Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 18 March 2026
Revised 27 April 2026
Accepted 16 May 2026

Natural Resources for Human Health 2026, Vol. 6 (3): 304–309

KEYWORDS:
Artificial intelligence; Phytochemistry; Natural products; Drug discovery; Machine learning; Deep learning; Molecular docking; Virtual screening; ADMET prediction; Precision medicine.

Integrating Phytochemistry and Artificial Intelligence for Natural Product Drug Discovery

Dr. Jai Vignesh[1], Dr. P. Grace Kanmani Prince[2], Ms Bhavani Pentela[3], Nipun Setia[4], Satish N. Pawar[5], Yogendra Bhati[6], Sitora Pimkulova[7], Rajashri CK[8]

[1] Assistant Professor, Department of Pharmacology, Vinayaka Missions Medical College, Karaikal, Vinayaka Mission's Research Foundation (DU), Tamil Nadu, India. jaivignesh110@gmail.com 0000-0002-9913-4097

[2] Associate Professor, Department of Biomedical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, grace.enc@sathyabama.ac.in, 0000-0002-2595-9635

[3]Assistant Professor, Department of Pharmacology, Noida Institute of Engineering and Technology (Pharmacy Institute), Greater Noida, Uttar Pradesh, India, pentelabhavani.pharmacy@niet.co.in, 0000-0002-4130-070X

[4]Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India. nipun.setia.orp@chitkara.edu.in https://orcid.org/0009-0005-8635-6802

[5]Department of Commerce & Management, Sri Balaji Society's Balaji College of Arts, Commerce & Science, Tathawade, Pune, Maharashtra- 411033, India. satishnpawar2024@gmail.com

[6]Department of Department of Pharmacology, Noida international University, Greater Noida, Uttar Pradesh 203201, India., yogendra.bhati@niims.edu.in

[7]Senior lecturer, Department of Foreign Languages, Tashkent International University of Chemistry,Uzbekistan
email: sitoraprimkulova@gmail.com, 0009-0003-9392-8659

[8]Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, rajashrick@maher.ac.in

ABSTRACT: Natural products continue to be an essential source of bioactive molecules for pharmaceutical development because of their remarkable structural diversity and wide range of biological activities. Traditional phytochemical drug discovery, however, is limited by the time consuming experimental process, the cost and the poor screening efficiency. The combination of phytochemistry and artificial intelligence (AI) has proved to be an effective approach to help identify, characterize and optimize therapeutic compounds in plants. The recent progress in AI-driven methodologies such as machine learning, deep learning, virtual screening, molecular docking, ADMET prediction, and multi-omics integration in the discovery of novel phytochemicals and their pharmacological potential is highlighted in this review. Computational applications are used synergistically to improve target identification, predict bioactivity, optimize for leads, and precision medicine, and lower drug development resources and time. Moreover, the current data quality, model interpretability, experimental validation, and clinical translation issues are discussed, alongside the novel opportunities enabled by explainable AI, generative AI, and autonomous drug discovery platforms. AI's ability to rapidly analyze vast amounts of data and identify novel compounds from natural sources alongside its

ability to optimize their properties in a way that would be difficult for humans to achieve will create a paradigm shift in the study of natural products and in the development of safe, effective, and sustainable therapeutics in the future.

1. INTRODUCTION

Natural products have been a source of therapeutic agents since the dawn of medicine, and provide a structurally diverse pool of molecules that has resulted in the development of numerous clinically important drugs for the treatment of cancer, infectious diseases, metabolic disorders and cardiovascular disease (1). Approximately 50% of the currently approved drugs are natural products or direct derivatives or synthetic molecules modeled after naturally occurring bioactive molecules, proving the vital role that natural products play in the drug discovery process in the 21st century (2). The medicinal plants, microorganisms, marine organisms and fungi are a rich arsenal of secondary metabolites with highly valuable medicinal properties, such as alkaloids, flavonoids, terpenoids, phenolic acids, glycosides and saponins. The science of identification, isolation, structural characterization and biological evaluation of such compounds forms the basis for the discovery of new therapeutic candidates and mechanism of action (phytochemistry). However, traditional natural product research has the drawbacks of slow and tedious extraction processes, complex chemical characterization, substantial need for biological screening, high costs of development and a long process of experimental validation, and only about 15% success rate in the identification of potential lead compounds (4). Recently, the emergence of artificial intelligence (AI) has transformed this landscape, making it possible to leverage computational tools that can quickly process enormous amounts of phytochemical data, forecast biological functions, identify molecular targets, perform virtual screening, gain insights into how the molecule might behave in the body, and optimize lead compounds with unprecedented efficiency and accuracy. With the help of machine learning and deep learning, network pharmacology, molecular docking, and generative AI models, researchers can save valuable time, resources, and money by narrowing down the vast array of natural compounds to those with the greatest potential for further investigation, thereby boosting the chances of a successful drug development process (5). The combination of phytochemistry and AI also makes it easier to integrate multi-omics data, systems biology, and precision medicine approaches, which will help to comprehension of the complex interactions in biological systems and speed the translation of natural products into clinically viable therapeutics. This review aims to provide a comprehensive account of the phytochemical resources, developments in the fields of artificial intelligence and computational methods, and their synergism in natural product drug discovery, alongside highlighting current challenges and future opportunities to harness natural products as medicines, which are safe, effective and sustainable.

