Network-Based Computational Identification of Bioactive Compound Combinations for Personalized Cancer Therapy

Antonietta Ackermann
Department of Computer Science, University of Alabama at Birmingham, Birmingham, AL, USA.
antoniettamail@uab.edu

Tane Herd
Department of Computer Science, Binghamton University, Binghamton, NY, USA.
taneherd598@binghamton.edu

Nienke Gerritse
School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
contactnienke@uc.edu

Abstract

Personalized cancer therapy must select compound combinations that act on the specific dysregulated modules of a tumor while limiting collateral network perturbation. The combinatorial space of bioactive compounds is too large for exhaustive experimental testing, and pair-level synergy models often ignore patient-specific network context, target redundancy, and annotation uncertainty. This manuscript develops a network-based computational framework for identifying and ranking compound combinations for individual tumors. The framework constructs a patient-specific multilayer interaction network, infers dysregulated target modules, maps compounds to protein targets and transcriptional effects, and scores a pair by its excess perturbation of a tumor module beyond an independence reference. A selectivity term penalizes combinations whose targets lie close to essential or healthy-tissue modules, and a robust objective accounts for edge-weight and target-annotation uncertainty. Candidate combinations are ranked under interpretability and safety constraints, producing a reproducible decision rule rather than a single unqualified score. The analytical assessment clarifies when network-community structure improves combination discovery, how mechanistic plausibility differs from statistical association, and which failure modes arise from incomplete interactomes, off-target annotations, and confounding between proliferation and drug response. The framework is intended as a transparent hypothesis-generation and prioritization layer that connects multi-omics evidence, compound-target resources, and pathway-level reasoning. Its practical value lies in narrowing experimental search spaces, exposing assumptions, and producing auditable combination rankings for translational evaluation.

Keywords

network pharmacology; drug synergy; personalized oncology; target modules; robust optimization; multi-omics integration

1. Introduction

Cancer therapy increasingly depends on combinations of bioactive compounds because single agents often produce partial responses and acquired resistance. Resistance can emerge through bypass signaling, target mutation, pathway reactivation, and adaptive rewiring, so effective treatment must often perturb more than one vulnerable node or module in the tumor network [8]. The theoretical case for combination therapy is strong, but the practical search problem is severe. Even a modest library of approved and investigational compounds generates thousands of pair candidates, and each pair may act differently across tumor genotypes, tissue contexts, and prior treatments. Computational prioritization is therefore essential, not as a replacement for experimental pharmacology but as a mechanism for concentrating limited assay capacity on combinations with coherent biological rationale.

Network biology provides a natural language for this problem because molecular lesions do not act as isolated genes but as perturbations of interaction systems [1]. Network pharmacology has extended this view to drug discovery by treating compounds as multi-target perturbations rather than single-key inhibitors [2]. Drug-target networks show that approved drugs often engage several proteins and that target proximity can explain therapeutic effects and adverse events [3]. Dynamic network analyses further emphasize that disease states involve changes in edge activity, community structure, and pathway cross-talk rather than static connectivity alone [4]. These ideas are especially relevant to cancer, where tumor heterogeneity means that the same pathway may be dysregulated through different molecular events in different patients.

Synergy has usually been formalized through dose-response models. The Chou-Talalay framework provides a widely used basis for quantifying combination effects and designing experimental combination studies [5], while Loewe additivity defines an independence reference for drugs with similar mechanisms [6]. Such models remain indispensable for experimental validation, but they are not direct guides to personalized candidate selection when the relevant targets are unknown or when the tumor module is defined by network topology rather than by a single measurable endpoint. Combinatorial therapy in the post-genomic era requires integration of genomic lesions, pathway dependencies, and pharmacological evidence [7]. A network formulation can supply that integration while remaining explicit about its assumptions.

The gap addressed here is the lack of a coherent, auditable computational procedure that converts patient-specific multi-omics data and compound-target evidence into ranked pairs by reasoning over network modules, target coverage, selectivity, and uncertainty. Many predictive models optimize pair response in cell-line panels, but cell-line aggregate models may not transfer to individual tumors because they collapse heterogeneity into a global response function. Conversely, purely mechanistic network scores can be unstable when interactomes are incomplete or when target annotations are sparse. The contribution of this manuscript is a methodological framework that combines module-level network inference, compound-induced perturbation scoring, excess-over-independence synergy, safety-aware selectivity, and robust ranking under constrained decision rules. The framework is designed for hypothesis generation: it narrows candidate spaces, exposes why a combination is prioritized, and separates evidence strength from biological plausibility.

