Single-Cell Transcriptomics-Guided Network Pharmacology of Traditional Chinese Medicine in Tumor Microenvironment Regulation

Niceles Wetsan
School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR, USA.
contactnicolas@oregonstate.edu

Ishaan Freeman
Department of Computer Science, University of North Texas, Denton, TX, USA.
contactishaan@unt.edu

Dominik Nearris
Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
dominik.norris@missouri.edu

Abstract

The tumor microenvironment is now understood as a heterogeneous, dynamically regulated ecosystem whose cellular states, molecular gradients, and spatial architectures jointly determine tumor progression and therapeutic response. Traditional Chinese Medicine formulations offer multi-component interventions that may modulate this ecosystem through coordinated action on multiple targets, yet their mechanistic opacity has limited clinical translation. This paper develops a systems-oriented framework in which single-cell transcriptomics functions as a high-resolution observational layer, while network pharmacology supplies an inferential architecture for mapping herbal constituents to cellular regulatory programs. The integration of these layers enables the identification of cell-type-specific responses, ligand-receptor perturbations, and multicellular network shifts associated with TCM exposure. We examine structural trade-offs among data granularity, computational complexity, model interpretability, and clinical generalizability. The discussion addresses modular architectures for data harmonization, network community detection, signature matching, and machine learning-based prioritization, together with considerations of data provenance, reproducibility, fairness, and regulatory oversight. Rather than proposing a single computational protocol, the paper emphasizes the governance of evidence generation, the sustainability of data infrastructures, and the policy conditions under which single-cell-guided network pharmacology can evolve into a trustworthy translational instrument. The resulting perspective situates TCM research within broader precision oncology and systems medicine agendas, highlighting both the promise and the infrastructural burdens of integrating traditional therapeutics with modern high-resolution molecular science.

Keywords

single-cell transcriptomics; network pharmacology; traditional Chinese medicine; tumor microenvironment; systems medicine; computational governance; data infrastructure.

1. Introduction

Cancer is increasingly conceptualized not as a disorder of malignant cells alone but as a systemic dysregulation of tissue organization in which transformed cells, immune populations, fibroblasts, endothelial cells, and extracellular matrix components form a co-evolving microenvironment. The tumor microenvironment is characterized by metabolic competition, immunosuppressive signaling, and spatial heterogeneity that together shape tumor progression and therapeutic vulnerability [1]. A growing body of clinical and experimental evidence indicates that durable therapeutic responses often depend on re-educating this microenvironment rather than simply eliminating malignant cells [2]. This shift has motivated interest in therapeutic modalities that act across multiple cellular compartments and regulatory pathways.

Single-cell transcriptomics has emerged as a transformative measurement technology for resolving the cellular composition and dynamic states of tumor ecosystems. By profiling thousands of individual cells in parallel, single-cell RNA sequencing enables the decomposition of tissue into discrete cell types, transitional states, and lineage trajectories that would remain hidden in bulk transcriptomic measurements [3]. The technology has proven especially valuable for mapping immune cell heterogeneity, identifying rare regulatory subsets, and reconstructing intercellular communication circuits that sustain the malignant niche [4]. As a result, single-cell transcriptomics provides an observational foundation for understanding how therapeutic agents perturb the tumor microenvironment at an unprecedented level of resolution.

In parallel, network pharmacology has developed as a systems-oriented alternative to single-target drug discovery. Network pharmacology treats disease as a perturbation of molecular interaction networks and seeks interventions that restore functional balance through multiple synergistic or compensatory nodes [5]. This perspective aligns with the multi-component, multi-target character of many natural product-derived medicines and traditional formulations. Network biology offers formalisms for representing protein-protein interactions, transcriptional regulation, and signaling pathways as interconnected graphs whose topological properties can be linked to disease modules and drug action [6]. Within this framework, traditional Chinese medicine is especially suited for network-based investigation because its formulations typically contain numerous phytochemical constituents that may collectively influence multiple targets.