2. PHYTOCHEMISTRY AND BIOACTIVE NATURAL PRODUCTS

This study of the chemical components of plants and their biological activities is called phytochemistry and it is the basis for the finding of new drugs from natural sources. The bioactive compounds of plant origin are generally divided into primary and secondary metabolites, the latter of which are the main contributors to the pharmacological activity, such as alkaloids, flavonoids, phenolic acids, terpenoids, tannins, saponins, glycosides, lignans, coumarins, and essential oils (6). These phytochemicals are part of plants' defense and adaptation mechanisms and are present in various medicinal herbs, edible crops, spices, fruits, vegetables, sea herbs, and fungi as well as other natural resources. Curcuma longa, Withania somnifera, Camellia sinensis, Azadirachta indica, Panax ginseng, and Moringa oleifera, among others, are noted for their abundance of bioactive metabolites with significant pharmaceutical potential (6). Efficient extraction and characterization methods like solvent extraction, Soxhlet extraction, ultrasound-assisted extraction, microwave-assisted extraction, supercritical fluid extraction, HPLC, GC–MS, LC–MS/MS, FTIR, and NMR spectroscopy are essential for the successful use of these compounds, which allow the accurate identification, purification, and elucidation of phytochemicals (7). The wide range of pharmacological studies has revealed that plant-derived bioactive compounds have antioxidant, anti-inflammatory, antimicrobial, antiviral, anticancer, antidiabetic, cardioprotective, hepatoprotective, neuroprotective, and immunomodulatory properties by regulating various molecular targets and signaling pathways. They have the potential to simultaneously target oxidative stress, inflammation, apoptosis, metabolic dysfunction, and cellular signaling, which makes them very appealing molecules to use in the treatment of complex chronic diseases. The diversity of their structure, compatibility with the biological system, and relatively safer safety profiles continue to render phytochemicals as valuable lead molecules for modern drug development, nutraceuticals, and precision medicine. Thus, a study of phytochemistry continues to be an essential field for the discovery of new bioactive

molecules and plays a vital role in the use of computational methods and artificial intelligence for natural product-based drug development to yield effective clinically relevant drugs (8).

3. ARTIFICIAL INTELLIGENCE IN NATURAL PRODUCT RESEARCH

Artificial intelligence (AI) has become a game-changer in natural product research, allowing for the swift analysis of vast phytochemical databases and the hastening of drug discovery processes.Rapidly analyzing vast amounts of phytochemical data and speeding up the drug discovery process, artificial intelligence (AI) is emerging as a game-changer in natural product research. Machine learning (ML) algorithms such as supervised, unsupervised, and reinforcement learning models can detect complex relationships between chemical structures and biological activities, which aids in compound classification, target prediction, and quantitative structure–activity relationship (QSAR) modeling (8). Finally, the use of deep learning (DL) with the use of multilayer neural networks further improves the predictive performance by learning complex molecular features in a self-supervised learning manner, which improves the accuracy of bioactivity prediction and compound prioritization (9). AI-enabled virtual screening can be used to quickly screen thousands of natural products against a biological target, thus saving valuable time and money compared to the time-consuming and expensive experimental screening of thousands of compounds. This computational approach is an effective method to screen candidate lead molecules for a drug discovery pipeline and it has become a key tool in the modern drug discovery process (10). Molecular docking is a complementary technique to virtual screening and helps in determining the binding mode, binding energy, stability of the phytochemical and predicting the molecular recognition and mechanism of action in the active site of the target protein. Furthermore, AI-driven absorption, distribution, metabolism, excretion and toxicity (ADMET) prediction models predict the pharmacokinetic features and safety profiles of a drug at early stages of development, helping researchers to weed out any compounds with unfavorable properties before undergoing costly lab or clinical testing. Additionally, AI can aid in lead optimization by proposing structural modifications that boost activity, selectivity, metabolic stability, and bioavailability while reducing toxicity. The implementation of machine learning, deep learning, molecular docking, virtual screening and ADMET prediction has significantly enhanced the efficiency, reliability and success rate of drug discovery from natural products. With the continuous development of computational power, biological databases, and explainable AI models, AI-driven methodologies are likely to become more important in identifying novel phytochemical therapeutics and precision medicine, as well as enabling faster, more accurate, and economical drug development processes, going forward (11).

