



# Graph-Based Inference of Cell-Cell Communication Networks from Single-Cell Transcriptomic Profiles for Disease Mechanism Discovery and Biomarker Prioritization across Tissues

Daniel C. Young

Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, UT 84112, USA

## Abstract

*The advent of single-cell transcriptomic technologies has revolutionized our understanding of cellular heterogeneity, offering unprecedented resolution into the molecular states of individual cells. However, tissues are not mere collections of isolated cells; they are highly coordinated ecosystems driven by intricate networks of cell-cell communication. Traditional analyses often overlook the complex topological structures of these interaction networks. This paper presents a comprehensive framework for the graph-based inference of cell-cell communication networks using single-cell transcriptomic profiles. By mapping ligand-receptor expression patterns onto mathematically rigorous graph structures, we facilitate the discovery of novel disease mechanisms and the prioritization of clinical biomarkers across diverse tissue types. This methodology conceptualizes cells as nodes and molecular interactions as weighted edges, enabling the application of advanced graph theory algorithms to decode spatial and functional signaling dynamics. The study explores the theoretical foundations of transcriptomic network inference, evaluates algorithmic approaches for graph construction, and demonstrates the utility of network topology in identifying pathological rewiring in complex diseases such as cancer and autoimmune disorders. Furthermore, centrality metrics and modularity analyses are deployed to pinpoint keystone cellular populations and prioritize robust biomarkers for clinical translation. Ultimately, this graph-centric paradigm bridges the gap between single-cell molecular profiling and systems-level tissue biology, providing a powerful analytical engine for translational medicine and therapeutic target discovery.*

**Keywords:** Single-Cell Transcriptomics, Cell-Cell Communication, Graph Theory, Biomarker Discovery, Systems Biology

## 1. Introduction

### 1.1 Background and Motivation

Biological tissues function through the orchestrated interactions of diverse cell types, maintaining homeostasis and responding to environmental perturbations through dynamic signaling networks. Cell-cell communication is the fundamental mechanism underlying these processes, mediating everything from embryonic development and immune responses to tissue repair and metabolic regulation. Historically, the study of cellular signaling was largely restricted to bulk tissue analysis, which averages the molecular profiles of millions of cells, thereby obscuring the critical contributions of rare cell populations and the spatial context of localized interactions. The emergence of single-cell ribonucleic acid sequencing technologies has systematically dismantled this barrier, providing a high-fidelity snapshot of the transcriptomic state of individual cells [1]. By quantifying gene expression at the single-cell level, researchers can now identify distinct cellular subpopulations, trace developmental lineages, and uncover

the immense cellular heterogeneity that characterizes complex tissues [2].

Despite these advancements, the initial wave of single-cell analysis primarily focused on characterizing isolated cellular states rather than the interactions between them. Recognizing that cellular function is highly context-dependent, the field has recently shifted its focus toward reconstructing the intercellular communication networks that coordinate tissue-level behavior [3]. These networks are typically inferred by examining the co-expression of specific signaling molecules, namely ligands and their corresponding receptors, across different cell types. However, inferring meaningful communication from static transcriptomic data poses significant computational and theoretical challenges [4]. The interactions are inherently multiplexed, highly non-linear, and contextually varied across different physiological states. Consequently, standard statistical correlation methods often fail to capture the hierarchical and distributed nature of biological signaling systems [5]. To overcome these limitations, researchers have increasingly turned to graph theory as a formal mathematical framework for modeling, analyzing, and interpreting cell-cell communication networks [6].

### **1.2 Challenges in Current Approaches**

The transition from single-cell transcriptomic profiles to robust cell-cell communication networks is fraught with methodological hurdles. One of the primary challenges is the pervasive issue of technical noise and dropout events inherent in single-cell sequencing. Due to the minute amount of genetic material available in a single cell, the capture efficiency of transcriptomic protocols is relatively low, resulting in many expressed genes being falsely recorded as having zero expression. This sparsity disproportionately affects the detection of signaling molecules, which are often expressed at lower baseline levels compared to structural or housekeeping genes [7]. Consequently, simple co-expression metrics can lead to massive underestimations of cellular connectivity or, conversely, produce false-positive interactions when imputation algorithms overcorrect for dropout events. Furthermore, many existing computational tools for inferring cellular communication treat interactions as isolated bipartite pairs, ignoring the broader network topology. Biological signaling, however, rarely occurs in a vacuum. A single ligand can bind to multiple distinct receptors, and a single receptor can be activated by an array of different ligands, creating complex signaling cascades and feedback loops. Traditional analytical frameworks struggle to integrate these multi-layered dependencies, often failing to identify secondary or tertiary signaling effects that are crucial for understanding disease propagation [8]. Moreover, when attempting to scale these analyses across multiple tissues or temporal stages, the dimensional complexity of the data increases exponentially. Comparing static networks between healthy and diseased states requires robust alignment techniques that can account for biological variance while isolating pathological network rewiring. Without a cohesive structural methodology, the sheer volume of putative ligand-receptor interactions generated by single-cell pipelines can overwhelm biological interpretation, leaving researchers with long lists of candidate genes but little actionable insight into tissue mechanics or systemic disease mechanisms [9].

