Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

GRAPH CONTRASTIVE LEARNING WITH AUTOENCODERS FOR RECONSTRUCTING CLUSTERS IN PERIODONTAL REGENERATION WITH PLURIPOTENT STEM CELLS

Dinesh, Devika[1]; **Pradeep Kumar Yadalam**[1]; **Carlos M. Ardila**[1,2]

1. Department of Periodontics, Saveetha Dental College & Hospital, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai-600077, Tamil Nadu, India..
2. Ph.D. Postdoctoral Researcher. Basic Sciences Department, Biomedical Stomatology Research Group, Faculty of Dentistry, Universidad de Antioquia U de A, Medellín, Colombia.

EMAIL: martin.ardila@udea.edu.co

ABSTRACT

Introduction: Because induced pluripotent stem cells (iPSCs) can differentiate into different periodontal lineages, they have promising regenerative potential. However, the identification of regenerative subpopulations is hampered by the heterogeneity in iPSC-derived populations and the shortcomings of traditional clustering techniques. **Objective:** This work presents a graph contrastive learning (GCL) framework combined with autoencoders to find biologically significant clusters in high-dimensional gene expression data from iPSC-derived cells. Methods: The GEO dataset GSE283726's gene expression data underwent preprocessing,

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

filtering, and normalization (log2(x+1). Cosine similarity was used to create a k-nearest neighbour graph (k=10). A symmetrical autoencoder used mean squared error as the loss function to reconstruct the original inputs after compressing features into a 128-dimensional latent space. GraphSAGE and InfoNCE loss increased agreement between augmented views with edge dropout (0.2) and feature masking (10%). **Results:** Reconstruction loss was reduced to 0.0001 after the autoencoder converged in 20 epochs. With an ideal silhouette score of 0.694, clustering analysis revealed two transcriptionally distinct populations. All experimental groups showed high expression of Cluster 1, which most likely represented fibroblastic or osteogenic lineages. The low expression in Cluster 0 was indicative of progenitor or quiescent states. PCA validated the intrinsic data structure, with PC1 accounting for 96.63% of the variance. **Conclusions:** The suggested GCL-autoencoder framework provides a scalable, unsupervised method for identifying regenerative subtypes for periodontal therapy by successfully reconstructing lineage-specific clusters in iPSC-derived cell populations.

KEYWORDS: Gene expression; Induced pluripotent Stem Cells; Machine learning; Periodontal regeneration.

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

APRENDIZAJE CONTRASTIVO DE GRAFOS CON AUTOENCODERS PARA LA RECONSTRUCCIÓN DE AGRUPAMIENTOS EN LA REGENERACIÓN PERIODONTAL CON CÉLULAS MADRE PLURIPOTENTES

RESUMEN

Introducción: Debido a que las células madre pluripotentes inducidas (iPSC) pueden diferenciarse en distintos linajes periodontales, tienen un gran potencial regenerativo. **Objetivos:** Presentar un marco de aprendizaje contrastivo de grafos (GCL) combinado con autoencoders para identificar agrupaciones biológicamente relevantes en datos de expresión génica de alta dimensionalidad provenientes de células derivadas de iPSC. **Métodos:** Los datos de expresión génica del conjunto GEO GSE283726 fueron preprocesados, filtrados y normalizados (log2(x+1)). Se utilizó la similitud del coseno para construir un grafo de vecinos más cercanos (k=10). Un autoencoder simétrico empleó el error cuadrático medio como función de pérdida para reconstruir las entradas originales después de comprimir las características en un espacio latente de 128 dimensiones. GraphSAGE y la pérdida InfoNCE mejoraron la concordancia entre vistas aumentadas con dropout de aristas (0,2) y enmascaramiento de características (10%). **Resultados:** La pérdida de reconstrucción se redujo a 0,0001 después de que el autoencoder convergiera en 20 épocas. El análisis de

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

DOI:

agrupamiento reveló dos poblaciones transcripcionalmente distintas, con un puntaje de silueta ideal de 0,694. El Cluster 1, que probablemente representa linajes fibroblásticos u osteogénicos, mostró alta expresión en todos los grupos experimentales. La baja expresión en el Cluster 0 sugirió estados progenitoros o quiescentes. El PCA validó la estructura intrínseca de los datos, con el PC1 explicando el 96,63% de la varianza. **Conclusiones:** El marco GCL-autoencoder propuesto ofrece un método escalable y no supervisado para identificar subtipos regenerativos útiles en terapia periodontal, al reconstruir exitosamente agrupaciones específicas de linaje en poblaciones celulares derivadas de iPSC.

PALABRAS CLAVE: Expresión génica; células madre pluripotentes inducidas; aprendizaje automático; regeneración periodontal.

