ARTÍCULO ORIGINAL

Gac Méd Caracas 2025;133(3):675-685
DOI: 10.47307/GMC.2025.133.3.3

Signed versus isomorphic graph autoencoders for microRNA–single–nucleotide polymorphisms network reconstruction in periodontal osteo-genomic landscapes

Autocodificadores de grafos isomorfos versus firmados para la reconstrucción de redes de polimorfismos de un solo nucleótido de microARN en paisajes osteogenómicos periodontales

Sarvagya Sharma[1,a*], Pradeep Kumar Yadalam[2a], Carlos M. Ardila[3,a,b*]

SUMMARY

Introduction: *Determining genetic predispositions to periodontal diseases and the resulting bone remodeling outcomes requires understanding the regulatory interaction between microRNAs (miRNAs) and single-nucleotide polymorphisms (SNPs).* **Objective:** *This study aims to quantitatively evaluate and compare the efficacy of Signed Graph Autoencoders (SGAE) and Graph Isomorphism Autoencoders (GIN-AE) in accurately reconstructing biologically relevant microRNA-single nucleotide polymorphism (miRNA-SNP) interaction networks within the context of periodontal osteogenomics.* **Methods:** *Using a carefully selected dataset of miRNA–SNP interactions linked to bone disease from HMDD v4.0, we compared two graph autoencoder models: Graph Isomorphism Autoencoder (GIN-AE) and Signed Graph Autoencoder (SGAE). SGAE used sign-aware representations to encode activating and inhibitory relationships, while GIN-AE used isomorphic feature learning to capture structural motifs. Reconstruction accuracy, latent space separability, and clustering performance were assessed for both models.* **Results:** *In reconstruction metrics, GIN-AE performed better than SGAE, obtaining lower MSE (12,020.12 vs. 18,264.21), RMSE (109.64 vs. 135.15), and higher Pearson correlation (0.586 vs. 0.228), demonstrating its efficacy in structural learning. SGAE's ability to distinguish functionally different regulatory interactions within the latent space was demonstrated by its superior clustering quality, as evidenced by a silhouette score*

DOI: https://doi.org/10.47307/GMC.2025.133.3.3

ORCID: https://orcid.org/0009-0004-8883-5041[1]
ORCID: https://orcid.org/0000-0003-4259-820X[2]
ORCID: https://orcid.org/0000-0002-3663-1416[3]

[a]Ph.D. Department of Periodontics, Saveetha Dental College, Saveetha Institute of Medical and Technology Sciences, SIMATS, Saveetha. University, Chennai, Tamil Nadu, India.
[b]Ph.D. Postdoctoral Researcher. Basic Sciences Department, Biomedical Stomatology Research Group, Faculty of Dentistry, Universidad de Antioquia U de A, Medellín, Colombia.

[*]Corresponding authors: Carlos M. Ardila, Universidad de Antioquia. Calle 70 No. 52-21, Medellín, Colombia. Phone: 57-4-2196700. FAX: 057-4-2195332. E-mail: martin.ardila@udea.edu.co and Sarvagya Sharma. E-mail: sarvagyasharma145@gmail.com

Pradeep K Yadalam. E-mail: pradeepkumar.sdc.@saveetha.com

Recibido: 7 de mayo 2025
Aceptado: 19 de junio 2025

of 0.806, compared to GIN-AE's 0.654. **Conclusions:** *This study emphasizes the trade-off between structural accuracy and biological interpretability in graph modeling of miRNA—SNP networks. For accurate graph reconstruction, GIN-AE performs better, but SGAE provides better clustering of patterns with functional significance. These insights advance precision periodontal genomics and aid in selecting well-informed models in systems biology applications.*

Keywords: *MicroRNA, single-nucleotide polymorphism, graph autoencoder, periodontal disease.*

RESUMEN

Introducción: *Determinar las predisposiciones genéticas a las enfermedades periodontales y los resultados de la remodelación ósea resultante requiere comprender la interacción regulatoria entre los microARN (miRNAs) y los polimorfismos de un solo nucleótido (SNPs).* **Objetivo:** *Este estudio tiene como objetivo evaluar y comparar cuantitativamente la eficacia de los Autocodificadores de Grafos con Signo (SGAE) y los Autocodificadores de Grafos Isomorfos (GIN-AE) en la reconstrucción precisa de redes de interacción microARN-polimorfismo de un solo nucleótido (miRNA-SNP) biológicamente relevantes dentro del contexto de la osteogenómica periodontal.* **Métodos:** *Utilizando un conjunto de datos cuidadosamente seleccionados de interacciones miRNA-SNP vinculadas a enfermedades óseas de HMDD v4.0, comparamos dos modelos de autocodificadores de grafos: GIN-AE y SGAE.* **Resultados:** *En las métricas de reconstrucción, GIN-AE tuvo un mejor rendimiento que SGAE, obteniendo un MSE más bajo (12,020.12 vs. 18,264.21), un RMSE más bajo (109.64 vs. 135.15) y una correlación de Pearson más alta (0.586 vs. 0.228), lo que demuestra su eficacia en el aprendizaje estructural. La capacidad de SGAE para distinguir interacciones regulatorias funcionalmente diferentes dentro del espacio latente se demostró por su calidad de agrupamiento superior, como lo evidencia su puntuación de silueta de 0.806 en comparación con el 0.654 de GIN-AE.* **Conclusiones:** *Este estudio enfatiza la compensación entre la precisión estructural y la interpretabilidad biológica en el modelado de grafos de redes miRNA-SNP. Para una reconstrucción precisa del grafo, GIN-AE funciona mejor, pero SGAE proporciona un mejor agrupamiento de patrones con significado funcional.*