### **1.3 Objectives and Contributions**

This paper aims to systematically address these limitations by formalizing a graph-based approach to the inference and analysis of cell-cell communication networks from single-cell transcriptomic data. By representing complex biological tissues as multi-layered graphs, where nodes denote distinct cell types or individual cells and edges represent quantified ligand-receptor interactions, we provide a mathematically rigorous foundation for systems-level biological discovery. The primary objective is to delineate how graph-theoretic principles can be leveraged to extract biologically meaningful patterns from noisy transcriptomic datasets, moving beyond simple pairwise interactions to uncover global signaling architectures. A key contribution of this work is the comprehensive exploration of how network topology can be utilized for disease mechanism discovery. By analyzing how graph structures are rewired in pathological states compared to healthy controls, we can identify systemic disruptions in cellular communication that drive disease progression. Additionally, this paper establishes a framework for biomarker prioritization based on network centrality metrics, proposing that the most clinically relevant biomarkers are often those that occupy strategic topological positions within the communication network, rather than merely those with the highest absolute expression levels. Through detailed methodological discussions and theoretical case studies across diverse tissues, this paper provides a definitive guide for

leveraging graph-based inference to accelerate translational research and precision medicine.

## **2. Literature Review and Background Concepts**

### **2.1 Single-Cell Transcriptomics**

The evolution of transcriptomic analysis has been marked by a continuous drive toward higher resolution and broader throughput. Bulk sequencing, while foundational for understanding general gene expression trends, fundamentally relies on the assumption of a homogenous cell population, an assumption that is severely violated in complex organs such as the brain, liver, or tumor microenvironments. The development of microfluidic droplet-based technologies and split-pool barcoding has enabled the simultaneous profiling of tens of thousands of individual cells in a single experiment [10]. These high-throughput methods have facilitated the construction of comprehensive cell atlases for various organisms and tissues, cataloging the transcriptomic signatures of both highly prevalent and exceedingly rare cell types [11]. However, the analytical paradigms for single-cell data have had to evolve rapidly to keep pace with the technology. Initial bioinformatics pipelines focused heavily on dimensionality reduction techniques and clustering algorithms to group cells based on transcriptomic similarity [12]. These clusters are typically annotated based on known marker genes to define cell populations. While these methods excel at characterizing cellular identity, they treat the tissue as a dissociated mixture, stripping away the spatial and relational context that dictates cellular function *in vivo*. Recognizing this gap, the scientific community has increasingly emphasized the need for computational techniques that can reconstruct the lost relational architecture from dissociated single-cell data. This has led to the development of sophisticated integration techniques that combine single-cell profiles with spatial transcriptomics, attempting to map identified cell clusters back to their original physical coordinates within the tissue architecture [13].

### **2.2 Principles of Cell-Cell Communication**

Cellular communication is orchestrated through a vast vocabulary of signaling molecules. The most common and computationally tractable form of this communication is biochemical signaling, characterized by the secretion of a ligand by a sender cell and the subsequent binding of that ligand to a specific cognate receptor on a receiver cell. This binding event triggers an intracellular signaling cascade that alters the transcriptional state, metabolic activity, or physical behavior of the receiving cell [14]. The biological principles governing these interactions dictate that communication is highly specific, concentration-dependent, and tightly regulated by spatial proximity. Autocrine signaling occurs when a cell targets itself, paracrine signaling affects nearby cells, and endocrine signaling involves ligands traveling through the circulatory system to distant tissues. In the context of transcriptomic inference, tools have been developed to query established databases of known ligand-receptor pairs against the expression data of different single-cell clusters. If cluster A highly expresses a specific ligand and cluster B highly expresses the corresponding receptor, a putative communication link is established [15]. Early iterations of this approach utilized simple thresholding or permutation tests to assess the statistical significance of these co-expression patterns. While valuable, these early models often assumed that messenger ribonucleic acid expression perfectly correlates with functional protein abundance, an assumption that overlooks post-transcriptional modifications, protein degradation rates, and the requirement for multi-subunit receptor complexes [16]. Consequently, modern theoretical frameworks have begun incorporating more complex scoring functions that account for heteromeric receptor complexes, competitive binding by antagonist molecules, and the presence of necessary intracellular downstream signaling components to increase the confidence of inferred interactions [17].

