Medical Research
Volume 7, Issue 4, 2025, Page 41-57
https://doi.org/10.6913/mrhk.070405

Medical Research

Integrating Artificial Intelligence with Traditional Wisdom: The Rise of Computational Ethnopharmacology for TCM Discovery and Interaction Prediction

Xin HUANG
Chinese Society for the Preservation and Innovative Development of Traditional Chinese Medicine

Abstract

Traditional Chinese Medicine (TCM) faces a persistent "black box" problem arising from multi-component, multi-target, and nonlinear synergistic mechanisms that challenge reductionist biomedical paradigms. This review synthesizes major breakthroughs (2024 2025) and argues that computational ethnopharmacology represents a paradigm shift beyond classical network pharmacology. We propose a four-layer architecture—data, semantic, topological, and structural layers—integrating evidence-oriented databases (e.g., HERB 2.0), GraphRAG-enabled LLM mining of classical texts, hypergraph/heterogeneous GNN modeling of formula compatibility, and AlphaFold 3 driven structural inference coupled with inverse docking and molecular dynamics. Focusing on clinically actionable herb drug interaction prediction, we further outline validation and reporting checklists to improve reproducibility and safety translation, and highlight dry wet closed-loop directions for future research.

Keywords Computational ethnopharmacology; Traditional Chinese Medicine; Artificial intelligence; Graph neural networks; AlphaFold

1 Introduction: The Rise of Computational Ethnopharmacology and Paradigm Shift

1.1 Epistemic Bottlenecks in the Modernization of Traditional Chinese Medicine and the "Black Box" Problem

Traditional Chinese Medicine (TCM), as a medical system refined through thousands of years of clinical practice within Chinese civilization, has developed a distinctive theoretical framework centered on syndrome differentiation and formula-based therapy. This system demonstrates unique advantages in addressing complex disorders, chronic metabolic diseases, and emerging infectious diseases. Nevertheless, integrating TCM into contemporary biomedical research frameworks poses profound epistemic challenges.

Modern drug discovery is largely governed by a reductionist paradigm characterized by the principle of "one disease–one target–one drug." In contrast, TCM formulae (*fufang*) exhibit extreme systemic complexity. A typical herbal formula may contain more than ten medicinal herbs and hundreds to thousands of phytochemicals. Once administered, these compounds may interact with thousands of biological

 Article History Received: August 5, 2025; Accpted: October 15, 2025; Published: December 15, 2025. ***Medical Research***, ISSN 2664-0333 (print), ISSN 2664-0341 (online), a Quarterly, founded on 2019, published by Creative Publishing Co., Limited. Email: cpcl@cpcl.hk, https://mrhk.cc, https://cpcl.hk.

targets, forming a nonlinear and dynamic biological network. Such complexity fundamentally challenges conventional experimental and analytical approaches.

For decades, TCM research has been constrained by three persistent bottlenecks: unclear material bases, insufficiently elucidated mechanisms of action, and difficulties in quality control. Traditional wet-lab screening strategies are not only costly and time-consuming, but also inherently limited in capturing multi-component synergistic effects. In many cases, certain constituents within a formula do not directly bind to disease-related targets; instead, they modulate gut microbiota, alter pharmacokinetic pathways, regulate metabolic enzymes, or reshape the cellular microenvironment to facilitate therapeutic efficacy. The epistemic gap between TCM holism and biomedical reductionism has therefore given rise to the so-called "black box" problem, wherein the systemic mechanisms of herbal medicine remain opaque and difficult to formalize.

1.2 From Network Pharmacology to Computational Ethnopharmacology: A Paradigm Transition

To address these challenges, network pharmacology emerged over the past two decades as an intermediary methodological framework. By constructing "drug–target–disease" networks and analyzing topological properties such as degree centrality and betweenness centrality, network pharmacology offered a systems-level perspective on herbal medicine. However, with increasing methodological scrutiny, several structural limitations of early network pharmacology have become evident.

First, most network pharmacology studies rely heavily on static databases and fail to account for temporal dynamics and dose-dependent effects inherent in biological systems. Second, traditional graph models are restricted to pairwise relationships, rendering them insufficient for modeling the group-level synergistic effects characteristic of multi-herb formulae. Third, the lack of quantitative pharmacokinetic integration contributes to elevated false-positive rates in target prediction and pathway inference.

Since 2024, driven by advances in high-throughput omics, large-scale biomedical data, and artificial intelligence, an emerging interdisciplinary field—**computational ethnopharmacology**—has gained formal academic recognition. Rather than representing a simple extension of network pharmacology, computational ethnopharmacology constitutes a substantive paradigm shift. It integrates multiple frontier technologies to enable predictive, mechanistic, and multiscale modeling of traditional medical knowledge.

Key technological pillars of this paradigm include: (1) *Generative artificial intelligence and large language models (LLMs)*, which facilitate structured knowledge extraction from classical medical texts and clinical records, enabling the construction of high-fidelity TCM knowledge graphs and semantic alignment across medical ontologies; (2) *Geometric deep learning and hypergraph modeling*, which transcend binary relationships and mathematically represent the nonlinear combinatorial logic of formula composition, such as the "Jun Chen Zuo Shi" principle; (3) *Structural biology oriented AI*, exemplified by AlphaFold 3, which enables the transition from sequence-based inference to structure-level prediction, thereby supporting inverse docking and proteome-wide target fishing for natural products.

1.3 Scope of This Article and Response to Reviewer Concerns

This article provides a comprehensive review of recent advances in computational ethnopharmacology, with particular emphasis on breakthroughs reported between 2024 and 2025. In response to reviewer feedback, several substantive revisions and enhancements have been implemented.

First, methodological analysis has been significantly deepened. Rather than enumerating applications at a descriptive level, this article interrogates the internal algorithmic mechanisms of models such as hypergraph neural networks and GraphRAG architectures, with explicit discussion of their mathematical suitability for TCM-specific data structures.

