hücrelerin (Kİ-MKH'ler) ise yapısal bütünleşme ve ortopedik onarımla daha uyumlu olduğunu göstermektedir.

Anahtar Kelimeler: *Mezenkimal kök hücre; Transkriptomik; Biyoinformatik analiz; Merkez genler; miRNA-hedef ağı; Rejeneratif tıp*

1. Introduction

Mesenchymal stem cells (MSCs) are attracting significant interest in regenerative medicine and immunotherapy due to their versatile differentiation capacities, immunomodulatory effects, and ability to secrete bioactive molecules that support tissue repair and regeneration (Zhang et al. 2022, Andrzejewska et al. 2019). These characteristics make MSCs highly attractive candidates for treating a wide range of diseases, including autoimmune disorders, cardiovascular diseases, neurodegenerative injuries, and traumatic injuries (Pittenger et al. 2019, Wang et al. 2014).

Among the various MSC sources, bone marrow (BM) and umbilical cord blood (UCB) are two distinct stem cell

*Sorumlu Yazar/Corresponding Author: Ahmet KARAKUŞ e-posta/e-mail: akarakus@bartin.edu.tr

sources that have been most extensively studied and clinically applied due to their accessibility and reliability (Jones and McGonagle 2008, Kuçi et al. 2019). Although they share basic stem cell characteristics, BM-MSCs and UCB-MSCs differ significantly in their biological behaviors such as proliferation rate, differentiation potential, paracrine activity, and immunomodulation, which directly impacts their therapeutic efficacy (Heo et al. 2016, Hou et al. 2021).

These source-specific differences are supported by distinct molecular signatures, including variations in gene expression profiles, transcription factor networks, and signaling pathways (Ding et al. 2015). For example, BM-MSCs typically exhibit strong osteogenic differentiation capacity, making them suitable for bone and cartilage repair applications (Wang et al. 2021). In contrast, UCB-MSCs exhibit enhanced angiogenic and immunosuppressive capabilities, particularly beneficial in the treatment of ischemic injuries and inflammatory diseases (Li et al. 2015). Therefore, understanding these molecular differences is critical for selecting the optimal MSC source tailored to specific clinical indications. Identification and validation of these molecular signatures are highly important for clinical applications. Precise characterization of MSC subpopulations at the transcriptomic level enables the development of biomarkers with predictable therapeutic outcomes, making more effective and personalized cell-based therapies possible (Nekanti et al. 2010).

In addition, elucidation of key regulatory genes and pathways provides novel targets for genetic or pharmacological manipulations to enhance MSC functionality and safety. With the emergence of high-throughput sequencing technologies and advanced bioinformatics approaches, in silico analyses have become indispensable for systematically investigating the molecular complexity and differences of MSCs from different tissue sources (Xu et al. 2024). This approach allows for comprehensive differential gene expression profiling, pathway enrichment, and network analysis, providing a holistic view of MSC biology.

This study addresses this gap through a comparative transcriptomic analysis of BM-MSCs and UCB-MSCs using an integrative bioinformatics approach. Specifically, the publicly available GSE6029 dataset was analyzed to identify differentially expressed genes, assess functional enrichment, construct protein interaction networks for hub gene discovery, and predict potential transcription factor and miRNA-target regulatory interactions unique to each MSC source. This analysis provides an understanding of the molecular profiles of BM-MSCs and UCB-MSCs, offering candidate insights that may inform future functional validation studies. To explore these molecular differences, comparative transcriptomic studies are essential. The analysis of publicly available datasets, such as GSE6029 (Markov et al. 2007), offers a valuable opportunity to identify underlying distinctions between these cell sources using contemporary bioinformatics tools.

2. Materials and Methods

2.1 Data Source and DEG Identification

The GSE6029 dataset, procured from the NCBI Gene Expression Omnibus (GEO) database, consists of microarray profiles of 9 human mesenchymal stem cell samples. Specifically, the dataset includes 3 samples of bone marrow-derived mesenchymal stem cells (BM-MSCs) and 6 samples of umbilical cord blood-derived mesenchymal stem cells (UCB-MSCs). According to the original study (Markov et al. 2007) and GEO metadata, all cell lines were established from healthy donors. Specific donor demographics such as age, sex, culture passage numbers, or detailed isolation protocols were unavailable in the dataset metadata and could not be factored into the analysis. The dataset was generated using the Affymetrix Human Genome U133A Array platform. Differentially expressed genes (DEGs) between BM-MSC and UCB-MSC groups were identified via the GEO2R web tool (Barrett et al. 2013), which employs the limma package (Ritchie et al. 2015) to statistically compare gene expression levels between predefined groups. DEGs were selected based on a cutoff of an absolute log2 fold change (log_2FC) > 0.8 (upregulated in BM-MSCs) or log_2FC < –0.8 (upregulated in UCB-MSCs) and a nominal p-value < 0.05. It is important to note that the GEO2R web interface was used without applying multiple testing correction; therefore, nominal p-values were utilized for DEG identification. Based on these thresholds, 365 genes were upregulated in BM-MSCs and 374 genes in UCB-MSCs.

2.2 Gene Ontology and Pathway Enrichment Analysis

To investigate the functional roles of DEGs, Gene Ontology (GO) enrichment was conducted across Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) categories using ShinyGO web tool (http://shinygo.jimmy.org) (Ge et al. 2020). In parallel, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed to identify significantly enriched signaling pathways. All terms and pathways with p-value < 0.05 were considered statistically significant.

2.3 Protein-Protein Interaction (PPI) Network Construction

To investigate protein-level interactions, DEGs were input into the STRING database (version 11.5) (Szklarczyk et al. 2021). Networks were constructed separately for upregulated genes in BM-MSCs and UCB-MSCs using a minimum required interaction score of 0.4 (medium confidence). The resulting networks were imported into Cytoscape software (version 3.8.2) for visualization and analysis.

2.4 Identification of Hub Genes

cytoHubba, a Cytoscape plugin (Chin et al. 2014) was used to identify hub genes based on network topology. The Degree algorithm was applied to rank nodes by connectivity, and the top 10 genes with the highest degrees in each group were designated as hub genes.

2.5 Prediction of Transcription Factors (TFs)

To predict transcriptional regulators of hub genes, the ChEA3 (ChIP-X Enrichment Analysis 3) platform was used (Keenan et al. 2019). Analyses were performed separately for BM-MSC and UCB-MSC-upregulated hub genes, and the top 10 transcription factors were selected based on mean rank scores across integrated reference databases. Network visualizations were generated in Cytoscape.

