N. Karahan and Ç. Bilgi / Hacettepe J. Biol. & Chem., 2025, 53 (2), 61-70

Review Article

Hacettepe Journal of Biology and Chemistry

journal homepage: https://hjbc.hacettepe.edu.tr/
Dergipark Akademik: https://dergipark.org.tr/en/pub/hjbc

Modern Approaches to the Separation and Purification of Natural Products

Doğal Ürünlerin Ayırma ve Saflaştırma İçin Modern Yaklaşımlar

Nilay Karahan[1], Çiğdem Bilgi[2]*

[1]Formerly of Faculty of Pharmacy, Ege University, Izmir, Türkiye.

[2]Department of Pharmacognosy, Faculty of Pharmacy, Istanbul University-Cerrahpasa, Istanbul, Türkiye.

ABSTRACT

Natural products are bioactive compounds derived from plants, animals, and microorganisms, holding great importance in medicine and pharmacy. Today, the term 'natural product' generally refers to secondary metabolites from plants with various biological activities. These products have addressed primary health issues for thousands of years and continue to treat many diseases. Given the low levels of bioactive substances in plants and the difficulty of isolating these molecules in high purity, the importance of separating and purifying natural products becomes evident. Traditional methods of isolation and purification are expensive, time-consuming, and labor-intensive. However, advancements in science and technology have led to new approaches that are fast, effective, low-cost, and highly applicable. These modern methods enable the isolation of substances, determination of their bioactivity, and allow for the monitoring of their interactions, potential synergistic effects, and toxicities. These methods also help reduce cost and effort and save time, which is crucial for academia and the pharmaceutical industry. Numerous studies highlight drug-like molecules discovered through these new approaches. This review explores modern approaches to separating and purifying natural products and their advantages over traditional methods.

Key Words
Bioactive compound, chromatography, spectroscopy, metabolomics, secondary metabolite.

ÖZ

Doğal ürünler, bitkiler, hayvanlar ve mikroorganizmalardan elde edilen, tıp ve eczacılık alanında büyük öneme sahip biyoaktif bileşiklerdir. Günümüzde "doğal ürün" terimi genellikle bitkilerden elde edilen ve çeşitli biyolojik aktivitelere sahip sekonder metabolitleri ifade etmektedir. Bu ürünler binlerce yıldır temel sağlık sorunlarına çözüm olmuştur ve halen birçok hastalığın tedavisinde kullanılmaya devam etmektedir. Bitkilerdeki biyoaktif maddelerin düşük seviyelerde bulunması ve bu moleküllerin yüksek saflıkta izole edilmesinin zorluğu, doğal ürünlerin ayrıştırılması ve saflaştırılmasının önemini ortaya koymaktadır. Geleneksel izolasyon ve saflaştırma yöntemleri pahalı, zaman alıcı ve yoğun emek gerektiren yöntemlerdir. Ancak bilim ve teknolojindeki ilerlemeler, hızlı, etkili, düşük maliyetli ve yüksek uygulanabilirliğe sahip yeni yaklaşımların ortaya çıkmasına yol açmıştır. Bu modern yöntemler sayesinde maddelerin izolasyonu, biyoaktivitelerinin belirlenmesi, etkileşimlerinin, potansiyel sinerjik etkilerinin ve toksisitelerinin izlenmesi mümkün hale gelmiştir. Ayrıca bu yöntemler maliyeti azaltmaya ve zaman tasarrufu sağlamaya yardımcı olur ki bu da akademi ve ilaç endüstrisi için oldukça önemlidir. Çok sayıda çalışma, bu yeni yaklaşımlar sayesinde keşfedilen ilaç benzeri molekülleri vurgulamaktadır. Bu derleme, doğal ürünlerin ayrıştırılması ve saflaştırılmasında kullanılan modern yaklaşımları ve geleneksel yöntemlere göre avantajlarını incelemektedir.

Anahtar Kelimeler
Biyoaktif bileşik, kromatografi, spektroskopi, metabolomik, sekonder metabolit.

Article History: Aug 13, 2024; Accepted: Nov 25, 2024; Available Online: Dec 11, 2024.

DOI: https://doi.org/10.15671/hjbc.1527647

Correspondence to: Ç. Bilgi, Department of Pharmacognosy, Faculty of Pharmacy, Istanbul University-Cerrahpasa, Istanbul, Türkiye.

E-Mail: cigdem.karakoyun@iuc.edu.tr

INTRODUCTION

Natural products derived from diverse sources like plants, animals, and microorganisms possess a remarkable structural diversity and exhibit a wide range of biological activities, making them invaluable for drug discovery [1]. However, traditional methods for their separation and purification, including crystallization, precipitation, extraction, and distillation, often coupled with chromatographic techniques like thin layer chromatography (TLC), column chromatography, and high performance liquid chromatography (HPLC), have been hindered by significant limitations [2]. These limitations include low efficiency, requiring multiple repetitions and extensive time for isolation and purification, high resource consumption, necessitating large sample volumes and solvents, and limited sensitivity and resolution, struggling to resolve complex mixtures and identify low-abundance bioactive compounds [3]. The emergence of modern technologies has transformed the landscape of natural product separation and purification, overcoming the limitations of traditional methods and accelerating the discovery of bioactive molecules.