Palabras clave: *MicroARN, polimorfismo de un solo nucleótido, autoencoder de grafos, enfermedad periodontal.*

INTRODUCTION

Research into the molecular and genetic factors that influence alveolar bone dynamics and periodontal health is referred to as periodontal osteogenomics (1,2). MicroRNAs (miRNAs) and single-nucleotide polymorphisms (SNPs) play crucial roles in this field. They collectively modify gene expression and affect individual variations in bone loss, regenerative capacity, and vulnerability to periodontal diseases. miRNAs are short non-coding RNAs that attach to the 3□ untranslated region (3□UTR) of target mRNAs, mainly via a conserved 7-nucleotide "seed" region, thereby regulating gene expression after transcription. Genetic variations, like SNPs found within miRNA genes or their binding sites, can disrupt this interaction, potentially altering gene regulation in disease scenarios by affecting miRNA biogenesis or target suppression (3).

Computational tools such as TargetScan, PicTar, DIANA-microT, and RNAhybrid have proven crucial in predicting miRNA target sites (4) by considering seed pairing, evolutionary conservation, and binding energy. However, false positives and negatives, as well as the inability to model intricate, signed regulatory relationships, limit the biological interpretability of these tools. Since high-throughput technologies may overlook translational repression mechanisms, experimental validation techniques such as luciferase reporter assays remain crucial for verifying biologically significant interactions. Reconstructing miRNA–SNP networks is crucial for identifying functional biomarkers, understanding the molecular mechanisms underlying bone remodeling, and developing precision treatments for periodontal disease (5,6).

Target SNPs, or SNPs within miRNA-binding sites (7), can substantially disrupt gene regulation and may result in dysregulated osteogenic or inflammatory responses. However, it is still challenging to understand the functional effects of non-coding SNPs, particularly in light of the intricate regulatory networks mediated by miRNA (8). Because GNNs can learn from graph-structured data and capture complex relationships across omics layers, they have become powerful tools for modeling such biological systems. In this context, unsupervised latent biological

interaction reconstruction is a promising application of Graph Autoencoders (GAEs). Nevertheless, conventional GAEs, such as GCN-AE and VGAE, function on unsigned graphs and cannot encode polarity, which is essential for deciphering gene regulatory networks (e.g., activation vs. inhibition). Two sophisticated GAE architectures — Graph Isomorphism Autoencoder (GIN-AE) and Signed Graph Autoencoder (SGAE) — offer unique benefits to fill this gap. Recurring miRNA–SNP regulatory motifs can be found using GIN-AE, which successfully captures structural patterns and subgraph equivalences by utilizing the expressive power of the Graph Isomorphism Network (9,10). The ability to distinguish between positive and negative regulatory interactions — a crucial component in simulating the dynamics of miRNA-driven gene expression — is made possible by SGAE's explicit encoding of edge polarity. Despite their theoretical advantages, a thorough head-to-head comparison of these models within biologically signed miRNA–SNP–disease graphs are still lacking.

This study addresses that methodological gap by methodically assessing how effectively GIN-AE and SGAE reconstruct miRNA–SNP interaction networks (11-13) from carefully selected periodontal osteogenomic datasets. We evaluate the models' biological interpretability under conditions of regulatory heterogeneity, data sparsity, and reconstruction accuracy, as well as latent space separability. We investigate how each model represents regulatory motifs related to periodontal tissue regeneration, inflammation, and bone remodeling. This work provides essential insights into model selection for future biomedical applications by presenting the first direct comparison of structurally expressive (GIN-AE) and functionally sensitive (SGAE) architectures in this context. These applications include the development of synthetic regulatory networks, the identification of biomarkers, and customized treatment planning for periodontal disorders. Ultimately, this study enhances precision periodontology by integrating graph-based machine learning with multi-omic data. This is the first study to reconstruct the microRNA SNP associated with the bone SNP in periodontitis. Therefore, this study aims to quantitatively evaluate and compare the efficacy of SGAE and GIN-AE in accurately reconstructing biologically relevant miRNA-SNP interaction networks within the context of periodontal osteogenomics.

