Supporting Information for

Knowledge Graph Modulated Deep Learning for Limited-Sample Clinical Data Analysis

Tracy Xue, Sakib Mostafa**, James Zou, Joseph Liao, Maximilian Diehn, Ash A. Alizadeh, Lei Xing, and Md. Tauhidul Islam**[1]

** Co-first author

Corresponding Author: Md. Tauhidul Islam and Lei Xing
E-mail: tauhid@stanford.edu and lei@stanford.edu

This PDF file includes:

Figs. S1 to S22

Supplementary Materials

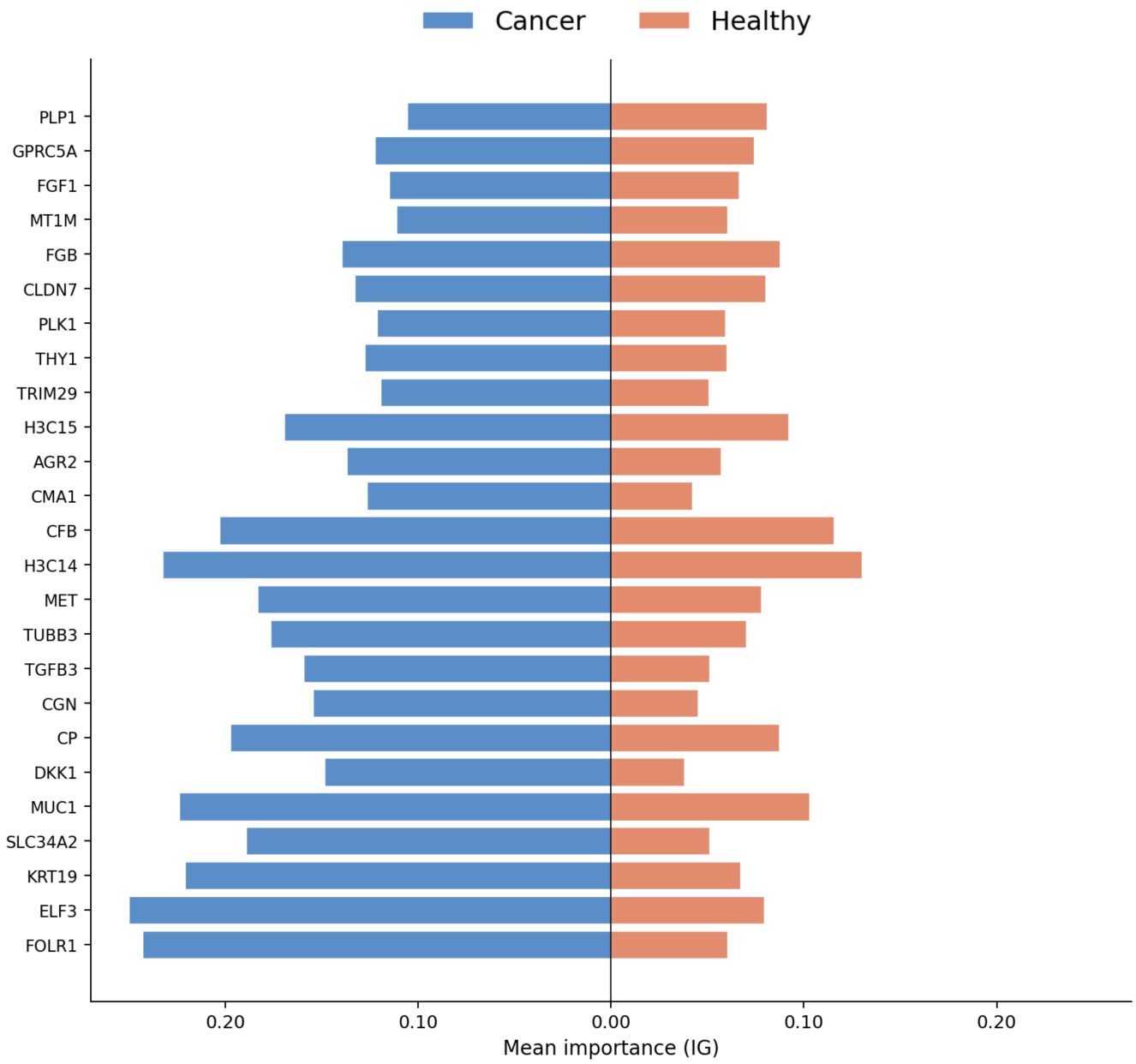

Fig. S1. Integrated Gradients attribution map for RARE-Seq binary classification. Class-specific Integrated Gradients (IG) attribution scores for the RARE-Seq cancer-versus-healthy classification task. Each bar represents the mean IG importance of a gene across samples, with cancer-associated attributions shown in blue and healthy-associated attributions shown in salmon. Bars are mirrored around zero to facilitate visual comparison between classes. The classifier assigns high cancer-associated attribution to genes including FOLR1, ELF3, KRT19, and MUC1, suggesting that GiG identifies discriminative molecular signals consistent with epithelial and tumor-associated transcriptional programs in circulating cfRNA.

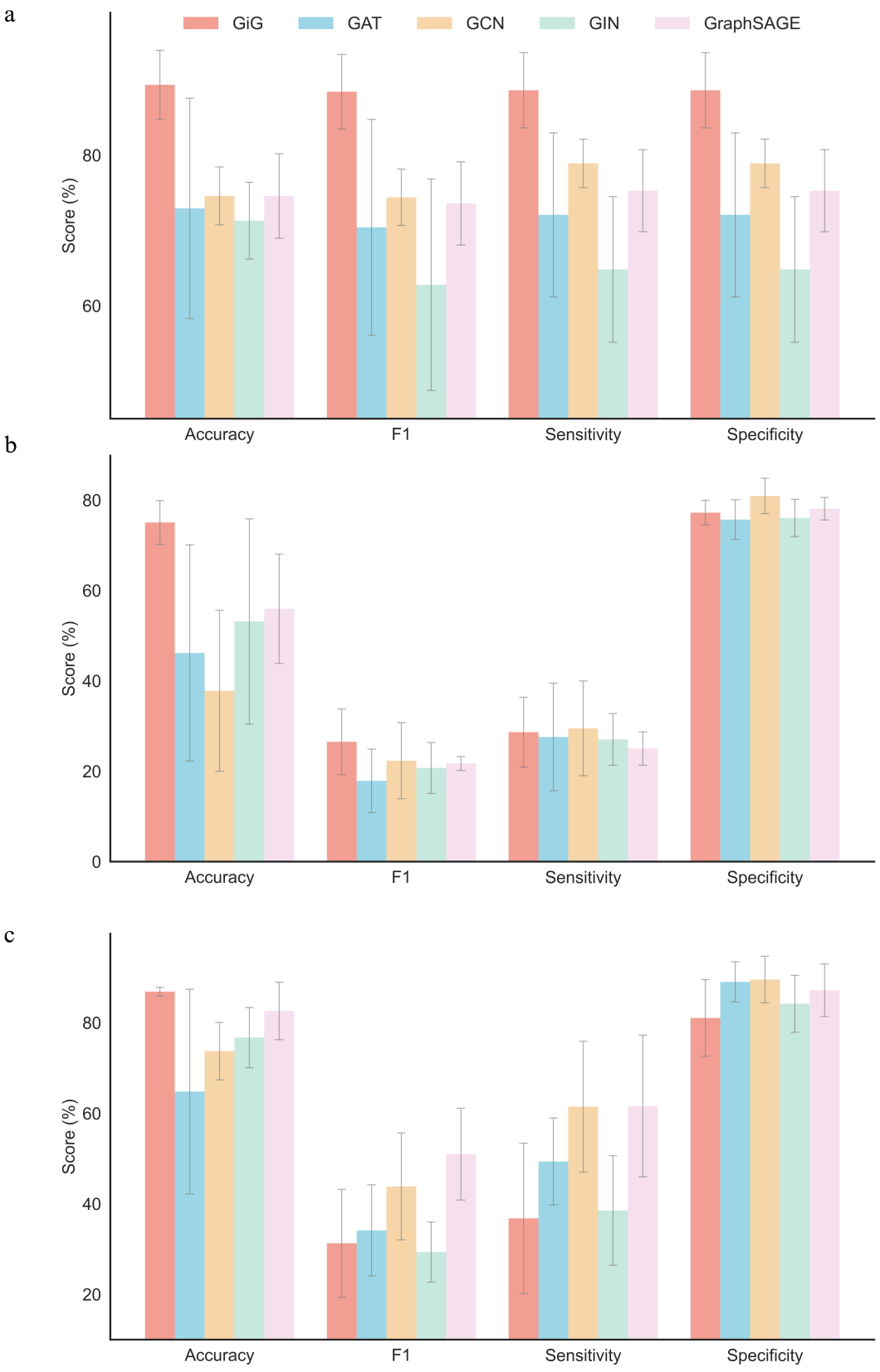

Fig. S2. RARE-Seq binary, stage, and subtype baseline performance. Cross-validation performance of GiG compared with node-level GNN baselines across three RARE-Seq classification tasks. a, Binary cancer-versus-healthy classification. b, Cancer stage classification. c, Cancer subtype classification. Bars indicate mean accuracy, macro-F1, sensitivity, and specificity across folds, and error bars indicate variability across folds. GiG shows the strongest overall performance in the binary detection task and maintains competitive performance across stage and subtype prediction, highlighting the benefit of incorporating patient-specific pathway graph structure into cfRNA-based classification.

