Abstract #1922

Application of an interpretable graph neural network to predict gene expression signatures associated with tertiary lymphoid structures in histopathological images

STUDY BACKGROUND

Tertiary lymphoid structures (TLS) are vascularized lymphocyte aggregates in the tumor microenvironment (TME) that correlate with better patient outcomes. Previous studies identified a 12-chemokine gene expression signature associated with disease progression and the type and degree of TLS[1]. These signatures could provide insight important for clinical decision making during pathologic evaluation.

Recently, deep learning (DL) approaches in digital pathology have been successfully piloted for cancer diagnosis; however, DL-based imputation of molecular phenotypes from pathology images has had mixed success. Predicting gene expression from whole slide images (WSI) may be impeded by low prediction accuracy and lack of interpretability[2].

To address these limitations, we developed an artificial intelligence (AI)-based, state-of-the-art workflow to predict the 12-chemokine TLS gene signature as well as the patient outcomes from lung and breast cancer WSIs, and to identify histological features relevant to model predictions. We also show that stratification based on GNN-inferred TLS gene expression has significant prognostic value for survival outcome.

GNNEXPLAINER MODEL INTERPRETATION

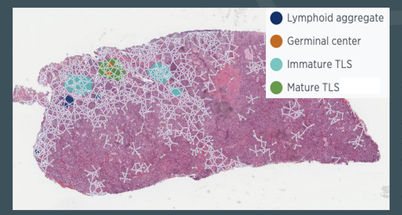

Figure 6. GNNExplainer interpretation of TLS gene model predictions. Example explained subgraph and predicted TLS regions from our previous segmentation models overlaid onto LUAD WSI. The selected subgraph (light grey lines) overlaps with TLS regions, but also extends beyond these regions, suggesting that TLS may affect gene expression and TME in regions beyond their immediate vicinity. Compared to the subgraphs produced for the CD8 model, the subgraphs for TLS models cover a larger area of the slides, indicating that cell types other than lymphocytes might be associated with the TLS gene expression (for example, plasma cells).

METHODS

We used a graph neural network (GNN) model, which takes into account the spatial arrangement of cells and tissue, to predict gene expression. We trained convolutional neural networks (CNNs) to classify cell and tissue features on entire WSIs[3], and used these outputs as nodes and features in the GNN model (**Fig. 1a**). To interpret the GNN models, we used the GNNExplainer[4], which simultaneously identifies a subgraph and a subset of node features important for predictions (**Fig. 1b**).

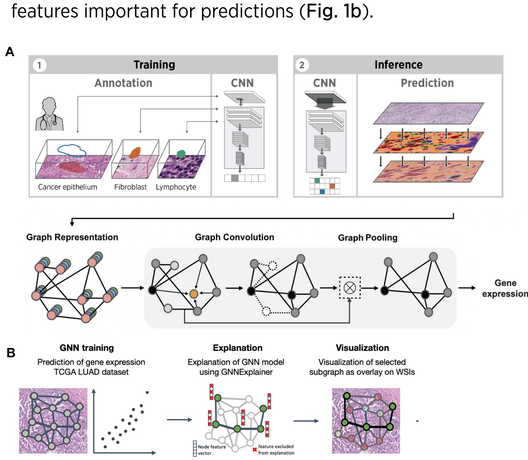

Figure 1. Workflow of the interpretable GNN model to predict gene expression. A) WSIs with pathologist annotations are used to train CNN based cell- and tissue-classifiers. Cell and tissue predictions are represented as graphs and used to train a GNN for prediction of gene expression. B) The GNNExplainer is used to interpret GNN model predictions by identifying the subgraph and node features most relevant for model predictions.

We validated the performance of the GNNExplainer by applying it on a GNN model trained to predict CD8A gene expression on TCGA (The Cancer Genome Atlas) LUAD (lung adenocarcinoma) dataset (**Fig. 2a**). As CD8 is a surface marker for cytotoxic T cells, the top selected node features were, as expected, lymphocyte related, including mean and standard deviation of lymphocyte cell model predictions, and the fraction of lymphocytes in cancer stroma (**Fig. 2b**). The selected subgraphs cover regions that are lymphocyte-dense (**Fig. 2c**).

RESULTS

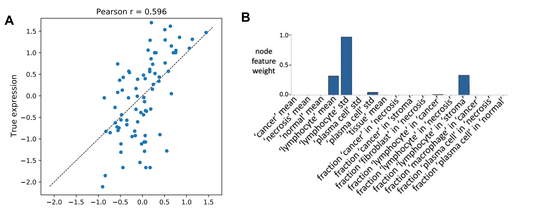

Figure 2. Validation of the GNNExplainer. A) GNN model can predict CD8A expression (normalized log RNAseq count) with significant Pearson correlation on a held-out test set. B) The selected node features by the GNNExplainer (high weight) are lymphocyte related. C) The selected subgraphs (light grey lines) selectively cover lymphocyte-dense tumor regions.