4. AI-INTEGRATED DRUG DISCOVERY PIPELINE

AI and phytochemistry have built a complete and effective drug discovery pipeline for the systematic identification, evaluation, and optimization of bioactive compounds derived from natural sources, especially plants (11). The AI-integrated workflow starts with phytochemical database mining to identify and gather data from large public databases like NPASS, ChEMBL, PubChem, COCONUT, TCMSP and KNApSAcK, which include information on chemical structure, physicochemical properties, biological activities and ethnopharmacological evidence. Machine learning algorithms can be used to help researchers identify promising compounds for further study by curating the data, annotating compounds, and recognizing patterns, all of which can be achieved quickly. (12) After compound selection, target identification uses AI-based prediction models, network pharmacology, reverse docking and graph neural networks to identify potential protein targets and disease-associated pathways, which enhances the understanding of the molecular mechanisms involved in phytochemical activities (13). Bioactivity prediction is the next step, where machine learning and deep learning algorithms predict the therapeutic value of compounds based on quantitative structure–activity relationships, molecular fingerprints, and other biological descriptors. Such computational models are designed for molecules of high affinity with reduced need for large experimental screening and therefore are less time consuming and expensive to develop (14). The pipeline also includes multi-omics integration, where genomic, transcriptomic, proteomic, metabolomic, and epigenomic data are integrated to yield a systems-level understanding of disease mechanisms and compound–target interactions, as depicted in Figure 1. These combined, diverse data sets are utilized by AI algorithms to uncover biomarkers, signaling pathways, and therapeutic targets, which help direct precision drug development. In the era of precision medicine, genomic profiles, pharmacogenomic data, and clinical features can be used together to guide the use of the most effective natural product to treat each patient, while minimizing the risk of adverse drug reactions (15). Last but not least, computational validation uses molecular docking, molecular dynamics, free-energy calculations, and AI-powered ADMET predictions to validate binding affinity, molecular stability, pharmacokinetic properties, and toxicity prior to lab testing. This computational system is a significant improvement in the efficiency and reliability of lead identification and optimization

and significantly increases the speed with which phytochemicals can be translated into clinically useful drug candidates. The overall AI-powered drug discovery pipeline outlined in Figure 1 illustrates the potential of cutting-edge computational tools to enhance each step of the natural product drug discovery process, paving the way to greater speed, accuracy, and reduced cost for the development of novel and effective therapeutics, and also for progress in precision medicine and sustainable drug development (16)