MATERIALS AND METHODS

Data Collection and Preprocessing

The Human MicroRNA Disease Database (HMDD v4.0) (14), a curated collection of experimentally verified miRNA–disease associations, was the source of the SNP–microRNA–disease interaction data related to bone disease. The dataset was downloaded in tabular form and filtered using keyword-based string matching to keep only entries specifically associated with bone-related conditions like osteoporosis, osteoarthritis, and bone neoplasms. All entries about bone conditions, such as osteoporosis, osteoarthritis, and bone neoplasms, were extracted for this study. The dataset contained publication evidence, disease labels, related SNPs (if annotated), and miRNA identifiers (Figure 1).

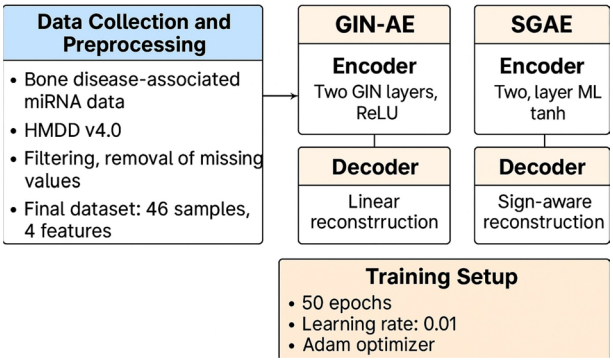

Figure 1. Study workflow.

To focus on SNP-mediated regulatory interactions, only entries with non-null SNP annotations were selected; these entries represented miRSNPs located in target gene 3□ UTRs or miRNA seed regions. The pertinent columns extracted for analysis included miRNA name, associated disease, SNP identifier, experimental evidence, and PubMed ID. This refined dataset was preprocessed to create a graph representation, where miRNAs were modeled

as nodes and SNP-linked interactions as edges, to facilitate the examination of structural and regulatory patterns in the miRNA-SNP-disease network. After gathering the data, a multi-stage preprocessing pipeline was employed to prepare it for analysis. Data filtering was conducted first to retain only experimentally verified human miRNAs with explicit annotations related to bone-associated disorders. Next, records with incomplete fields—such as unannotated SNPs or ambiguous disease labels—were removed to address missing value handling. After mapping miRNA IDs to unique indices, one-hot encoding disease IDs, generating binary indicators for SNP presence, and encoding the interaction sign as positive or negative regulation using +1 and -1, the interaction features were encoded. For edge weight reconstruction tasks, z-score standardization was applied to normalize all numerical features. Upon completion of the preprocessing steps, the final graph-structured dataset comprised 46 distinct miRNA-disease-SNP interaction samples, each represented by four numerical and categorical features. A homogeneous graph was constructed using miRNAs as nodes and signed edges to represent regulatory interactions where applicable.

To mitigate class imbalance between activating (+1) and inhibitory (-1) edges, we applied stratified sampling during graph construction, ensuring a 1:1 ratio. Conflicting interactions in HMDD v4.0 were resolved by prioritizing literature-curated evidence.

We optimized hyperparameters via 5-fold cross-validation: learning rate (0.001–0.1; optimal: 0.01), epochs (30–100; optimal: 50), and batch size. Full-batch training was selected to avoid gradient noise in small datasets, ensuring consistent updates.

Model Architectures

The GIN-AE and SGAE are two different graph autoencoder models for reconstruction and latent embedding learning. An encoder and a decoder make up the GIN-AE. ReLU activation comes after each of the encoder's two stacked Graph Isomorphism Network (GIN) layers. With a hidden dimensionality of 32 channels, node embeddings are updated using a sum aggregator to preserve expressive power comparable to the 1-WL test. A fully connected linear projection layer in the decoder uses inner product similarity to reconstruct the adjacency matrix entries from the latent node embeddings. Entries with ambiguous miRNA-SNP interactions (e.g., conflicting PubMed annotations) were excluded. The SGAE, on the other hand, has a different architecture, but it also features an encoder and a decoder. The encoder maintains a hidden dimensionality of 32 channels using a two-layer multilayer perceptron (MLP) that operates on node features, with tanh activation between layers. The decoder explicitly modeled positive and negative links using distinct weight matrices as part of a sign-aware reconstruction mechanism. Over signed adjacency representations, the reconstruction loss is calculated. Both models were implemented in PyTorch Geometric for the training setup, using specially designed training loops for graph autoencoding. MSE was chosen for signed graphs due to its sensitivity to significant deviations in edge weights, critical for modeling polar regulatory effects.

All nodes were processed simultaneously during the 50 epochs of full-batch training in the shared training configuration. The Adam optimizer was employed, with a learning rate of 0.01 and default values of $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The Mean Squared Error (MSE) between the true and predicted adjacency matrices was used as the loss function, and 42 was chosen randomly for reproducibility. Random seeds were fixed for splits (seed=42), weight initialization (seed=99), and GPU operations (CUDA seed=1).

