Network Pharmacology and Molecular Docking-Based Investigation of Herbal Compounds Against Cancer-Related Signaling Pathways

Bnaind Maliek
Department of Computer Science, University of Houston, Houston, TX, USA.
anand.work@uh.edu

Parth Kapoor
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
parthmail@unh.edu

Krish Trivedi
Department of Computer Science, University of North Texas, Denton, TX, USA.
trivedi369@unt.edu

Abstract

Cancer is a heterogeneous family of diseases characterized by the progressive dysregulation of intracellular and intercellular signaling networks. The limitations of single-target therapeutic strategies, including acquired resistance and systemic toxicity, have motivated the search for multicomponent interventions that can modulate multiple nodes within cancer-related pathways. Herbal medicines represent a rich source of structurally diverse compounds with potential polypharmacological activity, yet their mechanistic complexity poses substantial challenges for evidence-based investigation. This paper examines the integration of network pharmacology and molecular docking as a systems-level computational framework for investigating herbal compounds against cancer-related signaling pathways. The discussion is oriented toward architectural design, data integration, computational trade-offs, validation robustness, reproducibility, governance, and translational sustainability. Network pharmacology provides a high-level topological representation of compound-target-disease relationships, while molecular docking contributes structural filters for evaluating binding plausibility. The combined pipeline enables prioritization of herbal constituents for experimental and clinical follow-up, but it also introduces systemic vulnerabilities related to data bias, pathway incompleteness, scoring function uncertainty, and regulatory ambiguity. This paper argues that the utility of such platforms depends not only on algorithmic accuracy but also on transparent data provenance, interoperable infrastructure, rigorous external validation, and ethically aligned deployment. Future developments should emphasize adaptive learning architectures, federated data governance, cross-domain knowledge graphs, and robust frameworks for translating computational predictions into sustainable and equitable cancer therapeutic research.

Keywords

network pharmacology, molecular docking, cancer signaling, herbal compounds, systems architecture, reproducibility, governance, fairness.

1. Introduction

Cancer is increasingly conceptualized as a systems disease in which malignant phenotypes emerge from the coordinated dysregulation of signaling, metabolic, and immune-related networks. The traditional reductionist model of drug discovery, which seeks highly selective ligands for a single molecular target, has produced clinically valuable agents, but its limitations are now widely recognized in the context of tumor heterogeneity, pathway redundancy, and adaptive resistance. A systems-oriented perspective therefore emphasizes the importance of understanding how multiple molecular perturbations propagate through cellular networks and how therapeutic interventions may be designed to restore regulatory balance rather than simply inhibit one protein. This shift in perspective has significant implications for the investigation of herbal medicines, which typically contain numerous chemical constituents that may interact with multiple targets in a weak yet potentially synergistic manner. Network pharmacology has been advanced as a paradigm that aligns well with this multicomponent and multitarget reality [1]. By combining network science, cheminformatics, omics data, and systems biology, network pharmacology enables the construction of disease-relevant interaction maps and the projection of herbal compounds onto those maps for mechanistic inference. At the same time, the cellular network perspective has deep roots in the broader field of network biology, which has demonstrated that biological systems are organized according to scale-free topologies, modular architectures, and hub-like regulatory proteins [2]. These structural properties are highly relevant to cancer, because oncogenic alterations frequently affect high-centrality nodes whose perturbation can have widespread downstream consequences.

Herbal medicines have been used for centuries in multiple medical traditions, and their potential to contribute to cancer care is increasingly investigated through modern computational and experimental approaches. However, the translation of traditional empirical knowledge into rigorous biomedical evidence is complicated by the chemical complexity of herbal extracts, variability in preparation methods, and the absence of standardized annotations for many bioactive constituents. Network pharmacology addresses these challenges by representing herbal formulas as collections of compounds and by mapping their predicted targets onto disease-associated signaling networks. The resulting models can identify candidate mechanisms, rank compounds by network impact, and suggest combinations that may act through complementary pathways. Molecular docking adds a structural dimension to this analysis by estimating whether a given compound can bind to a predicted target with favorable geometry and energy. Docking does not replace network-level inference, but it provides an essential physical plausibility filter that reduces the proportion of false-positive network predictions. The integration of network pharmacology and molecular docking therefore constitutes a layered computational architecture in which topological, cheminformatic, and structural evidence are combined to prioritize compounds for further validation.