INTRODUCTION

Alveolar bone, periodontal ligament (PDL), and cementum are among the tissues lost due to periodontitis, and the goal of periodontal regeneration[1–3] is to restore their structural and functional integrity. Autografts, scaffolds made of biomaterials, and guided tissue regeneration are examples of traditional regenerative therapies that have shown mixed and limited results, especially in complex defects. Stem cell-based strategies have presented encouraging prospects for consistent periodontal regeneration in recent years. The ability of induced pluripotent stem cells (iPSCs) to self-renew

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

and differentiate into all somatic cell types, including mesenchymal cells essential for periodontal repair, has made them a revolutionary tool among these. This hybrid approach improves scaffold design, cell sourcing, and therapeutic precision in iPSC-based periodontal regeneration. It also guarantees improved modeling of differentiation dynamics, improved interpretability, and accuracy in cell state classification.[4]

By reprogramming somatic cells using specific transcription factors (such as Oct4, Sox2, Klf4, and c-Myc), iPSCs remove the moral dilemmas surrounding embryonic stem cells and enable patient-specific treatments. The ability of iPSC-derived mesenchymal stem cells (iPSC-MSCs)[4] to differentiate into osteoblasts, cementoblasts, and fibroblasts has been shown in preclinical research, which could aid in the regeneration of complicated periodontal defects. Additionally, iPSCs provide a platform that can be expanded and customized for scaffold-based delivery systems, gene editing, and disease modelling.[5] As translational research progresses, the future of periodontal regenerative medicine may be completely transformed by combining iPSC-based treatments with bioactive compounds, 3D scaffolds, and tissue engineering techniques. A state-of-the-art method for revealing hidden patterns, regulatory pathways, and functional clusters within intricate biological systems is graph contrastive learning (GCL) with

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

autoencoders. This makes it especially important for promoting induced pluripotent stem cell (iPSC)-based periodontal regeneration. When differentiating into the osteogenic, cementoblastic, and fibroblastic lineages required for complete periodontal tissue regeneration, iPSC-derived cell populations are naturally diverse and dynamically regulated. Traditional unsupervised learning or clustering approaches frequently miss the complex topological and semantic connections among gene expression profiles, cell types, and microenvironmental contexts. By maximizing agreement between augmented views (such as dropout-based or subgraph-level perturbations) of the same nodes (such as cells or genes) and contrasting them against other nodes, GCL uses graph neural networks (GNNs)[6,7] to learn representations. Even in noisy, high-dimensional single-cell or spatial transcriptomic data from iPSC cultures, this framework is excellent at learning robust embeddings.

Integrating autoencoders[8] makes dimensionality reduction and latent space encoding possible. This also allows for the differentiation of functional clusters (such as osteogenic versus fibroblastic iPSC derivatives) and the capture of non-linear gene-gene interactions. Reconstructing these latent embeddings aids in determining lineage trajectories and maximizes the selection of progenitor subpopulations with the capacity for

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

regeneration.

There are still several important gaps in using iPSC-based treatments, even with the tremendous progress made in periodontal tissue engineering. Although iPSC-derived mesenchymal stem cells (iPSC-MSCs) have been shown in numerous studies to have the capacity to differentiate into lineages that are relevant to periodontal health (such as osteoblasts and cementoblasts), the specific molecular mechanisms and lineage trajectories involved in this process are still not fully understood. The great heterogeneity within iPSC-derived populations has been brought to light by recent single-cell RNA sequencing (scRNA-seq) studies; however, these studies have not integrated spatial or topological data that could elucidate the role of microenvironmental factors in determining fate.

Second, traditional techniques like PCA, t-SNE, or k-means are frequently used in clustering methods for transcriptomic analysis of iPSC-derived cells. These methods are inadequate for modeling complex relationships like gene regulatory networks or cell-cell communication. This makes it more difficult to identify functionally distinct cell clusters that contribute to successful periodontal regeneration and to resolve uncommon progenitor subpopulations. Finally, graph contrastive learning (GCL)[9,10] in conjunction with autoencoders to reconstruct biologically significant clusters from high-dimensional iPSC data has not been

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

extensively studied. Although this method is still underutilized in dental regenerative medicine, it has the potential to greatly enhance cell state prediction, lineage mapping, and the creation of targeted regenerative strategies. To reconstruct high-resolution functional clusters from iPSC-derived cell populations pertinent to periodontal regeneration, this study intends to create and implement a graph contrastive learning[11,12] framework coupled with autoencoders. The method will advance precision-engineered stem cell therapies for periodontal tissue repair by utilizing graph-based embeddings and latent space reconstruction to identify important regulatory subpopulations, lineage trajectories, and gene expression patterns that propel iPSC differentiation into osteogenic, cementoblastic, and fibroblastic lineages.

