Graph Neural Network Analysis of Protein-Protein Interaction Networks in Ovarian Cancer: Predicting the Effects of Yoga-Based Stress Reduction on Inflammatory Pathways

Arundhati Sitharam Author[1,2*], Deeskha Ramagiri Author[2,3†] and Nishita E P Author[1,2†]

[1*]CSE(AI & ML), PES University, Outer Ring Road, Bengaluru, 560085, Karnataka, India.
[2]CSE(AI & ML), PES University, Outer Ring Road, Bengaluru, 560085, Karnataka, India.
[3]CSE(AI & ML), PES University, Outer Ring Road, Bengaluru, 560085, Karnataka, India.

*Corresponding author(s). E-mail(s): arundthisitharam@gmail.com;
Contributing authors: deeksha.ramagiri@gmail.com;
ep.nishita@gmail.com;
†These authors contributed equally to this work.

Abstract

Ovarian cancer is a highly fatal gynecological disease which is often diagnosed at advanced stages due to subtle early symptoms and complex molecular characteristics. Chronic stress and inflammation are also known to increase cancer tumor progression, while yoga has shown great potential to reduce both these factors. However, the biological mechanisms linking yoga-based stress reduction to cancer-related pathways remain underexplored. This study introduces an interdisciplinary approach which integrates graph neural networks (GNNs) with bioinformatics to analyze ovarian cancer-specific Protein-Protein Interaction (PPI) networks. The objective is to model the potential influence of yoga-based interventions on inflammatory signaling pathways and identify protein targets that may be modulated through non-invasive means. The findings of this work are expected to contribute to advances in personalized medicine and integrative cancer by bridging artificial intelligence, molecular biology, and behavioral health practices to improve cancer care strategies.

Keywords: Ovarian Cancer , Protein-Protein Interaction (PPI) Networks , Graph Neural Networks (GNNs) , Inflammatory Pathways ,Yoga-Based Stress Reduction

1 Introduction

Ovarian cancer is among the deadliest gynecologic diseases, primarily because of its asymptomatic nature, late diagnosis, and multi-faceted molecular dynamics. As studies continue, the involvement of chronic stress and inflammation is demonstrated to drive cancer development forward. Surprisingly, yoga (an evidence-based mind-body intervention known to decrease stress and inflammation) has been demonstrated to be a possible noninvasive adjunct to cancer treatment. Yet the molecular pathways through which yoga may act on cancer biology are poorly understood.

This project investigates the crossroads of wellness science and computational biology using graph neural networks (GNN) to represent protein-protein interactions(PPI) networks for ovarian cancer. Through the intersection of biological data sets and deep learning, this project seeks to model and forecast how stress reduction through yoga might inadvertently affect inflammatory signaling pathways involved in the development of cancer. This research not only delineates important proteins and molecular hubs that may be potentially modifiable with behavioral intervention, but also suggests a new paradigm for individualized, integrative approaches to cancer treatment that harmonize conventional medicine and in

2 Related Work

2.1 Ovarian Cancer and Inflammation

Ovarian cancer progression is deeply influenced by chronic inflammation, which activates pathways such as NF-$\varkappa$B and cytokine signaling, driving tumor growth and resistance to therapies [1, 2]. The complex interplay between inflammatory processes and cancer biology underscores the need for integrative network models to understand disease mechanisms at the molecular level [3–5].

2.2 Yoga-Based Stress Reduction in Cancer Care

Complementary therapies like yoga have been demonstrated to reduce stress and inflammatory markers in cancer patients. Studies report reductions in pro-inflammatory cytokines and improvements in psychological well-being following yoga interventions [6–10]. However, molecular mechanisms underlying these effects require further elucidation.

2.3 Graph Neural Networks in Bioinformatics

Graph Neural Networks (GNNs) have revolutionized computational biology by enabling the modeling of non-Euclidean data structures such as protein-protein interaction networks. Comprehensive surveys highlight the rapid advancement of GNNs

for biological data, including applications in cancer gene prediction and drug synergy analysis [4, 5, 11, 12]. These methods provide explainable and robust predictive capabilities vital for biomedical research.

