Moraveji et al. *Studies in Medical Sciences* (2025) 36(3):219-230
https://doi.org/10.61882/umj.36.3.219
Studies is Medical Sciences

ORIGINAL RESEARCH ARTICLE

Convergence of Signaling Pathways and Bioinformatics Analysis in Mesodermal Differentiation of iPSCs: Focus on KDR+/PDGFRα+ Populations for Cardiovascular Regeneration

Seyedeh-Faezeh Moraveji[1], Siamak Rezaeiani [2], Ay Sooda Mosafer Yadegari[3]

Published: 27 September 2025

Abstract

Background Cardiovascular diseases continue to be the leading cause of death worldwide. The use of iPSCs holds great promise for repairing heart and blood vessel tissues. Generation of cardiovascular progenitors requires precise modulation of these cells through signaling pathways. This study highlights the KDR and PDGFRα markers in guiding iPSCs toward mesodermal progenitors, specifically the KDR+/PDGFRα+ populations, which have enormous clinical promise for cardiovascular applications.

Methods Gene transcript analysis involved obtaining data from the GEO database with accession number GSE90000. The GEO2R tool was used to identify genes with significant changes, defined as p-values < 0.05 and absolute log-fold changes > 2.
Functional classification of genes was performed to identify biological processes and signaling pathways using GO analysis with the DAVID tool.
Protein-protein networks were analyzed by simulating protein interactions using the STRING database, which helped identify key genes such as EOMES.
Signaling pathway analysis used tools including Cytoscape, Reactome, and X2K to analyze pathways involved in iPSC differentiation into cardiomyocytes.

Results Our studies on KDR+/PDGFRα+ cells derived from iPSC differentiation revealed 1,635 genes that were significantly downregulated during cardiomyocyte formation, with p-values < 0.05 and |log-FC| ≥ 2. These genes include COCH, CYP26A1, and TUNAR. Using protein-protein interaction analysis, we identified EOMES (p-value 0.0026, |log-FC| -6.357) as a central transcription factor. Moreover, pathway enrichment analysis revealed a gradual downregulation of genes involved in cardiac disease, suggesting potential therapeutic applications.

Conclusion Integrating bioinformatics tools (GEO2R, STRING, Reactome) with multi-marker strategies (CD13, ROR2, APLNR) enhances the purity of cardiovascular progenitors, ultimately improving therapeutic applications in the treatment of cardiovascular diseases.

Keywords Bioinformatics, Cardiovascular diseases, Differentiation, Regenerative medicine, Signaling pathways

✉ Siamak Rezaeiani
siamak.cytology@gmail.com

1. Department of Stem Cells and Developmental Biology, Cell Science Research Center, Royan Institute for Stem Cell Biology and Technology, ACECR, Tehran, Iran
2. Cellular and Molecular Research Center, Cellular and Molecular Medicine Research Institute, Urmia University of Medical Sciences, Urmia, Iran
3. Univercity of Helth Scineces: Istanbul, Kadi Koy Turkey

Received: 27June 2025 / Revised: 5 August 2025 / Accepted: 7 September 2025

1 Introduction

Recent reports show that cardiovascular diseases continue to be the number one cause of death globally. In 2022, an estimated 19.8 million people died from cardiovascular diseases, accounting for 32% of all deaths around the world. Of these deaths, 85% were the result of heart attacks and strokes. It is estimated that deaths due to cardiovascular diseases will increase to 35.6 million in 2050, representing a 73.4% increase compared to 2025.[1] In the past decade, Regenerative medicine, with the discovery of induced pluripotent stem cells, has entered a new dimension. This breakthrough enables the reversion of somatic cells to a pluripotent state, making them suitable for tissue regeneration. The discovery of iPSCs by Yamanaka opened a new avenue of innovative therapeutic approaches but also raised essential questions about how cell fate is controlled during embryogenesis and its implications for the treatment of human diseases.[2] This finding is relevant to investigations into mesodermal derivatives, including cardiac and skeletal muscles, blood vessels, bone tissue, and blood. Many pathologies involve one or more of these elements, creating great expectations for novel treatments. Recent advances in tissue engineering and artificial organs emphasize the importance of this technology. Producing functional cells for each tissue type requires a thorough knowledge of differentiation pathways and the heterogeneity of precursor populations.

All three germ layers-ectoderm, mesoderm, and endoderm-can be differentiated, but how to direct them specifically to mesoderm remains a challenge. Mesodermal differentiation is precisely controlled by a network of signals, including Activin/Nodal, Wnt, FGF, and BMP, which act in sequence (Figure 1). Activin/Nodal commits the pluripotent cell to the mesodermal lineage, while Wnt and FGF support mesodermal differentiation by inhibiting GSK3β and PI3K/ERK; BMP, on the other hand, plays a role in mesoderm patterning.[3,4] This convergence of pathways reflects the strong dependence of mesodermal differentiation on precise timing and specific dosages of growth factors, such that even small changes in stimulus levels can shift cell fate from heart to blood vessels or to skeletal muscle. Clinically, precise control of these signals enables the generation of "therapeutic efficacy," or, in other words, the production of a target population free of contaminating unwanted cells. This is critical for clinical applications because inappropriate cells could give rise to tumorigenesis or unexpected functions.

Although iPSC generation has been successful, variability within mesodermal progenitor populations remains a significant challenge to overcome before translation into the clinic. For example, during cardiac differentiation, the cellular progression is from early mesoderm-expressing EOMES to a KDR+/PDGFRA+ stage, where both markers are coexpressed.[5] This KDR+/PDGFRA+ stage is transient, and during this point in time, cells can further differentiate into cardiomyocytes, mesenchyme, and endothelium. Beyond this, KDR expression declines, and ISL1 and NKX2-5 are upregulated to demarcate cells that enter a cardiac-specific trajectory. These dynamics underscore the importance of choosing the right timing for isolating and enriching the progenitor populations. Delays or premature harvests may lead to loss of multipotency or reduced purity of desired cell types.[6]

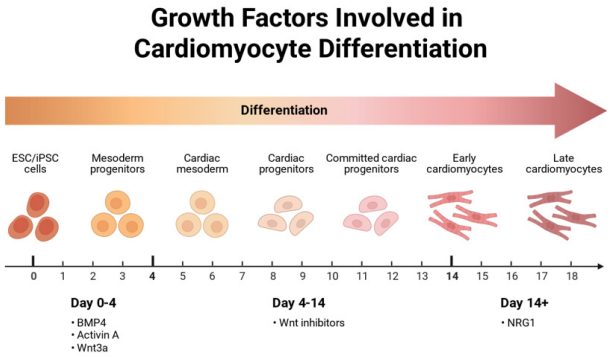

Figure 1 The figure is created with the Biorender/Available at http://app.biorender.com