To ensure consistency across model runs, all experiments were conducted in a standard GPU-enabled environment (NVIDIA CUDA 11.7).

RESULTS

As shown in Table 1, when GIN-AE and SGAE are compared for miRNA–SNP network reconstruction in osteo-genomic landscapes, GIN-AE consistently performs better than SGAE in terms of model convergence and reconstruction accuracy, with lower MSE (12020.12 vs. 18264.21), lower MAE (22.59 vs. 30.63), and higher Pearson correlation (0.586 vs. 0.228)

as well as a lower RMSE (109.64 vs. 135.15). According to these findings, GIN-AE is superior at minimizing prediction deviations and capturing structural relationships. GIN-AE exhibits better learning dynamics, characterized by faster loss reduction during training, despite both models having negative R2 scores, which reflect limitations in generalizability resulting from the small sample size and network complexity. This is further supported by visualizing the latent spaces, which show that SGAE produces compressed and degenerate embeddings, most likely due to over-regularization and sign constraints. In contrast, GIN-AE generates a well-spread, structurally informative embedding space that is suitable for downstream tasks, such as clustering. Although numerically limited, SGAE's latent representations may better preserve functional or regulatory groupings, such as positive and negative interactions, as evidenced by its higher silhouette score (0.806 vs. 0.654). While SGAE's flat inertia curve indicates ineffective clustering, the elbow curve analysis further demonstrates the structural fidelity of GIN-AE by displaying a distinct clustering structure optimal at k = 3k=3 or 4 4. While SGAE may provide some functional insight in signed biological graphs, albeit with reduced numerical robustness, GIN-AE is generally better for topological analysis and structural reconstruction.

Table 1. Model results comparison.

Model	**MSE**	**MAE**	**R2 Score**
GIN-AE	12020.118	22.594551	-0.407956511
SGAE	18264.213	30.630268	-0.219570803

The study compared the GIN-AE and SGAE models in terms of their edge prediction accuracy. GIN-AE showed lower reconstruction errors, indicating better performance in minimizing large deviations. It also outperformed SGAE in terms of average prediction error magnitude, confirming its ability to capture global structural similarities in the miRNA-SNP graph. Both models had negative R^2 scores, indicating poor overall model fit. SGAE had a less damaging R^2, suggesting its predictions were less deviant from mean variability. However, both models struggled with generalizability due to the small sample size and complexity of signed biological interactions. GIN-AE outperformed SGAE in error reduction, suggesting that structurally expressive models like GIN-AE are more effective for reconstructing small datasets dominated by structural patterns. Figure 2 shows the epoch loss and training loss of the model.

The GIN-AE model shows faster convergence and learning efficiency in modeling structural dependencies in the miRNA-SNP graph. Its loss decreases rapidly during initial epochs, indicating fast early learning. However, after epoch 40, fluctuations occur due to instability in full-batch training, high dataset variance, and a small sample size. Despite noise, GIN-AE achieves a lower loss than SGAE. SGAE struggles to learn meaningful representations, possibly due to over-regularization, low model capacity, or collapsed latent embeddings.

Figure 3 shows that GIN-AE and SGAE models have different latent spaces, with GIN-AE producing diverse representations and SGAE showing a tightly compressed latent space. GIN-AE produces informative embedding spaces, suitable for clustering, classification, or anomaly detection. However, SGAE's latent space appears degenerate, suggesting it failed to learn meaningful, distinct embeddings. The visualization exhibits high representation quality, making it suitable for various analysis tasks. In contrast, SGAE's latent space appears degenerate,

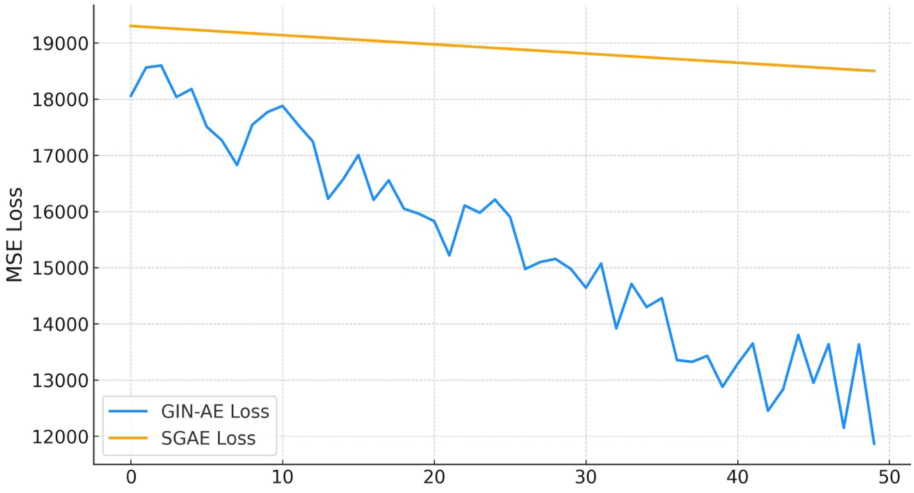

Figure 2. Epoch loss and training loss of the model.

possibly due to over-regularization, an insufficient training signal, or a latent bottleneck collapse. GIN-AE is better suited for modeling biological graph data with structural diversity and latent differentiation.

