https://doi.org/10.33472/AFJBS.6.Si3.2024.2509-2526

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Artificial Intelligence-Guided Discovery of Anticancer Compounds from Marine Plants: A Focus on Sargassum SPP

Ashu Mittal[1], Deepa Bajetha[2], Mekhala[3], Anil Kumar[4], Tanmay Ghosh[5], Surbhi Kamboj[6], Mathews T Thelly[7*]

[1]Professor of Pharmaceutics, KIET School.of Pharmacy (KIET Group of institutions), Muradnagar, Delhi NCR 201206
[2]Associate Professor, IIMT College of Pharmacy, Greater NOIDA
[3]Assistant Professor, G.L. Bajaj Institute of Management and Technology
[4]Head & Assistant Professor, Department of Chemistry (PG), Sahibganj College Sahibganj, Jharkhand, India
[5]Assistant Professor, Department of Microbiology, Dinabandhu Andrews College, Baishnabghata, South 24 Parganas, Kolkata, West Bengal, India
[6]KIET School of Pharmacy, KIET Group of Institutions, Ghaziabad
[7*]Associate Professor, Head and Research Guide, Department of Botany, Kuriakose Elias College Mannanam, Kottayam, Kerala, India

Article Info

Volume 6, Issue Si3, June 2024

Received: 27 April 2024

Accepted: 03 June 2024

Published: 29 June 2024

doi: 10.33472/AFJBS.6.Si3.2024.2509-2526

ABSTRACT:

The exploration of marine plants as sources of novel anticancer compounds has gained significant attention due to their diverse biochemical compositions and unique ecological adaptations. Among these marine organisms, Sargassum spp. stands out for its rich bioactive profile and potential therapeutic applications in oncology. This abstract provides an overview of the recent advancements in Artificial Intelligence (AI)-guided discovery of anticancer compounds from Sargassum spp., emphasizing the integration of computational approaches to accelerate drug discovery processes. Sargassum spp., a genus of brown algae widely distributed in coastal regions and oceanic currents, has been traditionally used in folk medicine and increasingly studied for its pharmacological properties. AI technologies, including machine learning and molecular docking simulations, have revolutionized the identification of bioactive compounds from natural sources. By leveraging large-scale datasets from chemical databases, bioactivity assays, and scientific literature, AI facilitates the efficient screening and prediction of potential anticancer agents within Sargassum spp. The AI-guided discovery process begins with data collection and preprocessing, where diverse chemical structures and biological activities of Sargassum spp. compounds are curated and standardized. Predictive models are developed using quantitative structure-activity relationships (QSAR) and molecular docking studies, leveraging machine learning algorithms. These models evaluate how compounds interact with cancer-specific targets, forecast their therapeutic effectiveness, and enhance their drug-like properties.

Keywords: Artificial intelligence, drug discovery, anticancer compounds, marine plants, Sargassum spp.

1. Introduction

Many health challenges continue to be daunting worldwide, such as uncontrolled cell growth and the risk of metastasis. Despite major strides in medical science, cancer remains a leading cause of death and illness globally [1,2]. Traditional treatments such as surgery, chemotherapy, and radiation therapy, while effective to some extent, are often accompanied by severe side effects and the risk of resistance development in cancer cells[3]. Thus, there is a critical need for new anticancer compounds that can offer more effective, targeted, and less toxic treatment options. The complexity and heterogeneity of cancer necessitate a multifaceted approach to drug discovery. Novel therapeutic agents must be capable of targeting specific cancer pathways and mechanisms without causing significant damage to normal cells[4]. This ongoing quest for new anticancer compounds drives researchers to explore diverse and unconventional sources, including marine ecosystems, which have proven to be a treasure trove of bioactive substances. Marine plants, particularly algae, have gained significant attention in the field of drug discovery due to their unique chemical diversity and the presence of novel bioactive compounds[5]. Marine environments, with their distinct ecological niches, offer a rich source of secondary metabolites that are often absent in terrestrial plants. These metabolites have evolved as part of

the plants' defense mechanisms against predators, pathogens, and environmental stressors, making them potent candidates for pharmaceutical applications[6,7].

Sargassum spp., a genus of brown macroalgae, is one such marine plant that has shown great promise in the discovery of anticancer compounds. Found predominantly in warm waters, Sargassum spp. are well-known for their rich chemical composition, including polysaccharides, phenolic compounds, flavonoids, terpenoids, and steroids[8]. These compounds exhibit a wide range of biological activities, including antioxidant, anti-inflammatory, antiviral, and notably, anticancer properties. The significance of Sargassum spp. in drug discovery is underscored by numerous studies demonstrating its potential to inhibit cancer cell proliferation, induce apoptosis, and prevent metastasis. For instance, fucoidan, a sulfated polysaccharide extracted from Sargassum spp., has been extensively studied for its anticancer activity against various cancer cell lines[9,10]. The diverse bioactive profile of Sargassum spp. makes it an invaluable resource in the ongoing search for new and effective anticancer agents. Artificial Intelligence (AI) has revolutionized numerous fields, and pharmaceutical research is no exception. AI technologies encompass a broad spectrum of computational techniques and tools, including machine learning (ML), deep learning (DL), natural language processing (NLP), and data mining[11]. These technologies enable the analysis and interpretation of large datasets, uncovering patterns and insights that are often beyond human capability. As part of the drug discovery process, AI is used in identifying and validating targets, as well as screening and optimizing compounds. To predict the activity and toxicity of potential drug candidates, machine learning algorithms can analyze large amounts of biological and chemical data[12]. The application of deep learning in image analysis and molecular structure analysis comes from its ability to recognize complex patterns in data[13]. The integration of AI in the discovery of anticancer compounds offers several significant benefits. Firstly, AI accelerates the drug discovery process by automating data analysis and reducing the time required for identifying potential drug candidates[14]. Traditional methods of drug discovery are often time-consuming and resource-intensive, whereas artificial intelligence can swiftly screen millions of compounds, identifying those most likely to be effective. Additionally, AI in drug discovery can predict biological activity and side effects with impressive precision using existing data. AI models learn from this data, enhancing accuracy and efficiency. This predictive capability minimizes the risk of failure in subsequent stages of drug development, thereby increasing the overall success rate[15,16]. Thirdly, AI facilitates the discovery of novel compounds that might be overlooked using conventional methods[17]. Advanced algorithms can identify unique molecular features and interactions that contribute to anticancer activity, leading to the identification of unconventional and innovative therapeutic agents. For example, AI has been instrumental in identifying marine-derived compounds with potent anticancer properties, including those from Sargassum spp[18].

