Pharmaco-Informatics for Natural Product Repurposing: Disease Similarity, Target Networks, Molecular Signatures, and Therapeutic Hypothesis Generation

William Taylor[1]*, Emily Davis[1], Thomas Wilson[2]

[1]Department of AI for New Zealand Native Plant Bioactives, School of Pharmacy, University of Auckland, Auckland, New Zealand.

[2]Department of Cheminformatics for Anti-Cancer Terpenes, School of Pharmacy, University of Otago, Dunedin, New Zealand.

ABSTRACT

Natural product repurposing has become an important topic in computational pharmacology because plant-derived compounds, phytochemicals, and natural product candidates often interact with multiple targets, pathways, and biological contexts. However, repurposing logic can be overstated when computational associations are interpreted as evidence of efficacy, safety, or clinical readiness. This article proposes an original pharmaco-informatics framework for natural product repurposing that treats disease similarity, target networks, molecular signatures, pathway relationships, compound–target evidence, safety signals, exposure plausibility, and evidence provenance as components of therapeutic hypothesis generation. The framework positions disease ontology and phenotype similarity as entry points for defining disease relevance, target-network mapping as a means of assessing mechanistic plausibility, and molecular-signature analysis as a perturbational layer for identifying alignment or reversal hypotheses. Candidate ranking is framed as evidence integration rather than therapeutic prioritisation, requiring explicit attention to compound identity, target confidence, pathway coherence, omics context, ADMET and toxicity evidence, herb–drug interaction risk, uncertainty grading, and validation readiness. The main contribution is a transparent decision-support logic that connects natural product informatics with repurposing hypothesis generation while preserving interpretation boundaries. The framework is intended to guide research prioritisation, experimental planning, and cautious translation rather than to establish therapeutic efficacy, clinical usefulness, or regulatory acceptability.

Key Words: *Pharmaco-informatics, Natural products, Drug repurposing, Disease similarity, Target networks, Molecular signatures*

eIJPPR 2025; 15(5):1-12

HOW TO CITE THIS ARTICLE: Taylor W, Davis E, Wilson T. Pharmaco-Informatics for Natural Product Repurposing: Disease Similarity, Target Networks, Molecular Signatures, and Therapeutic Hypothesis Generation. Int J Pharm Phytopharmacol Res. 2025;15(5):1-12. https://doi.org/10.51847/TPLl52Sy4S

INTRODUCTION

Drug repurposing is attractive because it can redirect existing pharmacological knowledge toward new therapeutic hypotheses, but repurposing remains a staged evidentiary process rather than a shortcut to clinical proof. For natural product research, this distinction is especially important because computationally suggested disease relevance, target proximity, or molecular-signature reversal can support research prioritisation but cannot establish therapeutic effectiveness without appropriate validation [1].

Natural products and plant-derived compounds remain central to drug discovery because they provide chemically diverse scaffolds, biologically active metabolites, and historically important pharmacological leads. Their

Corresponding author: William Taylor
Address: Department of AI for New Zealand Native Plant Bioactives, School of Pharmacy, University of Auckland, Auckland, New Zealand.
E-mail: ✉ william.taylor@auckland.ac.nz
Received: 14 May 2025; **Revised:** 29 August 2025; **Accepted:** 01 September 2025

repurposing potential, however, requires careful treatment of chemical identity, source context, preparation variability, bioactivity evidence, developability constraints, and safety interpretation, because natural origin does not by itself establish therapeutic suitability [2].

The continued contribution of natural products to pharmacological discovery supports sustained interest in phytochemical and natural product chemical space. Nevertheless, the historical success of natural products as drug sources should not be interpreted as evidence that any individual compound, extract, or preparation is suitable for a new disease context without compound-specific, mechanism-aware, exposure-aware, and safety-aware evaluation [3].

This article develops an original pharmaco-informatics framework for natural product repurposing. The framework integrates disease similarity, target networks, molecular signatures, pathway interpretation, candidate ranking, evidence provenance, validation planning, and translational caution, reflecting the broader computational repurposing literature while restricting the proposed contribution to transparent hypothesis generation rather than implemented prediction, experimental validation, clinical recommendation, or regulatory decision-making [4].

Natural product repurposing as an informatics challenge

Natural product repurposing begins with the informatics problem of identifying what is being evaluated. Curated natural product resources can support compound identity, species-source tracing, activity annotation, and bioactivity provenance, but such resources should be used to define candidates and evidence layers rather than to infer therapeutic readiness from database presence alone [5].

Traditional medicine and natural product databases can help connect ingredients, formulae, targets, and disease associations, creating a starting point for hypothesis generation. Yet database-derived associations often combine heterogeneous evidence types, so repurposing logic must distinguish traditional-use context, chemical annotation, predicted targets, experimental bioactivity, and disease relevance before generating candidate hypotheses [6].