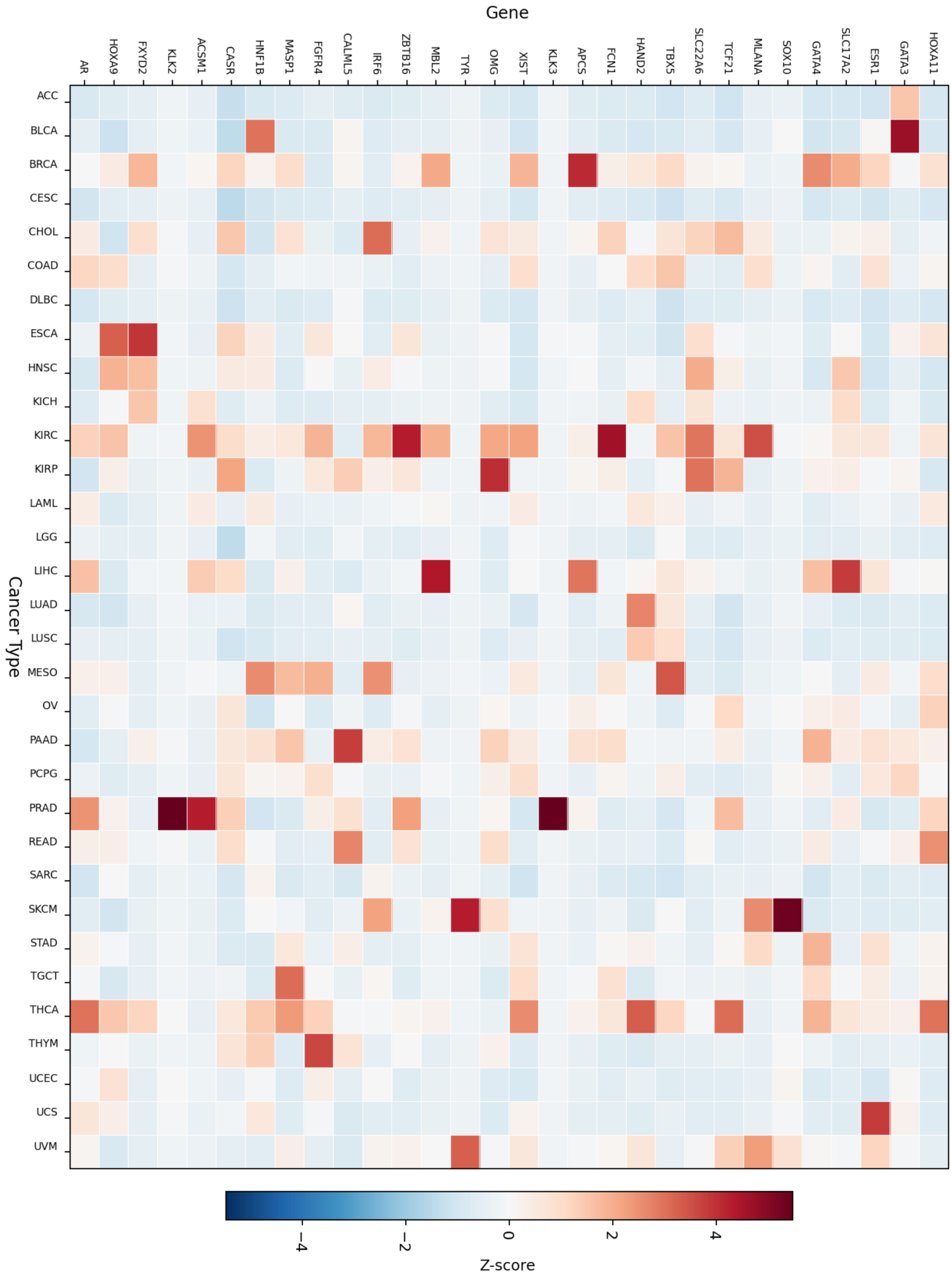

Fig. S3. Pan-cancer Integrated Gradients attribution heatmap. Z-scored class-wise Integrated Gradients (IG) attribution scores for the top-ranked genes in the pan-cancer classification task. Rows represent genes and columns represent cancer types. Red indicates higher relative attribution for a given cancer class, while blue indicates lower relative attribution. The heatmap shows that GiG assigns class-dependent importance to distinct gene subsets rather than relying on a single shared transcriptional signature across all tumor types.

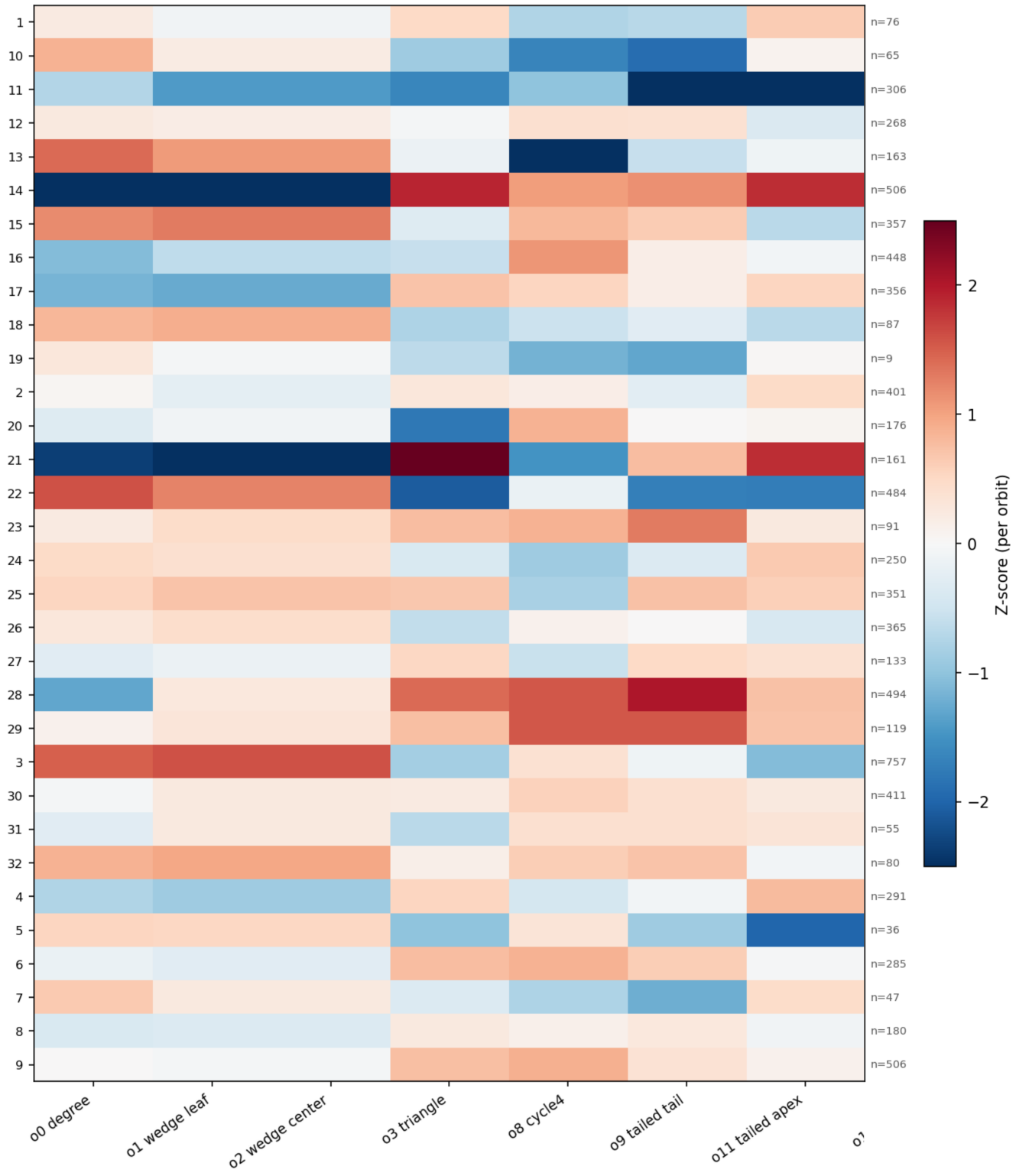

Fig. S4. Cancer type–specific graphlet orbit signatures in pan-cancer pathway graphs. Mean ORCA-derived graphlet orbit signatures for real patient-specific pan-cancer pathway graphs, aggregated by cancer type and z-scored within each orbit. Columns correspond to selected 4-node graphlet orbits, including degree, wedge, triangle, cycle, tailed-triangle, and clique-related positions. Each row represents a cancer type, with sample size shown on the right. The heatmap demonstrates that patient-specific pathway graphs differ not only in gene composition but also in local network topology across cancer classes.