### **2.3 Graph Theory in Biological Networks**

Graph theory, a branch of discrete mathematics focused on the study of interacting networks, provides an ideal paradigm for modeling complex biological systems. In a biological graph, the fundamental entities are represented as nodes, and the relationships or interactions between them are represented as edges [18]. This abstraction has been highly successful in various biological domains, including protein-protein interaction networks, gene regulatory networks, and metabolic pathways. When applied to cell-cell communication, the graph representation allows researchers to capture the distributed and interconnected nature of tissue signaling.

The power of the graph-based approach lies in its ability to quantify the topological properties of the network at multiple scales. At the microscopic level, properties of individual nodes, such as the number of incoming and outgoing edges, can identify which cell types act as primary signaling hubs or critical receivers within the tissue [19]. At the mesoscopic level, graph clustering algorithms can detect highly connected subgraphs or modules, which often correspond to specialized functional niches within the tissue, such as localized immune-regulatory environments or stem cell maintenance hubs. At the macroscopic level, global network properties, such as average path length and overall network density, provide insights into the robustness and efficiency of the tissue-wide communication system [20]. By transitioning from a localized, pairwise perspective to a global, graph-theoretic perspective, researchers can apply advanced network science algorithms, including random walk simulations, network propagation, and spectral graph theory, to uncover hidden structural rules that govern physiological health and disease pathogenesis.

### 3. Methodology for Graph-Based Network Inference

#### 3.1 Data Preprocessing and Integration

The accurate inference of a cell-cell communication graph begins with the rigorous preprocessing of raw single-cell transcriptomic data. The fundamental goal of this stage is to separate genuine biological signals from technical artifacts. Raw read counts must first be normalized to account for variations in sequencing depth across different cells, ensuring that the measured expression levels are comparable. Following normalization, robust batch correction techniques are essential, particularly when integrating datasets across different patients, tissue samples, or experimental platforms. Uncorrected batch effects can severely confound network inference, leading to the creation of artificial signaling edges based on technical variance rather than biological reality [21].

*Table 1: Comparison of Common Data Preprocessing Strategies for Network Inference*

| Preprocessing Strategy     | Primary Objective                           | Impact on Graph Topology                                       | Limitation                                                   |
|----------------------------|---------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------|
| Library Size Normalization | Adjust for sequencing depth variations      | Prevents dense highly-sequenced nodes from dominating edges    | May distort lowly expressed receptor signals                 |
| Variance Stabilization     | Stabilize variance across expression levels | Enhances detection of rare but critical signaling edges        | Computationally intensive for large multi-tissue datasets    |
| Batch Integration          | Remove platform or temporal artifacts       | Ensures edges represent biological rather than technical links | Risk of over-correction removing genuine biological variance |

Once the data is normalized and batch-corrected, handling the sparsity of single-cell data becomes paramount. Dropout events, where a gene is expressed but not detected, can sever critical communication edges in the inferred graph. While some researchers employ imputation algorithms to estimate missing values based on the expression profiles of similar cells, this approach requires extreme caution. Aggressive imputation can introduce false-positive interactions by artificially inflating the co-expression of ligand-receptor pairs, resulting in densely connected but biologically meaningless graphs. Instead of broad imputation, many advanced methodologies rely on cluster-level aggregation, wherein the average or probabilistic expression of signaling molecules is calculated across robustly defined cell clusters [22]. This strategy leverages the statistical power of the collective cluster to mitigate individual cell dropouts while preserving the overall biological architecture.

#### 3.2 Graph Construction and Weighting

The construction of the cell-cell communication graph involves defining the nodes, establishing the edges, and assigning appropriate weights to represent interaction strength. Nodes can be defined at different levels of granularity. In a macro-level graph, each node represents a distinct cell type or state cluster

identified during the preprocessing phase. In a micro-level or cell-resolution graph, each node represents an individual cell. The choice of node resolution significantly impacts the complexity of the subsequent analysis; cluster-level graphs are computationally efficient and easier to interpret, whereas cell-resolution graphs provide deeper insights into continuous developmental trajectories and intra-cluster heterogeneity [23]. Edges are established by mapping the transcriptomic data against curated databases of biochemical interactions. An edge is drawn from a sender node to a receiver node if the sender expresses a specific ligand and the receiver expresses the corresponding receptor. Because a single sender can communicate with a receiver via multiple different ligand-receptor pairs, the resulting structure is often a multi-graph, where multiple parallel edges can exist between the same two nodes. To quantify the strength or probability of these interactions, sophisticated weighting functions are employed. These functions typically consider the product of the average ligand expression in the sender and the average receptor expression in the receiver, often incorporated with a statistical permutation test to ensure the observed co-expression is significantly higher than expected by random chance [24]. Furthermore, advanced weighting schemes may incorporate spatial constraints, penalizing or eliminating edges between cell populations that are known to be physically separated by impermeable tissue barriers.