Natural products are essential resources for developing new lead compounds and structural scaffolds. Despite the inherent difficulties in harvesting adequate plant material and standardizing isolation methods, the therapeutic promise of natural products fuels ongoing efforts to develop new drugs from botanical sources [4]. The diversity of natural products worldwide is vast, and only a tiny portion has been studied for bioactivity to date. Given that the majority of plant biodiversity has not yet been screened, the potential biological activities and the possibility of each molecule theoretically acting on multiple targets indicate that natural products will continue to play an essential role in drug development in the future [5].

This review compiles modern approaches that offer solutions to these problems and play a crucial role in separating and purifying natural products. The advantages of modern approaches over existing methods are discussed.

Modern Methods Used in the Separation and Purification of Natural Products

In recent years, modern approaches like One Strain Many Compounds (OSMAC), high-throughput screening (HTS), and omics approaches like proteomics, transcriptomics, and metabolomics have become increasingly prominent in the field. These methods, especially beneficial for target-specific isolation and identifying complex mixtures, have accelerated drug discovery, led to the identification of novel compounds, and fostered a deeper understanding of natural product diversity.

OSMAC Approach

OSMAC is a simple yet effective approach frequently employed to increase production yields, generate a wider variety of natural products from a single microorganism, and activate silent metabolic pathways by manipulating culture conditions [6]. This method involves modulating factors such as temperature, pH, and nutrient sources within the culture medium, leading to the activation of previously silent gene clusters and the production of novel natural products. Notably, microorganisms, in their natural environment, compete for limited resources and develop diverse defense mechanisms, which often serve as the source of the bioactive molecules targeted by researchers [7].

While the concept of OSMAC gained widespread acceptance around two decades ago, its roots extend back to the 1960s, where it has been routinely utilized in industrial microbiology [8]. Traditional methods in microbial natural product research typically involve identifying the microorganism species, culturing it, and extracting the culture medium. Bioactivity-guided isolation and structure elucidation are then conducted on the extracts to uncover potential drug candidates. However, this approach frequently leads to the re-isolation of previously characterized natural compounds [9,10]. This is attributed to the fact that certain metabolic pathways remain inactive under consistent conditions.

Over the past two decades, research has demonstrated that various nutrients, trace elements, physical parameters (e.g., pH and temperature), and chemical elicitors (such as non-lethal antibiotic concentrations) can effectively influence secondary metabolite production [11]. By modifying culture conditions through alterations in cultivation parameters, co-culturing, the addition of enzyme inhibitors, and other strategies, researchers aim to activate silent genes and manipulate the secondary metabolite profiles of microorganisms [12].

Stress factors have also been shown to play a crucial role in triggering the expression of secondary metabolite genes. Studies indicate that modifying a single factor

within the culture environment and the resulting stress response can induce the production of novel secondary metabolites. Common methods for inducing stress responses include heat shock and ethanol shock [13]. Zeeck and colleagues reported that by altering cultivation parameters such as temperature, salinity, aeration, and even the shape of the culture vessels, the fungus *Aspergillus ochraceus*, previously known to produce only the metabolite aspinone, could produce 15 additional metabolites [14].

The production of the secondary metabolite jadomycin B by *Streptomyces venezuelae* exemplifies the role of stress responses. Jadomycin B is a benzoxazolophenanthridine antibiotic with potent activity against Gram-positive pathogens like *Staphylococcus aureus* and *S. epidermidis*. Under normal culture conditions, *Streptomyces venezuelae* produces relatively small amounts of jadomycin B. However, increasing the temperature from 27°C to 42°C results in significantly higher yields of this compound. Similarly, exposing cultures to 6% ethanol induces jadomycin B yields of up to 30 µg/ml without requiring an increase in temperature [13]. These alterations in cultivation processes allow the organism to respond broadly by adjusting its cellular mechanisms in response to diverse growth conditions [11].

As research into OSMAC continues to evolve, it holds immense promise for the discovery and development of novel bioactive compounds, paving the way for new therapeutic strategies.

High Throughput Screening (HTS)

High-throughput screening (HTS) is a pivotal technique in drug discovery research, enabling the simultaneous analysis of numerous compounds. It offers significant advantages in terms of time efficiency, labor savings, and cost-effectiveness.

Despite varying strategies in natural product research, such as sample selection and collection, isolation techniques, structural determination, and dereplication, the rate of new natural product discovery has decreased. This is due to the high cost and time loss associated with traditional methods in natural products. To solve this problem, the pharmaceutical industry uses modern HTS to produce new drugs [15].

Drugs fail mainly for two reasons: lack of efficacy and safety issues. Therefore, one of the most critical steps in developing a new drug is target identification and validation. A target is a broad term that can include proteins, genes, and RNA, applicable to various biological entities. A good target must be effective, safe, meet clinical and commercial needs, and, above all, be "druggable." Approximately 500 molecular targets are followed by drug discovery programs. According to this analysis, cell membrane receptors, primarily G protein-coupled receptors (GPCRs), constitute the largest subgroup at 45%. Enzymes account for the next largest share at 28%, followed by hormones (11%), unknowns (7%), ion channels (5%), nuclear receptors (2%), and DNA (2%) [16].