2.6 Identification of miRNA-Hub Gene Interactions

Hub genes were further analyzed for microRNA (miRNA) regulation using miRNet 2.0 (Chang et al. 2022), an integrated platform that combines multiple miRNA–target interaction databases. The hub gene lists were uploaded to miRNet, and human-specific interactions were extracted using default settings. Networks were visualized within miRNet, and key regulatory miRNAs were identified based on node degree.

3. Results and Discussion

3.1 Differentially Expressed Genes (DEGs)

A total of 739 DEGs were identified between UCB-MSCs and BM-MSCs, using thresholds of $\log_2FC > 0.8$ and $p < 0.05$. Of these, 365 genes were upregulated in BM-MSCs ($\log_2FC > 0.8$) and 374 genes were upregulated in UCB-MSCs ($\log_2FC < -0.8$). A volcano plot illustrates the distribution, significance, and magnitude of expression changes (Figure 1).

3.2 Functional Enrichment Analysis

To uncover functional differences between UCB-MSCs and BM-MSCs, DEGs were analyzed for enrichment in Gene Ontology (GO) terms including Biological Process (BP), Cellular Component (CC), Molecular Function (MF), as well as KEGG pathways.

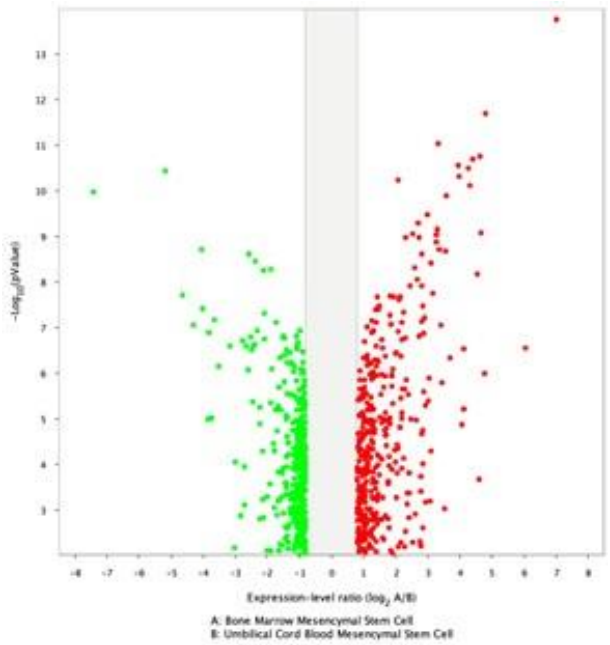

Figure 1. Volcano plot of differentially expressed genes between BM-MSCs and UCB-MSCs. Red dots represent genes significantly upregulated in BM-MSCs ($\log_2FC > 0.8$), green dots indicate genes upregulated in UCB-MSCs ($\log_2FC < -0.8$).

Downregulated genes in BM-MSCs (i.e., upregulated in UCB-MSCs; n=374) were enriched in GO Biological Processes such as cellular component biogenesis, protein-containing complex assembly, RNA processing, and mRNA splicing, suggesting heightened RNA-related biosynthetic activity in UCB-MSCs (Figure 2A). GO Cellular Component (GO-CC) analysis indicated localization in the nuclear lumen, nucleoplasm, catalytic complex, extracellular vesicles, and extracellular organelles (Figure 2B). GO Molecular Function (GO-MF) terms included nucleic acid binding, RNA binding, nucleotide binding, and ATP binding (Figure 2C).

Conversely, genes upregulated in BM-MSCs (n=365) were significantly enriched in GO- Biological Process terms such as signaling receptor binding, structural molecule activity, cell adhesion molecule binding system development, animal organ development, cellular development, and tissue development, indicating roles in morphogenesis and structural organization (Figure 2D). GO-CC terms included the extracellular region, vesicles, extracellular space, extracellular organelles, and cell junctions, reflecting features of extracellular matrix organization and intercellular communication (Figure 2E). GO-MF analysis highlighted metal ion binding, signal receptor binding, structural molecular activity, and cell adhesion molecule binding, consistent with signaling and adhesive properties of BM-MSCs (Figure 2F).

Upregulated genes in UCB-MSC, KEGG pathway analysis for these genes revealed enrichment in Huntington's disease, Alzheimer's disease, spliceosome, and

proteasome pathways (Table 1), indicating involvement in neurodegeneration-associated and proteolytic mechanisms. Upregulated genes in BM-MSC, KEGG pathway analysis showed enrichment in pathways in cancer, PI3K-Akt signaling pathway, and ECM–receptor interaction, underscoring survival, proliferation, and ECM communication mechanisms (Table 2).