Moreover, heterogeneity does not relate just to timing. There is also a challenge posed by the multitude of mesodermal subpopulations expressing distinct markers, and along with the KDR+/PDGFRA+ population, widely recognized as a cardiac mesodermal marker, a new subpopulation has been reported, characterized by triple positivity for KDR, CD56, and APLNR (defined as KNA+). A KNA+ population shows high clonogenic and proliferative capacity capable of forming endothelial colonies. Following transfer into diabetic animal models, these cells restored blood flow and repaired damaged retinal vessels.[7]

Notably, only a subset of these KDR+ or PDGFRA+ cells is functional, and the KNA+ population possesses higher regenerative ability.[8] These observations highlight how the use of multiple markers has been critical for more accurate prediction of function in mesodermal populations and for precisely defining subpopulation identity. Similarly, in another related experiment, the APLNR marker was more effective than both KDR and PDGFRα for cardiac progenitor isolation, because KDR and PDGFRα are also expressed in pluripotent and endodermal cell clusters.[9] Therefore, the identification and use of new markers to enhance the precision of isolation and reduce heterogeneity are vital to the progress made so far in the field. One of the recent remarkable works in regenerative medicine has

been the systematic application of growth factors and pathway inhibitors to drive cells into the endothelial and smooth muscle lineages. Studies have shown that combining BMP4 and GSK3 inhibitors with VEGF165 and Notch inhibitors rapidly differentiates iPSC-derived cells into endothelial progenitors. These progenitors are characterized by the expression of specific markers, such as VEC, CD31, CD34, and KDR^high, but are negative for CD14. Moreover, they possess high angiogenic and colony-proliferation capabilities.[10] In this manner, this approach targets two critical pathways, BMP/Wnt and VEGF/Notch, simultaneously. This "convergent" method, which drives cells toward differentiation via multiple pathways, has achieved higher efficiency than previously described methods that rely on a single factor. For vascular smooth muscle cells (VSMCs), PDGF-BB combined with TGF-β1 has been commonly used. These same studies showed that it was possible to modulate the synthetic and contractile phenotype by changing the doses of these factors and altering the physical properties of the substrate on which the cells attach.[11] The combination of an optimum amount of molecular stimulators with the biomechanical properties of the environment has facilitated the advancement of artificial vascular tissues. Some researchers noted that although the efficiency of differentiation is improving, achieving higher cardiomyocyte purity and maturation requires careful control of cell density at different developmental stages. According to a 2025 study, reseeding at critical time points during the early mesoderm stage, the KDR+/PDGFRA+ stage, and the ISL1+/NKX2-5+ stage results in approximately 10% higher cardiomyocyte purity.[5] This finding shows that densely populated culture environments that contain self-renewal signals and cell-to-cell contact are essential in driving the differentiation pathway. Freezing and storing cells at various stages also affects cells differently: storage at early stages reduces differentiation potential, whereas storage at advanced stages can improve final purity.[12,13] Information obtained from such research is vital for therapeutic cell-producing centers, as it indicates the optimal times for cell storage or passaging to achieve maximum yield. Recent advances in genetic engineering have allowed the engineering of specific vascular and smooth muscle progenitors. In 2024, it was observed that the temporal expression of the NKX3.1 gene in iPSCs can render these cells as mural cell progenitors, or iMPCs. These progenitors differentiate into pericytes and smooth muscle cells that exhibit contractility, secrete extracellular matrix, and support vascular networks in animal models. The same study identified NKX3.1's action in a complex with other transcription factors as integral to the conversion of cells into smooth muscle phenotypes.[14] The above example illustrates that manipulating a single gene is sufficient to achieve clinical efficacy.

Co-development of the cardiovascular system in the embryo involves interactions between different cell types. Therefore, several studies have focused on establishing protocols for co-differentiation of multiple lineages. For instance, one study used a combination of Activin A and CHIR-99021 to yield KDR+/PDGFRα+ cardiac mesoderm with an efficiency of nearly 95%. Multi-response models can also be used to control the ratios of mesenchymal, endothelial, and cardiac cells in culture. The results suggest that co-differentiation of cardiomyocytes with related cell types leads to better maturation than individual differentiation of cardiomyocytes. This underlines the role of cell-cell interaction during heart development. However, the proportion and temporal consideration of each cell type during these protocols are significant challenges to manage. These complicating factors necessitate the use of data analysis and modeling tools.[6]

Recent advances in laboratory techniques, single-cell data collection, and bioinformatics tools have remarkably enhanced our understanding of mesodermal differentiation. Among them, the public release of a dataset containing more than 60,000 single-cell transcriptomic profiles from iPSCs across the three germ layers is a significant development. Such a large dataset enables deep investigation of differentiation trajectories and comparisons across different protocols. Integrating these single-cell data into epigenetic, proteomic, and metabolomic information will facilitate the construction of a comprehensive regulatory landscape. This integration enables the discovery of new biomarkers and unidentified signaling pathways. Furthermore, the application of machine learning and network models to analyze these datasets is another critical example of how biology and data science are integrated to discover hidden patterns and predict cellular behavior.[15] Integration of multilayered data involves several challenges: there are no uniform standards for data processing, establishing causal relationships is difficult, and algorithms need to be optimized to minimize noise. Despite such impressive advances in research toward generating highly efficient cells, many important questions remain open. One crucial question is which signals regulate the dynamic expression of KDR and PDGFRα within mesodermal populations. Moreover, the interplay among the Wnt, Notch, FGF, and TGF-β pathways needs to be identified to determine cell fate.[7] Furthermore, the heterogeneity of progenitor populations, such as KDR+/PDGFRα+, KNA+, and APLNR+, results in uncertain differentiation trajectories, and no universal marker has yet been reported that can accurately identify functional subpopulations.[16,17] Again, though new protocols have reached significant milestones, several critical challenges

must be overcome for clinical use, including genomic stability, long-term safety, and the associated costs of mass production. Additionally, the complexity of co-differentiation modeling and the need to standardize multi-omics data are critical issues that require attention. There are numerous opportunities for advancement in this field. The combination of "convergence", the use in tandem of signaling pathways, genetic engineering, environmental modulation, and multi-tiered data analysis, along with "translational potential", the particular specific targeting of cell populations to treat specific diseases, and "surgical precision," or the ability to yield pure and functional cell populations, may allow for the eventual development of cell-based therapies. Specifically, the use of subpopulations, such as KNA+, with recognized regenerative vascular potential, in conjunction with single-cell maps for precise identification of signaling pathways, may provide a "blueprint" for developing personalized regenerative cells.[7] This article presents an analytical study that underlines the role of bioinformatics tools and the convergence of signaling pathways in generating functional cardiovascular progenitors from iPSCs. Single pathways, such as BMP4, FGF, and TGFβ/Nodal, demonstrate potency in guiding iPSCs toward cardiovascular fates, particularly in heart and blood vessel cell differentiation.