This paper provides a long-form academic examination of the systems-level issues arising from the application of network pharmacology and molecular docking to herbal compounds in cancer research. Rather than focusing narrowly on algorithmic details or specific case studies, the discussion emphasizes architectural design, structural trade-offs, data infrastructure, reproducibility, robustness, fairness, governance, and translational sustainability. The central argument is that the scientific value of such computational platforms depends on the careful management of uncertainty across multiple layers. A network model may be computationally elegant but biologically misleading if the underlying interaction data are incomplete or biased. A docking score may be mathematically precise but mechanistically unreliable if the scoring

function cannot capture solvation effects, protein flexibility, or allosteric modulation. Similarly, a promising prediction may fail to translate if the supporting data are not interoperable, the computational environment is not reproducible, or the regulatory framework does not accommodate complex multicomponent interventions. By treating network pharmacology and molecular docking as elements of a broader socio-technical infrastructure, this paper identifies critical bottlenecks and proposes a forward-looking agenda for more trustworthy, equitable, and sustainable computational investigation of herbal medicines in oncology.

2. Conceptual Foundations of Network Pharmacology in Cancer Research

Network pharmacology is best understood as a conceptual and methodological response to the failure of single-target assumptions in complex diseases such as cancer. In this framework, disease is not merely the result of a single dysfunctional protein but rather the emergent property of a perturbed molecular network. The therapeutic goal shifts from complete inhibition of one target to the modulation of a set of targets whose combined perturbation can restore system-level function or preferentially impair cancer cell viability. This perspective is especially relevant for herbal medicines because their multicomponent nature may produce therapeutic effects that are distributed across many low-affinity interactions rather than concentrated in one high-affinity binding event. The theory and methodology of network pharmacology have been extensively articulated in the context of traditional medicine, where multiple active substances are often assumed to act through complementary mechanisms [3]. The conceptual alignment between herbal polypharmacology and network modulation has made network pharmacology a powerful lens for investigating traditional formulas, but it also raises difficult questions about how to define causality, how to distinguish direct from indirect effects, and how to validate network predictions in experimental systems.

The construction of network pharmacology models depends on pathway databases and functional annotation resources that organize molecular interactions into interpretable biological contexts. Pathway databases provide curated representations of signaling cascades, metabolic routes, and disease-associated modules [4]. These resources are essential for linking compound targets to cancer-related pathways such as phosphatidylinositol 3-kinase and AKT signaling, mitogen-activated protein kinase cascades, WNT and beta-catenin signaling, and p53-mediated stress responses. However, pathway databases are themselves partial and evolving artifacts. They reflect the accumulated emphasis of experimental research, which may overrepresent well-studied oncogenes and tumor suppressors while underrepresenting less characterized regulatory factors. Consequently, network pharmacology models built on such resources are subject to structural bias that can distort centrality estimates and pathway enrichment results. A systems-level evaluation must therefore treat pathway annotations not as fixed ground truth but as probabilistic and historically contingent knowledge that requires continuous curation and cross-validation.