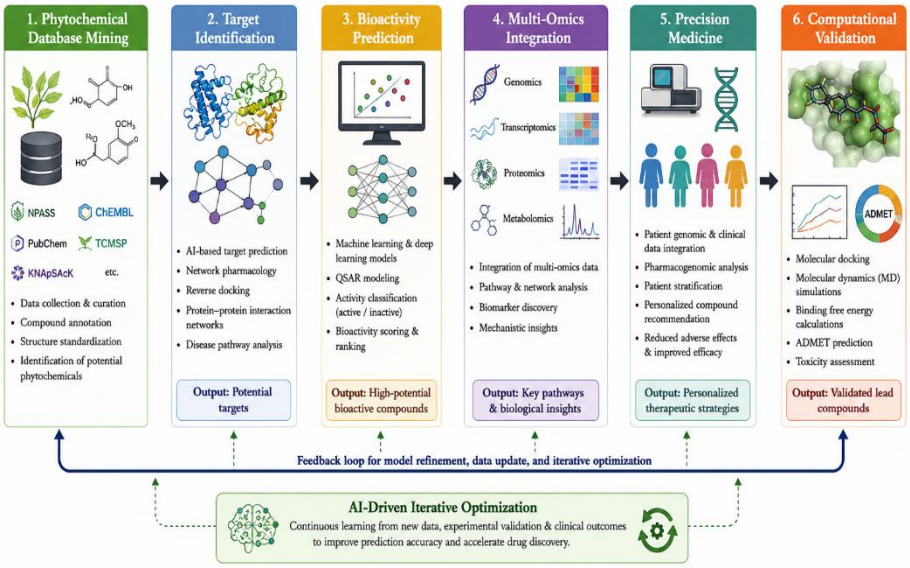

Figure 4. AI-Enabled Drug Discovery Pipeline Incorporating Phytochemical Database Mining, Target Prediction, Multi-Omics, and Molecular Validation.

5. APPLICATIONS, CHALLENGES, AND FUTURE PERSPECTIVES

Artificial intelligence (AI) is rapidly becoming a game-changer in the field of phytochemical drug discovery, revolutionizing the identification, optimization and validation of natural bioactive compounds for various therapeutic purposes. AI-driven computational platforms have been able to take disease-specific drug discovery to the next level by screening phytochemicals against molecular targets linked to various diseases, such as cancer, cardiovascular diseases, diabetes, neurodegenerative diseases, infectious diseases, inflammatory diseases, and others, with swiftness and lower experimental cost, which has led to more efficient identification of leads and reduced experimental effort (17). The combination of machine learning, deep learning, network pharmacology and multi-omics analyses have also enabled the development of personalized therapeutics, with the integration of genomic, transcriptomic, metabolomic and even pharmacogenomic information to recommend natural compounds with improved efficacy and reduced adverse effects for individual patients (18). Even though these significant advances have been made, there are still many challenges to overcome in the field of data integration, such as the diversity of phytochemical databases, the lack of consistency in experimental methods, the lack of biological annotations, and the limited availability of high-quality standardized datasets, that can impact the robustness and reproducibility of the AI models (19). Moreover, the growing sophistication of deep learning algorithms has sparked a need for explainable AI (XAI) methods that offer interpretable and explainable predictions, which can further help researchers and clinicians uncover the molecular mechanisms behind AI-driven decisions and increase the confidence in computational predictions and their regulatory acceptance (20). In addition, detailed experimental validation of AI-predicted phytochemicals is essential to ensure clinical translation, involving molecular, cellular, animal, and human studies, followed by a thorough pharmacokinetic, toxicological and clinical analysis to substantiate safety, efficacy, and therapeutic potential prior to regulatory approval and commercialization (21). To realize the translation from computational predictions to clinical use, the field must enter a new era of interdisciplinary cooperation between the phytochemists, computational biologists, pharmacologists, clinicians and data scientists (22). In the future, the field of natural product research is poised for a transformation due to the integration of generative AI, large language models, digital twins, robotics, automated laboratories, federated learning and quantum computing, which will allow for autonomous drug discovery processes, continual model updates, and extremely efficient lead optimization (23). With the

continued expansion of phytochemical databases and the growing precision medicine ecosystems, these technologies are expected to further expedite the identification of safe, effective, and eco-friendly plant-derived therapeutics and make AI a key driver of the future of global healthcare innovation and pharmaceutical research (24).

6. CONCLUSION

The synergy of phytochemistry and artificial intelligence revolutionizes drug discovery from natural sources by leveraging the diversity of bioactives found in plants and the power of computation. Machine learning, deep learning, virtual screening, molecular docking, and ADMET prediction are just a few examples of AI technologies that have greatly enhanced the efficiency of therapeutic candidate identification, prioritization, and optimization, while also decreasing time and development costs. This interdisciplinary approach will improve the comprehension of complex biological mechanisms, aid in precision medicine and speed up the translation of phytochemicals to clinically relevant drugs. However, with the rapid progress of AI algorithms, multi-omics integration, and explainable AI, these limitations are anticipated to be addressed and overcome. Standardized databases, strong clinical validation and academic-industry collaboration should be focused in future studies. These advances will enable sustainable pharmaceutical innovation, support industrial-scale natural product research, and speed the development of safe, effective and personalized therapeutics for global healthcare.

DECLARATIONS

Funding: No specific funding was received for this work.

Conflict of Interest: The authors declare no conflict of interest.