Good target identification provides confidence in the relationship between the target and the disease and information on whether target modulation will lead to mechanism-based side effects [17]. After identifying the target, the target validation stage determines if it is relevant to the disease being studied. Following the target validation process, compound screening assays are performed during the "hit" identification and lead discovery phase of the drug discovery process. The meaning of a "hit" molecule can vary among researchers but is generally defined as a compound that has the desired activity in a screen and whose activity is confirmed upon retesting. Various screening paradigms are available to identify hit molecules. HTS involves directly screening the entire compound library against the drug target or a complex assay system dependent on target activity, which requires subsequent secondary analyses [18]. Laboratory automation is used during these screenings.

HTS assumes no prior knowledge about the nature of the compound likely to be effective on the target protein. Subsets of small molecules likely to be active on the target protein are selected from chemical libraries. Compound libraries are assembled to include small molecules with chemical parameters that comply with the Lipinski Rule of Five [19]. Molecules typically have molecular weights less than 400 Da and logP values (expressing lipophilicity, the partition coefficient between octanol and water) of less than 4. Compounds with molecular weights ≤350 and logP <3 are called "drug-like" [17]. The number of hydrogen bond donors and acceptors (associated with bioavailability) and the absence of highly reactive chemical groups (unstable or toxic to liver enzymes) are important in compound selection.

Numerous assay formats have been enabled to support

compound screening [20]. The choice of assay format depends on the biology of the drug target protein, the equipment infrastructure in the laboratory where the study will be conducted, the experience of the scientists in that laboratory, whether an inhibitor or activator molecule is being sought, and the scale of the compound screening.

HTS is a method that shows the effects of a large number of substances on different parameters simultaneously and in a short period. It uses high-capacity laboratory equipment, allowing thousands of compounds to be screened simultaneously [21]. Using this technique, metabolic, pharmacokinetic, and toxicological data related to new drugs can be characterized.

Today, high throughput methods are in high demand in drug development studies. Their goals are generally to accelerate drug discovery by screening large libraries of hundreds of thousands of drug candidates (combinatorial chemistry, genomics, protein, and peptide libraries). With HTS, it is possible to screen up to 10,000 compounds per day. Ultra High Throughput Screening (UHTS) can conduct up to 100,000 tests per day [22].

HTS plays a significant role in the early phase of drug development. It provides qualitative and quantitative characterization of compound libraries and analytical support for preclinical and clinical ADME (Absorption, Distribution, Metabolism, and Excretion)) studies [23]. Therefore, HTS enables the early elimination of unsuitable compounds in drug discovery studies. Robotic systems, data processing and control software, automated devices for liquid sample preparation, and sensitive detectors are used during the process. These studies provide insights into drug design and understanding the roles and interactions of biochemical processes.

Omics Approaches

The term omics was derived from the word “genome” by Hans Winkler in the 1920s. The suffix “-ome”, from Greek, gives a sense of completeness to the word it is attached to [24]. This term was introduced as genomics in the 1980s and became widely used in the 1990s. In the 2000s, omic technologies, with genomics as their foundation, expanded into various fields such as transcriptomics, proteomics, and metabolomics [25].

Omics technologies allow us to see the differences in DNA, RNA, proteins, and other cellular molecules between species and among individuals of a species. Molecular profiles can vary depending on exposure to chemicals, drugs, cells, or tissues, so omics systems are also used in toxicological assessments. Omics experiments provide a wealth of data about functional or structural changes within cells. They play a crucial role in measuring molecular responses resulting from cell or tissue damage and understanding the complexities in functional cellular systems [26]. Omics platforms allow comprehensive analysis of genes (genomics), mRNA (transcriptomics), proteins (proteomics), and metabolites (metabolomics) [27].

Genomics

The complete set of genes within an organism is referred to as its genome. Genomics examines the interactions and communications of the genes with each other and their environment in terms of time, place, and amount, and controls their production and activation [28]. Through genomics, information about the types, numbers, and functions of proteins produced by organisms can be obtained, evolutionary similarities among organisms can be investigated, and genetic information from different organisms can be compared. The data resulting from genomic research are processed and stored in computer databases [28].

Genomic studies have facilitated the identification of genes related to diseases and physiological processes, the elucidation of structural-functional interactions between genes, the roles and expression profiles of genes in development, and the comparison of different organisms on a genetic basis [29]. Genomics, with its ability to evaluate the entire genome in a single study, differs from other genetic disciplines that focus on individual genes.

Genome-Wide Association Studies (GWAS) are a strategy proposed to identify genetic factors involved in the development of human diseases [30]. Modern genetic variation is termed single nucleotide polymorphism (SNP). SNPs are single base pair changes in the DNA sequence that occur at a high frequency in the human genome. They occur approximately every 200 to 300 base pairs and are effectively used to identify the locations of certain genes associated with various diseases. In genetic studies, SNPs are typically used as markers of a genomic region, and the majority have minimal impact on biological systems. SNPs can cause amino acid changes, changes in mRNA transcript stability, and changes

in transcription factor binding affinity, representing the most common form of genetic diversity in the human genome [31]. Genomic research allows for the identification of new genetic risk factors for many diseases. Companies examine the presence of previously identified associated SNPs for such risk factors. Knowing an individual's genetic risks can be beneficial in treatments. Such hereditary tests reflect relationships and provide predictions for other patients.