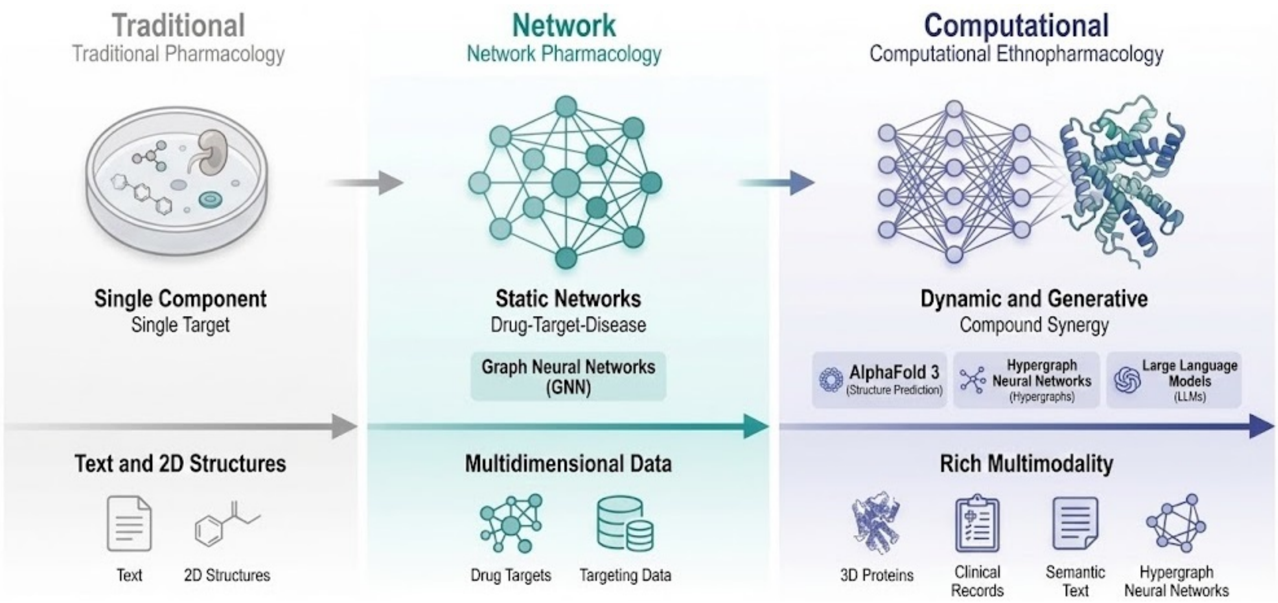

Figure 1: The Paradigm Evolution from Network Pharmacology to Computational Ethnopharmacology *This diagram illustrates the three main stages of the TCM research paradigm. The first stage is traditional reductionist pharmacology, focusing on single components and single targets. The second stage is network pharmacology, introducing static "drug-target-disease" networks. The third stage is the current computational ethnopharmacology, characterized by the integration of AlphaFold 3 structure prediction, Hypergraph Neural Networks, and Large Language Model (LLM)-driven semantic mining, realizing deep simulation of formula synergy and dynamic biological processes.*

Second, issues of data validation and standardization have been systematically addressed. A dedicated section has been added to examine database selection criteria, model evaluation metrics, and standardized reporting checklists for herb drug interaction (HDI) prediction, thereby enhancing methodological transparency and reproducibility.

Third, the English abstract has been fully revised to align with international journal standards, explicitly articulating the conceptual contributions and technical innovations of this review.

The remainder of this article is organized around four core technological layers of computational ethnopharmacology—namely the data layer, semantic layer, topological layer, and structural layer—followed by an analysis of clinical application scenarios such as HDI risk prediction. Through this multilevel framework, the article aims to provide both a conceptual foundation and a methodological roadmap for the AI-driven modernization and global translation of traditional medical knowledge.

2 Data Layer: Establishing an Evidence-Based Foundation for the Digitalization of Traditional Chinese Medicine

The performance upper bound of artificial intelligence models is fundamentally constrained by the quality, scale, and dimensional richness of the underlying data. In the field of Traditional Chinese Medicine (TCM), data sources are abundant yet highly heterogeneous, encompassing classical medical literature, modern pharmacopoeias, chemical constituent databases, high-throughput omics datasets, and real-world clinical records. In recent years, the construction of TCM databases has undergone a critical transition—from the simple aggregation of chemical components toward intelligent, evidence-oriented platforms integrating multi-omics data and clinical validation.

Table 1: A four-layer landscape of computational ethnopharmacology for TCM: tasks, representative methods, typical inputs/outputs, and validation practices.

Layer	Core Tasks	Representative Methods	Typical Inputs/Outputs	Validation
Data Layer	Data integration; compound–target linking; evidence curation	Evidence-aware databases; multi-omics linkage; ADMET/toxicity screening models	Inputs: herbs/formulas, compounds, targets, omics, clinical records; Outputs: curated associations, ADME/Tox profiles, candidate lists	Curation rules; evidence grading; external database cross-check
Semantic Layer	Classical text mining; case record structuring; ontology alignment; KG construction	Domain LLMs; RAG/GraphRAG; NER/RE pipelines; ontology engineering	Inputs: classical corpora, EMRs, guidelines; Outputs: entities/relations, triples, ontologies, explainable retrieval contexts	Expert-in-the-loop auditing; provenance tracing; hallucination control
Topological Layer	Formula compatibility modeling; DTI/DDI/HDI prediction; mechanism subgraph discovery	(Heterogeneous) GNNs; hypergraph learning for formulas; contrastive/self-supervised learning	Inputs: herb–compound–target–disease graphs; hyperedges as formulas; Outputs: interaction scores, key subgraphs, synergy/antagonism hypotheses	Benchmarks on held-out links; ablation studies; external validation when possible
Structural Layer	Target fishing; binding mode inference; structure-based screening	Protein structure prediction; docking/inverse docking; MD refinement	Inputs: protein sequences/structures, ligand structures; Outputs: complex structures, binding poses, stability metrics	Docking scores + MD stability; experimental assays (SPR/ITC/cell) when available

2.1 Generational Evolution of TCM Databases: From Physicochemical Properties to Clinical Evidence

Early-generation TCM databases primarily focused on the digitalization of herbal chemical constituents and target associations derived from text mining. While these resources played a foundational role in systems-level TCM research, they were limited by the absence of robust experimental and clinical evidence.

For instance, widely adopted screening criteria such as oral bioavailability and drug-likeness thresholds have been extensively used as practical filters for candidate compounds. However, these physicochemical rules have also been criticized for overlooking key biological factors, including gut microbiota mediated metabolism and context-dependent pharmacokinetics. Entering the period of 2024 2025, next-generation TCM databases achieved qualitative advances in data coverage, evidentiary confidence, and algorithmic integration.

2.1.1 HERB 2.0: A High-Throughput, Evidence-Driven Platform

The release of HERB 2.0 marks a decisive shift toward evidence-based TCM informatics. By integrating large-scale high-throughput experimental data with structured clinical evidence, the platform substantially enhances data reliability and translational relevance.

