Multi-Omics Approaches to Biomarker Discovery: Integration of Genomics, Proteomics, And Metabolomics for Precision Medicine - A Comprehensive Review

Sandeep Thapa[1*], Santosh Khanal[2], Govardhan Joshi[3], Dilip Bhattarai[4]

[1]*Kathmandu Center for Genomics and Research Laboratory, Lalitpur, Nepal, Email:sandeep@cils.org.np, https://orcid.org/0000-0002-0543-0398*

[2]*Kathmandu Center for Genomics and Research Laboratory, Lalitpur, Nepal, Email:ask.san@gmail.com, https://orcid.org/0000-0001-7738-352X*

[3]*Kathmandu Center for Genomics and Research Laboratory, Lalitpur, Nepal, Email:gobu2001@gmail.com, https://orcid.org/0000-0003-3819-2303*

[4]*Department of Biotechnology, Himalayan WhiteHouse International College, Kathmandu, Nepal, Email: dilip.bhattarai11@gmail.com, https://orcid.org/0009-0008-7991-0360*

**Corresponding Author*

Volume 1, Issue no. 1, May 2026
Received Date: March 4, 2026
Revised Date: April 3, 2026
Accepted Date: May 3, 2026
ISSN: 3149-7144 (Print)
ISSN: 3149-7152 (Online)

Abstract

The advent of high-throughput molecular profiling technologies has revolutionized biomedical research, allowing scientists to explore complex biological systems at multiple levels:genomic, proteomic, and metabolomic. These combined "multi-omics" approaches are proving instrumental in identifying reliable biomarkers essential for the success of precision medicine. This review explores the current landscape of multi-omics integration, from technological advances to computational frameworks, and delves into their real-world applications in disease diagnostics and prognostics. We discuss how integrating diverse datasets enhances clinical insights and outline the challenges in standardization, data harmonization, and clinical translation. Emerging tools, including artificial intelligence, are highlighted for their potential to manage complexity and propel discoveries from the lab to the clinic.

Keywords:Computational frameworks, Genomics, Multi-omics integration, Precision medicine, Proteomics

1. Introduction

Precision medicine is changing the way we treat people by using their unique genetic information to guide healthcare decisions. At the heart of this approach lies the identification of biomarkers, which are also known as biological indicators that can inform disease risk, progression,therapeutic response, or emerging approaches to disease treatment (Banappagoudar et al., 2023).

Historically, biomarker research has leaned heavily on single-layer omics data, such as genomics or proteomics. However, these isolated views often fall short of capturing the intricate web of biological interactions that underpin complex diseases. (Thiele et al., 2024; Wu & Xie, 2025). As a solution, multi-omics approaches—integrating genomic (DNA), proteomic (proteins), and metabolomic (metabolites) data—offer a more holistic picture of human biology. (Luo et al., 2024).

This review presents a comprehensive look at how multi-omics data integration is transforming biomarker discovery, addressing both methodological developments and real-world clinical applications.

1.1 Objectives

- To explore foundational principles and recent advances in genomics, proteomics, and metabolomics related to biomarker identification
- To evaluate computational strategies for combining diverse omics data
- To illustrate how multi-omics has been successfully applied in disease research
- To discuss barriers to implementation and propose future directions

2. Genomic Insights into Disease Signatures

Genomics provides a foundational layer in molecular medicine, revealing DNA-level variations such as single-nucleotide polymorphisms (SNPs), insertions, deletions, and structural changes (Splus, 2001). Techniques such as whole-genome sequencing (WGS), whole-exome sequencing (WES), and genome-wide association studies (GWAS) have uncovered numerous genetic associations with diseases (Dehghan & Salehi, 2023; Tan et al., 2024).