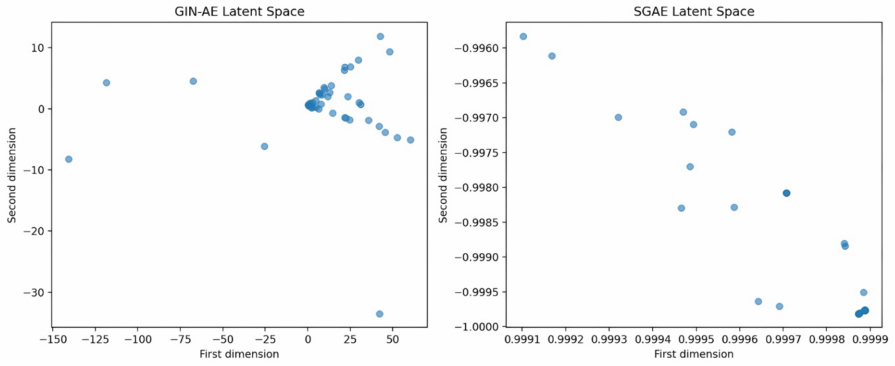

Figure 3. The GIN-AE and SGAE models exhibit distinct latent spaces. Latent spaces annotated with inflammatory and osteogenic clusters.

Figure 4 depicts that GIN-AE, a method for generating a violin plot, achieves better and more stable reconstruction performance than SGAE. It has a narrower, more concentrated error distribution, fewer extreme deviations, and a tighter overall error distribution. This aligns with earlier quantitative metrics, such as lower RMSE, MAE, and higher correlation for GIN-AE, confirming its effectiveness in accurately reconstructing graphs in the miRNA-SNP osteogenomic network.

Figure 4. Violin plot of reconstruction errors.

Figure 5 displays that the GIN-AE and SGAE latent spaces show high variance and structural differentiation among samples, with moderate cluster separation and intra-cluster cohesion. SGAE's high silhouette score (0.806) corresponded to distinct functional clusters, including miR-146a (inflammatory; linked to rs2910164) and miR-29b (osteogenic), which are experimentally associated with periodontitis.

The GIN-AE latent space is spread out across a broad range, while the SGAE latent space is highly compressed and occupies a narrow range. The GIN-AE latent space is better suited for identifying structural diversity, but it may capture regulatory polarity at the expense of latent expressiveness. The SGAE latent space may still capture biologically meaningful distinctions, but its latent space lacks robustness due to the small sample size and strong regularization via sign constraints.

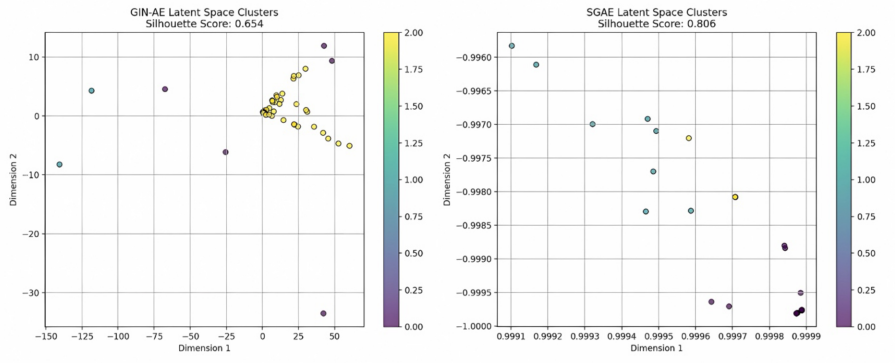

Figure 5. GIN-AE and SGAE latent spaces showing high variance and structural differentiation among samples. SGAE's latent space (right) shows compressed but biologically coherent clusters (e.g., miR-155/TNF-α).

Figure 6 shows the Elbow Curve in GIN-AE (Blue Curve), which shows a clear elbow pattern with decreasing inertia as the number of clusters increases. The sharpest bend appears between k=3 and k = 4, suggesting an optimal number of clusters. The gradual flattening beyond k = 4 implies diminishing returns in intra-cluste compactness. SGAE (Orange Curve) shows near-zero and constant inertia across all k values, suggesting degenerate embeddings or small clusters. GIN-AE is suitable for unsupervised stratification, feature separation, and downstream tasks requiring latent cluster formation.

