Physical Education, Health and Social Sciences

https://journal-of-social-education.org

E-ISSN: 2958-5996

P-ISSN: 2958-5988

Cancer Bioinformatics: Integrating Multi-Omics Data for Precision Oncology

Mateen Muzafar[1], Sidra Riaz[2], Lawang Gurganari[3], Sohaib Usman[4], Kiran Ghafoor[5]

[1] Department of Bioinformatics, Université of Paris Scalay, Evry Campus, France
mateenmuzafar006@gmail.com

[2] Ministry of Science and Technology, Islamabad.
Corresponding Author: drsidraawan@gmail.com

[3] Provisional Institute of Teacher Education. Lawang.uob86@gmail.com

[4] Department of Bioscience, Comsats University Islamabad Sahiwal Campus.
s.usman8@outlook.com

[5] Microbiology and Molecular Genetics, University of Okara.
kiran.ghafoor.microbiologist@gmail.com

DOI: https://doi.org/10.63163/jpehss.v4i1.982

Abstract

Cancer bioinformatics has evolved as a pivotal discipline in precision oncology, shifting from single-omics analyses to integrated multi-omics approaches that encompass genomics, epigenomics, transcriptomics, proteomics, metabolomics, and radiomics. This review explores the foundational biological hierarchy, computational methodologies (including early, intermediate, and late integration strategies), and advanced artificial intelligence techniques such as variational autoencoders, graph convolutional networks, and explainable AI to decode tumor heterogeneity and therapeutic vulnerabilities. Key innovations like the CancerSD model address incomplete data challenges, while single-cell and spatial omics technologies reveal intra-tumoral dynamics and microenvironment interactions. Despite barriers in data regulation, computational infrastructure, and ethical considerations, emerging trends in federated learning, quantum computing, and digital twins promise transformative clinical applications. By synthesizing multi-omics data, this framework advances from population-based to individualized cancer care, enhancing biomarker discovery, drug response prediction, and patient outcomes.

Keywords Cancer Bioinformatics, Multi-Omics Integration, Precision Oncology, Artificial Intelligence, Deep Learning, Single-Cell Omics, Spatial Biology, Tumor Microenvironment, Data Imputation, Federated Learning