For example, mutations in the tumor protein p53 (TP53) gene, breast cancer susceptibility genes 1 and 2 (*BRCA1/2*), and Kirsten rat sarcoma virus oncogene homolog (*KRAS*) are established biomarkers in oncology. However, genetic information alone often fails to fully account for phenotypic variability, as gene expression is influenced by epigenetics, environmental exposures, and regulatory networks—highlighting the need for a multidimensional molecular approach. (Kim et al., 2023; Singh et al., 2025)

3. Proteomics: Linking Genes to Function

Proteomics focuses on the proteins that carry out the biological instructions encoded in DNA. Unlike genomics, which is relatively static, proteomics provides dynamic insights into protein expression, post-translational modifications, and interactions under various physiological or pathological conditions. (Dogiparthi et al., 2025; Qoronfleh, 2025).

Advancements in mass spectrometry (MS)—such as tandem MS, isobaric tagging (iTRAQ, TMT), and data-independent acquisition (DIA)—have made it possible to conduct large-scale protein profiling with high precision. Clinically, protein biomarkers like prostate-specific antigen (PSA) and tau proteins have become key tools in diagnosing prostate cancer and Alzheimer’s disease, respectively. (Fröhlich et al., 2024; Moulder et al., 2018).

By combining proteomic and genomic data, researchers can better understand the downstream effects of genetic variants and identify potential therapeutic targets.

4. Metabolomics: The Final Layer of Molecular Information

Metabolomics captures the small-molecule byproducts of cellular activity, providing a snapshot of an organism’s physiological state. Techniques such as nuclear magnetic resonance (NMR) and mass spectrometry (coupled with chromatography methods like LC-MS and GC-MS) enable detailed metabolite profiling. (Koek et al., 2011; Nafie et al., 2025).

Alterations in metabolic pathways often occur early in disease development, making metabolites valuable biomarkers. For instance, elevated branched-chain amino acids have been linked to insulin resistance, while lipidomic changes can signal cardiovascular risks. Because metabolites are downstream of genomic and proteomic regulation, they offer a near-phenotypic readout that is highly responsive to both internal and external influences. (Cheng et al., 2017).

5. Integrating Multi-Omics Data

5.1 Challenges in Integration

Combining diverse omics datasets presents several hurdles:

- Variability in data formats and measurement scales
- Inconsistent sample sizes and missing values
- Batch effects introduced by different experimental protocols
- High dimensionality and low sample-to-variable ratios

5.2 Strategies for Data Integration

- **Early Integration**: Raw datasets are merged at the outset. While comprehensive, this method risks overfitting due to dimensional complexity.
- **Intermediate Integration**: Key features are extracted independently from each omics layer before merging, improving interpretability.
- **Late Integration**: Predictive models are built separately on each data type and then combined, often using ensemble learning.

5.3 Computational Approaches

- **Statistical Models**: Techniques like Canonical Correlation Analysis (CCA), Partial Least Squares (PLS), and Multi-Omics Factor Analysis (MOFA) identify correlations across data layers (Athieniti & Spyrou, 2023).
- **Network-Based Tools**: Protein interaction networks and gene regulatory maps contextualize data within known biological pathways.
- **Machine Learning**: Algorithms like random forests and support vector machines uncover complex, non-linear relationships.
- **Deep Learning**: Neural networks, including autoencoders and graph neural networks (GNNs), can model intricate, high-dimensional data structures.

Multi-omics biomarker discovery follows a sequential process that includes biological sample collection, generation of omics data, data integration, and analysis for clinical interpretation. The overall workflow of this process is shown in Figure 1. Common computational methods used for integrating multi-omics data, along with their typical applications, are listed in Table 1.

6. Real-World Applications in Disease Research

6.1 Oncology

Cancer has been at the forefront of multi-omics research. The Cancer Genome Atlas (TCGA) integrated genomics, epigenomics, transcriptomics, and proteomics to redefine tumor classification and identify therapeutic targets. In breast cancer, multi-omics integration of PAM50 gene expression panels with proteomic and metabolomic profiles has refined subtype characterization and predicted drug resistance.

In Nepal, where breast and cervical cancers are among the leading causes of female morbidity and mortality, multi-omics can provide localized molecular insights to improve early detection and treatment strategies. Additionally, hepatocellular carcinoma (HCC), linked to hepatitis B and C infections endemic in South Asia, has benefited from multi-omics analyses revealing metabolic reprogramming and immune microenvironment alterations, informing biomarker-driven therapies. (Cai et al., 2022).