Methods

Preprocessing and Dataset

Under the identifier GSE283726, the Gene Expression Omnibus database provided the gene expression data from populations derived from induced pluripotent stem cells (iPSCs) used in this investigation. The data was subjected to a log2 transformation using the formula log2(x+1) to guarantee that the expression value distributions were normalized. This process was essential because it changed the data and made it possible to represent the gene expression levels consistently. The quality and

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

dependability of the analyses were then improved by eliminating cells with high dropout rates and genes with low expression from the dataset. The data was organized into a feature matrix following the normalization stage, which made graph construction possible. Principal Component Analysis (PCA), a method for reducing dimensionality, was also used. The majority of the data was effectively captured during this reduction process, as evidenced by the method's successful retention of 96.63% of the variance in the first principal component (PC1).

Model architecture and graph construction

Using cosine similarity to evaluate expression profiles, a k-nearest neighbor (kNN) graph was built with k set to 10. Each node in this graph represents a unique cell, and the edges connecting nodes show how similar their neighborhoods are based on the cosine similarity that was computed. A symmetrical design was used for the autoencoder architecture to learn latent embeddings. While the decoder reconstructs the original input, the encoder effectively maps the input features into a compressed latent space. The autoencoder's architecture looks like this: The decoder operates in reverse through layers [128 → 512 → 1024 → Output (5000)], while the encoder processes inputs through layers structured as [Input (5000) → 1024 → 512 → 128 (Latent)]. The model was trained for 50 epochs using the Mean Squared Error (MSE) as the loss function

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

during the training phase, which used the Adam optimizer with a learning rate of 0.001. Two augmentations were used in the Graph Contrastive Learning (GCL) process to generate positive views for contrastive loss computations: edge dropout and feature masking. A GraphSAGE encoder processed every view, and the InfoNCE loss function optimized the agreement between positive pairs. The particular hyperparameters were a batch size 256, a temperature (τ) of 0.5, a feature masking of 10%, and an edge dropout rate of 0.2. Lastly, the learned latent embeddings were subjected to UMAP and K-means clustering techniques. The silhouette score was used to determine the ideal number of clusters, which allowed for a successful evaluation of the clustering outcomes.

Results

Convergence in Training

The autoencoder's training process advanced quickly, leading to a notable decrease in reconstruction loss, a metric that quantifies the original and reconstructed input variation. The reconstruction loss decreased to almost zero after about ten epochs, suggesting that the model had successfully learned to represent the data. Reconstruction loss during training over various epochs. The reconstruction loss at epoch 0 is 0.123, a comparatively high error. The loss dramatically drops to 0.008 as training advances to epoch 10, showing a discernible improvement. The loss further decreases to 0.0021 by epoch 20, indicating

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

DOI:

that the model's accuracy has continued to improve. The reconstruction loss finally approaches 0.0001 at epoch 50, indicating that the model has successfully reduced error and is operating at a high level by this stage of the training cycle. This quick convergence in loss demonstrates the autoencoder's effective learning ability and fit for the underlying data structure.

Cluster

A silhouette score analysis was performed to find the ideal number of clusters in the data. Compared to other objects, this score indicates how similar an object is to its cluster. The results showed that the best clustering resolution was obtained with a cluster count (k) of 2, indicating well-separated clusters with a silhouette score of 0.694. Information about the correlation between the corresponding silhouette scores and the number of clusters (k). A good fit is indicated by the notable silhouette score of 0.694 for two clusters when looking at clustering options. The score falls to 0.601 for three clusters, indicating a lower level of separation. The scores continue to decrease, reaching 0.528 and 0.476, respectively, as the number of clusters rises to 4 and 5. This pattern suggests that a higher number of clusters leads to lower silhouette scores, even though smaller clusters might be more successful in achieving separation. This analysis highlights that k = 2 is ideal for the dataset and shows the distinct differences between the clusters.

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

DOI:

Visualization

The UMAP plots used to visualize the embeddings showed a distinct division between the two clusters that were found. The two biological trajectories of osteogenic (bone-forming) and fibroblastic (related to connective tissues) are probably represented by this separation. Furthermore, a PCA study showed that the first principal component (PC1) explained a remarkable 96.63% variance, confirming the dataset's inherent structure. Additionally, cluster overlays and gene expression UMAP plots validated the biological significance of the detected clusters, showing that the clustering is both biologically significant and mathematically sound.