Experiment- and reference-guided herbal informatics resources can extend this evidence base by connecting herbs, compounds, target information, disease links, and expression-related evidence. These connections are valuable for framework construction because they support multi-layer evidence mapping, but they also require provenance checks so that literature-derived, computationally inferred, and experimentally generated signals are not treated as equivalent [7].

Artificial intelligence and machine-learning approaches can assist natural product discovery by supporting chemical representation, target prediction, activity modelling, and candidate prioritisation. In a repurposing framework, these tools should be interpreted as decision-support components whose outputs depend on training data, annotation quality, interpretability, applicability domain, and validation readiness, rather than as autonomous evidence of natural product efficacy [8].

Disease similarity and target networks

Disease similarity analysis requires the disease entity to be represented with curated and computable terminology. Disease-gene platforms provide structured disease-associated targets and evidence provenance, allowing a natural product repurposing workflow to define disease modules before comparing them with compound-associated targets or pathway-level signals [9].

Phenotype similarity provides an additional disease-representation layer when disease labels alone are too broad or heterogeneous. Structured phenotype ontologies allow computational comparison of disease manifestations, but phenotype similarity must be interpreted as contextual evidence for hypothesis generation rather than as proof that two diseases share a therapeutically actionable mechanism [10].

Compound-associated target mapping is a central bridge between natural product identity and disease biology. Heterogeneous network integration can support drug-target interaction prediction and computational repositioning by connecting compounds, targets, diseases, and other biomedical entities, but inferred targets must remain separate from experimentally supported compound–target evidence [11].

Network proximity and pathway relationships can strengthen repurposing hypotheses when disease-associated targets and compound-associated targets occupy biologically plausible regions of an interactome. Network-based repurposing can support candidate prioritisation while still requiring validation beyond the computational signal [12]. Complementarity and separation within biological networks can also inform multi-target or pathway-level plausibility without proving therapeutic benefit [13]. Disease-module approaches can generate repurposing opportunities by connecting disease biology and candidate pharmacology under explicit uncertainty [14]. Interactome-based mapping can further connect disease-relevant host or cellular targets with candidate compounds, although such mapping remains dependent on the completeness and confidence of the underlying network [15].

Table 1 summarises disease-similarity and target-network approaches that can support natural product repurposing hypotheses.

Table 1. Disease Similarity and Target-Network Approaches for Natural Product Repurposing: Disease Ontologies, Phenotype Similarity, Shared Targets, Network Proximity, Pathway Convergence, Mechanistic Plausibility, Interpretation Limits, and Validation Needs

Informatics approach	Core question	Required input data	Repurposing hypothesis contribution	Main uncertainty	Risk of false inference	Validation requirement	Framework role
Disease ontology mapping	What disease concept is being evaluated?	Standardised disease identifiers, synonyms, disease-gene annotations	Defines the disease context for candidate exploration	Disease labels may be broad or inconsistently mapped	Treating a broad disease category as a precise biological state	Expert review of disease definition and disease-model relevance	Establishes the disease-representation layer
Phenotype similarity	Do diseases share structured clinical or biological phenotypes?	Phenotype ontology terms and disease-phenotype annotations	Suggests disease contexts where biological overlap may be explored	Phenotype annotations may be incomplete or unevenly curated	Assuming phenotype resemblance equals shared therapeutic mechanism	Phenotype-relevant disease model and endpoint selection	Adds contextual disease similarity
Disease-associated target mapping	Which targets are linked to the disease?	Disease-gene associations, omics evidence, curated literature	Defines disease modules for comparison with compound targets	Disease-associated targets may be non-causal	Treating association as mechanism	Target-level disease validation and functional assays	Provides disease-side target evidence
Compound-associated target mapping	Which targets are known or predicted for the compound?	Bioassay data, target prediction, binding evidence, literature	Creates a compound target profile for network comparison	Predicted targets may be false positives	Treating predicted targets as confirmed mechanisms	Binding, engagement, and functional target assays	Provides compound-side target evidence
Target overlap	Do compound and disease target sets intersect?	Disease targets and compound-associated targets	Identifies direct target-level alignment	Overlap may reflect annotation bias	Assuming overlap proves therapeutic activity	Target engagement in disease-relevant context	Screens for direct mechanistic plausibility
Network proximity	Are compound targets close to disease modules?	Interactome, disease targets, compound targets	Suggests biological closeness between candidate and disease module	Interactome coverage and edge confidence vary	Equating proximity with efficacy	Perturbation experiments and disease-model testing	Evaluates network-level plausibility
Pathway convergence	Do disease and compound targets converge on relevant pathways?	Pathway annotations, target sets, enrichment logic	Supports pathway-level therapeutic hypothesis generation	Generic pathways may appear enriched across many contexts	Overinterpreting broad pathway enrichment	Pathway-specific functional validation	Filters mechanistic coherence
Pathway complementarity	Could multiple targets affect complementary disease processes?	Target networks, pathway maps, biological process annotations	Supports multi-target hypothesis formation	Complementarity can be speculative	Assuming multi-target action is beneficial	Multi-endpoint mechanistic assays	Supports systems pharmacology interpretation