6.2 Metabolic Disorders

Type 2 diabetes mellitus (T2DM) and metabolic syndrome are major public health concerns in Nepal and globally. Multi-omics integration has uncovered complex interactions among genetic predisposition (e.g., variants in TCF7L2), proteomic alterations in insulin signaling pathways, and metabolomic signatures, such as branched-chain amino acids, that predict insulin resistance and progression to diabetes. (Bazzazzadehgan et al., 2025).

Recent studies incorporating environmental factors, such as pesticide exposure common in rural Nepalese farming communities, have revealed gene-environment-metabolite networks that influence metabolic disease risk, highlighting the value of multi-omics in public health interventions. (Bugingo et al., 2025; Kafle et al., 2021).

6.3 Neurological Diseases

Neurodegenerative diseases, including Alzheimer's and Parkinson's, pose significant challenges due to their multifactorial etiology. Multi-omics studies have mapped networks involving genetic risk loci (APOE, SNCA), proteomic biomarkers (tau, alpha-synuclein), and metabolomic alterations linked to mitochondrial dysfunction and oxidative stress. (Bazzazzadehgan et al., 2025; Cocoș& Popescu, 2024).

Psychiatric disorders like schizophrenia and bipolar disorder are being explored through multi-omics lenses to identify biomarkers predictive of treatment response and disease trajectory. Such approaches are nascent but promising, particularly in LMIC contexts where mental health resources are limited. (Dhieb & Bastaki, 2025).

6.4 Infectious Diseases

Nepal's burden of infectious diseases like tuberculosis (TB), dengue, and COVID-19 highlights the urgent need for precise diagnostics and prognostics. Multi-omics integration has revealed host-pathogen interaction signatures—such as combined transcriptomic and metabolomic profiles distinguishing latent from active TB.

Proteomic identification of viral and host proteins during dengue infection has informed vaccine development and the identification of disease severity markers. The COVID-19 pandemic accelerated the global adoption of multi-omics applications, enabling the identification of immune and metabolic biomarkers predictive of severe disease and long COVID syndromes. (Agamah et al., 2024; Han et al., 2019).

6.5 cardiovascular diseases (CVD)

CVD remains the leading cause of mortality worldwide, including Nepal. Multi-omics approaches integrating genomics (e.g., variants at the 9p21 locus), proteomics (inflammatory and coagulation factors), and metabolomics (lipidomics) have advanced the understanding of atherosclerosis, heart failure, and hypertension. (Zhan et al., 2023).

Biomarkers derived from multi-omics data facilitate risk stratification and personalized therapy, including prediction of statin responsiveness, which is crucial for optimizing resource allocation in healthcare systems with limited capacity.

7. Outlook and Challenges: Implementing Multi-Omics in LMICs

7.1 Data Standardization and Sharing

The absence of uniform data generation and analysis protocols remains a major bottleneck. Adopting international standards such as Minimum Information About a Microarray Experiment. (MIAME)For microarray data, FAIR principles for data sharing are critical. Developing Nepal-specific repositories and encouraging open-access policies will enhance reproducibility and foster regional collaborations.

7.2 Infrastructure and Capacity Building

Robust computational infrastructure and trained personnel are prerequisites for multi-omics research. Nepal faces challenges, including limited high-performance computing resources and bioinformatics expertise. Strategic investments in cloud computing, regional data centers, and interdisciplinary training programs are imperative to build sustainable capabilities.

7.3 Ethical, Legal, and Social Implications (ELSI)

Given the sensitivity of omics data, especially genomics, rigorous ethical frameworks must safeguard participant privacy, ensure informed consent, and promote equitable benefit sharing. Cultural considerations unique to Nepal's diverse ethnic groups require tailored engagement strategies to build trust and avoid exploitation.