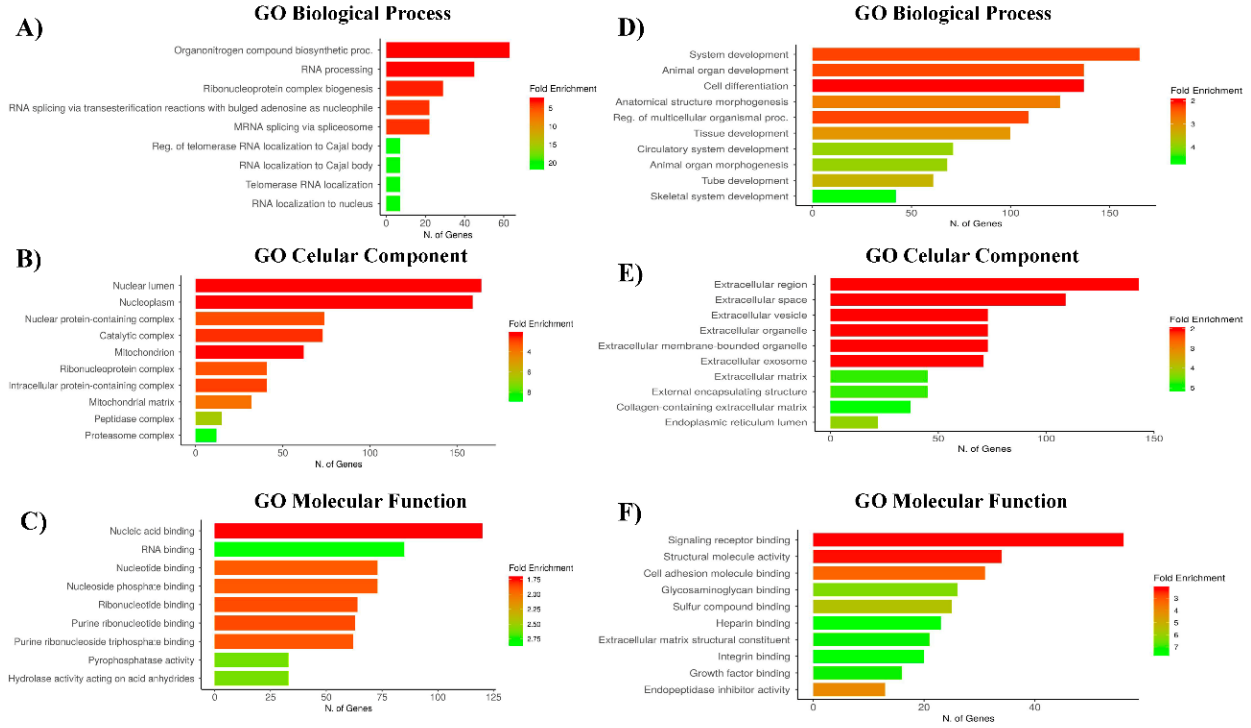

Figure 2. GO enrichment analysis of DEGs. (A–C) Enriched GO terms for genes upregulated in UCB-MSCs (downregulated in BM-MSCs): (A) Biological Process (BP), (B) Cellular Component (CC), (C) Molecular Function (MF). (D–F) Enriched GO terms for genes upregulated in BM-MSCs: (D) BP, (E) CC, (F) MF. Top 10 significantly enriched terms shown for each category (p-value < 0.05).

Table 1. KEGG Pathway Enrichment Analysis of Genes Upregulated in UCB-MSCs.

Pathway	Genes	p-Value
Metabolic pathways	*DHCR24, ADPGK, APIP, ATP5F1A, ATP5F1D, ATP5V0D1, NDUFS8, NDUFV1, NDUFA10, NDUFS5, NME1, NME4, ACAT1, ACO1, ALDH2, AKR1A1, AKR1B1, MGAT2, BLVRB, CANT1, CAMKMT, CDO1, DCTPP1, DGKZ, FUT8, FAH, GLA, GLUD1, GOT2, QRSL1, GCSH, GMPS, HMBS, ISYNA1, LIAS, MDH2, MOCOS, MTMR14, MTMR4, NEU1, ODC1, PRDX6, PISD, PTDSS1, PGD, PLOD3, RXLYT1, SMPD4, SYNJ2, UCK2*	3.2E-2
Spliceosome	*DDX23, LSM4, LSM7, U2AF1, ACIN1, NCBP1, PUF60, PQBP1, PRPF38B, PRPF4, SNRPD1, SNRNP40, SNRPA1, SF3A3, SF3B4, SF3B5*	2.5E-4
Huntington disease	*ATP5F1A, ATP5F1D, NDUFS8, NDUFV1, NDUFA10, NDUFS5, POLR2J3, TBPL1, CREB3L2, DNALI1, PSMA3, PSMB2, PSMB4, PSMD2, PSMC5, PSMD11, PSMD14, PSMD3, TUBB3, TUBB*	6.3E-5
Alzheimer disease	*ATP5F1A, ATP5F1D, BID, HRAS, NDUFS8, NDUFV1, NDUFA10, NDUFS5, APH1A, CAPN1, PSMA3, PSMB2, PSMB4, PSMD2, PSMC5, PSMD11, PSMD14, PSMD3, SLC11A2, TUBB3, TUBB*	5.1E-4
Proteasome	*PSMA3, PSMA5, PSMB10, PSMB2, PSMD2, PSMC5, PSMD11, PSMD14, PSMD3, PSME2, PSME4*	8.3E-9

Table 2. KEGG Pathway Enrichment Analysis of Genes Upregulated in BM-MSCs.

Pathway	Genes	p-Value
PI3K-Akt signaling pathway	*MET, TEK, ANGPT1, COMP, COL4A1, FGF7, FGFR2, IGF2, ITGA6, ITGA7, ITGB5, LAMA4, NGF, THBS2, THBS3*	4.7E-2
Proteoglycans in cancer	*FAS, MET, TIAM1, TIAMP3, EZR, FZD1, HSPB2, HIF1A, ITPR3, IGF2, ITGB5, LUM, PLCE1, TWIST1*	1.0E-3
Rap1 signaling pathway	*MET, TEK, TIAM1, ADCY7, ANGPT1, FGF7, FGFR2, ITGB2, NGF, PLCB4, PLCE2, PRKD1*	1.1E-2
TGF-beta signaling pathway	*SKIL, SMAD6, SMAD7, ACVR2A, FBN1, GREM2, INHBB, ID2, ID3, ID4, PITX2, TGFBR1*	4.4E-5
ECM-receptor interaction	*CD36, COMP, COL4A1, ITGA6, ITGA7, ITGB5, LAMA4, THBS2, THBS3*	1.1E-3

3.3 Protein-Protein Interaction (PPI) Network Analysis and Hub Genes

PPI networks for gene sets upregulated by both cells were constructed using STRING. The downregulated PPI network (genes upregulated in UCB-MSCs) included 374 nodes and 1681 edges, with an average node degree of 8.99 and a highly significant enrichment (PPI p-value = 1.0E-16). Hub genes such as *HSPA4, HSPA9, ATP5F1A, TUFM, NHP2, RUVBL1, RUVBL2, MYC, MRPS7*, and *NCBP* were identified and highlighted (Figure 3A). The upregulated PPI network (genes upregulated in BM-MSCs) comprised 365 nodes and 771 edges, average node degree of 4.22, also significantly enriched (PPI p-value < 1.0E-16). Hub genes included *BGN, COL5A1, FBN1, LUM, FGF7, MET, HIF1A, TIMP3, THBS2,* and *COL4A1* (Figure 3B). STRING networks in Cytoscape were simplified to reduce network complexity and improve clarity.