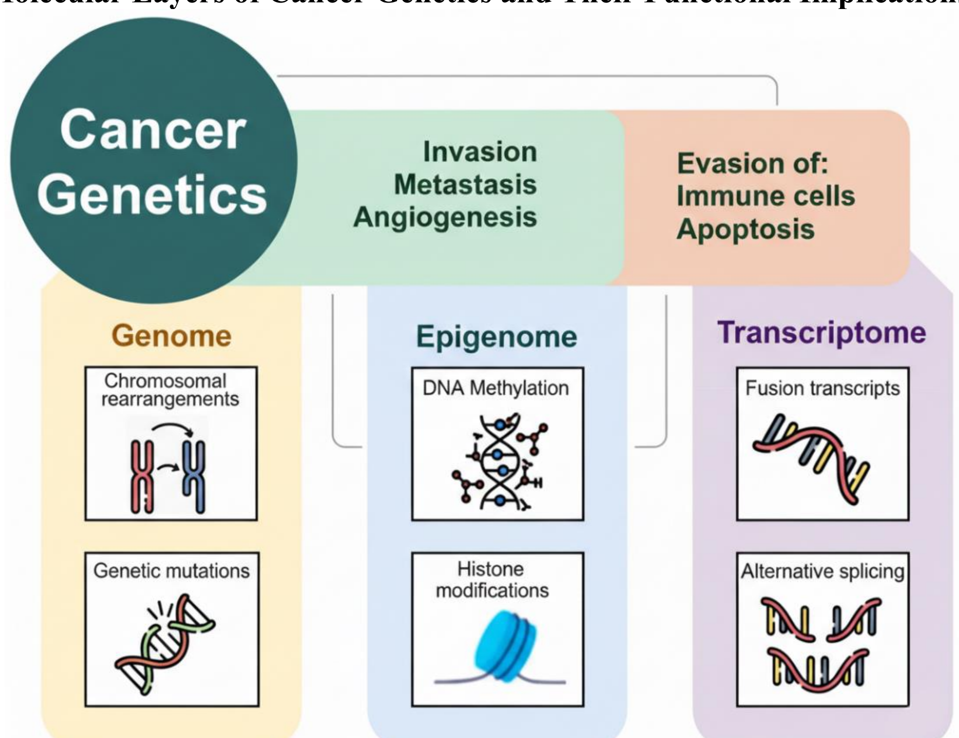
Introductions

The paradigm of oncological research and clinical practice has undergone a fundamental shift from a traditional, organ-centric classification of disease to a complex, molecularly driven framework. (Correa et al., 2022) This transition is largely fueled by the recognition that cancer is a profoundly heterogeneous disease, characterized by staggering molecular complexity that manifests across multiple biological scales. (Hsu et al., 2025) This biological complexity arises from dynamic interactions across genomic, transcriptomic, epigenetic, proteomic, and metabolomics strata, where alterations at one level such as a somatic mutation or a methylation change propagate through cellular networks to influence functional phenotypes and therapeutic vulnerabilities. (Gao et al., 2025) While single-omics approaches provided the initial map of the cancer landscape, they often offer a fragmented and sometimes contradictory view of tumor biology. (Salhi, 2025) Consequently, the field of cancer bioinformatics has increasingly focused on the integration of these disparate data layers to provide a holistic and systems-level understanding of cancer, moving precision oncology from a reactive, population-based approach toward proactive and individualized care. (Acharya et al., 2024)

Foundations of Multi-Omics and Systems Biology in Oncology

Multi-omics research represents more than a mere aggregation of datasets; it is a synergistic framework designed to decode the emergent properties of malignant systems. By integrating orthogonal molecular and phenotypic data, researchers can recover systemic signals such as spatial subclonality, microenvironment interactions, and metabolic immune crosstalk that are inherently missed by single-modality studies (Cheng et al., 2025). The promise of multi-omics integration is thus to provide a more complete perspective of complex bio systems by considering different functional levels simultaneously (Baião et al., 2025). This approach is particularly critical in understanding the metastatic tumor microenvironment, which functions as a highly dynamic ecosystem where genomic mutations, epigenetic alterations, and metabolic reprogramming converge to promote cancer cell colonization and immune escape (Ushijima et al., 2021). As shown in figure 1.1 Overview of the genomic, epigenomic, and transcriptomic alterations in cancer. The genome layer includes chromosomal rearrangements and genetic mutations; the epigenome encompasses DNA methylation and histone modifications; the transcriptome captures fusion transcripts and alternative splicing events. These molecular changes collectively drive cancer progression, including invasion, metastasis, angiogenesis, and evasion of immune surveillance and apoptosis.

Figure 1 Molecular Layers of Cancer Genetics and Their Functional Implications



The Biological Hierarchy and Molecular Strata

To appreciate the complexity of integration, one must first delineate the distinct roles of the biological layers involved. Genomics serves as the foundational stratum, identifying the genetic

characteristics and driver mutations such as single nucleotide polymorphisms (SNPs) and copy number variations (CNVs) that initiate carcinogenesis. (Frontiers in Molecular Biosciences, 2022; Sartori et al., 2025) Epigenetics adds a regulatory dimension, investigating DNA methylation and histone modifications that govern gene expression without altering the underlying sequence. These epigenetic patterns are often tissue-specific and highly sensitive to environmental factors, providing a link between the external environment and cellular behavior. (Gouru, 2025)
Transcriptomics provides a high-resolution snapshot of global gene expression patterns, revealing molecular subtypes and dynamic responses to therapy that genomic data alone may not fully capture (Zhu et al., 2023). Proteomics and metabolomics represent the functional end points of this hierarchy. Proteomics quantifies the primary effectors of cellular signaling, while metabolomics provides a real time reflection of biochemical activity, capturing the metabolic rewiring necessary for tumor growth and survival (Satrio et al., 2024). The integration of these layers allows for a panoramic dissection of driver mutations, signaling pathways, and the systemic dysregulation associated with specific cancer phenotypes (Yuan et al., 2024).