Transcriptomics

Transcriptomics is a discipline that links the genome, proteome, and cellular phenotype by simultaneously studying mRNA transcripts produced from the cellular genome through transcription, creating expression profiles [32].

Thanks to next-generation sequencing platforms, RNA material can be examined in parallel and simultaneously using transcriptome sequencing methods. Transcriptomics aims to measure the expression levels of all or a selected subset of genes in a sample based on the amount of RNA present, commonly using microarrays and DNA chips [33,34]. The resulting expression profiles are referred to as the transcriptome, representing all gene transcripts in a cell or tissue at a given time. Through transcriptome sequencing, it is possible to examine all genes and isoforms present in a transcript, identify the 5' and 3' ends of genes, SNPs, insertions, deletions, translocations, fusion genes, and the expressions of these genes, as well as non-coding RNAs [34].

As different functions are carried out within the cell, the transcription of genes associated with these functions increases, and since the transcriptome encompasses all mRNAs in the cell, it reflects the genes active at a given time. mRNAs are synthesized from a DNA template via transcription, transporting protein-coding information to ribosomes, the site of protein synthesis. Unlike the stable genome in a cell, the transcriptome can change in response to environmental factors such as nutrient variety, temperature changes, pH changes, and signals from other cells. Uncovering changes in gene expression in response to environmental factors is important for systems biology, especially in understanding environment-system interactions [35].

Global transcriptome profile analysis represents the first level of the dynamic state of a biological system, providing insights into potentially active processes under specific biological conditions [36]. It allows for the analysis of expression under different conditions, understanding plant response mechanisms, and regulating various biological processes. This enables the functional characterization of enzymes involved in the biosynthesis of valuable secondary metabolites, leading to advancements in the development of novel drugs and therapeutic agents.

Proteomics

Proteomics is defined as the total set of proteins expressed by a genome and present in a particular time and space. In this context, "space" refers to the state of proteins in different cell compartments or types, while "time" denotes processes such as aging, various developmental stages, environmental conditions, and disease states. Thus, the proteome is a dynamic entity that varies according to cells, internal and external stimuli, cell cycle stages, tissues, and environmental conditions. It is defined as the quantitative analysis technology of proteins in cells, tissues, or body fluids under different conditions [37]. The term "proteome" was first proposed by Australian researcher Marc Wilkins at a meeting in Siena in 1994 [38].

Proteins, the fundamental units of the proteome, are composed of amino acids and function in their active forms as three-dimensional structures. Although the amino acid sequences of proteins are determined by specific genes, genetic information alone does not provide complete information about a protein. All cells in an organism carry the same genome, but different cell types can synthesize thousands of different proteins. Unlike the one-dimensional genomic information expressed by combinations of four nucleotides, the information encoded in proteins is not limited to amino acid sequences. Proteins undergo numerous modifications by cells. The dynamic nature of the proteome is not solely dependent on environmental factors. The importance of proteomic studies has increased because proteins interact more with other elements in the organism than other molecules and carry out cellular activities, despite nucleic acids being the primary regulators [39].

Proteomics is a branch of functional genomics that studies proteomes. It examines all protein structures synthesized by the genome at a specific time and place, including their quantities, localizations, post-translational modifications, functions in tissues and cells, and interactions with other proteins or macromolecules.

Proteomic analysis provides critical information on identifying existing proteins, determining their quantities, localizations, three-dimensional structures, post-translational modifications, and protein-protein interactions. Proteomic research also enables the identification of biochemical changes by controlling protein groups potentially related to toxicity [40]. Comparative proteomics allows the comparison of expression differences between two conditions (e.g., normal vs. diseased, old vs. young) [37]. The occurrence of pathological conditions in an organism is directly related to changes in the proteome, and proteomic techniques are used to understand these changes and their causes [41]. With its ability to decipher the protein landscape of cells and organisms, proteomics has emerged as a powerful tool for advancing our understanding of biological processes and developing new diagnostic and therapeutic strategies.

Metabolomics

The Human Genome Project, completed in 2003, revealed that approximately 99.9% of the approximately 30,000 genes found in the human body are the same across individuals, with only a 0.1% difference [42]. This seemingly small 0.1% difference plays a significant role in explaining why some individuals are more prone to certain diseases, why disease progressions vary among individuals, and why some individuals respond better to medications. After the identification of genes, research into their functions began to address these remaining unknowns. Transcriptomics and proteomics studies were conducted for this purpose. However, the information obtained from these studies was insufficient to explain clinical phenotypes. Later research revealed that the key information determining clinical phenotypes was hidden in the metabolites produced by cells [38].

Metabolomics is the identification, quantification, and characterization of small-molecule metabolites, including lipids, carbohydrates, vitamins, hormones, and other cellular components, in tissues, cells, and physiological fluids using high-throughput technologies. Small molecules can include metabolites such as peptides, oligonucleotides, sugars, nucleosides, organic acids, ketones, aldehydes, amines, amino acids, lipids, steroids, alkaloids, drugs, and human and bacterial products. Their molecular weights are below 1,500 Da [43]. In natural product research, metabolomics is considered a large-scale analysis of an organism's metabolites during various physiological states.