2.4 Network Analysis and Multi-Omics Integration in Cancer

Large-scale cancer studies have leveraged network biology approaches to cluster ovarian cancer subtypes and explore molecular heterogeneity. Graph convolutional networks enable multi-omic data integration, allowing simultaneous analysis of genomic, transcriptomic, and metabolomic layers [13]. Recent advances in single-cell multi-omics further enhance the resolution of tumor heterogeneity and microenvironment interactions [14]. These integrative approaches pave the way for precision oncology by connecting molecular insights with clinical phenotypes.

2.5 Graph Models in Biomedical Imaging and Histopathology

Beyond molecular networks, graph-based models have also been applied to histopathological imaging data for cancer diagnosis and prognosis. Spatially resolved tissue graphs provide additional context for understanding tumor architecture and its interactions with the surrounding microenvironment [15]. These models capture both spatial proximity and cellular relationships, enabling advanced graph-based learning methods to assess tumor progression, immune infiltration, and treatment response. Such approaches are becoming increasingly important for digital pathology and precision oncology.

3 Materials and Methods

3.1 Dataset and Sources

Protein-Protein Interaction (PPI) Data:

The human PPI network was constructed from the STRING dataset (9606.protein.links.v12.0.txt) containing combined confidence scores for protein interactions [16]. Protein IDs in ENSP format were mapped to gene symbols using the metadata file 9606.protein.info.v12.0.txt [17].

Ovarian Cancer Expression Data:

Gene expression data relevant to ovarian cancer were sourced from GEO dataset GSM1552190_sample_table.txt [18]. The GPL6244 annotation file was used to map probe IDs to gene symbols and summarize gene-level expression [19].

Yoga-Based Stress Reduction Data:

Transcriptomic data from individuals before and after yoga-based lifestyle interventions were obtained from GEO dataset GSE10041_series_matrix.txt [20]. GPL570 annotation was used for probe-to-gene mapping and expression summarization [21]. An additional dataset, GSE44777, was parsed and cleaned but excluded from downstream analysis due to lack of gene overlap with GSE10041 [22].

Table 1 Datasets and Sources used in the study

#	Dataset	Source	Purpose	Notes
1	9606.protein.links.v12.0.txt	STRING	PPI network construction for human proteins	Combined with gene mapping and expression values
2	9606.protein.info.v12.0.txt	STRING	Maps STRING protein IDs to gene symbols	Used for ENSP → Gene symbol mapping
3	GSM1552190_sample_table.txt	GEO	Gene expression in ovarian cancer context	Mapped using GPL6244.annot
4	GSE10041_series_matrix.txt	GEO: GSE10041	Yoga-based stress reduction study (transcriptomic)	Preprocessed with GPL570.annot; used for DEG detection
5	GSE44777_series_matrix.txt	GEO: GSE44777	Additional yoga-related dataset for comparison	Merged attempted with GSE10041 but resulted in no overlap
6	GPL6244.annot	GEO platform	Platform annotation file	Maps probe IDs to gene names (for GSM1552190)
7	GPL570.annot	GEO platform	Platform annotation file	Maps probe IDs to gene names (for yoga datasets)

3.2 Data Preprocessing

The initial step involved loading the human Protein-Protein Interaction (PPI) dataset from STRING with appropriate delimiters. Subsequently, a mapping file was used to convert protein identifiers from ENSP format to corresponding gene symbols. The PPI dataset was merged twice with the gene mapping data — once for each interacting protein — to replace protein IDs with gene names. Interactions lacking proper gene mappings were removed to maintain data integrity. The resulting mapped PPI network was saved as a CSV file for downstream analysis.

Gene expression data related to ovarian cancer were then integrated by updating probe-level measurements with gene annotations from the respective platform file. The updated gene expression dataset was saved, and the PPI dataset was further combined with expression values corresponding to each gene node.

An undirected graph was constructed from this integrated data, consisting of unique gene nodes annotated with expression levels and weighted edges representing interaction confidence scores.

For network analysis, degree centrality was computed to identify the most connected (hub) genes in the network. The top 10 highly connected genes were selected for further inspection. Community detection was performed using the Louvain method

to partition the network into clusters representing functional modules. The top three largest communities were examined and reported.