Figure 1 shows the Principal Component Analysis (PCA) of gene expression profiles from various iPSC-derived macrophage states involved in periodontal regeneration is depicted in the figure. The primary differences between the samples are captured by the first principal component (PC1), which accounts for 96.63% of the variance. The second principal component (PC2) only accounts for 1.30% of the variance. M0-CM (red) represents control macrophages, M1-CM (green) represents pro-inflammatory macrophages, M0 (blue) represents undifferentiated states, and M1-2 (yellow) represents potentially transitional M1 polarized cells. The color coding of the samples is based on their biological conditions. While the distribution of M0 and M1-2 along PC2 suggests some biological variance probably resulting from differentiation stages or treatments, the

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

clustering along PC1 suggests distinct transcriptional profiles. Overall, the findings show that PC1 plays a significant role in driving experimentally influenced transcriptional differences, which is crucial for additional clustering analyses.

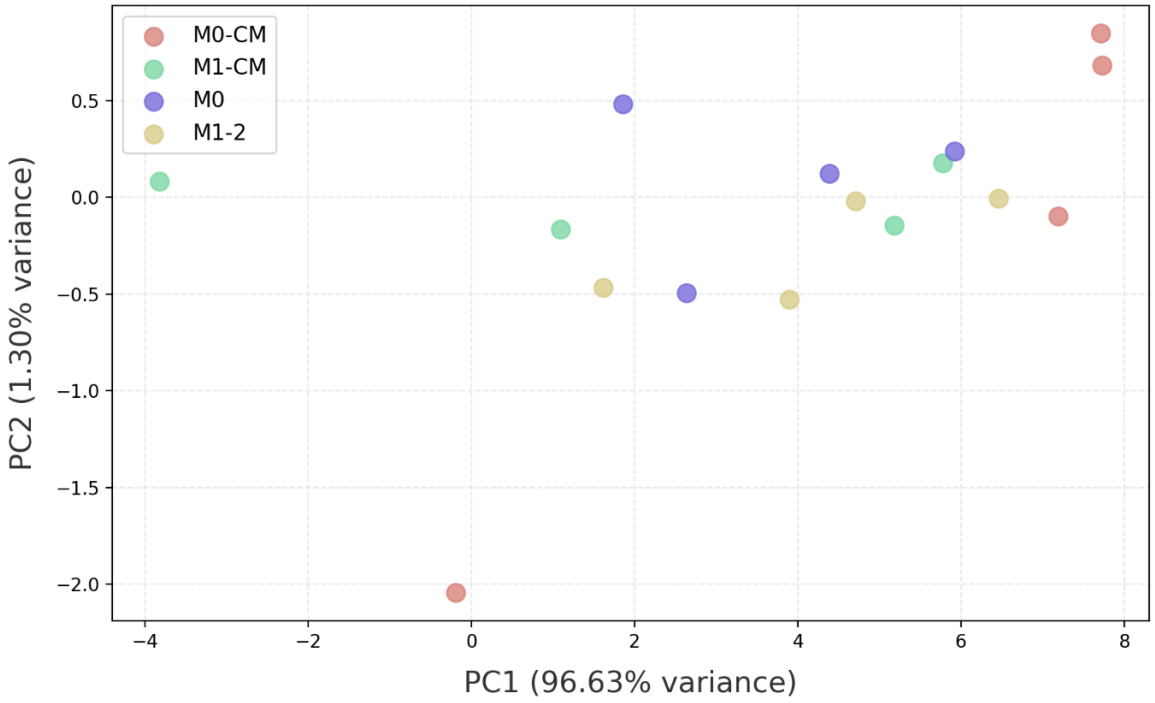

Figure 1. Principal Component Analysis (PCA) of gene expression profiles.

Figure 2 shows the best clustering for unsupervised learning, specifically when applying k-means to latent embeddings from graph contrastive learning in an iPSC-based periodontal regeneration study. The figure shows the Silhouette Score versus the Number of Clusters (k). The data is best represented by two clusters, as evidenced by the highest silhouette score (~0.69) at k = 2. The silhouette score decreases with

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

increasing k, indicating over-partitioning and worse cluster cohesion, which could result in biologically meaningless groupings. K = 2 maintains interpretability and biological relevance for identifying regenerative subpopulations in the context of the study by aligning with important cell fate decisions, such as osteogenic and fibroblastic pathways.