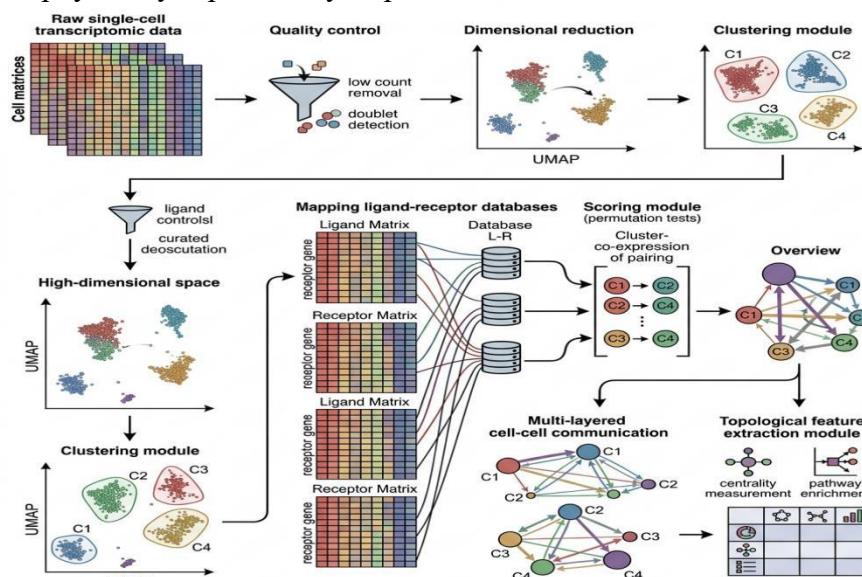


Figure 1: Comprehensive Schematic of Graph

### 3.3 Network Inference Algorithms

Once the weighted graph is constructed, various network inference algorithms are utilized to distill the complex web of interactions into interpretable biological patterns. One prominent approach involves the use of information theory to identify the most critical signaling pathways. By calculating the mutual information or utilizing directed information flow algorithms, researchers can trace the primary routes of signal transduction through the tissue network. This helps identify dominating pathways that coordinate major physiological responses, distinguishing them from baseline homeostatic signaling chatter [25]. Another critical class of algorithms involves graph partitioning and modularity optimization. These techniques seek to divide the network into distinct sub-networks or communities, where the internal connectivity among nodes within a community is significantly higher than the connectivity between different communities. In the context of cell-cell communication, these modules often represent specific functional microenvironments, such as the invasive front of a tumor or a localized inflammatory niche. By identifying these modules, researchers can map the spatial and functional compartmentalization of the tissue. Additionally, modern machine learning approaches, particularly graph neural networks, are increasingly being applied to these structures. Graph neural networks can learn low-dimensional embeddings of the nodes that capture both their transcriptomic identity and their topological context within the communication network. These embeddings can subsequently be used to predict hidden interactions, classify disease states based on network structure, or simulate the downstream effects of

targeted receptor inhibition [26].

#### 4. Discovery of Disease Mechanisms Across Tissues

##### 4.1 Cross-Tissue Communication Dynamics

Disease processes rarely remain confined to isolated cellular populations; they typically propagate through the alteration of systemic communication networks that span multiple tissues. Graph-based inference offers a unique vantage point for investigating these cross-tissue dynamics. By generating communication graphs for multiple interconnected organs, researchers can identify endocrine signaling axes that mediate systemic physiological responses. For instance, the metabolic crosstalk between adipose tissue, the liver, and skeletal muscle is critical for maintaining glucose homeostasis. Single-cell transcriptomic profiling of these tissues, combined with graph-based network integration, allows for the precise mapping of circulating adipokines and hepatokines to their respective receptor complexes on distant target cells [27]. Analyzing these multi-tissue graphs requires specialized computational techniques capable of handling the disparate baseline expression levels and cellular compositions inherent to different organs. Network alignment algorithms are employed to match structurally similar subgraphs across different tissues, facilitating the identification of conserved signaling motifs. Furthermore, temporal network analysis can be applied to time-series single-cell data to observe how cross-tissue communication evolves over the course of a biological process, such as aging or the progression of an infection. By tracking changes in edge weights and node centralities over time, researchers can pinpoint the exact temporal window during which systemic signaling breaks down, providing crucial insights into the sequence of pathogenic events.