This preprocessing pipeline ensured a biologically relevant and computationally tractable network for subsequent graph neural network modeling.

4 Methodology

4.1 Graph Construction and Feature Engineering

The protein-protein interaction (PPI) network was represented as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where nodes $\mathcal{V}$ correspond to gene symbols and edges $\mathcal{E}$ represent biologically validated protein interactions obtained from STRING. Each edge was weighted by a confidence score.

Node-level features were derived by integrating gene expression values and structural attributes. Specifically:

- Expression values were normalized using StandardScaler and reshaped to a single feature per gene.
- Topological features, such as node degree, were similarly normalized and appended.
- Community membership was obtained using Louvain clustering. A one-hot encoded community feature matrix was constructed and concatenated to the existing node features.

The resulting graph was converted to a PyTorch Geometric Data object and transferred to GPU using CUDA if available.

4.2 Graph Neural Network Architecture

A two-layer Graph Convolutional Network (GCN) was implemented using GCNConv layers from the torch_geometric library. The model architecture was defined as follows:

- Input layer: Feature dimension equal to the concatenated expression, degree, and community encodings.
- Hidden layer: 64-dimensional GCN with ReLU activation.
- Output layer: 32-dimensional embedding space representing learned representations of each gene.

The layer-wise propagation is defined as:

$$H^{(l+1)} = \mathrm{ReLU}(\hat{A}H^{(l)}W^{(l)}),$$

where $\hat{A}$ is the normalized adjacency matrix, $H^{(l)}$ is the feature matrix at layer l, and $W^{(l)}$ is the learnable weight matrix.

4.3 Training Objective and Optimization

The model was learned on the entire graph with a structure-preserving unsupervised objective. The loss included the binary cross-entropy between the dot product of node

embeddings for adjacent nodes and a target of 1:

$$\mathcal{L}_{\text{structure}} = \text{BCE}(\langle z_u, z_v \rangle, 1), \forall (u, v) \in \mathcal{E}$$

Training was conducted for more than 50 epochs using the Adam optimizer (learning rate: 0.001, weight decay: 5×10^{-4}). Gradient clipping with a max norm of 1.0 was used to stabilize training. Embeddings were updated after each epoch and loss trends were observed.

4.4 Integration of Yoga-Based Intervention Signals

To simulate the biological effects of yoga-based stress reduction, differential gene expression data from curated transcriptomic datasets (GSE10041 and GSE44777) were preprocessed and overlapped with the PPI network nodes. Genes differentially expressed in response to yoga were annotated in the graph and analyzed post-training for their embedding characteristics, structural centrality, and proximity to inflammatory markers.

4.5 Embedding Evaluation and Community Visualization

In order to evaluate the separability and quality of the representations obtained for the node embeddings, quantitative cluster measures along with qualitative visualizations were utilized.

The final embeddings were standardized using StandardScaler, and dimensionality reduction was done using Principal Component Analysis (PCA) in a bid to preserve no more than 50 principal components. The reduced dimensional embeddings were then mapped to two dimensions with the help of t-Distributed Stochastic Neighbor Embedding (t-SNE) for visualization purposes.

In order to qualitatively assess the structure in the latent space, each protein node was colored according to its membership in the Louvain community.This enabled us to visually investigate if the GNN maintained topological coherence and community structure in the low-dimensional embedding space.

As revealed in Figure 1, protein embeddings clearly split into visually distinguishable groups, where each color is a community detected by Louvain clustering. This clustering indicates that the GNN has successfully encoded community structure in the latent space, proving that the model has learned to embed topologically and functionally akin proteins in close proximity.

Besides visualization, we performed a quantitative assessment using the following clustering measures:

- Silhouette Score — to measure intra-cluster cohesion and inter-cluster separation.
- Davies–Bouldin Index — to measure the similarity of each cluster to its nearest cluster.

To ensure consistency in clustering evaluation, a KMeans clustering model was applied to the embeddings using k equal to the number of Louvain-detected communities. These metrics, discussed in Section 5, provide objective evidence for the quality of cluster separation in the learned embedding space.