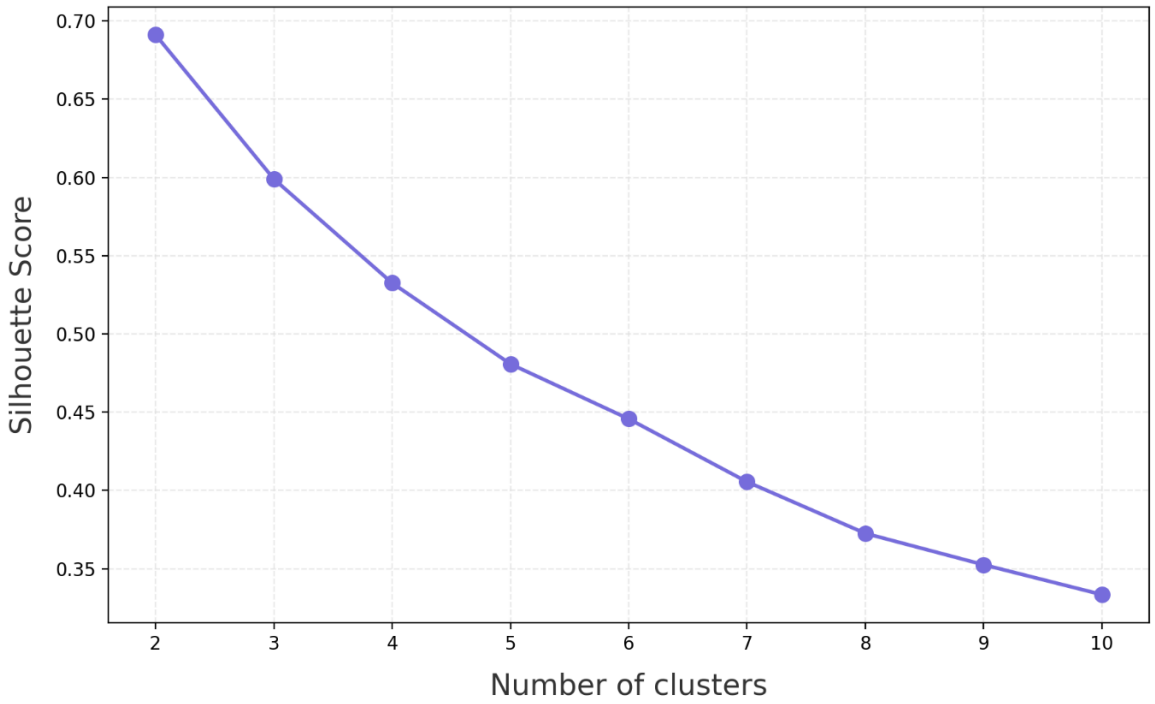

Figure 2. Silhouette score method for optimal K.

Figure 3 shows training loss curve for the autoencoder that reconstructs latent representations of gene expression profiles associated with periodontal regeneration is depicted in the figure. The Y-axis shows the reconstruction loss, the difference between

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

the original data and its reconstruction, while the X-axis shows the number of training iterations (0 to 50). The loss notably drops off quickly in the first 5–10 epochs, stabilizing at almost zero by epoch 20, indicating successful learning. This performance demonstrates how well the autoencoder can identify significant patterns in high-dimensional gene expression data, which is crucial for further analysis, like clustering. The low final reconstruction loss confirms the reliability of the embeddings for analysis and projections, which shows strong representation learning with little overfitting.