The conceptual foundations of network pharmacology also intersect with principles of network medicine, which emphasize the role of network proximity between drug targets and disease modules. In cancer, disease-associated proteins are not randomly distributed in the human interactome; they tend to cluster into neighborhoods that correspond to distinct biological processes. Herbal compounds may exert therapeutic effects by targeting proteins located within or adjacent to these disease modules, even when individual binding affinities are modest. This modular perspective reframes the traditional notion of target specificity. Instead of asking whether a compound binds one target with high selectivity, network-based

approaches ask whether a compound modulates a coherent set of targets that collectively influence a disease module. Such a perspective is particularly suitable for herbal medicines, whose constituents may act as weak but network-coherent perturbations. Nevertheless, the modularity assumption must be tested empirically, because network proximity does not guarantee functional efficacy, and the direction of network effects can vary depending on cellular context, post-translational modifications, and feedback compensation.

3. Systems Architecture for Herbal Compound Screening

A robust computational platform for investigating herbal compounds against cancer-related pathways requires a layered architecture that separates data ingestion, network construction, compound-target prediction, structural filtering, and result interpretation. This architecture must balance flexibility with standardization, because the underlying data sources are heterogeneous and evolve rapidly. Protein-protein interaction databases provide the backbone for network construction and enable the projection of compound targets onto larger cellular maps [5]. Herbal compound databases contribute knowledge about the chemical constituents of medicinal plants and their known or predicted targets [6]. Drug repositories provide pharmacological annotations, structural metadata, and approved drug comparisons that support the interpretation of compound candidates [7]. The integration of these resources requires careful identifier mapping, because the same gene, protein, or compound may be represented differently across databases. Without comprehensive identifier harmonization, downstream network analyses can suffer from fragmented nodes and spurious loss of associations. The architectural challenge is therefore not simply to build a pipeline but to design a data governance layer that maintains provenance, tracks updates, and resolves semantic conflicts.

The screening architecture must also accommodate the chemical diversity of herbal compounds. Medicinal plants contain thousands of secondary metabolites, many of which have not been isolated, characterized, or tested in disease models. A screening platform should incorporate modules for compound filtering, drug-likeness assessment, and target prediction before docking or network scoring. These modules introduce trade-offs between recall and precision. Highly permissive filtering retains a larger number of candidate compounds but increases the computational burden and false-positive rate. Highly stringent filtering improves efficiency but may remove unusual natural products whose chemical scaffolds diverge from synthetic drug-like chemical space. The systems architecture should therefore support configurable filtering policies and allow researchers to inspect the consequences of different thresholds. In addition, the platform should preserve negative results, because the systematic documentation of compounds that fail docking or network filters is essential for future model improvement. Many current databases and workflows are biased toward positive findings, which creates a distorted view of the evidence landscape.

Another architectural consideration is the need to separate statistical prediction from mechanistic interpretation. A network model may predict that a compound influences a cancer-related pathway because its predicted targets are enriched in that pathway. However, enrichment does not prove that the compound exerts a functionally relevant effect in a living cancer cell. The architecture should therefore include explicit uncertainty propagation and validation layers that distinguish between target prediction confidence, network association strength, and docking-based structural plausibility. For example, a compound may have high network score but poor docking performance, suggesting that the network prediction may reflect indirect or non-binding interactions. Conversely, a compound may bind strongly in a

docking simulation but have limited network relevance, indicating that its target is not connected to the disease module under study. By retaining these different evidence types in separate layers, the system avoids premature conflation and supports more nuanced candidate prioritization.

4. Network Construction and Signaling Pathway Representation

Network construction is the central computational step that transforms disconnected data records into a relational model of cancer-related signaling. The quality of this network is determined by coverage, curation, and contextual relevance. Protein-protein interaction resources provide large-scale interaction maps, but these maps are often assembled from heterogeneous experimental techniques, including yeast two-hybrid assays, affinity purification, co-immunoprecipitation, and text mining. Each technique has distinct false-positive and false-negative profiles, and the aggregation of these data can introduce systematic distortions. Moreover, cancer signaling is not a static network but a dynamic process regulated by post-translational modifications, subcellular localization, and temporal feedback. Static interaction maps cannot fully capture these dynamic features, yet they remain widely used because they offer computational tractability and broad coverage. The systems-level investigator must therefore interpret network metrics with an awareness of these limitations and, where possible, augment static maps with context-specific transcriptomic, proteomic, or phosphoproteomic data.