##### 4.2 Pathological Rewiring of Networks

A core application of graph-based inference is the comparative analysis of communication networks between healthy and diseased states. Disease pathogenesis is frequently characterized not merely by the presence or absence of specific cell types, but by the pathological rewiring of the interactions between them. In this context, rewiring refers to the gain of anomalous signaling edges, the loss of essential homeostatic edges, or a drastic shift in the interaction weights within the network graph [28]. By computing the differential network topology, researchers can systematically identify the cellular communication failures that drive disease.

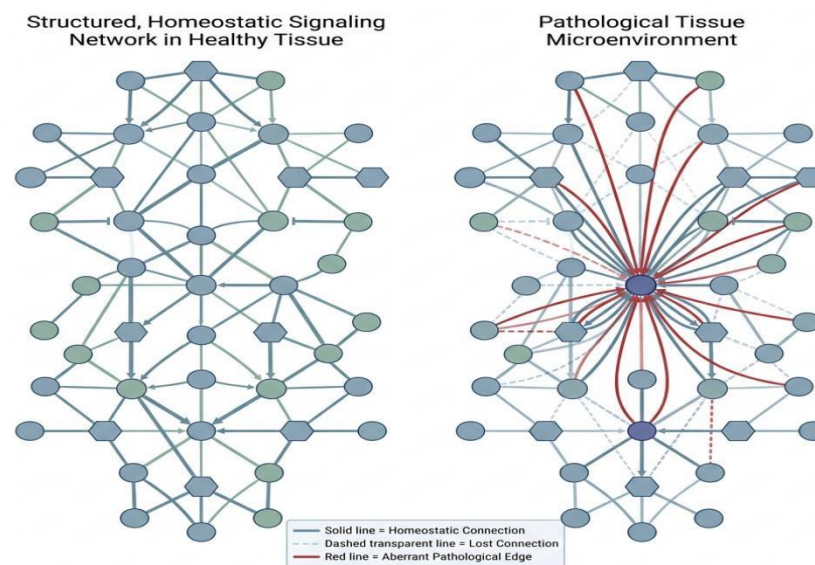


Figure 2: Differential Network Rewiring Map between Healthy and Pathological Tissue Microenvironments

For example, in fibrotic diseases, the homeostatic communication between epithelial cells and resident fibroblasts is severely disrupted. Graph analysis of fibrotic tissue often reveals the emergence of a hyper-connected, self-sustaining signaling module dominated by pro-fibrotic ligands and their receptors. This pathological module operates somewhat autonomously, driving continuous extracellular matrix deposition independently of normal regulatory signals. By quantifying the structural differences between the healthy and fibrotic networks, algorithms can identify the specific ligand-receptor pairs responsible

for bridging the transition into the pathological state. This approach shifts the analytical focus from individual dysregulated genes to the dysregulated relational architecture, offering a more holistic understanding of the disease mechanism.

#### 4.3 Case Studies in Complex Diseases

The utility of graph-based network inference is most vividly demonstrated through its application to complex, highly heterogeneous diseases such as cancer and autoimmune disorders. In the context of oncology, the tumor microenvironment is a complex ecosystem comprising malignant cells, immune infiltrates, stromal cells, and vasculature. Tumors frequently hijack normal cellular communication pathways to evade immune surveillance, stimulate angiogenesis, and promote metastasis. Graph analysis of single-cell tumor data has repeatedly revealed that malignant cells often position themselves as central hubs in the communication network, emitting immunosuppressive signals while simultaneously heavily receiving growth-promoting factors from cancer-associated fibroblasts [29]. The topology of this malignant network often correlates strongly with clinical outcomes; highly integrated, modular tumor networks tend to be more resilient to therapeutic intervention. In autoimmune diseases, the pathology stems from a failure of the immune system to recognize self-antigens, leading to chronic inflammation and tissue destruction. Graph-based inference allows researchers to map the intricate dialogue between resident tissue cells and infiltrating autoreactive immune cells. Analysis of tissues afflicted by conditions such as rheumatoid arthritis or inflammatory bowel disease often shows a massive expansion of inflammatory signaling edges, accompanied by the collapse of regulatory feedback loops that normally suppress immune overactivation. By mapping these dynamic changes onto a mathematical graph, researchers can pinpoint the precise cellular interactions that initiate the inflammatory cascade and identify the specific receptor pathways that fail to provide adequate inhibitory signals, thereby uncovering deep mechanistic insights into autoimmune pathogenesis.