Table 1: Overview of Omics Layers, Biological Focus, Analytical Resolution, and Clinical Utility

Omics Layer	Primary Biological Focus	Analytical Resolution	Clinical Utility
Genomics	DNA sequence, mutations, CNVs	High (nucleotide-level)	Risk stratification, pan-cancer classification
Epigenomics	DNA methylation, chromatin state	High (CpG site-level)	Identifying drug targets, environmental links
Transcriptomics	RNA expression, splicing	High (isoform-level)	Subtype discovery, pathway analysis
Proteomics	Protein abundance, PTMs	Moderate (peptide-level)	Identifying signaling hubs, biomarkers
Metabolomics	Small molecules (less than or equal to 1500 Da)	Moderate (pathway-level)	Early diagnosis, metabolic-immune crosstalk
Radiomics	Imaging features (CT, MRI)	Low (tissue-level)	Non-invasive phenotype monitoring

Infrastructure and Database Ecosystems for Precision Oncology

The advancement of multi-omics is predicated on the availability of high-quality, large-scale datasets and standardized bioinformatics pipelines. Several international consortia have generated invaluable resources that serve as the bedrock for modern cancer research. (Baião et al., 2025) The Cancer Genome Atlas (TCGA) is the largest and most widely utilized pan-cancer multi-omics database for adult cancers, containing clinical annotations and sequencing data across dozens of tumor types. Similarly, the International Cancer Genome Consortium (ICGC) represents one of the most ambitious biomedical efforts since the Human Genome Project, hosting whole-genome annotated alterations from thousands of samples worldwide. (Vandereyken et al., 2023).
For pediatric oncology, the TARGET initiative and the Treehouse Childhood Cancer Initiative provide critical multi-omics datasets tailored to the unique molecular drivers of childhood malignancies (Mani et al., 2025). Beyond patient samples, resources like the Cancer Cell Line Encyclopedia (CCLE) and the Genomics of Drug Sensitivity in Cancer (GDSC) provide the multi-omics characterization of cell lines paired with drug screening data, which is essential for training models aimed at drug response prediction (Abhang et al., 2025). The Genomic Data Commons (GDC) portal serves as a centralized repository and computational platform for these datasets,

employing state of the-art bioinformatics workflows to align sequencing reads and generate high-level derived data for the research community. (Bu, 2024)

Computational Methodologies for Multi-Omics Integration

Integrating multi-omics data presents significant technical challenges, primarily due to the high dimensionality of the data, the inherent heterogeneity between different platforms, and the frequency of missing values. Bioinformatics strategies are generally categorized by the timing of their execution: early, intermediate, and late integration. (Vahabi et al., 2022)

Early Integration and Concatenation

Early integration, also known as concatenation, is the simplest approach, involving the merging of features from each omics layer into a single matrix prior to analysis. While this allows for a comprehensive initial view, it often disregards the technical heterogeneity between platforms such as differences in dynamic range or measurement units and can lead to models dominated by the noisier or more high-dimensional datasets (Baião et al., 2025). To mitigate this, normalization methods such as max-min scaling or the standardization of concatenated data (mean of zero and variance of one) are frequently applied. (EskandariNasab et al., 2024)

Intermediate and Transformation-Based Integration

Intermediate integration, or "middle integration," uses sophisticated machine learning models to consolidate data without simple concatenation. This strategy often involves transforming each omics layer into a reduced set of new variables, such as latent factors or components, which are then modeled together. (Kang et al., 2022). Canonical Correlation Analysis (CCA) and its sparse and generalized variants (sGCCA, RGCCA) are classical statistical methods used to explore relationships between two or more sets of variables by maximizing their covariance. These methods are particularly effective for joint dimensionality reduction in studies where multiple data types are collected from the same set of samples (Ray et al., 2021).

Similarity Network Fusion (SNF) is another prominent intermediate method, which uses a "message passing" approach to cluster patients (Gliozzo et al., 2025). SNF builds similarity networks for each omics type and merges them iteratively until the graphs converge into a single network that emphasizes common and complementary information while discarding noise. This is particularly useful for identifying novel cancer subtypes that may not be apparent in single-omics analysis (Chierici et al., 2020).

Late Integration and Ensemble Modeling

Late integration involves performing modeling and analysis on each omics layer separately and then merging the results or the predictive models at the end. While this strategy is robust to missing data in specific modalities and respects the unique properties of each platform, it can fail to capture

Table 2: Multi-Omics Integration Strategies: Timing, Advantages, and Limitations

Integration Strategy	Timing of Fusion	Primary Advantage	Key Limitation
Early (Concatenation)	Before analysis	Simple implementation, holistic view	Disregards platform heterogeneity
Intermediate (Joint Modeling)	During analysis	Captures inter-layer interactions	High computational complexity
Late (Ensemble)	After analysis	Robust to missing data, flexible	Ignores synergistic interactions

complex, non-linear interactions and synergies between different biological levels (Zhou et al., 2024). Despite these limitations, late integration methods like MOLI (Multi Omics Late Integration), which is based on deep neural networks, have shown high predictive power in external validations for drug response prediction. (Sartori et al., 2025)