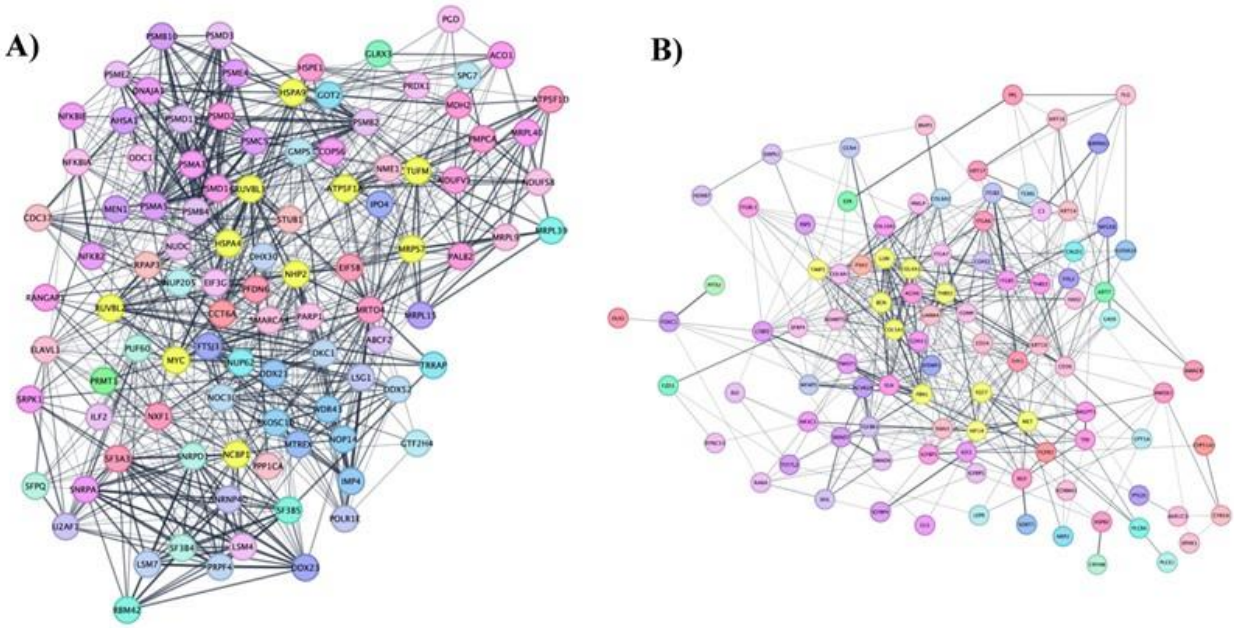

Figure 3. Protein–protein interaction (PPI) networks of DEGs. (A) PPI network of genes upregulated in UCB-MSCs. (B) PPI network of genes upregulated in BM-MSCs. Networks constructed using STRING database and visualized in Cytoscape.

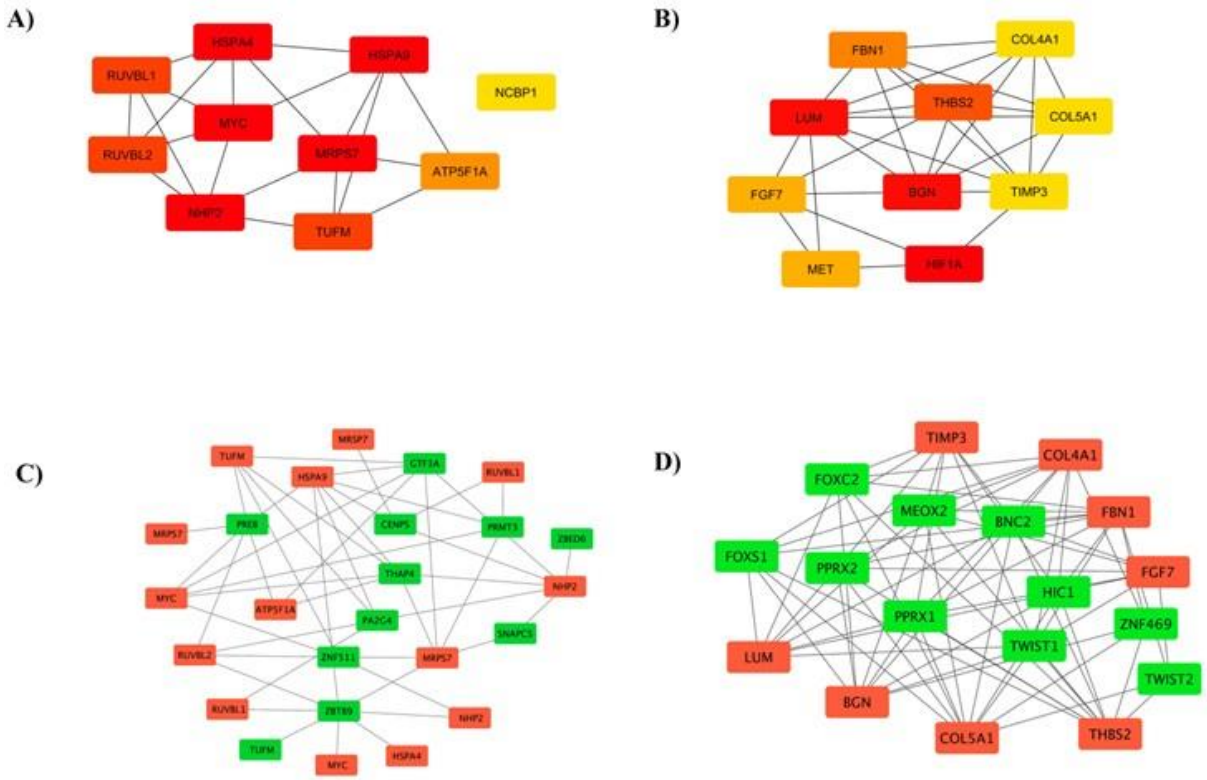

Figure 4. (A) Subnetwork of hub genes upregulated in UCB-MSCs. (B) Subnetwork of hub genes upregulated in BM-MSCs. (C) Top 10 TFs associated with UCB-MSC-upregulated hub genes. (D) Top 10 TFs regulating BM-MSC-upregulated hub genes.

In both networks, hub genes showed dense interconnectivity, forming hub-to-hub subnetworks (Figure 4A and 4B). For UCB-MSC-upregulated hubs (Figure 4A), roles in mitochondrial function and RNA processing are suggested. Node color gradients from red (most central) to yellow represent interaction strength. Notably, *NCBP* appeared disconnected from the other hubs in the UCB-MSC-upregulated network (Figure 4A), implying functional independence. For BM-MSC-upregulated hubs (Figure 4B), the observed dense connectivity implies coordinated roles in extracellular matrix (ECM) remodeling, cell adhesion, and tissue organization.

3.4 Transcription Factor (TF) Prediction

ChEA3-based TF prediction identified upstream regulators of the hub genes. For UCB-MSC-upregulated hub genes, predicted TFs included PA2G4, ZBTB9, PREB, THAP4, PRMT3, SNAPC5, ZNF511, ZBED6, GTF3A, and CENPS mostly involved in RNA processing and transcription regulation (Figure 4C). For BM-MSC-upregulated hub genes, TFs such as PRRX1, FOXS1, ZNF469, PRRX2, FOXC2, TWIST2, BNC2, HIC1, TWIST1, and MEOX2 were identified, many of which regulate development, EMT, and ECM organization (Figure 4D). Networks were visualized in Cytoscape with hub genes in red and TFs in green. Furthermore, the functional roles and regulatory effects of top transcription factors associated with hub genes upregulated in UCB-MSCs and BM-MSCs are presented in Table 3 and Table 4, respectively.

Table 3. Functional Roles and Regulatory Effects of Top Transcription Factors Associated with Hub Genes Upregulated in UCB-MSCs.

