Scientific Research Publishing

Health, 2024, 16, 837-848
https://www.scirp.org/journal/health
ISSN Online: 1949-5005
ISSN Print: 1949-4998

# Exploring the Unknown: The Application and Prospects of Artificial Intelligence in Genomics and Bioinformatics

**Qigang Feng[1,2,3], Jie Li[1,2], Qing Zhang[1,2,3*]**

[1]Department of Gastroenterology, First Hospital of Yangtze University, Jingzhou, China
[2]Digestive Disease Research Institution of Yangtze University, Jingzhou, China
[3]Clinical Medical College, Yangtze University, Jingzhou, China
Email: *2364420754@qq.com

**How to cite this paper:** Feng, Q.G., Li, J. and Zhang, Q. (2024) Exploring the Unknown: The Application and Prospects of Artificial Intelligence in Genomics and Bioinformatics. *Health*, **16**, 837-848.
https://doi.org/10.4236/health.2024.169059

**Received:** August 26, 2024
**Accepted:** September 20, 2024
**Published:** September 23, 2024



## Abstract

This review comprehensively explores the core application of artificial intelligence (AI) in the fields of genomics and bioinformatics, and deeply analyzes how it leads the innovative progress of science. In the cutting-edge fields of genomics and bioinformatics, the application of AI is propelling a deeper understanding of complex genetic mechanisms and the development of innovative therapeutic approaches. The precision of AI in genomic sequence analysis, coupled with breakthroughs in precise gene editing, such as AI-designed gene editors, significantly enhances our comprehension of gene functions and disease associations . Moreover, AI's capabilities in disease prediction, assessing individual disease risks through genomic data analysis, provide robust support for personalized medicine. AI applications extend beyond gene identification, gene expression pattern prediction, and genomic structural variant analysis, encompassing key areas such as epigenetics, multi-omics data integration, genetic disease diagnosis, evolutionary genomics, and non-coding RNA function prediction. Despite challenges including data privacy, algorithm transparency, and bioethical issues, the future of AI is expected to continue revolutionizing genomics and bioinformatics, ushering in a new era of personalized medicine and precision treatments.

## Keywords

AI, Genomics, Disease Prediction, Gene Editing, Multi-Omics Data Fusion

## 1. Introduction

AI is reshaping our understanding and application of genomics and bioinformatics

with its revolutionary technology. AI's powerful computational power has become a key driver of tasks such as gene sequence analysis, gene editing and disease prediction, greatly enhancing our insight into genetic mechanisms and opening new paths to innovative treatments (Figure 1(a)). In this review, we will focus on the main application of AI in the field of genomics, covering multiple aspects from gene identification and classification to expression pattern prediction, from genome structure variation analysis to gene editing, as well as personalized medicine, epigenetics analysis, multiple omics data fusion, genetic disease diagnosis, evolutionary genomics research and non-coding RNA function prediction (as shown in Figures 1(b)-(d)). We aimed to present how AI functions in these key areas and to explore its future potential in driving advances in genomic science.