Pathway enrichment is frequently used to evaluate whether the predicted targets of herbal compounds concentrate in cancer-related signaling pathways. This approach calculates whether a set of targets overlaps with a predefined pathway more than expected by chance. Pathway resources such as KEGG have been foundational for this purpose, but their boundaries are often based on canonical representations that may not reflect cell-type-specific signaling [4]. In cancer, pathway rewiring is common, and therapeutic relevance may depend on noncanonical branches that are not well represented in reference pathway maps. A network construction strategy that relies solely on canonical pathways may therefore underestimate the impact of compounds that act through alternative or newly discovered signaling routes. To mitigate this risk, platforms increasingly integrate cancer-specific genomic and transcriptomic data to identify disease-associated modules that may extend beyond canonical pathway boundaries. This integration allows the network model to reflect the molecular alterations observed in actual tumors rather than relying exclusively on generic pathway annotations.

Topological analysis provides quantitative descriptors for nodes and edges within the constructed network. Measures such as degree, betweenness, closeness, and eigenvector centrality are used to identify hub proteins and bottleneck interactions that may be particularly influential in signal propagation. In principle, targeting high-centrality nodes may produce stronger therapeutic effects, but it may also increase toxicity because hub proteins often participate in multiple physiological processes. Herbal compounds that modulate multiple lower-centrality nodes may achieve a more selective disease effect by exploiting the combinatorial logic of the cancer network. The representation of signaling pathways as directed graphs further allows the modeling of feedback loops and cross-talk, which are prominent in cancer drug resistance. However, the construction of directed networks requires detailed mechanistic evidence that is often unavailable. As a result, many network pharmacology studies rely on undirected interaction maps, which can obscure the directionality of signal flow and limit the inference of causal relationships. This trade-off

between mechanistic richness and data availability is a recurring theme in the design of network pharmacology systems.

5. Molecular Docking as a Structural Filter in Systems Pipelines

Molecular docking provides a structural complement to network-level target prediction by estimating the preferred orientation and binding affinity of a small molecule within a protein binding site. The availability of experimentally determined three-dimensional protein structures through public repositories has been foundational for docking-based virtual screening [8]. Docking algorithms such as AutoDock Vina have become widely adopted because they balance sampling efficiency with scoring accuracy [9]. In the context of herbal compound investigation, docking is used to test whether a compound predicted to target a cancer-related protein can form plausible interactions with the intended binding pocket. This structural filter is important because network pharmacology predictions are often based on similarity, literature co-occurrence, or machine learning classifiers that may not account for the physical constraints of molecular recognition. Docking therefore reduces the space of candidates that advance to costly experimental validation.

Despite its utility, docking is subject to significant technical limitations that must be addressed at the systems level. Scoring functions approximate the free energy of binding and often fail to rank compounds accurately across diverse chemical scaffolds [10]. The accuracy of docking depends on the quality of the protein structure, the definition of the binding site, the treatment of protein flexibility, and the assignment of protonation states. Herbal natural products frequently contain complex stereochemistry, macrocyclic rings, and multiple hydrogen bond donors and acceptors, which can challenge conventional docking programs. In addition, docking scores are not direct measures of biological activity, because they do not capture membrane permeability, metabolic stability, or intracellular target engagement. A systems pipeline should therefore use docking not as a definitive predictor but as one component of a multi-attribute decision framework. Recent modeling approaches have attempted to capture synergistic relations among herbal components by embedding docking outcomes within network community structures, thereby linking structural binding evidence with network-level cooperation [11]. This integration illustrates the broader trend toward hybrid models that combine structural and topological information.