Deep integration of clinical evidence. HERB 2.0 incorporates thousands of clinical trial records and meta-analyses. Through the combination of natural language processing techniques and expert manual curation, explicit conclusions are extracted from clinical studies, directly linking TCM interventions to clinical outcomes. This systematic causal mapping addresses long-standing concerns regarding the lack of high-quality clinical evidence in TCM research.

Multi-omics integration and "formula-to-syndrome" inference. The platform further integrates high-throughput experimental datasets across a wide range of diseases and provides automated gene expression analysis interfaces. Notably, it supports reverse inference workflows that allow users to upload transcriptomic data, which are then mapped onto herbal target networks. This capability enables the prediction of potentially effective herbs or formulae based on molecular phenotypes, laying the groundwork for genotype-informed precision TCM interventions.

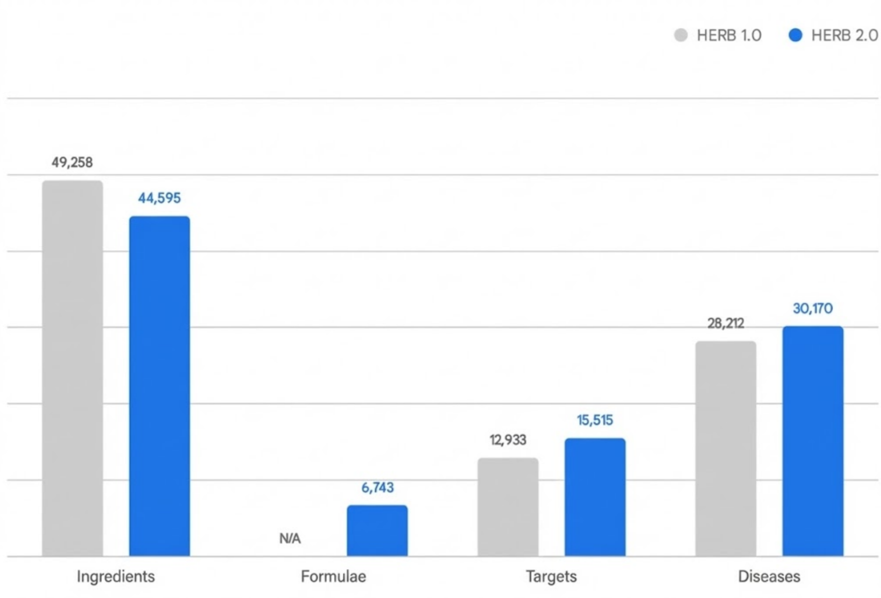

Figure 2: Comparison of Data Scale between HERB 1.0 and HERB 2.0 *Note: HERB 2.0 significantly surpasses its predecessor in both data breadth and depth. In particular, it has added 6,743 new formulae data and increased the target coverage to 15,515. Furthermore, HERB 2.0 has introduced 8,558 clinical trial data points not included in the previous generation (the comparison of clinical trials is not shown in the figure; only the growth in entities is displayed), greatly enhancing the clinical credibility of the data. Data sources: Nucleic Acids Research, Database Commons*

2.1.2 SymMap v2: Bridging Phenotypes and Genotypes

A persistent challenge in integrative medicine lies in the semantic mismatch between TCM and biomedicine. While TCM conceptualizes disease through syndromes, modern medicine operates in terms of genes, pathways, and molecular mechanisms. SymMap v2 addresses this gap by constructing comprehensive mappings among herbs, chemical components, molecular targets, clinical symptoms, and TCM syndromes.

Through expert-driven manual review and semantic alignment, SymMap v2 establishes high-confidence herb symptom association matrices. Phenotypic symptoms such as headache can be linked to specific herbal components and their corresponding molecular networks. These mappings provide a molecular basis for interpreting classical TCM principles such as treating the same disease with different methods or treating different diseases with the same method, enabling biological interpretation of syndrome differentiation.

2.1.3 ADMETlab 3.0: AI-Driven Pharmacokinetic and Toxicological Assessment

Although not exclusively dedicated to TCM, ADMETlab 3.0 plays a critical role in evaluating the druggability and safety of natural products. The latest version introduces a multi-task directed message-passing neural network architecture, significantly improving predictive accuracy for toxicity and metabolic stability.

A major advance lies in the precise prediction of cardiotoxicity and drug-induced liver injury, both of which are common causes of failure in natural product drug development. Given the compositional complexity of herbal medicines, early identification of high-risk compounds is particularly important. ADMETlab 3.0 supports batch screening through programmatic interfaces, enabling rapid in silico evaluation of thousands of herbal constituents and substantially reducing trial-and-error costs during early-stage drug discovery.

2.2 Innovative Strategies in Knowledge Graph Construction: Addressing Data Sparsity

Conventional relational databases struggle to represent the rich semantic relationships inherent in TCM knowledge, such as mutual promotion and inhibition or hierarchical formula composition. Knowledge graphs, which store information in graph structures composed of nodes and edges, offer a natural solution for capturing these high-dimensional relationships. However, a central challenge in TCM knowledge graph construction is data sparsity, as experimentally validated compound target interactions remain limited.

To address this issue, a neighbor-diffusion strategy has been proposed in the construction of multi-dimensional TCM knowledge graphs.

Algorithmic rationale. The strategy is grounded in the biological principle that similar compounds tend to interact with similar targets. Rather than relying solely on directly validated interactions, the algorithm allows information to propagate through neighboring nodes within the graph, enabling indirect inference of potential associations.

Practical outcomes. By applying neighbor diffusion, target coverage can be dramatically expanded, effectively densifying the knowledge graph. This improvement not only fills critical knowledge gaps but also provides richer input features for downstream graph neural network models, enabling activity prediction for underexplored or less-studied herbal medicines.

3 Semantic Layer: Large Language Models for Decoding Classical TCM Texts and Intelligent Diagnosis

The core knowledge of Traditional Chinese Medicine (TCM) is primarily preserved in natural language form within thousands of years of classical medical literature. These texts are typically written in classical Chinese, characterized by extreme conciseness, heavy metaphorical expression, and the absence of punctuation. Such features pose substantial challenges for modern computational semantic understanding. Although general-purpose large language models demonstrate strong generalization capabilities, they frequently exhibit semantic drift or hallucination when processing domain-specific TCM terminology. As a result, domain-specialized LLMs combined with retrieval-augmented generation (RAG) architectures have emerged as a central research focus in 2025.

3.1 OpenTCM: Deep Knowledge Retrieval Based on GraphRAG

The OpenTCM system represents a state-of-the-art implementation of LLM-based semantic decoding for TCM. The platform is designed to extract structured knowledge from classical medical texts and to support downstream intelligent diagnostic applications.