2. Problem Formulation and Network Representation

The starting point is a patient-specific molecular interaction graph. Nodes represent genes, proteins, or protein complexes. Edges represent physical interactions, pathway membership, or functional associations. For patient p, the graph is written as $G_p = (V,E,W_p)$, where V is the shared node set, E is the reference edge set, and W_p contains patient-adjusted edge weights. Edge weights are not assumed to be directly measured physical affinities; they express the strength of functional coupling inferred from prior interaction evidence and patient-specific molecular state. A logistic transformation keeps weights bounded and allows heterogeneous evidence sources to be combined without treating them as commensurate measurements.

$$w_{ij}^{(p)} = \sigma\left(\alpha w_{ij}^{0} + \beta a_{ij} + \gamma x_{ij}^{(p)}\right)$$

Here $w_{ij}^{(p)}$ is the patient-specific weight between nodes i and j, σ is a bounded monotone transform, w_{ij}^{0} is the reference interaction weight, a_{ij} encodes pathway co-membership or curated association, and $x_{ij}^{(p)}$ summarizes patient-specific co-expression or co-alteration. The coefficients α, β, and γ control the relative contribution of reference topology, pathway context, and patient state. This representation is deliberately modular: if a data source is missing for one patient, the corresponding term can be set to zero rather than forcing imputation across all modalities.

Data sources are integrated at the level of evidence types rather than pooled into a single similarity score. Somatic mutations, copy-number alterations, transcript abundance, and methylation can all inform the patient layer; large public tumor atlases provide reference distributions for standardization [9]. Compound-target annotations provide the candidate drug layer and distinguish primary targets from secondary or off-target interactions [10]. Protein interaction and functional association resources provide the reference edge set and pathway context [11]. The key assumption is that a cancer-relevant module should be visible as a connected subgraph whose nodes are both topologically coherent and molecularly dysregulated in the patient.

The output of this stage is not a single network but a family of patient-conditioned networks indexed by modeling choices. That family is important for uncertainty analysis. If a combination is ranked highly only for one edge-weight calibration, the result is fragile. If it remains highly ranked across plausible calibrations, the hypothesis is more robust. This distinction is central to the later robust objective and prevents the framework from treating network reconstruction as a solved preprocessing step.

3. Methodological Framework

3.1 Personalized target-module inference

The first inferential task is to identify modules that represent actionable vulnerabilities. A module M is a connected node set enriched for patient-specific alterations and for genes whose perturbation is plausibly related to tumor maintenance. Rather than selecting individual mutation-bearing genes, the framework scores modules by aggregating standardized molecular deviations and prior biological importance. The module score is normalized by size so that large modules do not dominate solely because they contain more nodes.

$$S_p(M) = \sum_{i} \in M\ d_i^{(p)} b_i / \sqrt{|M|}$$

In this expression $S_p(M)$ is the dysregulation score of module M in patient p, $d_i^{(p)}$ is the standardized molecular deviation of node i, b_i is a prior weight reflecting cancer relevance or

essentiality, and |M| is the number of nodes in the module. The square-root normalization helps control inflation from module size while preserving the contribution of coordinated alterations. Candidate modules can be obtained by community detection, pathway enrichment, or seeded propagation from driver lesions; the scoring rule is agnostic to the discovery method and can be applied consistently across alternatives.

The interpretation of $S_p(M)$ matters. A high score indicates coordinated dysregulation, not proven dependency. Some modules may be highly altered because they reflect proliferation or tissue lineage rather than a tumor-specific vulnerability. For that reason, the framework compares module scores against matched normal or lineage reference distributions where available and flags modules whose signal is dominated by broad proliferation programs. This step is a confounder-control mechanism, not a statistical afterthought. Personalized ranking without such controls tends to prioritize generic cytotoxic sensitivity rather than tumor-selective mechanisms.

3.2 Compound-target perturbation modeling

Once modules are inferred, each compound is represented by its target set and by the direction and magnitude of its expected effect. Compound-target evidence may come from curated databases, transcriptional connectivity signatures, or docking predictions. Transcriptional profiling resources such as the L1000 platform provide broad perturbation signatures that can be used to infer pathway-level effects when direct target binding data are incomplete [12]. Structural docking can add target-level hypotheses for compounds with sparse annotation, but docking scores should be treated as weak evidence unless supported by experimental affinity data [13]. The framework therefore separates binding evidence from transcriptional evidence and allows each to contribute through calibrated weights.