Marine Plants and Anticancer Compounds

Importance of Marine Biodiversity in Pharmaceutical Research

Marine biodiversity is a vital asset in pharmaceutical research, providing a vast reservoir of bioactive compounds that hold potential for developing new medicines. The ocean, which covers more than 70% of the Earth’s surface, is home to a diverse array of organisms, many of which have evolved unique chemical defenses and metabolic pathways due to the extreme and varied conditions of their habitats[19]. These unique compounds, often not found in terrestrial environments, can offer novel mechanisms of action and structural frameworks for drug development. Marine plants, particularly algae, are significant for their production of various secondary metabolites with diverse biological properties, such as antimicrobial, antiviral, antifungal, and anticancer effects. The structural diversity and complexity of these metabolites provide numerous therapeutic opportunities. Preclinical studies have shown the bioactivity of

compounds derived from marine algae, including fucoidans, phlorotannins, and carotenoids[20,21]. The exploration of marine biodiversity has already led to the discovery of several clinically useful drugs. For example, cytarabine, an anticancer drug, and ziconotide, a pain management drug, are both derived from marine organisms. These successes underscore the potential of marine ecosystems as a critical resource for pharmaceutical innovation, encouraging ongoing and future research into marine-derived compounds for therapeutic applications[22].

Historical Use of Marine Plants in Traditional Medicine

The use of marine plants in traditional medicine dates back centuries, with various cultures recognizing their therapeutic potential long before the advent of modern scientific methods. Coastal communities across the world have utilized marine plants for their medicinal properties, incorporating them into remedies for a variety of ailments[23]. In traditional Chinese medicine (TCM), marine algae have been used to treat conditions such as goiter, edema, and respiratory issues. Similarly, in Japan, seaweeds like kombu and wakame are integral to the diet and are believed to confer health benefits, including improved digestion and cardiovascular health[24,25]. Indigenous cultures in the Pacific Islands and along the coasts of Europe have also harnessed the medicinal properties of marine plants for wound healing, anti-inflammatory purposes, and as nutritional supplements. These historical uses are supported by contemporary research, which has validated many of the traditional claims and uncovered new potential applications[26]. For instance, the antimicrobial properties of certain seaweeds align with their traditional use in treating infections. This blending of historical knowledge with modern science enhances the understanding and utilization of marine plants in contemporary medicine, paving the way for innovative therapeutic approaches[27].

Sargassum spp.

Sargassum spp. is a genus of brown macroalgae found predominantly in tropical and subtropical oceans[28]. These seaweeds are characterized by their distinctive air bladders, which help them float, and their leafy structures, which provide a habitat for various marine organisms. Sargassum species can be found in both coastal regions and open oceans, with some species forming large floating mats, such as those seen in the Sargasso Sea[29,3]. The distribution of Sargassum is widespread, with notable concentrations in the Atlantic Ocean, particularly in the Gulf of Mexico and the Caribbean Sea[5]. These floating mats are dynamic ecosystems, supporting a diverse range of marine life and playing a crucial role in marine ecology. The adaptability of Sargassum to different environmental conditions contributes to its extensive distribution and ecological importance[30,7,21].

Traditional and Contemporary Uses of Sargassum Spp. In Medicine

Traditionally, Sargassum spp. have been used in various forms of medicine across different cultures. In TCM, Sargassum is known as "hai zao" and is used to treat thyroid conditions, particularly goiter, due to its high iodine content. It is also employed in the treatment of edema, testicular swelling, and other inflammatory conditions[31]. The use of Sargassum in traditional remedies highlights its perceived therapeutic value and historical significance. In contemporary medicine, research has focused on the bioactive compounds found in Sargassum spp. and their potential health benefits[15]. Studies have shown that Sargassum extracts possess a range of pharmacological activities, including antioxidant, anti-inflammatory, antiviral, and anticancer properties. These findings have sparked interest in Sargassum as a source of novel therapeutic agents, particularly for the treatment of cancer[32,33].

Overview of known bioactive compounds in Sargassum spp.

Sargassum spp. are rich in a variety of bioactive compounds, contributing to their therapeutic potential. Some of the key bioactive compounds include:

- **Fucoidans**: Sulfated polysaccharides that exhibit significant anticancer activity. Fucoidans have been shown to induce apoptosis in cancer cells, inhibit tumor growth, and enhance immune responses[22].
- **Phlorotannins**: Polyphenolic compounds with strong antioxidant properties. Phlorotannins can scavenge free radicals and protect against oxidative stress, which is linked to cancer progression.
- **Terpenoids**: Known as terpenoids, these compounds exhibit a variety of biological effects, including anti-inflammatory and anticancer properties. In cancer cell studies, terpenoids from Sargassum have shown to inhibit cell proliferation and arrest cell cycle progression[19].
- **Steroids**: Compounds that can modulate immune responses and exhibit anticancer properties. Steroidal compounds from Sargassum have been investigated for their potential to suppress tumor growth[19].
- **Carotenoids**: Pigments with antioxidant and anticancer activities. Carotenoids can protect cells from oxidative damage and modulate signaling pathways involved in cancer development[21].

The rich chemical diversity of Sargassum spp. underscores their potential as a source of novel anticancer compounds. Ongoing research aims to isolate and characterize these bioactive substances, further exploring their mechanisms of action and therapeutic applications[15]. This growing body of evidence highlights the importance of Sargassum spp. in the quest for new anticancer drugs and reinforces the value of marine plants in pharmaceutical research[31].

AI Technologies in Drug Discovery

Machine learning, deep learning, and other relevant AI methodologies

Artificial intelligence (AI) enables computers to perform tasks typically done by humans in various ways. In drug discovery, AI primarily employs machine learning (ML) and deep learning (DL) methodologies. These techniques have revolutionized the field by simplifying the analysis and interpretation of complex biological data[32].

Machine Learning (ML): Machine learning (ML) identifies patterns, makes predictions, and makes decisions without explicit programming by training algorithms on large datasets[17]. In drug discovery, ML helps identify potential drug targets, optimize drug formulations, and predict the biological activity of compounds. Various algorithms, including random forests, support vector machines (SVM), and k-nearest neighbors (k-NN), are used in this process. These algorithms enable scientists to uncover correlations, such as those between chemical structures and biological assays, that traditional analysis methods might miss[33].

Deep Learning (DL): A subset of ML, DL employs neural networks with multiple layers (hence "deep") to model complex relationships within data. DL excels at tasks involving large, unstructured datasets, such as images, sequences, and text[34]. In drug discovery, DL can analyze high-throughput screening data, model protein-ligand interactions, and predict molecular properties. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are among the DL architectures frequently utilized. CNNs, for example, are particularly effective in image-based drug screening, while RNNs can be used for sequence data, such as predicting the activity of peptides[19,4].

Other Relevant AI Methodologies: Beyond ML and DL, other AI methodologies also play crucial roles in drug discovery. Natural language processing (NLP) helps in mining scientific literature and patents for relevant information[21]. Reinforcement learning (RL) is used to optimize drug design by iteratively testing and refining compounds. Transfer learning, which involves using pre-trained models on new tasks, helps in leveraging existing knowledge to accelerate the drug discovery process[35].

