

Tracing Mechanistic Hypotheses from Networks to Experiments in Dieda Qili Tablet Research

Cameron Griffin, Jordan Palmer, Austin Wallace

Abstract

Network pharmacology can organize possible relationships among a multicomponent formulation, molecular targets, and a disease process. It cannot by itself identify which relationship explains a treatment effect. This article uses research on Dieda Qili Tablet and fracture healing to distinguish candidate generation, expression discrimination, molecular compatibility, and causal mechanism. The motivating study combines network analysis, expression-based machine learning, docking, and molecular simulation. Those components answer different questions and can share prior information that makes apparent convergence less independent than it appears. We develop a claim-by-claim evidence framework covering target provenance, network degree, expression contrasts, nested model evaluation, enrichment backgrounds, and physical target engagement. Efficient language-model adaptation may assist evidence extraction, but its output remains another measurement requiring verification. The article presents a methodological synthesis and a prospective validation design, not a new analysis of patient expression data or a demonstration of clinical efficacy.

Keywords: network pharmacology; fracture healing; target prioritization; expression classification; molecular docking; causal interpretation.

1. Several kinds of convergence

A complex formulation can generate a long list of plausible molecular relationships before any particular mechanism has been tested experimentally. Exploring the mechanism of Dieda Qili Tablet on fracture healing based on network pharmacology combined with machine learning models offers a timely setting for examining how those relationships are assembled (S. Sun et al., 2026). Its workflow joins ingredient-related target information, disease-related information, interaction networks, expression analysis, machine learning, and structural modeling. The resulting candidates are useful hypotheses. The inferential question is whether agreement among these components supplies independent evidence of a treatment mechanism or repeatedly reflects the same initial assumptions.

This distinction matters because a candidate can pass several computational filters without being altered by the administered formulation in the relevant tissue. It might be a well-studied disease-associated protein, a highly connected network node, and a strong discriminator of injured and healthy samples. Those properties make it interesting, but none establishes that a constituent reaches it at an effective concentration or that changing its activity improves healing. A mechanistic claim needs a chain connecting exposure, target engagement, pathway change, tissue response, and an outcome. Each computational component should be located on that chain rather than counted as an interchangeable vote.

Network pharmacology was developed to examine drug action in systems of interacting biological components (Hopkins, 2008). Its systems perspective is especially useful when a one-compound, one-target description would be implausibly restrictive. However, moving from a simple model to a network does not remove the need for identification. A denser representation can increase the number of explanations compatible with the observations. This article therefore evaluates the evidence structure rather than attempting to reproduce the motivating study's numerical rankings. No raw expression matrix was downloaded or reanalyzed, and no docking or molecular dynamics calculation was performed for this manuscript.

2. A provenance ledger before an interaction graph

The first useful object is a ledger connecting every candidate constituent to its source, identity, and evidentiary status. Chemical names, database identifiers, stereochemistry, and formulation descriptions can refer to overlapping but nonidentical objects. A reported ingredient is not automatically an absorbed circulating molecule, and a predicted metabolite is not necessarily present in the target tissue. The ledger should distinguish a measured constituent from a database association or a predicted property. Without these distinctions, an attractive network can quietly combine experimentally observed compounds with hypothetical exposures and then treat every edge as equally grounded.

Target provenance deserves the same care. A direct binding assay, a transcriptional association, a computational prediction, and a text-mined relationship offer different support. STRING is a major resource for protein association networks, with evidence channels that should be interpreted in relation to the question being asked (Szklarczyk et al., 2023). An association between two proteins is not necessarily a direct physical interaction, and neither necessarily implies a direction of regulation. Thresholding a combined score can improve tractability while concealing those differences. A reviewable analysis should retain evidence types and report how conclusions change when uncertain or purely text-derived relationships are removed.

Visualization software supports inspection but does not supply the missing biological identification. Cytoscape provides a platform for displaying and analyzing molecular interaction networks (Shannon et al., 2003). A central node can be visually prominent because it has many documented connections, not because it is the formulation's most important target. Degree, betweenness, and related centrality measures depend on the selected network and its observation process. Proteins that have attracted extensive research may acquire more recorded edges. A candidate selected by centrality should therefore be compared with an appropriate network background and described as structurally prominent within that constructed graph, not as a proven causal controller of healing.

Shared databases can also create circularity across stages. Disease-target resources may already incorporate expression studies, and an interaction resource may include literature that motivated the formulation hypothesis. A later enrichment result or expression overlap can then repeat evidence introduced earlier in the workflow. This does not make the resources unusable. It changes the meaning of convergence. The ledger should identify shared source studies where possible and separate analyses based on new observations from transformations of existing annotations. Counting the number of software packages used is not a measure of how many independent lines of evidence support a candidate.