ISSN (Online) 2249-6084 (Print) 2250-1029

Mechanistic plausibility synthesis	Do disease, target, network, and pathway evidence align?	Integrated disease, target, network, and pathway evidence	Converts isolated signals into a testable hypothesis	Evidence layers may conflict	Hiding weak evidence through broad synthesis	Explicit uncertainty grading and experimental plan	Links disease similarity to validation planning

Molecular signatures and hypothesis generation

Molecular signatures add a perturbational dimension to natural product repurposing by comparing disease-associated molecular states with compound-induced changes. Large-scale perturbational transcriptomic resources support the use of gene-expression signatures for candidate generation, but the resulting signals should be read as context-sensitive hypotheses shaped by cell type, dose, exposure duration, and assay conditions [16].

Perturbation-response data portals make signature evidence more accessible by linking compounds, transcriptional responses, metadata, and analytical resources. For natural product repurposing, such resources can support signature matching or reversal hypotheses only when compound identity, experimental context, and provenance are sufficiently clear to avoid treating a decontextualised expression profile as a general therapeutic claim [17].

Connectivity-based methods can support pharmacogenomic interpretation by identifying compounds whose perturbational profiles resemble or oppose disease-associated signatures. However, connectivity scoring requires careful methodological selection, because different scoring approaches can produce discordant candidate signals [18]. Reconciliation of multiple connectivity scores is therefore important for transparent candidate ranking and uncertainty grading [19]. Reproducibility limitations further reinforce that molecular-signature reversal should not be interpreted as proof of therapeutic effect, especially when disease models, cell states, and perturbation contexts do not match the intended biological setting [20].

Gene-expression similarity can generate repurposing hypotheses by linking compound-induced signatures with disease-relevant molecular states, but such hypotheses still require target, pathway, safety, and validation evidence before they can support translational claims [21]. Annotated repurposing libraries can strengthen this layer by providing standardised compound information and enabling more consistent interpretation of perturbation experiments, although drug-centric libraries may not fully cover the diversity and preparation complexity of natural product candidates [22].

Table 2 maps molecular-signature approaches for generating therapeutic hypotheses in natural product repurposing.

Table 2. Molecular-Signature Approaches for Natural Product Repurposing: Disease Signatures, Compound Perturbation Profiles, Signature Matching, Signature Reversal, Omics Evidence, Pathway Context, Hypothesis Generation, and Interpretation Boundaries

Signature approach	Primary data source	Computational logic	Potential hypothesis signal	Biological interpretation	Main limitation	Validation need	Repurposing relevance
Disease molecular signature	Disease transcriptomic, proteomic, or metabolomic profile	Identify disease-associated molecular changes	Defines molecular state to be compared with compound perturbation	Represents disease biology in a measured context	Disease signatures vary by tissue, model, stage, and platform	Independent disease-context confirmation	Provides disease-side molecular evidence
Compound perturbation profile	Compound-induced omics profile	Measure molecular response after compound exposure	Shows how a candidate perturbs biological systems	Indicates context-specific biological response	Profile depends on dose, exposure, cell type, and assay	Repeated perturbation profiling in relevant systems	Provides compound-side molecular evidence
Signature matching	Compare disease and compound signatures for similarity	Similarity scoring between molecular profiles	Shared molecular-state signal	May suggest overlapping biology or mechanism	Similarity may reflect non-specific stress or assay artefacts	Mechanistic validation of shared pathways	Supports hypothesis generation through molecular resemblance
Signature reversal	Compare disease up/down signatures	Connectivity or reversal scoring	Candidate may oppose disease-	Suggests a therapeutic	Reversal may be non-reproducible or	Disease-model perturbation and	Adds perturbational rationale for

	with opposite compound effects		associated molecular state	hypothesis, not efficacy	context-inappropriate	phenotypic assays	candidate testing
Pathway-context interpretation	Map signatures to pathways or biological processes	Enrichment or pathway mapping of signature genes	Coherent pathway-level signal	Connects omics changes to biological mechanisms	Pathway databases may be incomplete or generic	Functional pathway validation	Links molecular signatures to mechanism plausibility
Multi-omics triangulation	Transcriptomic, proteomic, metabolomic, or other omics evidence where available	Compare evidence across molecular layers	Convergent molecular evidence	Strengthens biological plausibility when layers agree	Omics layers may disagree or be unavailable	Cross-platform and endpoint-specific validation	Reduces reliance on a single molecular evidence type
Candidate-library annotation	Curated compound libraries and perturbation metadata	Link compound identity with profile and metadata	More interpretable perturbation evidence	Supports provenance-aware signature analysis	Natural product coverage may be limited	Compound identity and assay-context verification	Improves evidence traceability
Uncertainty-aware signature scoring	Multiple scoring methods and reproducibility checks	Compare and reconcile connectivity outputs	Robust signal across methods	Highlights stable or unstable candidate hypotheses	Scores may conflict across methods or datasets	Independent replication and sensitivity analysis	Prevents overclaiming from single-score outputs