Metabolomic technologies are widely used in drug discovery and development. The technology's advancement is highly significant for the pharmaceutical industry, as it facilitates the identification of hundreds of endogenous and exogenous metabolites from urine, plasma, and tissue samples. Metabolomics has the potential to have a strong impact on preclinical drug development studies, including the identification of new drug targets, elucidation of the mechanism of action of new drugs, development of safety and efficacy profiles, and the ADME of new drugs. Toxicity is a cause of attrition in all stages of the drug development process. Metabolic studies allow the identification of toxicity early in the drug discovery process, saving time and money for pharmaceutical companies [44].

Metabolites are a result of the interaction between an organism's genome and its environment. Therefore, metabolomics allows scientists to explore gene-environment interactions. While genomics and proteomics studies tell researchers "what could be," metabolomics provides information about "what actually is". Therefore, metabolomic studies are a valuable method for investigating disease diagnosis or the effects of toxic agents on phenotype [25].

Metabolomics originated from a concept called metabolic profiling, which refers to the analysis of a group of metabolites. Its goal is to qualitatively and quantitatively analyze all metabolites present in an organism at a given time and under specific conditions. This approach is a powerful tool because it reflects the biochemical activity of cells and tissues. It can be used to optimize biosynthetic pathways for the selective production of biologically active secondary metabolites [45].

Molecular Networking

In natural product research, identifying known molecules beforehand minimizes the time, effort, and cost of the research process. Therefore, developing algorithms that allow for the systematic screening and identification of structurally diverse natural product compounds is crucial [46]. Molecular networking is a computer-based strategy that visualizes complex data obtained from MS (mass spectroscopy) analyses, making it easier to understand and interpret findings [47]. This dereplication technique, introduced in 2012, revolutionized the isolation of natural products, shifting from the traditional

"grind and find" model to hypothesis-driven targeting [48]. As an accessible and adaptable method, molecular networking enables the visualization and targeted analysis of natural products, supporting biological research and biotechnological applications across various fields [49].

It rapidly identifies previously unreported natural products in complex mixtures and systematically categorizes thousands of specialized metabolites based on their chemical structures. This process prevents redundant purification of known compounds (dereplication) and allows for focused studies on new compounds [50].

Molecular networking visually analyzes all ions detected during the fragmentation of molecules in MS analysis and the chemical relationships between these ions. It is a graph-based computational algorithm that aims to organize large MS datasets by revealing spectral similarities between MS-MS fragmentation patterns. These techniques are essential tools used in the identification, discovery, and characterization of secondary metabolites [51].

Molecular networks built using tandem MS-MS data are based on the principle that structurally similar molecules produce similar MS-MS fragmentation patterns [52]. MS-MS data is represented by a graph where each node represents an ion with a corresponding fragmentation spectrum, and connections between nodes provide insight into the similarities between spectra. Molecular networking utilizes a vector-based computational algorithm at this point to compare the degree of spectral similarity [52].

Because structurally similar natural products produce similar MS-MS fragmentation patterns, molecular families tend to cluster together within a network. The dissemination of structural information within the network highlights unknown but structurally related molecules, enabling successful dereplication [46]. These networks also allow for simultaneous visual analysis of the same molecules, analogs, or compound families obtained from single or multiple datasets and diverse biological sources [53,54].

The molecular networking approach, based on the assumption that related molecules produce similar MS-MS fragmentation, generates an MS-MS spectral similarity map that allows for the visualization of structurally related molecules. The primary strength of this approach lies in its applicability for exploring thousands and millions (potentially billions) of MS-MS spectra without any prior knowledge about the chemical composition of the samples [55]. Additionally, if MS-MS spectra are available in open-access MS-MS spectral libraries, molecular networks can perform an automated search for known molecules [56].

The widespread adoption of molecular networking has led to the development of the GNPS web platform, a valuable resource for data sharing and analysis. GNPS enables users to upload and store MS-MS data, generate molecular networks, and visualize molecular families online [56]. For more comprehensive visualization, researchers can utilize tools like Cytoscape to analyze complete networks and adjust network attributes for better data interpretation [57].

Molecular networking has revolutionized natural product research by enabling rapid identification, dereplication, and visualization of complex mixtures of natural products. This powerful tool continues to transform the field, driving the discovery of novel bioactive compounds and advancing our understanding of the diverse chemical world of nature.

RESULTS and CONCLUSION

Natural products have long been a rich source of successful drug molecules. With high structural diversity and a wide range of bioactivities, natural products continue to hold promise for the discovery of novel pharmacological effects and new structural scaffolds.

Traditional methods used in the separation and purification of natural products are labor-intensive, time-consuming, expensive, and often result in unsuccessful studies. These challenges present disadvantages for the application of natural product-derived compounds in academia and industry.

