Proceedings of the 6th International Conference on Biological Engineering and Medical Science
DOI: 10.54254/2753-8818/2026.33707

Applications of Artificial Intelligence–Driven Multi-Omics in Tumor Target Identification

Xiyue Wang

Department of Integrated Traditional Chinese and Western Medicine, Liaoning University of Traditional Chinese Medicine Xinglin College, Shenyang, China
ww2178421@gmail.com

Abstract. Tumor target identification is a key issue in anti-tumor drug development and precision therapy. In recent years, the rapid development of genomics, transcriptomics, and proteomics has provided a rich data foundation for the discovery of critical tumor molecules. However, the inherent high dimensionality, heterogeneity, and complexity of omics data also pose challenges to traditional analytical methods. Artificial intelligence, with its advantages in feature extraction, heterogeneous data integration, and prioritization of candidate targets, has gradually become an important analytical tool for tumor target discovery. This paper reviews the research progress of AI-driven omics analysis in tumor target identification, focusing on the role of genomics in the identification of driver genes and driver mutations, the application of artificial intelligence in genomic feature modeling and functional annotation, as well as the main advances and key challenges faced by AI models in multi-omics integration.

Keywords: Artificial intelligence, Cancer, Target identification, Multi-omics integration

1. Introduction

Tumorigenesis involves multi-level molecular alterations, including genetic mutations, dysregulation of transcriptional control, and aberrant protein function. Accordingly, target identification has become a critical step and focus in the development of anticancer therapeutics and precision medicine [1, 2]. With the rapid development of high-throughput sequencing and mass spectrometry, genomics, transcriptomics, and proteomics have provided important information resources for identifying key tumor molecules. However, a single omics layer often reflects only local characteristics of tumor biological processes, making it difficult to fully delineate the continuous cascade from molecular abnormalities to functional alterations and ultimately to the formation of potential therapeutic targets [3]. At the same time, the high dimensionality, heterogeneity, and complexity of omics data also make traditional analysis methods face challenges [4]. In recent years, the application of artificial intelligence in high-dimensional feature extraction, heterogeneous data integration, and candidate target prioritization has been increasing, and multi-omics integration has gradually become an important research direction for tumor target discovery [5, 6]. This paper reviews the application of AI in tumor target identification, summarizes the role

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.33707

and limitations of genomics in target discovery, and outlines the key advances of various AI models in multi-omics integration and candidate target identification.

2. Genomics

2.1. The role of genomics in cancer target identification

Genomics characterizes the global landscape of tumor genomes through large-scale sequencing and data integration, identifying driver events from numerous somatic mutations and thus laying a foundation for the discovery and prioritization of candidate therapeutic targets [2, 7].

Currently, tumor genomics research and clinical molecular diagnostics predominantly employ next-generation sequencing technologies, including whole-genome sequencing (WGS), whole-exome sequencing (WES), and targeted gene panels [8, 9]. Pan-cancer cohort analyses enable the construction of cancer genomic landscapes and the delineation of shared oncogenic pathways across tumor types [10, 11].

Vogelstein et al. systematically introduced the concepts of driver mutations and passenger mutations [2]. The former confers a selective growth advantage on tumor cells and drives disease progression. A central goal of tumor genomics is to identify bona fide driver genes and map them to established signaling pathways, such as RTK-RAS-MAPK and PI3K-AKT-mTOR pathways, to determine which molecular modules are recurrently dysregulated in specific cancer types [12, 13]. Such alterations often implicate potential therapeutic targets or critical regulatory nodes within these pathways. Nevertheless, not all driver genes are amenable to direct pharmacological targeting, and their translational potential must be assessed in light of clinical evidence and druggability [7].

Although genomics provides a foundational framework for driver event identification, the extensive heterogeneity of cancer genomes means that low-frequency driver mutations, complex structural variants, and non-coding alterations remain difficult to characterize with precision. Accordingly, artificial intelligence has been increasingly applied to driver event identification, functional interpretation, and candidate target prioritization, offering the potential to enhance both analytical efficiency and detection performance.