3.1.1 Corpus Scale and Text Processing

OpenTCM integrates approximately 3.73 million classical Chinese characters drawn from sixty-eight gynecology monographs included in the *Encyclopedia of Traditional Chinese Medicine*. The corpus covers a large number of herbal prescriptions, medicinal substances, and clinical case records. To enable effective model comprehension of these highly condensed and context-dependent texts, customized prompt engineering strategies are employed. The LLM is explicitly instructed to function as a *TCM knowledge extraction assistant*, with outputs constrained to structured JSON formats. This design enables the automated transformation of unstructured classical narratives into structured databases, significantly reducing manual annotation costs.

3.1.2 Innovations in GraphRAG Architecture

The core methodological innovation of OpenTCM lies in its adoption of GraphRAG, a graph-based retrieval-augmented generation framework that differs fundamentally from conventional RAG approaches.

Conventional RAG. Traditional RAG systems retrieve text chunks based on vector similarity. In the TCM context, this approach often leads to semantic fragmentation, such as retrieving an herb name while losing its prescription context or combinatorial constraints.

GraphRAG. In contrast, GraphRAG performs retrieval at the level of structured knowledge graphs. When a query is issued, the system first identifies a relevant entity subgraph from the TCM knowledge graph, which contains tens of thousands of entities and relationships. This structured subgraph is then provided as contextual input to the LLM.

Performance implications. By anchoring generation to explicit graph-based knowledge structures, GraphRAG significantly reduces hallucination risk. The hybrid integration of explicit symbolic knowledge and implicit neural reasoning enables the generation of diagnostic outputs that are both interpretable and traceable, establishing a reliable semantic foundation for intelligent TCM applications.

3.2 Automated Medical Case Mining and Ontology Construction

Clinical case records represent the primary interface between TCM theory and real-world practice. Traditionally, the extraction and organization of medical case knowledge has been highly labor-intensive and dependent on expert interpretation. Large language models are fundamentally transforming this process.

In recent studies focusing on classical physicians' case records, LLMs have been employed to automatically extract key entities, including symptoms, pathogenesis patterns, therapeutic principles, and prescriptions, followed by semantic normalization and alignment. The extracted information is organized into structured quadruples, forming the basis for computable clinical knowledge representations.

Empirical evidence suggests that a human-in-the-loop paradigm currently offers the most robust solution. In this workflow, LLMs perform preliminary named entity recognition and relation extraction, after which domain experts conduct validation and correction. This collaborative approach substantially improves efficiency while preserving semantic accuracy, thereby enabling the construction of standardized TCM ontologies suitable for downstream reasoning, data integration, and intelligent diagnostic systems.

4 Topological Layer: Graph Neural Networks and Mathematical Modeling of Herbal Formula Compatibility

If knowledge graphs provide a structured representation of Traditional Chinese Medicine (TCM) knowledge, graph neural networks (GNNs) function as the computational engine that extracts latent patterns embedded within these structures. Herbal formulas are intrinsically combinatorial systems, characterized by nonlinear interactions among multiple herbs, compounds, targets, and biological pathways. GNN-based

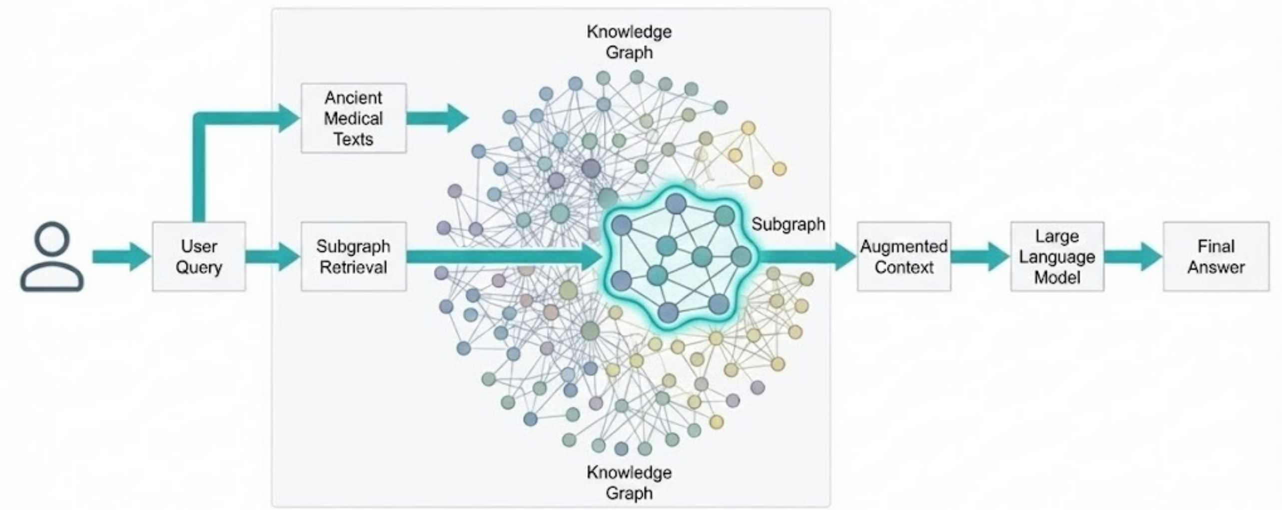

Figure 3: GraphRAG Workflow in the OpenTCM System *Schematic diagram of the GraphRAG architecture in OpenTCM. The system first structures ancient medical texts into a Knowledge Graph. When a user initiates a query, the retrieval module not only matches keywords but also extracts relevant entity sub-graphs, preserving the structural relationships between entities. These sub-graphs are input into the Large Language Model as augmented context, thereby generating highly accurate and interpretable diagnostic suggestions.*

models are particularly well suited to this domain, as they enable representation learning over relational structures while preserving topological dependencies.

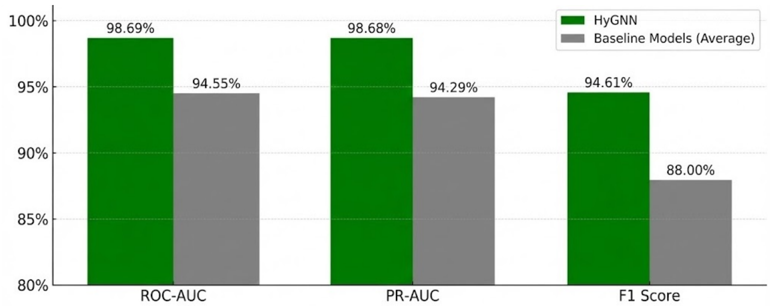

Figure 4: Performance Comparison of HyGNN and Baseline Models in DDI Prediction *Note: Performance of the HyGNN model in the Drug-Drug Interaction (DDI) prediction task. The data shows that HyGNN outperforms traditional baseline models on key metrics such as ROC-AUC (98.69%) and F1 Score (94.61%), demonstrating the advantages of hypergraph structures in capturing complex drug relationships. Source: Data source: NSF Public Access Repository*

4.1 From Homogeneous to Heterogeneous Graph Neural Networks

Early computational studies in network pharmacology primarily relied on homogeneous graph representations, treating all nodes as structurally equivalent. In such models, herbs, compounds, targets, and diseases are embedded within a single node space, and information propagation occurs uniformly across edges. While computationally convenient, this abstraction overlooks the intrinsic semantic heterogeneity of biomedical systems.