The structural filtering stage also requires careful validation against known active compounds and decoy sets. Retrospective benchmarking can reveal the sensitivity and specificity of docking protocols, but the choice of benchmark influences the conclusions. A docking protocol that performs well on drug-like synthetic compounds may perform poorly on natural products because the physicochemical properties of herbal constituents differ from those of typical drug libraries. The use of multiple docking programs and consensus scoring can improve robustness by reducing the influence of any single scoring function. Ensemble docking against multiple protein conformations can partially account for receptor flexibility, but it increases computational cost and introduces new challenges in combining results. These structural trade-offs are not merely technical details; they shape the credibility and translational utility of the platform. Consequently, the governance of docking workflows should include explicit documentation of software versions, grid parameters, protonation assignments, and scoring normalization methods.

6. Data Infrastructure, Interoperability, and Reproducibility

The reliability of network pharmacology and docking workflows depends on the quality and interoperability of the underlying data infrastructure. Cheminformatics toolkits provide essential capabilities for structure standardization, descriptor calculation, and file format conversion, yet differences in chemical representation can propagate errors across integrated pipelines [13]. Herbal compound databases often contain duplicate records, ambiguous stereochemistry, and inconsistent identifiers, which complicate the construction of high-quality molecular libraries. Before docking or network projection, compound structures should be normalized through canonicalization and filtered for problematic substructures that interfere with assays [14]. Drug-likeness criteria such as Lipinski guidelines are commonly applied, although many natural products violate these criteria while retaining biological activity [15]. The systems-level challenge is to apply filtering without systematically excluding promising herbal compounds solely because they deviate from synthetic drug-like norms.

Reproducibility has become a central concern in computational pharmacology. The FAIR principles emphasize that data and software should be findable, accessible, interoperable, and reusable [16]. In practice, many network pharmacology studies lack sufficient metadata to reproduce their workflows, including exact database versions, network construction parameters, docking software settings, and scoring thresholds. The absence of this information undermines confidence in published predictions and limits the ability of independent researchers to validate or extend the work. Cancer research has benefited from large-scale public data initiatives that provide well-characterized molecular profiles across tumor types [17]. These resources enable context-specific network analyses but also introduce new responsibilities for data harmonization and ethical use. The integration of clinical and genomic data with herbal compound information must respect patient privacy, consent, and data-sharing agreements, particularly when cross-border collaboration is involved.

Interoperability also requires common ontologies and controlled vocabularies that can represent genes, pathways, chemical entities, and disease phenotypes. Without such standards, data from different sources cannot be reliably merged. Pathway enrichment methods rely on gene set definitions that must be versioned and traceable [18]. Bioactivity databases contribute high-throughput screening results and medicinal chemistry annotations, but their coverage is biased toward compounds that have already been studied [19]. This creates a feedback loop in which well-characterized compounds are repeatedly predicted and tested, while less studied herbal constituents remain invisible to computational models. A more sustainable data infrastructure should allocate resources to fill knowledge gaps rather than merely exploiting existing data abundance. The design of such infrastructure must also consider long-term maintenance, community governance, and the alignment of incentives across academic, industrial, and traditional medicine stakeholders.

7. Robustness, Bias, and Fairness in Predictive Network Models

The predictive models embedded in network pharmacology and docking systems are vulnerable to multiple forms of bias that can reduce their robustness and fairness. Chemical space coverage is uneven, with commercially available compound libraries and synthetic drug collections overrepresented relative to natural products [20]. This imbalance affects machine learning models trained on existing bioactivity data and docking benchmarks that assume drug-like chemical space. Herbal compounds that occupy underrepresented regions of chemical space may receive systematically lower confidence scores, not because they lack biological activity, but because the training data do not support reliable extrapolation.

Fairness in this context requires that computational models be evaluated for performance across different chemical classes, target families, and disease subtypes, rather than relying on aggregate accuracy metrics that may hide systematic failures.

