



Views & Comments

Phenotype–Target Coupled Drug Screening: A High-Efficiency Framework for Innovative Drug Discovery from CHMs

Wei Zhou^{*}, Yue Gao^{*}

Beijing Institute of Radiation Medicine, Beijing 100850, China

Chinese herbal medicines (CHMs) serve as the cornerstone of traditional Chinese medicine (TCM) practices and are vital sources of inspiration for novel drug discovery. Many landmark drugs, including artemisinin, ephedrine, bicyclol, berberine, and *dl*-3-*n*-butylphthalide, originated from CHMs. Nevertheless, only 23.5% of the new drugs approved by the US Food and Drug Administration (FDA) over the past four decades have stemmed from botanical drugs, natural products, or their derivatives [1]. The primary reason for this disparity lies in the diverse and complex chemical composition of natural medicines such as CHMs combined with poorly characterized *in vivo* multicomponent pharmacokinetics and multitarget mechanisms of action. These factors collectively not only contribute to the therapeutic “black-box” effect of CHMs but also present great challenges for TCM-inspired new drug research and development (R&D). However, in recent decades, drug discovery has prioritized target-based drug discovery (TDD), a paradigm that emphasizes well-defined drug–target interactions (DTIs). Advances in artificial intelligence (AI) and related technologies have significantly enhanced target identification capabilities, particularly in terms of resolving three-dimensional target structures. Consequently, target-guided drug screening, synthesis, and design have become dominant approaches in new drug R&D [2]. This target-centric framework inherently disadvantages TCM resources, where active constituents and molecular targets often remain uncharacterized. Thus, developing novel drug discovery strategies specifically tailored to the unique properties of CHMs is imperative.

1. Phenotype–target coupled drug screening (PTDS): A strategic framework for innovative drug discovery from CHM

The phenotypic drug discovery (PDD) approach, which involves identifying drugs by anchoring screening to specific characteristic states or symptom changes, was historically an empirical mainstay in drug R&D. Most FDA-approved first-in-class drugs were discovered via this target-agnostic method, driving its resurgence since 2011 [3]. Since CHMs exert therapeutic effects through integrated regulation involving multiple components, multiple targets, and

multiple pathways, a drug screening method prioritizing phenotypic outcomes—such as PDD, which functions as an end-to-end approach transcending complex chemical mixtures and undefined mechanisms of action—appears to be especially compatible with new drug discovery from CHM sources. That said, this does not suggest that TDD should be completely disregarded in CHM-based drug R&D. Through the stepwise screening and evaluation of phenotypic responses across hierarchical levels (from molecular and cellular to tissue and whole organism), researchers can expedite the localization of active single compounds or compound groups. Moreover, this layered approach yields crucial clues for deeply interpreting the pharmacological mechanisms, ultimately enabling efficient identification of the key molecular target(s) underlying efficacy and further promoting drug rediscovery via the TDD method. Furthermore, compounds identified via target deconvolution can be systematically verified and prioritized through multi-level, stepwise phenotypic screening, significantly increasing TDD success rates. Evidently, a PTDS strategy employing progressive, macro-/micro-integrated evaluation offers a novel pathway for pioneering CHM-based drug discovery.

2. Advanced technologies facilitate the novel architecture of high-throughput PTDS paradigms

Notably, advances in AI, life multiomics, and microphysiological models have enabled the PTDS strategy to overcome critical technical bottlenecks, including panoramic analysis of disease feature networks, comprehensive profiling of CHM (particularly prescription) components at target organs, and precise identification of microphenotypic signatures. Supported by these novel methodologies, PTDS undergoes continuous refinement and iterative optimization, establishing a resource-efficient, high-yield discovery pipeline for innovative CHM-derived drugs (Fig. 1).

2.1. Drug screening via dynamically corrected dysregulation molecular networks

Given the spatiotemporal complexity of disease pathogenesis, longitudinal multiomic integration analysis drives a paradigm shift from identifying static molecular targets at isolated timepoints or omic layers toward mapping multidimensional dysregulation molecular networks (e.g., gene–protein–metabolism–microbiome

^{*} Corresponding authors.

E-mail addresses: zhouweisy1802@163.com (W. Zhou), gaoyue@bmi.ac.cn (Y. Gao).