4.1.1 Heterogeneous Graph Construction

Modern TCM-oriented models increasingly adopt heterogeneous graph neural networks (HGNNs), explicitly distinguishing node and edge types. A typical heterogeneous TCM graph may include multiple node categories (herbs, phytochemicals, protein targets, diseases) and relation types (contains, interacts with, treats, associates with). Message passing is then conditioned not only on node features but also on relation semantics.

By incorporating meta-path based aggregation mechanisms, HGNNs are able to differentiate distinct semantic trajectories. For example, a path of the form *herb compound target* encodes molecular mechanism information, whereas a path such as *herb herb* captures compatibility or co-occurrence relationships within classical formulas. This distinction significantly improves representation fidelity and interpretability.

4.1.2 Advantages over Homogeneous Models

Compared with homogeneous GNNs, HGNNs demonstrate superior performance in predicting compound target interactions and herb disease associations. More importantly, they allow researchers to decompose model predictions according to semantic pathways, providing mechanistic insights that align with the explanatory demands of medical research.

4.2 Hypergraph Learning: Modeling the Formula as a Higher-Order System

While heterogeneous graphs represent pairwise relationships effectively, they remain fundamentally limited in describing higher-order interactions. Traditional Chinese herbal formulas are not reducible to binary combinations; instead, they operate as integrated sets in which therapeutic effects emerge from collective behavior. This limitation motivates the introduction of hypergraph learning.

4.2.1 Hypergraph Representation of Herbal Formulas

In a hypergraph, a single hyperedge can connect multiple nodes simultaneously. This structure corresponds naturally to the conceptual organization of herbal formulas, where one prescription links several herbs as a functional unit. Unlike pairwise graphs, hypergraphs explicitly encode group-level interactions without decomposing them into multiple binary edges.

From a mathematical perspective, hypergraph convolution generalizes standard message passing by aggregating information across node sets rather than node pairs. This enables the model to learn representations that capture synergy, redundancy, and antagonism among herbal components.

4.2.2 Capturing Compatibility Principles

Hypergraph neural networks provide a formal computational framework for modeling classical compatibility principles such as "Jun Chen Zuo Shi" (sovereign minister assistant courier). Rather than encoding these roles as static labels, hypergraph learning allows role differentiation to emerge dynamically from interaction patterns and contribution weights within hyperedges. This data-driven formulation avoids rigid rule encoding while remaining faithful to traditional theoretical structures.

4.3 Contrastive Learning under Sparse and Biased Supervision

A persistent challenge in TCM data modeling is the scarcity of explicit negative samples. Historical medical records overwhelmingly document successful treatments, while ineffective or adverse outcomes are rarely recorded. This leads to strong survivorship bias and undermines supervised learning.

4.3.1 Self-Supervised Contrastive Objectives

Contrastive learning addresses this limitation by enabling self-supervised representation learning. By constructing multiple augmented views of the same graph—through node masking, edge perturbation, or subgraph sampling—models are trained to maximize agreement between representations of the same entity across views while minimizing similarity with others.

In the TCM context, contrastive learning allows herbal and compound representations to be learned independently of explicit failure cases. The resulting embeddings exhibit improved robustness, generalization, and resistance to noise inherent in historical data sources.

4.3.2 Implications for Model Reliability

Empirical studies indicate that contrastive objectives substantially enhance performance under low-label and high-noise conditions. From a clinical perspective, this translates into more stable predictions when extrapolating to under-studied herbs or novel formula combinations, thereby increasing the translational reliability of computational ethnopharmacology models.

4.4 Topological Interpretability and Medical Meaning

Beyond predictive accuracy, topological models offer a unique advantage in interpretability. Attention mechanisms and contribution scores within GNN and hypergraph frameworks allow the identification of influential nodes, edges, or hyperedges. These signals can be mapped back to herbs, compounds, or formula components, providing mechanistic hypotheses rather than opaque predictions.

This interpretability is particularly critical in medical contexts, where computational outputs must support expert reasoning rather than replace it. By aligning learned topological features with established pharmacological and theoretical constructs, graph-based models facilitate meaningful dialogue between artificial intelligence and traditional medical knowledge systems.

5 Structural Layer: AlphaFold 3 and the Transformation of Structure-Based Ethnopharmacology

Graph-based learning models excel at predicting statistical associations within high-dimensional biomedical networks; however, association alone is insufficient to establish molecular mechanism. Structure-based biology provides the complementary capacity to evaluate physical binding, conformational compatibility, and thermodynamic stability. The emergence of AlphaFold—particularly AlphaFold 3 (AF3) and AlphaFold-Multimer—has fundamentally reshaped structure-based drug discovery (SBDD), enabling unprecedented integration between artificial intelligence and molecular biophysics. For Traditional Chinese Medicine (TCM), this development offers a decisive methodological breakthrough by transforming previously opaque multi-component interventions into structurally interpretable mechanisms.

5.1 AlphaFold 3 in Target Discovery for Natural Products

A recurring challenge in ethnopharmacology is the presence of so-called "orphan compounds": bioactive phytochemicals with demonstrated therapeutic effects but unknown protein targets. Conventional molecular docking pipelines depend heavily on experimentally resolved protein structures, typically obtained via X-ray crystallography or cryo-electron microscopy. Despite decades of effort, a substantial proportion of the human proteome remains structurally unresolved, severely constraining target identification for natural products.

AlphaFold 3 addresses this limitation by providing near-experimental accuracy in predicting protein ligand and protein protein complex structures. Unlike earlier iterations focused primarily on monomeric folding, AF3 incorporates ligand-aware modeling and multi-chain interaction prediction, thereby extending structural inference from isolated proteins to functional complexes. The availability of high-confidence three-dimensional target structures enables systematic, proteome-scale virtual screening strategies that were previously infeasible in TCM research.