Candidate ranking and evidence integration

Candidate ranking in natural product repurposing should be treated as an evidence-integration process rather than a numerical declaration of therapeutic priority. Standardised repositioning databases show the value of organising disease, compound, indication, and evidence records in structured formats, supporting a ranking logic in which traceability and evidence type are visible before any candidate is advanced for experimental consideration [23]. Compound–target evidence is a key determinant of ranking quality because known bioactivity should be distinguished from predicted or text-mined associations. Curated bioassay repositories can support this distinction by linking compounds, assays, targets, and activity evidence, allowing a pharmaco-informatics workflow to assign greater interpretive weight to experimentally supported target relationships than to unvalidated computational predictions [24].

Compound-level annotation also matters because candidate ranking requires pharmacological, target, interaction, and chemical information that can be compared across evidence layers. Drug and compound knowledge resources can support this process by organising target, pharmacology, and interaction information, but the presence of an annotation should be interpreted as evidence for structured assessment rather than as confirmation of repurposing suitability [25].

Exposure plausibility, ADMET evidence, toxicity prediction, and interaction risk should function as constraints on candidate ranking rather than as post hoc additions. In silico pharmacokinetic assessment can help flag exposure and developability concerns [26]. Toxicity prediction can identify safety signals that require further investigation, but it cannot establish safety [27]. Integrated ADMET prediction can support multi-property filtering while remaining dependent on model applicability and input quality [28]. Knowledge-graph approaches for pharmacokinetic natural product–drug interactions further indicate that interaction evidence should be incorporated into ranking boundaries, especially when plant-derived compounds may affect enzymes, transporters, or co-administered medicines [29].

Table 3 summarises candidate-ranking and evidence-integration criteria for natural product repurposing.

Table 3. Candidate Ranking and Evidence Integration Criteria for Natural Product Repurposing: Disease Relevance, Target-Network Plausibility, Molecular-Signature Alignment, Pathway Coherence, Exposure Plausibility, Safety Evidence, Evidence Provenance, Validation Readiness, and Translational Caution

Ranking criterion	Core assessment question	Evidence required	High-priority signal	Caution signal	Uncertainty source	Validation need	Decision-support boundary
Disease relevance	Is the candidate linked to a clearly defined disease context?	Disease ontology, disease-gene evidence,	Specific disease module or phenotype-	Broad or ambiguous disease label	Disease heterogeneity and annotation bias	Disease-model confirmation	Supports research prioritisation only

		phenotype annotations	supported context				
Target-network plausibility	Are compound-associated targets biologically close to disease-associated targets?	Compound targets, disease targets, interactome evidence	Evidence-supported proximity or overlap	Target prediction without experimental support	Incomplete interactome and target uncertainty	Target engagement and perturbation assays	Does not prove mechanism
Pathway coherence	Do target and signature evidence converge on relevant pathways?	Pathway mapping, enrichment logic, biological process annotation	Disease-relevant pathway convergence	Generic pathway enrichment	Pathway database bias	Functional pathway assays	Supports mechanistic plausibility
Molecular-signature alignment	Does the compound perturb disease-relevant molecular states?	Disease signatures, compound perturbation profiles, metadata	Context-relevant matching or reversal	Single-context or irreproducible signature	Cell-type, dose, and platform effects	Independent omics validation	Does not prove therapeutic activity
Evidence provenance	Can each evidence layer be traced to its source?	Database records, assay metadata, literature links	Transparent source and evidence type	Untraceable association	Variable curation quality	Manual audit and reproducibility check	Prevents opaque ranking
Compound characterization	Is the candidate chemically and biologically defined?	Structure, identifiers, source context, preparation metadata	Well-characterised compound or standardised preparation	Ambiguous extract or synonym conflict	Identity and composition uncertainty	Analytical confirmation	Anchors natural product identity
Exposure plausibility	Could biologically relevant exposure be plausible?	ADME predictions, pharmacokinetic information, formulation context	Plausible exposure relative to target context	Activity only under implausible conditions	Prediction-model limits	Pharmacokinetic and bioavailability studies	Blocks implausible translation
Safety evidence	Are toxicity concerns visible early?	Toxicity predictions, known adverse signals, assay evidence	No major flagged concern with traceable evidence	Predicted or known toxicity risk	False-negative safety predictions	Toxicology testing	Does not establish safety
Herb–drug interaction risk	Could the candidate modify drug exposure or response?	Enzyme, transporter, pharmacokinetic, and interaction evidence	Low interaction concern with supporting evidence	CYP, transporter, or interaction signal	Sparse interaction evidence	Interaction assays and safety review	Protects translational caution
Validation readiness	Is the hypothesis experimentally testable?	Target, pathway, phenotype, and model-selection evidence	Clear validation endpoint and model	Vague mechanism or non-testable claim	Assay mismatch	Fit-for-purpose experimental plan	Converts ranking into validation planning
Translational relevance	Is there enough convergent evidence to justify further work?	Integrated disease, target, pathway, signature, safety, and exposure evidence	Multiple aligned evidence layers	Single-layer computational signal	Evidence imbalance	Preclinical validation	Supports research decision-making