2.2. AI-driven analysis for genomics

The general pipeline for AI-driven identification of genomic driver events involves first constructing multi-dimensional feature representations for genes or mutations from large-scale cancer genomic data, then integrating known driver events, weakly labeled samples, or statistical priors to train classification and ranking models that can prioritize candidate driver genes or mutations [14]. It is worth emphasizing that driver gene identification and driver mutation identification, while closely related, represent distinct analytical tasks [15]. While the two approaches share overlapping data sources and feature representations, the former treats genes as the unit of prediction whereas the latter targets individual variant sites. Driver gene identification does not conform to a standard binary classification paradigm; most methods therefore incorporate statistical testing, weak supervision, or network propagation strategies to mitigate the challenge of limited labeled data.

These data are mainly come from large cancer genome projects such as TCGA, ICGC, and TARGET, curated somatic mutation databases including COSMIC, and diverse clinical sequencing cohorts [16-20]. These resources supply somatic mutation calls, copy number alterations, and partial multi-omics and clinical annotation data, collectively supporting model training and evaluation while facilitating systematic investigation of driver events across cancer types.

DOI: 10.54254/2753-8818/2026.33707

In terms of feature construction, this paper summarizes these features into five categories: genomic variation features, such as point mutations, copy number variations, and sequence context, used to describe the types, frequencies, and distribution patterns of abnormalities at the DNA level in samples; functional impact features, used to evaluate the potential effects of mutations on protein structure or function [21]. Functional consequence features, used to examine whether DNA abnormalities are accompanied by observable changes at the expression, epigenetic, or protein levels; network features, used to describe the topological position of genes in protein interaction networks or signaling pathways, and to evaluate whether they are at key hub or bottleneck nodes [22]. Ontology and prior knowledge features, by integrating existing knowledge such as cancer gene databases, pathway annotations, and functional ontologies, are used to evaluate the strength of evidence of the gene's association with tumors. By combining these features, the model can describe driver genes from multiple dimensions such as structure, function, network, and knowledge.

At the method level, early studies mainly used traditional supervised learning methods such as random forest and support vector machine; later, frameworks that combine statistical evidence, frequency models, and ensemble learning were gradually developed, such as DriverML [23]. In recent years, with the wide use of protein interaction networks and pathway maps, graph learning methods have been introduced to integrate complex gene relationships and reduce the problems of insufficient labels and network noise [24]. This kind of model combines multi-dimensional features and multi-source evidence to give comprehensive scores and priority ranking for candidate genes or mutations. Its output forms include driver probability scores or statistical significance results [25]. These findings yield a prioritized candidate pool for downstream functional validation and drug target screening.

3. Multi-omics

3.1. Single-omics limitations and multi-omics integration

Single-omics analysis provides clues for target discovery from different dimensions such as genetic variation, transcriptional regulation, and protein function, but each has its own limitations [26].

At the genomic level, DNA alterations do not directly translate to protein-level functional changes, so inferring potential targets solely from gene mutations may easily overestimate the actual functional contributions of certain genes [27, 28]. Large-scale paired genome–proteome studies have shown that gene mutations and protein function are not directly correlated and the effects of some variations are mainly reflected at the level of protein abundance or modification [28]. This suggests that although genomic information can reveal potential driver sources, it is not enough to support the judgment of target functionality alone.

At the transcriptome level, RNA can reflect the activity of gene regulation and signaling pathways, but increased RNA expression does not necessarily equate to higher protein abundance, enhanced activity, or good druggability [29, 30]. Analysis based on large-scale tumor proteomics cohorts have demonstrated that RNA and protein are only moderately correlated. In addition, tumor classification results based on proteomics are not always consistent with those based on transcriptomics [30]. This shows that the transcriptome is more suitable for describing dynamic regulation and disease states, but without validation at the protein level, it may still lead to bias in target prioritization.