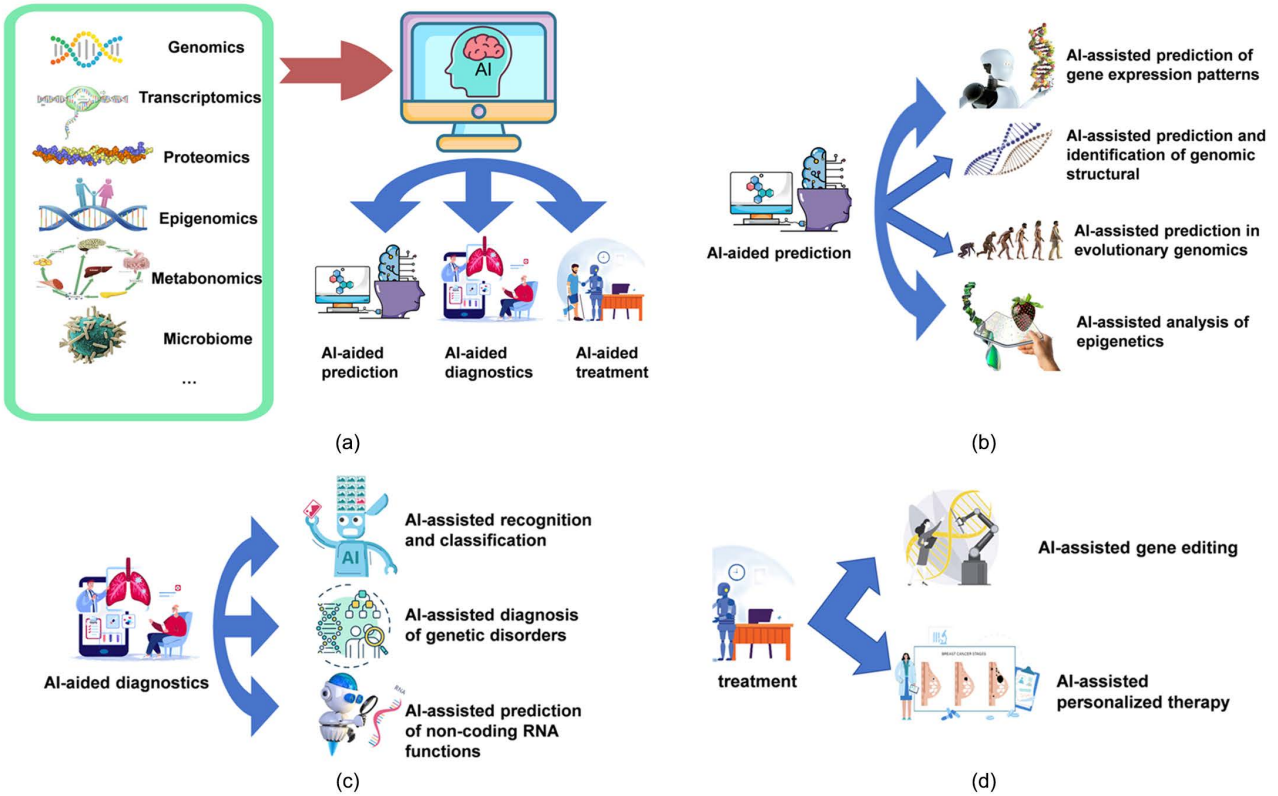


**Figure 1.** (a) Ai realizes the prediction, identification, and diagnosis of single-omics through the integration and analysis of multi-omics data; (b) Prediction: gene expression pattern, epigenetics, evolutionary genomics, non-coding RNA function; (c) Identification and diagnosis: gene sequence, genome structure variation, genetic disease; (d) Treatment: gene editing, personalized genomic medicine.

## 2. Application of AI in Genomics and Bioinformatics

### 2.1. Application of AI in Gene Sequence Identification and Classification

Deep learning has become a powerful tool in the field of gene sequence identification and classification, providing a new impetus for bioinformatics research. It helps to correct possible misannotation of long non-coding RNA (lncRNA) and

achieve remarkable results in predicting and identifying enhancer and promoter interactions (EPI) in gene sequences [1]. Deep learning also supports the analysis of RNA sequencing (RNA-Seq) technology, making processing and parsing large amounts of sequencing data more efficient [2]. Notably, some deep learning models, such as DeepECtransformer, have been successfully applied to predict enzymatic functions in microbial genomes and to classify [3]. The deep learning-based toolkit SPACEL performs well in analyzing spatial transcriptomic datasets, providing accurate 3 D tissue alignment, spatial domain identification, and batch effect elimination [4]. An important breakthrough is the ATGO method, which enables the accurate identification of protein function [5] by predicting the gene ontology (GO) properties of proteins. Deep learning also demonstrates great potential in single-cell RNA sequencing data analysis, especially in cell-type identification. The development of the novel deep learning model FLAN provides better interpretability while maintaining a high level of predictive performance [6]. Finally, The mOWL library acts as a bridge, translating biological terminology and rules into a mathematical language understandable by computers, in the form of vectors. By analyzing these vectors, mOWL can predict protein interactions and the relationship between genes and diseases, accelerating biological research and enabling scientists to uncover the secrets of the biological world more quickly [7]. These studies and applications demonstrate the diversity and profound impact of deep learning in gene sequence identification and classification, providing powerful tools for genomics research, while also bringing new opportunities for the future fields of bioinformatics and precision medicine.

### 2.2 Deep Learning Prediction of Gene Expression Patterns

Deep learning has shown remarkable breakthrough results in the prediction of gene expression patterns. First, and foremost, by using things such as Delta. Deep learning tools such as EPI have realized the comprehensive re-identification and review of enhancers in the human genome, and greatly improved the accuracy of prediction. In addition, the enhancer identification is further improved with the help of a multi-classifier stacking integration model, which is of extremely important value for deconstructing the gene expression regulatory network [8]. The DNA binding patterns of transcription factors revealed by the deep learning tool DeepTFactor give us a completely new perspective on understanding and predicting gene regulatory networks. Functional annotation of enzyme-coding genes through deep learning of fused transformer layers, while performing a genome-wide selective scan using a convolutional neural network, helps to reveal the expression pattern of genes in space [9]. In unsupervised gene expression analysis, principled feature attribution was performed by PAUSE and protein function prediction by heterogeneous network converter HNetGO. These methods and tools not only improve our understanding of gene expression and cell type, but also provide new solutions to more complex biological problems [10]. In the field of disease research, such as Parkinson's disease (PD), the combination of artificial