Tools and platforms commonly used in AI-guided drug discovery

Several tools and platforms have been developed to facilitate the application of AI in drug discovery, integrating various AI methodologies to streamline the process from data acquisition to compound validation[36]. One notable tool is DeepChem, an open-source library specifically designed for applying deep learning (DL) to drug discovery[3,9]. DeepChem provides a comprehensive suite of tools for tasks such as molecular modeling, data processing, and bioactivity prediction. It supports various neural network architectures, making it a popular choice for developing predictive models in the pharmaceutical industry[37].

Another significant platform is Atomwise, which utilizes DL to predict the binding affinity of small molecules to target proteins[32]. Atomwise's AtomNet platform employs convolutional neural networks (CNNs) to analyze 3D representations of molecules, aiding in the virtual screening of large compound libraries. This technology has been instrumental in identifying potential drug candidates for various diseases, including cancer, showcasing its utility in accelerating the drug discovery process[38,19].

Schrödinger's Maestro platform combines AI with molecular modeling to enhance drug discovery workflows[18]. It integrates machine learning (ML) models with traditional computational chemistry methods, such as molecular dynamics and quantum mechanics, providing accurate predictions of molecular properties and interactions. This integration allows researchers to optimize compounds more efficiently and predict their behavior in biological systems with higher precision[39]. Insilico Medicine is another leader in the field, using AI to accelerate every stage of the drug discovery process. Its platforms, such as Chemistry42 and PandaOmics, apply DL and reinforcement learning (RL) to tasks ranging from target identification to de novo drug design[40]. Insilico Medicine has been successful in generating novel compounds with high potential for therapeutic applications, demonstrating the power of AI in innovating and speeding up drug development[1,21].

IBM Watson for Drug Discovery leverages natural language processing (NLP) and ML to analyze scientific literature and biomedical data, uncovering new insights and identifying potential drug targets[29,5]. This AI-powered platform helps researchers prioritize compounds and understand mechanisms of action, thereby accelerating the drug development pipeline. By sifting through vast amounts of data, Watson for Drug Discovery can highlight promising areas for research that might otherwise be overlooked[41]. These tools and platforms exemplify how AI technologies are integrated into the drug discovery process, enhancing efficiency and accuracy while reducing time and costs[18]. By leveraging AI, researchers can more effectively identify bioactive compounds, predict their properties, and optimize their therapeutic potential, ultimately leading to faster and more precise drug development[42].

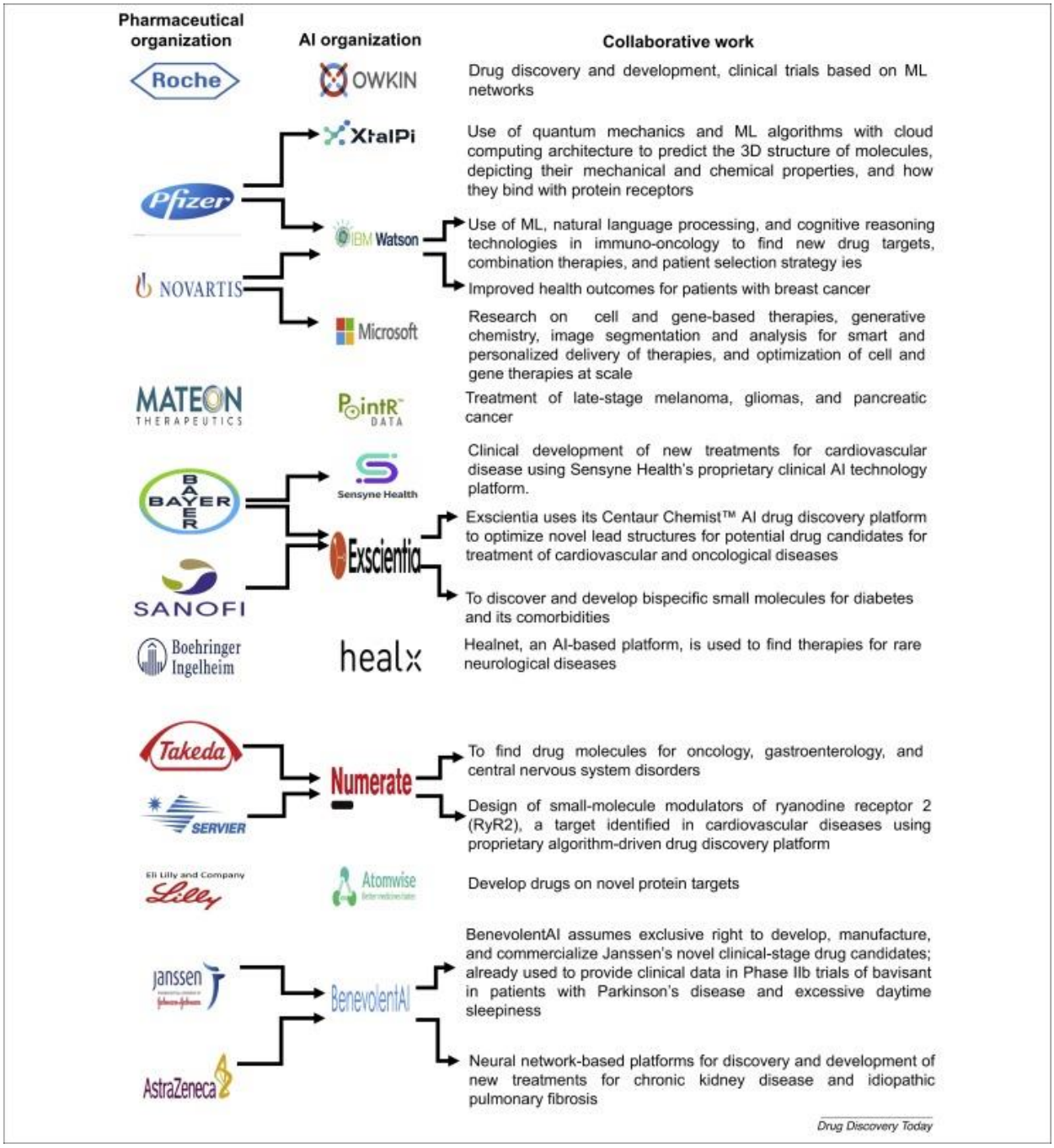

Fig.1 Leading pharmaceutical companies and their association with Artificial Intelligence (AI)

Applications of AI in Drug Discovery

Specific techniques used for predicting anticancer properties

AI techniques for predicting the anticancer properties of compounds involve several specialized approaches that enhance the efficiency and accuracy of drug discovery. One such approach is virtual screening, which employs machine learning (ML) and deep learning (DL) models to predict the activity of compounds against cancer targets[43]. By evaluating large chemical libraries, AI models can identify compounds with potential anticancer activity. These models are trained on datasets of known active and inactive compounds, enabling them to distinguish promising candidates effectively[44].

Quantitative Structure-Activity Relationship (QSAR) models are another critical technique in AI-guided drug discovery. QSAR models correlate chemical structure with biological activity, using ML algorithms like random forests and support vector machines to predict the anticancer activity of new compounds based on their molecular descriptors[45]. This approach helps prioritize compounds for experimental validation, thereby streamlining the drug development process[46].