Transcription Factor (TF)	Target Hub Genes	Functional Annotation / Biological Role	Regulator Effect of TF
PA2G4	*HSPA9, RUVBL2, RUVBL1, NHP2, TUFM*	Transcriptional repressor involved in cancer progression and gene silencing	Repressor
ZBTB9	*HSPA9, HSPA4, MYC, RUVBL2, RUVBL1, NHP2, MRPS7, TUFM*	Regulates cell motility, cytoskeleton organization, and differentiation	Activator / Repressor
PREB	*HSPA9, ATP5F1A, RUVBL2, MYC, MRPS7, TUFM*	Modulates inflammation and cell survival	Activator
THAP4	*HSPA9, ATP5F1A, RUVBL2, MYC, NHP2, MRPS7, TUFM*	Regulates metabolism, lipid homeostasis, and DNA damage response	Repression
PRMT3	*HSPA9, MYC, RUVBL1, MRPS7*	Critical for inflammation and cell differentiation	Activator
SNAPC5	*NHP2, MRPS7*	Functions as transcriptional activator and repressor; chromatin remodeling	Activator / Repressor
ZNF511	*RUVBL2, MYC, NHP2, MRPS7, TUFM*	Oncogenic TF regulating cell cycle, apoptosis, and metabolism	Activator
ZBED6	*NHP2*	Stress-inducible TF modulating apoptosis and cell adaptation	Repression
GTF3A	*HSPA9, ATP5F1A, MYC, NHP2, MRPS7, TUFM*	Transactivator of MHC class II genes; immune system function	Activator
CENPS	*HSPA9, HSPA4, RUVBL1, NHP2, MRPS7*	Forms heterodimers with MYC; regulates growth and differentiation genes	Activator

Table 4. Functional Roles and Regulatory Effects of Top Transcription Factors Associated with Hub Genes Upregulated in BM-MSCs.

Transcription Factor (TF)	Target Hub Genes	Functional Annotation / Biological Role	Regulator Effect of TF
PRRX1	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Regulation of myogenic differentiation and extracellular matrix organization	Activator
FOXS1	*COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Epigenetic regulation of EMT and transcriptional repression	Repressor
ZNF469	*COL5A1, COL4A1, BGN, THBS2, FBN1*	Control of tissue morphogenesis and ECM-related pathways	Activator
PRRX2	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Modulation of cell proliferation and developmental signaling cascades	Activator
FOXC2	*COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Maintenance of genome stability and regulation of transcriptional activity	Activator
TWIST2	*FGF7, COL5A1, THBS2, FBN1*	Regulation of mesenchymal development and lineage specification	Activator
BNC2	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Chromatin organization and genome architecture modulation	Repressor
HIC1	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Transcriptional repression of neuronal genes and regulation of differentiation	Repressor
TWIST1	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	EMT regulation and RNA Pol II-mediated gene expression	Activator
MEOX2	*FGF7, COL5A1, LUM, COL4A1, BGN, TIMP3, THBS2, FBN1*	Transcriptional regulation and DNA replication via cohesin complex	Activator

3.5 miRNA-Hub Gene Interaction Analysis

To investigate post-transcriptional regulation, miRNA–hub gene interactions were analyzed using miRNet 2.0. Separate networks were constructed for hub genes upregulated in UCB-MSCs and BM-MSCs. As indicated in Table 5, the top miRNAs were identified based on degree (number of target hub genes) and betweenness centrality (regulatory influence within the network).

In the upregulated UCB-MSC hub genes, hsa-miR-20b-5p, hsa-miR-423-5p, hsa-let-7a-5p, hsa-miR-34a-5p, and hsa-miR-196a-5p showed the highest degree values (9), indicating strong regulatory potential. Members of the let-7 family (let-7a/b/c/e/i-5p) were also prominent, suggesting roles in differentiation and extracellular matrix (ECM) regulation (Figure 5A).

For upregulated BM-MSC hub genes, hsa-miR-15a-5p and hsa-miR-26a-5p were the top miRNAs, both with high degree (9) and the highest betweenness scores (5.53). Other key regulators included let-7a/b/c/e/f-5p, hsa-miR-17-5p, and hsa-miR-20a-5p, which are known to regulate RNA processing, apoptosis, and stem cell activity (Figure 5B).

Table 5. Top miRNAs Interacting with Upregulated Hub Genes in UCB-MSCs and BM-MSCs.

Upregulated UCB-MSC miRNA and Hub Genes Interaction			Upregulated BM-MSC miRNA and Hub Genes Interaction		
miRNA	**Degree**	**Betweenness**	**miRNA**	**Degree**	**Betweenness**
hsa-miR-20b-5p	9	1.950646	hsa-miR-15a-5p	9	5.528808
hsa-miR-423-5p	9	1.950646	hsa-miR-26a-5p	9	5.528808
hsa-let-7a-5p	9	1.950646	hsa-let-7a-5p	8	1.716699
hsa-let-7b-5p	8	1.649777	hsa-let-7b-5p	8	1.716699
hsa-let-7c-5p	9	1.950646	hsa-let-7c-5p	8	1.716699
hsa-miR-196a-5p	9	1.950646	hsa-let-7e-5p	8	1.716699
hsa-let-7i-5p	9	1.950646	hsa-let-7f-5p	8	1.716699
hsa-miR-34a-5p	9	1.950646	hsa-miR-17-5p	8	1.716699
hsa-let-7e-5p	8	1.638378	hsa-miR-19a-3p	8	4.36873
hsa-miR-15a-5p	8	1.649777	hsa-miR-20a-5p	8	4.36873