ISSN (Online) 2249-6084 (Print) 2250-1029

Clinical claim boundary	What claim can responsibly be made?	Evidence grade, validation status, safety profile	Hypothesis or preclinical rationale only	Language implying efficacy or clinical use	Overinterpretation	Clinical evaluation before clinical claims	Prevents treatment recommendation

Proposed pharmaco-informatics framework

The proposed framework begins with natural product identity and evidence provenance as mandatory entry points. Knowledge-graph resources for drug repurposing demonstrate how diseases, compounds, targets, indications, and evidence relations can be organised into computable structures, supporting a framework in which natural product candidates are represented through traceable identifiers, compound characterization, source context, and evidence type before disease or mechanism claims are considered [30].

The second stage integrates disease representation, disease similarity, target-network mapping, and evidence weighting. Systematic biomedical knowledge integration can support drug repurposing prioritisation by connecting heterogeneous biomedical relationships, but such prioritisation remains a computational ranking process rather than a validation result [31]. In the proposed framework, disease ontology, phenotype similarity, disease-associated targets, compound-associated targets, target overlap, network proximity, and pathway convergence are therefore interpreted as evidence layers that can increase or decrease hypothesis plausibility.

The third stage links molecular-signature evidence with mechanism description and candidate interpretation. Knowledge-graph machine-learning frameworks can connect treatment prediction with mechanism-oriented explanation, supporting the idea that a repurposing output should be accompanied by interpretable disease, target, pathway, or molecular relationships rather than presented as an isolated candidate label [32]. In this manuscript's framework, signature matching and signature reversal are incorporated only as context-sensitive hypothesis signals that require metadata review, pathway interpretation, and independent validation.

The fourth stage introduces auditability, uncertainty grading, safety-aware filtering, and decision-support boundaries. Explainable path-based knowledge-graph approaches show the importance of making repurposing logic inspectable through interpretable paths and supporting evidence [33]. For natural product repurposing, this means that candidate ranking should display why a compound is prioritised, which evidence layers support the hypothesis, which layers are missing, whether safety or interaction concerns are present, and what validation would be required before any translational claim could be made.

The final stage is validation planning rather than therapeutic assertion. Computational approaches can streamline discovery by prioritising molecules, mechanisms, and experiments, but they remain part of a broader discovery pipeline that requires biological confirmation, safety assessment, and translational evaluation [34]. The proposed framework therefore ends with therapeutic hypothesis generation, experimental validation priorities, safety review, and decision-support boundaries, not with a clinical recommendation, validated repurposed indication, or regulatory conclusion.

Table 4 presents the proposed pharmaco-informatics framework for natural product repurposing and therapeutic hypothesis generation.

Table 4. Proposed Pharmaco-Informatics Framework for Natural Product Repurposing: Natural Product Identity, Disease Similarity, Target Networks, Molecular Signatures, Safety Filtering, Evidence Integration, Candidate Ranking, Therapeutic Hypothesis Generation, and Validation Planning

Framework stage	Core purpose	Required evidence	Computational method or logic	Hypothesis-generation output	Interpretation boundary	Validation requirement	Failure risk	Research decision value
Natural product identity	Define the evaluated entity	Compound identifiers, source context, preparation metadata, structural information	Identifier normalisation and compound/preparation distinction	Traceable candidate definition	Identity does not establish activity	Analytical confirmation and reproducible characterization	Misidentified or poorly defined candidate	Prevents ambiguous starting points