intelligence algorithms and biomarkers has successfully revealed different patterns of progression in the PD patient population. This deep understanding of the heterogeneity not only helps to decipher the complexity of the disease, but also provides an important reference for optimizing the clinical trial design and the development of more accurate and effective personalized treatment methods [11]. Finally, the development of a novel deep learning method, scGeneRAI, has successfully inferred the gene regulatory network [12] from static single-cell RNA sequencing data of single cells. These studies and applications fully embody the wide application and far-reaching impact of deep learning in biomedical research, and provide a new tool and perspective for future disease research and treatment.

### 2.3. The Role of AI in the Analysis of Genome Structural Variation

The application of artificial intelligence is showing its importance, especially in identifying complex genomic rearrangements, detection of copy number variations (CNVs), and analysis of short tandem duplicated repeats (STRs). The introduction of deep learning models opens a new chapter in our understanding and prediction of genetic variation. Take the study of Libiseller-Egger *et al.* [13], for example, they skillfully used deep learning models to predict cardiovascular age, and revealed the genetic basis closely related to cardiovascular age through genomic association studies (GWAS). This study has not only provided a breakthrough in identifying complex genomic rearrangements, but also provides new perspectives on our understanding of the genetic risk of cardiovascular disease. These results fully demonstrate the outstanding role of AI in resolving genome structural variation and revealing the genetic mechanisms of diseases, providing powerful tools and methods for future genomics research and precision medicine.

### 2.4. Application of AI in Gene Editing Technology

In the broad field of genomics research, the application of artificial intelligence is rapidly emerging, leading a series of innovative research paths. A striking example is PRIDICT, a deep learning model, specifically used to predict the efficiency of Prime Editing, a high-precision gene editing tool. Although the optimization process of Prime Editing requires a large amount of time and energy, the PRIDICT model has been able to effectively predict the efficiency of Prime Editing by training on a large amount of human pathological mutation data, thus providing new possibilities for the application of gene editing technology [14]. Moreover, the combination of AI and gene editing technologies has also made important breakthroughs in designing and achieving broad-spectrum disease resistance, which opens a new path for personalized gene therapy design [15]. Another tool of interest is DeepTFactor, a deep learning-based tool for predicting transcription factors. DeepTFactor The DNA binding patterns of transcription factors can be learned and predicted from a large amount of genomic data, help researchers understand the expression and regulation mechanisms of genes, and even design specific transcription factors to change the expression of specific genes [9]. These

research results not only further reveal the great potential of AI in genomics research, but also provide valuable tools and methods for future genomics research and precision medicine, indicating that AI will play an increasingly important role in the field of genomics.

### 2.5. AI-Assisted Personalized Genomic Medicine

Recently, AI and machine learning technology have shown significant application value in the field of genomic medicine, which has greatly promoted the innovation and progress in this field. For example, the algorithm NIAPU developed using machine learning technology accurately classified [16] of disease-related genes; the powerful analysis ability of machine learning is also applied to [17] in gene prediction of congenital renal tract malformation. In addition, the application of AI technology in next-generation sequencing data processing has improved the accuracy of cancer risk prediction, early diagnosis, and biomarker discovery, [18]. Studies combined with gene expression data predict the necessity of genes through machine learning, which has important implications for the identification of drug targets for cancer therapy and the understanding of genetic diseases [19]. In the field of drug interaction prediction, deep learning methods have also achieved important breakthroughs in [20]. Machine learning is further applied to predict the survival probability of triple negative breast cancer patients and revealed clinical and genetic factors closely related to their survival. Recent computational strategies bring new strategies and tools [21] to reveal associations between genes and diseases and disease gene exploration through knowledge graph embeddings and graph neural networks. These research achievements highlight the powerful potential of AI and machine learning in genomic medicine, providing new perspectives and possibilities for the future development of precision medicine.