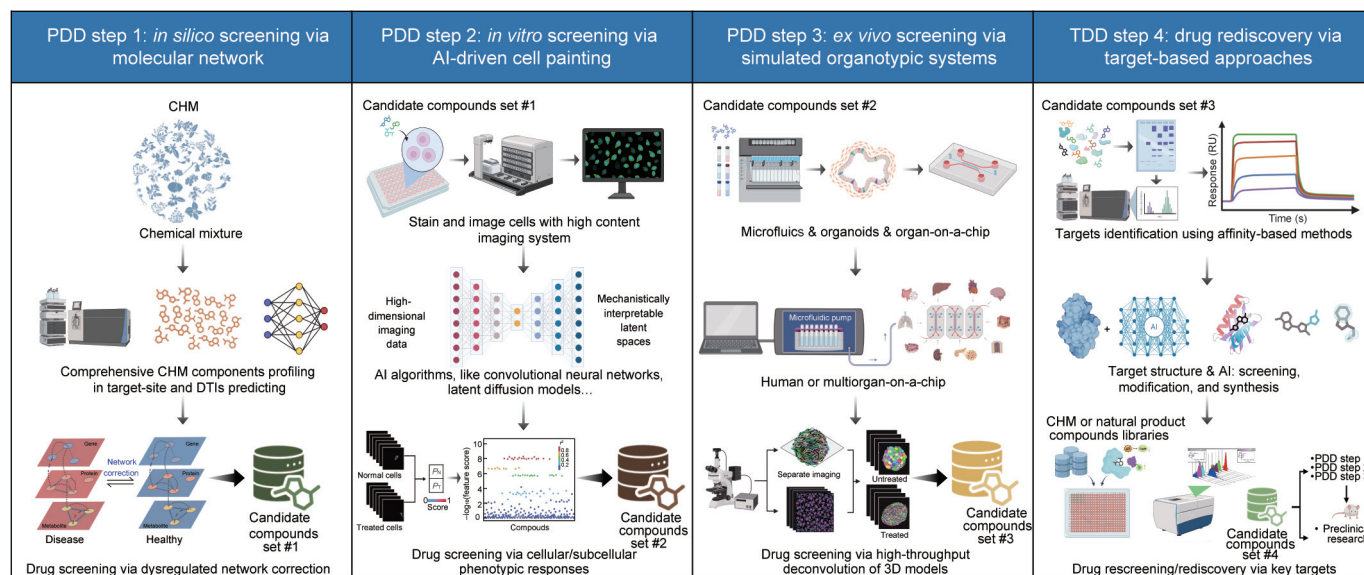


Fig. 1. PTDS: a stepwise and high-efficiency workflow for innovative drug discovery from CHMs. 3D: three-dimensional; P_N : the probability of normal cells; P_T : the probability of treated cells.

interactomes) that bridge macro- and micro-scale phenotypes. Representative integration strategies include matrix factorization approaches (e.g., nonnegative matrix factorization), Bayesian network inference, and similarity network fusion (SNF), which enable the identification of coherent molecular modules across heterogeneous omic layers [4]. Moreover, dynamic network biomarker (DNB) algorithms—such as the DNB theory, time-series network entropy analysis, and temporal clustering of differential coexpression—allow real-time monitoring of the spatiotemporal trajectory of a specific phenotype-associated molecular network and prediction of critical transitions from normal to disease states [5]. Moreover, many *in silico* DTI prediction methods, particularly AI-driven computational algorithms, can help predict interactions and binding affinities between compounds derived from CHMs and targets in molecular networks and eventually achieve qualitative and quantitative assessments of CHM efficacy through network correction potential (driving the characteristic molecular network trajectory to return from the pathological state to the normal state). In addition, this network-based virtual drug screening strategy can provide preliminary information on the important active ingredients and molecular targets of the candidate CHM. Critically, compounds used for DTI prediction must represent the bioactive fractions of candidate CHMs that reach target organs/tissues to exert therapeutic effects (i.e., bioactive components at target sites).

Advances in high-resolution mass spectrometry (MS)-based CHM metabolomics and AI algorithms now enable comprehensive CHM component profiling at target sites. Briefly, multiple MS-based instruments (e.g., gas chromatography–tandem MS (GC–MS/MS), liquid chromatography (LC)–MS/MS, and LC–electrospray ionization (ESI)–ion mobility (IM)–MS) are generally used to directly characterize fragment ions of raw CHM materials and their bioactive components (including absorbed prototype compounds and related metabolites) present in target tissues in depth without considering complex *in vivo* metabolic processes. Next, the tailored MS compound databases derived from reference compound libraries, metabolites predicted *in silico*, and chemical profile-based pseudomolecular ions are employed to enhance the identification of the targeted components of CHMs [6]. In addition, the bioactive components known to reach target organs for certain CHMs can be found in existing studies or specialized databases such as database of constituents absorbed into the blood and metabolites of TCM

(DCABM-TCM) [7]. Furthermore, spatially resolved MS imaging has emerged as a powerful tool for visualizing the tissue distribution of CHM-derived active constituents, thereby circumventing interference from CHM-related compounds circulating in the blood. Together, these complementary approaches provide a feasible strategy for the rapid and comprehensive profiling of active CHM fractions that effectively reach target sites.