With advancements in technology, rapid, novel, targeted, and highly sensitive techniques have been developed in the field of separation and purification. The new approaches are bioactivity-focused and are utilized in many aspects of the process, from the synthesis of the active compound from its natural source to obtaining the final product. These approaches have the potential to revitalize natural product-based drug discovery

research. This review presents a compilation of novel approaches employed in the separation and purification of natural products, along with examples of their application.

Among the innovative approaches mentioned above, OSMAC is a simple and effective approach used to increase production yields by modifying culturing conditions, obtaining a wider variety of natural products from a single microorganism, and activating silent metabolic pathways.

HTS is another modern approach that plays a significant role in the early stages of drug development. With HTS, the effects of numerous substances on various parameters can be analyzed simultaneously in a short time. The ability to screen a large number of molecules in a single experimental set is possible thanks to this approach. Today, it accelerates the screening of combinatorial libraries, saving cost and time in drug development and "drug repurposing" studies.

Moreover, innovative omics approaches enable the assessment of the genetic potential of producer strains, the expression of biosynthetic gene clusters (BGCs) that encode secondary metabolites, and the more effective identification of previously unknown metabolites. Significant advancements in omics technologies allow for the high-throughput monitoring of various molecular and organismal processes. These techniques are widely applied to identify biological variants (e.g., biomarkers), characterize complex biochemical systems, and investigate pathophysiological processes. The integration of omics datasets through robust statistical tests provides information used to create system predictions and metabolic models.

Molecular networking is an approach that enables the very rapid identification of previously unreported natural products in complex mixtures. With this method, analyzes can be performed on the metabolome collected from organisms, allowing metabolites to be quickly categorized based on their chemical structures. This approach allows for targeted studies of new compounds, and it saves time by preventing the repeated purification of known compounds.

As described above, drug discovery and development efforts based on natural products are a complex endeavor requiring a highly integrated interdisciplinary approach. However, the presented recent scientific/technological advancements and innovative approaches provide solutions to challenges in this process and facilitate the work of researchers. Current developments in the field of separation and purification clearly indicate that natural products, as one of the most important sources of new drug candidates, will continue to play a significant role in pharmaceutical science.

These advancements hold immense promise for the future of medicine, particularly in addressing unmet needs. By accelerating drug discovery and uncovering hidden biosynthetic potential, these strategies pave the way for the development of more effective, safer, and targeted therapies. Furthermore, their application in areas like rare diseases, often neglected due to high development costs, could usher in a new era of personalized treatments. It is essential that these innovative approaches continue to be integrated and refined alongside emerging technologies, ensuring that the full potential of natural products is harnessed for the benefit of human health.

Acknowledgments

This research paper is based on an undergraduate research project conducted as a bachelor's thesis at Ege University, Faculty of Pharmacy, authored by Nilay Karahan under the supervision of Dr. Çiğdem Bilgi.

References

1. D.D. Baker, M. Chu, U. Oza, V. Rajgarhia, The value of natural products to future pharmaceutical discovery, *Nat. Prod. Rep.*, 24:6 (2007) 1225-1255, 2007, doi: 10.1039/b602241n.
2. O. Sticher, Natural product isolation, *Nat. Prod. Rep.*, 25:3 (2008) 517-554.
3. Q.W. Zhang, L.G. Lin, W.C. Ye, Techniques for extraction and isolation of natural products: a comprehensive review, *Chin. Med.*, 13:1 (2018) 26.
4. A. G. Atanasov, B.Waltenberger, E-M. Pferschy-Wenzig, T. Linder, C. Wawrosch, P. Uhrin, V. Temml, L. Wang, S. Schwaiger, E.H. Heiss, J.M. Rollinger, D. Schuster, J.M. Breuss, V. Bochkov, M.D. Mihovilovic, B. Kopp, R. Bauer, V.M. Dirsch, H. Stuppner, Discovery and resupply of pharmacologically active plant-derived natural products: A review, *Biotechnol. Adv.*, 33:8 (2015) 1582–1614.
5. S. Bernardini, A. Tiezzi, V. Laghezza Masci, E. Ovidi, Natural products for human health: an historical overview of the drug discovery approaches, *Nat. Prod. Res.*, 32:16 (2018) 1926–1950.
6. J. S. Zarins-Tutt, T.T. Barberi, H. Gao, A. Mearns-Spragg, L. Zhang, D.J. Newman, R.J.M. Goss, Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products, *Nat. Prod. Rep*. 33:1 (2016) 54–72.