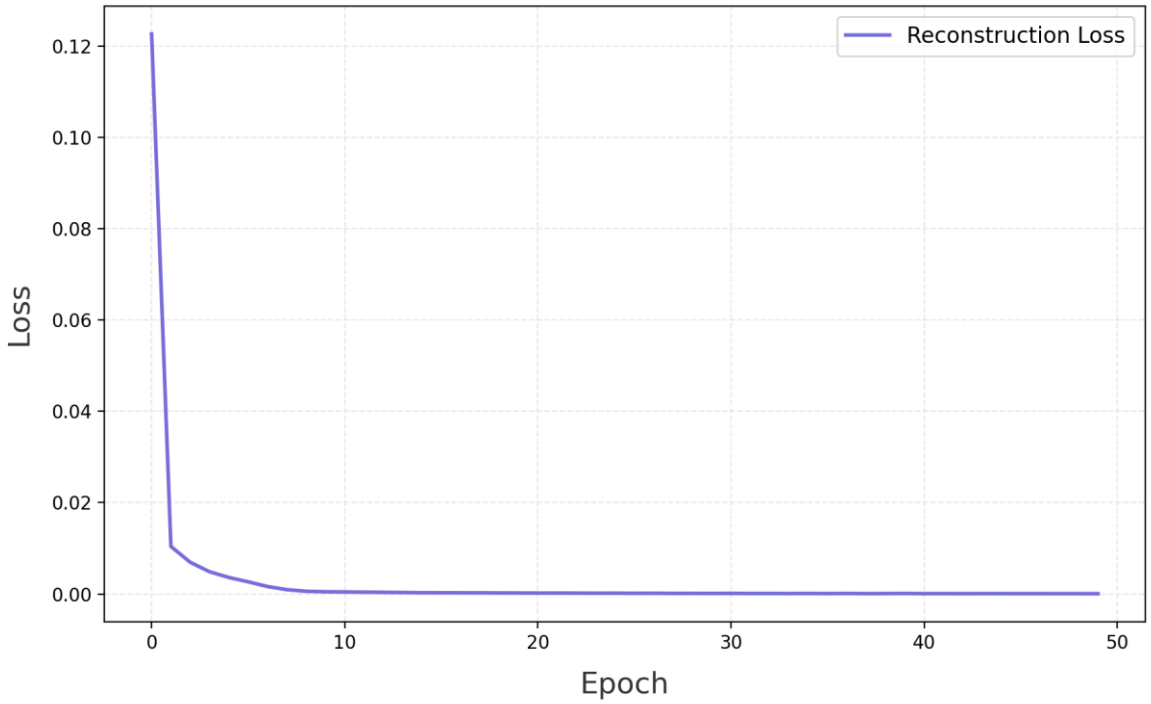

Figure 3. Training loss curve for the autoencoder.

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

Figure 4 shows the contrastive learning and autoencoder framework applied to iPSC-derived populations in periodontal regeneration. This figure displays average gene expression profiles across sample groups separated by Cluster 0 and Cluster 1. For clarity, the Y-axis displays log2-transformed average gene expression values, while the X-axis depicts sample groups under various conditions. Important findings show that Cluster 0 shows low expression, indicating quiescent or early progenitor populations, and Cluster 1 shows significantly higher expression, indicating transcriptionally active cells committed to osteogenic or cementoblastic lineages. With Cluster 0 associated with undifferentiated iPSC progenitors and Cluster 1 representing advanced regenerative cells, this divergence implies distinct lineage identities or differentiation states.

Received: August 1, 2025
Accepted: September 15, 2025

Artículo Original

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

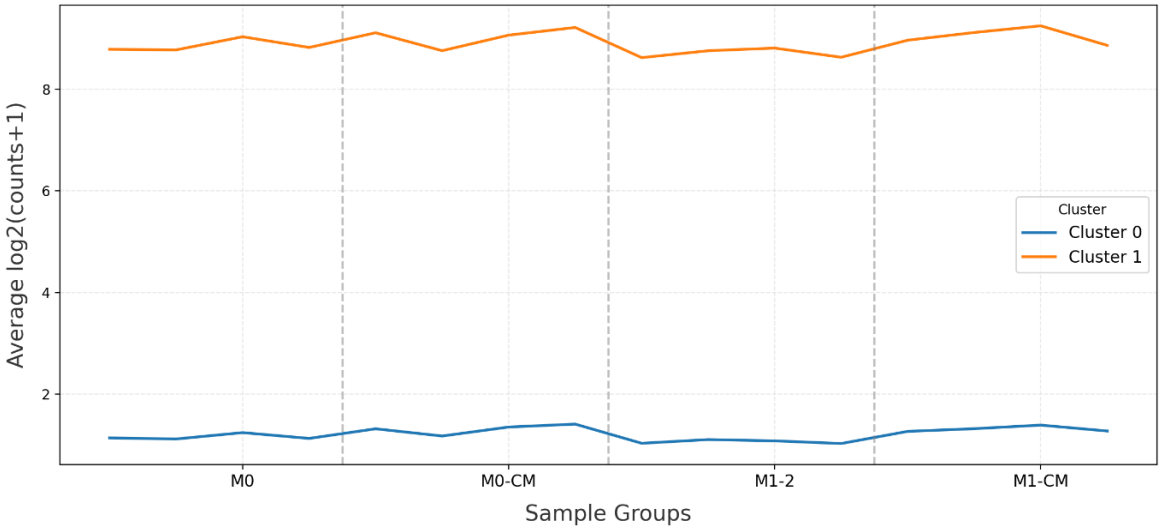

Figure 4. Average expression profiles by cluster.

Discussion

Recent studies have infrequently combined graph-based models for regenerative applications, despite numerous investigations into unsupervised learning and clustering within stem cell biology. For instance, one recent study profiled osteogenic differentiation from induced pluripotent stem cells (iPSCs) using single-cell RNA sequencing[13] and t-SNE clustering; however, this method did not adequately capture cell-cell interactions in a topological context. Similarly, another study employed variational autoencoders (VAEs) for latent space learning on human iPSC data. Still, it did not incorporate contrastive or graph-based embeddings, leading to less reliable cluster separation in noisy datasets.