Artificial Intelligence and Deep Learning in Multi-Omics

The integration of artificial intelligence (AI), particularly deep learning (DL), has revolutionized the field by enabling scalable, non-linear integration of disparate omics layers into clinically actionable insights. Deep learning models excel at identifying subtle, high-dimensional patterns that traditional statistical methods might miss, such as complex regulatory motifs or subtle morphological patterns in digital pathology (Jin et al., 2024).

Variational Autoencoders and Generative Models

Variational Autoencoders (VAEs) have emerged as a powerful tool for cancer bioinformatics since 2020. VAEs use encoder-decoder architectures to map high-dimensional multi omics data into a low dimensional latent space. (Rahmanian et al., 2024). These models are extensively used for data imputation, augmentation, and creating joint embeddings that facilitate cancer type classification and survival analysis. For example, the VAE-Surv approach integrates multi-omics and clinical data to predict survival outcomes and perform genetic-based clustering in complex diseases like myelodysplastic syndromes (Rollo et al., 2025).

Network-Based Deep Learning and Graph Convolutional Networks

Graph Convolutional Networks (GCNs) and other network-based deep learning methods integrate biological prior knowledge such as known protein-protein interactions directly into the model architecture (Cui et al., 2025). By modeling biological entities as nodes in a graph, GCNs can capture the spatial and functional relationships between molecular features, improving the accuracy of cancer subtype classification and driver gene identification. These models are particularly effective at revealing the molecular landscape of individual cells and the communication networks within the tumor microenvironment (Xiao et al., 2019).

Explainable AI (XAI) for Clinical Transparency

As AI models become more complex, the need for transparency and interpretability in clinical settings has grown. Explainable AI (XAI) methodologies, such as SHapley Additive exPlanations (SHAP) and LIME, are used to provide clear clinical decision support by identifying which molecular features contribute most to a given prediction (Adeniran et al., 2024). This is essential for translating "black box" deep learning models into tools that clinicians can trust for tasks like therapy selection or identifying biomarkers for chemotherapy toxicity (Kumar et al., 2024).

Tackling Incomplete and Limited Data with CancerSD

A significant bottleneck in clinical precision oncology is the difficulty of obtaining complete multi-omics datasets for all patients. Many existing algorithms require perfectly paired data across all modalities, which often leads to the exclusion of valuable clinical samples. To address this, the CancerSD model was developed as a flexible integrative framework to diagnose cancer subtypes using limited samples with incomplete data (Acharya et al., 2024).

CancerSD employs a sophisticated multi-module pipeline:

- **Incomplete Data Imputation:** Utilizes contrastive learning to extract cross-omics consistent features from available data, which are then fed into a generator to reliably impute missing modalities. (Bu, 2024)

- **Masking-and-Reconstruction:** Employs omics-specific generators with masking mechanisms to ensure that the recovered data is biologically meaningful and maintains high authenticity. (Cao et al., 2024)
- **Knowledge Transfer:** Uses a meta-learning paradigm and category-level contrastive loss to mine domain-specific knowledge from external, large-scale datasets (like TCGA), effectively pretraining the model for use with smaller clinical cohorts. (Azher et al., 2023)

Experimental results on gastric, lung, and breast cancer datasets demonstrate that CancerSD significantly outperforms existing integration methods, identifying discriminative molecules and molecular characteristics that are highly correlated with clinical phenotypes and patient prognosis. (Nakach et al., 2024)

The Single-Cell and Spatial Revolution

The transition from bulk sequencing to single-cell and spatial multi-omics resolution has provided unprecedented insights into the cellular diversity of tumors. Bulk methods often mask the genomic variations and rare cell populations that drive therapeutic resistance and disease progression (Vandereyken et al., 2023).