The perturbation score for compound c on module M is a weighted coverage of module nodes by compound targets and their downstream effects. It is bounded between zero and one and can be interpreted as the fraction of patient-relevant module importance that the compound is expected to engage. The weighting includes both target affinity and patient-specific module importance, so that a compound does not receive credit for targeting a node that is present but not dysregulated.

$$P_p(c, M) = \sum_i \in T(c) \cap Mw(c, i)h_i^(\frac{p)}{\sum_j \in Mh_j^(p)}$$

Here $P_p(c,M)$ is the perturbation score of compound c on module M for patient p, T(c) is the annotated target set of c, w(c,i) is a normalized target evidence weight, and $h_i^{(p)}$ is the patient-specific importance of node i, derived from module membership and dysregulation. The denominator normalizes by total module importance, preventing compounds that target many unimportant nodes from receiving inflated scores. The numerator rewards coverage of high-importance nodes and is reduced when target annotations are uncertain.

This formulation does not assume that hitting more targets is always better. A compound with broad low-affinity activity may achieve moderate P_p scores while producing unacceptable off-target effects. The score is therefore a module engagement estimate, not a therapeutic utility function. It becomes useful only after combination excess and safety terms are introduced.

3.3 Network synergy and combination scoring

The central quantity is the excess perturbation of a tumor module by a pair beyond what would be expected from independent single-agent effects. Classical synergy models provide the conceptual reference. If two compounds act independently on a module, their expected combined coverage follows an inclusion-exclusion form. The network excess is the difference between the observed or predicted joint module perturbation and that independent reference. Positive values suggest complementary targeting; negative values suggest redundancy or antagonistic overlap. Community-level formulations of synergistic combination discovery reinforce the idea that combinations should be evaluated against the organization of complex systems rather than only against pairwise target overlap [14]. Graph-based combination prediction has also shown that network relationships among drug targets can carry predictive information beyond expression signatures alone [15].

$$\Delta_p(a,b) \;=\; P_p(\{a,b\},M) \;-\; [P_p(a,M) + P_p(b,M) - P_p(a,M)P_p(b,M)]$$

In this expression $\Delta_p(a,b)$ is the network excess for compounds a and b, $P_p(\{a,b\},M)$ is the joint perturbation score, and the bracketed term is the independence reference. The joint score can be computed by taking the union of target effects and propagating them through the patient network, or by using a saturating combination rule when two compounds converge on the same node. The framework should not double-count redundant target engagement. If both compounds inhibit the same node with the same direction, the joint perturbation saturates rather than adding linearly.

A high Δ_p alone is insufficient for prioritization. A pair may strongly perturb a tumor module while also perturbing essential or healthy-tissue modules. The framework therefore introduces a selectivity penalty based on network distance between combination targets and a safety set H, which may include essential genes, normal tissue expression modules, or known toxicity pathways. The penalty is not a clinical risk estimate; it is a topological proxy for collateral engagement. A pair whose targets lie very close to essential nodes receives a larger penalty than a pair with similar tumor-module coverage but greater separation from the safety set.

$$D_p(a,b) \;=\; min_h \in H\; dist(T(a) \cup T(b), h)$$

Here $D_p(a,b)$ is the minimum network distance from the union of the two target sets to any node or module in the safety set H, and dist is a shortest-path or diffusion distance on the patient network. Smaller values indicate greater proximity and therefore greater concern. The penalty enters a scalar utility together with the excess score and an interpretability term $I_p(a,b)$ that rewards combinations with documented pathway explanations. A simple additive form is $F_p(a,b;\theta)=\Delta_p(a,b;\theta)-\lambda D_p(a,b;\theta)+\eta I_p(a,b)$, where λ and η are trade-off parameters and θ denotes network or annotation parameters. This utility is not a probability of response; it is a ranking device whose components remain auditable.

3.4 Robust personalization and uncertainty

Network models contain multiple sources of uncertainty: edge weights are inferred, target annotations are incomplete, and module boundaries are not unique. A combination ranked highest under one reconstruction may fall sharply under another. The robust objective addresses this by evaluating the utility over a plausible set of model parameters and penalizing solutions whose utility is highly variable. This is not a substitute for experimental replication, but it prevents overconfident rankings based on a single interactome version.

$$R_p(a,b) \;=\; E\theta[F_p(a,b;\theta)] \;-\; \kappa\sqrt{Var\theta[F_p(a,b;\theta)]}$$

In this expression $R_p(a,b)$ is the robust ranking score, $E\theta$ is the expectation over plausible parameter settings θ, $Var\theta$ is the corresponding variance, and κ controls risk aversion. A larger κ favors combinations with stable utility across network reconstructions, even if their mean utility is slightly lower. The parameter set can include edge-weight scaling, target-annotation confidence thresholds, and module-resolution choices. The robust score is especially useful when candidate pairs have similar mean utilities but different sensitivity to missing edges or uncertain targets.