At the protein level, proteomics is closer to the endpoint of functional execution, and can directly reflect protein abundance, post-translational modifications, and the activation state of signaling pathways [27]. However, protein-level abnormalities do not constitute actionable targets, as not all

DOI: 10.54254/2753-8818/2026.33707

dysregulated proteins exhibit favorable druggability [31]. Whether a target is suitable for drug development still needs to be evaluated by combining factors such as structural accessibility, binding features, and safety [31, 32].

Therefore, integrating genomics, transcriptomics, and proteomics information is important for tumor target discovery. Multi-omics integration builds cross-level evidence chains that support each other, which helps to improve the reliability of disease classification, mechanism analysis, and candidate target prioritization, and it has gradually become an important direction to improve the accuracy and translational value of tumor target identification [26, 33].

3.2. AI-driven strategies and models for multi-omics integration

In multi-omics-driven target discovery, AI models serve not only to boost predictive performance but also to undertake distinct tasks at different analytical stages. Existing methods can be roughly divided into deep learning, network methods, clustering methods, feature extraction, and factorization, and their common goal is to extract disease-related information from complex, high-dimensional multi-omics data for patient stratification, biomarker selection, and drug target discovery [34]. In practical applications, appropriate models should be selected based on specific research tasks [35]. Tree models and ensemble learning methods are often used in the backend ranking stage of target discovery. When multi-omics data have been organized into structured feature matrices, these methods can score and rank candidate genes, and can stably integrate tabular information such as somatic mutations, copy number alterations, DNA methylation, and gene expression. For example, the DF-CAGE method proposed by Yang et al. uses deep forest to integrate TCGA multi-omics data and achieves the identification of druggable genes in cancer [36]. This suggests that ensemble learning offers considerable practical utility in large-scale candidate gene screening.

Deep learning models such as autoencoders are often used in the representation learning stage at the front end of the pipeline. These methods are mainly used for dimensionality reduction, denoising, and latent feature extraction of high-dimensional data, while mitigating differences in scale, noise, and distribution across omics layers and preserving complementary information. For example, in multi-omics survival prediction of breast cancer, ConcatAE and CrossAE are proposed to model omics-specific and cross-omics shared representations [37]. some studies also integrated autoencoders with similarity network fusion to analyze prostate cancer multi-omics data for biomarker identification related to prostate cancer [38].These methods mainly show the ability of AI in multi-omics feature extraction, integrated representation, and downstream risk prediction, and can inform subsequent target screening.

When the research focus shifts to molecular mechanism analysis, graph models, especially graph neural networks, show more obvious advantages. These methods can represent protein interaction networks, signaling pathways, and co-expression relationships as graph structures, enabling network-level analysis of gene functional positions and potential regulatory roles [39]. For example, GNN-SubNet maps multi-omics features to the PPI network to identify disease-related subnetworks [40]. MDMNI-DGD further integrates multi-omics data with multi-view gene networks to predict druggable genes [41]. Graph neural networks help to reveal functional relationships between genes and network topological features, thus providing clues for mechanism interpretation in target discovery.

In addition, Transformer and attention-based models are a type of method that has received attention in recent years in multi-omics integration research. These methods can more effectively capture complex interactions between different omics, and can enhance the interpretability of the

DOI: 10.54254/2753-8818/2026.33707

model by introducing pathway or regulatory network information. For example, DeePathNet combines pathway information with the Transformer architecture, and can be used for both drug response prediction and biomarker discovery [42]. However, the application of these methods in target discovery is still at a developing stage.

Overall, different AI models take different roles in multi-omics target discovery: ensemble learning is often used for candidate gene ranking, autoencoders are often used for dimensionality reduction and representation integration, graph neural networks are often used for mechanism analysis at the network level, and Transformer shows potential in modeling complex cross-modal relationships. Although these methods have advanced multi-omics-driven target identification, they still face challenges including data heterogeneity, integration complexity, and limited model generalization and translational utility. Future efforts are still needed to improve algorithm design, external validation, and clinical applicability [43].