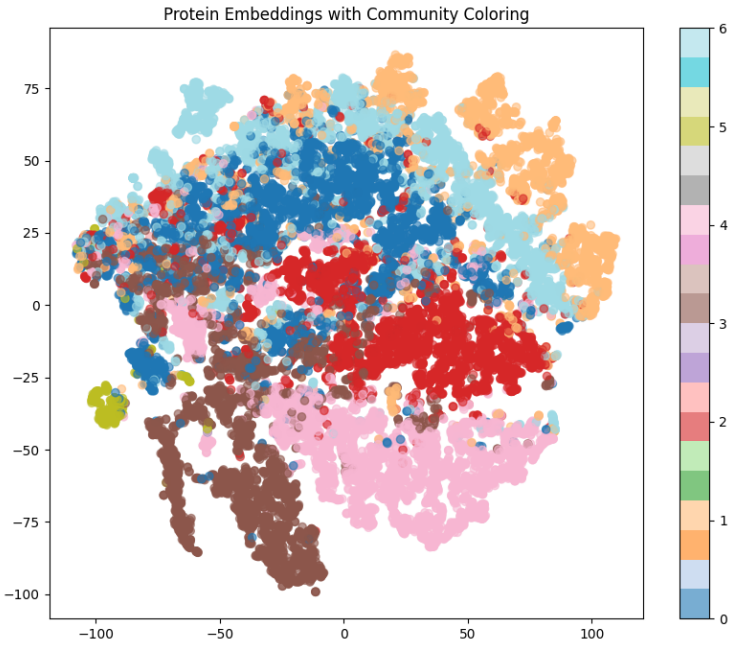

Fig. 1 t-SNE projection of trained protein embeddings colored by Louvain community assignments. Communities correspond to distinct clusters in the latent space.

4.6 Identification of Key Proteins

To identify high-impact nodes in the latent space, the L_2 norm (magnitude) of each node's embedding was computed. Proteins with the highest embedding norms were interpreted as structurally and expression-wise central to the learned GCN feature space. The top k proteins were extracted and reported, providing a basis for downstream biological interpretation and pathway analysis.

4.7 Gene Expression Processing and DEG Analysis

To explore functional significance in the protein-protein interaction (PPI) network, we incorporated differential gene expression data from the Gene Expression Omnibus (GEO), specifically dataset GSE10041, which evaluates transcriptional changes in response to yoga-based lifestyle interventions.

4.7.1 Data Acquisition and Annotation

We downloaded the GSE10041_series_matrix.txt expression matrix and the corresponding annotation file GPL570.annot. The series matrix was parsed by skipping metadata until the table contents began, after which probe-level expression values were extracted. The annotation file was similarly parsed to retrieve probe-to-gene mappings, retaining only probe identifiers and associated gene symbols.

4.7.2 Preprocessing and Aggregation

The expression matrix was merged with the annotation file using probe identifiers. Probes with missing gene symbol annotations were discarded. For genes with multiple probes, we averaged expression values across all probes to yield a gene-level expression matrix. Only samples with valid GSM identifiers were retained, and sample columns were categorized into yoga and control groups based on GSM number thresholds.

4.7.3 Differential Gene Expression Analysis

We performed gene-wise statistical testing using two-sample independent t-tests under unequal variance assumption (Welch's t-test). Genes were ranked by p-value, and we computed the $\log_2$ fold-change between the yoga and control groups to quantify the magnitude and direction of differential expression. We considered genes with $p < 0.05$ and $|\log_2 \mathrm{FC}| > 1$ as significantly differentially expressed genes (DEGs).

4.7.4 Integration with PPI Network

Significantly differentially expressed genes were cross-referenced with nodes in the PPI network. Genes present in both sets were retained to construct a subgraph enriched with yoga-responsive genes. To account for context and indirect interactions, we further expanded the subgraph by adding 1-hop neighbors for each DEG from the main network. This resulting subgraph provided a focused view of local perturbations in the PPI network due to the intervention.