Ranking is then performed under constraints that encode minimal biological and safety requirements. The constraints prevent the optimizer from selecting a pair with high excess but negligible tumor-module coverage or unacceptable proximity to the safety set. They also allow the decision rule to be transparent: a pair is eligible only if it satisfies coverage and selectivity conditions, and among eligible pairs the robust score determines priority. This separation between eligibility and ranking makes the framework easier to audit than a single unconstrained objective.

$$(a,b)* \;=\; argmax\, R_p(a,b)\; subject\; to\; D_p(a,b) \;\leq\; \delta$$

Here $(a,b)*$ is the selected pair, argmax denotes the pair maximizing the robust score, and δ is a tolerance for safety-set proximity. Additional constraints can be imposed on minimum module coverage, maximum target overlap, or required mechanistic diversity. The constraint set should be pre-specified for a given analysis to avoid post hoc selection. If no pair satisfies the constraints, the framework returns the constraint set as part of the output rather than silently relaxing thresholds.

3.5 Computational workflow and baselines

The complete workflow has six stages. First, patient molecular data are standardized and mapped to the reference network. Second, module candidates are generated and scored. Third, compounds are annotated with target, transcriptional, and structural evidence. Fourth, single-agent perturbation scores are computed for each compound-module pair. Fifth, pair scores are computed with a saturating joint perturbation rule, safety penalty, and robust evaluation. Sixth, eligible pairs are ranked and reported with component scores and sensitivity intervals.

Baselines are necessary to establish whether network reasoning adds value beyond simpler heuristics. A target-overlap baseline counts shared targets and favors pairs with little overlap. An expression-only baseline predicts response from tumor expression signatures without network propagation. A network-proximity baseline ranks pairs by target separation or pathway complementarity [15]. A deep synergy baseline learns pair-level response patterns from combination screening data [16]. A drug-sensitivity baseline predicts single-agent response and combines predictions additively [17]. The proposed framework should be compared with these baselines on ranking stability, recovery of known mechanistic combinations, and calibration of uncertainty, not only on aggregate predictive accuracy. Because the framework is designed for hypothesis generation, its primary evaluation metric should include the coherence of the biological explanation attached to each high-ranked pair.

4. Analytical Assessment and Design Trade-offs

4.1 Mechanistic plausibility versus statistical association

Network scores can appear mechanistic while remaining statistical associations. An edge in a protein interaction network does not prove that perturbation of one node changes the other in a tumor. A module score does not prove dependency. The framework must therefore

distinguish three levels of evidence: molecular association, network-based mechanistic hypothesis, and experimentally demonstrated combination effect. Large combination screens provide important benchmarks for the third level, but they also reveal that many nominally synergistic pairs are context-specific and that effective combinations vary by tissue lineage and molecular subtype [18]. Unbiased oncology compound screens further show that combination effects cannot be inferred reliably from single-agent potency alone [19]. These findings argue for cautious interpretation of any purely computational ranking.

The proposed framework responds by keeping component scores separate. A pair may have high Δ_p because of dense target coverage, but if its target annotations are weak or its transcriptional signature is inconsistent, the evidence level should be marked as low. A pair may have strong literature support for individual targets but no network excess if the targets converge redundantly on the same node. The interpretability term I_p is intended to reward mechanistic diversity only when it is supported by pathway evidence, not when it merely increases the number of annotated targets.

4.2 Network community structure and combination novelty

Community structure is a central reason network methods can find non-obvious combinations. Two compounds may have little target overlap yet perturb adjacent communities that are coupled through a disease module. A pair that bridges two dysregulated communities may produce greater module-level effect than a pair that hits two nodes in the same community. However, community detection is sensitive to resolution parameters and edge weights. If resolution is too coarse, distinct vulnerabilities are merged; if too fine, the same pathway is split into redundant pieces. The robust objective mitigates this instability by averaging over module-resolution settings and penalizing pairs whose rank depends on a single partition.

Novelty is not the same as plausibility. A pair that connects two distant communities may be novel but biologically implausible if the communities are not functionally coupled. The framework therefore requires every high-ranked pair to have a connecting path or shared downstream mediator supported by the patient network. This requirement reduces the risk of recommending combinations that are topologically novel but mechanistically disconnected.