4. Conclusion

In summary, tumor target identification is undergoing a systematic shift from single-omics descriptive analysis to AI-driven multi-omics integration. Genomics, transcriptomics, and proteomics provide key information for candidate targets from the levels of genetic variation, dynamic expression regulation, and protein functional execution, while artificial intelligence, with methods such as machine learning, deep learning, graph neural networks, and Transformer, shows clear advantages in high-dimensional feature extraction, heterogeneous data integration, network relationship analysis, and candidate target prioritization, and promotes the systematic improvement of driver event identification and target screening efficiency. However, current research still faces numerous challenges: multi-omics data heterogeneity and uneven sample quality limit model cross-cohort generalizability; poor model interpretability compromises the reliability of biological conclusions; and the translation of computational predictions into functional validation and clinical practice remains limited. In the future, it is necessary to further strengthen the construction of high-quality multi-omics data, develop models with both prediction performance and biological interpretability, and promote the deep combination of computational analysis with functional validation, pharmacological evaluation, and clinical research, so as to improve the reliability and translational value of tumor target discovery.

References

[1] Murciano-Goroff YR, Suehnholz SP, Drilon A, Chakravarty D. Precision Oncology: 2023 in Review. Cancer Discov. 2023 Dec 12; 13(12): 2525-2531. doi: 10.1158/2159-8290.CD-23-1194. PMID: 38084089; PMCID: PMC10715685.

[2] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013 Mar 29; 339(6127): 1546-58. doi: 10.1126/science.1235122. PMID: 23539594; PMCID: PMC3749880.

[3] Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics Data Integration, Interpretation, and Its Application. Bioinform Biol Insights. 2020 Jan 31; 14: 1177932219899051. doi: 10.1177/1177932219899051. PMID: 32076369; PMCID: PMC7003173.

[4] Baião AR, Cai Z, Poulos RC, Robinson PJ, Reddel RR, Zhong Q, Vinga S, Gonçalves E. A technical review of multi-omics data integration methods: from classical statistical to deep generative approaches. Brief Bioinform. 2025 Jul 2; 26(4): bbaf355. doi: 10.1093/bib/bbaf355. PMID: 40748323; PMCID: PMC12315550.

[5] Fountzilas E, Pearce T, Baysal MA, Chakraborty A, Tsimberidou AM. Convergence of evolving artificial intelligence and machine learning techniques in precision oncology. NPJ Digit Med. 2025 Jan 31; 8(1): 75. doi: 10.1038/s41746-025-01471-y. PMID: 39890986; PMCID: PMC11785769.

