Global Journal of Biology
Vol. 1 No.1 2026 Page 1-10
E-ISSN: XXXX-XXXX, DOI: 10.XXXXX/XXXX

Integration of Transcriptomics and Proteomics in Systems Biology-Based Modeling of Gene Expression Regulation

Niken Gunawan[1]*, Rahmat Febrianti[2], Anisa Ningsih[3]
[1,2] Department of Biotechnology, School of Life Sciences and Technology, Institut Teknologi Bandung, Bandung, Indonesia
[3] Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, Indonesia

Article Info

Received : 23 January 2026
Revised : 28 March 2026
Accepted : 22 April 2026

Keyword:

Transcriptomics;
Proteomics;
Systems biology;
Gene expression regulation.

ABSTRACT

This study aims to analyze the integration of transcriptomics and proteomics in systems biology-based modeling of gene expression regulation. The issue examined in this study concerns the gap between RNA expression and protein abundance, which often creates challenges in interpreting biological mechanisms through a single omics layer. This research used a qualitative approach with an exploratory case study design. Data were collected through semi-structured interviews, limited observation, and scientific documentation involving researchers, lecturers, postgraduate students, bioinformatics analysts, and laboratory practitioners with experience in transcriptomic, proteomic, or multi-omics analysis. The data were analyzed using thematic analysis through data condensation, coding, theme development, and conclusion verification. The findings revealed five main themes: RNA-protein inconsistency, the interpretive role of computational tools, limitations in data quality and infrastructure, the need for interdisciplinary collaboration, and the construction of biological meaning from integrated data. These findings show that transcriptomics-proteomics integration is not merely a computational procedure, but also a scientific interpretation process shaped by biological reasoning, data context, and methodological decisions. This study contributes to the understanding of gene expression regulation as a multilayered and context-dependent phenomenon. The findings imply the need for stronger bioinformatics capacity, improved proteomic infrastructure, and further qualitative studies on multi-omics research practices in different scientific settings.

***) Corresponding Author:**

Niken Gunawan,
Department of Biotechnology, School of Life Sciences and Technology, Institut Teknologi Bandung, Bandung, Indonesia
Email: niken.gunawan071@gmail.com

1. Introduction

The introduction serves as a crucial opening for the research paper. It should capture the reader's attention and provide essential context for understanding the content of the research. The author(s) must clearly address the research problem, offer relevant background information, and establish their work within the existing body of research. Additionally, the introduction should justify the study's significance and outline the paper's structure and contributions.

Recent developments in molecular biology show that gene expression regulation is a complex process that cannot be explained through a single omics layer. Transcriptomics provides

information about RNA dynamics, while proteomics explains how gene products are produced, modified, and function at the protein level. Veenstra (2021) states that multi-omics approaches are an important part of systems biology because they help connect molecular changes with biological functions in a more comprehensive way. Sanches et al. (2024) show that the integration of transcriptomics, proteomics, and metabolomics can be conducted through combined modeling, correlation-based analysis, and machine learning to interpret complex biological systems. Hernández-Lemus and Ochoa (2024) also explain that multi-omics integration is important because biological phenomena involve interactions among genes, proteins, molecular pathways, and cellular environments.

A major problem in the study of gene expression regulation arises when RNA data do not always align with protein data. Increased mRNA expression does not always produce an increase in protein abundance in the same direction or magnitude. This condition occurs because gene expression does not end at transcription. It continues through translation, protein degradation, post-translational modification, and molecular stability regulation. Zhou et al. (2020) show that cell surface protein abundance cannot always be predicted directly from transcriptomic data. Lakkis et al. (2022) developed a deep learning approach to bridge CITE-seq and single-cell RNA-seq data to improve cell surface protein prediction. Chen et al. (2025) further support this direction through a transformer-based model designed to map the relationship between RNA expression and protein expression at the single-cell level.

In the Indonesian context, the integration of transcriptomics and proteomics is important because Indonesia has high biodiversity and human genetic diversity. This diversity requires gene regulation research to consider local biological contexts, rather than relying only on general models developed from specific populations or organisms. Natri et al. (2022) show that the genetic architecture of gene regulation in Indonesian populations is related to global and local ancestry variation. Ibeh et al. (2024) also identify alternative splicing variation in samples from the Indonesian archipelago, which indicates that post-transcriptional regulation plays an important role in understanding gene function in diverse populations. However, national studies still tend to emphasize genomics and transcriptomics, while integration with proteomics within a systems biology framework remains underexplored.