4.3 Evaluation logic and failure modes

Evaluation of a personalized combination framework must address ranking, calibration, and decision utility. Ranking metrics assess whether known effective pairs appear near the top of the candidate list. Calibration assesses whether predicted uncertainty matches observed variability. Decision utility assesses whether the ranking helps select assays or treatments under resource constraints. These are different objectives, and optimizing one can degrade another. Machine learning in drug discovery has shown that predictive performance depends heavily on data representation, task definition, and validation design [20]. The same caution applies here: a network score that performs well on one tumor type may fail on another because the relevant modules, normal-tissue constraints, and compound annotations differ.

Several failure modes are predictable. Incomplete interactomes can make a coherent module appear disconnected, causing false negatives. Overly dense target annotations can make many compounds appear to cover a module, causing false positives. essentiality annotation errors can invert the safety penalty. Confounding by proliferation can make cytotoxic combinations rank highly without tumor selectivity. Finally, data leakage can occur if compound-target annotations are derived from the same response data used for evaluation.

The framework should use external annotation sources and pre-specified validation splits, and it should report component scores so that reviewers can diagnose which term drove a ranking.

5. Technical Discussion and Clinical Translation

Clinical translation requires more than a ranked list. A useful decision-support system must expose why a combination is proposed, which patient features drive the module score, which targets are engaged, and how sensitive the ranking is to annotation uncertainty. The framework's component scores provide this audit trail. A clinician or translational scientist can ask whether a high-ranked pair depends on a marginal module, an uncertain target, or a large safety penalty. This transparency is especially important because cancer combination therapy often involves overlapping toxicities that are not captured by molecular networks alone.

Calibration is the next requirement. The robust score R_p is not a probability of clinical benefit. It orders hypotheses under an explicit utility model. Any mapping from R_p to expected response would require prospective outcome data and a separate calibration model. Until such calibration exists, the framework should be used to prioritize experimental or organoid testing, not to justify treatment directly. The constraint set also needs clinical input: essential genes, normal tissue modules, and toxicity pathways should be defined with domain experts rather than inferred solely from generic networks.

The framework has boundary conditions. It is most appropriate when a tumor has at least one coherent dysregulated module and when compound-target annotations are available for the candidate library. It is less appropriate for immunotherapies, antibody-drug conjugates, or cell therapies whose mechanisms depend on cellular interactions not represented as intracellular protein networks. It is also limited when the therapeutic goal is to modulate the tumor microenvironment rather than tumor-cell-intrinsic signaling. These boundaries should be stated explicitly in any application, because extending the method beyond its representational scope would produce rankings that are technically precise but biologically misleading.

From a computational perspective, the main design trade-off is between model richness and identifiability. A richer patient-specific network can capture more biology but also more parameters that cannot be estimated from a single tumor sample. The multilayer formulation mitigates this by sharing reference topology across patients and learning only a small number of patient-specific weights. The robust objective further prevents the model from relying on unstable edge estimates. The cost is that the framework may miss highly individualized mechanisms that are not represented in the reference network. This is an acceptable trade-off for hypothesis prioritization, but it argues for combining network scoring with orthogonal evidence such as functional genomics or pharmacogenomic screening.

6. Conclusion

This manuscript formulates a network-based computational framework for identifying bioactive compound combinations in personalized cancer therapy. The framework represents each tumor as a patient-conditioned molecular network, infers dysregulated target modules, maps compounds to target and transcriptional effects, and scores pairs by the excess perturbation of tumor modules beyond an independence reference. A selectivity term based on proximity to essential or healthy-tissue modules penalizes collateral network engagement, while a robust objective accounts for edge-weight, module-resolution, and target-annotation uncertainty. Constrained ranking produces an auditable list of candidate pairs with component

scores that distinguish biological rationale from evidence strength. Analytical assessment shows that network community structure can support non-obvious combination hypotheses, but it also exposes failure modes related to incomplete interactomes, annotation bias, proliferation confounding, and overinterpretation of topological proximity. The framework is therefore best understood as a prioritization layer for experimental and translational investigation, not as a validated clinical predictor. Its scientific value lies in making assumptions explicit, narrowing combinatorial search spaces, and linking multi-omics evidence to mechanism-level combination hypotheses in a reproducible way.