Traditional Chinese medicine network pharmacology has therefore become an active research field aimed at explaining how herbal formulas exert therapeutic action across molecular networks, disease modules, and physiological systems [7]. Early computational strategies adapted transcriptional signature matching approaches such as the Connectivity Map to infer relationships between herbal compounds, gene expression states, and disease phenotypes [8]. Gene set enrichment analysis has similarly enabled the interpretation of transcriptomic signatures in terms of known biological pathways and regulatory programs [9]. A network community-based model for identifying synergistic combinations from herbal medicines and complex systems [10] illustrates the continuing evolution of computational methods designed to handle the combinatorial complexity of herbal interventions. These tools make it possible to formulate testable hypotheses about which cell types and signaling circuits respond to a given herbal formula.

This paper examines the integration of single-cell transcriptomics and network pharmacology as a systems architecture for investigating TCM regulation of the tumor microenvironment. The analysis proceeds from the premise that high-resolution observational data and network-

based mechanistic inference are not merely complementary, but structurally interdependent. The goal is not to provide a single computational recipe, but to articulate the architectural trade-offs, governance requirements, and translational conditions that determine whether this integration can generate robust, reproducible, and clinically meaningful knowledge.

2. The Tumor Microenvironment as a Complex Adaptive System

The tumor microenvironment is more than a static collection of supporting cells. It is a complex adaptive system in which cancer cells and host cell populations continuously sense and respond to one another through paracrine factors, metabolic intermediates, extracellular vesicles, and biomechanical cues. Malignant cells can recruit monocytes, polarize macrophages, suppress cytotoxic T cells, activate fibroblasts, and induce angiogenesis, thereby constructing a niche that supports invasion and therapeutic resistance [11]. Each of these processes involves feedback loops that are nonlinear and context dependent. A perturbation in one cell type can propagate through intercellular communication networks to alter the state of distant cellular compartments.

Understanding TCM action in this context requires moving beyond a single-pathway perspective. Herbal formulas often contain compounds with weak individual affinities but broad collective effects that may simultaneously influence immune activation, stromal remodeling, vascular permeability, and tumor metabolism. In a complex adaptive system, such distributed perturbations can produce emergent outcomes that are difficult to infer from reductionist assays. For example, a compound that mildly inhibits tumor cell proliferation may become therapeutically meaningful only when combined with changes in macrophage polarization or T cell recruitment. The relevant question is therefore not simply which molecular target a compound binds, but how the entire microenvironment reorganizes after exposure.

Single-cell transcriptomics is uniquely suited to capture this reorganization because it records the state of numerous interacting cell populations in a shared coordinate space. When applied to treated and untreated tissues, single-cell profiling can reveal shifts in cell type abundance, transcriptomic state, and inferred intercellular communication. These observations can then be interpreted through network pharmacology to identify candidate active constituents and their probable targets. The analytical challenge is that the tumor microenvironment is highly heterogeneous across patients, tumor types, anatomical sites, and treatment histories. Any computational framework must therefore be designed to accommodate biological variance while still extracting reproducible regulatory patterns.

The complex adaptive systems perspective also has implications for model interpretation. Static network diagrams may overstate the determinism of therapeutic action if they ignore dynamic feedback and stochastic transitions. A robust systems framework should therefore combine snapshots from single-cell measurements with longitudinal or perturbation-based inference, enabling the reconstruction of state transitions rather than fixed pathways. In TCM research, this means that network pharmacology should be used not as a final mechanistic explanation but as a hypothesis-generating engine whose outputs are iteratively tested against high-resolution cellular observations. This iterative loop is essential for translating multi-component pharmacology into actionable insights.

3. Single-Cell Transcriptomics as a High-Resolution Observational Layer

Single-cell transcriptomics provides a high-dimensional representation of cellular identity that is particularly valuable for dissecting the tumor microenvironment. Unlike bulk RNA

sequencing, which averages gene expression across all cells in a tissue, single-cell methods preserve the molecular distinctions among malignant clones, T cell subsets, myeloid populations, fibroblasts, and endothelial cells. This resolution enables researchers to identify rare immunosuppressive populations, exhausted effector cells, and transitional states that may be disproportionately important for therapeutic response [3]. Such cell states are often obscured in aggregate measurements yet may serve as key targets for TCM formulations.

The technological pipeline for single-cell transcriptomics involves tissue dissociation, cell capture, library preparation, sequencing, and computational preprocessing. Each step introduces potential biases, including cell-type loss during dissociation, gene dropout, batch effects, and doublet contamination [4]. These technical sources of variation must be carefully managed before biological conclusions can be drawn. Data integration methods, batch correction algorithms, and rigorous quality control procedures are therefore foundational components of the observational layer. Without such procedures, single-cell data can produce misleading cell clusters, pseudo-transitions, and spurious ligand-receptor interactions.