From a methodological perspective, AF3 shifts target discovery from a data scarcity regime to a hypothesis generation regime. Instead of asking whether a structure exists, researchers can now prioritize candidate targets based on predicted binding feasibility, structural complementarity, and interaction topology. This paradigm significantly lowers the entry barrier for mechanistic exploration of complex herbal constituents.

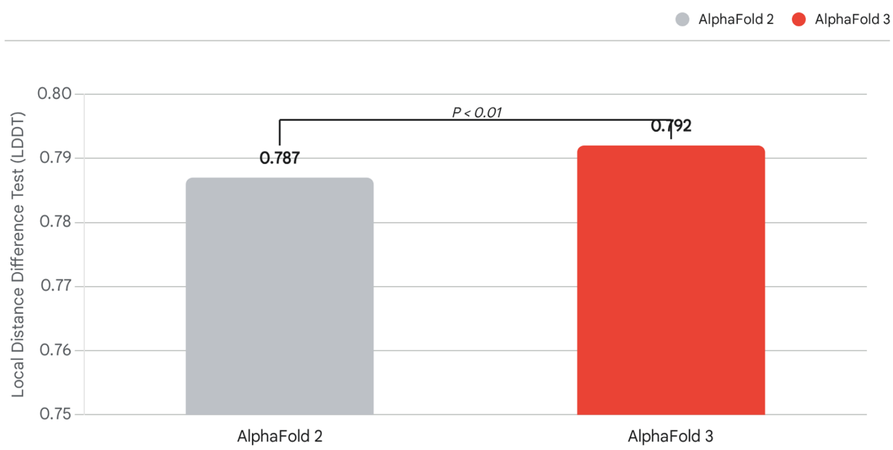

Figure 5: Comparison of Structural Prediction Accuracy between AlphaFold 3 and AlphaFold 2 (Subtitle: Average LDDT Score Comparison) *Note: AlphaFold 3 achieved a score of 0.792 on the Local Distance Difference Test (LDDT), realizing a statistically significant improvement (P < 0.01) compared to AlphaFold 2 (0.787). This increase in precision is crucial for accurately predicting the binding conformations of small molecules in Traditional Chinese Medicine (TCM) with protein complexes. Source: Data sources: NCBI PMC*

5.2 Inverse Docking as a Proteome-Wide Fishing Strategy

Traditional docking workflows typically involve screening large chemical libraries against a single protein target. While effective for single-target drug development, this approach is poorly suited to multi-target pharmacology, which is central to TCM. Inverse docking reverses this logic by systematically docking a single compound against thousands of potential protein targets to identify its interaction spectrum.

When combined with AlphaFold-generated structures, inverse docking becomes a powerful hypothesis engine. Predicted binding affinities and interaction geometries allow researchers to construct ranked target profiles for individual phytochemicals. These profiles can then be integrated with pathway analysis to identify clusters of functionally related targets, revealing coordinated modulation of biological processes rather than isolated effects.

In applied settings, inverse docking has proven particularly informative for metabolic and inflammatory disorders, where disease phenotypes arise from distributed network dysregulation. For example, in studies of metabolic-associated fatty liver disease (MAFLD), AlphaFold-based docking has been used to uncover multi-target engagement patterns along the AMPK/SIRT1 TFEB signaling axis, supporting the hypothesis that herbal compounds exert therapeutic effects through convergent regulation of lipid metabolism, autophagy, and mitochondrial homeostasis.

5.3 Molecular Dynamics Simulation as a Stability Filter

While docking predicts plausible binding poses, it represents a static approximation of molecular interaction. Biological systems, by contrast, operate under dynamic conditions characterized by thermal fluctuation, solvent effects, and conformational flexibility. Molecular dynamics (MD) simulation provides a physics-based framework for validating and refining AI-predicted complexes by modeling their temporal evolution.

In the context of ethnopharmacology, MD simulation serves three critical functions. First, it assesses the structural stability of predicted ligand protein complexes over time, distinguishing transient contacts from sustained interactions. Second, it enables estimation of binding free energy, offering quantitative comparison between candidate compounds. Third, it reveals conformational adaptation processes, such as induced fit or allosteric modulation, which are especially relevant for multi-target natural products.

Empirical workflows increasingly adopt a sequential strategy: network pharmacology for candidate prioritization, AlphaFold-based docking for structural hypothesis generation, and MD simulation for energetic and dynamic validation. This layered approach reduces false positives and aligns computational prediction more closely with experimental observability.

5.4 Toward a Dry Wet Closed-Loop Paradigm

The integration of AlphaFold, inverse docking, and MD simulation supports the emergence of a dry wet closed-loop research paradigm. In this framework, computational predictions generate testable hypotheses that guide experimental validation, while experimental outcomes—both positive and negative—are fed back into model refinement.

For TCM research, this paradigm is particularly transformative. It enables rational selection of active components from complex formulas, prioritization of mechanistically coherent target sets, and design of focused wet-lab experiments such as surface plasmon resonance assays, isothermal titration calorimetry, or cellular functional assays. Over time, iterative feedback improves both model accuracy and biological understanding.

More broadly, the structural layer elevates computational ethnopharmacology from pattern recognition to mechanistic science. By anchoring network-level associations in physically plausible molecular interactions, AlphaFold-centered workflows bridge the epistemic gap between holistic medical traditions and reductionist molecular biology, offering a unified framework for credible, explainable, and translatable discovery.

6 Application Layer: Multimodal Learning for Herb–Drug Interaction Prediction

With the rapid aging of global populations, polypharmacy has become increasingly prevalent, particularly among patients with chronic and metabolic diseases. In this context, the concurrent use of traditional herbal medicines and Western pharmaceuticals is no longer exceptional but routine. While such combinations may offer therapeutic complementarity, they also introduce significant safety concerns. Herbal components can interfere with drug-metabolizing enzymes, transporters, or downstream signaling pathways, potentially leading to reduced efficacy or increased toxicity. Consequently, the prediction and mechanistic interpretation of herb–drug interactions (HDIs) represent one of the most clinically actionable applications of computational ethnopharmacology.

6.1 Multimodal Deep Learning Frameworks

Single-modality approaches—relying solely on chemical structure or textual descriptions—are inherently insufficient to capture the multifaceted nature of drug interactions. Recent advances therefore emphasize multimodal learning frameworks that integrate heterogeneous data sources, most notably molecular structure information and biomedical text.