This study presents a synopsis of recent studies on iPSC differentiation into mesoderm, with a focus on KDR+/PDGFRα+ populations. Further, we have emphasized how bioinformatics tools have enhanced our understanding of cardiac progenitor differentiation and the screening of suitable candidates for drug research and therapy development. Advanced bioinformatics enables the identification of genes silenced or downregulated during cardiomyocyte differentiation, addressing knowledge gaps in signaling pathways, marker heterogeneity, and the integration of multi-omics data. From this analysis, it would appear that interventions targeting these silenced genes could not only enhance differentiation but also promote tissue regeneration. We conclude by recommending that experimental and computational approaches be combined to realize the full potential of iPSCs in therapeutic applications.

2 Methods

Analysis of Microarray

To study the transcriptomes of KDR+ and PDGFRα+ cells, a review of relevant articles was conducted, and a dataset was downloaded from the public Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo) under accession GSE90000. Kanki Wai and colleagues deposited this microarray dataset profiling the expression of KDR+ and PDGFRα+ progenitor cells derived from iPSCs and their differentiation into cardiomyocytes at days 3, 5, and 19. The data were re-analyzed using the GEO2R tool,[18] and transcripts with significant changes (p-value < 0.05 and $|\log\text{-fold change (FC)}| \geq 2$) were identified and selected.

Functional Classification of Differentially Expressed Transcripts

A gene ontology (GO) enrichment analysis was performed to identify the biological functions of differentially expressed genes (DEGs) with reduced expression. The GO of DEGs was evaluated using the GO enrichment analysis database.[19] To discover enriched signaling pathways, the Database for Annotation, Visualization, and Integrated Discovery (DAVID) was used. A p-value of < 0.05 was set as the significance threshold for selecting GO terms and signaling pathway entries.

Prediction of Transcription Factors and Construction of Protein-Protein Interaction (PPI) Networks

To interpret the regulators of DEGs, the gene list of transcripts with significant downregulation and their transcriptional regulators was used to construct a PPI network using the STRING database.[20] This web-based prediction tool offers evidence of PPIs through a combination of predictions and validated experimental results. It utilizes various methods, including text mining, experimental data, databases, co-expression analysis, neighborhood information, gene fusion, and co-occurrence data. Each interaction is assigned a confidence score, which categorizes it as low confidence (0.15), medium confidence (0.4), high confidence (0.7), or highest confidence (0.9). For visualizing networks, we focused on interactions with a confidence score greater than 0.7, indicating high-evidence interactions.

Biological and Signaling Pathway Analysis Using Software

To analyze biological pathways and identify signaling convergences, the tools Cytoscape, Reactome, and X2K were used.

Cytoscape 3.10.4:[21] Cytoscape is an advanced software tool for visualizing and analyzing biological networks, particularly for proteomic and genomic data. This platform enables researchers to simulate relationships among proteins, genes, and signaling pathways using complex networks, thereby allowing them to study interactions among these molecules. In this study, Cytoscape was employed to map and analyze PPI networks and to simulate interactions among key genes involved in the differentiation of KDR+/PDGFRα+ cells into cardiomyocytes.

Reactome:[22] A biological database designed for the analysis of signaling pathways and biological processes.

This database visually represents information related to biochemical reactions, signaling pathways, and cellular processes. It is used to simulate complex biological functions and assess molecular interactions. In this study, Reactome was used to identify and evaluate signaling pathways involved in mesodermal cell differentiation into cardiomyocytes.

X2K:[23] This tool is designed to predict transcription factors and identify gene networks involved in various cellular processes. Using gene expression profile data highlights key transcription factors involved in distinct cellular pathways. In this study, X2K was used to identify transcription factors that may regulate the expression of key genes during differentiation.

These tools employ complex analyses to identify new biological and genetic pathways, clarifying the role of transcription factors in directing stem cell differentiation into specialized cells. By combining these methods with an understanding of signaling pathways and multi-layered data analysis, we can effectively predict cellular behavior and develop new therapeutic approaches.

3 Results

Here, we reanalyzed microarray data from KDR+ and PDGFRα+ cells generated by cardiac differentiation of iPSCs. To identify the gene expression signature specific to KDR+ and PDGFRα+ cells, their gene expression data were compared with that of differentiated cardiomyocytes from days 3, 5, and 19, which were pooled for analysis. This comparison identified 8,494 differentially expressed transcripts (p-value < 0.05 and $|\text{log-FC}| \geq 1$), showing significant differential expression in KDR+ and PDGFRα+ cells compared to various differentiated cardiomyocyte types at the designated time points. Of those, 1,635 transcripts were downregulated in cardiomyocytes, as defined in the Methods section. Among the notably lower-expressed genes were COCH, CYP26A1, and TUNAR (Table 1).

Next, using the DAVID bioinformatics tool, the downregulated genes were analyzed for GO terms, including biological processes (Table 2), cellular components (Table 3), and cellular functions (supplementary Table 1), among the 100 genes with the most significant downward regulation. It was found that the most dramatic alterations in chromatin structure, as judged by RNA polymerase II-specific and DNA-binding transcription factor activity, occurred in progenitor cells during the course of evolutionary change toward cellular aggregation. These results provide insight into how the protein expression machinery of progenitor cells is altered to become more like that of cardiomyocytes.