For TCM studies, the observational layer must also address experimental design questions. The choice of tumor model, dose regimen, sampling time point, and treatment duration can profoundly affect observed cell state changes. Because herbal formulations often act slowly and through multiple mechanisms, short-term in vitro exposure may fail to capture the full spectrum of microenvironmental remodeling. Longitudinal single-cell profiling in suitable preclinical models, and eventually in patient-derived specimens, is needed to distinguish direct drug effects from compensatory or adaptive responses. The sampling design should be matched to the expected time course of immune infiltration, stromal remodeling, and transcriptional reprogramming.

The computational analysis of single-cell data involves dimensionality reduction, unsupervised clustering, trajectory inference, and differential state analysis. These methods generate cell type annotations, state transition graphs, and gene regulatory network predictions that can be linked to network pharmacology databases. A central architectural challenge is that single-cell analyses often produce hundreds of candidate cell states and thousands of differentially expressed genes. Prioritizing which of these states and genes are most relevant to TCM action requires a principled integration with prior knowledge about herbal compounds, target proteins, and disease modules. This is precisely the function of the network pharmacology layer.

4. Network Pharmacology of Traditional Chinese Medicine: Architectural Foundations

Network pharmacology originated as a response to the limitations of linear, single-target drug discovery. It represents biological systems as networks of molecular entities and uses graph-theoretic concepts to analyze disease mechanisms and drug effects [5]. The network perspective is especially natural for TCM because a single herbal formula may contain dozens or hundreds of constituents, each with multiple potential targets. Network pharmacology transforms this complexity into a computationally tractable problem by constructing herb-compound-target-disease networks and then identifying modules that may be responsible for therapeutic activity [6]. These modules can be visualized and analyzed using software platforms that support network construction, topological analysis, and integrative visualization [12].

Pathway databases provide a curated substrate for interpreting network relationships. Resources such as the Kyoto Encyclopedia of Genes and Genomes organize molecular

interactions into signaling, metabolic, and immune pathways that can be mapped onto gene expression changes [13]. This knowledge base enables researchers to infer whether a set of herbal targets converges on particular cellular programs, such as inflammatory signaling, metabolic reprogramming, or apoptosis regulation. When such pathway enrichment is combined with single-cell data, it becomes possible to ask cell-type-specific questions: which ligands, receptors, and downstream transcription factors are most affected in specific tumor or immune cell populations.

The chemical diversity of natural products adds further complexity. Many TCM compounds are derived from plant, fungal, or mineral sources and have been optimized by evolutionary selection for interactions with biological macromolecules. Natural products have historically contributed a large fraction of approved anticancer agents and remain a rich source of bioactive scaffolds [14]. However, the pharmacokinetic properties of these compounds, their bioavailability, and their metabolic transformations in vivo are often poorly characterized. Network pharmacology must therefore incorporate not only molecular target information but also compound availability, tissue distribution, and metabolism when translating network predictions into biological hypotheses.

Traditional Chinese medicine network pharmacology has developed several computational paradigms to address these challenges. Target prediction, network construction, module detection, and pathway enrichment are commonly used to generate mechanistic hypotheses from herbal formula composition [7]. Transcriptional signature comparison methods such as the Connectivity Map allow researchers to relate compound-induced gene expression changes to disease signatures [8]. Gene set enrichment analysis provides a statistical framework for linking these signatures to curated biological processes [9]. The identification of synergistic herbal combinations can be further supported by community detection and network-based scoring approaches that evaluate how multiple compounds jointly influence different parts of a disease network [10]. Together, these methods form a modular architecture for converting chemical and genomic data into testable mechanistic claims.

5. Integrative System Architecture for Multi-Modal Mechanistic Inference

The integration of single-cell transcriptomics and network pharmacology can be understood as a layered information-processing architecture. The first layer ingests raw single-cell data and produces cell state representations. The second layer maps herbal compounds to potential targets and constructs molecular interaction networks. The third layer projects cell state changes onto these networks to infer which cell types and signaling programs are most strongly associated with treatment. The fourth layer uses machine learning and statistical modeling to prioritize compounds, targets, and cell states for further validation [15].