### 5. Biomarker Prioritization Strategies

#### 5.1 Topological Analysis of Candidate Biomarkers

The translation of transcriptomic discoveries into clinical utility heavily depends on the identification of robust, reliable biomarkers for disease diagnosis, prognosis, and therapeutic monitoring. Traditional biomarker discovery often relies on differential gene expression analysis, selecting targets based solely on the magnitude of their upregulation in diseased tissues. However, high expression does not necessarily equate to high biological importance; a highly expressed gene may be a downstream bystander rather than a driving force of the pathology. Graph-based inference introduces a paradigm shift by prioritizing biomarkers based on their topological significance within the cell-cell communication network [30]. Centrality metrics derived from graph theory are instrumental in this prioritization process. Degree centrality, which counts the total number of incoming and outgoing signaling edges for a given node, helps identify cellular populations that act as primary communicators. Genes encoding ligands or receptors that mediate these high-degree connections are strong biomarker candidates because their perturbation would disproportionately disrupt the network. Betweenness centrality measures how frequently a specific node or interaction acts as a bridge along the shortest signaling path between other cell types. Biomarkers with high betweenness centrality are critical bottlenecks in tissue communication; targeting these pathways can effectively dismantle pathological signaling cascades even if the genes themselves are not the most highly expressed in the tissue. Furthermore, algorithms like PageRank can be adapted to biological graphs to identify genes that hold structural influence by virtue of being connected to other highly influential nodes.

*Table 2: Network Centrality Metrics for Biomarker Prioritization*

| Centrality Metric      | Network Interpretation                     | Biological and Clinical Relevance                                     |
|------------------------|--------------------------------------------|-----------------------------------------------------------------------|
| Degree Centrality      | Total volume of communication for a target | Identifies major signaling hubs and highly active therapeutic targets |
| Betweenness Centrality | Control over information flow              | Pinpoints bottleneck receptors                                        |

|                        |                                                  |                                                                               |
|------------------------|--------------------------------------------------|-------------------------------------------------------------------------------|
|                        | between modules                                  | crucial for disrupting pathological cascades                                  |
| Eigenvector Centrality | Influence based on the connectivity of neighbors | Highlights regulatory genes controlling critical secondary signaling pathways |

## 5.2 Clinical Relevance and Validation

The ultimate goal of prioritizing biomarkers via network topology is to identify targets with high clinical relevance and translatability. Biomarkers selected through graph-based inference often exhibit superior robustness across heterogeneous patient cohorts because they represent structural dependencies of the disease rather than transient expression spikes. For instance, a receptor identified as a key topological bridge in a tumor microenvironment network is likely to remain a crucial vulnerability even if the tumor mutates downstream pathways, provided the overarching network architecture remains intact. However, computational prioritization must be coupled with rigorous experimental and clinical validation. Candidates identified through single-cell graph analysis require subsequent verification using orthogonal techniques such as multiplexed immunofluorescence, spatial transcriptomics, or functional *in vitro* assays to confirm the physical presence and functional necessity of the predicted interactions. Spatial validation is particularly critical, as the computational graphs inferred from dissociated cells must ultimately map onto the physical geometry of the intact tissue. When effectively validated, these topologically prioritized biomarkers offer a significant advantage for precision medicine, providing a more rational basis for the design of targeted therapies, combination treatments, and diagnostic panels that accurately reflect the underlying systems biology of the disease.

## 6. Conclusion and Future Directions

### 6.1 Summary of Findings

This paper has detailed the transformative potential of graph-based inference for decoding cell-cell communication networks from single-cell transcriptomic data. By conceptualizing complex tissues as interconnected mathematical graphs, this methodological framework transcends the limitations of traditional pairwise expression analysis. The formalization of cellular populations as nodes and specific ligand-receptor interactions as weighted edges allows for the application of advanced network science algorithms. Through this lens, the structural and functional nuances of tissue ecosystems become quantifiable and analytically tractable. We have demonstrated how analyzing the differential topology of these networks provides unparalleled insights into disease mechanisms across various tissues. Pathological conditions are rarely localized anomalies; they manifest as systemic rewirings of communication architectures, characterized by the emergence of malignant signaling hubs, the disruption of critical homeostatic bridges, and the functional isolation of regulatory cell types. Furthermore, this graph-centric approach revolutionizes biomarker prioritization. By utilizing network centrality metrics such as degree, betweenness, and eigenvector centrality, researchers can identify keystone molecules and pivotal interactions that govern disease states, ensuring that clinical targets are selected based on their structural importance rather than mere expression abundance.