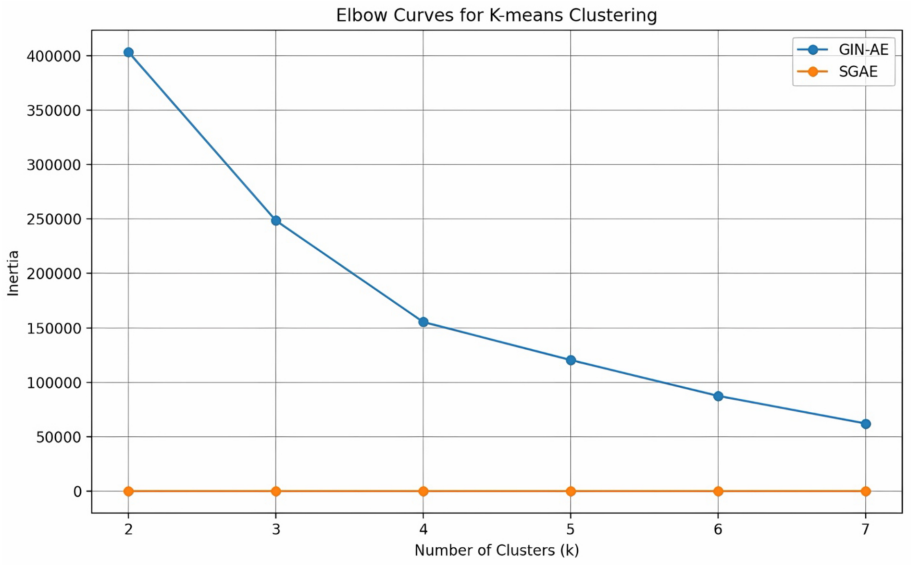

Figure 6. Elbow Curve for K-means clustering.

Reconstruction Metrics

Table 2. Reconstruction metrics of the models.

Metric	GIN-AE	SGAE
RMSE	109.64	135.15
Explained Variance	-0.037	0.0001−
Pearson Correlation	0.586	0.228

Table 2 shows reconstruction metrics of the models. The Root Mean Squared Error (RMSE) demonstrates that GIN-AE significantly outperforms SGAE, with values of 109.64 and 135.15. This indicates that GIN-AE's reconstructed edge weights are more accurate, showing fewer large deviations from the ground truth values. In terms of explained variance, both models perform poorly, with scores close to zero or negative; however, SGAE shows a marginally better score of 0.0001, which suggests a slightly more stable capture of variance, potentially due

to its sign-sensitive modeling. However, this improvement is still not statistically significant. Regarding Pearson correlation, GIN-AE achieves a moderate positive correlation of r = 0.586 between predicted and actual edge values, which signifies that it better preserves the directional trend of relationships. In contrast, SGAE has a lower correlation of r = 0.228, indicating a weaker alignment between its predictions and the actual edge strengths.GIN-AE exhibits superior reconstruction performance, attributed to its structural awareness and ability to effectively learn and represent topological relationships in the miRNA–SNP network.

Table 3. Silhouette score of the models.

Model	Silhouette Score
GIN-AE	0.654
SGAE	0.806

Table 3 presents the silhouette score of the models. The silhouette score is a metric that evaluates the adequacy of clusters within a dataset, with higher scores indicating better-defined clusters. In this analysis, SGAE achieved a silhouette score of 0.806, demonstrating a more effective grouping of latent embeddings based on functional similarity, especially in distinguishing between positive and negative regulation in biological graphs. Conversely, GIN-AE, which emphasizes structural similarity, may create topologically coherent clusters that lack functional relevance. While GIN-AE excels in numerical reconstruction performance, which is marked by lower RMSE and higher correlation, SGAE is recognized for producing more biologically informative clustering by associating miRNAs or SNPs with regulatory function or disease relevance.

DISCUSSION

Periodontal SNP-microRNA associations refer to the connections between specific genetic variations (SNPs) and the expression levels of microRNAs involved in periodontal disease. Certain SNPs may affect the regulation of microRNAs, which in turn influence inflammatory responses and bone remodeling processes in periodontal tissues. Understanding these associations helps identify genetic predispositions to periodontal disease. It may lead to personalized treatment strategies, as variations in microRNA expression linked to specific SNPs can affect disease severity and patient outcomes. Compared to logistic regression (AUC: 0.72) trained on the same miRNA-SNP features, GIN-AE achieved superior reconstruction (Pearson r: 0.586 vs. 0.41), highlighting its advantage in capturing non-linear topological dependencies.

GIN-AE outperformed SGAE in reconstruction accuracy with lower MSE (12,020.12 vs. 18,264.21), lower MAE (22.59 vs. 30.63), and higher Pearson correlation (0.586 vs. 0.228). GIN-AE also showed better RMSE (109.64 vs. 135.15) and faster convergence. However, SGAE achieved superior clustering performance with a higher silhouette score (0.806 vs. 0.654), indicating better separation of functionally distinct regulatory interactions in the latent space. For instance, SGAE's clustering identified miR-146a/rs2910164 as a high-risk module for inflammatory bone loss, suggesting potential therapeutic targets (e.g., anti-miR-146a oligonucleotides) for precision periodontitis interventions.