AI-driven molecular docking and simulation techniques predict how small molecules interact with target proteins. DL models, such as convolutional neural networks (CNNs), analyze the 3D structures of proteins and ligands to predict binding affinities[3,9]. When coupled with molecular dynamics simulations, these models provide insights into the stability and dynamics of protein-ligand interactions, aiding in the identification of potent anticancer compounds. This method allows researchers to understand better and optimize interactions critical for drug efficacy[47]. Omics data integration is another powerful AI application in drug discovery. AI models integrate multi-omics data, including genomics, transcriptomics, proteomics, and metabolomics, to identify biomarkers and therapeutic targets[48]. By analyzing cancer-specific omics profiles, AI can predict which compounds are likely to be effective against particular cancer subtypes[22]. This personalized approach enhances the precision of anticancer drug discovery, ensuring that treatments are tailored to the unique molecular characteristics of different cancer types[49]. Ensuring the safety of anticancer compounds is crucial, and AI models play a vital role in predictive toxicology. These models analyze chemical structures and biological data to predict potential toxicity, helping identify compounds with favorable safety profiles. This predictive capability reduces the risk of adverse effects in clinical trials, making the drug development process safer and more efficient[50]. AI technologies and methodologies are transforming drug discovery by enhancing the efficiency and accuracy of identifying bioactive compounds[52]. From virtual screening to predictive toxicology, AI applications in drug discovery are paving the way for the development of new and effective anticancer therapies. The integration of AI-driven tools and platforms continues to accelerate the drug discovery process, offering hope for faster and more targeted treatment options for cancer and other diseases[53].

AI-Guided Discovery of Anticancer Compounds from Sargassum spp.

1. Data Collection and Preprocessing

Sources of Data for AI Models

AI-guided discovery of anticancer compounds from Sargassum spp. relies heavily on extensive and high-quality data, with the primary sources of data for AI models including chemical databases, bioactivity assays, and scientific literature[54]. Chemical databases serve as repositories of chemical information, encompassing molecular structures, properties, and activities. Key databases in this context include PubChem, ChEMBL, and ZINC[55]. PubChem provides a vast collection of information on chemical compounds, including their bioactivity data, while ChEMBL contains data on bioactive molecules with drug-like properties[56]. ZINC offers a comprehensive catalog of commercially available compounds for virtual screening, making it a valuable resource for identifying potential anticancer agents. Bioactivity assays provide experimental data on the biological activity of compounds, which is crucial for understanding their potential efficacy and toxicity[57]. High-throughput screening (HTS) assays are particularly valuable in this regard, as they test large libraries of compounds against various biological targets. The data generated from these assays can be used to train AI models to recognize patterns associated with anticancer activity, enhancing the models' ability to predict the efficacy of new compounds[58]. Scientific literature, including published research articles and patents, is another rich source of information on the bioactivity of compounds

derived from marine plants, including Sargassum spp. Natural language processing (NLP) techniques can be employed to mine this literature for relevant data, such as previously identified bioactive compounds and their mechanisms of action[23]. By leveraging these diverse sources of data, AI models can be developed and refined to identify promising anticancer compounds with greater accuracy and efficiency[59].

Methods for Preprocessing and Curating Data for AI Analysis

Preprocessing and curating data are critical steps to ensure the accuracy and effectiveness of AI models in the discovery of anticancer compounds. This process involves several stages, each aimed at refining and optimizing the dataset for reliable analysis[12]. The first stage, data cleaning, addresses the common issues found in raw data, such as errors, missing values, and inconsistencies[60]. To produce a reliable dataset, chemical structures may need to be standardized to a common format, duplicate entries must be removed, and any erroneous data should be corrected. This step ensures that the data used for AI model training is accurate and consistent[61].

After cleansing, data normalization is performed to ensure consistency across the dataset. Molecular descriptors, such as molecular weights and hydrophobicities, might be scaled to between zero and one[62]. Normalization is crucial for machine learning algorithms to perform well, as it ensures all features contribute equally to the training process and prevents any single feature from disproportionately affecting the model results. Feature selection is another important step in data preprocessing[63]. Not every feature is equally informative for predicting anticancer activity. Feature selection identifies and retains the most relevant features that significantly contribute to the prediction task. Techniques like principal component analysis (PCA) and recursive feature elimination (RFE) are commonly used to reduce the dataset's dimensionality[64]. These methods help to focus the model on the most informative variables, improving its predictive accuracy and efficiency while discarding irrelevant or redundant data[65]. To address the issue of limited data, especially in the context of rare compounds from marine plants, data augmentation techniques can be employed. Data augmentation involves generating synthetic data points to expand the dataset[32]. Methods such as the Synthetic Minority Over-sampling Technique (SMOTE) create additional synthetic examples by interpolating between existing data points, thus balancing the dataset[18]. Additionally, generative adversarial networks (GANs) can be used to create new molecular structures based on existing data, further enhancing the dataset's diversity and comprehensiveness[66]. The final stage in data preprocessing is splitting the data for model training and evaluation.

Typically, the dataset is divided into three subsets: training, validation, and test sets. The training set is used to train the AI model, enabling it to learn the underlying patterns in the data[31]. The validation set is used to tune hyperparameters and prevent overfitting, ensuring that the model generalizes well to new data. Finally, the test set assesses the model's performance on unseen data, providing an unbiased evaluation of its predictive capabilities[28]. This systematic approach to data preprocessing and curation ensures that the AI models developed are robust, accurate, and effective in identifying potential anticancer compounds from Sargassum spp.