A representative strategy combines graph neural networks (GNNs) with language models specialized for biomedical corpora. In this framework, GNNs encode the topological and physicochemical features of drug molecules, while transformer-based language models process unstructured text such as drug labels, pharmacological literature, and clinical guidelines. The semantic representations extracted from text provide contextual cues regarding metabolic pathways, regulatory effects, and documented adverse interactions that are not directly inferable from molecular structure alone.

The fusion of structural and semantic modalities enables the model to reason across complementary knowledge domains. For example, textual signals indicating enzyme induction or inhibition can be aligned with structural motifs known to interact with specific cytochrome P450 isoforms. This multimodal integration substantially enhances predictive robustness, particularly for complex interaction scenarios where structural similarity alone is insufficient. Importantly, such frameworks also improve interpretability by allowing predictions to be traced back to both molecular features and linguistic evidence.

6.2 Pharmacokinetic and Pharmacodynamic Mechanism Stratification

Beyond interaction detection, a critical goal of HDI modeling is mechanistic stratification. Contemporary AI models increasingly attempt to distinguish whether an interaction is driven primarily by pharmacokinetic (PK) or pharmacodynamic (PD) mechanisms, as these categories imply different clinical management strategies.

6.2.1 Pharmacokinetic Interactions

PK-driven HDIs arise when herbal components alter the absorption, distribution, metabolism, or excretion of co-administered drugs. The most common mechanisms involve modulation of metabolic enzymes, particularly cytochrome P450 isoforms, as well as drug transporters. Computational ADMET platforms equipped with deep learning architectures can predict whether specific phytochemicals are likely to inhibit or induce key enzymes such as CYP3A4 or CYP2D6.

From a clinical perspective, PK interaction prediction enables early risk screening. For instance, flavonoid compounds frequently present in herbal formulations may suppress CYP3A4 activity, leading to elevated plasma concentrations of calcium channel blockers or statins. By identifying such risks in silico, clinicians and researchers can prioritize monitoring strategies, dose adjustment, or alternative therapies before adverse events occur.

6.2.2 Pharmacodynamic Interactions

PD-driven HDIs occur at the level of drug targets and signaling pathways, involving synergistic, additive, or antagonistic effects. Unlike PK interactions, which often influence systemic exposure, PD interactions modulate therapeutic outcomes through network-level biological processes.

Network-based modeling plays a central role in PD interaction analysis. By constructing integrated target networks that include both herbal and pharmaceutical agents, AI models can identify pathway crosstalk and shared regulatory nodes. Such analyses are particularly relevant for complex diseases characterized by immune dysregulation or metabolic imbalance. In disease contexts such as diabetic nephropathy, computational models have revealed that certain herbal components may converge with pharmaceutical interventions on innate immune signaling pathways, suggesting potential synergy rather than conflict.

6.3 Toward Clinical Decision Support

The ultimate value of HDI prediction lies in its translation into clinical decision support systems. Multimodal AI frameworks provide a foundation for stratifying interaction risk, distinguishing between PK and PD mechanisms, and generating mechanistically informed recommendations. Rather than issuing binary warnings, future systems can offer nuanced guidance—such as heightened monitoring for PK interactions or therapeutic optimization for PD synergy.

In this sense, HDI prediction exemplifies the broader promise of computational ethnopharmacology: transforming traditional empirical knowledge into a data-driven, explainable, and clinically actionable intelligence layer that enhances patient safety while preserving therapeutic diversity.

7 Protocol Design: Data Validation and Reporting Standards (Response to Reviewers)

To enhance methodological rigor, reproducibility, and transparency in computational ethnopharmacology, and in direct response to reviewers' requests for standardized validation and reporting schemes, this section proposes a structured protocol covering database selection, model evaluation, and clinical reporting standards. The proposed framework is intended to reduce methodological heterogeneity across studies and to facilitate cross-study comparison and replication.

To enhance transparency, reproducibility, and clinical interpretability of herb drug interaction studies, we summarize a practical reporting checklist covering data provenance, modeling choices, evaluation protocols, and safety translation (Table X). This checklist is not intended as a rigid standard, but as a pragmatic reference aligned with current best practices in computational pharmacology and medical AI.

Table 2: A practical checklist for herb–drug interaction (HDI) modeling and reporting.

Item	Recommended Reporting Practice
Interaction definition	Specify whether HDI is PK-based (CYP/transporters), PD-based (targets/pathways), or mixed.
Data provenance	Provide sources for herb components, drug labels, enzyme/target annotations, and clinical evidence; include version/date.
Negative samples	Describe how negatives are constructed (random, time-split, debiased); report sensitivity analyses.
Model modality	Clarify whether structure-only, text-only, graph-based, or multimodal; justify modality choice.
Evaluation protocol	Use time-split or external test sets where feasible; report AUROC/AUPRC and calibration.
Explainability	Provide subgraph explanations, attention maps, or mechanistic pathways; avoid "black-box" claims.
Safety translation	Map model outputs to clinical actions: contraindication flags, monitoring suggestions, risk stratification.

7.1 Table Scheme 1: Comparative Guide for Traditional Medicine Databases

This table is designed to assist researchers in selecting appropriate databases according to specific research objectives, evidence requirements, and modeling tasks.

7.2 Table Scheme 2: Performance Evaluation Metrics for AI Models

Different prediction tasks in computational ethnopharmacology require task-specific evaluation metrics. The following standardized metric combinations are recommended to ensure robust and comparable model

Table 3: Comparative Characteristics of Major Traditional Medicine Databases

Database	Core Data Types	Data Scale	Evidence Level	Key Algorithms / Features	Recommended Applications
HERB 2.0	Clinical trials, multi-omics	8,558 trials; 15,515 targets	Very high (RCTs, meta-analyses)	Transcriptomic reverse mapping, knowledge graph	Evidence-based pharmacology, formula validation, prescription discovery
TCMSP	Phytochemicals, ADME parameters	~500 herbs	Low (in silico prediction)	OB/DL filtering rules	Preliminary compound screening, ADME profiling
SymMap v2	Symptom–gene associations	Tens of thousands of mappings	Medium (expert curation)	Phenotype–genotype alignment	Integrative medicine, syndrome mechanism studies
OpenTCM	Classical texts, entity relations	48,000+ entities; 3.73M characters	High (graph-structured validation)	GraphRAG, prompt engineering	Formula mining, intelligent clinical QA systems
ADMETlab 3.0	Toxicity and PK parameters	Large-scale API access	Predictive (cross-validated models)	Multi-task DMPNN	Toxicity alert (hERG/DILI), drug-likeness assessment

assessment.