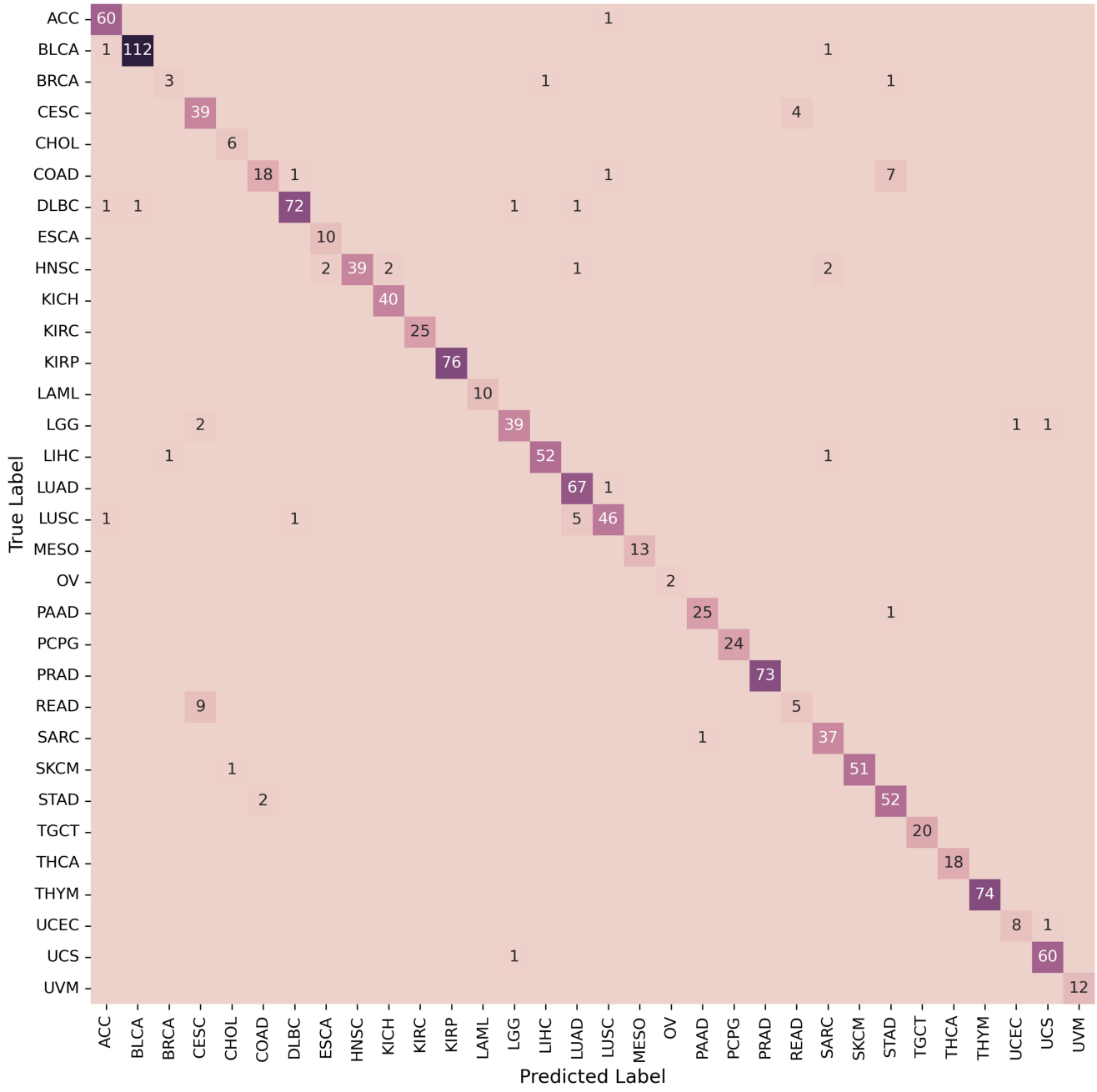

Fig. S5. Pan-cancer full-gene confusion matrix. Confusion matrix for the pan-cancer classification task using the full-gene input representation. Rows indicate true cancer labels and columns indicate predicted labels, with values corresponding to sample counts. Darker diagonal entries represent correctly classified samples, while off-diagonal entries indicate misclassifications between cancer types. The strong diagonal structure shows that the model accurately distinguishes most cancer classes, with limited confusion concentrated among a small subset of tumor types.

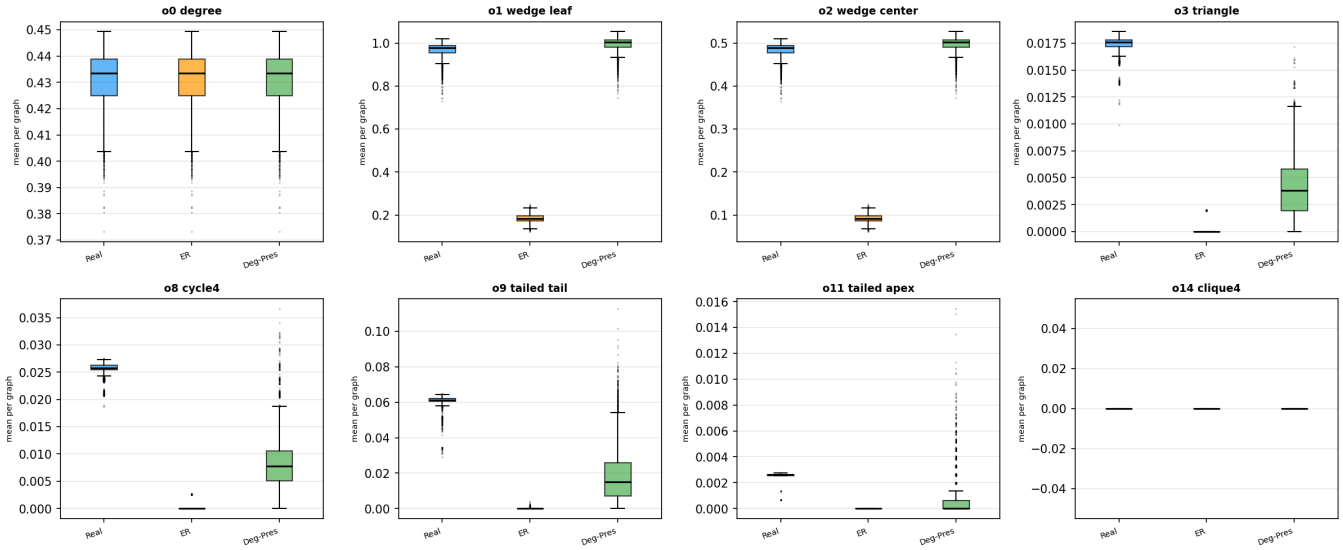

Fig. S6. Pan-cancer pathway graph topology compared with random graph controls: Distributions of selected graphlet orbit features in real pan-cancer pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot shows the mean orbit count per graph for a specific local topology. Real pathway graphs show distinct enrichment patterns across wedge, triangle, cycle, and tailed-triangle orbits, indicating that the observed patient-specific graphs retain non-random pathway structure.

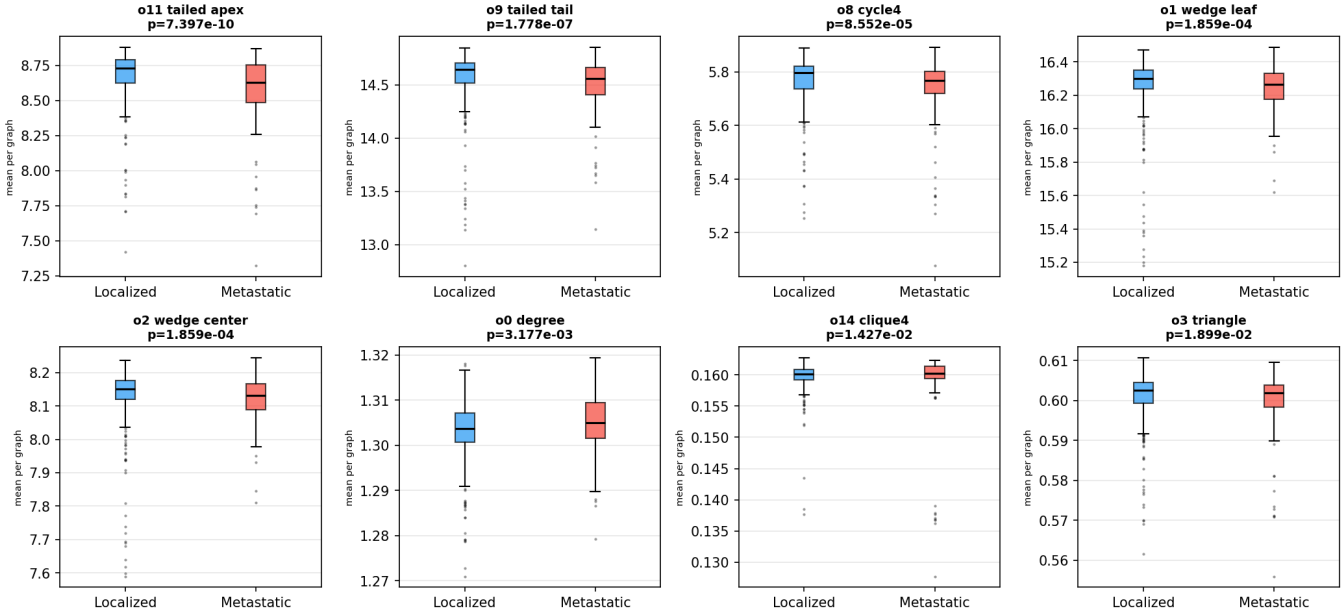

Fig. S7. Graphlet orbit differences between localized and metastatic prostate cancer. Top graphlet orbit features distinguishing localized and metastatic prostate cancer pathway graphs, ranked by statistical significance. Boxplots compare mean orbit counts per graph between clinical groups for selected ORCA-derived orbit positions. Differences across tailed-triangle, wedge, cycle, degree, clique, and triangle-related orbits suggest that disease progression is associated with measurable changes in local pathway graph organization.