This layered architecture has several structural advantages. It separates observational noise from mechanistic inference, allowing each layer to be improved independently. It also supports modular replacement of computational components: a new batch correction method can be inserted without redesigning the network projection layer, and an improved target prediction algorithm can be incorporated without altering the single-cell processing pipeline. At the same time, the architecture creates challenges of error propagation and interpretability. Errors in cell type annotation or target prediction can cascade through subsequent layers, producing plausible but incorrect mechanistic hypotheses.

A key design decision is the granularity at which cell states should be mapped onto network nodes. One approach treats each cell type as a node and connects cell types through ligand-

receptor interactions inferred from single-cell data. Another approach treats individual cells as nodes in a cell-cell communication graph, then overlays compound-target information to identify which interactions are most likely perturbed. The choice of granularity involves a trade-off between biological interpretability and computational cost. Coarse cell type-level networks are easier to visualize and validate, but they may miss rare transitional states. Fine-grained cell-level networks preserve heterogeneity but are more sensitive to technical noise and require substantial computational resources.

The network projection step itself must address the multi-scale nature of biological organization. Molecular pathways inside a cell, intercellular communication between adjacent cells, and tissue-level gradients all contribute to therapeutic response. A comprehensive architecture should therefore represent intracellular regulatory networks, cell-cell interaction networks, and higher-level tissue modules as distinct but interacting layers. TCM formulations may act at multiple scales simultaneously, for example by altering an intracellular kinase pathway in tumor cells while also changing cytokine secretion that affects neighboring immune cells. Integrative models that explicitly account for these cross-scale interactions are more likely to capture the full effect of herbal therapy.

Machine learning has become increasingly important for integrating these heterogeneous data types. Supervised models can be trained to predict treatment response from single-cell features, while unsupervised models can identify latent cell states or network modules associated with drug sensitivity [15]. However, the high dimensionality of single-cell data and the small size of most TCM experimental cohorts create a risk of overfitting. Regularization, cross-validation, and independent experimental validation are therefore essential. The most useful machine learning applications in this domain are those that generate interpretable biological hypotheses rather than black-box predictions, because the ultimate goal is mechanistic understanding that can guide formulation design and clinical development.

6. Data Infrastructure, Interoperability and Computational Governance

The integration of single-cell transcriptomics and TCM network pharmacology depends on a data infrastructure capable of handling heterogeneous data types, metadata standards, and analytical workflows. Single-cell datasets vary widely in sequencing platform, cell capture method, reference genome, and annotation schema. Herbal formula databases vary in compound nomenclature, botanical source, preparation method, and dosage information. Integrating these resources requires shared vocabularies, ontologies, and application programming interfaces that allow researchers to trace a compound from its botanical source through its predicted targets to the cell states observed in a single-cell experiment.

The FAIR principles for scientific data management provide a useful governance framework for this integration. Data should be findable, accessible, interoperable, and reusable across institutional and disciplinary boundaries [16]. In practice, this means that single-cell datasets should be deposited in repositories with standardized metadata describing the experimental conditions, species, tissue type, and treatment. Herbal compound databases should expose robust identifiers that can be linked to chemical structures, target predictions, and published pharmacological evidence. Without such interoperability, integrative analyses are difficult to reproduce, and the value of publicly funded research is reduced.

Computational governance also encompasses the provenance of analytical results. A network pharmacology prediction that links a specific herbal constituent to a cell state in the tumor microenvironment is the product of many sequential decisions: quality control thresholds,

normalization parameters, clustering resolution, target prediction algorithms, and network construction rules. If these decisions are not documented, the resulting hypothesis cannot be independently audited or replicated. Workflow management systems, containerization, and semantic provenance tracking are therefore important infrastructural components. They enable a reviewer to reconstruct the entire analytical chain and assess whether a conclusion is robust to alternative parameter choices.

The infrastructure should also support iterative model refinement. Single-cell data may reveal that a predicted target is not expressed in the expected cell type, prompting revision of the compound-target network. Conversely, network pharmacology may identify a pathway that is not directly measured by standard single-cell transcriptomics, prompting additional targeted assays. This iterative cycle requires data platforms that can accept new experimental results, update network models, and recalculate predictions without disrupting the underlying architecture. Such platforms are more than databases; they are computational ecosystems that enable distributed collaboration across pharmacology, genomics, oncology, and traditional medicine.