[6] Hsu CY, Askar S, Alshkarchy SS, et, al. AI-driven multi-omics integration in precision oncology: bridging the data deluge to clinical decisions. Clin Exp Med. 2025 Nov 21; 26(1): 29. doi: 10.1007/s10238-025-01965-9. PMID: 41266662; PMCID: PMC12634751.
[7] Dienstmann R, Jang IS, Bot B, Friend S, Guinney J. Database of genomic biomarkers for cancer drugs and clinical targetability in solid tumors. Cancer Discov. 2015 Feb; 5(2): 118-23. doi: 10.1158/2159-8290.CD-14-1118. PMID: 25656898; PMCID: PMC4825796.
[8] Isaic A, Motofelea N, Hoinoiu T, Motofelea AC, Leancu IC, Stan E, Gheorghe SR, Dutu AG, Crintea A. Next-Generation Sequencing: A Review of Its Transformative Impact on Cancer Diagnosis, Treatment, and Resistance Management. Diagnostics (Basel). 2025 Sep 23; 15(19): 2425. doi: 10.3390/diagnostics15192425. PMID: 41095644; PMCID: PMC12523276.
[9] Zalis M, Viana Veloso GG, Aguiar PN Jr, Gimenes N, Reis MX, Matsas S, Ferreira CG. Next-generation sequencing impact on cancer care: applications, challenges, and future directions. Front Genet. 2024 Jul 9; 15: 1420190. doi: 10.3389/fgene.2024.1420190. PMID: 39045325; PMCID: PMC11263191.
[10] Cooper LA, Demicco EG, Saltz JH, Powell RT, Rao A, Lazar AJ. PanCancer insights from The Cancer Genome Atlas: the pathologist's perspective. J Pathol. 2018 Apr; 244(5): 512-524. doi: 10.1002/path.5028. Epub 2018 Feb 22. PMID: 29288495; PMCID: PMC6240356.
[11] ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature. 2020 Feb; 578(7793): 82-93. doi: 10.1038/s41586-020-1969-6. Epub 2020 Feb 5. Erratum in: Nature. 2023 Feb; 614(7948): E39. doi: 10.1038/s41586-022-05598-w. PMID: 32025007; PMCID: PMC7025898.
[12] Regad T. Targeting RTK Signaling Pathways in Cancer. Cancers (Basel). 2015 Sep 3; 7(3): 1758-84. doi: 10.3390/cancers7030860. PMID: 26404379; PMCID: PMC4586793.
[13] Peng Y, Wang Y, Zhou C, Mei W, Zeng C. PI3K/Akt/mTOR Pathway and Its Role in Cancer Therapeutics: Are We Making Headway? Front Oncol. 2022 Mar 24; 12: 819128. doi: 10.3389/fonc.2022.819128. PMID: 35402264; PMCID: PMC8987494.
[14] Andrades R, Recamonde-Mendoza M. Machine learning methods for prediction of cancer driver genes: a survey paper. Brief Bioinform. 2022 May 13; 23(3): bbac062. doi: 10.1093/bib/bbac062. PMID: 35323900.
[15] Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, et, al. Comprehensive Characterization of Cancer Driver Genes and Mutations. Cell. 2018 Apr 5; 173(2): 371-385.e18. doi: 10.1016/j.cell.2018.02.060. Erratum in: Cell. 2018 Aug 9; 174(4): 1034-1035. doi: 10.1016/j.cell.2018.07.034. PMID: 29625053; PMCID: PMC6029450.
[16] Ding L, Bailey MH, Porta-Pardo E, et, al; Cancer Genome Atlas Research Network. Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell. 2018 Apr 5; 173(2): 305-320.e10. doi: 10.1016/j.cell.2018.03.033. PMID: 29625049; PMCID: PMC5916814.
[17] Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, Liang Y, Rivkin E, Wang J, Whitty B, Wong-Erasmus M, Yao L, Kasprzyk A. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database (Oxford). 2011 Sep 19; 2011: bar026. doi: 10.1093/database/bar026. PMID: 21930502; PMCID: PMC3263593.
[18] Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ, Forbes SA. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019 Jan 8; 47(D1): D941-D947. doi: 10.1093/nar/gky1015. PMID: 30371878; PMCID: PMC6323903.
[19] National Cancer Institute. Origins and Outcomes of the TARGET Childhood Cancer Program [EB/OL]. Bethesda: National Cancer Institute, 2022-06-21 [2026-03-23]. available from: https: //www.cancer.gov/ccg/research/genome-sequencing/target/about
[20] Sweet-Cordero EA, Biegel JA. The genomic landscape of pediatric cancers: Implications for diagnosis and treatment. Science. 2019 Mar 15; 363(6432): 1170-1175. doi: 10.1126/science.aaw3535. PMID: 30872516; PMCID: PMC7757338.
[21] Carter H, Karchin R. Predicting the functional consequences of somatic missense mutations found in tumors. Methods Mol Biol. 2014; 1101: 135-59. doi: 10.1007/978-1-62703-721-1_8. PMID: 24233781.
[22] Creixell P, Reimand J, Haider S, Wu G, Shibata T, Vazquez M, Mustonen V, Gonzalez-Perez A, Pearson J, Sander C, Raphael BJ, Marks DS, Ouellette BFF, Valencia A, Bader GD, Boutros PC, Stuart JM, Linding R, Lopez-Bigas N, Stein LD; Mutation Consequences and Pathway Analysis Working Group of the International Cancer Genome Consortium. Pathway and network analysis of cancer genomes. Nat Methods. 2015 Jul; 12(7): 615-621. doi: 10.1038/nmeth.3440. PMID: 26125594; PMCID: PMC4717906.
[23] Han Y, Yang J, Qian X, Cheng WC, Liu SH, Hua X, Zhou L, Yang Y, Wu Q, Liu P, Lu Y. DriverML: a machine learning algorithm for identifying driver genes in cancer sequencing studies. Nucleic Acids Res. 2019 May 7;