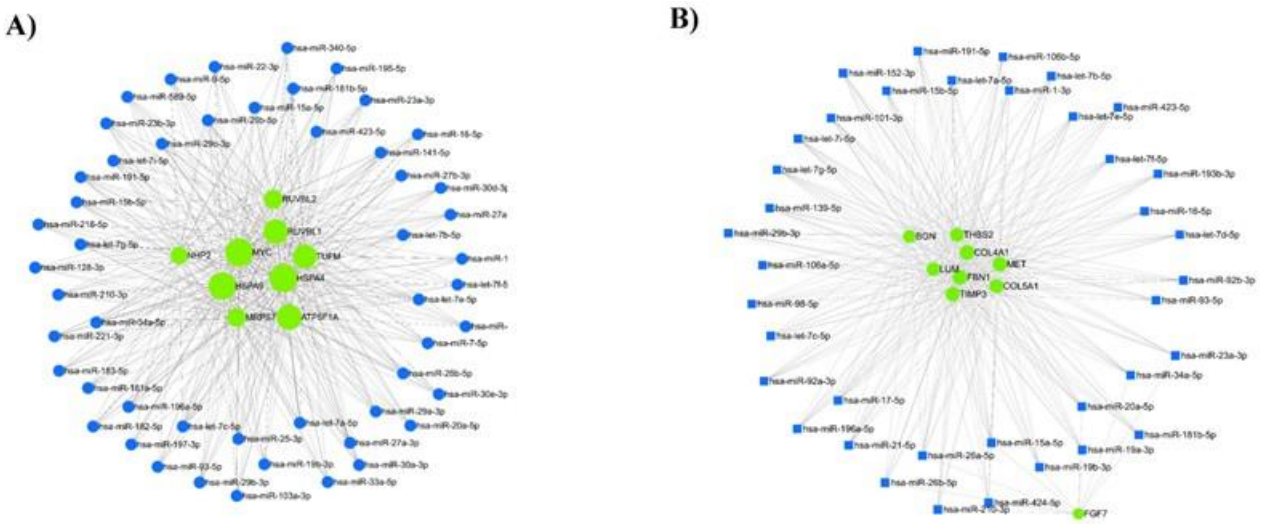

Figure 5. miRNA–hub gene interaction networks. (A) Predicted miRNAs targeting hub genes upregulated in UCB-MSCs. (B) Predicted miRNAs targeting hub genes upregulated in BM-MSCs. Networks were generated using miRNet platform.

3.6 *Comparison with Original Dataset Study*

To provide a direct and systematic cross-validation of the present reanalysis against the foundational study by Markov et al. (2007), Table 6 summarizes the key findings reported in the original study alongside the results recovered under the current analytical pipeline. This comparison is intended to clarify which findings are reproduced, which are newly identified, and where discrepancies emerge along with their probable methodological explanations.

As shown in Table 6, several key biological themes originally identified by Markov et al. (2007) are directionally consistent with the present findings. The superior proliferative and adipogenic capacity of UCB-MSCs, originally attributed to *HGF, SDF1/CXCL12,* and *PPARG* expression, is now understood at a systems level: *MYC* and *HSPA4* emerge as central hub genes coordinating RNA processing, proteasome, and spliceosome pathway enrichment, suggesting a broader transcriptional program underpinning this phenotype. For

BM-MSCs, the chondrogenic signature originally linked to *COMP, AGC1,* and *COL10A1* is partially recovered — *COMP* is represented within the PI3K-Akt signaling gene set (Table 2), and *COL4A1 and MET* emerge as dominant hub genes in ECM-receptor and structural pathways. The absence of *AGC1* and *COL10A1* among the top-ranked hub genes under the current thresholds likely reflects differences in the analytical strategy: the present pipeline applies a log2FC > 0.8 cutoff combined with nominal p-values and network topology-based hub ranking, which prioritizes highly connected genes rather than markers of highest fold-change.

Table 6. Structured comparison of findings between Markov et al. (2007) and the present re-analysis of the GSE6029 dataset

Finding Category	Markov et al. (2007) Original Findings	Present Re-analysis (GSE6029, Current Thresholds)
UCB-MSC proliferation markers	High expression of *HGF, SDF1/CXCL12, PPARG*; high growth rates and adipogenic potential	Consistent: *MYC* and *HSPA4* identified as hub genes; systemic enrichment of RNA processing, proteasome, and spliceosome pathways corroborates superior proliferative capacity. *HGF/SDF1/PPARG* not among top hub genes under current log2FC > 0.8 threshold, though directionally consistent.
BM-MSC chondrogenic markers	*COMP, AGC1, COL10A1* associated with chondrogenic superiority	Partially consistent: *COL4A1* and *MET* are top hub genes in BM-MSCs; ECM-receptor interaction and PI3K-Akt pathways enriched. *COMP* is present in PI3K-Akt gene list (Table 2). *AGC1* and *COL10A1* not recovered as hub genes under current thresholds — likely due to differences in analytical pipeline and FDR criteria.
Regulatory networks	Not investigated at systems level; marker-level gene expression reported	Novel contribution: TF networks (PRRX1, FOXS1, TWIST1/2 for BM-MSC; PA2G4, ZBTB9 for UCB-MSC) and miRNA regulatory layers (miR-20b-5p in UCB-MSC; miR-15a-5p, miR-26a-5p in BM-MSC) identified — not present in original study.
Pathway analysis	Not performed; individual gene markers characterized	Novel contribution: KEGG pathway enrichment reveals spliceosome, proteasome, Huntington/Alzheimer disease pathways in UCB-MSCs; PI3K-Akt, ECM-receptor interaction, TGF-beta pathways in BM-MSCs.
Discrepancy note	-	Discrepancies arise from: (1) updated bioinformatics pipeline (limma/GEO2R vs. original methods); (2) log2FC threshold of 0.8 applied here vs. unconstrained in original; (3) nominal p-value used without FDR correction, which may inflate DEG count; (4) systems-level network topology prioritizes connectivity over single-gene significance.

Importantly, the present study extends the original findings by mapping post-transcriptional and epigenetic regulatory layers — specifically, transcription factor networks and miRNA–hub gene interactions — that were not investigated by Markov et al. (2007). These constitute a genuinely novel contribution, rather than a mere replication, of the original dataset analysis.

4. Discussion

The transcriptomic divergence between UCB-MSCs and BM-MSCs observed in this study is inherently linked to the developmental and microenvironmental contexts of their respective niches (Hass et al. 2011). The results indicate that despite sharing fundamental stem cell properties, these cell types exhibit distinct molecular signatures that may reflect functional specialization (Dominici et al. 2006).

A key finding of this study is the enrichment of RNA metabolism and proteostasis pathways in UCB-MSCs, suggesting a heightened state of transcriptional plasticity (Batsali et al. 2019). The identification of MYC as a central hub gene in UCB-MSCs is noteworthy, as it may underpin the superior expansion rates often reported for neonatal sources (Wilson et al. 2008, Nagamura-Inoue 2014). Furthermore, the upregulation of heat shock proteins like HSPA4 suggests a robust stress adaptation mechanism, which is critical for cell survival in inflammatory environments (Yuan et al. 2025).

In contrast, BM-MSCs exhibited a program heavily oriented toward structural development. The enrichment of the PI3K-Akt signaling pathway and hub genes such as *COL4A1* is consistent with the well-established role of BM-MSCs in orthopedic regeneration (Infante and Rodriguez 2018). Post-transcriptional regulation via miRNAs like hsa-miR-26a-5p is consistent with the involvement of BM-MSCs in structural repair (Han and Fan 2021).