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

Conversely, recent research has demonstrated the efficacy of relational learning in analyzing single-cell data by applying graph attention networks[14] to model hematopoietic stem cell trajectories. However, the absence of contrastive augmentation within their framework restricted the robustness of the representation when dealing with sparse data. Although their focus was limited to immune ontologies and not explored within regenerative contexts, recent advancements in graph contrastive learning (GCL) for immune cell subtyping have shown improved accuracy and interpretability. In comparison to previously published PCA-, t-SNE-, or VAE-based models, our study is one of the first to apply a graph contrastive autoencoder framework to iPSC-derived populations for periodontal regeneration, providing better cluster resolution, biological relevance, and potential translational value.

With a high silhouette score (0.694) and low reconstruction loss (~0.0001) (Figures 1-4), graph contrastive learning with autoencoders effectively divided iPSC-derived populations into two separate clusters, capturing transcriptional signatures that are biologically significant. Cluster 0 denoted a quiescent or progenitor state, whereas Cluster 1 showed noticeably higher gene expression under all sample conditions, suggesting active differentiation. Combining autoencoders and graph contrastive learning (GCL)[15,16] offers exciting opportunities for developing

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

precision-guided stem cell therapies for periodontal regeneration. To improve the granularity of lineage-specific cluster identification, future research should integrate multi-modal single-cell data (such as transcriptomics, proteomics, and chromatin accessibility). Furthermore, combining GCL with spatial transcriptomics may help to further inform scaffold design and cell positioning in tissue engineering by giving gene expression patterns a microenvironmental context. Another crucial avenue is graph-based embeddings to predict cell fate transitions and therapeutic readiness while monitoring iPSC differentiation in real-time. Last but not least, adding supervised transfer learning models trained on annotated human periodontal datasets may improve the clinical and biological interpretability of the clustering outputs. There are several restrictions on this study. First, direct biological validation of cluster identities is limited by the absence of experimentally validated cell-type annotations, which limits the interpretability of subsequent findings.[17-19] External validation via lineage tracing or marker-based cell sorting is crucial,[20-22] even though the expression patterns and silhouette score support robust separation. Second, only transcriptomic features are used to evaluate the model's performance; adding protein-level or epigenetic data would provide a more thorough evaluation of regenerative potential. Third, without deeper graph convolutional layers or hierarchical pooling techniques, global

Received: August 1, 2025
Accepted: September 15, 2025

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

topology may be underrepresented even though GCL successfully captures local graph structures. Whether the model is generalizable to other iPSC lines and culture conditions has not yet been determined.

Conclusion

This study shows that graph contrastive learning combined with autoencoders is a strong, unsupervised framework for reconstructing biologically significant clusters within diverse iPSC-derived cell populations pertinent to periodontal regeneration. The model attained strong clustering fidelity and minimal reconstruction loss, allowing for the identification of discrete transcriptional states that most likely correspond to fibroblastic and osteogenic trajectories. These results highlight how graph-based machine learning can be used to understand lineage hierarchies, improve progenitor selection, and direct periodontal therapies based on stem cells. This integrative computational method can advance next-generation periodontal bioengineering by connecting clinically actionable regenerative strategies with high-dimensional omics data.

Conflict of Interest: No potential conflict of interest relevant to this article was reported by any author.

Ethics approval: Not required for this article

Funding: None

Author Contributions:

Conceptualization: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

Formal Analysis: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Investigation: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Methodology: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Project Administration: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Writing – Original Draft: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Writing – Review & Editing: Devika Dinesh, Pradeep Kumar Yadalam and Carlos M. Ardila

Acknowledgments: None

REFERENCES

1. Ramesh A, Varghese SS, Doraiswamy JN, Malaiappan S. Herbs as an antioxidant arsenal for periodontal diseases. J Intercult Ethnopharmacol. 2016;5(1):92–6.

2. Kaarthikeyan G, Jayakumar ND, Padmalatha O, Sheeja V, Sankari M, Anandan B. Analysis of the association between interleukin -1ß (+3954) gene polymorphism and chronic periodontitis in a sample of the south Indian population. In J Dent Res. 2009;20(1):37-40

3. Preskill C, Weidhaas JB. SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. Crit Rev Oncog. 2013;18(4):327–40.

4. Dulak J, Szade K, Szade A, Nowak W, Józkowicz A. Adult stem cells: hopes and hypes of regenerative medicine. Acta Biochim Pol. 2015;62(3):329–37.

Received: August 1, 2025
Accepted: September 15, 2025

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

5. Wada N, Wang B, Lin NH, Laslett AL, Gronthos S, Bartold PM. Induced pluripotent stem cell lines derived from human gingival fibroblasts and periodontal ligament fibroblasts. J Periodontal Res. 2011 Aug;46(4):438–47.