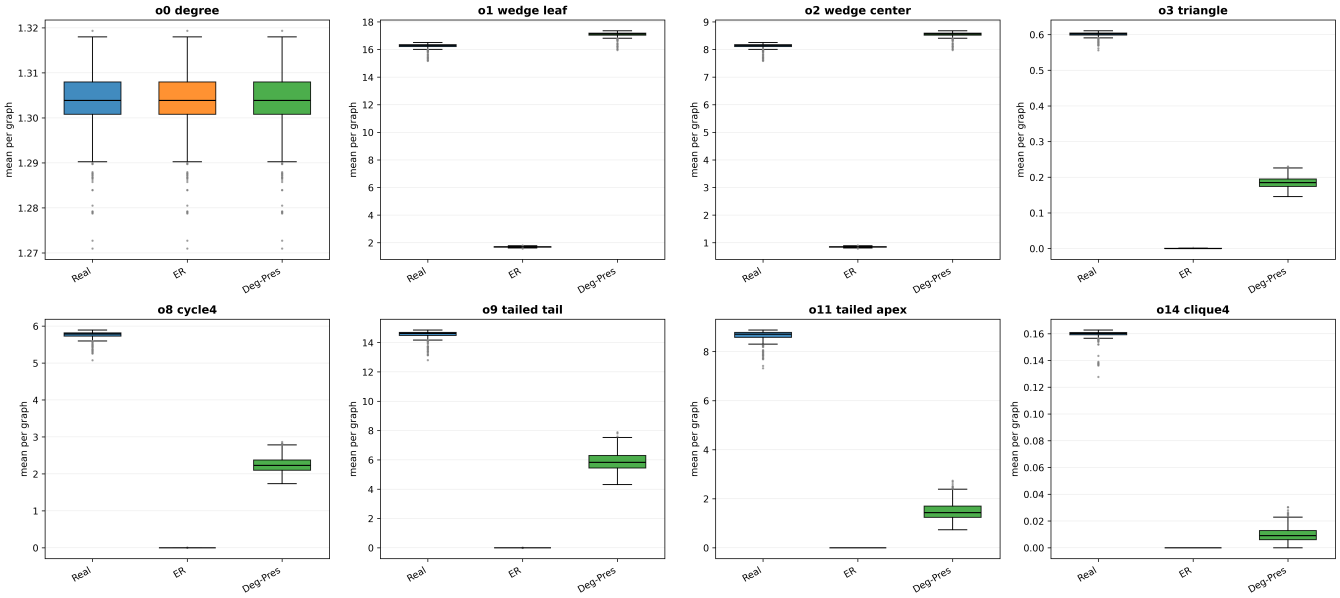

Fig. S8. Prostate cancer pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real prostate cancer patient-specific pathway graphs compared with Erdős–Rényi, degree-preserving, and fully connected graph controls. Real pathway graphs display orbit distributions that differ strongly from random and fully connected controls, supporting that curated pathway topology contributes structured, biologically constrained graph organization rather than arbitrary connectivity.

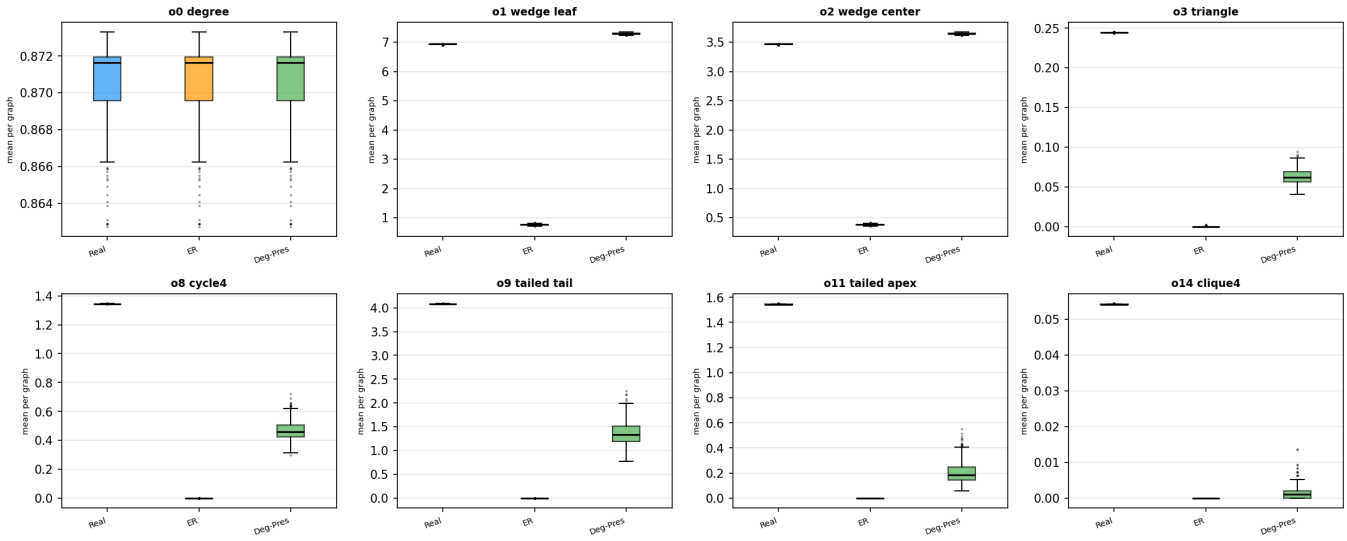

Fig. S9. RareSeq pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real RareSeq patient-specific pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot shows the mean orbit count per graph for a specific local topology. Real pathway graphs retain elevated counts of closed-motif orbits including triangles (o3), 4-cycles (o8), tailed triangles (o9, o11), and 4-cliques (o14). Erdős–Rényi rewiring collapses these orbits to near zero. Degree-preserving rewiring recovers part of the closed-motif structure but remains well below real-graph values. The degree sequence alone does not reproduce the local clustering present in curated pathway graphs.

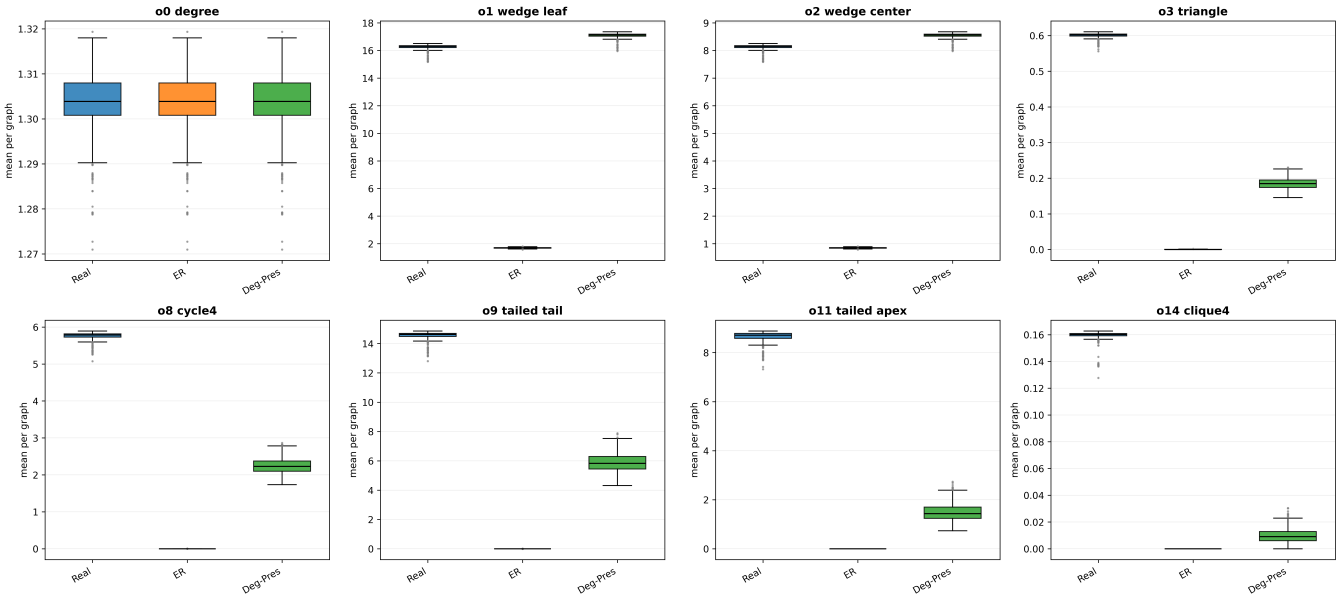

Fig. S10. Prostate cancer pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real prostate cancer patient-specific pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot shows the mean orbit count per graph for a specific local topology. Real pathway graphs show large enrichments across closed-motif orbits including 4-cycles (o8) and tailed triangles (o9), and both random controls collapse these orbits toward zero. Degree-preserving rewiring partially recovers cycle and tailed-triangle counts but remains well below real values. Curated prostate pathway graphs carry non-random local connectivity that cannot be reproduced by edge count or degree sequence alone.