We applied this workflow to predict the 12-chemokine TLS gene signature. GNN models were trained using both the TCGA LUAD dataset (538 cases with paired lung cancer WSI and mRNA-seq expression data) and the TCGA breast invasive carcinoma (BRCA) dataset (1125 cases with paired breast cancer WSI and mRNA-seq expression data). GNN predictions of TLS signature genes were compared with the predictions of models trained using hand-crafted, task-specific features (HIFs LASSO models) describing the number, size, and cellular composition of identified TLS.

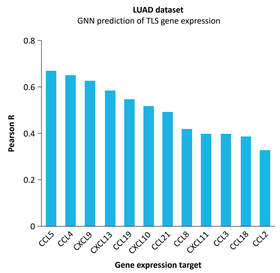

Figure 3. GNN model performance on prediction of TLS signature genes in the LUAD dataset. GNN models significantly predicted mRNA expression of all 12 genes ($p<0.05$).

The Pearson correlation coefficient was used to assess the accuracy of GNN and TLS feature model predictions. GNN models predict 6/12 genes with $r>0.5$ in the LUAD dataset, and 5/12 genes with $r>0.5$ in the BRCA dataset, and the GNN approach outperforms the HIFs LASSO model (**Figs. 3,4**).

RESULTS

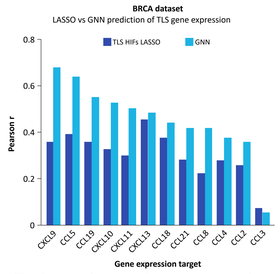

Figure 4. GNN model performance on prediction of TLS signature genes in the BRCA dataset and comparison with HIFs LASSO approach. GNN models significantly predicted mRNA expression of 11/12 genes ($p<0.05$) in the BRCA dataset. Individual GNN models outperformed models trained using hand-crafted TLS feature models in all significantly predicted 12-chemokine TLS signature genes.

To investigate the prognostic value of GNN-inferred TLS gene expression, we first grouped all cases in the held-out test set into two groups using hierarchical clustering (**Fig 5a**). This stratification based on inferred gene expression has significant prognostic value for overall survival ($p=0.03$, log-rank test; **Fig. 5b**).

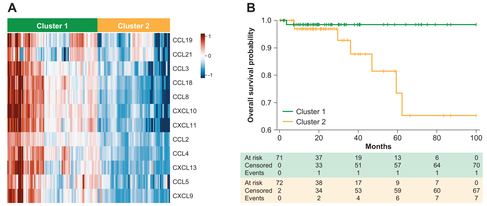

Figure 5. Prognostic value of GNN-inferred TLS gene expression. A) Hierarchical clustering of all cases in the held-out test set. B) KM plots for test cases separated by cluster assignment.

Application of the GNNExplainer on TLS-predicting GNN models identified some node features common among all models (for example, mean and standard deviation (std) of lymphocyte counts), while other selected node features were found to be unique to particular models (for example, plasma cell std, tissue std, macrophage std). Qualitative evaluation of subgraphs identified by GNNExplainer reveals that human-annotated TLS regions as well as small substructures in other regions (**Fig. 6**) contribute to model predictions, suggesting that information on TLS-induced genes might come from regions beyond TLSs.

CONCLUSIONS

Here we report an interpretable GNN-based approach to predicting TLS gene expression from lung and breast cancer WSIs. Stratification based on GNN-inferred TLS gene expression was further found to have prognostic value for survival outcome in breast cancer. The GNN-based model outperforms the HIF LASSO model, and predicts histopathology features relevant to TLS that may be used to inform patient prognosis and treatment. These methods could be applied to predict additional clinically relevant transcriptomic signatures.

AUTHORS

Ciyue Shen[1,2], Collin Schlager[1,3], Deepta Rajan[1], Victoria Mountain[1], Mary Lin[1], Maryam Pouryahya[1], Ilan Wapinski[1], Amaro Taylor-Weiner[1], Benjamin Glass[1], Andrew Beck[1], & Robert Egger[1]

[1]PathAI, Boston, MA; [2]Harvard University, Cambridge, MA; [3]Stanford University, Palo Alto, CA

contact: c_shen@g.harvard.edu; robert.egger@pathai.com

REFERENCES

1. Zhu, G, et. al. Front Immunol. 2017; 8:767.
2. Schmauch, B, et al. Nat. Commun. 2020; 11: 3877.
3. Diao, JA, et al. Nat. Commun. 2021; 12: 1613.
4. Ying, R, et al. arXiv: 1903.03894.

ACKNOWLEDGMENTS

The results shown here are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

We thank Bioscience Communications for expert assistance in data visualization. https://bioscicom.net

PathAI

AACR Annual Meeting 2022, April 8-13, 2022, New Orleans, LA, USA