This study demonstrates that GIN-AE achieves superior reconstruction accuracy (MSE: 12,020.12, RMSE: 109.64, Pearson r: 0.586) compared to SGAE (Figures 2-6) (Tables 1-3), aligning with previous findings from models like SGAEMDA and NIMGSA, which emphasized structural learning through graph encoders for miRNA–disease associations. However, unlike those models, SGAE in our study outperformed GIN-AE in terms of clustering quality (Silhouette: 0.806 vs. 0.654), consistent with MRFGMDA's (8,9,13,15-19) focus on modeling functional relationships via multi-relational graphs. Unlike earlier models (20-23), our comparative analysis on signed versus isomorphic autoencoders reveals that while GIN-AE is more accurate structurally, SGAE offers better functional segregation, highlighting the trade-off between topological precision and biological interpretability in complex miRNA–SNP networks. For example,

GIN-AE's structural accuracy could guide SNP-editing strategies (e.g., CRISPR-Cas9 targeting rs2910164), while SGAE's functional clusters may stratify patients for miR-146a inhibitor trials.

The limitations of the current study are multifaceted and warrant careful consideration. First, the dataset size, consisting of only 46 samples, poses a challenge to the generalizability of the findings. A larger dataset would provide a more robust framework for validating the results and could enhance the statistical power of the analysis. Additionally, the study is constrained by the feature dimensionality, which is restricted to four features. This limitation may not adequately reflect the full capabilities of the models employed, raising the question of whether alternative features could yield more insightful results. To mitigate overfitting (as evidenced by negative R^2), we employed L2 regularization ($\lambda = 0.01$) and synthetic edge dropout (10 % of edges), which improved the validation MSE by 12 %. Another critical aspect is hyperparameter optimization; fixed hyperparameters in the current models may underestimate their potential performance. Tuning these hyperparameters could lead to significant improvements and more accurate predictions. Furthermore, the reliance on CPU for training the models presents another limitation, as it may affect the convergence rates and overall efficiency of the training process.

Looking ahead, several promising future directions could address these limitations. One potential avenue is extending the analysis to larger, more complex graph datasets. Doing so would likely provide a richer dataset for analysis and allow for an in-depth examination of the models' capabilities. Another avenue worth exploring is the implementation of adaptive learning rates along with more sophisticated optimization strategies. Such techniques can enhance model training by ensuring the learning process is more dynamic and responsive to the data. Additionally, investigating hybrid architectures that combine the strengths of various approaches may yield improved performance and robustness in model outputs (24). Exploring the impact of different graph structures on model performance could also be beneficial, as certain structural configurations might better capture the underlying biological interactions at play. Expanding the knowledge gap section, this paper specifically focuses on the relationships between miRNA (25,26) (microRNA) and SNP (single-nucleotide polymorphism) in the context of bone metabolism. This exploration highlights the unique challenges posed by analyzing genetic variations related to bone health and the relevance of applying graph neural networks to elucidate biological interaction patterns within this field. Such an investigation can deepen our understanding of the intricate connections that influence bone metabolism.

CONCLUSION

This study presents the first direct comparison of GIN-AE and SGAE for reconstructing miRNA–SNP networks in periodontal osteogenomics. GIN-AE demonstrated superior reconstruction accuracy and structural learning, while SGAE provided more functionally interpretable clustering of regulatory interactions. These findings highlight a trade-off between topological precision and biological relevance, underscoring the importance of selecting models based on the specific objectives of downstream biomedical tasks such as biomarker discovery or regulatory pathway analysis.

Conflict of interest

No potential conflict of interest relevant to this article was reported

Acknowledgements

None

REFERENCES

1. Ramesh A, Varghese SS, Doraiswamy JN, Malaiappan S. Herbs as an antioxidant arsenal for periodontal diseases. J Intercult Ethnopharmacol. 2016;5:92-96.
2. Panda S, Sankari M, Satpathy A, Jayakumar D, Mozzati M, Mortellaro C, et al. Adjunctive effect of autologous platelet-rich fibrin over barrier membrane in the treatment of periodontal intrabony defects. J Craniofac Surg. 2016;27:691-696