3. Expression discrimination is not a treatment response

An expression contrast can identify genes whose measured abundance differs between groups. The biological meaning depends on the contrast: fracture versus healthy tissue, different healing stages, treated versus untreated samples, or patients with different outcomes. These comparisons cannot be exchanged without changing the claim. The motivating study reports a public expression component associated with fracture-related comparisons. Its use here is limited to that reported design. We do not infer an effect of Diedo Qili Tablet from an expression dataset that was not a randomized administration of that formulation. A disease marker may be a consequence of injury rather than a modifiable driver of repair.

Public repositories such as GEO make independent inspection of study descriptions and expression data possible (Edgar, 2002). Repository availability is only the start of that inspection. The analyst needs the sample annotations, platform, tissue, collection times, and relationships among repeated or matched samples. Differences in cell composition, batch, age, or injury severity can contribute to a group contrast.

If a dataset mixes these variations, a predictive classifier can exploit them successfully while answering a different question from the proposed mechanism. The source of a label matters as much as the numerical values in the expression matrix.

Methods for expression analysis, including the limma framework, help fit specified contrasts and account for particular aspects of the data structure (Ritchie et al., 2015). Their validity still depends on a meaningful design matrix and appropriate preprocessing. A contrast should be written in words before it is encoded, and any adjustment should be justified by the causal or sampling structure. Adding covariates without considering their role can remove part of a biological pathway or introduce instability in a small sample. Conversely, omitting major technical differences can produce a long list of apparently disease-related genes. Software cannot decide which of these problems defines the study.

Machine learning selection adds another inferential layer. Agreement among LASSO, random forest, and support vector machine candidates can identify a useful shortlist, but the models may be trained on the same samples and the same initial gene set. Their intersection is not three independent replications. If feature selection uses the full dataset before cross-validation, the test folds have already influenced the representation. Nested selection is necessary when the reported performance is intended to describe the complete selection procedure (Varma and Simon, 2006). An apparent improvement after choosing a shared feature can otherwise reflect information that has leaked across the intended evaluation boundary.

Even an honestly evaluated classifier answers a predictive question. High discrimination between injured and healthy samples does not imply that pharmacologically changing the selected gene will move a patient from the injured state to a healed state. A causal driver, a compensatory response, and a passive correlate can all be predictive. The motivating study's emphasis on JAK2 can therefore be interpreted as candidate prioritization for further work. Determining direction, timing, and the relevant cell population would require observations or interventions designed for those questions. The statistical importance of a feature is not a prescription for increasing or inhibiting its activity.

4. Enrichment needs a meaningful comparison universe

Pathway enrichment asks whether a selected set contains more members of an annotation category than expected under a specified reference process. The choice of reference universe is therefore substantive. Comparing measured genes with every annotated human gene can be inappropriate if the experiment could only select a restricted subset. Comparing a network-derived shortlist with a background that ignores network degree can favor pathways containing heavily connected proteins. An analysis should state which genes were eligible for selection and how missing annotations were handled. Otherwise a small enrichment probability can reflect the construction of the candidate pool rather than a distinctive treatment mechanism.

Multiple testing also belongs in the design. Testing many pathways, changing thresholds, and reporting only familiar biological terms can produce a persuasive narrative from a broad candidate list. False discovery rate procedures provide a framework for controlling a defined family of tests under their stated conditions (Benjamini and Hochberg, 1995). They do not account automatically for every unreported choice of database, gene cutoff, or network filter. The study should preserve the complete tested set, annotation version, and adjustment rule. A pathway that remains prominent across documented alternative specifications may be a more stable hypothesis, but stability across related annotations is still not direct evidence of biological activation.

Pathways often overlap extensively. Several significant labels may refer to many of the same proteins, and a single multifunctional candidate can generate a long thematic list. Reporting those labels as independent

discoveries exaggerates the breadth of evidence. A clearer interpretation groups overlapping annotations and identifies the genes supporting each theme. It then asks whether the proposed tissue and time window make the relationship plausible. Fracture repair includes changing cellular environments, so a pathway's general association with inflammation or growth is insufficient to establish when a formulation acts or whether its net effect is beneficial.

5. Structural compatibility and physical engagement

Docking evaluates candidate poses under a model of molecular interaction. AutoDock Vina is a widely used tool for efficient docking and scoring (Trott and Olson, 2010). Its output can prioritize structures for further investigation, but a favorable score is not a measured binding affinity. The result depends on the selected protein structure, protonation, ligand preparation, search region, flexibility assumptions, and scoring function. A comparison should use matched preparation and explain the biological relevance of the modeled site. Ranking compounds prepared under incompatible conditions can produce a numerical ordering without a coherent physical interpretation.