7. J. Wiese, B. Ohlendorf, M. Blümel, R. Schmaljohann, J. F. Imhoff, Phylogenetic Identification of Fungi Isolated from the Marine Sponge Tethya aurantium and Identification of Their Secondary Metabolites, *Mar. Drugs*, 9:4 (2011) 561–585.
8. G. Atanasov, S. B. Zotchev, V. M. Dirsch, C. T. Supuran, "Natural products in drug discovery: advances and opportunities," *Nat Rev Drug Discov*, vol. 20, no. 3, pp. 200–216, Mar. 2021, doi: 10.1038/s41573-020-00114-z.
9. K. Scherlach, C. Hertweck, Triggering cryptic natural product biosynthesis in microorganisms, *Org. Biomol. Chem.*, 7:9 (2009) 1753.
10. Z. Wasil, K.A.K. Pahirulzaman, C. Butts, T.J. Simpson, C.M. Lazarus, R.J. Cox, One pathway, many compounds: heterologous expression of a fungal biosynthetic pathway reveals its intrinsic potential for diversity, *Chem. Sci.*, 4:10 (2013) 3845.
11. R. Pan, X. Bai, J. Chen, H. Zhang, H. Wang, Exploring Structural Diversity of Microbe Secondary Metabolites Using OSMAC Strategy: A Literature Review, *Front. Microbiol.*, 10 (2019).
12. D. M. Selegato, I. Castro-Gamboa, Enhancing chemical and biological diversity by co-cultivation, *Front. Microbiol.*, 14 (2023) 1117559.
13. V. Yoon, J.R. Nodwell, Activating secondary metabolism with stress and chemicals, *J. Ind. Microbiol. Biotechnol.*, 41:2 (2014) 415–424.
14. H.B. Bode, B. Bethe, R. Höfs, A. Zeeck, Big Effects from Small Changes: Possible Ways to Explore Nature's Chemical Diversity, *ChemBioChem.*, 3:7 (2002) 619.
15. K.P. Mishra, L. Ganju, M. Sairam, P.K. Banerjee, R.C. Sawhney, A review of high throughput technology for the screening of natural products, *Biomedicine & Pharmacotherapy*, 62:2 (2008) 94–98.
16. J.A. Landro, I.C.A. Taylor, W.G. Stirtan, D.G. Osterman, J. Kristie, E.J. Hunnicutt, P.M.M. Rae, P.M. Sweetnam, HTS in the new millennium, *J. Pharmacol. Toxicol. Methods*, 44:1 (2000) 273–289.
17. J. Hughes, S. Rees, S. Kalindjian, K. Philpott, Principles of early drug discovery, *Br. J. Pharmacol.*, 162:6 (2011).
18. S. Fox, S. Farr-Jones, L. Sopchak, A. Boggs, H.W. Nicely, R. Khoury, M. Biros, High-Throughput Screening: Update on Practices and Success, *SLAS Discovery*, 11:7 (2006) 864–869.
19. C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development, *Adv. Drug. Deliv. Rev.*, 46 (2001) 1–3.
20. V. Blay, B. Tolani, S.P. Ho, M.R. Arkin, High-throughput screening: today's biochemical and cell-based approaches, *Drug. Discov. Today*, 25:10 (2020) 1807–1821.
21. O. Osakwe, The Significance of Discovery Screening and Structure Optimization Studies, *Social Aspects of Drug Discovery, Development and Commercialization*, (2016) 109–128.
22. P. Szymański, M. Markowicz, E. Mikiciuk-Olasik, Adaptation of High-Throughput Screening in Drug Discovery—Toxicological Screening Tests, *Int. J. Mol. Sci.*, 13:1 (2011) 427–452.
23. S. N. R. Gajula, N. Nadimpalli, and R. Sonti, "Drug metabolic stability in early drug discovery to develop potential lead compounds," *Drug Metab Rev*, vol. 53, no. 3, pp. 459–477, 2021.
24. S.H. Bal, F. Budak, Genomik, Proteomik Kavramlarına Genel Bakış ve Uygulama Alanları, *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 39:1 (2013) 65–69.
25. Ş. Ö.Budak, S. Dönmez, Gıda Biliminde Yeni Omik Teknolojileri, *The Journal of Food*, 37:3 (2012) 173–179.
26. M.J. Aardema, J.T. MacGregor, Toxicology and genetic toxicology in the new era of 'toxicogenomics': impact of '-omics' technologies, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 499:1 (2002) 13–25.
27. K. Wanichthanarak, J.F. Fahrmann, D. Grapov, Genomic, Proteomic, and Metabolomic Data Integration Strategies, *Biomark Insights*, 10:4 (2015) 29511.
28. B.S. Yarman, H. Gürkan, Ü. Güz, B. Aygün, A New Modelling Method of the ECG Signals Based on the use of an Optimized Predefined Functional Database, *Acta. Cardiol*, 58:3 (2003) 59–61.
29. P.A. Hohenlohe, W.C. Funk, O.P. Rajora, Population genomics for wildlife conservation and management, *Mol. Ecol.*, 30:1 (2021) 62–82.
30. Ö. Gök, A. Aslan, O. Erman, İnsan ENCODE, HapMap ve 1000 Genom Projeler, *Erciyes University Journal of Natural and Applied Sciences*, 33:2 (2017) 35–42.
31. W.S. Bush, J.H. Moore, Chapter 11: Genome-Wide Association Studies, *PLoS Comput Biol*, 8:12 (2012) 1002822.
32. L. Cocolin, K. Rantsiou, Molecular Biology | Transcriptomics, *Encyclopedia of Food Microbiology*, Elsevier, Netherlands, 2014.
33. K-H. Liang, Transcriptomics. *Bioinformatics for Biomedical Science and Clinical Applications*, Elsevier, Netherlands, 2013.
34. M. Sarıman, S.S. Ekmekci, N. Abacı, A. Çakiris, F. Paçal, Z. Emrence, D. Üstek, Transcriptome Analysis Using Next Generation Sequencing, *Experimed.*, 10:4 (2015) 51–59.
35. A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gilette, A. Paulovic, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles, *Proceedings of the National Academy of Sciences*, 102: 43 (2005) 15545–15550.
36. Z. Miao, B.D. Humphreys, A.P. McMahon, J. Kim, Multi-omics integration in the age of million single-cell data, *Nat. Rev. Nephrol.*, 17:11 (2021) 710–724.
37. G. Marko-Varga, Proteomics principles and challenges, *Pure and Applied Chemistry*, 76:4 (2004) 829–837.
38. E .Başaran, S. Aras, General Outlook and Applications of Genomics, Proteomics and Metabolomics, *Turkish Bulletin of Hygiene and Experimental Biology*, 67:2 (2010) 85–96.
39. H. Kasap, P. Pazarcı, Sistem Biyolojisi, *Arşiv,*19:25 (2010) 25–35.
40. C. B. Kasurka, Moleküler Toksikoloji ve 'Omik' Teknolojileri, *Nevşehir Bilim ve Teknoloji Dergisi*, 8:1 (2019) 42–55.
41. M. Moreira, D. Schrama, A.P. Farinha, M. Cerqueira, Fish pathology research and diagnosis in aquaculture of farmed fish; a proteomics perspective, *Animals*, 11:1 (2021) 125.
42. J.C. Venter, A Part of the Human Genome Sequence, *Science,* 299:5610 (2003) 1183–1184.
43. R. Goodacre, Metabolomics – the way forward, *Metabolomics*, 1:1 (2005) 1–2.
44. Q. Zhao, J.L. Zhang, and F. Li, Application of Metabolomics in the Study of Natural Products, *Nat. Prod. Bioprospect*, 8:4 (2018) 321–334.
45. L. Harvey, R. Edrada-Ebel, R.J. Quinn, The re-emergence of natural products for drug discovery in the genomics era, *Nat. Rev. Drug Discov.*, 14:2 (2015) 111–129.
46. Y. Gao, Y. Wang, H. Chung, K. Chen, T. Shen, C. Hsu, Molecular networking as a dereplication strategy for monitoring metabolites of natural product treated cancer cells, *Rapid Communications in Mass Spectrometry*, 34:S1 (2020).