The findings of this study provide a critical molecular framework that may inform the strategic selection of MSC sources in clinical applications. The distinct transcriptomic profiles identified here suggest that the therapeutic use of MSCs should not be viewed as a 'one-size-fits-all' approach; instead, the source may be deliberately matched to the clinical objective (Han et al. 2019). For instance, the metabolic and RNA-processing signature of

UCB-MSCs positions them as a candidate source for conditions requiring rapid cellular response, such as systemic inflammatory diseases (Kuçi et al. 2019). Conversely, the structural and ECM-oriented signature of BM-MSCs suggests potential utility in reconstructive therapies and bone tissue engineering (Kostadinova et al. 2025). By delineating these source-specific programs, this study emphasizes that achieving optimal clinical efficacy depends on aligning the biological strengths of the MSC source with the pathological demands of the patient.

Building on the marker-level distinctions originally characterized by Markov et al. (2007) — UCB-MSCs linked to growth factors such as *HGF, SDF1/CXCL12,* and *PPARG,* and BM-MSCs to chondrogenic markers including COMP, *AGC1*, and *COL10A1* — the present re-analysis reframes these source-specific differences at the systems level. The proliferative capacity of UCB-MSCs appears associated with systemic activation of RNA processing, proteasome, and spliceosome pathways centered on hub genes such as *MYC* and *HSPA4*, whereas the orthopedic potential of BM-MSCs is reflected in enrichment of PI3K-Akt signaling and ECM-receptor interactions converging on *COL4A1* and *MET*.

The present study further extends this framework by mapping transcription-factor and miRNA-based regulatory layers. Notably, miR-20b-5p emerges as a candidate regulator within the proliferative profile of UCB-MSCs, while miR-15a-5p and miR-26a-5p appear linked to the structural-repair features of BM-MSCs.

Taken together, the present analysis expands upon the original marker-level characterization by Markov et al. (2007) by elevating isolated gene-level observations to a

systems-biology perspective. BM-MSCs and UCB-MSCs exhibit distinct transcriptomic patterns and candidate regulatory features that could inform future studies on source-specific applications. However, it must be emphasized that, as an entirely in silico study relying on a single microarray dataset with a limited sample size, the present findings are exploratory and hypothesis-generating in nature. The identification of hub genes, transcription factor networks, and miRNA regulatory interactions represents a foundation for future experimental investigations rather than a basis for definitive clinical recommendations. Rigorous validation — including independent cohort replication, qRT-PCR, Western blot, and functional gene knockdown/overexpression studies — is required before these transcriptomic distinctions can be translated into actionable clinical guidelines for MSC source selection.

The GSE6029 dataset was selected due to its unique comparative design encompassing both BM-MSCs and UCB-MSCs from healthy donors, despite limitations inherent to 2007 microarray platforms. This provided a homogeneous and well-characterized platform crucial for directly comparing the transcriptomic profiles of these two distinct MSC types at that time. However, microarray technology inherently offers lower resolution and a more limited dynamic range compared to contemporary RNA-sequencing (RNA-seq) or single-cell RNA sequencing (scRNA-seq) approaches. Future investigations employing advanced methodologies, especially scRNA-seq, are warranted to resolve MSC heterogeneity at single-cell resolution, thereby complementing these macro-level findings.

In addition, the differential gene expression analysis conducted via the GEO2R web tool relied on nominal p-values without multiple testing corrections, such as False Discovery Rate (FDR) adjustment. While GEO2R allows for rapid and accessible analysis, multiple testing correction was not applied in the present pipeline. Therefore, the presented DEG lists and subsequent downstream analyses (e.g., hub genes, regulatory networks) should be interpreted with caution, recognizing the potential impact of an unadjusted p-value on the false-positive rate. Future studies, ideally incorporating robust statistical pipelines with FDR corrections, are essential to validate and refine these specific findings.

Several limitations of this study warrant consideration. First, the analysis is based on a single publicly available dataset (GSE6029) with a limited sample size (n = 9), which may not fully capture the biological variance inherent to MSC populations (Shakya et al. 2019). Second, the GEO2R-based differential expression analysis relied on nominal p-values, as multiple testing correction was not applied through this interface; this increases the potential for false-positive findings, and all downstream results should be interpreted with this caveat in mind. Third, the predictions of PPI networks, transcription factor binding, and miRNA–target interactions are entirely computational and derived from curated databases; they do not constitute experimental evidence of functional interactions. Fourth, donor-specific variables such as age and sex, which can influence the MSC transcriptome, were not available within the dataset metadata (Rojewski and Weber 2008). Future studies employing RNA-sequencing or single-cell RNA-sequencing approaches with larger, well-characterized cohorts and rigorous statistical frameworks will be essential to validate and extend the candidate regulators and pathways proposed here.

5. Conclusion

In conclusion, this integrative bioinformatics analysis of the GSE6029 dataset provides insights into the distinct transcriptomic profiles of BM-MSCs and UCB-MSCs.

Specifically, profiles of differentially expressed genes, enriched biological pathways, potential hub genes, and regulatory networks were mapped to underscore the molecular heterogeneity between these two MSC sources.

While these findings indicate that BM-MSCs exhibited transcriptomic features associated with bone regeneration, and UCB-MSCs in immunomodulation, certain limitations must be noted, including the inherent constraints of microarray platforms, small sample sizes, and reliance on nominal p-values.

These computational predictions warrant rigorous experimental validation and further investigation using contemporary high-throughput sequencing technologies and larger cohorts. Nevertheless, this study proposes a molecular framework that can contribute to the informed selection of MSC sources for specific regenerative medicine applications and guides future research directions into their therapeutic potential.

Declaration of Ethical Standards

The authors declare that they comply with all ethical standards.

Credit Authorship Contribution Statement

Author-1: Conceptualization, investigation, methodology and software, visualization and writing – original draft.

Declaration of Competing Interest

The authors have no conflicts of interest to declare regarding the content of this article.

Data Availability Statement

All data generated or analyzed during this study are included in this published article.

References

Andrzejewska, A., Lukomska, B. and Janowski, M., 2019. Concise review: Mesenchymal stem cells: from roots to boost. *Stem Cells*, 37(7), 855-864.

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S. & Soboleva, A., 2013. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Research*, 41(D1), D991-D995.

Batsali, A. K., Kastrinaki, M. C., Papadaki, H. A. and Pontikoglou, C., 2019. Mesenchymal stem cells derived from Wharton's Jelly of the umbilical cord: biological properties and emerging clinical applications. *Current Stem Cell Research & Therapy*, 14, 58-71.

Chang, L., Zhou, G., Soufan, O. and Xia, J., 2022. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Research*, 50, W217-W227.

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T. and Lin, C. Y., 2014. CytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Systems Biology*, 8, S11.