Molecular dynamics can explore how a selected modeled complex behaves under a specified force field and simulation setup. Persistence of a contact within that trajectory is conditional on the starting structure and model. It does not establish that the complex forms in living tissue, that the compound reaches the site, or that binding changes the relevant signaling process. A trajectory also needs adequate sampling and a comparison appropriate to the claim. An attractive stable-looking segment should not stand in for a quantified assessment of alternative states. This article makes no new simulation claim and does not convert the motivating study's structural calculations into experimental target engagement.

The next empirical step should match the missing link. If direct engagement is uncertain, a suitable binding or cellular engagement experiment is more informative than another network visualization. If the target is engaged but its role in healing is uncertain, perturbation in a relevant system can test whether the downstream response depends on it. Rescue and orthogonal perturbation designs can help distinguish pathway dependence from nonspecific toxicity. These are proposed validation strategies, not completed experiments or recommendations for patient treatment. Dose, timing, and exposure measurements would be essential before interpreting any biological response as evidence for the formulation's mechanism.

6. Efficient evidence extraction without automated validation

The literature underlying a multicomponent network can be large and inconsistent in terminology. A language model could assist by extracting compound identities, assay types, target relationships, and experimental conditions into a structured ledger. The Hadamard Tensor Ring method is relevant as a candidate way to reduce the cost of adapting such a model (H. Tong et al., 2026). That connection concerns the construction of an evidence-processing tool, not the pharmacological validity of its output. No language model was adapted in this work, and no claimed molecular relationship is accepted merely because an automated extractor can produce it in a standardized format.

An evaluation of that tool should use independently annotated documents and hold out source studies rather than random sentences from the same paper. Otherwise repeated terminology and duplicated descriptions can make extraction appear more general than it is. The tool should distinguish a measured interaction from a prediction and retain the sentence or table supporting each field. Efficiency measurements would include adaptation cost and the work required for human correction. A compact model that systematically upgrades speculative language into experimental fact would create a cleaner-looking ledger while weakening the mechanism argument. Provenance and uncertainty should survive the transformation.

7. A claim that can be tested

The most useful output of an integrated computational study is a short, explicit set of hypotheses with their unresolved links. For each prioritized target, the record should state the proposed constituent exposure, relevant tissue, expected direction of change, and observable consequence for repair. It should also state which existing observation supports each link and which observation could contradict it. This format makes it possible to select an efficient next experiment. It avoids treating a sequence of related computational rankings as if it had already completed the entire causal chain.

The Dieda Qili Tablet study illustrates how network pharmacology, expression modeling, and structural calculations can organize a difficult biological question. Their value is strongest when their distinct evidentiary roles remain visible. A hub is a feature of a constructed graph; a classifier feature is a predictor under a specified sampling design; a docking pose is a model of possible compatibility. A mechanism requires additional connections among exposure, engagement, pathway response, and outcome. Keeping those distinctions explicit turns an attractive candidate list into a research program that can be challenged, refined, and eventually supported by the right observations.

References

- Sun, Shulong, Hua Li, Yan Li, Liubing Yang, Juanjuan Zhang, and Yujing Cao. "Exploring the mechanism of Dieda Qili Tablet on fracture healing based on network pharmacology combined with machine learning models." *Scientific Reports*, vol. 16, no. 1, 2026. <https://doi.org/10.1038/s41598-026-48781-z>.
- Tong, H., Xu, G., & Chen, Y. (2026, May). Hadamard Tensor Ring for Efficient Low-Rank Fine-Tuning. In ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
- Hopkins, Andrew L. "Network pharmacology: the next paradigm in drug discovery." *Nature Chemical Biology*, vol. 4, no. 11, 2008, pp. 682-690. <https://doi.org/10.1038/nchembio.118>.
- Szklarczyk, Damian, et al. "The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest." *Nucleic Acids Research*, vol. 51, no. D1, 2023, pp. D638-D646. <https://doi.org/10.1093/nar/gkac1000>.
- Shannon, Paul, et al. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." *Genome Research*, vol. 13, no. 11, 2003, pp. 2498-2504. <https://doi.org/10.1101/gr.1239303>.
- Edgar, R. "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository." *Nucleic Acids Research*, vol. 30, no. 1, 2002, pp. 207-210. <https://doi.org/10.1093/nar/30.1.207>.
- Ritchie, Matthew E., et al. "limma powers differential expression analyses for RNA-sequencing and microarray studies." *Nucleic Acids Research*, vol. 43, no. 7, 2015, pp. e47-e47. <https://doi.org/10.1093/nar/gkv007>.
- Varma, Sudhir and Richard Simon. "Bias in error estimation when using cross-validation for model selection." *BMC Bioinformatics*, vol. 7, no. 1, 2006. <https://doi.org/10.1186/1471-2105-7-91>.
- Benjamini, Yoav and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 57, no. 1, 1995, pp. 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Trott, Oleg and Arthur J. Olson. "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." *Journal of Computational Chemistry*, vol. 31, no. 2, 2010, pp. 455-461. <https://doi.org/10.1002/jcc.21334>.