7.3 Methodological Significance

Together, these protocol designs establish a unified methodological baseline for computational ethnopharmacology. By standardizing database selection, evaluation metrics, and reporting practices, this framework directly addresses reviewer concerns regarding reproducibility, comparability, and clinical relevance. More broadly, it provides a bridge between exploratory AI modeling and evidence-oriented biomedical research, facilitating the responsible translation of computational predictions into experimental and clinical contexts.

8 Conclusions and Future Perspectives

Computational ethnopharmacology is undergoing a profound transformation. The field has evolved from early-stage static network analysis toward an integrated intelligent framework that combines large language models for theoretical interpretation, knowledge graphs for data organization, hypergraph neural networks for modeling prescription compatibility, and AlphaFold-based structural biology for molecular-level validation. Together, these advances are progressively enabling the systematic decoding of the long-standing "black box" problem in traditional Chinese medicine.

Importantly, the significance of this transformation does not lie merely in technological accumulation, but in a fundamental shift in scientific reasoning. High-dimensional mathematical representations are increasingly used to formalize the holistic worldview of traditional medicine, while structure-resolved computational biology provides a mechanistic bridge to modern biomedical science. Through the construction of high-quality digital infrastructures and the adoption of multimodal deep learning paradigms, compu-

Table 4: Recommended Evaluation Metrics for Core Prediction Tasks

Task Type	Suggested Architecture	Key Metrics	Target Reference (SOTA)	Validation Strategy
Drug–Drug Interaction (DDI)	Hypergraph GNN (HyGNN)	ROC-AUC, F1-score	ROC-AUC > 0.98	Five-fold cross-validation + external test set
Drug–Target Interaction (DTI)	Heterogeneous GNN (HDCTI)	AUPR	AUPR > 0.85 (sparse data)	Negative sample augmentation + wet-lab validation (SPR/ITC)
Structure prediction and docking	AlphaFold 3	LDDT, binding energy (ΔG)	LDDT > 0.75; $\Delta G < -7$ kcal/mol	Molecular dynamics simulation (RMSD/RMSF)
Toxicity classification	Multi-task DMPNN	Balanced accuracy, sensitivity	Accuracy > 0.85	Validation using known positive control compounds

tational ethnopharmacology is becoming a viable pathway for translating traditional medical knowledge into modern drug discovery and precision medicine.

Beyond individual methodological innovations, the emerging paradigm offers a new epistemic framework in which traditional empirical knowledge and contemporary computational science can be integrated within a unified analytical system. In this sense, computational ethnopharmacology functions not only as a technical toolkit, but also as a conceptual bridge connecting traditional medical wisdom with evidence-oriented biomedical research.

Looking forward, several directions are critical for the maturation of the field. First, computational prediction must be embedded within a rigorous dry–wet loop. Algorithmic outputs should be treated as testable hypotheses rather than conclusions, and systematic experimental validation using biophysical assays, cellular models, or organoid systems should be incorporated as a standard research component. Equally important, negative experimental results should be explicitly reported and fed back into model retraining to reduce confirmation bias and improve predictive robustness.

Second, the application of large language models in medical and pharmacological contexts requires the mandatory introduction of a fact-checking layer. Knowledge-graph-enhanced retrieval mechanisms are essential to ensure that every AI-generated diagnostic suggestion or prescription recommendation can be traced back to explicit textual sources or empirical evidence, thereby preventing uncontrolled hallucination in clinical decision support scenarios.

Third, the long-standing survivorship bias in traditional medicine data must be addressed systematically. Because historical records overwhelmingly document successful treatments while omitting ineffective attempts, future computational frameworks should prioritize negative sample engineering, contrastive learning strategies, and unlabeled data augmentation. These approaches are necessary for training models that generalize beyond well-studied herbs and canonical formulas.

In summary, computational ethnopharmacology is transitioning from exploratory modeling toward a mature, evidence-oriented research paradigm. By integrating algorithmic innovation with experimental validation and standardized reporting practices, the field is positioned to contribute not only to the modernization and internationalization of traditional medicine, but also to the broader development of AI-assisted drug discovery and systems pharmacology. Ultimately, its most enduring contribution may lie in demonstrating how ancient medical knowledge can be rigorously interpreted, validated, and extended within a contemporary scientific framework.

References

[1] Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology*, 4(11), 682–690. https://doi.org/10.1038/nchembio.118

[2] Gao, R., Wang, S., Chen, Y., et al. (2025). HERB 2.0: An updated database integrating clinical and experimental evidence for traditional Chinese medicine. *Nucleic Acids Research*, 53(D1), D1404–D1412. https://doi.org/10.1093/nar/gkae1046

[3] Xiong, G., Wu, Z., Yi, J., et al. (2024). ADMETlab 3.0: An updated comprehensive online ADMET prediction platform with enhanced coverage and decision support. *Nucleic Acids Research*, 52(W1), W422–W431. https://doi.org/10.1093/nar/gkae236

[4] Wu, Y., Zhang, F., Yang, K., et al. (2019). SymMap: An integrative database of traditional Chinese medicine enhanced by symptom mapping. *Nucleic Acids Research*, 47(D1), D1110–D1117. https://doi.org/10.1093/nar/gky1021

[5] Lewis, P., Perez, E., Piktus, A., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9470.

[6] Edge, D., Trinh, H., Cheng, N., et al. (2024). From Local to Global: A Graph RAG approach to query-focused summarization. *arXiv preprint*, arXiv:2404.16130.

[7] Chen, H., Guo, Y., Li, Z., et al. (2025). OpenTCM: A GraphRAG-empowered large language model system for traditional Chinese medicine knowledge retrieval and diagnosis. *arXiv preprint*, arXiv:2504.20118.

[8] Yang, K., Hou, Y., Li, S., et al. (2023). Uncovering the combinatorial rules of traditional Chinese medicine formulae with hypergraph knowledge representation. *Journal of Ethnopharmacology*, 302, 115904. https://doi.org/10.1016/j.jep.2022.115904

[9] Abramson, J., Adler, J., Dunger, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630, 493–500. https://doi.org/10.1038/s41586-024-07487-w

[10] Zhou, L., Li, Z., Li, J., et al. (2024). Artificial intelligence in traditional Chinese medicine: Theory, application, and future developments. *Pharmacological Research*, 205, 107238. https://doi.org/10.1016/j.phrs.2024.107238

To Cite This Article Xin HUANG.(2025). Integrating Artificial Intelligence with Traditional Wisdom: The Rise of Computational Ethnopharmacology for TCM Discovery and Interaction Prediction. *Medical Research*, 7(4), 41-57. https://doi.org/10.6913/mrhk.070405

(Editor: Kai ZHENG, Email: cpcl@cpcl.hk)