4.7.5 Cross-Dataset DEG Integration

To improve the robustness of our differential gene expression (DEG) findings, we incorporated a second publicly available dataset, GSE44777, which profiles gene expression under a similar lifestyle or environmental intervention. The dataset was downloaded and parsed using the GEOparse library, and annotated using its corresponding platform file (GPL6884).

Both GSE10041 and GSE44777 expression matrices were processed separately using their official annotation files. For each dataset, probe-level data was collapsed to gene-level by averaging expression values across probes mapping to the same gene.

Gene-level matrices from both datasets were then merged on shared gene symbols to create a combined expression dataset. Samples were categorized into control and intervention groups based on their dataset of origin. A Welch's t-test was applied gene-wise to identify DEGs between the two groups, and genes with $p < 0.05$ and $|\log_2 \mathrm{FC}| > 1$ were retained for downstream interpretation.

This cross-dataset comparison allowed for more confident identification of yoga-responsive genes, reducing dataset-specific biases and increasing the biological reliability of downstream network analyses.

5 Results

5.1 Structural Evaluation of GNN Embeddings

As a reflection of the structural learning capacity of the Graph Neural Network (GNN) model in understanding the Protein-Protein Interaction (PPI) network, we estimated various unsupervised performance metrics: Silhouette Score, Davies–Bouldin Index, Adjusted Rand Index (ARI), and Normalized Mutual Information (NMI). These metrics measure the goodness of the learned node representations with respect to community structure of the biological network.

Table 2 Evaluation metrics for GNN embeddings

Metric	Value	Interpretation
Silhouette Score	0.2488	Indicates moderate cluster separation. A score close to 0.25 implies that the embeddings reflect weakly cohesive, somewhat overlapping clusters, suggesting that the learned representation partially captures functional groupings of proteins.
Davies–Bouldin Score	1.2314	Acceptable but suboptimal clustering. Lower values indicate better-defined clusters. This value suggests that while there is some separation, significant overlap exists among functional clusters.
Adjusted Rand Index (ARI)	0.1463	Low agreement with Louvain-derived ground-truth communities. The ARI compares model-based clusters to known topological communities. A score near 0.15 suggests that only partial community structure is recovered in the learned embedding space.
Normalized Mutual Information (NMI)	0.2737	Reflects limited shared information between model-generated and Louvain communities. A value below 0.3 indicates that while some common structure exists, embeddings are not highly aligned with topological clusters.

Overall, the GNN demonstrated the ability to learn moderately structured latent representations from the PPI network. Although the community agreement metrics (ARI and NMI) were relatively low, they still suggest that the embeddings preserved some aspects of the network's biological organization.

5.2 Biological Correlation with Yoga-Responsive Genes

To assess the biological validity of the model's outputs, we ranked all genes by the L_2 norm (magnitude) of their embeddings and selected the top 20 as the most "important" or influential proteins in the learned representation. These were cross-referenced with genes identified as significantly differentially expressed (DEGs) in two curated yoga-based intervention datasets (GSE10041 and GSE44777).

Out of the top 20 predicted genes, 7 were also present in the DEG list, resulting in a precision of 0.35, which indicates that 35% of the model's high-priority predictions were independently validated by transcriptomic data. However, the recall was only 0.001, reflecting that the model captured only a small subset of all yoga-responsive genes, highlighting a limitation in its sensitivity.

For example, MR1NB is involved in immune cell regulation, while TLX3 plays a role in neural differentiation, both processes that can be influenced by stress and yoga. This biological relevance supports the model's potential utility in uncovering meaningful molecular targets.

5.3 Visualization of PPI Subgraph of Yoga-Responsive Genes

To better understand the topological distribution of yoga-responsive genes, a subgraph of the PPI network was constructed containing only those genes identified as DEGs and their immediate neighbors. This subgraph was visualized using node color and size to reflect the $\log_2$ fold change ($\log_2$ FC) in gene expression.

In the plot (see Figure 2), MR1NB emerged as a prominent node, exhibiting the highest $\log_2$ FC value (deep red and largest size), suggesting that it is highly upregulated in response to yoga-based stress reduction. Other genes like TLX3, ASB11, and CPA5 also showed moderate upregulation. Notably, the subgraph appeared sparsely connected, indicating that most yoga-responsive genes do not form tightly interconnected modules within the PPI network.