Compound characterization	Establish chemical and bioactivity context	Structure, compound class, bioassay records, target annotations	Chemical annotation and bioactivity mapping	Compound evidence profile	Bioactivity does not establish disease relevance	Assay review and compound-quality assessment	Overgeneralised bioactivity claims	Supports evidence-weighted screening
Disease representation	Define the disease context	Disease ontology, phenotype terms, disease-gene evidence	Ontology mapping and phenotype annotation	Computable disease model	Disease labels are not mechanisms	Disease-model and endpoint review	Ambiguous disease scope	Anchors disease relevance
Disease similarity analysis	Identify related disease contexts	Disease-gene, phenotype, pathway, and comorbidity evidence where available	Similarity comparison across structured disease data	Disease-transfer hypothesis	Similarity does not prove repurposing suitability	Disease-context validation	False similarity from annotation bias	Supports hypothesis expansion
Disease-associated target mapping	Define disease module	Disease-target records, omics evidence, literature support	Target-set construction and evidence weighting	Disease target profile	Association is not causality	Functional disease-target validation	Non-causal target inclusion	Provides disease-side mechanism evidence
Compound-associated target mapping	Define compound target profile	Bioassay evidence, target prediction, binding or functional data	Separate known targets from predicted targets	Compound target profile	Predicted targets are not confirmed mechanisms	Target engagement assays	False target assignment	Provides compound-side mechanism evidence
Target-network mapping	Assess biological proximity	Interactome, disease targets, compound targets	Target overlap, network distance, module proximity	Network plausibility signal	Proximity is not efficacy	Perturbation and disease-model testing	Interactome incompleteness	Prioritises mechanistic hypotheses
Pathway interpretation	Assess pathway coherence	Pathway annotations, enrichment evidence, biological process data	Pathway convergence and complementarity mapping	Pathway-level rationale	Enrichment is hypothesis-generating	Functional pathway validation	Generic pathway overinterpretation	Filters mechanistic noise
Molecular-signature integration	Compare disease and compound molecular states	Disease signatures, compound perturbation profiles, omics metadata	Signature matching, reversal, and context review	Perturbational hypothesis	Signature reversal is not therapeutic proof	Independent omics and phenotypic validation	Context mismatch or poor reproducibility	Adds molecular-state evidence
Safety and ADMET filtering	Identify translational constraints	ADME prediction, toxicity evidence,	Safety flagging and exposure plausibility review	Candidate risk profile	Prediction is not safety proof	Toxicology, pharmacokinetic, and interaction testing	False reassurance or false alarm	Prevents unsafe overprioritisation

		interaction signals						
Evidence grading	Make uncertainty explicit	Evidence source, study type, assay quality, prediction confidence	Evidence-tier assignment and uncertainty labelling	Transparent confidence profile	Grading is not validation	Independent evidence audit	Overweighting weak data	Supports responsible interpretation
Candidate ranking	Prioritise research candidates	Integrated disease, target, pathway, signature, safety, exposure, and provenance evidence	Multi-criteria ranking with uncertainty penalties	Research-priority list	Ranking is not treatment recommendation	Prospective experimental testing	Score-driven overclaiming	Guides resource allocation
Therapeutic hypothesis generation	Produce a testable claim	Integrated evidence summary and validation gaps	Convert convergent evidence into explicit hypothesis	Mechanism-supported research hypothesis	Hypothesis is not efficacy	In vitro and preclinical validation as appropriate	Vague or non-testable hypothesis	Defines the next experiment
Validation planning	Define next evidence needs	Candidate rationale, disease model, target/pathway endpoints, safety concerns	Assay prioritisation and decision boundary setting	Experimental roadmap	Planning is not validation	Fit-for-purpose validation studies	Wrong model or endpoint	Converts informatics into research action

Figure 1 illustrates the proposed pharmaco-informatics framework connecting disease similarity, target networks, molecular signatures, evidence integration, candidate ranking, and therapeutic hypothesis generation for natural product repurposing.

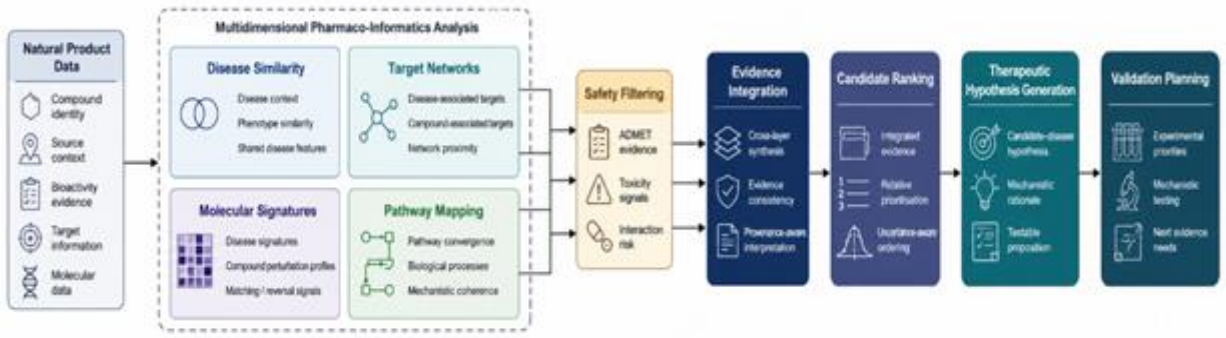

Figure 1. Pharmaco-Informatics Framework for Natural Product Repurposing

Alt-text description

A conceptual framework diagram showing natural product data connected to disease similarity, target networks, molecular signatures, pathway mapping, safety filtering, evidence integration, candidate ranking, therapeutic hypothesis generation, and validation planning.

Research implications

For computational pharmacologists, natural product researchers, phytopharmacologists, and systems pharmacologists, the framework provides a structured way to separate candidate discovery from therapeutic assertion.