Resolving Intra-Tumoral Heterogeneity (ITH)

Single-cell multimodal techniques allow for the concurrent measurement of genomics, transcriptomics, and epigenomics within individual cells. (Kang et al., 2022). Modern platforms like DNTR-seq (gDNA-mRNA) and G&T-seq have revolutionized the ability to trace lineage relationships and map cell fate decisions. In neuroblastoma research, these technologies resolved how early aneuploid subclones transition into malignant states, a discovery that would be impossible with traditional methods (Krawczyk et al., 2023)

Mapping the Regulatory Landscape

Techniques such as snATAC-seq and snRNA-seq, when integrated using platforms like the 10x Multiome, allow researchers to map complex regulatory networks and chromatin remodeling events. In osteosarcoma, this integration revealed that chromatin remodeling acts as a primary driver of progression, establishing a causal link between epigenetic programming and transcriptomic output.

Spatial Biology and the Tumor Microenvironment

The emerging field of spatial multi-omics provides a physical dimension to molecular data by analyzing transcripts and proteins across specific tissue regions. (Gao et al., 2025) Tools like NanoString GeoMx Digital Spatial Profiling (DSP) can analyze over 1800 transcripts and dozens

Table 3: Single-Cell and Spatial Multi-Omics Technologies and Their Clinical Applications

Technology Platform	Integrated Modalities	Primary Application	Clinical Advantage
DNTR-seq	DNA + RNA	Resolving clonal evolution	Tracing origin of malignant cells
10x Multiome (WNN)	ATAC + RNA	Mapping regulatory networks	Identifying epigenetic drivers
RAID-seq	RNA + Proteins	Proteogenomics	Identifying functional biomarkers
GeoMx DSP	RNA + Proteins	Spatial Profiling	Mapping immune escape in TME

of proteins, revealing how the spatial distribution of immune checkpoints, such as PD-1 and CTLA-4, influences the response to immunotherapy. (Cheng et al., 2025). This "3D spatial biology" accelerates drug development and improves clinical diagnostics by characterizing the intricate crosstalk between malignant cells and the surrounding stroma (Zhu et al., 2023).

Clinical Implementation and Translational Barriers

Despite the rapid technological progress, the integration of multi-omics into routine clinical practice faces significant hurdles spanning data regulation, computational cost, and ethical equity (Mani et al., 2025).

Regulatory and Ethical Challenges

The clinical adoption of AI-powered multi-omics is often complicated by a lack of standardized imaging and sequencing protocols. Heterogeneous file formats and missing clinical metadata across different institutions make cross-site validation difficult. (Abhang et al., 2025). Furthermore, ethical hurdles around patient consent and data privacy remain a major concern. Federated Learning (FL) has emerged as a promising solution, allowing for the training of AI models on distributed datasets without the need to centralize sensitive raw patient information. (Bashir et al., 2023). The Oncology Federated Network (OFN) demonstrated a 30% improvement in predicting immunotherapy response while maintaining data sovereignty across multiple centers (Nasajpour et al., 2025).

Practical and Infrastructural Barriers

The sheer volume of data generated by modern technologies with whole-genome sequencing and spatial proteomics producing over 10 TB per patient necessitates powerful analytical tools and advanced exascale computing resources (Yuan et al., 2024). This leads to substantial financial and logistical challenges for most healthcare infrastructures. Additionally, there is an urgent need for rigorous longitudinal validation to ensure the reproducibility and real-world utility of identified biomarkers (Ushijima et al., 2021).

The Clinical Roadmap for Precision Oncology

To overcome these barriers, a comprehensive clinical roadmap is essential:

1. **Domain Integration:** Triangulating signals across genomics, epigenomics, transcriptomics, proteomics, and metabolomics to establish causal links (Srivastava et al., 2024)
2. **Computational Modeling:** Utilizing advanced machine learning and PKPD modeling to simulate drug-tumor interactions and predict resistance. (Salhi, 2025)
3. **Standardization:** Developing centralized repositories and standardized validation pipelines for multi-institutional collaboration. (Abhang et al., 2025).
4. **Patient-Centric Models:** Incorporating "N-of-1" statistical models, such as digital twins, to simulate disease progression and personalize therapy in real-time. (Vahabi et al., 2022)
5. **Future Perspectives and Emerging Trends**

The future of cancer bioinformatics is poised for a paradigm shift driven by several nascent technologies. Quantum computing holds the potential to handle the exponential complexity of multi-omics data, offering unprecedented speed in drug target identification and molecular modeling (Rahmanian et al., 2024). Federated and distributed learning models will continue to facilitate large-scale collaboration while preserving patient privacy through technologies like blockchain and homomorphic encryption (EskandariNasab et al., 2024)

The integration of multi-omics with electronic health records (EHR) and digital pathology will enable a "panoramic dissection" of cancer, linking molecular variations directly to clinical outcomes and lifestyle data. (Ray et al., 2021). As single-cell and spatial resolution become more affordable and standardized, they will move from specialized research tools into routine clinical diagnostics, providing a roadmap for highly personalized treatment strategies that anticipate resistance and optimize therapy in real time (Abhang et al., 2025).