7. Robustness, Uncertainty and Validation

A central concern in single-cell-guided network pharmacology is the reliability of high-dimensional inferences. Single-cell data are noisy, sparse, and sensitive to experimental conditions. Network pharmacology predictions are built on incomplete target annotations and simplified interaction models. When these two uncertainty sources are combined, the resulting mechanistic hypotheses can be highly variable. Robustness analysis should therefore be a formal component of the computational framework. This may involve perturbing clustering parameters, subsampling cells, repeating target prediction with different algorithms, and testing whether the core network modules remain stable under such perturbations.

The distinction between statistical significance and biological reproducibility is especially important. A cell state that is differentially abundant after TCM treatment may be statistically significant in one experiment but fail to replicate in another due to batch effects or biological variability. Similarly, a predicted compound-target interaction may be supported by computational docking but lack experimental confirmation. Machine learning models trained on small single-cell cohorts can produce optimistic performance estimates if validation is not carefully designed [17]. Independent validation in additional cohorts, experimental models, or orthogonal measurement platforms is therefore necessary before mechanistic claims can be considered credible.

The tumor microenvironment also evolves over time and under selective pressure. Cancer cells may develop resistance to immunomodulatory signals, and immune cells may become exhausted or suppressive. These dynamic processes mean that a single snapshot cannot fully capture the therapeutic mechanism. Longitudinal sampling and trajectory inference can help distinguish transient drug effects from durable reprogramming. In TCM research, where treatment may involve prolonged exposure or cycling between formulations, the timing of sampling is an essential variable. A network module that appears activated at one time point may be replaced by a compensatory module at a later point.

Uncertainty quantification should be communicated honestly in publications and translational documents. A common failure mode in integrative studies is to present the most plausible mechanistic narrative while omitting alternative interpretations. To avoid this, computational workflows should produce not only ranked hypotheses but also measures of confidence,

robustness, and independent evidence [18]. By making uncertainty explicit, the field can avoid overclaiming and build a cumulative evidence base that is more resilient to refutation. This is particularly important for TCM, where historical use and cultural authority may sometimes substitute for rigorous mechanistic validation.

8. Fairness, Ethics and Regulatory Governance

The integration of single-cell genomics and TCM pharmacology raises ethical and regulatory questions that extend beyond technical validation. One issue is the representativeness of training data. Single-cell atlases are often derived from selected patient populations and tumor types, which may not reflect the genetic, environmental, and cultural diversity of patients who might receive TCM-based therapies. If predictive models are trained primarily on data from one population, their performance may be uneven when deployed in another. Fairness-aware machine learning and inclusive cohort design are therefore relevant to the development of single-cell-guided network pharmacology [19].

Another concern is the governance of traditional knowledge. TCM formulations are not simply chemical libraries; they are embedded in cultural, historical, and medical traditions. When computational models identify active constituents or synergistic combinations, questions arise about intellectual property, benefit sharing, and the rights of communities that have preserved this knowledge. A purely extractive approach, in which modern genomics and network pharmacology mine traditional formulations without reciprocal engagement, may reproduce patterns of epistemic injustice. Governance frameworks should include mechanisms for acknowledging traditional knowledge holders, negotiating benefit sharing, and avoiding the misappropriation of cultural heritage [19].

Regulatory pathways for herbal medicines are heterogeneous across jurisdictions. In some countries, TCM formulas are regulated as traditional medicines with less stringent clinical evidence requirements. In others, they must meet the same efficacy and safety standards as conventional pharmaceuticals. Single-cell and network pharmacology evidence could support regulatory submissions by providing mechanistic rationale, dose-response relationships, and biomarkers of response. However, regulators are unlikely to accept computational predictions as primary evidence of efficacy. The computational layer must be paired with well-designed clinical studies that use transcriptomic or immune profiling as exploratory endpoints [20].

Transparency and accountability are essential for trustworthy deployment. Algorithms that rank herbal combinations or predict patient-specific responses should be explainable to clinicians, patients, and regulators. Black-box models may offer high predictive accuracy in retrospective data but are difficult to audit for safety and fairness. Interpretable models that identify specific cell states, pathways, or targets provide a more defensible basis for clinical decisions. The choice between predictive performance and interpretability is a governance trade-off that should be made explicit in any deployment strategy.