7.4 Cost and Clinical Translation

Translating multi-omics biomarkers into routine clinical practice demands cost-effective assays and a clear demonstration of clinical utility. Partnerships with government health agencies, international funders, and local stakeholders will be essential for developing accessible diagnostics tailored to Nepal's healthcare system.

7.5 Future Directions: Emerging Omics and Technologies

Expanding to transcriptomics, epigenomics, microbiomics, and single-cell analyses will deepen biological understanding. Additionally, leveraging wearable sensors, liquid biopsies, and AI-driven predictive models can revolutionize real-time disease monitoring and precision medicine delivery in resource-limited settings.

8. Conclusion

The integration of multi-omics data marks a transformative step toward unraveling complex disease mechanisms and advancing precision medicine globally. While high-resource settings continue to drive methodological innovation, it is critical to adapt these advances to the needs and constraints of LMICs such as Nepal. Addressing infrastructural, ethical, and socio-cultural challenges through capacity building, collaborative networks, and inclusive research policies will ensure that the benefits of multi-omics reach diverse populations and contribute meaningfully to global health equity.

Figure and Table

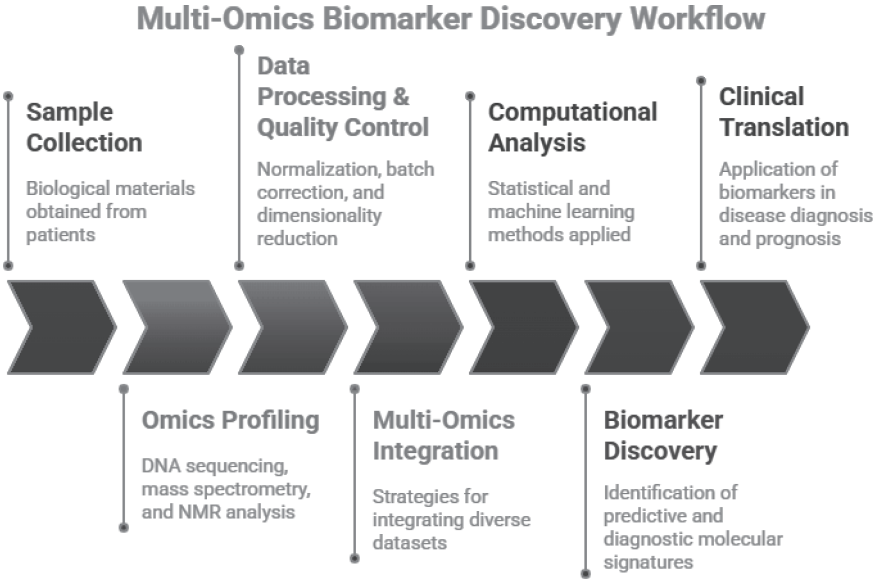

Figure 1: *Workflow of multi-omics biomarker discovery, from sample collection to clinical application.*

***Table 1:** Overview of common multi-omics integration tools and their applications.*

Tool/Method	Type	Description	Application Example
MOFA	Statistical	Latent factor modeling across omics layers	Cancer subtype classification
Random Forest	Machine Learning	Non-linear classification of multi-omics data	Diabetes biomarker prediction
Graph Neural Networks	Deep Learning	Network-based integration of omics data	Integrated biomarker discovery in cancer

Acknowledgement

The authors would like to acknowledge all colleagues for their steering support and encouragement during the preparation of this manuscript, as well as for their insightful discussions and constructive feedback.

Conflict of Interest

The authors declare no conflict of interest.

References

Agamah, F. E., Ederveen, T. H. A., Skelton, M., Martin, D. P., Chimusa, E. R., & 't Hoen, P. A. C. (2024). Network-based Multi-omics Disease–Drug Associations Reveal Drug Repurposing Candidates for Covid-19 Disease Phases. *Drug Repurposing*, *1*(1), 1–22. https://doi.org/10.58647/drugrepo.24.1.0007

Athieniti, E., & Spyrou, G. M. (2023). A guide to multi-omics data collection and integration for translational medicine. *Computational and Structural Biotechnology Journal*, *21*, 134–149. https://doi.org/10.1016/j.csbj.2022.11.050

Banappagoudar, S. B., Santhosh, S. U., Thangam, M. M. N., David, S., & Shamili, G. S. (2023). The Advantages of Precision Medicine over Traditional Medicine in Improving Public Health Outcomes. *Journal for ReAttach Therapy and Developmental Diversities*, *6*(5), 272–277.