Conclusions

Cancer bioinformatics, through the integration of multi-omics data, has fundamentally altered our understanding of carcinogenesis and therapeutic response. The shift from single-modality studies to a holistic, systems-level approach allows for the identification of complex molecular signatures and regulatory networks that define the malignant state. While the technical, regulatory, and financial barriers to clinical implementation are significant, the development of advanced AI methodologies, privacy-preserving collaborative frameworks, and high-resolution spatial technologies provides a clear trajectory for the future of oncology. By harnessing the full potential of multi-omics integration, precision medicine can finally transition from a reactive population-based science to a truly individualized, proactive discipline, improving survival rates and the quality of life for cancer patients globally. The successful translation of these bioinformatics innovations depends on continued interdisciplinary collaboration and the rigorous validation of predictive models within real-world clinical workflows.

Acknowledgment

We would like to express my deepest gratitude to my parents for their unconditional love, constant encouragement, and unwavering support throughout my academic journey. Their guidance and prayers have always been a source of strength and motivation for me. We are also sincerely thankful to my teachers, whose dedication, knowledge, and mentorship have shaped my learning and inspired me to strive for excellence. Their continuous support and valuable guidance played an essential role in the completion of this work.

References

Abhang, M. V. K., & Gunjal, B. L. (2025). Deep Graph Ensemble Convolutional Neural Networks for Drug Response Prediction from Multi-Omics Cancer Cell Lines. COMPUTER, 25(4).

Acharya, D., & Mukhopadhyay, A. (2024). A comprehensive review of machine learning techniques for multi-omics data integration: challenges and applications in precision oncology. *Briefings in functional genomics*, *23*(5), 549-560.

Acharya, D., & Mukhopadhyay, A. (2024). A comprehensive review of machine learning techniques for multi-omics data integration: challenges and applications in precision oncology. Briefings in functional genomics, 23(5), 549-560.

Adeniran, A. A., Onebunne, A. P., & William, P. (2024). Explainable AI (XAI) in healthcare: Enhancing trust and transparency in critical decision-making. World Journal of Advanced Research and Reviews, 23(3), 2647-2658.

Ali Sultan, Q., & Wahab, S. (2023). Essential oils affect the development of apricot brown rot during post-harvest storage. Horticulture, Environment, and Biotechnology, 64(4), 643-654.

Aqeel, A., Amin, A., Sultan, Q. A., Yousaf, I., & Fatima, K. MICROBIAL CONTAMINATION OF STREET FOOD VENDORS'HANDS: PREVALENCE OF STAPHYLOCOCCUS AUREUS AND OTHER PATHOGENS.

Aqeel, A., Jabeen, S., Kainat, T., Ahmed, M., Azhar, I. A., Arain, H., & Khanam, Y. (2025). ROLE OF MICROBIAL CONTAMINANTS IN THE DEVELOPMENT OF MALNUTRITION IN PEDIATRIC PATIENTS WITH GASTROINTESTINAL DISEASES. Frontier in Medical and Health Research, 3(8), 1215-1225.

Azher, Z. L., Suvarna, A., Chen, J. Q., Zhang, Z., Christensen, B. C., Salas, L. A., ... & Levy, J. J. (2023). Assessment of emerging pretraining strategies in interpretable multimodal deep learning for cancer prognostication. BioData Mining, 16(1), 23.

Baião, A. R., Cai, Z., Poulos, R. C., Robinson, P. J., Reddel, R. R., Zhong, Q., Vinga, S., & Gonçalves, E. (2025). A technical review of multi-omics data integration methods: From classical statistical to deep generative approaches. *Briefings in Bioinformatics*, *26*(4), bbaf355. https://doi.org/10.1093/bib/bbaf355

Bashir, A. K., Victor, N., Bhattacharya, S., Huynh-The, T., Chengoden, R., Yenduri, G., ... & Liyanage, M. (2023). Federated learning for the healthcare metaverse: Concepts, applications, challenges, and future directions. IEEE Internet of Things Journal, 10(24), 21873-21891.