2.2. Drug discovery via AI-driven cell painting

Cell painting was initially a quantitative morphological characterization method that combined multichannel fluorescence staining with high-content microscopic imaging. This method typically involves the characterization of eight cellular structures using six dyes to capture holistic cellular architecture, including nuclei, cytoplasmic features, and cytoskeletal organization. Along with the development of fluorophore design, optical imaging systems and high-throughput automated liquid handling have expanded cell painting beyond cellular morphology to enable quantitative profiling of subcellular processes [8]. This evolution now permits the measurement and integration of critical subcellular biological events, such as mitochondrial membrane potential dynamics, endoplasmic reticulum stress responses, and DNA damage mechanisms, into multidimensional phenotypic profiles. The constant emergence of AI algorithms, particularly deep learning and representation learning frameworks such as convolutional neural networks, self-supervised representations, and latent diffusion models, has enabled the compression of high-dimensional imaging data into mechanically interpretable latent spaces. Hence, AI-driven cell painting based on cellular/subcellular phenotypic responses could be pivotal for new drug discovery from CHMs for which the underlying mechanism and molecular targets are often elusive.

In particular, recent advances in this PDD approach have progressively addressed several critical practical challenges. Data standardization has been notably improved through initiatives such as the joint undertaking for morphological profiling–cell painting (JUMP-CP) consortium, which introduced the quantitatively refined “Cell Painting v3” protocol. This update standardizes staining, imaging, and quality control procedures across different laboratories, mitigating batch effects and ensuring consistent

inputs for modeling [9]. Reproducibility and model validation have been bolstered by the public release of large, annotated datasets—most prominently the JUMP-CP dataset and the Cell Painting Gallery. These resources offer harmonized image collections and morphological profiles at scale, supporting robust training and external validation of machine learning algorithms [10]. In terms of interpretability, representation learning techniques customized for cell painting have enhanced feature quality for downstream predictive tasks; moreover, explainable-AI (XAI) toolkits, such as single-cell phenotypic parsing and feature-to-function mapping, strengthen interpretability by connecting image features to biological mechanisms [11]. Collectively, these developments not only support the quantitative prediction of phenotypic outcomes induced by drugs and genetic perturbations but also yield actionable insights into pharmacological and molecular mechanisms within an image-based PDD framework [12].

2.3. Drug screening at tissue/organ resolution via simulated organotypic phenotypic systems

Unlike two-dimensional cell cultures, three-dimensional (3D) physiological systems, including organoids, organs-on-a-chip, and other 3D models, faithfully replicate native tissue/organ architecture and function. These advanced models more accurately recapitulate complex physiological or pathological processes, thereby generating higher-value preclinical evidence for drug research [13]. By modeling key pathophysiological processes in disease progression, researchers have developed 3D lung-on-a-chip systems that simulate acute pulmonary edema through blood–air barrier disruption and brain region-specific organoids that replicate the pathology of Alzheimer’s disease via $A\beta$ deposition. Through integrated microfluidics, these systems enable the efficacy evaluation of CHM-derived compounds initially screened at the molecular network and cellular/subcellular levels. This tiered approach significantly reduces failure risks and development costs in subsequent drug R&D pipelines. Additionally, supported by 3D high-content imaging and AI, this phenotypic screening approach retains the biological complexity inherent in organoids while maintaining the experimental controllability of a high-throughput platform. However, the interbatch variability represents an inherent challenge in organoid models. To mitigate its impact and enhance the robustness of simulated organotypic phenotypic systems, it is necessary to implement an automated digital microfluidic platform that contains standardized laboratory procedures and rigorous quality control throughout organoid culture and high-throughput manipulation and analysis [14]. To further reinforce repeatable and reliable data sharing, full traceability and findable, accessible, interoperable, and reusable (FAIR) practices should be adopted in this workflow.