Table 1 Genes with the most significant downregulation in the differentiation of cardiac progenitor cells into cardiomyocytes (Not all 1,635 genes are displayed due to page limitations)

ID	P-value	logFC	Gene.symbol
1554242_a_at	0.000407	-9.82059	COCH
206424_at	0.0007	-9.65363	CYP26A1
205229_s_at	0.000433	-9.41135	COCH
232111_at	0.00044	-9.32213	TUNAR
202888_s_at	0.000954	-8.95353	ANPEP
207197_at	0.000647	-8.92876	ZIC3
223642_at	0.020353	-8.91771	ZIC2
206230_at	0.002368	-8.88323	LHX1
231731_at	0.002328	-8.83373	OTX2
215145_s_at	0.003129	-8.77486	CNTNAP2
205818_at	0.000583	-8.40491	BRINP1
210503_at	0.001373	-8.03938	MAGEA11
239205_s_at	0.005618	-7.99874	CR1L///CR1
1554593_s_at	0.000962	-7.95261	SLC1A6
220448_at	0.001021	-7.55939	KCNK12
219537_x_at	0.005455	-7.52257	DLL3
1562713_a_at	0.002335	-7.4821	NETO1
1554592_a_at	0.00118	-7.43685	SLC1A6
1554340_a_at	0.000854	-7.39966	DRAXIN
208062_s_at	0.004435	-7.36649	NRG2
214532_x_at	0.005066	-7.35787	POU5F1B
204416_x_at	0.001171	-7.26147	APOC1
207854_at	0.009767	-7.21563	GYPE
211821_x_at	0.002546	-7.13821	GYPA
239823_at	0.001138	-7.10852	LOC101927841
1555613_a_at	0.002012	-7.05844	ZAP70
210905_x_at	0.011724	-6.97057	POU5F1P4
242128_at	0.003833	-6.93162	OTX2
1569023_a_at	0.001148	-6.91205	LINC00458
1553652_a_at	0.005197	-6.91178	C18orf54
206528_at	0.011966	-6.81034	TRPC6
232985_s_at	0.001597	-6.78157	DPPA4
214407_x_at	0.016783	-6.74098	GYPB
220384_at	0.005471	-6.6537	NME8
219823_at	0.001049	-6.64983	LIN28A
213592_at	0.001705	-6.62265	APLNR
227282_at	0.003556	-6.62188	PCDH19
1553060_at	0.001603	-6.59653	PSKH2
229233_at	0.001162	-6.59268	NRG3
229273_at	0.001355	-6.56631	SALL1
1560291_at	0.005136	-6.56384	RIPPLY1

Table 2 Biological processes associated with the 100 genes that show the most significant downregulation during the conversion of cardiac progenitors into mature cardiomyocytes. These processes are inhibited as progenitors transition to mature heart cells.

Category	Term	Genes	Pop total	P-value
GOTERM_BP_1	Developmental process	56%	19416	3.63E-08
GOTERM_BP_1	Multicellular organismal process	57%	19416	1.7E-07
GOTERM_BP_1	Biological regulation	71%	19416	0.0127
GOTERM_BP_1	Biological process involved in intraspecies interaction between organisms	4%	19416	0.0188
GOTERM_BP_1	Growth	8%	19416	0.0242
GOTERM_BP_1	Reproductive process	15%	19416	0.0268
GOTERM_BP_1	Homeostatic process	14%	19416	0.0416

Table 3 Impact of the 100 downregulated genes on the cell during the conversion of cardiac progenitors into cardiac cells

Category	Term	Genes	Pop total	P-value
GOTERM_CC_DIRECT	Chromatin	15%	20795	0.00213
GOTERM_CC_DIRECT	Glutamatergic synapse	10%	20795	0.00489
GOTERM_CC_DIRECT	Presynaptic membrane	5%	20795	0.0161
GOTERM_CC_DIRECT	Perikaryon	5%	20795	0.0195
GOTERM_CC_DIRECT	Nucleus	41%	20795	0.0198
GOTERM_CC_DIRECT	Ankyrin-1 complex	3%	20795	0.0281
GOTERM_CC_DIRECT	Paranode region of the axon	3%	20795	0.0419
GOTERM_CC_DIRECT	Juxtaparanode region of the axon	3%	20795	0.0453
GOTERM_CC_DIRECT	GABAergic synapse	4%	20795	0.0567
GOTERM_CC_DIRECT	Extracellular region	18%	20795	0.0634
GOTERM_CC_DIRECT	Calyx of Held	3%	20795	0.0722
GOTERM_CC_DIRECT	Dendrite	6%	20795	0.084

Protein Interactions

One of the most critical concerns in cellular processes is PPIs. Identifying an upstream protein in these molecular processes can provide a valuable molecular target to enhance the progression of converting progenitor cells into mature cardiac cells. To explore this, the STRING database was used to examine protein interactions associated with the 100 most downregulated genes. The analysis revealed that the EOMES protein plays a crucial role in this process (Figure 2). Further investigation of EOMES revealed that it is specific to eukaryotes (Supplementary Figure 1). Since EOMES is downregulated during the transition from progenitors to cardiomyocytes, it could serve as a potential candidate to help maintain progenitor cell properties in laboratory cultures or to guide them toward becoming specialized cardiomyocytes.

Main Activities Altered in the Differentiation Pathway

We needed to understand how the routine activities of a cell are impacted during the differentiation of progenitor cells into cardiac cells. To analyze the data, we utilized the Reactome database. The results (Figure 3) indicated that as we progress toward cardiomyocytes, there is a significant downregulation of genes associated with cardiac disease. This finding could provide a valuable model for treating heart disease through the use of genetically engineered progenitor cells or heart transplantation.

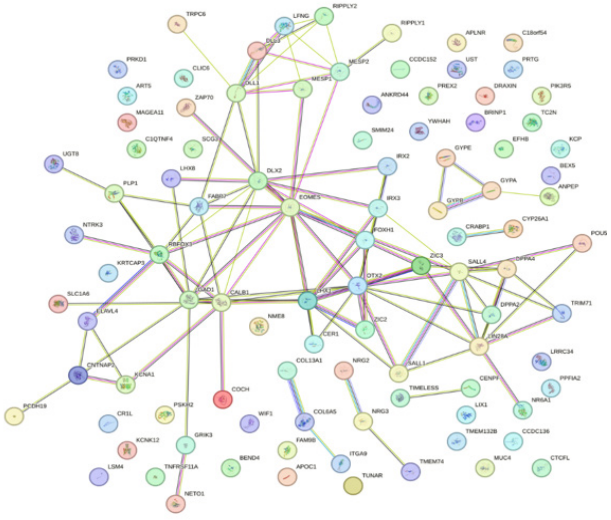

Figure 2 This shows the interaction network of downregulated DEGs and their transcriptional regulators. Line colors indicate interaction types: blue (databases), purple (experimental), green (gene neighborhood), red (gene fusion), dark blue (co-expression), yellow (text mining), and black (co-occurrence). The figure was generated based on bioinformatics analysis using the STRING database.