ISSN (Online) 2249-6084 (Print) 2250-1029

Academic repurposing efforts are vulnerable to overstated interpretations when computational outputs are treated as translational evidence, so the proposed workflow emphasises validation readiness, evidence provenance, safety review, and claim boundaries before candidate hypotheses are advanced [35].

For bioinformaticians and AI-enabled drug discovery platform designers, the framework highlights the need for interoperable disease ontologies, standardised compound identifiers, confidence-scored compound–target evidence, omics-signature metadata, exposure-aware filtering, and explainable evidence paths. These requirements are especially important for natural product candidates because chemical identity, biological activity, source context, and safety signals may be less standardised than for many approved small-molecule drugs [36].

For journal reviewers and translational scientists, the framework can support more consistent evaluation of natural product repurposing manuscripts. A defensible study should state whether it is presenting disease similarity, phenotype similarity, target overlap, network proximity, pathway convergence, signature matching, signature reversal, candidate ranking, safety screening, validation planning, or a validated therapeutic claim; the framework makes these categories explicit so that computational plausibility is not mistaken for clinical usefulness.

CONCLUSION

This article presents a pharmaco-informatics framework for natural product repurposing that integrates natural product identity, disease representation, disease similarity, target networks, pathway interpretation, molecular signatures, safety and ADMET filtering, evidence provenance, candidate ranking, uncertainty grading, therapeutic hypothesis generation, and validation planning. The framework supports research prioritisation by making computational reasoning transparent and by distinguishing disease similarity, network proximity, molecular-signature reversal, and candidate ranking from validated therapeutic evidence. Its intended role is to guide hypothesis generation and experimental planning while preventing premature clinical, safety, or regulatory claims.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

[1] Pushpakom S, Iorio F, Eyers PA, Escott KJ, Hopper S, Wells A, et al. Drug repurposing: Progress, challenges and recommendations. Nat Rev Drug Discov. 2019;18(1):41-58. doi:10.1038/nrd.2018.168

[2] Atanasov AG, Zotchev SB, Dirsch VM, International Natural Product Sciences Taskforce, Supuran CT. Natural products in drug discovery: Advances and opportunities. Nat Rev Drug Discov. 2021;20(3):200-16. doi:10.1038/s41573-020-00114-z

[3] Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod. 2020;83(3):770-803. doi:10.1021/acs.jnatprod.9b01285

[4] Dotolo S, Marabotti A, Facchiano A, Tagliaferri R. A review on drug repurposing applicable to COVID-19. Brief Bioinform. 2021;22(2):726-41. doi:10.1093/bib/bbaa288

[5] Zeng X, Zhang P, He W, Qin C, Chen S, Tao L, et al. NPASS: Natural product activity and species source database for natural product research, discovery and tool development. Nucleic Acids Res. 2018;46(D1):D1217-22. doi:10.1093/nar/gkx1026

[6] Huang L, Xie D, Yu Y, Liu H, Shi Y, Shi T, et al. TCMID 2.0: A comprehensive resource for TCM. Nucleic Acids Res. 2018;46(D1):D1117-20. doi:10.1093/nar/gkx1028

[7] Fang S, Dong L, Liu L, Guo J, Zhao L, Zhang J, et al. HERB: A high-throughput experiment- and reference-guided database of traditional Chinese medicine. Nucleic Acids Res. 2021;49(D1):D1197-206. doi:10.1093/nar/gkaa1063

[8] Saldívar-González FI, Aldas-Bulos VD, Medina-Franco JL, Plisson F. Natural product drug discovery in the artificial intelligence era. Chem Sci. 2022;13(6):1526-46. doi:10.1039/D1SC04471K

[9] Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res. 2020;48(D1):D845-55. doi:10.1093/nar/gkz1021

[10] Köhler S, Gargano M, Matentzoglu N, Carmody LC, Lewis-Smith D, Vasilevsky NA, et al. The Human Phenotype Ontology in 2021. Nucleic Acids Res. 2021;49(D1):D1207-17. doi:10.1093/nar/gkaa1043

[11] Luo Y, Zhao X, Zhou J, Yang J, Zhang Y, Kuang W, et al. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. Nat Commun. 2017;8(1):573. doi:10.1038/s41467-017-00680-8

[12] Cheng F, Desai RJ, Handy DE, Wang R, Schneeweiss S, Barabási AL, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. Nat Commun. 2018;9(1):2691. doi:10.1038/s41467-018-05116-5

[13] Cheng F, Kovács IA, Barabási AL. Network-based prediction of drug combinations. Nat Commun. 2019;10(1):1197. doi:10.1038/s41467-019-09186-x

[14] Morselli Gysi D, do Valle Í, Zitnik M, Ameli A, Gan X, Varol O, et al. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. Proc Natl Acad Sci U S A. 2021;118(19):e2025581118. doi:10.1073/pnas.2025581118