9. Clinical Translation and Deployment Pathways

Translating single-cell-guided network pharmacology findings into clinical practice requires a staged approach that begins with preclinical mechanism elucidation and moves through biomarker validation, formulation optimization, and clinical trial design. Preclinically, single-cell profiling can identify the immune and stromal cell states that are most responsive to a TCM formula. Network pharmacology can then link these cell states to candidate active constituents and propose which combinations are most likely to be synergistic. This

knowledge can guide the development of more standardized extracts or fractionated formulations with better defined composition and bioavailability.

Clinical deployment poses additional challenges. TCM formulations are often individualized and adjusted over time, which complicates the use of fixed molecular biomarkers. Single-cell profiling of tumor biopsies is invasive, costly, and not routinely available in many clinical settings. A pragmatic deployment pathway may therefore use single-cell discoveries to derive lower-dimensional gene signatures that can be measured by bulk RNA sequencing, targeted panels, or circulating immune profiling. These derived signatures can serve as practical biomarkers in clinical trials and routine care. The single-cell layer thus functions as a discovery engine, while the deployment layer relies on more accessible assays.

The regulatory environment for combination therapies also matters. A TCM formula that modulates the tumor microenvironment may be most useful in combination with immune checkpoint inhibitors, chemotherapy, or radiotherapy. Yet herbal medicines can influence drug metabolism, immune activation, and toxicity profiles. Clinical translation must therefore address herb-drug interactions and safety pharmacology. Network pharmacology can help predict potential interactions by mapping herbal constituents to metabolic enzymes and transport proteins. Single-cell transcriptomics can reveal whether a formula affects immune checkpoints or inflammatory pathways that might potentiate or antagonize conventional therapy [21].

The path to clinical adoption depends on evidence quality and stakeholder trust. Oncologists, pharmacists, and patients may be skeptical of TCM unless its mechanisms are articulated in a language that aligns with contemporary molecular medicine. Single-cell transcriptomics and network pharmacology can provide that language by converting traditional therapeutic concepts into cell states, ligand-receptor circuits, and pathway modules. However, the computational evidence must be integrated with clinical observation, patient-reported outcomes, and real-world data. A hybrid evidence architecture that respects both traditional knowledge and modern measurement is more likely to gain acceptance than computational modeling alone.

10. Sustainability, Policy and Global Health Implications

The long-term viability of single-cell-guided network pharmacology for TCM depends on sustainable data, computational, and human resources. Single-cell data generation is expensive and technically demanding, while TCM chemical and pharmacological databases require continuous curation. Without sustained funding, these resources may become fragmented or obsolete. Academic consortia, public repositories, and international collaborations can help distribute costs and establish shared standards. Policy makers should support open data infrastructures that enable secondary analysis, while also protecting patient privacy and traditional knowledge rights.

The global health dimension is significant. TCM and other traditional medical systems are widely used in many regions, yet they are often excluded from formal precision oncology initiatives. Integrating traditional medicines into systems pharmacology frameworks could broaden the therapeutic options available in resource-limited settings, where novel immunotherapies are frequently inaccessible. Single-cell transcriptomics is unlikely to be routinely available in all settings, but the mechanistic insights derived from it can inform the development of affordable, standardized botanical preparations. This translational model requires careful attention to quality control, supply chain integrity, and equitable access.

Policy instruments can encourage responsible innovation. Research funding agencies may require data sharing, provenance documentation, and independent validation for integrative TCM studies. Regulatory agencies may develop guidance for computational evidence in traditional medicine applications. Health technology assessment bodies may define criteria for evaluating multi-component herbal therapies within precision oncology frameworks. These policies should balance the need for scientific rigor with the recognition that traditional medicines operate within distinct cultural and clinical contexts.

International collaboration is essential for harmonizing standards and avoiding duplication. Single-cell atlases, herbal compound databases, and network pharmacology tools are increasingly developed by distributed teams. Governance arrangements should ensure that contributors from low- and middle-income countries receive appropriate credit and access. The design of federated data systems can enable cross-border analysis without requiring centralized data sharing, thereby addressing privacy and data sovereignty concerns. These infrastructural choices are not merely technical; they shape who participates in knowledge production and who benefits from resulting therapies.