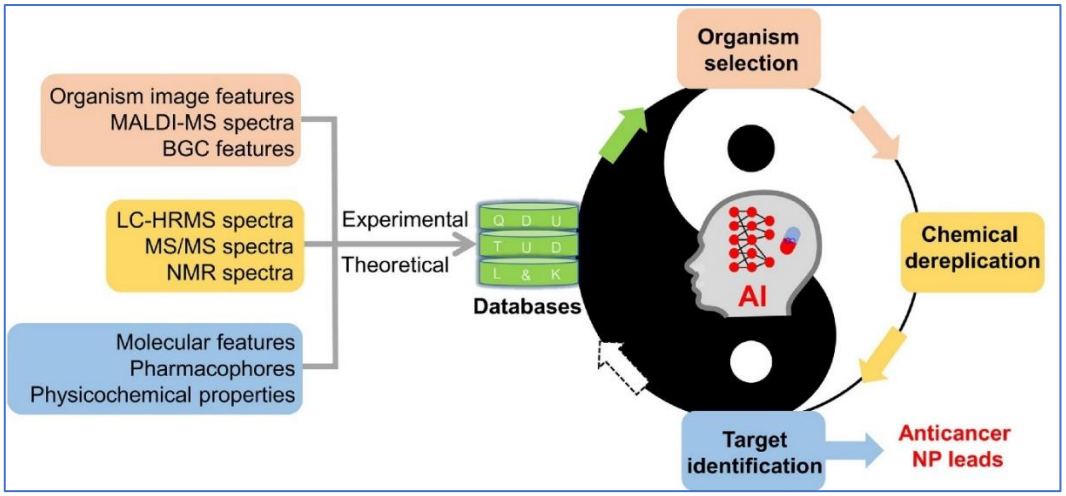

Fig.2 Artificial intelligence guided discovery of anticancer lead compound

AI Model Development

selection of AI Algorithms and Their Rationale

It is possible to handle different types of data and analytical requirements using various machine learning and deep learning algorithms. A number of machine learning algorithms are commonly used to model quantitative structure-activity relationships (QSARs) such as random forests, support vector machines (SVMs), and gradient boosting [67]. Using these algorithms, chemical structures and biological activities can be modeled in complex ways thanks to their ability to handle structured data. As a result of their robustness and capability to handle large datasets, they can be used to identify patterns and predict the efficacy of new compounds[68]. Unstructured data, such as images and sequences, can be easily analyzed with deep learning algorithms, such as convolutional neural networks and recurrent neural networks. CNNs are especially useful in molecular docking studies, where the 3D structures of molecules are analyzed to predict binding affinities[69]. This capability is crucial for understanding how compounds interact with target proteins at a molecular level. Autoencoders and generative adversarial networks (GANs) also play a significant role in generating new compounds and optimizing molecular properties, thereby expanding the potential chemical space for discovering novel anticancer agents[70]. Natural language processing (NLP) techniques, including models like BERT and GPT, are employed to mine scientific literature and patents for relevant information[18]. These NLP models are designed to extract valuable insights from textual data, such as known interactions between compounds and cancer targets. By processing and interpreting vast amounts of textual information, NLP models help identify promising compounds and elucidate their mechanisms of action, thereby guiding experimental efforts in a more targeted manner[71].

Training, Validation, and Testing of AI Models

In order to develop effective AI models for drug discovery, a rigorous training, validation, and testing process is required. A training dataset is used to train the chosen AI model. In order to minimize the prediction error, the model parameters are adjusted [72,5]. To ensure the model generalizes well to new data, cross-validation is used. In cross-validation, data are partitioned into subsets, the model is trained on some of these subsets while it is validated on others, and this process is repeated repeatedly to ensure robust performance across different data splits. Hyperparameters are tuned during validation and overfitting is prevented. Overfitting occurs when the model performs well on training data but poorly on unknown data. Regularization techniques in neural networks, such as L1/L2 regularization and dropout, are used to reduce overfitting. By penalizing large coefficients, L1/L2 regularization encourages simpler models,

while dropout eliminates neurons from training randomly to enhance generalization [73]. Test sets represent unseen data that were not used during training and validation and are used for final evaluation of AI models. An unbiased evaluation of the model's performance is provided by this assessment, which examines key metrics such as accuracy, precision, recall, F1 score, and area under receiver operating characteristic curves. The accuracy of a model determines whether or not it is accurate in its predictions, whereas its precision and recall tell us how accurate it is in predicting true positives and false positives[18,9]. An AUC-ROC curve shows the model's discriminatory ability at various threshold settings when comparing true positive rate and false positive rate. F1 scores provide a balanced measure of the model's prediction capability[74].

Identification and Screening of Compounds

Process of Virtual Screening and Identification of Potential Anticancer Compounds

Virtual screening is a pivotal process in AI-guided drug discovery, leveraging AI models to evaluate extensive libraries of compounds and pinpoint those with potential anticancer activity. The process begins with library preparation, where a virtual library of compounds is assembled[33]. These compounds are often sourced from chemical databases like ZINC, which provides a comprehensive collection of commercially available compounds. Additionally, novel structures can be generated in silico through AI techniques, expanding the library with potential new drug candidates[54]. This combination of known and novel compounds ensures a broad scope for identifying promising anticancer agents[75]. The next step is target selection, which involves identifying biological targets such as specific cancer-related proteins for screening. These targets are chosen based on their relevance to cancer pathways and their potential as drug targets, ensuring that the screening process is focused on the most promising areas for therapeutic intervention[22]. Docking simulations are then conducted using AI-driven molecular docking techniques to predict the binding affinities of compounds to the selected targets. Deep learning models, particularly convolutional neural networks (CNNs), analyze the 3D structures of the compounds and targets to identify potential interactions[29]. Compounds that demonstrate high predicted binding affinities are prioritized for further evaluation, as these interactions suggest a stronger potential for therapeutic efficacy. The identified compounds are then subjected to scoring and ranking based on their predicted binding affinities and other crucial properties, such as drug-likeness and toxicity[30]. The compounds are scored to reflect their overall potential as drug candidates, and the highest-ranking compounds are selected for subsequent experimental validation. This prioritization ensures that resources are focused on the most promising candidates, streamlining the development process[76].

Validation Techniques for the Identified Compounds

Following the virtual screening, potential anticancer compounds undergo rigorous validation to confirm their efficacy and safety. The initial validation is conducted through in vitro studies, where the compounds are tested on cancer cell lines. These assays assess various parameters, including the compounds' cytotoxicity, ability to induce apoptosis, and effects on cell proliferation[74,34]. High-throughput screening techniques facilitate the rapid evaluation of a large number of compounds, allowing researchers to identify those with significant anticancer activity. Compounds that show promise in vitro are further evaluated through in vivo studies using animal models[29]. These studies provide a more comprehensive assessment of the compounds' pharmacokinetics, bioavailability, and therapeutic efficacy within a complex biological system[2,8]. In vivo studies also offer valuable insights into the potential toxicity and side effects of the compounds, which are critical for determining their suitability for further development[18,32]. To gain a deeper understanding of the compounds' mechanisms of action, detailed mechanistic studies are conducted. These studies investigate how the compounds

interact with their targets and influence cancer pathways[77]. Techniques such as RNA sequencing and proteomics are employed to examine changes in gene and protein expression in response to the treatment. This information helps elucidate the biological processes affected by the compounds, providing a clearer picture of their therapeutic potential[29]. The most promising compounds advance to clinical trials, where they are tested in humans. Clinical trials are conducted in multiple phases, beginning with Phase I trials to assess safety and dosage[24]. Phase II and III trials follow, focusing on evaluating efficacy and monitoring for adverse effects. AI models play a significant role in designing and optimizing these trials by predicting patient responses and identifying potential biomarkers for treatment stratification[12,45]. This integration of AI ensures that clinical trials are more efficient and targeted, increasing the likelihood of successful outcomes[78].

Future Perspectives

Advancements in AI and Drug Discovery

Artificial Intelligence (AI) is continually evolving, introducing new technologies that hold promise for revolutionizing drug discovery. Among the most impactful emerging AI technologies are quantum computing, federated learning, and reinforcement learning[54].