### 2.6. Analysing Epigenetic Data with Deep Learning

The progress of deep learning in predicting gene expression patterns marks a big leap forward in this field. The researchers conducted a detailed analysis of the organization and evolution of enhancers, highlighting the influence of multiple factors on these relationships. Technological advances in machine learning and synthetic biology have provided us with new insights into the enhancer complexity [22]. In the development of lung adenocarcinoma (LUAD), the role of miRNA and methylation sites has been focused, and the investigators have developed models to predict remote metastasis in LUAD patients. Also, in parallel, the cell types in the immune microenvironment were meticulously explored, revealing [23], a key gene closely related to the progression of LUAD. Further, by applying an integrated gradient approach to resolve the inference mechanism of the deep learning tool DeepTFactor, the researchers demonstrated the ability of AI to understand the DNA binding patterns of transcription factors, even in [9] in the absence of direct training information. These achievements highlight the great potential of deep learning to reveal the laws of gene expression and drive genomics

research and precision medicine.

### 2.7. Application of AI in the Integration of Multi-Omics Data

AI is leading a new trend in the integration and analysis of multi-omics data, covering genomics, transcriptomic and proteomics data, thus promoting the progress of precision medicine. The multimodal integration (MMI) method enables AI technology to predict gene mutation status, thus further improving the accuracy of disease prediction [24]. In addition, multi-omics data affinity AI algorithms relying on graph convolution networks have been successfully applied to improve the prediction accuracy of non-small cell lung cancer (NSCLC) by integrating mRNA expression, DNA methylation and DNA sequencing data. This includes not only the training and validation of the model, but also involves using functional annotation and pathway analysis to deeply study [25], the biomarker of NSCLC. In the study of chronic obstructive pulmonary disease (COPD), the new deep learning method uses the graph convolutional neural network (ConvGNN) on the protein-protein interaction network, combining single-omics or multi-omics data, and provides a new way to reveal the molecular mechanism of COPD and find key genes and proteins related to COPD [26]. AI technology also plays an important role in the deconvolution of cell types from spatially resolved transcriptomics (SRT) data, including a new technique called SpaDecon, using semi-supervised learning that combines gene expression, spatial location, and histological information, [27]. Moreover, the application of AI in RCC pathology and genomics pioneered new pathway [28] for RCC research by identifying unique gene patterns of RCC subtypes and ranks and improving survival prediction models. Overall, the application of AI in multi-omics data integration is bringing unprecedented innovation to biomedical research and clinical medical practice.

### 2.8. Application of Machine Learning in the Diagnosis of Genetic Diseases

AI plays an indispensable role in predicting and identifying gene signature biomarkers in tumors or genetic diseases. A common strategy is to incorporate WGCNA and multiple machine learning models to mine and validate potential biomarkers. This process involves screening out differentially expressed genes (DEGs) associated with specific diseases, performing weighted gene correlation network analysis (WGCNA) and enrichment analysis, followed by a machine learning Wayne algorithm to obtain feature genes. This method has been widely used in the research of breast cancer [29], diabetic [30] [31], polycystic ovary syndrome [32], Parkinson's disease [33] [34], Alzheimer's disease [35], systemic lupus erythematosus [36], diabetic nephropathy [37], lung adenocarcinoma [38], glioma [39] and other diseases. In addition, it is able to identify the characteristic genes [40] in different parts of the same disease, or to reveal the potential hub gene [41] between the two diseases, providing new perspectives for the research and treatment of genetic diseases. Meanwhile, the independence tests based on non-

linear regression, innovative feature engineering methods, and the ontology theory of sets have also been widely used in biomedical research. These methods all showed significant effects in optimizing the search process, improving the efficiency of model trainings, and improving gene-disease association prediction [42]. In general, the application of machine learning in the diagnosis of genetic diseases is mainly reflected in the identification of genetic disease markers, the screening of rare disease variants, and the development of genetic risk assessment model, which opens up new possibilities for the diagnosis and treatment of genetic diseases.

### 2.9. Contribution of AI in Evolutionary Genomics Research

Neural network technology has proved its advantages in revealing the evolutionary relationships of species, the evolutionary process of functional genes, and the analysis of gene family history. The revolutionary tool ASDEC, by directly using raw sequence data for accurate classification, not only greatly improves the efficiency of genome-wide selective clearance detection, but also shows excellent robustness in response to the complexity of biology. ASDEC Demonexcellent performance in identifying selective clearance, pinlocating selection targets, and deeply evaluating the impact of positive selection. Its successful discovery of candidate genes on human chromosome 1 further confirms its potential application in functional gene evolution analysis. Taken together, ASDEC not only demonstrates the great ability of neural networks to deal with complex problems in biology, but also opens up a new perspective for future research in evolutionary genomics [43].