Nevertheless, the systemic complexity of the human body, which includes immune function, neuroendocrine regulation, and multiorgan crosstalk, has been difficult to replicate using organoids or organs-on-a-chip alone for a long time. Consequently, compounds showing efficacy in these models may still fail during subsequent *in vivo* validation or clinical trials. To maximize the reliability of these *in vitro* alternative models, the US National Institutes of Health (NIH) is actively advancing the interconnection of multiorgan chips to create “human-on-a-chip” systems for simulating drug metabolism (e.g., absorption, distribution, metabolism, excretion, and toxicity (ADMET)) processes [15]. Remarkably, up to 18 organ microphysiological systems connected by vascular networks have been created to more accurately mimic the physiological complexity of the human body [16]. Moreover, the integration of organoid technology with organ-on-a-chip platforms has led to the development of “organoids-on-a-chip” systems that model physiological multiorgan interactions at higher levels of organiza-

tion [14]. Thus, these microengineered multiorgan or/and multi-organoid systems could serve as vital and innovative preclinical models that significantly increase the efficiency and predictive accuracy of *in vitro* efficacy and safety assessments for candidate drugs derived from CHMs, although they are currently unable to fully mimic the complex physiological landscape of a body.

2.4. Drug rediscovery via key targets

Once promising compounds or combinations are identified through multilevel phenotypic screening for subsequent *in vivo* evaluation, their molecular targets can be leveraged to perform additional high-throughput virtual screening of compound libraries derived from CHMs or natural products. Integrating this target-based approach with the stepwise phenotypic screening described above enables the discovery of more candidate compounds. Accurate identification of drug targets is a prerequisite for this approach. As noted previously, the screening process can provide initial clues for target identification. However, precise target determination typically requires techniques such as affinity chromatography (e.g., cell membrane chromatography (CMC) and cell membrane affinity chromatography (CMAC)) and affinity MS (e.g., limited proteolysis (LiP)–MS and drug affinity responsive target stability (DARTS)) [17]. Structural biology techniques can be used to elucidate the spatial structures of novel targets and characterize their binding interactions with compounds. Following target identification and AI-assisted screening of target-specific compounds, experimental validation of drug–target binding is still needed. Given that traditional ligand fishing techniques, such as surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR), are less suitable for efficient drug repurposing because of throughput and cost constraints, biological phenotypes arising during drug–target binding can also be leveraged to develop high-throughput drug screening methods. For instance, fluorescence changes resulting from compound binding to kinases have been detected using high-throughput screening to validate inhibitors [18].

Currently, the implementation of the PTDS framework still faces some dilemmas. First, inherent heterogeneity across multimodal data types, such as omic data, high-dimensional imaging, and organotypic models, introduces difficulties in seamless integration of multimodal data (e.g., linking omic network modules to cell painting phenotypes or 3D model functions) using current AI tools. The development of new AI algorithms that can handle multiple data types simultaneously is highly important for bridging the gap in multimodal data integration. Second, while AI drives key steps in this workflow (e.g., DTI prediction and imaging feature compression), bottlenecks in interpretability and validation remain, even with the application of XAI toolkits. For instance, linking AI-derived latent space features to concrete biological mechanisms remains imprecise. Additionally, AI models for CHM DTI prediction rely on limited high-quality target-site component analysis, which limits their generalizability. Third, organoids and organoids-on-a-chip now provide sophisticated alternative multi-organ interaction models but are generally limited by inadequate vascularization. Unlike connected engineered channels, the presence of a functional microvascular network fosters a more physiologically complex microenvironment that supports intricate immune and metabolic processes. This, in turn, significantly expands the functional capacity of the organoid, leading to more faithful simulation of human organs. Additionally, the current framework primarily supports static multilevel data integration and lacks the capacity for dynamic spatiotemporal alignment. This prevents the synchronization of time-resolved data streams derived from longitudinal omics, progressive 3D models, and real-time imaging, which is critical for capturing the biological

relevance of disease progression and CHM therapeutic dynamics. To this end, achieving a cohesive understanding of CHM drug action necessitates the development of advanced tools.

In summary, while AI has accelerated the development of TDD, PDD retains substantial untapped value for CHM-based drug development, as it is not typically applied to decipher the “black boxes” of CHMs or organisms. Critically, PTDS, as an integrated strategy, fully leverages the principles of phenotypic screening in conjunction with TDD technologies. By placing greater emphasis on utilizing disease-relevant phenotypic changes across multiple levels during the discovery process, this strategy employs a stepwise workflow that progressively narrows the scope and even potentially directly identifies the compound basis underlying CHM efficacy. Furthermore, PTDS provides crucial information for subsequent research into the pharmacological mechanisms and targets of candidate drugs, as well as for target-driven drug rediscovery. Importantly, discovered candidate compounds can undergo effective preliminary screening via the aforementioned high-throughput PDD models. Recent breakthroughs in AI algorithm interpretability and organoid model vascularization have facilitated the generation and analysis of more objective phenotypic data. These advances significantly enhance the identification of potential CHM-derived drugs while reducing failure risk during preclinical and clinical research phases. Consequently, PTDS represents a new paradigm for novel drug discovery from CHMs that is characterized by high efficiency and minimal resource expenditure.