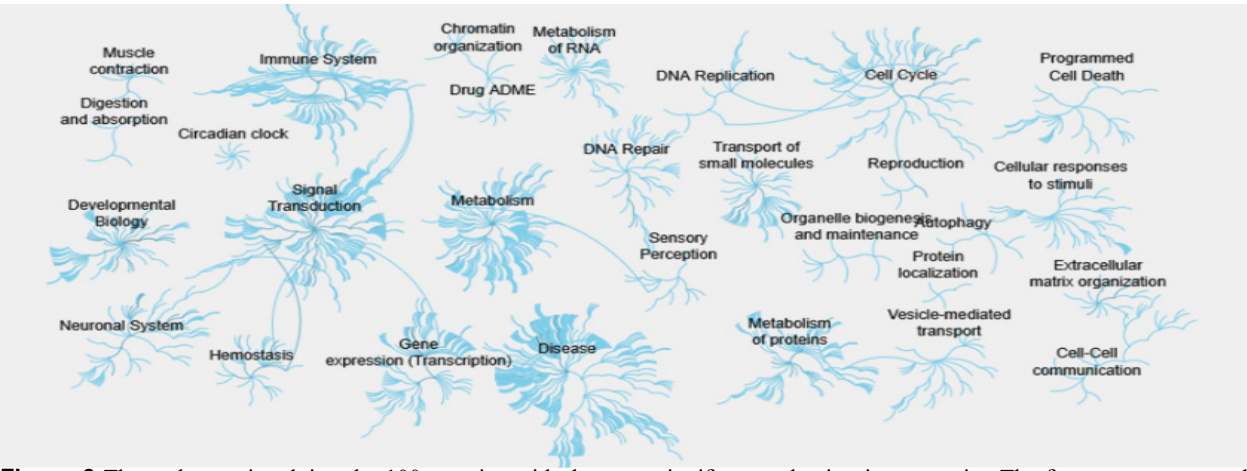

Figure 3 The pathways involving the 100 proteins with the most significant reduction in expression.The figure was generated based on bioinformatics analysis using the REACTOME database.

Examination of the Presence of Kinases and Transcription Factors Among Selected Genes

Given the essential role of kinases and transcription factors in cellular and molecular pathways and differentiation, we examined the 100 most downregulated genes for the presence of these proteins using the X2K tool. The analysis identified 16 kinases and transcription factors among these genes, highlighting their significant importance and potential impact. This is illustrated schematically in the accompanying figure (Figure 4), suggesting that they may be suitable candidates for cardiac studies.

Examination of All Genes

We identified downregulated genes using stringent criteria: a p-value < 0.05 and a log-fold change (FC) of $|\geq 2|$. A total of 100 of the most downregulated genes were further evaluated. However, when these strict criteria were not applied, we found that 54,676 genes were affected (either upregulated or downregulated) during the conversion of cardiac progenitors into mature cardiac cells. All 54,676 genes were examined in the Cytoscape database version 3.10.3.

Our analysis revealed that the GP160 (Probable G-protein coupled receptor 160 protein "gene GPR160 ") emerged as a central hub in the PPI network. Despite showing modest expression changes ($log_2FC = -0.17$, $p = 0.72$), GP160 occupied the highest upstream position in the signaling cascade, with no regulators identified above it in the network hierarchy (Figure 5). As a core G-protein subunit, GP160 may coordinate multiple downstream pathways during cardiac progenitor-to-cardiomyocyte transition, warranting further experimental validation. This discovery could serve as a starting point for investigating the role of this protein in cardiac cell differentiation.

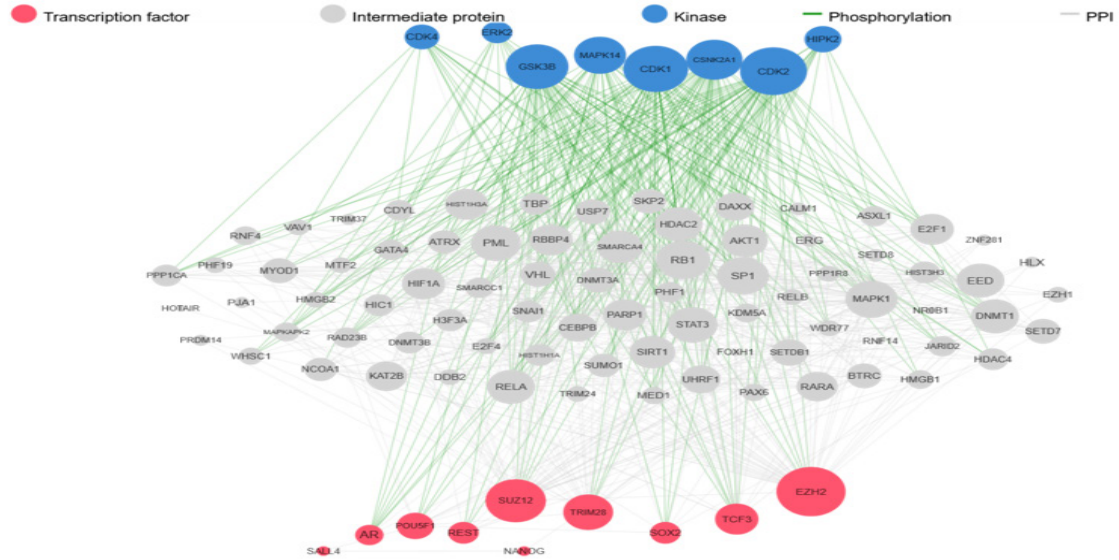

Figure 4 Kinase proteins and transcription factors identified among the 100 downregulated genes.The figure was generated based on bioinformatics analysis using the X2K database.