[15] Sadegh S, Matschinske J, Blumenthal DB, Galindez G, Kacprowski T, List M, et al. Exploring the SARS-CoV-2 virus-host-drug interactome for drug repurposing. Nat Commun. 2020;11(1):3518. doi:10.1038/s41467-020-17189-2

[16] Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A next generation Connectivity Map: L1000 platform and the first 1,000,000 profiles. Cell. 2017;171(6):1437-52.e17. doi:10.1016/j.cell.2017.10.049

[17] Stathias V, Turner J, Koleti A, Vidovic D, Cooper D, Fazel-Najafabadi M, et al. LINCS Data Portal 2.0: Next generation access point for perturbation-response signatures. Nucleic Acids Res. 2020;48(D1):D431-9. doi:10.1093/nar/gkz1023

[18] Musa A, Ghoraie LS, Zhang SD, Glazko G, Yli-Harja O, Dehmer M, et al. A review of Connectivity Map and computational approaches in pharmacogenomics. Brief Bioinform. 2018;19(3):506-23. doi:10.1093/bib/bbw112

[19] Samart K, Tuyishime P, Krishnan A, Ravi J. Reconciling multiple connectivity scores for drug repurposing. Brief Bioinform. 2021;22(6):bbab161. doi:10.1093/bib/bbab161

[20] Lim N, Pavlidis P. Evaluation of Connectivity Map shows limited reproducibility in drug repositioning. Sci Rep. 2021;11(1):17624. doi:10.1038/s41598-021-97005-z

[21] Chen Y, Liu Y, Li W, Xu L, Yang Y, Wang B, et al. Drug repurposing based on the similarity of gene expression profiles: A framework for quinacrine repositioning. Front Chem. 2023;11:1250043. doi:10.3389/fchem.2023.1250043

[22] Corsello SM, Bittker JA, Liu Z, Gould J, McCarren P, Hirschman JE, et al. The Drug Repurposing Hub: A next-generation drug library and information resource. Nat Med. 2017;23(4):405-8. doi:10.1038/nm.4306

[23] Brown AS, Patel CJ. A standard database for drug repositioning. Sci Data. 2017;4:170029. doi:10.1038/sdata.2017.29

[24] Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, et al. ChEMBL: Towards direct deposition of bioassay data. Nucleic Acids Res. 2019;47(D1):D930-40. doi:10.1093/nar/gky1075

[25] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: A major update to the DrugBank database. Nucleic Acids Res. 2018;46(D1):D1074-82. doi:10.1093/nar/gkx1037

[26] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi:10.1038/srep42717

[27] Banerjee P, Eckert AO, Schrey AK, Preissner R. ProTox-II: A webserver for the prediction of toxicity of chemicals. Nucleic Acids Res. 2018;46(W1):W257-63. doi:10.1093/nar/gky318

[28] Xiong G, Wu Z, Yi J, Fu L, Yang Z, Hsieh C, et al. ADMETlab 2.0: An integrated online platform for accurate and comprehensive predictions of ADMET properties. Nucleic Acids Res. 2021;49(W1):W5-14. doi:10.1093/nar/gkab255

[29] Taneja SB, Callahan TJ, Paine MF, Kane-Gill SL, Kilicoglu H, Joachimiak MP, et al. Developing a knowledge graph for pharmacokinetic natural product-drug interactions. J Biomed Inform. 2023;140:104341. doi:10.1016/j.jbi.2023.104341

[30] Boudin M, Diallo G, Drancé M, Mougin F. The OREGANO knowledge graph for computational drug repurposing. Sci Data. 2023;10(1):871. doi:10.1038/s41597-023-02757-0

[31] Himmelstein DS, Lizee A, Hessler C, Brueggeman L, Chen SL, Hadley D, et al. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. Elife. 2017;6:e26726. doi:10.7554/eLife.26726

[32] Ma C, Zhou Z, Liu H, Koslicki D. KGML-xDTD: A knowledge graph-based machine learning framework for drug treatment prediction and mechanism description. Gigascience. 2023;12:giad057. doi:10.1093/gigascience/giad057

[33] Jiménez A, Merino MJ, Parras J, Zazo S. Explainable drug repurposing via path based knowledge graph completion. Sci Rep. 2024;14(1):16587. doi:10.1038/s41598-024-67163-x

[34] Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. Nature. 2023;616(7958):673-85. doi:10.1038/s41586-023-05905-z

[35] Begley CG, Ashton M, Baell J, Bettess M, Brown MP, Carter B, et al. Drug repurposing: Misconceptions, challenges, and opportunities for academic researchers. Sci Transl Med. 2021;13(612):eabd5524. doi:10.1126/scitranslmed.abd5524

[36] Jourdan JP, Bureau R, Rochais C, Dallemagne P. Drug repositioning: A brief overview. J Pharm Pharmacol. 2020;72(9):1145-51. doi:10.1111/jphp.13273