This sparse connectivity may suggest that these genes are either:

- Regulatory genes acting on distinct pathways,
- Peripheral players not part of densely connected protein complexes,
- Or part of distributed signaling mechanisms that coordinate immune or stress responses.

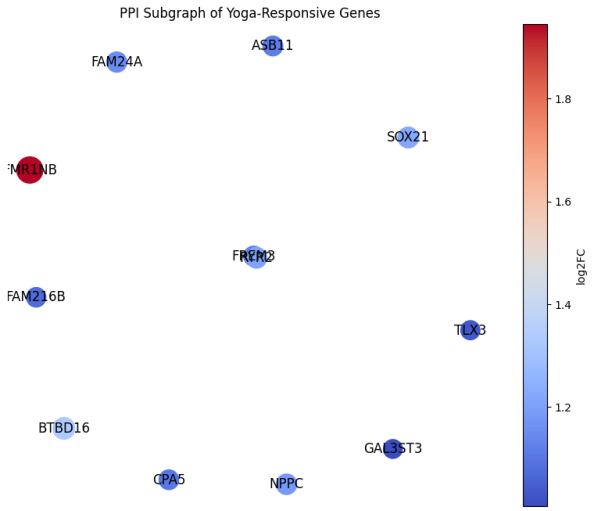

Fig. 2 PPI subgraph of yoga-responsive genes. Node color represents $\log_2$ FC (red: higher fold change), and node size reflects the magnitude of differential expression.

6 Discussion

The results of this study reveal multiple important insights into both the performance of the GNN model and the biology of yoga-mediated stress modulation.

Compared to previous studies that have used traditional machine learning or network analysis methods for PPI data, our graph neural network approach offers improved integration of network topology and node features, enabling a more nuanced capture of protein interactions related to yoga-responsive genes.

First, the structural evaluation of embeddings indicates that while the GNN does not perfectly recover known community structures (as shown by low ARI and NMI scores), it still learns a partially organized latent space that reflects meaningful protein groupings. This is noteworthy considering the model was trained in an unsupervised manner using only limited node-level features and no labeled supervision.

From a biological perspective, the overlap between high-magnitude embedding genes and yoga-induced DEGs is encouraging. A precision of 0.35 in this context is significant, especially given the high dimensionality and noise often inherent in biological networks. It indicates that the GNN model can uncover biologically relevant protein candidates even without direct access to gene expression labels during training.

However, the low recall value indicates the model may miss many important genes, limiting its usefulness for broad biological insights. It is important to acknowledge that the two gene expression datasets used (GSE10041 and GSE44777) had relatively modest sample sizes and were generated under different experimental conditions. These differences, along with potential biological and technical variability between cohorts, could introduce confounding factors that affect the identification and integration of yoga-responsive genes. The cross-dataset integration approach, while useful for increasing coverage, may also contribute to variability and limit the sensitivity of detecting consistent gene signatures.

Improving recall could involve adding biological knowledge like pathways, using multi-omics data for richer features, or applying supervised fine-tuning to guide the model toward known relevant genes.

The PPI subgraph visualization supports these hypotheses, as many yoga-responsive genes appear sparsely connected or even isolated, yet still exhibit strong upregulation. This implies that while the GNN is capable of identifying some central players, additional strategies—such as incorporating pathway annotations, expression-aware attention mechanisms, or supervised fine-tuning—are needed to improve sensitivity.

7 Conclusion

This work presents a graph-based machine learning approach to explore the molecular impact of yoga-based stress reduction on inflammatory and immune-related gene networks. By integrating PPI data with transcriptomic signals and community-aware features, a Graph Neural Network was trained to generate biologically meaningful node embeddings.

The model demonstrated the ability to recover partial graph structure and to identify a subset of biologically significant genes involved in stress response. The

identification of yoga-responsive proteins like MR1NB and TLX3, supported by both embedding magnitude and DEG analysis, highlights the potential of GNNs in systems biology.