47(8): e45. doi: 10.1093/nar/gkz096. Erratum in: Nucleic Acids Res. 2021 Apr 19; 49(7): 4196. doi: 10.1093/nar/gkab193. PMID: 30773592; PMCID: PMC6486576.
[24] Xu J, Hao J, Liao X, Shang X, Li X. SSCI: Self-Supervised Deep Learning Improves Network Structure for Cancer Driver Gene Identification. Int J Mol Sci. 2024 Sep 26; 25(19): 10351. doi: 10.3390/ijms251910351. PMID: 39408682; PMCID: PMC11476395.
[25] Tokheim C, Karchin R. CHASMplus Reveals the Scope of Somatic Missense Mutations Driving Human Cancers. Cell Syst. 2019 Jul 24; 9(1): 9-23.e8. doi: 10.1016/j.cels.2019.05.005. Epub 2019 Jun 12. PMID: 31202631; PMCID: PMC6857794.
[26] Heo YJ, Hwa C, Lee GH, Park JM, An JY. Integrative Multi-Omics Approaches in Cancer Research: From Biological Networks to Clinical Subtypes. Mol Cells. 2021 Jul 31; 44(7): 433-443. doi: 10.14348/molcells.2021.0042. PMID: 34238766; PMCID: PMC8334347.
[27] Mani DR, Krug K, Zhang B, Satpathy S, Clauser KR, Ding L, Ellis M, Gillette MA, Carr SA. Cancer proteogenomics: current impact and future prospects. Nat Rev Cancer. 2022 May; 22(5): 298-313. doi: 10.1038/s41568-022-00446-5. Epub 2022 Mar 2. PMID: 35236940; PMCID: PMC12404316.
[28] Liu Y, Elmas A, Huang KL. Mutation impact on mRNA versus protein expression across human cancers. Gigascience. 2025 Jan 6; 14: giae113. doi: 10.1093/gigascience/giae113. PMID: 39775839; PMCID: PMC11702362.
[29] Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet. 2012 Mar 13; 13(4): 227-32. doi: 10.1038/nrg3185. PMID: 22411467; PMCID: PMC3654667.
[30] Arad G, Geiger T. Functional Impact of Protein-RNA Variation in Clinical Cancer Analyses. Mol Cell Proteomics. 2023 Jul; 22(7): 100587. doi: 10.1016/j.mcpro.2023.100587. Epub 2023 Jun 7. PMID: 37290530; PMCID: PMC10388586.
[31] Beis G, Serafeim AP, Papasotiriou I. Data-driven analysis and druggability assessment methods to accelerate the identification of novel cancer targets. Comput Struct Biotechnol J. 2022 Nov 24; 21: 46-57. doi: 10.1016/j.csbj.2022.11.042. PMID: 36514341; PMCID: PMC9732000.
[32] López-Cortés A, Cabrera-Andrade A, Echeverría-Garcés G, Echeverría-Espinoza P, Pineda-Albán M, Elsitdie N, Bueno-Miño J, Cruz-Segundo CM, Dorado J, Pazos A, Gonzáles-Díaz H, Pérez-Castillo Y, Tejera E, Munteanu CR. Unraveling druggable cancer-driving proteins and targeted drugs using artificial intelligence and multi-omics analyses. Sci Rep. 2024 Aug 21; 14(1): 19359. doi: 10.1038/s41598-024-68565-7. PMID: 39169044; PMCID: PMC11339426.
[33] Du P, Fan R, Zhang N, Wu C, Zhang Y. Advances in Integrated Multi-omics Analysis for Drug-Target Identification. Biomolecules. 2024 Jun 14; 14(6): 692. doi: 10.3390/biom14060692. PMID: 38927095; PMCID: PMC11201992.