### 2.10. Application of Deep Learning in Non-Coding RNA Function Prediction

In particular, in target prediction of small RNA (miRNA) and functional annotation of long non-coding RNA (lncRNA), deep learning shows significant advantages. A new deep learning scheme called sequence pre-trained Graph Neural Network (SPGNN) is used to predict associations between lncRNA and miRNA, graphically represented [44] from RNA sequences and existing interactions. In miRNA target prediction, deep learning models combined with microfluidic technologies enable accurate and efficient detection of miRNA biomarkers, thus providing a powerful tool for early diagnosis and prognostic analysis of cancer. This method not only improves the prediction accuracy, but also greatly improves the prediction speed of [45]. In terms of lncRNA functional annotation, the deep learning model was successfully applied to identify lncRNA that might be misannotated. Through deep learning coding and training models for RNA sequences to distinguish coding and non-coding transcripts, researchers revealed some lncRNA [1] that may be misannotated as non-coding but actually have coding potential. This computational approach, which relies on nucleotide sequence, helps to reveal hidden proteomes and provides high-quality datasets for building coding

potential predictors. These results not only demonstrate the strength of deep learning in the prediction of noncoding RNA function, but also provide innovative tools and methods for future genomics research.

## 3. Challenges and Limitations

In the fusion of artificial intelligence technology in genomics and bioinformatics practice, we encountered a series of challenges: how to ensure the high quality and accessibility of data, how to meet the huge data analysis for computing resources, how to improve the AI model interpretability, how to effectively integrate multiple omics data, how to enhance the model of generalization ability of new data, how to deal with genetic diversity, how to obey ethical and legal boundaries in research, and how to promote the collaboration between different disciplines. Overcoming these challenges requires not only continuous technological innovation, improvements in data sharing policies, but also the development of new algorithms, and enhanced inter-disciplinary collaboration. Solving AI ethical issues is of paramount importance, which requires a multifaceted strategy including the development and refinement of artificial intelligence ethical guidelines and regulations, establishing accountability mechanisms, strengthening ethical oversight of research and development activities, enhancing the transparency and explainability of algorithms, building effective ethical review and supervision systems, and ensuring the responsible and sustainable development of AI technology.

## 4. Prospects and Future Direction

For the application of artificial intelligence in the field of genomics and bioinformatics challenges, the future development aimed at several core goals: improve data management and standardization, optimize computing resources, strengthen AI model interpretation, integration of multi-source biological data, enhance the model of generalization ability to adapt to genetic diversity, strictly abide by the ethics and compliance, and actively promote interdisciplinary collaboration. Through these strategic efforts, we anticipate major breakthroughs in precision medicine, personalized therapy, and gene function understanding, leading genomics into a new era of more efficient, reliable, and inclusive technologies.

## 5. Conclusion

After in-depth discussion, we can see the huge potential and vast prospects of artificial intelligence in the application field of genomics and bioinformatics. Through refined data management strategies, efficient computational resource allocation, enhanced model interpretability, integration of multi-dimensional omics data, and improvement of model generalization ability, we can expect to make revolutionary progress in precision medical diagnosis and customized chemotherapy in the future. At the same time, adhering to the ethical bottom line of ethics and legal requirements, as well as cultivating the spirit of interdisciplinary collaboration, is the key to ensure the healthy and sustainable development of this field. Overall,

AI plays a crucial role in advancing innovation in genomics and bioinformatics, and in addressing related challenges.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Nabi, A., Dilekoglu, B., Adebali, O. and Tastan, O. (2022) Discovering Misannotated Lncrnas Using Deep Learning Training Dynamics. *Bioinformatics*, **39**, btac821. https://doi.org/10.1093/bioinformatics/btac821

[2] Chen, J., Shrestha, L., Green, G., Leier, A. and Marquez-Lago, T.T. (2023) The Hitchhikers' Guide to RNA Sequencing and Functional Analysis. *Briefings in Bioinformatics*, **24**, bbac529. https://doi.org/10.1093/bib/bbac529

[3] Kim, G.B., Kim, J.Y., Lee, J.A., Norsigian, C.J., Palsson, B.O. and Lee, S.Y. (2023) Functional Annotation of Enzyme-Encoding Genes Using Deep Learning with Transformer Layers. *Nature Communications*, **14**, Article No. 7370. https://doi.org/10.1038/s41467-023-43216-z

[4] Xu, H., Wang, S., Fang, M., Luo, S., Chen, C., Wan, S., *et al.* (2023) SPACEL: Deep Learning-Based Characterization of Spatial Transcriptome Architectures. *Nature Communications*, **14**, Article No. 7603. https://doi.org/10.1038/s41467-023-43220-3