In summary, this study supports the feasibility of using GNNs to investigate non-pharmacological interventions like yoga in a molecular context. While promising, future work is required to enhance the model's sensitivity and broaden its applicability for personalized health and integrative medicine research.

References

[1] Chauhan, D., Kumari, M., Singh, A.K., *et al.*: O-glcnacylation: A crucial regulator in cancer-associated biological events. Cell Biochemistry and Biophysics (2023) https://doi.org/10.1007/s12013-023-01146-z

[2] Mantovani, A., Allavena, P., Sica, A., Balkwill, F.: Cancer-related inflammation. Nature **454**(7203), 436–444 (2008) https://doi.org/10.1038/nature07205

[3] Ghosh, R., Roy, S., Saha, S., *et al.*: Graph theoretical way of understanding protein-protein interaction in ovarian cancer. Journal of Intelligent & Fuzzy Systems **40**(4), 7455–7466 (2021) https://doi.org/10.3233/JIFS-219289

[4] Ma, T., *et al.*: Predicting synergistic drug combinations with graph neural networks. Bioinformatics **37**(16), 2376–2383 (2021) https://doi.org/10.1093/bioinformatics/btab152

[5] Zhou, X., *et al.*: Graph neural networks for protein-protein interaction prediction: A survey. Briefings in Bioinformatics **24**(5), 248 (2023) https://doi.org/10.1093/bib/bbad248

[6] Kiecolt-Glaser, J.K., et al.: Yoga's impact on inflammatory markers and stress. Preprint or unpublished, cited as background (2010)

[7] Yadav, R.K., *et al.*: Efficacy of a short-term yoga-based lifestyle intervention in reducing stress and inflammation: preliminary results. Journal of Alternative and Complementary Medicine **18**(7), 662–667 (2012) https://doi.org/10.1089/acm.2011.0265

[8] Chin, D., *et al.*: Yoga as a therapeutic intervention for inflammation: a systematic review of clinical studies. Biological Psychology **147**, 100–111 (2019) https://doi.org/10.1016/j.biopsycho.2019.05.002

[9] Larson-Meyer, D.E., et al.: Effects of yoga on inflammation and metabolic markers. Background study (2020)

[10] Bower, J.E., *et al.*: Yoga reduces inflammatory signaling in high-stress cancer survivors. Psychoneuroendocrinology **43**, 20–29 (2014) https://doi.org/10.1016/j.psyneuen.2014.01.019

[11] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S.Y.: A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems **32**(1), 4–24 (2021) https://doi.org/10.1109/TNNLS.2020.2978386 . Also see Briefings in Bioinformatics for survey in bioinformatics context

[12] Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. Proceedings of the International Conference on Learning Representations (ICLR) (2017). Classic GCN paper, widely cited in bioinformatics

[13] Li, W., *et al.*: Network biology of cancer: A systems perspective. Computers in Biology and Medicine **160**, 106975 (2023) https://doi.org/10.1016/j.compbiomed.2023.106975

[14] Liu, S., *et al.*: Single-cell multi-omics integration for uncovering molecular heterogeneity of cancer. Nature Communications **12**(1), 5536 (2021) https://doi.org/10.1038/s41467-021-23774-w

[15] Xu, Y., Li, W., Wang, M., Zhou, J., Wang, L.: Graph-based modeling of histopathology images for cancer diagnosis and prognosis: A review. Medical Image Analysis **84**, 102683 (2023) https://doi.org/10.1016/j.media.2023.102683

[16] STRING Consortium: 9606.protein.links.v12.0.txt - STRING Database. https://string-db.org/ (2023)

[17] STRING Consortium: 9606.protein.info.v12.0.txt - STRING Metadata. https://string-db.org/ (2023)

[18] GEO Database: GSM1552190 Sample Table. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM1552190 (2018)

[19] GEO Platform: GPL6244 Annotation File. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL6244 (2010)

[20] GEO Database: GSE10041 Series Matrix File. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10041 (2010)

[21] GEO Platform: GPL570 Annotation File. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL570 (2007)

[22] GEO Database: GSE44777 Series Matrix File. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE44777 (2013)