Quantum Computing: Quantum computing is poised to significantly accelerate AI's capabilities in drug discovery. Unlike classical computers, which use bits as the smallest unit of information, quantum computers use quantum bits or qubits, which can represent and process more complex data sets simultaneously[4]. This enables quantum computers to solve optimization problems and simulate molecular structures far more efficiently than classical computers[42]. In the context of drug discovery, quantum computing can expedite the process of molecular docking and interaction studies, allowing researchers to explore vast chemical spaces more comprehensively and rapidly[34,28].

Federated Learning: Federated learning is an emerging AI technique that enables multiple institutions to collaborate on model training without sharing their data directly. This approach addresses privacy concerns and enables the utilization of diverse datasets from different sources, leading to more robust and generalizable AI models[13,19]. In drug discovery, federated learning can facilitate collaboration between pharmaceutical companies, research institutions, and hospitals, combining their data to enhance the identification of potential anticancer compounds without compromising data privacy[20].

Reinforcement Learning (RL): Reinforcement learning is another promising AI technology, particularly useful in de novo drug design. RL algorithms learn optimal strategies through trial and error by interacting with an environment[22]. In drug discovery, RL can be used to design novel compounds by iteratively generating and testing new molecular structures, optimizing for desired properties such as high binding affinity and low toxicity. This approach can significantly enhance the efficiency of discovering new drug candidates[79].

Future Trends in AI-Guided Discovery of Marine Natural Products

The integration of AI into the discovery of marine natural products, including those derived from Sargassum spp., is expected to grow. Key trends include the increasing use of multi-omics data, the development of AI-driven natural product libraries, and enhanced AI models for predicting bioactivity[12,5].

Multi-Omics Integration: Future trends will likely see a deeper integration of multi-omics data (genomics, proteomics, metabolomics) into AI models. By combining data from various

biological layers, researchers can gain a holistic understanding of how marine natural products interact with biological systems. This comprehensive approach can lead to the discovery of novel anticancer compounds with unique mechanisms of action[19,3].

AI-Driven Natural Product Libraries: The creation of AI-driven natural product libraries is another anticipated trend. These libraries will utilize AI to predict the bioactivity of marine natural products, enabling more efficient screening and prioritization of compounds for experimental validation. By leveraging AI, these libraries can continuously evolve, incorporating new data to refine predictions and enhance the discovery process[33,2].

Enhanced Predictive Models: As AI technologies advance, we can expect the development of more sophisticated predictive models that can accurately forecast the bioactivity and toxicity of marine natural products. These models will incorporate advanced algorithms and larger, more diverse datasets, leading to improved predictions and a higher success rate in identifying viable anticancer compounds[80].

Research Directions for Sargassum spp.

Potential Areas of Research for Discovering New Anticancer Compounds

Research on Sargassum spp. holds considerable potential for uncovering new anticancer compounds. Several areas warrant further exploration, including bioactive compound isolation, mechanism of action studies, and synergy with existing treatments[4].

Bioactive Compound Isolation: One crucial area of research is the isolation and characterization of bioactive compounds from Sargassum spp. Advanced techniques such as high-performance liquid chromatography (HPLC) and mass spectrometry (MS) can be employed to identify and purify compounds with potential anticancer activity. Once isolated, these compounds can be screened for their efficacy against various cancer cell lines, providing a foundation for further drug development[20,3,33].

Mechanism of Action Studies: Understanding the mechanisms of action of bioactive compounds from Sargassum spp. is essential for their development as anticancer agents[44]. Research should focus on elucidating how these compounds interact with cellular targets and pathways to exert their anticancer effects. Techniques such as RNA sequencing, proteomics, and CRISPR-Cas9 gene editing can be used to investigate the molecular mechanisms underlying the anticancer activity of Sargassum-derived compounds[18].

Synergy with Existing Treatments: Another promising area of research is the potential synergy between compounds from Sargassum spp. and existing cancer treatments[29]. Studies can explore how these compounds enhance the efficacy of conventional therapies, such as chemotherapy and radiation, or reduce their side effects. This synergistic approach could lead to more effective combination therapies, improving patient outcomes[81].

Integration of AI with Other Technologies

The integration of AI with other cutting-edge technologies such as genomics and metabolomics can further enhance the discovery of anticancer compounds from Sargassum spp. This multidisciplinary approach leverages the strengths of each technology to provide deeper insights and more robust predictions[48,2].

Genomics: The integration of AI with genomics can provide a more comprehensive understanding of the genetic basis of cancer and the potential therapeutic targets of Sargassum-

derived compounds. AI can analyze genomic data to identify mutations and pathways associated with cancer, and then predict how Sargassum compounds may interact with these targets. This approach can guide the development of personalized therapies based on an individual's genetic profile[2,9,11].

Metabolomics: Metabolomics, the study of small molecules (metabolites) in biological systems, can be combined with AI to identify metabolic pathways affected by Sargassum compounds[20]. AI algorithms can analyze metabolomic data to uncover changes in metabolic profiles in response to treatment, providing insights into the compounds' mechanisms of action. This integration can also help identify biomarkers for monitoring treatment response and efficacy[12].

High-Throughput Screening (HTS): AI can enhance high-throughput screening by analyzing large datasets generated from HTS assays[38]. By predicting the activity of Sargassum compounds based on HTS data, AI can prioritize compounds for further testing, streamlining the drug discovery process[3]. Machine learning models can also identify patterns and relationships in HTS data that may not be apparent through traditional analysis methods[82].

2. Conclusion

Research and development opportunities for anticancer compounds from Sargassum spp. are bright with emerging AI technologies and multidisciplinary integration. In the near future, advances in AI, such as quantum computing, federated learning, and reinforcement learning, will revolutionize drug discovery, improving its efficiency and effectiveness. A greater emphasis will likely be placed on multi-omics integration, artificial intelligence-driven natural product libraries, and enhanced prediction models in the future, all contributing to the discovery of new anticancer agents. In parallel, continued research on Sargassum spp., focusing on bioactive compound isolation, mechanism of action studies, and synergy with existing treatments, will further our understanding of its potential as a source of new cancer therapies. By leveraging the power of AI and integrating it with other technologies, researchers can accelerate the development of effective anticancer treatments, ultimately improving outcomes for patients worldwide.

3. References

1. D. Eckelmann, "Spatial profiling of maytansine during the germination process of Maytenus senegalensis seeds," Fitoterapia, 2017.
2. N. Ziemert, "The evolution of genome mining in microbes - a review," Nat. Prod. Rep., 2016.
3. H. Russell et al., "Genome mining strategies for ribosomally synthesised and post-translationally modified peptides," Comput. Struct. Biotechnol. J., 2020.
4. Q. Wang, "Characterization of CYP76M5-8 indicates metabolic plasticity within a plant biosynthetic gene cluster," J. Biol. Chem., 2012.
5. D. Prihoda, "The application potential of machine learning and genomics for understanding natural product diversity, chemistry, and therapeutic translatability," Nat. Prod. Rep., 2021.
6. P. Cimermancic, "Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters," Cell, 2014.
7. P. Agrawal, "RiPPMiner-Genome: a web resource for automated prediction of crosslinked chemical structures of RiPPs by genome mining," J. Mol. Biol., 2021.