Network bias is another pervasive issue. Protein-protein interaction databases are enriched for highly studied proteins, and cancer research has historically prioritized certain oncogenic pathways. As a result, network-based prioritization may systematically favor compounds targeting well-characterized proteins while overlooking compounds that act through less studied but equally relevant mechanisms. This bias is not necessarily intentional, but it can reproduce existing inequities in research attention and therapeutic development. For example, rare cancers or cancers prevalent in underrepresented populations may be associated with fewer network annotations, leading to weaker model performance and fewer predicted candidates. Addressing this requires deliberate curation efforts, the integration of diverse omics datasets, and the use of computational methods that can account for missing data rather than treating absence of evidence as evidence of absence.

Fairness also extends to validation and deployment. In silico pharmacology models should be validated across multiple independent datasets and experimental contexts before they are used to guide resource-intensive studies [21]. The predictive value of a model may vary across cell lines, patient-derived models, and tumor microenvironments. Herbal medicine research in oncology must likewise consider the sociotechnical context in which traditional knowledge is incorporated [22]. There is a risk that computational frameworks extract value from traditional medical systems without appropriate acknowledgment, benefit sharing, or inclusion of the communities that developed and preserved that knowledge. A fairness-oriented governance approach would ensure that the benefits of network pharmacology research are distributed equitably, that local knowledge holders participate in priority setting, and that data governance respects both scientific openness and cultural sovereignty.

8. Governance, Policy, and Ethical Implications

The governance of network pharmacology and docking research must address scientific validity, data stewardship, intellectual property, and clinical translation. Regulatory agencies traditionally evaluate therapeutic products according to defined chemical entities, manufacturing consistency, and evidence from randomized controlled trials. Herbal medicines challenge this framework because they are often complex mixtures with variable composition and multiple active constituents. Computational predictions alone do not provide sufficient evidence for clinical use, and experimental validation remains essential. However, regulatory pathways for multicomponent interventions are still evolving, and the integration of network pharmacology into regulatory decision-making requires transparent standards for model qualification, data quality, and uncertainty communication. A governance framework that treats computational evidence as supportive but not definitive can help avoid both overinterpretation and premature dismissal.

Data governance is particularly important when network models incorporate patient-derived molecular data, traditional knowledge, and proprietary compound databases. The FAIR principles encourage data sharing, but open access must be balanced against privacy, consent, and the risk of misuse. Federated learning and privacy-preserving computation offer architectural strategies for developing predictive models without centralizing sensitive patient data. Such approaches can support multi-institutional collaboration while reducing the risk of data breaches and unauthorized re-identification. However, federated systems introduce new governance challenges related to model auditing, differential performance across institutions,

and the need for shared evaluation protocols. The implementation of effective governance must therefore be iterative and include participants from computer science, pharmacology, oncology, law, and community advocacy.

Ethical considerations also include the responsible communication of results. Network pharmacology predictions are frequently presented as evidence that a particular herbal compound or formula has anticancer potential. These claims can be amplified by media and commercial interests before they are adequately validated. Researchers have a responsibility to distinguish between computational hypotheses and established therapeutic findings. The overstatement of docking scores or network enrichment can lead to unsafe self-medication, financial exploitation, and erosion of public trust in both computational science and herbal medicine. Institutional review processes, preprint norms, and journal policies should encourage the publication of negative results and null findings, which are essential for an accurate evidence base. Governance mechanisms should also address the environmental sustainability of large-scale computational screening, including the energy and resource costs of high-throughput docking and machine learning pipelines, especially when these are deployed without careful experimental follow-up.

9. Deployment and Translational Sustainability

The deployment of network pharmacology and docking systems in real-world cancer research requires moving beyond proof-of-concept demonstrations toward sustainable, maintainable, and validated platforms. A common failure mode is the development of bespoke pipelines that work on a single dataset or a single herbal formula but cannot be reused for other queries. Sustainable deployment demands modular design, documented application programming interfaces, containerized computational environments, and continuous integration of updated databases. Without these technical practices, the platform decays rapidly as underlying data sources change and software dependencies evolve. Institutional support for long-term maintenance is often difficult to secure because academic funding prioritizes novel discoveries over infrastructure upkeep. Yet without stable infrastructure, the reliability of computational pharmacology at scale cannot be achieved.