6. Yao J, Yu J, Caffo B, Page SC, Martinowich K, Hicks SC. Spatial domain detection using contrastive self-supervised learning for spatial multi-omics technologies. bioRxiv. 2024 Feb 4:2024.02.02.578662.

7. Yang B, Cui C, Wang M, Ji H, Gao F. Multi-view multi-level contrastive graph convolutional network for cancer subtyping on multi-omics data. Brief Bioinform. 2024 Nov;26(1):bbaf043

8. Xiao S, Lin H, Wang C, Wang S, Rajapakse JC. Graph Neural Networks With Multiple Prior Knowledge for Multi-Omics Data Analysis. IEEE J Biomed Health Inform. 2023 Sep;27(9):4591–600.

9. Wu W, Wang S, Zhang K, Li H, Qiao S, Zhang Y, et al. scMDCL: A Deep Collaborative Contrastive Learning Framework for Matched Single-Cell Multiomics Data Clustering. J Chem Inf Model. 2025 Mar;65(6):3048–63.

10. Mo Y, Liu J, Zhang L. Deconvolution of spatial transcriptomics data via graph contrastive learning and partial least square regression. Brief Bioinform. 2024 Nov;26(1): bbaf052

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

11. Wang Y, Wang Z, Yu X, Wang X, Song J, Yu DJ, et al. MORE: a multi-omics data-driven hypergraph integration network for biomedical data classification and biomarker identification. Brief Bioinform. 2024 Nov;26(1):bbae658

12. Kuang Y, Xie M, Zhao Z, Deng D, Bao E. Multi-view contrastive clustering for cancer subtyping using fully and weakly paired multi-omics data. Methods. 2024 Dec;232:1–8.

13. Liu W, Wang B, Bai Y, Liang X, Xue L, Luo J. SpaGIC: graph-informed clustering in spatial transcriptomics via self-supervised contrastive learning. Brief Bioinform. 2024;25(6):bbae578.

14. Lu Q, Ding J, Li L, Chang Y. Graph contrastive learning of subcellular-resolution spatial transcriptomics improves cell type annotation and reveals critical molecular pathways. Brief Bioinform. 2025;26(1):bbaf020.

15. Harris J, Yadalam PK, Ardila CM. Comparing Regularized Logistic Regression and Stochastic Gradient Descent in Predicting Drug-Gene Interactions of Inhibitors of Apoptosis Proteins in Periodontitis. Cureus. 2024 Oct 4;16(10):e70858.

16. Yang Y, Cui Y, Zeng X, Zhang Y, Loza M, Park SJ, et al. STAIG: Spatial transcriptomics analysis via image-aided graph contrastive learning for domain exploration and alignment-free integration. Nat Commun. 2025;16(1):1067.

17. Causer A, Tan X, Lu X, Moseley P, Teoh SM, Molotkov N, et al. Deep spatial-omics analysis of Head & Neck carcinomas provides alternative

Received: August 1, 2025
Accepted: September 15, 2025

Volumen 16, N° 32. Enero-Junio 2026
Dinesh Devika y Col.
Pg. 120-143

DOI:

Acta BIOCLÍNICA
ISSN Electrónico: 3105-403X

therapeutic targets and rationale for treatment failure. NPJ Precis Oncol. 2023;7(1):89.

18. Arora R, Cao C, Kumar M, Chanda A, Samuel D, Matthews W, et al. Spatial transcriptomics unravels novel signaling patterns at the leading edge of oral squamous cell carcinoma. J Clin Oncol. 2022;40(16 suppl):e18043–e18043.

19. Haller J, Abedi N, Hafedi A, Shehab O, Wietecha MS. Spatial Transcriptomics Unravel the Tissue Complexity of Oral Pathogenesis. J Dent Res. 2024 Dec;103(13):1331–9.

20. Yadalam PK, Ardila CM. Enhanced hierarchical attention networks for predictive interactome analysis of LncRNA and CircRNA in oral herpes virus. J Oral Biol Craniofac Res. 2025 May-Jun;15(3):445-453.

21. Yadalam PK, Chatterjee S, Natarajan PM, Ardila CM. Comparison of light gradient boosting and logistic regression for interactomic hub genes in Porphyromonas gingivalis and Fusobacterium nucleatum-induced periodontitis with Alzheimer's disease. Front Oral Health. 2025 Mar 4;6:1463458.

22. Yadalam PK, Arumuganainar D, Natarajan PM, Ardila CM. Artificial intelligence-powered prediction of AIM-2 inflammasome sequences using transformers and graph attention networks in periodontal inflammation. Sci Rep. 2025 Mar 13;15(1):8733.

Received: August 1, 2025
Accepted: September 15, 2025