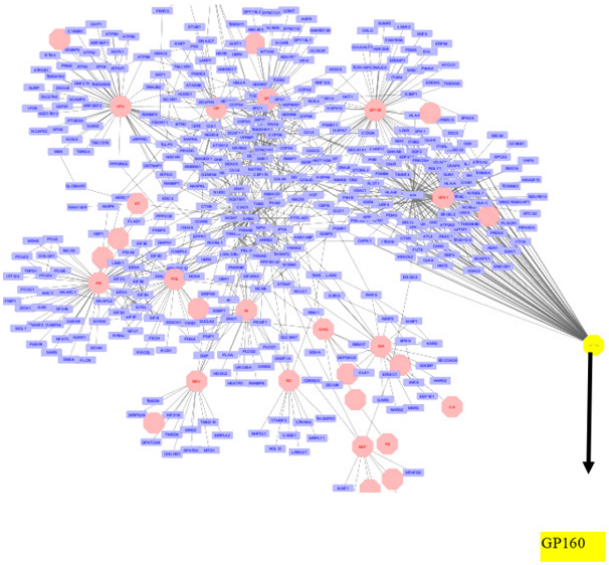

Figure 5 Among all the altered genes involved in converting progenitor cells into mature cardiac cells, the GP160 gene, a G protein, plays a crucial upstream role in gene expression. This Figure was generated based on bioinformatics analysis using the SRTING database (PP1 network)

In the end, the volcano plot displayed in Figure 6 illustrates the gene expression profile of KDR+/PDGFRα+ cells as they differentiate into cardiomyocytes. This plot shows significant changes in gene expression, highlighting several genes that are either substantially upregulated or downregulated. These findings provide valuable insights into the molecular mechanisms underlying differentiation.

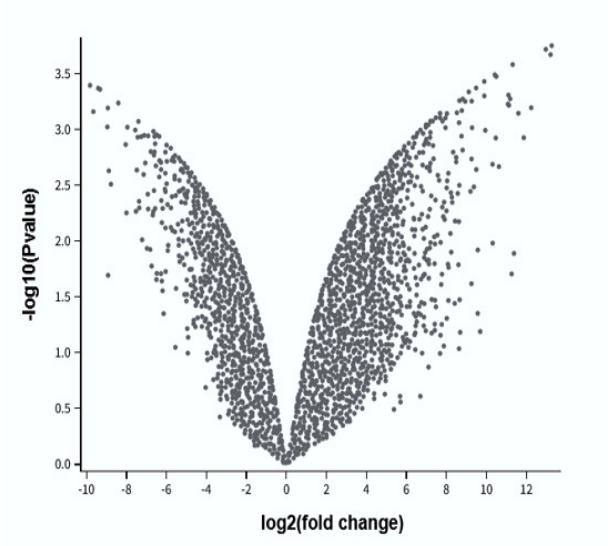

Figure 6 The volcano plot displays gene-expression changes in KDR+/PDGFRα+ cells differentiated into cardiomyocytes relative to controls, highlighting genes with $p < 0.05$ and $|\log2 \text{ fold change}| \geq 2$. It shows significant upregulation or downregulation during differentiation. Generated using the G2R Analysis tool.

4 Discussion

Analysis of the GSE90000 array set demonstrates that KDR+/PDGFRα+ cells undergo a dramatic shift in gene expression upon cardiomyocyte differentiation, exemplified by the repression of cell cycle and cell division gene sets and the induction of those involved in extracellular matrix production. These findings indicate a transition from a progenitor state toward a functional cardiac cell type. Such results are concordant with prior reports highlighting the use of BMP4, FGF, and TGF-β/nodal to form mesoderm and WNT to modulate cardiac fate.[24,25] Supporting this, this study showed that KDR repression, particularly on day five of differentiation, is accompanied by low GATA6 activity in the cardiac mesoderm. Perhaps the biggest challenge with the populations of KDR+/PDGFRα+ cells is the high degree of heterogeneity.[26] While traditional markers such as KDR and PDGFRα enable the enrichment of mesoderm, they do not ensure that all selected cells possess a cardiac or vascular fate.[27]

Such combined markers, such as CD13 and ROR2, or ALCAM and ISL1, have already demonstrated that more specific populations can be obtained, thereby improving their performance in animal models.[28] Bioinformatic analyses showed that many of the downregulated genes are associated with the WNT, Notch, and BMP signaling pathways. These observations are in line with reports indicating that temporal and dosage control of these pathways determines the final cell fate.[7]

Correspondingly, our bioinformatics analysis of KDR+/PDGFRα+ progenitors revealed downregulation of 1,635 genes during cardiac differentiation. We identified EOMES and GP160 as central regulators using functional enrichment analysis. Our bioinformatics analysis thus uncovered suppression of pluripotency-related pathways, including WNT and Notch, while cardiac-specific programs were activated to drive progenitor cells toward a cardiac lineage. EOMES is a eukaryote-specific T-box transcription factor that progressively showed downregulation during the different stages ($\log_2FC = -X$, $p < 0.001$). This study suggests EOMES has a key role in mesodermal differentiation. Moreover, suppression of EOMES may be necessary for cardiac lineage commitment. Future studies should examine whether forced expression of EOMES maintains progenitor self-renewal and whether it can further improve the efficiency of cardiac differentiation protocols while offering greater control over the process. Despite its expression change, GP160 maintained a central position in the PPI network, underscoring its role as a key G-protein subunit that controls GPCR signaling. Our data suggest that other markers, such as CD13 and ROR2, may help isolate more homogeneous progenitor subsets; however, this

requires validation. Single-cell RNA sequencing will be instrumental in resolving these subpopulations and will provide valuable insight into their differentiation processes. These findings are expected to help refine the strategies for generating more homogeneous populations relevant to therapy.

It is worth noting that our study was based on the data from iPSC (GSE90000), though the title states iPSCs. However, hESCs and iPSCs both have the same differentiation potential.[29] These findings should be confirmed in further research using iPSC datasets, especially patient-derived datasets, to assess variability and ensure these results are clinically relevant for personalized medicine.

In summary, while bioinformatics tools and functional analyses have provided critical insights into the molecular mechanisms of mesodermal specification, further experimental validation using iPSC-specific datasets is paramount to address limitations in sample size and population heterogeneity. The integration of powerful technologies, such as single-cell RNA sequencing and machine learning models, will be imperative for refining these differentiation protocols and advancing therapeutic outcomes with iPSC-derived cardiovascular progenitors. Their promise lies in further guiding the precise differentiation of pluripotent stem cells into functional cardiovascular progenitors and advancing regenerative medicine applications.

Clinical Implications and Future Perspectives

The present article demonstrates that combining growth factors, genetic engineering, and modified culture conditions increases the efficiency and purity of target cell differentiation. For instance, by applying KDR+/PDGFRα+ cells, which have been shown to generate blood vessels in diabetic models, while monitoring gene expression and using complementary markers, one could obtain tailored cells for patient treatment. It is necessary to create pure populations of cells free of unwanted cells to avoid tumorigenesis.