Ding, D. C., Chang, Y. H., Shyu, W. C. and Lin, S. Z., 2015. Human umbilical cord mesenchymal stem cells: a new era for stem cell therapy. *Cell Transplantation*, 24(3), 339-347.

Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., Deans, R., Keating, A., Prockop, Dj. and Horwitz, E., 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, 8(4), 315-317.

Ge, S. X., Jung, D. and Yao, R., 2020. ShinyGO: a graphical enrichment tool for animals and plants. *Bioinformatics*, 36, 2628-2629.

Han, X. and Fan, Z., 2021. MicroRNAs regulation in osteogenic differentiation of mesenchymal stem cells. *Frontiers in Dental Medicine*, 2, 747068.

Han, Y., Li, X., Zhang, Y., Han, Y., Chang, F. and Ding, J., 2019. Mesenchymal Stem Cells for Regenerative Medicine. *Cells*, 8(8), 886.

Hass, R., Kasper, C., Böhm, S. and Jacobs, R., 2011. Different populations and sources of human mesenchymal stem cells (MSC): A comparison of adult and neonatal tissue-derived MSC. *Cell Community Signaling*, 9, 12.

Heo, J. S., Choi, Y., Kim, H. O., 2016. Comparison of molecular profiles of human mesenchymal stem cells derived from bone marrow, umbilical cord blood, placenta, and adipose tissue. International Journal of *Molecular Medicine*, 37, 115-125.

Hou, W., Duan, L., Huang, C., Li, X., Xu, X., Qin, P., Hong, N., Wang, D. and Jin, W., 2021. Cross-tissue characterization of heterogeneities of mesenchymal stem cells and their differentiation potentials. *Frontiers in Cell and Developmental Biology*, 9, 781021.

Infante, A. and Rodriguez, C. I., 2018. Osteogenesis and aging: lessons from mesenchymal stem cells. *Stem Cell Research & Therapy*, 9(1), 244.

Jones, E. and McGonagle, D., 2008. Human bone marrow mesenchymal stem cells in vivo. *Rheumatology*, 47, 126-131.

Keenan, A. B., Torre, D., Lachmann, A., Leong, J., Wojciechowicz, K. and Ma'ayan, A., 2019. ChEA3:

transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Research*, 47, W212-W224.

Kostadinova, M., Raykovska, M., Simeonov, R., Lolov, S. and Mourdjeva, M., 2025. Recent advances in bone tissue engineering: enhancing the potential of mesenchymal stem cells for regenerative therapies. *Current Issues in Molecular Biology*, 47(4), 287.

Kuçi, S., Kuçi, Z., Schäfer, R., Spohn, G., Winter, S., Schwab, M., Salzmann-Manrique, E., Klingebiel, T. and Bader, P., 2019. Molecular signature of human bone marrow-derived mesenchymal stromal cell subsets. *Scientific Reports*, 9(1), 1774.

Li, C. Y., Wu, X. Y., Tong, J. B., Yang, X. X., Zhao, J. L., Zheng, Q. F., Zhao, G. B. and Ma, Z. J., 2015. Comparative analysis of human mesenchymal stem cells from bone marrow and adipose tissue under xeno-free conditions for cell therapy. *Stem Cell Research & Therapy*, 6(1), 55.

Markov, V., Kusumi, K., Tadesse, M. G., William, D. A., Hall, D. M., Lounev, V., Carlton, A., Leonard, J., Cohen, R. I., Rappaport, E. F., Saitta, B., 2007. Identification of cord blood-derived mesenchymal stem/stromal cell populations with distinct growth kinetics, differentiation potentials, and gene expression profiles. *Stem Cells and Development*, 16(1), 53-73.

Nagamura-Inoue, T. and He, H., 2014. Umbilical cord-derived mesenchymal stem cells: Their advantages and potential clinical utility. *World Journal of Stem Cells*, 6(2), 195-202.

Nekanti, U., Mohanty, L., Venugopal, P., Balasubramanian, S., Totey, S. and Ta, M., 2010. Optimization and scale-up of Wharton's jelly-derived mesenchymal stem cells for clinical applications. *Stem Cell Research*, 5(3), 244-254.

Pittenger, M. F., Discher, D. E., Péault, B. M., Phinney, D. G., Hare, J. M. and Caplan, A. I., 2019. Mesenchymal stem cell perspective: cell biology to clinical progress. NPJ Regenerative Medicine, 4(1), 22.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. & Smyth, G. K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, 43(7), e47.

Rojewski, M. T., Weber, B. M. and Schrezenmeier, H., 2008. Phenotypic characterization of mesenchymal stem cells from various tissues. *Transfusion Medicine and Hemotherapy*, 35(3), 168-184.

Shakya, M., Lo, C. C. and Chain, P. S. G., 2019. Advances and Challenges in Metatranscriptomic Analysis. *Frontiers in Genetics*, 10, 904.

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Jensen, L. J. and von Mering, C., 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, 49, D605-D612.

Wang, Q., Li, J., Wang, S., Deng, Q., Wang, K., Dai, X., An, Y., Dong, G., Ke, W., Chen, F. and Liu, L., 2021. Single-cell transcriptome profiling reveals molecular heterogeneity in human umbilical cord tissue and culture-expanded mesenchymal stem cells. *FEBS Journal*, 288(18), 5311-5330.

Wang, Y., Chen, X., Cao, W. and Shi, Y., 2014. Plasticity of mesenchymal stem cells in immunomodulation: pathological and therapeutic implications. *Nature Immunology*, 15(11), 1009-1016.

Wilson, A., Murphy, M. J., Oskarsson, T., Kaloulis, K., Bettess, M. D., Oser, G. M., Pasche, A. C., Knabenhans, C., Macdonald, H. R. and Trumpp, A., 2008. c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes & Development*, 22, 2118-2123.

Xu, X., Wang, Y., Tan, X., Hu, Y., Zhang, X., Zhang, M., Wang, Y., Wu, J. and Hou, Y., 2024. Intravenous application of human umbilical cord mesenchymal stem cells alleviate neuropathic pain by suppressing microglia activation in rats. *Heliyon*, 10(13), e33340.

Yuan, Y. L., Huang, Q., Li, S., Zhang, X. and Wang, W. L., 2025. The role of heat shock proteins in tumorigenesis and their potential as targets for anti-tumor therapy. *Stem Cell Research & Therapy*, 178379.

Zhang, C., Han, X., Liu, J., Chen, L., Lei, Y., Chen, K., Si, J., Wang, T., Zhou, H., Zhao, X. and Zhang, X., 2022. Single-cell transcriptomic analysis reveals the cellular heterogeneity of mesenchymal stem cells. *Genomics, Proteomics & Bioinformatics*, 20(1), 70-86.