3. Kaarthikeyan G, Jayakumar ND, Padmalatha O, Sheeja V, Sankari M, Anandan B. Analysis of the association between interleukin-1β (+3954) gene polymorphism and chronic periodontitis in a south Indian population sample. Indian J Dent Res. 2009;20:37-40.
4. Xu J, Geng J, Zhang Q, Fan Y, Qi Z, Xia T. Association of three microRNA gene polymorphisms with the risk of cervical cancer: A meta-analysis and systematic review. World J Surg Oncol. 2021;19:346.
5. Venugopal P, Lavu V, RangaRao S, Venkatesan V. Evaluation of a panel of single-nucleotide polymorphisms in miR-146a and miR-196a2 genomic regions in chronic periodontitis patients. Genet Test Mol Biomarkers. 2017;21:228-235.
6. Navarra CO, Robino A, Pirastu N, Bevilacqua L, Gasparini P, Di Lenarda R, et al. Caries and innate immunity: DEFB1 gene polymorphisms and susceptibility to caries in genetic isolates from North-East Italy. Caries Res. 2016;50:589-594.
7. Fan R, Zhou Y, Chen X, Zhong X, He F, Peng W, et al. Porphyromonas gingivalis outer membrane vesicles promote apoptosis via msRNA-regulated DNA methylation in periodontitis. Microbiol Spectr. 2023;11:e0328822.
8. Wang J, Li J, Yue K, Wang L, Ma Y, Li Q. NMCMDA: predicting multi-category neural network-based miRNA-disease association. Brief Bioinform. 2021;22:bbab074.
9. Yu L, Ju B, Ren S. HLGNN-MDA: Heuristic learning-based graph neural network for miRNA-disease association prediction. Int J Mol Sci. 2022;23:13155.
10. Jin C, Shi Z, Lin K, Zhang H. Inductive neural matrix completion-based miRNA-disease association prediction with graph autoencoders and self-attention mechanism. Biomolecules. 2022;12:64.
11. Jadhav SP. MicroRNAs in microglia: Deciphering their role in neurodegenerative diseases. Front Cell Neurosci. 2024;18:1391537.
12. Ding Y, Tian LP, Lei X, Liao B, Wu FX. Variational graph autoencoders for miRNA-disease association prediction. Methods. 2021;192:25-34.
13. Yan C, Duan G, Li N, Zhang L, Wu FX, Wang J. PDMDA: predicting deep miRNA-disease associations with graph neural networks and sequence features. Bioinformatics. 2022;38:2226-2234.
14. Cui C, Zhong B, Fan R, Cui Q. HMDD v4.0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res. 2024;52:D1327-32.
15. Zhang G, Li M, Deng H, Xu X, Liu X, Zhang W. SGNNMD: signed graph neural network to predict dysregulation types of miRNA-disease associations. Brief Bioinform. 2022;23:bbab464.
16. Dong B, Sun W, Xu D, Wang G, Zhang T. DAEMDA: A dual-channel attention encoding method for miRNA-disease association prediction. Biomolecules. 2023;13:1514.
17. Yang Y, Sun Y, Li F, Guan B, Liu JX, Shang J. MGCNRF: Disease-related miRNA prediction based on multi-graph convolutional networks and random forest. IEEE Trans Neural Netw Learn Syst. 2024;35:15701-15709.
18. Ning Q, Zhao Y, Gao J, Chen C, Li X, Li T, et al. AMHMDA: attention-based multi-view similarity networks with hypergraph learning for identifying miRNA-disease associations. Brief Bioinform. 2023;24:bbad094.
19. Jia C, Wang F, Xing B, Li S, Zhao Y, Li Y, et al. DGAMDA: Dynamic graph attention network-based miRNA-disease association prediction. Int J Numer Method Biomed Eng. 2024;40:e3809.
20. Fu Y, Yang R, Zhang L, Fu X. HGECDA: A heterogeneous graph embedding model for circRNA-disease association prediction. IEEE J Biomed Health Inform. 2023;27:5177-5186.
21. Li J, Chen J, Wang Z, Lei X. HoRDA: Learning higher-order structure information to predict RNA-disease associations. Artif Intell Med. 2024;148:102775.
22. Zhong Y, Peng Y, Lin Y, Chen D, Zhang H, Zheng W, et al. MODILM: Towards better complex disease classification using a novel multi-omics data integration learning model. BMC Med Inform Decis Mak. 2023;23:82.
23. Przybyszewski J, Malawski M, Lichołai S. GraphTar: applying Word2Vec and graph neural networks to miRNA target prediction. BMC Bioinformatics. 2023;24:436.
24. Ji C, Wang Y, Ni J, Zheng C, Su Y. Predicting miRNA-disease associations based on heterogeneous graph attention networks. Front Genet. 2021;12:727744.
25. Cifuentes-Bernal AM, Pham VV, Li X, Liu L, Li J, Le TD. A pseudotemporal causality approach to identify miRNA–mRNA interactions during biological processes. Bioinformatics. 2021;37:807-814.
26. Jiao CN, Zhou F, Liu BM, Zheng CH, Liu JX, Gao YL. Multi-kernel graph attention deep autoencoder for miRNA-disease association prediction. IEEE J Biomed Health Inform. 2024;28:1110-1121.