Bazzazzadehgan, S., Shariat-Madar, Z., & Mahdi, F. (2025). Distinct Roles of Common Genetic Variants and Their Contributions to Diabetes: MODY and Uncontrolled T2DM. *Biomolecules*, *15*(3), 1–40. https://doi.org/10.3390/biom15030414

Bugingo, C., Infantino, A., Okello, P., Pérez-hernández, O., & Petrovi, K. (2025). From Morphology to Multi-Omics: A New Age of Fusarium Research. *Preprints*, 0–48.

Cai, Z., Poulos, R. C., Liu, J., & Zhong, Q. (2022). Machine learning for multi-omics data integration in cancer. *IScience*, *25*(2), 103798. https://doi.org/10.1016/j.isci.2022.103798

Cheng, S., Shah, S. H., Corwin, E. J., Fiehn, O., Fitzgerald, R. L., Gerszten, R. E., Illig, T., Rhee, E. P., Srinivas, P. R., Wang, T. J., & Jain, M. (2017). Potential Impact and Study Considerations of Metabolomics in Cardiovascular Health and Disease: A Scientific Statement from the American Heart Association. *Circulation: Cardiovascular Genetics*, *10*(2), 1–13. https://doi.org/10.1161/HCG.0000000000000032

Cocoș, R., & Popescu, B. O. (2024). Scrutinizing neurodegenerative diseases: decoding the complex genetic architectures through a multi-omics lens. *Human Genomics*, *18*(1). https://doi.org/10.1186/s40246-024-00704-7

Dehghan, R., & Salehi, M. (2023). Whole Exome Sequencing and Whole Genome Sequencing for Investigation of the Genetic Basis of Obesity: A Rapid Review. *Bahrain Medical Bulletin*, *45*(2), 1492–1497.

Dhieb, D., & Bastaki, K. (2025). Pharmaco-Multiomics: A New Frontier in Precision Psychiatry. *International Journal of Molecular Sciences*, *26*(3). https://doi.org/10.3390/ijms26031082

Dogiparthi, L. K., Bukke, S. P. N., Thalluri, C., Thalamanchi, B., Vidya, K. P., Sree, G. N., Tatiparthi, H., Uppicherla, D., & Thummaginjala, K. P. (2025). The role of genomics and proteomics in drug discovery and its application in pharmacy. In *Discover Applied Sciences* (Vol. 7, Issue 6). Springer International Publishing. https://doi.org/10.1007/s42452-025-07155-2

Fröhlich, K., Fahrner, M., Brombacher, E., Seredynska, A., Maldacker, M., Kreutz, C., Schmidt, A., & Schilling, O. (2024). Data-Independent Acquisition: A Milestone and Prospect in Clinical Mass Spectrometry–Based Proteomics. *Molecular and Cellular Proteomics*, *23*(8), 100800. https://doi.org/10.1016/j.mcpro.2024.100800

Han, L., Ao, X., Lin, S., Guan, S., Zheng, L., Han, X., & Ye, H. (2019). Quantitative Comparative Proteomics Reveal Biomarkers for Dengue Disease Severity. *Frontiers in Microbiology*, *10*(December), 1–11. https://doi.org/10.3389/fmicb.2019.02836

Kafle, S., Vaidya, A., Pradhan, B., Jørs, E., & Onta, S. (2021). Factors associated with the practice of chemical pesticide use and acute poisoning among farmers in Chitwan district, Nepal. *International Journal of Environmental Research and Public Health*, *18*(8). https://doi.org/10.3390/ijerph18084194