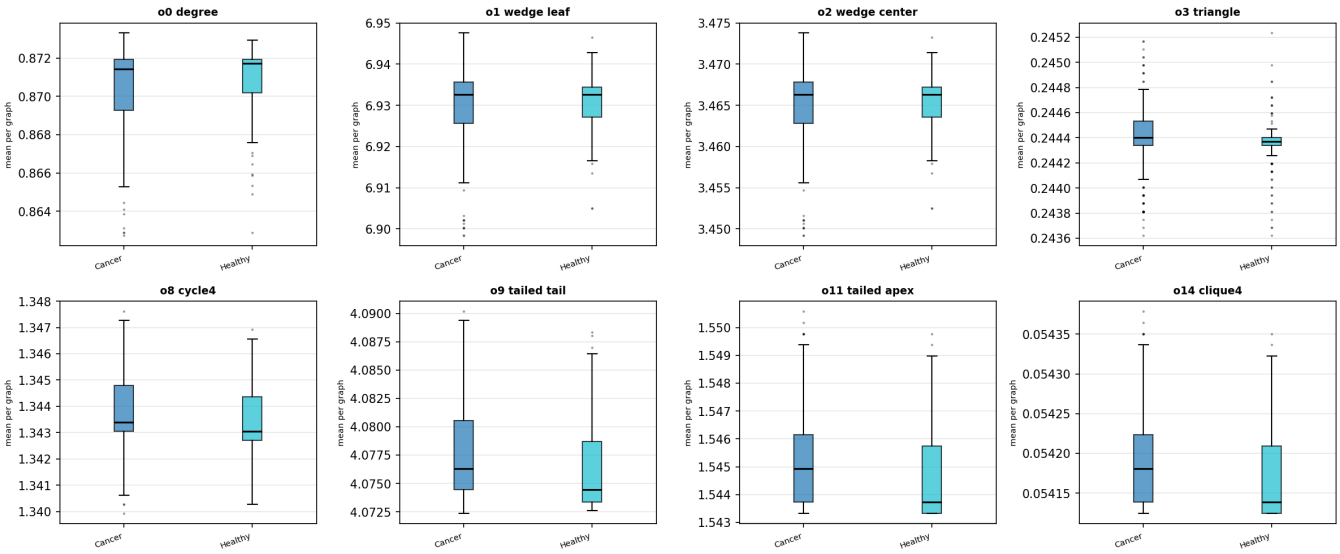

Fig. S11. Graphlet orbit differences between cancer and healthy RARE-Seq binary pathway graphs. Boxplots compare selected ORCA-derived graphlet orbit features between cancer and healthy patient-specific pathway graphs in the RARE-Seq binary classification task. Each panel shows the mean orbit count per graph for a distinct local topology, including degree, wedge, triangle, cycle, tailed-triangle, and clique-related orbit positions. The distributions indicate that cancer and healthy samples exhibit measurable differences in local pathway graph organization.

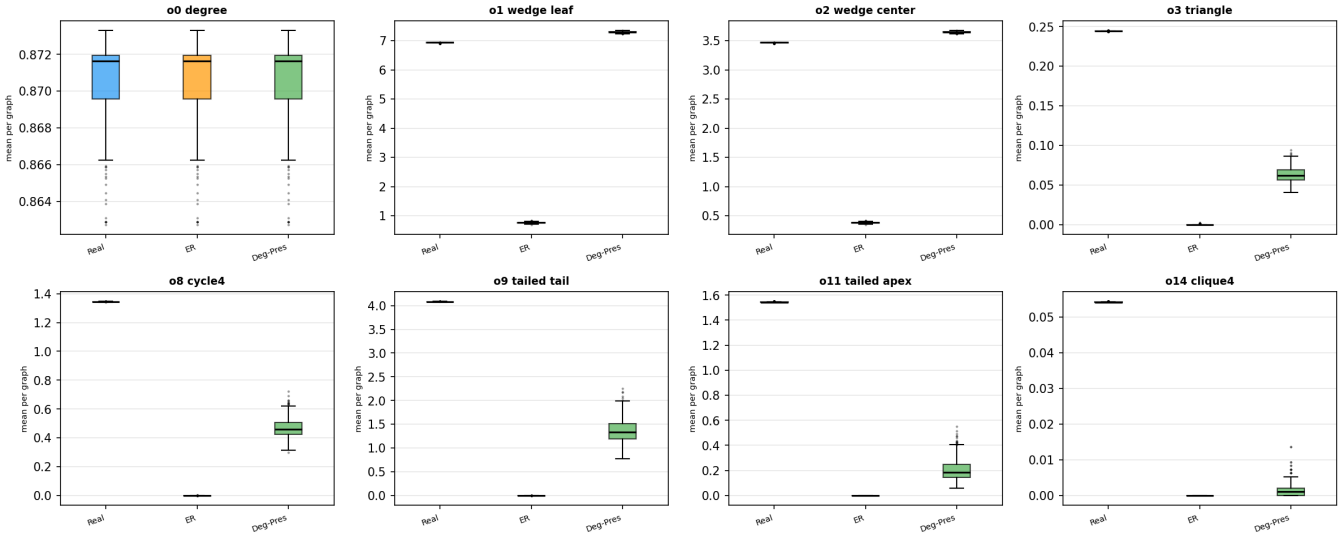

Fig. S12. RARE-Seq binary pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real RARE-Seq binary patient-specific pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot represents the mean orbit count per graph for a specific ORCA-derived local topology. Real pathway graphs show orbit distributions that differ from random controls, supporting that the constructed patient-specific graphs retain non-random pathway structure.

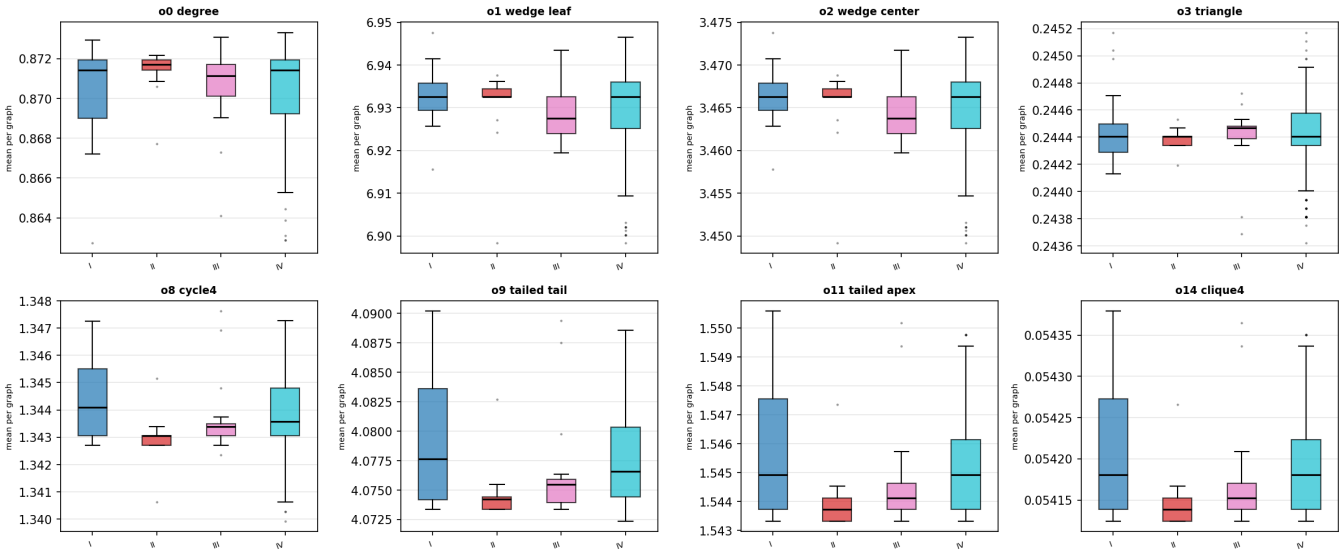

Fig. S13. Graphlet orbit differences across RARE-Seq cancer stages. Boxplots compare selected ORCA-derived graphlet orbit features across cancer stages in RARE-Seq patient-specific pathway graphs. Each panel shows the mean orbit count per graph for a distinct local topology, including degree, wedge, triangle, cycle, tailed-triangle, and clique-related orbit positions. The distributions suggest that local pathway graph organization varies across stage groups, reflecting stage-associated differences in patient-specific pathway structure.