8. J. Shi, "Comparative genome mining and heterologous expression of an orphan NRPS gene cluster direct the production of ashimides," Chem. Sci., 2019.
9. H. H. Wang, "Chamiside A, a cytochalasan with a tricyclic core skeleton from the endophytic fungus Chaetomium nigricolor F5," Org. Lett., 2019.
10. F. Ul Haq, "A high-throughput method for dereplication and assessment of metabolite distribution in Salvia species using LC-MS/MS," J. Adv. Res., 2020.
11. E. Fox Ramos, "Natural products targeting strategies involving molecular networking: different manners, one goal," Nat. Prod. Rep., 2019.
12. J. W. Kim, "Unusual bridged angucyclinones and potent anticancer compounds from Streptomyces bulli GJA1," Org. Biomol. Chem., 2020.
13. T. H. Kuo, "Mass spectrometry imaging guided molecular networking to expedite discovery and structural analysis of agarwood natural products," Anal. Chim. Acta, 2019.
14. P. Phuwapraisirisan, "Anhydrocochlioquinone A, a new antitumor compound from Bipolaris oryzae," Tetrahedron Lett., 2007.
15. G. Atanasov, "Natural products in drug discovery: advances and opportunities," Nat. Rev. Drug Discov., 2021.
16. Gao K.Y. Interpretable drug target prediction using deep neural representation. In: Lang Jérôme., editor. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence; IJCAI; 2018. pp. 3371–3377.
17. Feng Q. Padme: a deep learning-based framework for drug–target interaction prediction. arXiv. 2018 arXiv:1807.09741.
18. Karimi M. DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. Bioinformatics. 2019;35(18):3329–3338.
19. Pu L. eToxPred: a machine learning-based approach to estimate the toxicity of drug candidates. BMC Pharmacol. Toxicol. 2019;20:2.
20. Mayr A. DeepTox: toxicity prediction using deep learning. Frontiers Environ. Sci. 2016;3:80.
21. Basile A.O., Yahi A., Tatonetti N.P. Artificial intelligence for drug toxicity and safety. Trends Pharmacol. Sci. 2019;40(September (9)):624–635. doi: 10.1016/j.tips.2019.07.005. Epub 2019 Aug 2. PMID: 31383376; PMCID: PMC6710127.
22. Lysenko A. An integrative machine learning approach for prediction of toxicity-related drug safety. Life Sci. Alliance. 2018;1:YYY–ZZZ.
23. Basile A.O. Artificial intelligence for drug toxicity and safety. Trends Pharmacol. Sci. 2019;40:624–635.
24. Dharmendra Bhati et al., "FUSED AND SUBSTITUTED PYRIMIDINE DERIVATIVES AS POTENT ANTICANCER AGENTS," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 24, no. 1, Jan. 2024, doi: https://doi.org/10.51470/bca.2024.24.1.749.
25. K. J. Mangala et al., "NANOCELLULOSE: A VERSATILE AND SUSTAINABLE CARRIER FOR DRUG AND BIOACTIVE COMPOUNDS," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 24, no. 1, Jan. 2024, doi: https://doi.org/10.51470/bca.2024.24.1.553.
26. Rohit Kumar Trivedi et al., "REVOLUTIONIZING DRUG DELIVERY THROUGH 3D PRINTING TECHNOLOGY: CURRENT ADVANCES AND FUTURE DIRECTIONS," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 24, no. 1, Jan. 2024, doi: https://doi.org/10.51470/bca.2024.24.1.521.
27. H. Rastogi, P. Bhatt, S. Garg, S. Kamboj, V. Deva, and R. Goel, "EXPLORING THE POTENTIAL OF QUANTUM DOTS AS LUMINOUS PROBES FOR TARGETED

DRUG DELIVERY AND BIOIMAGING IN CLINICAL DIAGNOSTICS," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 24, no. 1, Jan. 2024, doi: https://doi.org/10.51470/bca.2024.24.1.457.

28. M. shama, "CRISPR-Cas9 gene editing in pharmaceuticals : Current applications and future prospects," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 23, no. S1, Dec. 2023, doi: https://doi.org/10.51470/bca.2023.23.s1.1655.
29. S. Arora, Saiphali, G. Dharmamoorthy Dharmendra Bhati, T. Gupta, and P. Bhatt, "Advancements in peptide-based therapeutics: Design, synthesis and clinical applications," Biochemical and Cellular Archives/Biochemical and cellular archives, vol. 23, no. S1, Dec. 2023, doi: https://doi.org/10.51470/bca.2023.23.s1.1415.
30. M. Singhal et al., "Formulation development and characterization of powder for oral suspension containing H2 blocker drug to combat GERD and peptic ulcer," NeuroQuantology, vol. 20, no. 11, pp. 1258, 2022.
31. S. Ahamed, P. Bhatt, S. J. Sultanuddin, R. Walia, M. A. Haque, and S. B. InayathAhamed, "An Intelligent IoT enabled Health Care Surveillance using Machine Learning," in 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). IEEE, 2022.
32. V. Ahmed, S. Sharma, and P. Bhatt, "Formulation and evaluation of sustained release tablet of diltiazem hydrochloride," International Journal of Pharmaceutical Sciences and Research, vol. 11, no. 5, pp. 2193–2198, 2020.
33. A. E. Al-Snafi, S. Singh, P. Bhatt, and V. Kumar, "A review on prescription and non-prescription appetite suppressants and evidence-based method to treat overweight and obesity," GSC biol pharm sci, vol. 19, no. 3, pp. 148–155, 2022.
34. B. Baskar, S. Ramakrishna, and A. Daniela La Rosa, Eds., Encyclopedia of green materials. Singapore: Springer Nature Singapore, 2022.
35. P. Bhatt et al., "Nanorobots recent and future advances in cancer or dentistry therapy- A review," Am J PharmTech Res, vol. 9, no. 3, pp. 321–331, 2019.
36. P. Bhatt et al., "Citrus Flavonoids: Recent Advances and Future Perspectives On Preventing Cardiovascular Diseases," in The Flavonoids, 2024, pp. 131-152.
37. P. Bhatt et al., "Functional and tableting properties of alkali-isolated and phosphorylated barnyard millet (Echinochloa esculenta) starch," ACS Omega, vol. 8, no. 33, pp. 30294–305, 2023.
38. P. Bhatt et al., "Plasma modification techniques for natural polymer-based drug delivery systems," Pharmaceutics, vol. 15, no. 8, p. 2066, 2023.
39. P. Bhatt et al., "Comparative study and in vitro evaluation of sustained release marketed formulation of aceclofenac sustained release tablets," Pharma Science Monitor, vol. 9, no. 2, 2018.
40. P. Bhatt et al., "Development and characterization of fast dissolving buccal strip of frovatriptan succinate monohydrate for buccal delivery," Int J Pharm Investig, vol. 11, no. 1, pp. 69–75, 2021.
41. P. Bhatt et al., "Artificial intelligence in pharmaceutical industry: Revolutionizing drug development and delivery," The Chinese Journal of Artificial Intelligence, 2023.
42. P. Bhatt et al., "Blockchain technology applications for improving quality of electronic healthcare system," in Blockchain for Healthcare Systems, 2021, pp. 97–113.
43. P. Bhatt, "Mouth Dissolving Tablets Challenges, Preparation Strategies with a Special Emphasis on Losartan Potassium–A Review," World J. Pharm. Pharm. Sci, vol. 7, no. 9, pp. 271-287, 2018.
44. C. Goyal et al., "Estimation of shelf-life of Balachaturbhadrika syrup containing different sweetening agents," Res J Pharm Technol, pp. 5078–5083, 2022.