For this, several approaches should be taken: CRISPR-based perturbation studies of EOMES and GP160, which will allow insight into genetic manipulation influencing the process of differentiation; RNA-seq of single cells, mapping differentiation trajectories for a clear insight into cell evolution; functional assays such as flow cytometry or patch-clamp techniques should be performed to test certain marker combinations, like KDR+/PDGFRα+/CD13+/ROR2+, to ensure the isolation of pure progenitor populations. Additionally, validating induced pluripotent stem cells using matched datasets will enhance the robustness and reproducibility of the differentiation protocols. Our analysis identified that appropriately timed cell separation, for example, day 5 of differentiation, combined with multiple markers, significantly enhances both cell purity and therapeutic efficacy. Taken together, these methods will improve the generation of high-quality, patient-specific cells that are suitable for regenerative medicine applications.[26]

One limitation of this study is the small sample size, containing only five replicates. This restricts the statistical power to detect individual variation and linear heterogeneity within the population. Another limitation is that it applies bulk RNA sequencing, which averages gene expression across the entire sample and thus cannot resolve intra-population heterogeneity. Further, many downregulated genes lacked functional annotations, complicating our understanding of their roles, as reflected in the GO analysis. Such would be better addressed by researchers considering bioinformatics tools that account for the hierarchical structure of GO terms, since this would yield more accurate functional insights.[30] Moreover, the results are purely computational and lack experimental validation. Cross-platform bias might also lead to the exclusion of low-abundance transcripts in microarray data.

Single-cell RNA sequencing will be necessary for uncovering hidden patterns of gene expression and cellular behavior. Integration of more layers of data, such as transcriptomes, epigenomes, and proteomes, will also be helpful. The development of machine learning algorithms that predict cell behavior or improve cell separation will significantly enhance the accuracy and impact of future studies in this field.

5 Conclusion

An integrative bioinformatics approach identifies novel regulators (EOMES, GP160) and highlights the challenges of heterogeneity in cardiac differentiation. Though promising, these findings need rigorous experimental validation before clinical translation. Another note is that integrating signaling pathways with surface markers and bioinformatics analyses provides a novel framework that guides iPSCs toward cardiac and vascular mesoderm development. The present study explicitly focuses on populations marked by KDR+/PDGFRα+ and underscores the need for careful regulation of both dosage and timing in growth factors. Moreover, such an approach also requires complementary markers, along with genome analysis tools and protein networks, to reduce heterogeneity within cell populations and generate more functional, targeted cell populations. This precision, so achieved, may enhance cell-based therapies and improve tissue regeneration. To address these challenges, our analysis used advanced bioinformatics tools, including GEO2R, STRING, and Reactome. This helped us identify key genes and

transcription factors involved in the differentiation process. These bioinformatics tools have shed light on complex networks of gene interactions and signaling pathways essential for mesodermal differentiation, a step toward the development of functional cardiovascular progenitors. In addition, bulk-cell RNA sequencing and machine learning models are increasingly crucial for predicting cellular behavior and optimizing differentiation protocols. Such progress carries hope for personalized treatments of CVDs in the near future. In this regard, future studies will rely on genetic engineering, growth factors, and culture conditions to increase the yield and purity of the target cell populations. Similarly, multi-omics data integration is a promising new technology for identifying novel biomarkers and regulatory pathways that could enhance regenerative therapies. Furthermore, advanced computational models will provide insights into the timing and dosage of signaling factors to produce functional, safe cells for clinical applications. In this regard, the present study represents an essential contribution to regenerative medicine by providing an inclusive roadmap for overcoming current challenges in iPSC differentiation. The integration of bioinformatics with experimental techniques opens new avenues for cardiovascular regeneration and thus paves the way for the development of more viable and personalized therapies for heart disease.

Declarations

Acknowledgments
The authors express their gratitude to all those who assisted them in this research.

Artificial Intelligence Disclosure
Artificial intelligence tools were used to enhance grammar and improve the manuscript's language.

Authors' Contributions
Siamak Rezaeiani conducted the topic selection, data analysis, and manuscript preparation. Seyedeh Faezeh Moraveji contributed to the development of the topic and the literature review.

Availability of Data and Materials
The datasets used in this study are publicly available through the Gene Expression Omnibus (GEO) database under the accession number GSE90000. Additional data and materials can be accessed upon request.

Conflict of Interest
The authors of this study declare that this work is the result of independent research and that there are no conflicts of interest with any organizations or individuals.

Consent for Publication
Not applicable.

Ethical Considerations
Not required as it was a bioinformatic analysis.

Funding
This research did not receive any specific funding. The authors declare no financial support for the research, authorship, or publication of this article

The supplementary information file has been submitted as a separate document and includes additional figures and tables that support the findings of this study.

References

1. Chong B, Jayabaskaran J, Jauhari SM, Chan SP, Goh R, Kueh MTW, et al. Global burden of cardiovascular diseases: projections from 2025 to 2050. Eur J Prev Cardiol. 2025;32(11):1001-15. doi: 10.1093/eurjpc/zwadXXX
2. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76. doi: 10.1016/j.cell.2006.07.024
3. Burdsal CA, Flannery ML, Pedersen RA. FGF-2 alters the fate of mouse epiblast from ectoderm to mesoderm in vitro. Dev Biol. 1998;198(2):231-44. doi: 10.1006/dbio.1998.8882
4. Carmeliet P. One cell, two fates. Nature. 2000;408(6808):43-5. doi: 10.1038/35040644
5. Feeney AK, Simmons AD, Peplinski CJ, Zhang X, Palecek SP. Enhancing human pluripotent stem cell differentiation to cardiomyocytes through cardiac progenitor reseeding and cryopreservation. iScience. 2025;28(5). doi: 10.1016/j.isci.2025.110123
6. Akiyama H, Katayama Y, Shimizu K, Honda H. Engineering a controlled cardiac multilineage co-differentiation process using statistical design of experiments. Stem Cell Res Ther. 2025;16(1):273. doi: 10.1186/s13287-025-03763-1
7. Gil C-H, Chakraborty D, Vieira CP, Prasain N, Calzi SL, Fortmann SD, et al. Specific mesoderm subset derived from human pluripotent stem cells ameliorates microvascular pathology in type 2 diabetic mice. Sci Adv. 2022;8(9):eabm5559. doi: 10.1126/sciadv.abm5559
8. Sugimura R, Jha DK, Han A, Soria-Valles C, Da Rocha EL, Lu Y-F, et al. Haematopoietic stem and progenitor cells from human pluripotent stem cells. Nature. 2017;545(7655):432-8. doi: 10.1038/nature22370
9. Lam Y-Y, Chan C-H, Geng L, Wong N, Keung W, Cheung Y-F. APLNR marks a cardiac progenitor derived with human induced pluripotent stem cells. Heliyon. 2023;9(7). doi: 10.1016/j.heliyon.2023.e17684