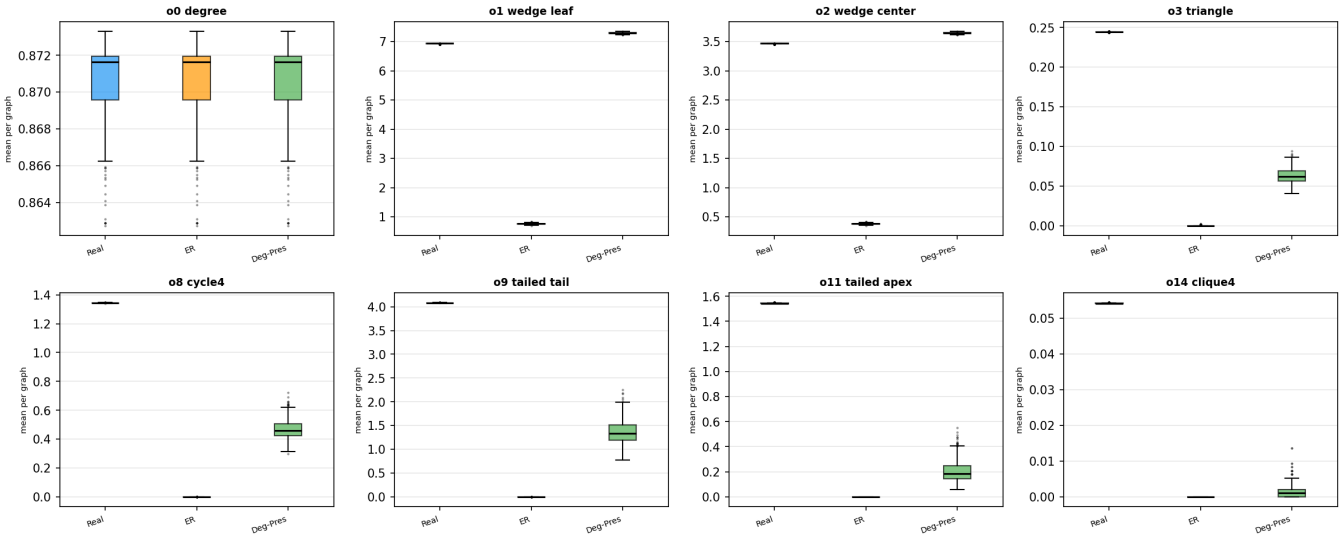

Fig. S14. RARE-Seq stage pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real RARE-Seq stage-task pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot represents the mean orbit count per graph for a specific ORCA-derived local topology. Real pathway graphs display orbit patterns that diverge from random controls, indicating that curated pathway topology contributes structured local graph organization in the stage classification setting.

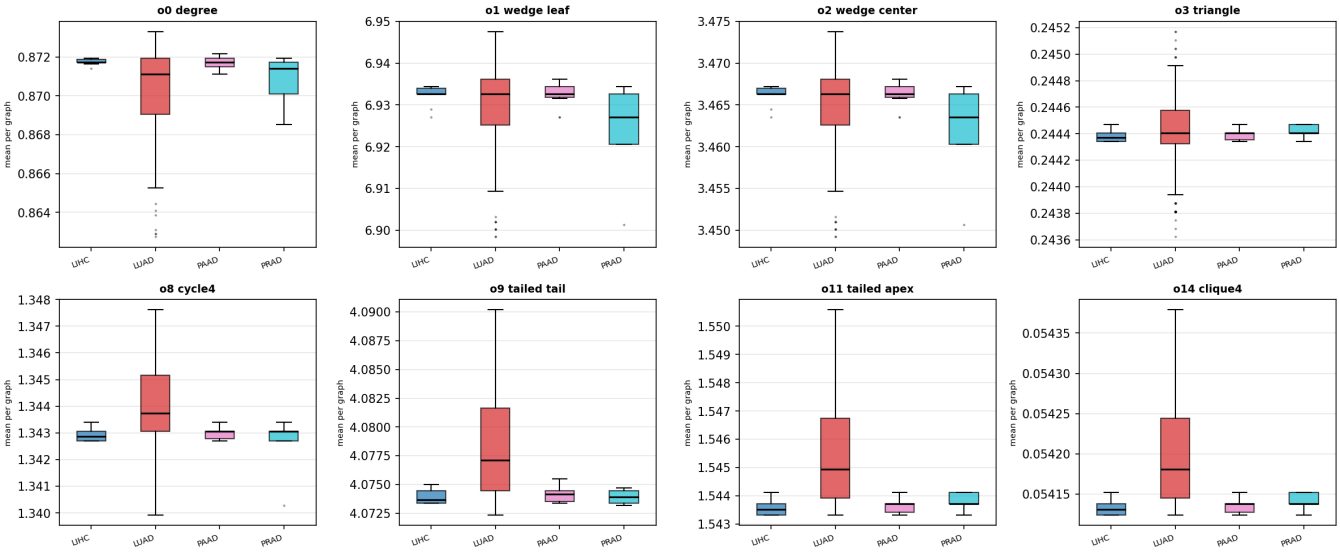

Fig. S15. Graphlet orbit differences across RARE-Seq cancer subtypes. Boxplots compare selected ORCA-derived graphlet orbit features across cancer subtypes in RARE-Seq patient-specific pathway graphs. Each panel shows the mean orbit count per graph for a distinct local topology, including degree, wedge, triangle, cycle, tailed-triangle, and clique-related orbit positions. The distributions indicate subtype-dependent variation in local pathway graph structure, consistent with differences in molecular organization across cancer types.

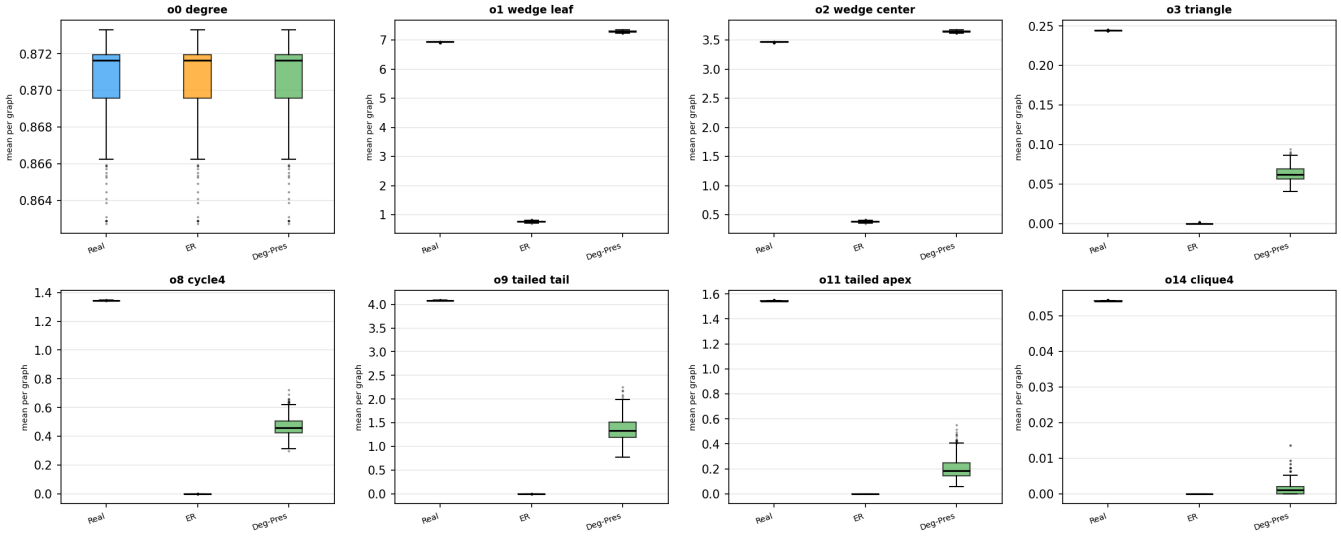

Fig. S16. RARE-Seq subtype pathway graph topology compared with random graph controls. Distributions of selected graphlet orbit features in real RARE-Seq subtype-task pathway graphs compared with Erdős–Rényi and degree-preserving random controls. Each boxplot represents the mean orbit count per graph for a specific ORCA-derived local topology. Real pathway graphs show distinct orbit enrichment and depletion patterns relative to random controls, supporting the presence of biologically constrained pathway structure in the subtype classification setting.