[34] Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated Multi-Omics Analyses in Oncology: A Review of Machine Learning Methods and Tools. Front Oncol. 2020 Jun 30; 10: 1030. doi: 10.3389/fonc.2020.01030. PMID: 32695678; PMCID: PMC7338582.
[35] Cai Z, Poulos RC, Liu J, Zhong Q. Machine learning for multi-omics data integration in cancer. iScience. 2022 Jan 22; 25(2): 103798. doi: 10.1016/j.isci.2022.103798. PMID: 35169688; PMCID: PMC8829812.
[36] Yang H, Gan L, Chen R, Li D, Zhang J, Wang Z. From multi-omics data to the cancer druggable gene discovery: a novel machine learning-based approach. Brief Bioinform. 2023 Jan 19; 24(1): bbac528. doi: 10.1093/bib/bbac528. PMID: 36515158.
[37] Tong L, Mitchel J, Chatlin K, Wang MD. Deep learning based feature-level integration of multi-omics data for breast cancer patients survival analysis. BMC Med Inform Decis Mak. 2020 Sep 15; 20(1): 225. doi: 10.1186/s12911-020-01225-8. PMID: 32933515; PMCID: PMC7493161.
[38] Wang TH, Lee CY, Lee TY, Huang HD, Hsu JB, Chang TH. Biomarker Identification through Multiomics Data Analysis of Prostate Cancer Prognostication Using a Deep Learning Model and Similarity Network Fusion. Cancers (Basel). 2021 May 21; 13(11): 2528. doi: 10.3390/cancers13112528. PMID: 34064004; PMCID: PMC8196729.
[39] Valous NA, Popp F, Zörnig I, Jäger D, Charoentong P. Graph machine learning for integrated multi-omics analysis. Br J Cancer. 2024 Jul; 131(2): 205-211. doi: 10.1038/s41416-024-02706-7. Epub 2024 May 10. PMID: 38729996; PMCID: PMC11263675.
[40] Pfeifer B, Saranti A, Holzinger A. GNN-SubNet: disease subnetwork detection with explainable graph neural networks. Bioinformatics. 2022 Sep 16; 38(Suppl_2): ii120-ii126. doi: 10.1093/bioinformatics/btac478. PMID: 36124793.
[41] Li J, Li B, Zhang X, Ma X, Li Z. MDMNI-DGD: A novel graph neural network approach for druggable gene discovery based on the integration of multi-omics data and the multi-view network. Comput Biol Med. 2025 Feb;

DOI: 10.54254/2753-8818/2026.33707

185: 109511. doi: 10.1016/j.compbiomed.2024.109511. Epub 2024 Dec 6. PMID: 39644579.
[42] Cai Z, Poulos RC, Aref A, Robinson PJ, Reddel RR, Zhong Q. DeePathNet: A Transformer-Based Deep Learning Model Integrating Multiomic Data with Cancer Pathways. Cancer Res Commun. 2024 Dec 1; 4(12): 3151-3164. doi: 10.1158/2767-9764.CRC-24-0285. PMID: 39530738; PMCID: PMC11652962.
[43] Picard M, Scott-Boyer MP, Bodein A, Périn O, Droit A. Integration strategies of multi-omics data for machine learning analysis. Comput Struct Biotechnol J. 2021 Jun 22; 19: 3735-3746. doi: 10.1016/j.csbj.2021.06.030. PMID: 34285775; PMCID: PMC8258788.