Bu, Y. (2024). CancerSD: A flexible integrative model to diagnose cancer subtype using limited samples with incomplete multi-omics data. *PLOS Computational Biology*, *20*(12), e1012710. https://doi.org/10.1371/journal.pcbi.1012710

Cao, Y., Zhao, X., Tang, S., Jiang, Q., Li, S., Li, S., & Chen, S. (2024). scButterfly: a versatile single-cell cross-modality translation method via dual-aligned variational autoencoders. Nature Communications, 15(1), 2973.

Cheng, X., Peng, T., Chu, T., Yang, Y., Liu, J., Gao, Q., ... & Wei, J. (2025). Application of single-cell and spatial omics in deciphering cellular hallmarks of cancer drug response and resistance. Journal of Hematology & Oncology, 18, 70.

Chierici, M., Bussola, N., Marcolini, A., Francescatto, M., Zandonà, A., Trastulla, L., ... & Furlanello, C. (2020). Integrative network fusion: a multi-omics approach in molecular profiling. Frontiers in oncology, 10, 1065.

Correa-Aguila, R., Alonso-Pupo, N., & Hernandez-Rodriguez, E. W. (2022). Multi-omics data integration approaches for precision oncology. *Molecular Omics*, *18*(6), 469-479.

Cui, J., Yang, S., Yi, L., Xi, Q., Yang, D., & Zuo, Y. (2025). Recent advances in deep learning for protein-protein interaction: a review. BioData Mining, 18(1), 43.

EskandariNasab, M., Hamdi, S. M., & Boubrahimi, S. F. (2024). Impacts of data preprocessing and sampling techniques on solar flare prediction from multivariate time series data of photospheric magnetic field parameters. The Astrophysical Journal Supplement Series, 275(1), 6.

Frontiers in Molecular Biosciences. (2022). Integrating multiple biological layers of omics studies brought oncology to a new paradigm: From tumor site classification to pan-cancer molecular classification. *Frontiers in Molecular Biosciences*, *9*, 962743. https://doi.org/10.3389/fmolb.2022.962743

Gao, Q., Wu, H., Duan, L., Lv, G., & Quan, D. (2025). Multi-omics analysis of the tumor microenvironment after metastasis: Advancing toward personalized immunotherapy and molecular targeted strategies. *Frontiers in Immunology*, *16*, 1648987. https://doi.org/10.3389/fimmu.2025.1648987

Gliozzo, J., Soto Gomez, M. A., Bonometti, A., Patak, A., Casiraghi, E., & Valentini, G. (2025). miss-SNF: a multimodal patient similarity network integration approach to handle completely missing data sources. Bioinformatics, 41(4), btaf150.

Gouru, G. (2025). Precision oncology: Computational methods for multi-omics data integration to improve drug response prediction. *Cambridge Prisms: Precision Medicine*, *3*, e7. https://doi.org/10.1017/pcm.2025.10003

Hsu, C.-Y., Askar, S., Alshkarchy, S., Nayak, P., Attabi, K., Khan, M., Mayan, J., Sharma, M., Islomov, S., & Soleimani Samarkhazan, H. (2025). AI-driven multi-omics integration in precision oncology: Bridging the data deluge to clinical decisions. *Clinical and Experimental Medicine*, *26*(1), 29. https://doi.org/10.1007/s10238-025-01965-9

International Journal of Molecular Sciences. (2025). Advances in single-cell multi-omics technologies have deepened our understanding of cancer biology by integrating genomic, transcriptomic, epigenomic, and proteomic data. *International Journal of Molecular Sciences*, *26*(6), 2447. https://doi.org/10.3390/ijms26062447

Ishaque, B., Ahmad, W., Sultan, Q. A., Aqeel, A., Imran, M., Khan, K., ... & Hasan, M. (2025). HIGH-PRESSURE PROCESSING (HPP) AND ITS IMPACT ON MICROBIAL SAFETY TEXTURE OF READY–TO-EAT PRODUCTS. Pakistan Journal of Medical & Cardiological Review, 4(3), 1390-1401.

Jin, J., & Kim, J. (2024). Multi-omics and Artificial Intelligence technology for elucidating disease pathophysiology.

Kang, M., Ko, E., & Mersha, T. B. (2022). A roadmap for multi-omics data integration using deep learning. Briefings in Bioinformatics, 23(1), bbab454.