References

1. Barabási, A.-L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. Nature Reviews Genetics, 5(2), 101–113. https://doi.org/10.1038/nrg1272

2. 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

3. Yıldırım, M. A., Goh, K.-I., Cusick, M. E., Barabási, A.-L., & Vidal, M. (2007). Drug-target network. Nature Biotechnology, 25(10), 1119–1126. https://doi.org/10.1038/nbt1338

4. Csermely, P., Korcsmáros, T., Kiss, H. J. M., London, G., & Nussinov, R. (2013). Structure and dynamics of molecular networks: A novel paradigm of drug discovery. Pharmacology & Therapeutics, 138(3), 333–408. https://doi.org/10.1016/j.pharmthera.2013.01.016

5. Chou, T.-C. (2006). Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. Pharmacological Reviews, 58(3), 621–681. https://doi.org/10.1124/pr.58.3.10

6. Loewe, S. (1953). The problem of synergism and antagonism of combined drugs. Arzneimittel-Forschung, 3(6), 285–290.

7. Al-Lazikani, B., Banerji, U., & Workman, P. (2012). Combinatorial drug therapy for cancer in the post-genomic era. Nature Biotechnology, 30(7), 679–692. https://doi.org/10.1038/nbt.2284

8. Holohan, C., Van Schaeybroeck, S., Longley, D. B., & Johnston, P. G. (2013). Cancer drug resistance: An evolving paradigm. Nature Reviews Cancer, 13(10), 714–726. https://doi.org/10.1038/nrc3599

9. Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., & Stuart, J. M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nature Genetics, 45(10), 1113–1120. https://doi.org/10.1038/ng.2764

10. Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., ... Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. Nucleic Acids Research, 46(D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037

11. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & von Mering, C.

(2019). STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

12. Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., Gould, J., Davis, J. F., Tubelli, A. A., Asiedu, J. K., Lahr, D. L., Hirschman, J. E., Liu, Z., Donahue, M., Julian, B., Khan, M., Wadden, D., Smith, I. C., Lam, D., ... Golub, T. R. (2017). A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell, 171(6), 1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049

13. Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31(2), 455–461. https://doi.org/10.1002/jcc.21334

14. HerbSyner_Finder: a network community-based model for identifying synergistic combinations from herbal medicines and complex systems. Targetome. 2026;2(2):0-0. doi: 10.48130/targetome-0026-001

15. Cheng, F., Kovács, I. A., & Barabási, A.-L. (2019). Network-based prediction of drug combinations. Nature Communications, 10, 1197. https://doi.org/10.1038/s41467-019-09186-x

16. Preuer, K., Lewis, R. P. I., Hochreiter, S., Bender, A., Bulusu, K. C., & Klambauer, G. (2018). DeepSynergy: Predicting anti-cancer drug synergy with deep learning. Bioinformatics, 34(9), 1538–1546. https://doi.org/10.1093/bioinformatics/btx806

17. Costello, J. C., Heiser, L. M., Georgii, E., Gönen, M., Menden, M. P., Wang, N. J., Bansal, M., Ammad-ud-din, M., Hintsanen, P., Khan, S. A., Mpindi, J.-P., Kallioniemi, O., Honkela, A., Aittokallio, T., Wennerberg, K., Collins, J. J., Gallahan, D., Singer, D., Saez-Rodriguez, J., ... Stolovitzky, G. (2014). A community effort to assess and improve drug sensitivity prediction algorithms. Nature Biotechnology, 32(12), 1202–1212. https://doi.org/10.1038/nbt.2877

18. Jaaks, P., Coker, E. A., Vis, D. J., Edwards, O., Carpenter, E. F., Leto, S. M., Dwane, L., Sassi, F., Lightfoot, H., Barthorpe, S., van der Meer, D., Yang, W., Beck, A., Moudgil, V., Hotta, T., Silva, D., Beck, A., ... Garnett, M. J. (2022). Effective drug combinations in breast, colon and pancreatic cancer cells. Nature, 603(7899), 166–173. https://doi.org/10.1038/s41586-022-04437-2

19. O'Neil, J., Benita, Y., Feldman, I., Chenard, M., Roberts, B., Liu, Y., Li, J., Kral, A., Lejnine, S., Loboda, A., Arthur, W., Cristescu, R., Haines, B. B., Winter, C., Zhang, T., Bloecher, A., & Stover, D. (2016). An unbiased oncology compound screen to identify novel combination strategies. Molecular Cancer Therapeutics, 15(6), 1155–1162. https://doi.org/10.1158/1535-7163.MCT-15-0843

20. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., & Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463–477. https://doi.org/10.1038/s41573-019-0024-5