### 6.2 Future Perspectives

While graph-based inference of transcriptomic networks has achieved significant milestones, several critical frontiers remain. The most pressing advancement is the integration of spatial transcriptomic data directly into the graph construction algorithms. Current methodologies largely infer interaction probabilities based on expression profiles alone, which can result in biologically implausible edges if the communicating cells are physically separated by impermeable barriers *in vivo*. Incorporating precise physical coordinates and distance constraints into the edge weighting functions will drastically enhance the fidelity of the inferred networks. Additionally, the transition toward multi-omic network inference represents a crucial evolutionary step for the field. Relying solely on transcriptomic data overlooks the essential roles of post-translational modifications, metabolite signaling, and epigenetic regulation in cellular communication. Future graph models must evolve into heterogeneous, multi-layered networks capable of simultaneously analyzing single-cell ribonucleic acid, chromatin accessibility, and spatial proteomic data. Finally, the integration of temporal dynamics through time-series single-cell sequencing

will enable the modeling of network trajectories, allowing researchers not only to diagnose the current state of a disease but to predict its future progression and simulate its response to various pharmacological interventions. By continuing to refine these mathematical and computational frameworks, graph-based systems biology will undoubtedly remain at the forefront of disease mechanism discovery and translational therapeutics.

## References

- Zhang, Y., & Chen, Y. (2025, July). The role of machine learning in reducing healthcare costs: The impact of medication adherence and preventive care on hospitalization expenses. In *Proceedings of the International Conference on Implementing Generative AI into Telecommunication and Digital Innovation 2025* (pp. 179-185).
- Tao, Xiaoyu, et al. "A structural roadmap to gut health: A comprehensive review of edible insect chitin–protein complexes as alternative protein-derived, structure-specific prebiotics." *International Journal of Biological Macromolecules* (2026): 151312.
- Yin, L., Tong, Y., Islam, Z. H., Zhang, K., Xie, R., Burger, J.,... & Wang, B. (2024). Targeted NAD<sup>+</sup> Delivery for intimal Hyperplasia and Re-endothelialization: a novel anti-restenotic therapy approach. *bioRxiv*, 2024-02.
- Wang, G., Zhong, Z., Zhou, Y., Li, Y., Wu, E., Cheah, W. P., ... & Shen, L. (2026). X-PCR: A Benchmark for Cross-modality Progressive Clinical Reasoning in Ophthalmic Diagnosis. *arXiv preprint arXiv:2604.20350*.
- Kabra, M., Shahi, P. K., Wang, Y., Sinha, D., Spillane, A., Newby, G. A., ... & Pattnaik, B. R. (2022). Nonviral base editing of KCNJ13 mutation preserves vision in an inherited retinal channelopathy. *bioRxiv*, 2022-07.
- Wang, G., Ren, J., Shen, L., Cheah, W. P., & Qu, R. (2026). M-MambaS: Multimodal Mamba for small lesion segmentation. *Pattern Recognition*, 113923.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., et al. (2019). Comprehensive integration of single-cell data. *Cell*, 177(7), 1888–1902.
- Liu, Z. Q., Xiong, Y. J., Yu, X., Lv, T., & Xu, H. Y. (2026). Female reproductive factors, genetic susceptibility and incident osteoporosis risk. *Climacteric*, 1-8.
- Xiao, Y., Xiao-yue, Z., Yue, W., Ruo-tao, L., Xiang-jie, L., Xing-yuan, W., ... & Zhen-yi, J. (2022). Use of computed tomography for the diagnosis of surgical sarcopenia: Review of recent research advances. *Nutrition in Clinical Practice*, 37(3), 583-593.
- Chen, Z., Zhu, Q., & Zeng, Y. (2026). MWED-Former: A Multi-Scale Weather-Aware Transformer for Emergency Department Demand Forecasting with Weather Bias Correction. *Fundamental Scientific Reports in Multidisciplinary Areas*, 2 (02), 132-145.
- Wang, G., Zhou, Y., Deng, K., Cheah, W. P., Qu, R., Ren, J., & Shen, L. (2026). SurgCoT: Advancing Spatiotemporal Reasoning in Surgical Videos through a Chain-of-Thought Benchmark. *arXiv preprint arXiv:2604.20319*.
- Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., et al. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- Zhao, K., Li, G., Li, C., Jiang, T., Wang, W., & Yang, Z. (2026). Identification Markers for *Salvia miltiorrhiza* and Its Close Relatives Based on Cell Wall Component Characteristics. *Engineered Science*, 40, 2122.
- Shi, M., & Chow, S. C. (2024). Statistical Analysis for Rating Scale in Clinical Trials. *Therapeutic Innovation & Regulatory Science*, 58(5), 930-945.
- Li, S., Yu, X., Ma, X., Wang, Y., Guo, J., Wang, J.,... & Yin, S. (2024). Optimizing a classification model to evaluate individual susceptibility in noise-induced hearing loss: Cross-sectional study. *JMIR Public Health and Surveillance*, 10, e60373.
- Xu, X., Bowtell, J., Young, W. R., Fong, D. T., & Williams, G. K. (2025). Distinct patterns in neuromuscular adaptation to repeated perturbations in chronic ankle instability. *Journal of NeuroEngineering and Rehabilitation*.
- GAO, T., LIANG, D., CHEN, X., & TAO, X. (2025). PREDICTIVE ANALYTICS IN PEDIATRIC