Kim, J., Jeong, K., Jun, H., Kim, K., Bae, J. M., Song, M. G., Yi, H., Park, S., Woo, G. un, Lee, D. W., Kim, T. Y., Lee, K. H., & Im, S. A. (2023). Mutations of TP53 and genes related to homologous recombination repair in breast cancer with germline BRCA1/2 mutations. *Human Genomics*, *17*(1), 1–10. https://doi.org/10.1186/s40246-022-00447-3

Koek, M. M., Jellema, R. H., van der Greef, J., Tas, A. C., & Hankemeier, T. (2011). Quantitative metabolomics based on gas chromatography mass spectrometry: Status and perspectives. *Metabolomics*, *7*(3), 307–328. https://doi.org/10.1007/s11306-010-0254-3

Luo, Y., Zhao, C., & Chen, F. (2024). Multiomics Research: Principles and Challenges in Integrated Analysis. *BioDesign Research*, *6*, 59. https://doi.org/10.34133/bdr.0059

Molder, R., Bhosale, S. D., Goodlett, D. R., & Lahesmaa, R. (2018). Analysis of the plasma proteome using iTRAQ and TMT-based isobaric labeling. *Mass Spectrometry Reviews*, *37*(5), 583–606. https://doi.org/10.1002/mas.21550

Nafie, M. S., Abu-Elsaoud, A. M., & Diab, M. K. (2025). A comprehensive review on computational metabolomics: Advancing multiscale analysis through in-silico approaches. *Computational and Structural Biotechnology Journal*, *27*(March), 3191–3215. https://doi.org/10.1016/j.csbj.2025.07.016

Qoronfleh, M. W. (2025). Precision Medicine and the Human Proteome in Disease, Diagnostics and Translation: Current Status and Future Prospects. *Biomedicines*, *13*(5). https://doi.org/10.3390/biomedicines13051130

Singh, S. R., Bhaskar, R., Ghosh, S., Yarlagadda, B., Singh, K. K., Verma, P., Sengupta, S., Mladenov, M., Hadzi-Petrushev, N., Stojchevski, R., Sinha, J. K., & Avtanski, D. (2025). Exploring the Genetic Orchestra of Cancer: The Interplay Between Oncogenes and Tumor-Suppressor Genes. *Cancers*, *17*(7). https://doi.org/10.3390/cancers17071082

Splus, S. (2001). R 2001 9. *Change*, 1–35.

Tan, J. K., Awuah, W. A., Ahluwalia, A., Sanker, V., Ben-Jaafar, A., Tenkorang, P. O., Aderinto, N., Mehta, A., Darko, K., Shah, M. H., Roy, S., Abdul-Rahman, T., & Atallah, O. (2024). Genes to therapy: a comprehensive literature review of whole-exome sequencing in neurology and neurosurgery. *European Journal of Medical Research*, *29*(1), 538. https://doi.org/10.1186/s40001-024-02063-4

Thiele, M., Villesen, I. F., Niu, L., Johansen, S., Sulek, K., Nishijima, S., Espen, L. Van, Keller, M., Israelsen, M., Suvitaival, T., Zawadzki, A. de, Juel, H. B., Brol, M. J., Stinson, S. E., Huang, Y., Silva, M. C. A., Kuhn, M., Anastasiadou, E., Leeming, D. J., … Melberg, H. O. (2024). Opportunities and barriers in omics-based biomarker discovery for steatotic liver diseases. *Journal of Hepatology*, *81*(2), 345–359. https://doi.org/10.1016/j.jhep.2024.03.035

Wu, Y., & Xie, L. (2025). AI-driven multi-omics integration for multi-scale predictive modeling of genotype-environment-phenotype relationships. *Computational and Structural Biotechnology Journal*, *27*(January), 265–277. https://doi.org/10.1016/j.csbj.2024.12.030

Zhan, C., Tang, T., Wu, E., Zhang, Y., He, M., Wu, R., Bi, C., Wang, J., Zhang, Y., & Shen, B. (2023). From multi-omics approaches to personalized medicine in myocardial infarction. *Frontiers in Cardiovascular Medicine*, *10*(October), 1–30. https://doi.org/10.3389/fcvm.2023.1250340