[5] Zhu, Y., Zhang, C., Yu, D. and Zhang, Y. (2022) Integrating Unsupervised Language Model with Triplet Neural Networks for Protein Gene Ontology Prediction. *PLOS Computational Biology*, **18**, e1010793. https://doi.org/10.1371/journal.pcbi.1010793

[6] Nguyen, A., Vasilaki, S. and Martínez, M.R. (2023) FLAN: Feature-Wise Latent Additive Neural Models for Biological Applications. *Briefings in Bioinformatics*, **24**, bbad056. https://doi.org/10.1093/bib/bbad056

[7] Zhapa-Camacho, F., Kulmanov, M. and Hoehndorf, R. (2022) Mowl: Python Library for Machine Learning with Biomedical Ontologies. *Bioinformatics*, **39**, btac811. https://doi.org/10.1093/bioinformatics/btac811

[8] Zhang, Y., Wang, H., Liu, J., Li, J., Zhang, Q., Tang, B., *et al.* (2023) Delta.EPI: A Probabilistic Voting-Based Enhancer-Promoter Interaction Prediction Platform. *Journal of Genetics and Genomics*, **50**, 519-527. https://doi.org/10.1016/j.jgg.2023.02.006

[9] Kim, G.B., Gao, Y., Palsson, B.O. and Lee, S.Y. (2020) DeepTFactor: A Deep Learning-Based Tool for the Prediction of Transcription Factors. *Proceedings of the National Academy of Sciences*, **118**, e2021171118. https://doi.org/10.1073/pnas.2021171118

[10] Janizek, J.D., Spiro, A., Celik, S., Blue, B.W., Russell, J.C., Lee, T., et al. (2023) PAUSE: Principled Feature Attribution for Unsupervised Gene Expression Analysis. *Genome Biology*, **24**, Article No. 81. https://doi.org/10.1186/s13059-023-02901-4

[11] Birkenbihl, C., Ahmad, A., Massat, N.J., Raschka, T., Avbersek, A., Downey, P., *et al.* (2023) Artificial Intelligence-Based Clustering and Characterization of Parkinson's Disease Trajectories. *Scientific Reports*, **13**, Article No. 2897. https://doi.org/10.1038/s41598-023-30038-8

[12] Keyl, P., Bischoff, P., Dernbach, G., Bockmayr, M., Fritz, R., Horst, D., *et al.* (2023) Single-Cell Gene Regulatory Network Prediction by Explainable AI. *Nucleic Acids Research*, **51**, e20-e20. https://doi.org/10.1093/nar/gkac1212

[13] Libiseller-Egger, J., Phelan, J.E., Attia, Z.I., Benavente, E.D., Campino, S., Friedman,

P.A., *et al.* (2022) Deep Learning-Derived Cardiovascular Age Shares a Genetic Basis with Other Cardiac Phenotypes. *Scientific Reports*, **12**, Article No. 22265. https://doi.org/10.1038/s41598-022-27254-z

[14] Mathis, N., Allam, A., Kissling, L., Marquart, K.F., Schmidheini, L., Solari, C., *et al.* (2023) Predicting Prime Editing Efficiency and Product Purity by Deep Learning. *Nature Biotechnology*, **41**, 1151-1159. https://doi.org/10.1038/s41587-022-01613-7

[15] Sha, G. and Li, G. (2023) Effector Translocation and Rational Design of Disease Resistance. *Trends in Microbiology*, **31**, 1202-1205. https://doi.org/10.1016/j.tim.2023.09.007

[16] Stolfi, P., Mastropietro, A., Pasculli, G., Tieri, P. and Vergni, D. (2023) NIAPU: Network-Informed Adaptive Positive-Unlabeled Learning for Disease Gene Identification. *Bioinformatics*, **39**, btac848. https://doi.org/10.1093/bioinformatics/btac848

[17] Kabir, M., Stuart, H.M., Lopes, F.M., Fotiou, E., Keavney, B., Doig, A.J., *et al.* (2023) Predicting Congenital Renal Tract Malformation Genes Using Machine Learning. *Scientific Reports*, **13**, Article No. 13204. https://doi.org/10.1038/s41598-023-38110-z

[18] Srivastava, R. (2022) Applications of Artificial Intelligence Multiomics in Precision Oncology. *Journal of Cancer Research and Clinical Oncology*, **149**, 503-510. https://doi.org/10.1007/s00432-022-04161-4

[19] Rosenski, J., Shifman, S. and Kaplan, T. (2023) Predicting Gene Knockout Effects from Expression Data. *BMC Medical Genomics*, **16**, Article No. 26. https://doi.org/10.1186/s12920-023-01446-6