Krawczyk, E., & Kitlińska, J. (2023). Preclinical Models of Neuroblastoma—Current Status and Perspectives. Cancers, 15(13), 3314.

Kumar, A., & Metta, D. (2024). AI-driven precision oncology: predictive biomarker discovery and personalized treatment optimization using genomic data. Int J Adv Res Publ Rev, 1(3), 21-38.

Mani, S., Lalani, S. R., & Pammi, M. (2025). Genomics and multiomics in the age of precision medicine. Pediatric Research, 1-12.

Nakach, F. Z., Idri, A., & Goceri, E. (2024). A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification. Artificial Intelligence Review, 57(12), 327.

Nasajpour, M., Pouriyeh, S., Parizi, R. M., Han, M., Mosaiyebzadeh, F., Xie, Y., ... & Batista, D. M. (2025). Advances in Application of Federated Machine Learning for Oncology and Cancer Diagnosis. Information, 16(6), 487.

Rahmanian, M., & Mansoori, E. G. (2024). MoVAE: Multi-Omics Variational Auto-Encoder for Cancer Subtype Detection. IEEE Access.

Ray, P., Reddy, S. S., & Banerjee, T. (2021). Various dimension reduction techniques for high dimensional data analysis: a review. Artificial Intelligence Review, 54(5), 3473-3515.

Rollo, C., Pancotti, C., Sartori, F., Caranzano, I., D'Amico, S., Carota, L., ... & Sanavia, T. (2025). VAE-Surv: A novel approach for genetic-based clustering and prognosis prediction in myelodysplastic syndromes. Computer Methods and Programs in Biomedicine, 261, 108605.

Salhi, M. (2025). Integrated multi-omics approaches in non-small cell lung cancer for biomarker and pathway discovery. *EMJ Oncology*, *13*(1), 102–115. https://doi.org/10.33590/emjoncol/OEHS9513

Sartori, F., Codicè, F., Caranzano, I., Rollo, C., Birolo, G., Fariselli, P., & Pancotti, C. (2025). A comprehensive review of deep learning applications with multi-omics data in cancer research. *Genes*, *16*(6), 648. https://doi.org/10.3390/genes16060648

Satrio, R. D., Fendiyanto, M. H., & Miftahudin, M. (2024). Tools and techniques used at global scale through genomics, transcriptomics, proteomics, and metabolomics to investigate plant stress responses at the molecular level. In Molecular dynamics of plant stress and its management (pp. 555-607). Singapore: Springer Nature Singapore.

Shahbaz, A., Tabassum, A., Mahmood, S., Aqeel, A., Sultan, Q. A., Khanam, Y., ... & Ghafoor, K. PRODUCTION AND CHARACTERIZATION OF TANNASE ENZYME BY ASPERGILLUS NIGER UNDER SOLID STATE FERMENTATION USING FRUIT PEELINGS (ORANGE AND BANANA PEELS).

Srivastava, U., Kanchan, S., Kesheri, M., Gupta, M. K., & Singh, S. (2024). Types of omics data: genomics, metagenomics, epigenomi

Ushijima, T., Clark, S. J., & Tan, P. (2021). Mapping genomic and epigenomic evolution in cancer ecosystems. Science, 373(6562), 1474-1479.

Vahabi, N., & Michailidis, G. (2022). Unsupervised multi-omics data integration methods: a comprehensive review. Frontiers in genetics, 13, 854752.

Vandereyken, K., Sifrim, A., Thienpont, B., & Voet, T. (2023). Methods and applications for single-cell and spatial multi-omics. Nature Reviews Genetics, 24(8), 494-515.

Xiao, Z., Dai, Z., & Locasale, J. W. (2019). Metabolic landscape of the tumor microenvironment at single cell resolution. Nature communications, 10(1), 3763.

Yuan, S., Almagro, J., & Fuchs, E. (2024). Beyond genetics: driving cancer with the tumour microenvironment behind the wheel. Nature Reviews Cancer, 24(4), 274-286.

Zhou, H., Zhou, F., Zhao, C., Xu, Y., Luo, L., & Chen, H. (2024). Multimodal data integration for precision oncology: Challenges and future directions. arXiv preprint arXiv:2406.19611.

Zhu, Z., Jiang, L., & Ding, X. (2023). Advancing breast cancer heterogeneity analysis: insights from genomics, transcriptomics and proteomics at bulk and single-cell levels. Cancers, 15(16), 4164.