45. T. Kaur and S. Singh, "Controlled release of bi-layered malvidin tablets using 3D printing techniques," J Pharm Res Int, pp. 70–78, 2021.
46. M. Kaurav et al., "In-depth analysis of the chemical composition, pharmacological effects, pharmacokinetics, and patent history of mangiferin," Phytomed Plus, vol. 3, no. 2, p. 100445, 2023.
47. A. Kumar, P. Bhatt, and N. Mishra, "Irritable bowel Syndrome with reference of Alosetron Hydrochloride and Excipient profile used in the manufacturing of Alosetron tablet-A review," J Chem Pharm Sci, vol. 12, no. 03, pp. 71–78, 2019.
48. M. K. Malik et al., "Significance of chemically derivatized starch as drug carrier in developing novel drug delivery devices," Nat Prod J, 2022.
49. M. K. Malik et al., "Preclinical safety assessment of chemically cross-linked modified mandua starch: Acute and sub-acute oral toxicity studies in Swiss albino mice," ACS Omega, vol. 7, no. 40, pp. 35506–35514, 2022.
50. M. K. Malik et al., "Phosphorylation of alkali extracted mandua starch by STPP/STMP for improving digestion resistibility," ACS Omega, vol. 8, no. 13, pp. 11750–11767, 2023.
51. Pankaj, "Anti-cancer cyclodextrin nanocapsules based formulation development for lung chemotherapy," J Pharm Res Int, pp. 54–63, 2021.
52. Pankaj, "Cyclodextrin modified block polymer for oral chemotherapy," J Pharm Res Int, pp. 21–29, 2021.
53. V. Raghuwanshi et al., "Recent Advances In Nanotechnology For Combating Against Corona Virus Infection," Journal of Pharmaceutical Negative Results, pp. 1811-1820, 2022.
54. K. K. Sahu et al., "Utility of nanomaterials in wound management," in Nanotechnological Aspects for Next-Generation Wound Management, 2024, pp. 101–130.
55. S. K. Sharma et al., "Combined therapy with ivermectin and doxycycline can effectively alleviate the cytokine storm of COVID-19 infection amid vaccination drive: A narrative review," J Infect Public Health, vol. 15, no. 5, pp. 566–572, 2022.
56. S. K. Sharma and P. Bhatt, "Controlled release of bi-layered EGCG tablets using 3D printing techniques," J Pharm Res Int, pp. 5–13, 2021.
57. S. K. Sharma and S. Singh, "Antimicrobial Herbal Soap Formulation," Journal of Pharmaceutical Research International, vol. 32, no. 36, pp. 82-88, 2022.
58. S. Singh et al., "Cardiovascular comorbidity of COVID-19 disease: A review," WJPMR, vol. 8, no. 4, pp. 216–225, 2022.
59. S. Singh et al., "Phytonutrients, Anthocyanidins, and Anthocyanins: Dietary and Medicinal Pigments with Possible Health Benefits," in Advances in Flavonoids for Human Health and Prevention of Diseases, 2024, pp. 23-46.
60. S. Singh et al., "Digital Transformation in Healthcare: Innovation and Technologies," in Blockchain for Healthcare Systems, 2021, pp. 61–79.
61. S. Singh et al., "Alginate based Nanoparticles and Its Application in Drug Delivery Systems," Journal of Pharmaceutical Negative Results, pp. 1463-1469, 2022.
62. R. Johari et al., "Artificial Intelligence and Machine Learning in Drug Discovery and Development," in 2023 12th International Conference on System Modeling & Advancement in Research Trends (SMART), 2023, pp. 556-561.
63. P. Bhatt et al., "Impact of Cross-Linking on the Physicochemical and Physiological Characteristics of Barnyard Millet (Echinochloa frumentacea) Grains Starch," Stärke/Starch, May 2024, doi: https://doi.org/10.1002/star.202300285.
64. V. Kumar et al., "Ultrasound assisted techniques for starch modification to develop novel drug delivery systems: A comprehensive study," Journal of bioactive and compatible polymers, May 2024, doi: https://doi.org/10.1177/08839115241249143.

65. D. J. Newman et al., "Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019," J. Nat. Prod., 2020.
66. G. Li et al., "Strategies to diversify natural products for drug discovery," Med. Res. Rev., 2018.
67. M. L. Seca et al., "Plant secondary metabolites as anticancer agents: successes in clinical trials and therapeutic application," Int. J. Mol. Sci., 2018.
68. Y. Ren, "Potential anticancer agents characterized from selected tropical plants," J. Nat. Prod., 2019.
69. L. Eyberger, "Endophyte fungal isolates from Podophyllum peltatum produce podophyllotoxin," J. Nat. Prod., 2006.
70. W. X. Wang, "Unraveling the chemical interactions of fungal endophytes for exploitation as microbial factories," Nat. Prod. Rep., 2011.
71. R. N. Kharwar, "Harnessing the phytotherapeutic treasure troves of the ancient medicinal plant Azadirachta indica (Neem) and associated endophytic microorganisms," Planta Med., 2020.
72. F. Bucar, "Natural product isolation - how to get from biological material to pure compounds," Nat. Prod. Rep., 2013.
73. D. Kinghorn, "Discovery of anticancer agents of diverse natural origin," Anticancer Res., 2016.
74. X. Yang, "Concepts of artificial intelligence for computer-assisted drug discovery," Chem. Rev., 2019.
75. S. Vatansever, "Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: state-of-the-arts and future directions," Med. Res. Rev., 2021.
76. Fleming N. How artificial intelligence is changing drug discovery. Nature. 2018;557:S55–S55.
77. Dana D. Deep learning in drug discovery and medicine; scratching the surface. Molecules. 2018;23:2384.
78. Zang Q. In silico prediction of physicochemical properties of environmental chemicals using molecular fingerprints and machine learning. J. Chem. Inf. Model. 2017;57:36–49.
79. Öztürk H. DeepDTA: deep drug–target binding affinity prediction. Bioinformatics. 2018;34:i821–i829.
80. Lounkine E. Large-scale prediction and testing of drug activity on side-effect targets. Nature. 2012;486:361–367.
81. Mahmud S.H. iDTi-CSsmoteB: identification of drug–target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE. IEEE Access. 2019;7:48699–48714.