[20] Gan, Y., Liu, W., Xu, G., Yan, C. and Zou, G. (2023) DMFDDI: Deep Multimodal Fusion for Drug-Drug Interaction Prediction. *Briefings in Bioinformatics*, **24**, bbad397. https://doi.org/10.1093/bib/bbad397

[21] Banu, A., Ahmed, R., Musleh, S., Shah, Z., Househ, M. and Alam, T. (2023) Predicting Overall Survival in METABRIC Cohort Using Machine Learning. In: *Studies in Health Technology and Informatics*, IOS Press, 632-635. https://doi.org/10.3233/shti230577

[22] Smith, G.D., Ching, W.H., Cornejo-Páramo, P. and Wong, E.S. (2023) Decoding Enhancer Complexity with Machine Learning and High-Throughput Discovery. *Genome Biology*, **24**, Article No. 116. https://doi.org/10.1186/s13059-023-02955-4

[23] Cheng, N., Liu, J., Chen, C., Zheng, T., Li, C. and Huang, J. (2023) Prediction of Lung Cancer Metastasis by Gene Expression. *Computers in Biology and Medicine*, **153**, Article 106490. https://doi.org/10.1016/j.compbiomed.2022.106490

[24] Shao, J., Ma, J., Zhang, Q., Li, W. and Wang, C. (2023) Predicting Gene Mutation Status via Artificial Intelligence Technologies Based on Multimodal Integration (MMI) to Advance Precision Oncology. *Seminars in Cancer Biology*, **91**, 1-15. https://doi.org/10.1016/j.semcancer.2023.02.006

[25] Park, M., Lim, J., Jeong, J., Jang, Y., Lee, J., Lee, J., *et al.* (2022) Deep-Learning Algorithm and Concomitant Biomarker Identification for NSCLC Prediction Using Multi-Omics Data Integration. *Biomolecules*, **12**, Article 1839. https://doi.org/10.3390/biom12121839

[26] Zhuang, Y., Xing, F., Ghosh, D., Hobbs, B.D., Hersh, C.P., Banaei-Kashani, F., *et al.* (2023) Deep Learning on Graphs for Multi-Omics Classification of COPD. *PLOS ONE*, **18**, e0284563. https://doi.org/10.1371/journal.pone.0284563

[27] Coleman, K., Hu, J., Schroeder, A., Lee, E.B. and Li, M. (2023) Spadecon: Cell-Type Deconvolution in Spatial Transcriptomics with Semi-Supervised Learning. *Communications Biology*, **6**, Article No. 378. https://doi.org/10.1038/s42003-023-04761-x

[28] Knudsen, J.E., Rich, J.M. and Ma, R. (2024) Artificial Intelligence in Pathomics and

Genomics of Renal Cell Carcinoma. *Urologic Clinics of North America*, **51**, 47-62. https://doi.org/10.1016/j.ucl.2023.06.002

[29] Mirza, Z., Ansari, M.S., Iqbal, M.S., Ahmad, N., Alganmi, N., Banjar, H., *et al.* (2023) Identification of Novel Diagnostic and Prognostic Gene Signature Biomarkers for Breast Cancer Using Artificial Intelligence and Machine Learning Assisted Transcriptomics Analysis. *Cancers*, **15**, Article 3237. https://doi.org/10.3390/cancers15123237

[30] Wang, X., Meng, L., Zhang, J., Zhao, Z., Zou, L., Jia, Z., *et al.* (2023) Identification of Ferroptosis-Related Molecular Clusters and Genes for Diabetic Osteoporosis Based on the Machine Learning. *Frontiers in Endocrinology*, **14**, Article 1189513. https://doi.org/10.3389/fendo.2023.1189513

[31] Li, W., Guo, J., Chen, J., Yao, H., Mao, R., Li, C., *et al.* (2022) Identification of Immune Infiltration and the Potential Biomarkers in Diabetic Peripheral Neuropathy through Bioinformatics and Machine Learning Methods. *Biomolecules*, **13**, Article 39. https://doi.org/10.3390/biom13010039

[32] Chen, W., Yang, Q., Hu, L., Wang, M., Yang, Z., Zeng, X., *et al.* (2023) Shared Diagnostic Genes and Potential Mechanism between PCOS and Recurrent Implantation Failure Revealed by Integrated Transcriptomic Analysis and Machine Learning. *Frontiers in Immunology*, **14**, Article 1175384. https://doi.org/10.3389/fimmu.2023.1175384