47. F. Vincenti, C. Montesano, F. Di Ottavio, A. Gregory, D. Compagnone, M. Sergi, P. Dorrestein, Molecular networking: a useful tool for the identification of new psychoactive substances in seizures by LC–HRMS, *Front. Chem.*, 8 (2020) 572952.
48. E. Fox Ramos, L. Evanno, E. Poupon, P. Champy, M.A. Beniddir, Natural products targeting strategies involving molecular networking: different manners, one goal, *Nat. Prod. Rep.*, 36:7 (2019) 960–980.
49. S. Allard, P. Allard, I. Morel, T. Gicquel, Application of a molecular networking approach for clinical and forensic toxicology exemplified in three cases involving 3-MeO-PCP, doxylamine, and chlormequat, *Drug. Test. Anal.*, 11:5 (2019) 669–677.
50. J. Hubert, J.M. Nuzillard, J.H. Renault, Dereplication strategies in natural product research: How many tools and methodologies behind the same concept?, *Phytochemistry Reviews*,16:1 (2017) 55–95.
51. D.D. Nguyen, C.H. Wu, W.J. Moree, A. Lamsa, M.H. Medema, X. Zhao, R.G. Gavilan, M. Aparicio et al., MS/MS networking guided analysis of molecule and gene cluster families, *Proceedings of the National Academy of Sciences*, 110:28 (2013).
52. R.A. Quinn, L.F. Nothias, O. Vining, M. Meehan, E. Esquenazi, P.C. Dorrestein, Molecular Networking As a Drug Discovery, Drug Metabolism, and Precision Medicine Strategy, *Trends Pharmacol Sci*, 38: 2 (2017) 143–154.
53. A. C. Pilon, H. Gu, D. Raftery, V.S. Bolzani, N.P. Lopes, I. Castro-Gamboa, F. Carnevale Neto, Mass Spectral Similarity Networking and Gas-Phase Fragmentation Reactions in the Structural Analysis of Flavonoid Glycoconjugates, *Anal Chem*, 91:16 (2019) 10413–10423.
54. S. Woo, K. Bin Kang, J. Kim, S.H. Sung, Molecular Networking Reveals the Chemical Diversity of Selaginellin Derivatives, Natural Phosphodiesterase-4 Inhibitors from Selaginella tamariscina, *J. Nat. Prod*, 82:7 (2019) 1820–1830.
55. N. Garg, C.A. Kapono, Y.W. Lim, N. Koyama, M.J.A. Vermeij, D. Conrad, F. Rohwer, P.C. Dorrestein, Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures, *Int. J. Mass. Spectrom.*, 377 (2015) 719–727.
56. M. Wang, J.J. Carver, V.V. Phelan, L.M. Sanchez, N. Garg, Y. Peng, D.D. Nguyen, J. Watrous et al., Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nat. Biotechnol*, 34:8 (2016) 828–837.
57. P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks, *Genome Res.*, 13:11 (2003) 2498–2504.