- MENTAL HEALTH: A MACHINE LEARNING APPROACH TO DEPRESSION DETECTION BY PHYSICAL ACTIVITY AND BEHAVIORAL HEALTH INDICATORS. INTERCIENCIA.
- Xu, X., Bowtell, J., Young, W. R., Fong, D. T., & Williams, G. K. (2026). Increased common corticospinal input during eyes-closed unilateral stance in people with chronic ankle instability. *Scientific Reports*.
- Verhoven, B., Tong, Y., Chlebeck, P., Zhong, W., Zeng, W., Jennings, H., ... & Al-Adra, D. P. (2025). Attenuating Ischemia and Reperfusion Injury Using NAD<sup>+</sup>-Loaded Nanoparticles in Mouse Kidneys. *Transplantation Direct*, 12(1), e1890.
- Yang Y, Li G, Zhong Y et al. scTenifoldXct: a semi-supervised method for predicting cell-cell interactions and mapping cellular communication graphs. *Cell Syst* 2023a;14:302–11.e4.
- Wang, W., Chen, Y., Xiong, Z., Sun, Z., Wang, Z., Ma, L., ... & Li, X. (2026). Network toxicology and multidimensional bioinformatics analysis reveal the shared mechanism of action of bisphenol A and phthalates in glioblastoma. *Ecotoxicology and Environmental Safety*, 309, 119650.
- Zhao, K., Wu, X., Han, G., Sun, L., Zheng, C., Hou, H., ... & Shi, Z. (2024). *Phyllostachys nigra* (Lodd. ex Lindl.) derived polysaccharide with enhanced glycolipid metabolism regulation and mice gut microbiome. *International Journal of Biological Macromolecules*, 257, 128588.
- Yang, Y., Li, G., Zhong, Y., Xu, Q., Chen, B.-J., Lin, Y.-T., Chapkin, R. S., & Cai, J. J. (2023). Gene knockout inference with variational graph autoencoder learning single-cell gene regulatory networks. *Nucleic Acids Research*, 51(13), 6578–6592.
- Li, Y., Liao, X., Yu, X., Xiao, Y., Tao, X., & Zhong, T. (2025). Mediating Role of the ANGPTL3/TFPI Protein Ratio in Regulating T-Cell Surface Glycoprotein CD5 Levels on Knee Osteoarthritis (KOA): A Mendelian Randomization Study. *International Journal of Molecular Sciences*, 26(10), 4471.
- Zhang, Y., & Yu, X. (2025). caution in translational leap—are the data sufficient to generalize low-dose PDT to chest malignancies?. *Journal for immunotherapy of cancer*, 13 (7), e013151.
- Ye, M. F., Chele, D., Calatrava, J., Alrmali, A., Huang, W. X., & Wang, H. L. (2026). Guided transpositional bone blocks in esthetic zone: Surgical technique and case report. *Clinical Advances in Periodontics*, 16(1), 197-204.
- Chai L, Zhang J, Fernando W G D, et al. Detection of blackleg resistance gene Rlm1 in double-low rapeseed accessions from Sichuan province, by kompetitive allele-specific PCR[J]. *The Plant Pathology Journal*, 2021, 37(2): 194.
- Xu, Z., Lu, Y., Tao, X., Cheng, W., Tian, J., Tang, W., & Wu, D. Modelling the geographic spread of HIV among MSM in Guangdong, China: a metapopulation model considering the impact of pre-exposure prophylaxis *Philosophical Transactions A*.
- Chen, Z., Gao, P., Lee, Y., Barthelemy, J., Zhou, L., & Wang, L. (2025, June). Optimizing Efficiency and Visual-Textual Alignment for LLM-Based Radiology Report Generation. In *2025 IEEE International Conference on Multimedia and Expo (ICME)* (pp. 1-6). IEEE.
- Poplin, R., Chang, P.-C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., et al. (2018). A universal SNP and small-indel variant caller using deep neural networks. *Nature Biotechnology*, 36(10), 983–987.
- Shen, Q., Sun, Z. H., Xu, Y. M., Hu, Q. B., Zhang, W. H., & Huang, S. A. (2025). Associations between complete-blood-count-derived inflammatory markers and gastric ulcer: A cross-sectional study. *World Journal of Gastroenterology*, 31(38), 111298.