Translational sustainability also involves the alignment of computational predictions with experimental and clinical validation pathways. Network pharmacology can generate a ranked list of herbal compounds or combinations, but the value of this list depends on the feasibility and rigor of subsequent assays. High-ranked candidates should be triaged through biochemical binding assays, cell viability experiments, pathway activity measurements, and pharmacokinetic profiling. The selection of validation models should reflect the complexity of the predicted mechanism. If the network model predicts modulation of multiple targets, single-target assays may not capture the intended polypharmacology. Organoid models, patient-derived xenografts, and systems-level omics readouts can provide more informative validation, but they also increase cost and complexity. Sustainability requires a staged approach in which computational prioritization reduces the search space while preserving the ability to detect network-level effects that would be missed by overly narrow validation.

From a global health perspective, translational sustainability includes capacity building in regions where herbal medicine is widely used but computational infrastructure may be limited. Collaborative platforms should support training, local data curation, and the development of regionally relevant cancer models. If computational resources remain concentrated in high-income institutions, the benefits of network pharmacology may not reach the populations that could benefit most from affordable and culturally accepted herbal interventions. Sustainable

deployment therefore requires attention to open source software, cloud-based access, standards for data exchange, and the inclusion of diverse research communities in governance. The long-term goal is not merely to produce individual predictions but to cultivate resilient research ecosystems that can continuously evaluate herbal compounds with methodological rigor and social accountability.

10. Future Directions and Cross-Domain Learning

Future developments in network pharmacology and docking for cancer will likely be shaped by advances in knowledge graphs, deep learning, experimental automation, and privacy-preserving computation. Knowledge graphs can integrate heterogeneous information about compounds, targets, pathways, diseases, and clinical outcomes in a machine-readable form. Unlike traditional relational databases, knowledge graphs support flexible reasoning over multi-hop relationships, which is particularly valuable for interrogating how herbal compounds may influence cancer signaling through indirect mechanisms. Deep learning models can learn representations of molecules, proteins, and networks that capture nonlinear relationships and improve target prediction and docking scoring. However, the deployment of these models must be accompanied by rigorous interpretability and uncertainty estimation, because black-box predictions are difficult to validate in biological systems and may obscure systematic biases.

Cross-domain learning from engineering, ecology, and social science can also inform the design of network pharmacology platforms. Engineering principles of modularity, fault tolerance, and redundancy are directly applicable to computational pipelines that must remain functional despite missing data and evolving databases. Ecological concepts of network robustness and resilience can illuminate how cancer signaling networks respond to perturbations and how multitarget interventions might avoid adaptive resistance. Social science perspectives on knowledge production can help researchers understand how biases enter datasets and how stakeholder participation can improve the relevance and legitimacy of predictive models. The integration of these perspectives may lead to platforms that are not only technically sophisticated but also organizationally adaptive and socially responsible.

The future of this field also depends on the development of multi-scale models that connect molecular interactions to cellular behavior and clinical outcomes. Current network pharmacology models often operate at a single scale, linking compounds to targets and pathways without simulating the spatiotemporal dynamics of tumor growth or immune response. Multi-scale integration will require new ways of coupling molecular docking results with kinetic models, cell state transition models, and tumor microenvironment simulations. This is a formidable challenge because each scale has different data requirements, time constants, and sources of uncertainty. Nevertheless, even partial integration can improve prioritization by filtering out compounds whose molecular effects are unlikely to propagate into meaningful phenotypic changes. Collaborative communities that share benchmark datasets, negative results, and model evaluation protocols will be essential for making progress in this direction.