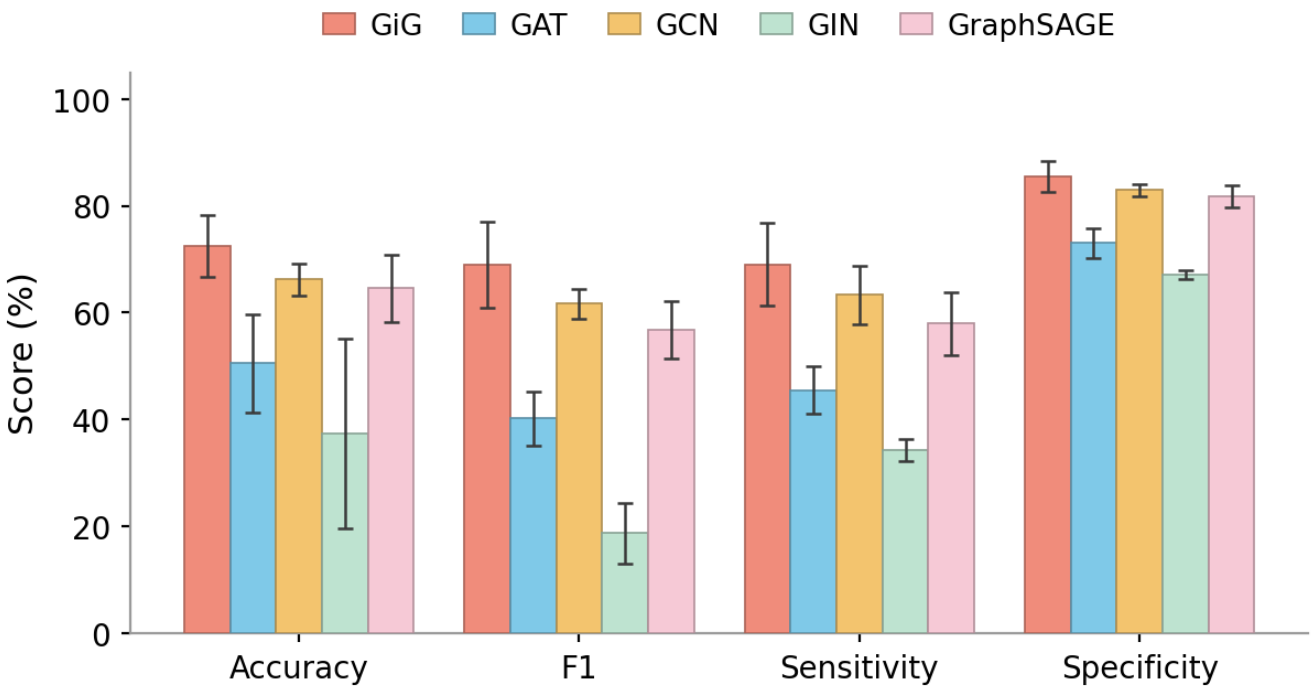

Fig. S17. GiG improves tumor-grade classification in urothelial carcinoma. GiG was compared with GAT, GCN, GIN, and GraphSAGE for predicting tumor grade from patient-specific pathway graphs. GiG achieved the highest mean accuracy, F1 score, sensitivity, and specificity compared to conventional GNN baselines.

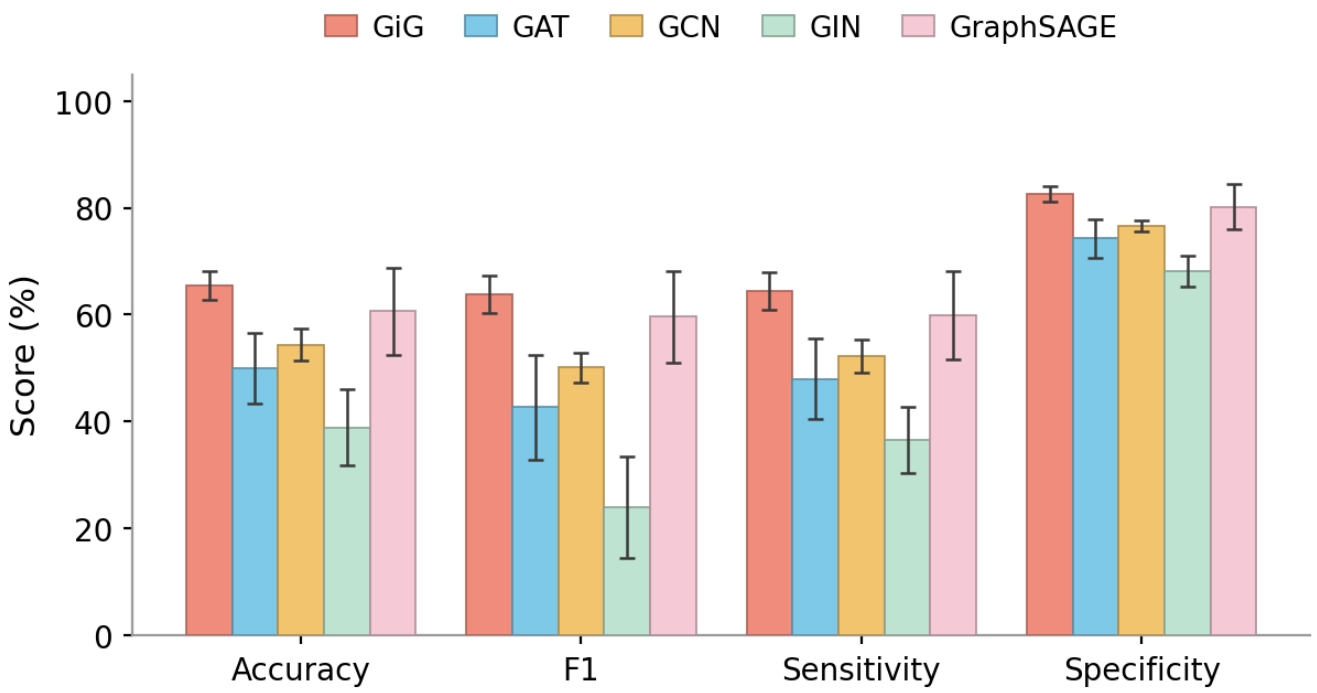

Fig. S18. GiG outperforms baseline models in classifying tumor-stage for urothelial carcinoma. Performance on the GSE32894 tumor-stage prediction task was evaluated using accuracy, F1 score, sensitivity, and specificity. GiG achieved the highest mean performance across all evaluated metrics.

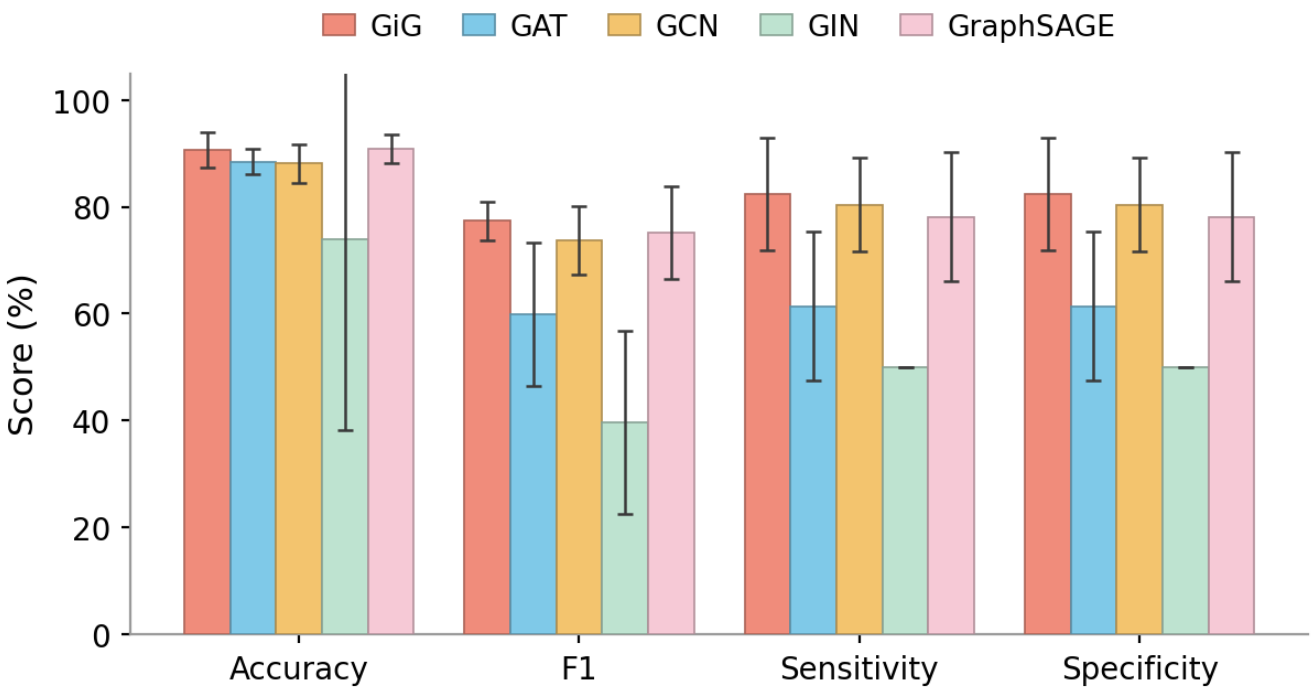

Fig. S19. GiG accurately predicts BRAF mutation status in colorectal cancer. GiG and four conventional GNN architectures were evaluated for BRAF mutation-status classification in GSE39582. Though GraphSAGE produced similarly high accuracy, GiG exhibited lower performance variability across evaluations whilst maintaining high accuracy, F1 score, sensitivity, and specificity.