Empirical findings from various fields show that the integration of transcriptomics and proteomics can reveal regulatory mechanisms that may not be visible through a single-layer approach. Li et al. (2022) demonstrate that integrated analysis of *Myricaria laxiflora* can explain molecular responses to flooding stress through dynamic changes in genes and proteins. Zhao et al. (2023) find that transcriptomics-proteomics integration can explain the relationship between alternative polyadenylation and the translation of inflammation-related proteins in diabetic nephropathy. Liu et al. (2024) also show that integrated analysis can identify molecular phenotypic alterations associated with colorectal cancer. These findings indicate that omics integration is not only an analytical technique but also a way to understand biological processes more meaningfully.

Although multi-omics research continues to grow rapidly, there remains a gap in studies that examine meaning-making, interpretation processes, and biological model construction from integrated transcriptomic and proteomic data. Many previous studies focus on algorithm accuracy, biomarker identification, molecular pathway mapping, and statistical validation. Argelaguet et al. (2020) introduce MOFA+ as a framework for multimodal data integration, but integrated results still require contextual biological interpretation. Sidhaye et al. (2023) show that transcriptome and proteome integration can reveal post-transcriptional regulation in human brain organoids. Patel et al. (2024) also explain that the relationship between the transcriptome and proteome can be understood through modularization, although this relationship is not always direct at the single-gene level. Sibilio et al. (2025) emphasize that multi-omics integration still faces major challenges because the data have high dimensionality, different scales, and strong biological heterogeneity.

Based on this background, this study aims to analyze the integration of transcriptomics and proteomics in systems biology-based modeling of gene expression regulation through a qualitative approach. The study focuses on scientific meaning-making processes, data interpretation patterns, integration challenges, and the contribution of multi-omics approaches in explaining the relationship among RNA, proteins, and biological functions. Theoretically, this study is expected to strengthen the understanding that gene expression regulation is a multilayered process involving transcription, post-transcriptional regulation, translation, and post-translational regulation. Practically, this study can provide a foundation for researchers, educators, and bioinformatics developers in designing integrative models that are more contextual, transparent, and relevant to the development of systems biology in Indonesia.

2. Materials and Methods

This study uses a qualitative approach with an exploratory case study design. This approach was chosen because the study does not aim to test statistical relationships between transcriptomic and proteomic variables. Instead, it seeks to understand the processes, experiences, and scientific considerations used by researchers in integrating these two omics data types. Lim (2025) explains that qualitative research is appropriate when researchers aim to understand meanings, contexts, processes, and experiences within a phenomenon. An exploratory case study design was selected because the integration of transcriptomics and proteomics in gene expression regulation modeling is still a developing research practice. Miller et al. (2023) state that qualitative case studies can be used to understand a phenomenon in depth through multiple and complementary data sources.

The research setting is located within molecular biology, bioinformatics, and omics research environments in Indonesia that have experience in using or interpreting transcriptomic and proteomic data. These settings include academic laboratories, bioinformatics research groups, omics research communities, and online scientific discussion spaces related to multi-omics analysis. The research is designed to be conducted over three months, covering instrument preparation, data collection, data validation, data analysis, and report writing. The research setting is flexible because the object of study is not limited to physical laboratories. It also includes data analysis practices, scientific documentation, bioinformatics repositories, and academic communication among researchers. Dahal (2025) emphasizes that qualitative analysis must consider the relationship between data, context, experience, and interpretation processes that develop during research.

The research subjects consist of researchers, lecturers, postgraduate students, bioinformatics analysts, and laboratory practitioners who have experience in transcriptomic analysis, proteomic analysis, or multi-omics integration. Informants are selected using purposive sampling based on several criteria. First, informants must have at least one year of experience in omics research. Second, informants must have been involved in RNA-seq data analysis, mass spectrometry-based proteomics, or biological network modeling. Third, informants must be willing to provide information openly according to the research needs. Campbell et al. (2020) explain that purposive sampling helps researchers select participants who are relevant to the research objectives. If the initial data do not show sufficient variation in experience, the researcher will use snowball sampling by asking initial participants to recommend additional informants.