11. Conclusion

Network pharmacology and molecular docking offer a powerful systems-level lens for investigating herbal compounds against cancer-related signaling pathways. Their integration enables the construction of disease-relevant network models, the projection of multicomponent herbal formulas onto those models, and the structural evaluation of

compound-target interactions. The resulting pipelines can prioritize candidates for experimental validation and generate mechanistic hypotheses that respect the polypharmacological character of herbal medicines. At the same time, these computational architectures are embedded in complex socio-technical systems shaped by data limitations, algorithmic bias, regulatory uncertainty, and uneven access to infrastructure. A mature research agenda must therefore address not only the improvement of network algorithms and docking scoring functions but also the governance, fairness, and sustainability of the platforms through which predictions are produced and used.

The systems-level perspective advanced in this paper emphasizes that computational predictions are provisional knowledge artifacts rather than definitive biological facts. Their value depends on the transparency of data provenance, the reproducibility of analytical workflows, the rigor of external validation, and the ethical alignment of research priorities. Robustness can be improved through consensus scoring, ensemble network construction, context-specific pathway modeling, and the integration of negative results. Fairness can be advanced by evaluating model performance across chemical classes, target families, cancer subtypes, and populations. Governance must balance open science with privacy, traditional knowledge with benefit sharing, and innovation with public safety. Ultimately, the successful application of network pharmacology and molecular docking to herbal cancer research will depend less on any single computational breakthrough than on the collective capacity to build trustworthy, interoperable, and sustainable research infrastructures that connect computational insight with meaningful therapeutic and public health outcomes.

References

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

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

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

4. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Research, 45(D1), D353-D361.

5. Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Research, 49(D1), D605-D612.

6. Ru, J., Li, P., Wang, J., Zhou, W., Li, B., Huang, C., Li, P., Guo, Z., Tao, W., Yang, Y., Xu, X., Li, Y., Wang, Y., & Yang, L. (2014). TCMSP: A database of systems pharmacology for drug discovery from herbal medicines. Journal of Cheminformatics, 6, 13.

7. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., Pon, A., Knox, C., & Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074-D1082.

8. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. Nucleic Acids Research, 28(1), 235-242.

9. Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31(2), 455-461.

10. Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. Nature Reviews Drug Discovery, 3(11), 935-949.

11. 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).

12. Pinzi, L., & Rastelli, G. (2019). Molecular docking: Shifting paradigms in drug discovery. International Journal of Molecular Sciences, 20(18), 4331.

13. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. Journal of Cheminformatics, 3, 33.

14. Baell, J. B., & Holloway, G. A. (2010). New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. Journal of Medicinal Chemistry, 53(7), 2719-2740.

15. Lipinski, C. A. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews, 46(1-3), 3-26.

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., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018.

17. Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., & Stuart, J. M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics, 45(10), 1113-1120.

18. 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.

19. Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., De Veij, M., Felix, E., Magarinos, M. P., Mosquera, J. F., Mutowo, P., Nowotka, M., Gordillo-Maranon, M., Hunter, F., Junco, L., Mugumbate, G., Rodriguez-Lopez, M., Atkinson, F., Bosc, N., Radoux, C. J.,

Segura-Cabrera, A., Hersey, A., & Leach, A. R. (2019). ChEMBL: Towards direct deposition of bioassay data. Nucleic Acids Research, 47(D1), D930-D940.

20. Irwin, J. J., & Shoichet, B. K. (2005). ZINC: A free database of commercially available compounds for virtual screening. Journal of Chemical Information and Modeling, 45(1), 177-182.

21. Ekins, S., Mestres, J., & Testa, B. (2007). In silico pharmacology for drug discovery: Methods for virtual ligand screening and profiling. British Journal of Pharmacology, 152(1), 9-20.

22. Hao, D. C., & Xiao, P. G. (2014). Network pharmacology: A Rosetta Stone for traditional Chinese medicine. Drug Development Research, 75(5), 299-312.