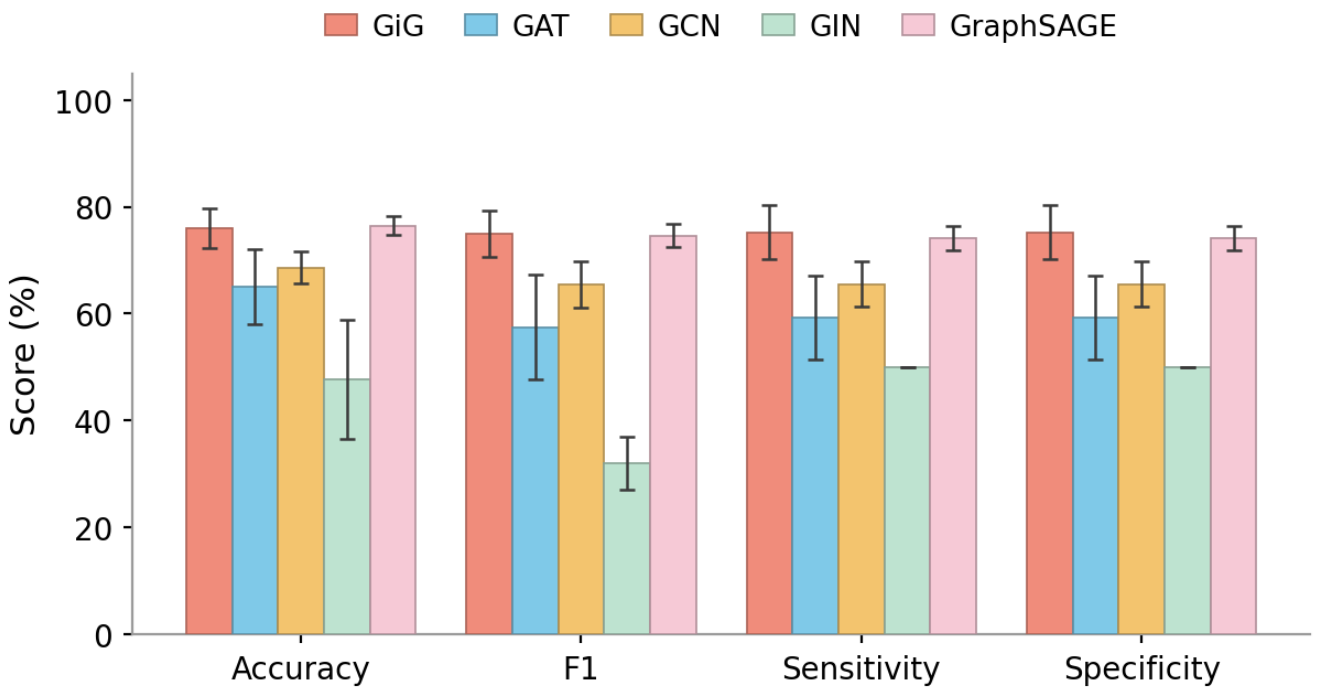

Fig. S20. GiG supports balanced prediction of colorectal tumor location. Classification performance for tumor location in GSE39582 was compared across GiG, GAT, GCN, GIN, and GraphSAGE. GiG achieved the strongest overall balance of F1 score, sensitivity, and specificity while maintaining accuracy comparable to GraphSAGE, the highest-performing baseline for accuracy.

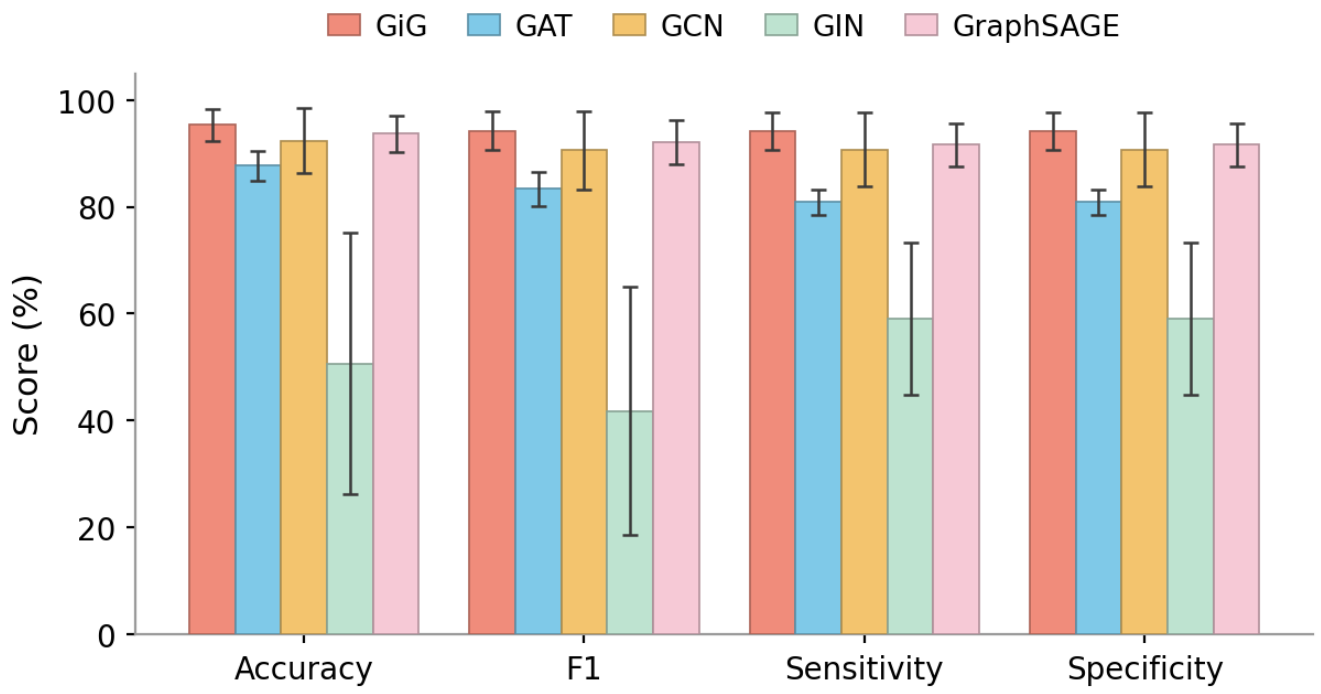

Fig. S21. GiG achieves high performance for gastric cancer EMP subgroup classification. GiG was evaluated against GAT, GCN, GIN, and GraphSAGE for predicting EMP subgroup labels in GSE62254. GiG achieved the highest mean accuracy, F1 score, sensitivity, and specificity, with GraphSAGE showing comparably strong but slightly lower performance. GIN exhibited markedly lower scores and greater variability, indicating less stable classification across evaluations.

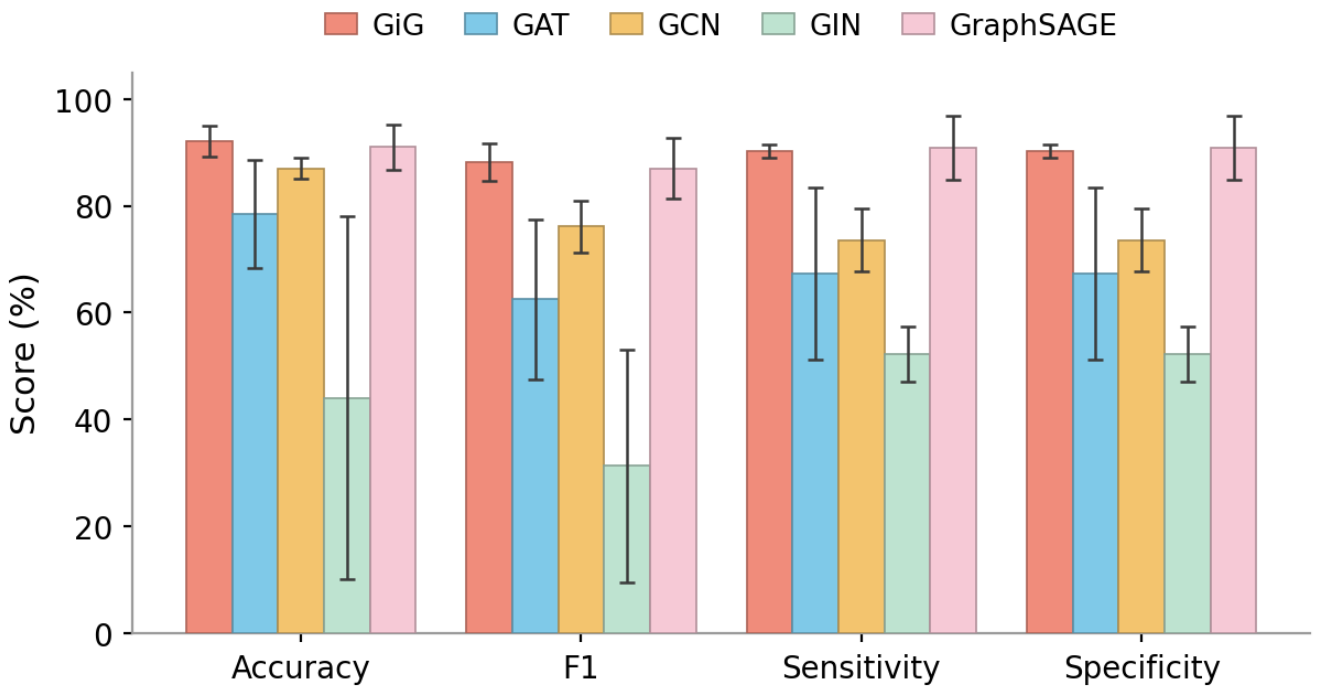

Fig. S22. GiG accurately distinguishes tumor-educated platelets from healthy platelets in a pan-cancer liquid-biopsy cohort. Models were evaluated on the GSE68086 platelet RNA-sequencing dataset for binary classification of tumor-educated platelet samples from patients with cancer versus platelet samples from healthy individuals. The cohort included 228 patients with six malignancies (non-small cell lung cancer, colorectal cancer, pancreatic cancer, glioblastoma, breast cancer, and hepatobiliary carcinoma) and 55 healthy controls. GiG achieved the highest mean accuracy and F1 score while maintaining high sensitivity and specificity comparable to GraphSAGE, demonstrating effective detection of pan-cancer transcriptomic signals in blood platelets.