CRediT authorship contribution statement

Wei Zhou: Writing – original draft, Funding acquisition, Conceptualization. **Yue Gao:** Writing – review & editing, Supervision, Conceptualization.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2025YFC3507500) and the National Natural Science Foundation of China (82274198).

References

- [1] 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.
- [2] Zhang K, Yang X, Wang Y, Yu Y, Huang N, Li G, et al. Artificial intelligence in drug development. *Nat Med* 2025;31(1):45–59.
- [3] Vincent F, Nueda A, Lee J, Schenone M, Prunotto M, Mercola M. Phenotypic drug discovery: recent successes, lessons learned and new directions. *Nat Rev Drug Discov* 2022;21(12):899–914.
- [4] Jiang W, Ye W, Tan X, Bao YJ. Network-based multi-omics integrative analysis methods in drug discovery: a systematic review. *Biodata Min* 2025;18(1):27.
- [5] Gao R, Yan J, Li P, Chen L. Detecting the critical states during disease development based on temporal network flow entropy. *Brief Bioinform* 2022;23(5):bbac164.
- [6] Xue Z, Shang Y, Yang L, Li T, Yang B. A tailored database combining reference compound-derived metabolite, metabolism platform and chemical characteristic of Chinese herb followed by activity screening: application to *Magnoliae Officinalis Cortex*. *J Pharm Anal* 2025;15(4):101066.
- [7] Liu X, Liu J, Fu B, Chen R, Jiang J, Chen H, et al. DCABM-TCM: a database of constituents absorbed into the blood and metabolites of traditional Chinese medicine. *J Chem Inf Model* 2023;63(15):4948–59.
- [8] Seal S, Trapotsi MA, Spjuth O, Singh S, Carreras-Puigvert J, Greene N, et al. Cell Painting: a decade of discovery and innovation in cellular imaging. *Nat Methods* 2025;22(2):254–68.
- [9] Cimini BA, Chandrasekaran SN, Kost-Alimova M, Miller L, Goodale A, Fritchman B, et al. Optimizing the Cell Painting assay for image-based profiling. *Nat Protoc* 2023;18(7):1981–2013.
- [10] Weisbart E, Kumar A, Arevalo J, Carpenter AE, Cimini BA, Singh S. Cell Painting Gallery: an open resource for image-based profiling. *Nat Methods* 2024;21(10):1775–7.
- [11] Sun DZ, Yang XR, Huang CS, Bai Z, Shen P, Ni Z, et al. CPHNet: a novel pipeline for anti-HAPE drug screening via deep learning-based Cell Painting scoring. *Respir Res* 2025;26(1):91.
- [12] Liu N, Kattan WE, Mead BE, Kummerlowe C, Cheng T, Ingabire S, et al. Scalable, compressed phenotypic screening using pooled perturbations. *Nat Biotechnol* 2025;43:1324–36.
- [13] Zhou L, Chen S, Liu J, Zhou Z, Yan Z, Li C, et al. When artificial intelligence (AI) meets organoids and organs-on-chips (OoCs): game-changer for drug discovery and development? *Innov Life* 2025;3(1):100115.
- [14] Park SE, Georgescu A, Huh D. Organoids-on-a-chip. *Science* 2019;364(6444):960–5.
- [15] Zhou L, Huang J, Li C, Gu Q, Li G, Li ZA, et al. Organoids and organs-on-chips: recent advances, applications in drug development, and regulatory challenges. *Med* 2025;6(4):100667.
- [16] Wang J, Zhang H, Qu Y, Yang Y, Xu S, Ji Z, et al. An eighteen-organ microphysiological system coupling a vascular network and excretion system for drug discovery. *Microsyst Nanoeng* 2025;11(1):89.
- [17] Prudent R, Lemoine H, Walsh J, Roche D. Affinity selection mass spectrometry speeding drug discovery. *Drug Discov Today* 2023;28(11):103760.
- [18] Sun T, Zhao H, Hu L, Shao X, Lu Z, Wang Y, et al. Enhanced optical imaging and fluorescent labeling for visualizing drug molecules within living organisms. *Acta Pharm Sin B* 2024;14(6):2428–46.