10. Elcheva I, Brok-Volchanskaya V, Kumar A, Liu P, Lee J-H, Tong L, et al. Direct induction of haematoendothelial programs in human pluripotent stem cells by transcriptional regulators. Nat Commun. 2014;5(1):4372. doi: 10.1038/ncomms5372

11. Jiao Y-C, Wang Y-X, Liu W-Z, Xu J-W, Zhao Y-Y, Yan C-Z, et al. Advances in the differentiation of pluripotent stem cells into vascular cells. World J Stem Cell. 2024;16(2):137. doi: 10.4252/wjsc.v16.i2.137

12. Grancharova T, Gerbin KA, Rosenberg AB, Roco CM, Arakaki JE, DeLizo CM, et al. A comprehensive analysis of gene expression changes in a high replicate and open-source dataset of differentiating hiPSC-derived cardiomyocytes. Sci Rep. 2021;11(1):15845. doi: 10.1038/s41598-021-95273-2

13. Lian X, Zhang J, Zhu K, Kamp TJ, Palecek SP. Insulin inhibits cardiac mesoderm, not mesendoderm, formation during cardiac differentiation of human pluripotent stem cells and modulation of canonical Wnt signaling can rescue this inhibition. Stem cells. 2013;31(3):447-57. doi: 10.1002/stem.1283

14. Lee U, Zhang Y, Zhu Y, Luo AC, Gong L, Tremmel DM, et al. Robust differentiation of human pluripotent stem cells into mural progenitor cells via transient activation of NKX3. 1. Nat Commun. 2024;15(1):8392. doi: 10.1038/s41467-024-48392-1

15. Shen S, Werner T, Chiu HS, Chen X, Nguyen Q, Palpant NJ. A pluripotent stem cell atlas of multilineage differentiation. Sci Data. 2025;12(1):1238. doi: : 10.1038/s41597-025-03792-4

16. Yang L, Soonpaa MH, Adler ED, Roepke TK, Kattman SJ, Kennedy M, et al. Human cardiovascular progenitor cells develop from a KDR+ embryonic-stem-cell-derived population. Nature. 2008;453(7194):524-8. doi: 10.1038/nature06894

17. Kattman SJ, Witty AD, Gagliardi M, Dubois NC, Niapour M, Hotta A, et al. Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. Cell stem. 2011;8(2):228-40. doi: 10.1016/j.stem.2010.12.008

18. NCBI GEO2R. Available from: https://www.ncbi.nlm.nih.gov/geo/geo2r/

19. Gene Ontology enrichment analysis. Available from: https://geneontology.org/

20. STRING database. Available from: https://string-db.org/

21. Cytoscape. Available from: https://cytoscape.org/

22. Reactome. Available from: https://reactome.org/

23. X2K. Available from: https://maayanlab.cloud/X2K/

24. Burridge PW, Thompson S, Millrod MA, Weinberg S, Yuan X, Peters A, et al. A universal system for highly efficient cardiac differentiation of human induced pluripotent stem cells that eliminates interline variability. PloS One. 2011;6(4):e18293. doi: 10.1371/journal.pone.0018293

25. Mummery CL, Zhang J, Ng ES, Elliott DA, Elefanty AG, Kamp TJ. Differentiation of human embryonic stem cells and induced pluripotent stem cells to cardiomyocytes: a methods overview. Circ Res. 2012;111(3):344-58. doi: 10.1161/CIRCRESAHA.110.227512

26. Bisson JA, Gordillo M, Kumar R, de Silva N, Yang E, Banks KM, et al. GATA6 regulates WNT and BMP programs to pattern precardiac mesoderm during the earliest stages of human cardiogenesis. eLife. 2025;13:RP100797. doi: 10.7554/eLife.100797

27. Skelton RJ, Brady B, Khoja S, Sahoo D, Engel J, Arasaratnam D, et al. CD13 and ROR2 permit isolation of highly enriched cardiac mesoderm from differentiating human embryonic stem cells. Stem Cell Reports. 2016;6(1):95-108. doi: 10.1016/j.stemcr.2015.11.012

28. Ghazizadeh Z, Fattahi F, Mirzaei M, Bayersaikhan D, Lee J, Chae S, et al. Prospective isolation of ISL1+ cardiac progenitors from human ESCs for myocardial infarction therapy. Stem Cell Reports. 2018;10(3):848-59. doi: 10.1016/j.stemcr.2018.01.017

29. Narsinh KH, Plews J, Wu JC. Comparison of human induced pluripotent and embryonic stem cells: fraternal or identical twins? Mol Ther. 2011;19(4):635-8. doi: 10.1038/mt.2011.41

30. Huang DW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):R183. doi: 10.1186/gb-2007-8-9-r183

Supplementary data

Table 1 supplementary Shows the cellular functions of the 100 downregulated genes during the process of converting cardiac progenitors into cardiac cells.

Category	Term	Genes	Pop Total	P-Value
GOTERM_MF_DIRECT	DNA-binding transcription factor activity, RNA polymerase II-specific	20%	19253	6.22E-05
GOTERM_MF_DIRECT	DNA-binding transcription factor activity	14%	19253	9.75E-05
GOTERM_MF_DIRECT	RNA polymerase II cis-regulatory region sequence-specific DNA binding	18%	19253	0.000431
GOTERM_MF_DIRECT	DNA-binding transcription activator activity, RNA polymerase II-specific	10%	19253	0.00275
GOTERM_MF_DIRECT	sequence-specific double-stranded DNA binding	10%	19253	0.00569
GOTERM_MF_DIRECT	sequence-specific DNA binding	8%	19253	0.00952
GOTERM_MF_DIRECT	DNA binding	15%	19253	0.0252
GOTERM_MF_DIRECT	signaling receptor activity	5%	19253	0.0557
GOTERM_MF_DIRECT	transcription cis-regulatory region binding	5%	19253	0.0636
GOTERM_MF_DIRECT	retinoic acid binding	3%	19253	0.0732
GOTERM_MF_DIRECT	Notch binding	3%	19253	0.0941

Figure 1 supplementary displays the expression of proteins from Figure 2 across different organisms. This indicates that the EOMES gene is specific to eukaryotes. The figure was generated based on bioinformatics analysis using the STRING database.