The integration of single-cell transcriptomics and network pharmacology also has implications for research training. The field requires investigators who can move between computational biology, pharmacology, oncology, and traditional medicine. Capacity building should therefore include interdisciplinary curricula, shared mentorship, and collaborative research environments. Without a new generation of researchers fluent in both systems modeling and traditional medicine, the translational potential of this integration will remain unrealized. The sustainability of the field ultimately depends on human infrastructure as much as on technical platforms.

11. Conclusion

Single-cell transcriptomics-guided network pharmacology offers a powerful systems framework for investigating how traditional Chinese medicine regulates the tumor microenvironment. The high-resolution observational capabilities of single-cell profiling reveal the cellular states and intercellular circuits that define the tumor ecosystem, while network pharmacology provides a structured approach for linking multi-component herbal interventions to those states. When integrated within a layered computational architecture, these methods enable researchers to move beyond single-target explanations and toward a multicellular, network-level understanding of TCM action.

The realization of this potential depends on more than algorithmic innovation. It requires robust data infrastructure, transparent provenance, careful uncertainty quantification, and governance that respects both scientific rigor and cultural knowledge. The structural trade-offs among data granularity, computational complexity, interpretability, and clinical feasibility must be managed explicitly. By treating single-cell-guided network pharmacology as a governance-sensitive systems engineering challenge, the field can produce evidence that is not only mechanistically plausible but also reproducible, fair, and translationally meaningful. In this way, traditional medicine and modern systems science can be brought into a productive alignment that serves the broader goals of precision oncology and global health.

References

1. Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. Cell, 144(5), 646–674.

2. Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., Coussens, L. M., Gabrilovich, D. I., Ostrand-Rosenberg, S., Hedrick, C. C., Vonderheide, R. H., Pittet, M. J., Jain, R. K., Zou, W., Howcroft, T. K., Woodhouse, E. C., Weinberg, R. A., & Krummel, M. F. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. Nature Medicine, 24(5), 541–550.

3. Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., III, Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive integration of single-cell data. Cell, 177(7), 1888–1902.

4. Papalexi, E., & Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. Nature Reviews Immunology, 18(1), 35–45.

5. Hopkins, A. L. (2008). Network pharmacology: The next paradigm in drug discovery. Nature Chemical Biology, 4(11), 682–690.

6. Barabási, A.-L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. Nature Reviews Genetics, 5(2), 101–113.

7. Li, S., & Zhang, B. (2013). Traditional Chinese medicine network pharmacology: Theory, methodology and application. Chinese Journal of Natural Medicines, 11(2), 110–120.

8. Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., Lerner, J., Brunet, J.-P., Subramanian, A., Ross, K. N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S. A., Haggarty, S. J., Clemons, P. A., Wei, R., Carr, S. A., Lander, E. S., & Golub, T. R. (2006). The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. Science, 313(5795), 1929–1935.

9. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences, 102(43), 15545–15550.

10. Wang, Y., Yao, J., Sui, Y., Jiang, H., Ma, B., Lai, S., ... & Tan, N. (2026). HerbSyner_Finder: a network community-based model for identifying synergistic combinations from herbal medicines and complex systems. Targetome, 2(2).

11. Joyce, J. A., & Pollard, J. W. (2009). Microenvironmental regulation of metastasis. Nature Reviews Cancer, 9(4), 239–252.

12. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498–2504.

13. Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research, 28(1), 27–30.

14. Newman, D. J., & Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. Journal of Natural Products, 83(3), 770–803.

15. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine

learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477.

16. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018.

17. Wang, Y. (2025, June). RAGNet: Transformer–GNN–Enhanced Cox–Logistic Hybrid Model for Rheumatoid Arthritis Risk Prediction. In Proceedings of the 2025 International Conference on Health Informatization and Data Analytics (pp. 90-94).

18. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

19. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679.

20. Zitvogel, L., Tesniere, A., & Kroemer, G. (2006). Cancer despite immunosurveillance: Immunoselection and immunosubversion. Nature Reviews Immunology, 6(10), 715–727.

21. Tsimberidou, A. M., Fountzilas, E., Nikanjam, M., & Kurzrock, R. (2020). Review of precision cancer medicine: Evolution of the treatment paradigm. Cancer Treatment Reviews, 86, 102019.