[33] Xing, N., Dong, Z., Wu, Q., Zhang, Y., Kan, P., Han, Y., *et al.* (2023) Identification of Ferroptosis Related Biomarkers and Immune Infiltration in Parkinson's Disease by Integrated Bioinformatic Analysis. *BMC Medical Genomics*, **16**, Article No, 55. https://doi.org/10.1186/s12920-023-01481-3

[34] Cai, L., Tang, S., Liu, Y., Zhang, Y. and Yang, Q. (2023) The Application of Weighted Gene Co-Expression Network Analysis and Support Vector Machine Learning in the Screening of Parkinson's Disease Biomarkers and Construction of Diagnostic Models. *Frontiers in Molecular Neuroscience*, **16**, Article 1274268. https://doi.org/10.3389/fnmol.2023.1274268

[35] Lai, Y., Lin, P., Lin, F., Chen, M., Lin, C., Lin, X., *et al.* (2022) Identification of Immune Microenvironment Subtypes and Signature Genes for Alzheimer's Disease Diagnosis and Risk Prediction Based on Explainable Machine Learning. *Frontiers in Immunology*, **13**, Article 1046410. https://doi.org/10.3389/fimmu.2022.1046410

[36] Zhao, X., Duan, L., Cui, D. and Xie, J. (2023) Exploration of Biomarkers for Systemic Lupus Erythematosus by Machine-Learning Analysis. *BMC Immunology*, **24**, Article No. 44. https://doi.org/10.1186/s12865-023-00581-0

[37] Gao, Q., Jin, H., Xu, W. and Wang, Y. (2023) Predicting Diagnostic Gene Biomarkers in Patients with Diabetic Kidney Disease Based on Weighted Gene Co Expression Network Analysis and Machine Learning Algorithms. *Medicine*, **102**, e35618. https://doi.org/10.1097/md.0000000000035618

[38] Cong, D., Zhao, Y., Zhang, W., Li, J. and Bai, Y. (2023) Applying Machine Learning Algorithms to Develop a Survival Prediction Model for Lung Adenocarcinoma Based on Genes Related to Fatty Acid Metabolism. *Frontiers in Pharmacology*, **14**, Article 1260742. https://doi.org/10.3389/fphar.2023.1260742

[39] Chen, G., He, Z., Jiang, W., Li, L., Luo, B., Wang, X., *et al.* (2022) Construction of a Machine Learning-Based Artificial Neural Network for Discriminating Panoptosis Related Subgroups to Predict Prognosis in Low-Grade Gliomas. *Scientific Reports*, **12**, Article No. 22219. https://doi.org/10.1038/s41598-022-26389-3

[40] Bao, W., Wang, L., Liu, X. and Li, M. (2023) Predicting Diagnostic Biomarkers

Associated with Immune Infiltration in Crohn's Disease Based on Machine Learning and Bioinformatics. *European Journal of Medical Research*, **28**, Article No. 255. https://doi.org/10.1186/s40001-023-01200-9

[41] Liu, J., Wu, P., Lai, S., Wang, J., Wang, J. and Zhang, Y. (2023) Identifying Possible Hub Genes and Biological Mechanisms Shared between Bladder Cancer and Inflammatory Bowel Disease Using Machine Learning and Integrated Bioinformatics. *Journal of Cancer Research and Clinical Oncology*, **149**, 16885-16904. https://doi.org/10.1007/s00432-023-05266-0

[42] Zhang, H., Yan, C., Xia, Y., Guan, J. and Zhou, S. (2023) Causal Gene Identification Using Non-Linear Regression-Based Independence Tests. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **20**, 185-195. https://doi.org/10.1109/tcbb.2022.3149864

[43] Zhao, H., Souilljee, M., Pavlidis, P. and Alachiotis, N. (2023) Genome-Wide Scans for Selective Sweeps Using Convolutional Neural Networks. *Bioinformatics*, **39**, i194-i203. https://doi.org/10.1093/bioinformatics/btad265

[44] Wang, Z., Liang, S., Liu, S., Meng, Z., Wang, J. and Liang, S. (2023) Sequence Pre-Training-Based Graph Neural Network for Predicting lncRNA-miRNA Associations. *Briefings in Bioinformatics*, **24**, bbad317. https://doi.org/10.1093/bib/bbad317

[45] Muthamilselvan, S., Ramasami Sundhar Baabu, P. and Palaniappan, A. (2023) Microfluidics for Profiling Mirna Biomarker Panels in AI-Assisted Cancer Diagnosis and Prognosis. *Technology in Cancer Research & Treatment*, **22**, Article 15330338231185284. https://doi.org/10.1177/15330338231185284