Author Contributions: All authors contributed equally to the conception, writing, revision, and approval of the manuscript.

Data Availability: No new data were generated or analyzed in this study.

Ethical Approval: Not applicable, as this study is based solely on published literature.

REFERENCES

[1] Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 1981 to 2019. J Nat Prod. 2020;83(3):770–803. doi:10.1021/acs.jnatprod.9b01285.
[2] Atanasov AG, Zotchev SB, Dirsch VM, et al. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 2021;20(3):200–216. doi:10.1038/s41573-020-00114-z.
[3] Harborne JB. Phytochemical Methods: A Guide to Modern Techniques of Plant Analysis. 3rd ed. London: Chapman & Hall; 1998.
[4] Dewick PM. Medicinal Natural Products: A Biosynthetic Approach. 3rd ed. Chichester: Wiley; 2009.
[5] Wink M. Modes of action of herbal medicines and plant secondary metabolites. Medicines. 2015;2(3):251–286. doi:10.3390/medicines2030251.
[6] Seca AML, Pinto DCGA. Plant secondary metabolites as anticancer agents: successes in clinical trials and therapeutic application. Int J Mol Sci. 2018;19(1):263. doi:10.3390/ijms19010263.
[7] Azmir J, Zaidul ISM, Rahman MM, et al. Techniques for extraction of bioactive compounds from plant materials: A review. J Food Eng. 2013;117(4):426–436. doi:10.1016/j.jfoodeng.2013.01.014.
[8] Chen H, Engkvist O, Wang Y, Oliveira I, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018;23(6):1241–1250. doi:10.1016/j.drudis.2018.01.039.
[9] Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463–477. doi:10.1038/s41573-019-0024-5.
[10] Ekins S, Puhl AC, Zorn KM, et al. Exploiting machine learning for end-to-end drug discovery and development. Nat Mater. 2019;18(5):435–441. doi:10.1038/s41563-019-0338-z.
[11] Schneider G. Automating drug discovery. Nat Rev Drug Discov. 2018;17(2):97–113. doi:10.1038/nrd.2017.232.
[12] Sorokina M, Merseburger P, Rajan K, Yirik MA, Steinbeck C. COCONUT online: Collection of Open Natural Products database. J Cheminform. 2021;13:2. doi:10.1186/s13321-020-00478-9.
[13] Zhou Y, Zhang Y, Lian X, et al. Therapeutic target database update 2022: facilitating drug discovery. Nucleic Acids Res. 2022;50(D1):D1398–D1407. doi:10.1093/nar/gkab953.

[14] Stokes JM, Yang K, Swanson K, et al. A deep learning approach to antibiotic discovery. Cell. 2020;180(4):688–702.e13. doi:10.1016/j.cell.2020.01.021.
[15] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56. doi:10.1038/s41591-018-0300-7.
[16] Paul D, Sanap G, Shenoy S, et al. Artificial intelligence in drug discovery and development. Drug Discov Today. 2021;26(1):80–93. doi:10.1016/j.drudis.2020.10.010.
[17] Zhavoronkov A. Artificial intelligence for drug discovery, biomarker development, and generation of novel chemistry. Mol Pharm. 2018;15(10):4311–4313. doi:10.1021/acs.molpharmaceut.8b00930.
[18] Bashyam V, Vargheese R, Hamide A. Artificial intelligence in precision medicine. J Assoc Physicians India. 2021;69(11):11–12.
[19] Walters WP, Murcko MA. Assessing the impact of generative AI on medicinal chemistry. Nat Biotechnol. 2020;38(2):143–145. doi:10.1038/s41587-020-0418-2.
[20] Holzinger A, Carrington A, Müller H. Measuring the quality of explanations: The system causability scale (SCS). KI-Künstliche Intelligenz. 2020;34:193–198. doi:10.1007/s13218-020-00636-z.
[21] Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. Drug Discov Today. 2019;24(3):773–780. doi:10.1016/j.drudis.2018.11.014.
[22] Paul SM, Mytelka DS, Dunwiddie CT, et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nat Rev Drug Discov. 2010;9(3):203–214. doi:10.1038/nrd3078.
[23] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–589. doi:10.1038/s41586-021-03819-2.
[24] Brown N, Ertl P, Lewis R, et al. Artificial intelligence in chemistry and drug design. J Comput Aided Mol Des. 2020;34(7):709–715. doi:10.1007/s10822-020-00317-4.