Data are collected through semi-structured interviews, limited observation, and scientific documentation. Semi-structured interviews are used to explore informants' experiences in understanding the relationship between mRNA expression and protein abundance, their strategies for choosing integration methods, and the challenges they face when transcriptomic results do not align with proteomic results. The interviews also explore how informants assess the biological validity of gene expression regulation models. Limited observation is conducted on data analysis workflows, team discussions, or demonstrations of bioinformatics procedures used by informants. Documentation includes scientific articles, analytical protocols, workflow notes, data integration diagrams, training documents, and technical records related to gene expression regulation

modeling. Morgan (2022) states that document analysis can serve as an important data source because it helps researchers identify patterns, contexts, and meanings from available documents.

The main research instrument is the researcher, supported by an interview guide, observation sheet, and document analysis matrix. The interview guide is developed based on the research focus, including informants' understanding of transcriptomics and proteomics integration, their experience with analytical tools, their methods for assessing biological model validity, and the technical and conceptual challenges in developing systems biology-based models. The observation sheet is used to record activities, technical terms, decision-making processes, and discussion patterns that emerge when informants explain their analytical procedures. The document analysis matrix is used to record document type, source, year, content focus, relevance to gene expression regulation, and contribution to research findings. Moilanen et al. (2022) explain that document analysis requires clear selection strategies, extraction matrices, preliminary testing, and credibility considerations so that written data can be analyzed systematically.

Data validation is conducted through source triangulation, method triangulation, member checking, peer debriefing, and an audit trail. Source triangulation is conducted by comparing data from laboratory researchers, bioinformatics analysts, lecturers, and postgraduate students. Method triangulation is conducted by comparing interview results, observation notes, and documents. Member checking is carried out by sending interview summaries or preliminary interpretations to informants to ensure that the researcher's interpretation reflects their experiences accurately. Peer debriefing is conducted through discussions with colleagues who understand qualitative methodology or bioinformatics. The audit trail is maintained by keeping records of the research process, coding decisions, theme revisions, and researcher reflections. Ahmed (2024) states that credibility, transferability, dependability, and confirmability can be strengthened through triangulation, rigorous documentation, member checking, and a clear audit trail.

Data analysis uses thematic analysis with the interactive model of Miles, Huberman, and Saldaña, which includes data condensation, data display, and conclusion drawing and verification. The first stage involves transcribing interview data, rereading observation notes, and selecting relevant documents. The second stage uses open coding to identify units of meaning, such as "RNA-protein inconsistency," "selection of integration algorithms," "biological validation," "database limitations," and "interpretation of gene regulatory networks." The third stage groups codes into broader categories and develops main themes that answer the research focus. Braun and Clarke (2022) emphasize that thematic analysis must be conducted consciously, consistently, and in line with the research objectives. Ahmed et al. (2025) also explain that thematic analysis provides a flexible framework for identifying, reviewing, naming, and reporting patterns of meaning in qualitative data.

The analysis is conducted iteratively until thematic saturation is reached. Saturation is indicated when additional interviews, observations, and documents no longer produce significant new codes or themes. Hennink and Kaiser (2022) show that qualitative studies with specific objectives can reach saturation with a limited number of interviews, especially when participants have relevant and focused experiences. Olmos-Vega et al. (2022) emphasize that reflexivity is important because qualitative researchers need to recognize how their position, experience, and methodological decisions influence the research process. Therefore, the researcher keeps.

3. Results and Discussions

The results and discussions section is the main part of the paper. The author(s) meticulously present their findings, providing a comprehensive analysis of the data collected. The author(s) elucidates the significant trends and patterns observed, correlating them with the hypotheses stated earlier in the study. The discussion integrates these findings with existing literature, highlighting agreements and discrepancies with prior research. The author(s) delve into the

implications of the results, considering both theoretical and practical perspectives. Furthermore, the author(s) should address potential limitations and suggest areas for future research to build upon their work, ensuring a holistic understanding of the topic under investigation.

The findings of this study show that the integration of transcriptomics and proteomics was understood by participants as a necessary step in interpreting gene expression regulation more comprehensively. Participants did not view transcriptomic data as a final explanation of biological activity, but as an initial layer that requires confirmation through protein-level analysis. One bioinformatics researcher stated, "RNA data can tell us that a gene is active, but it does not always tell us whether the protein really works in the system." This statement reflects a shared understanding among participants that gene expression regulation involves multiple biological stages, including transcription, translation, protein degradation, and post-translational modification. Documentation from workflow notes and published protocols also showed that researchers often began their analysis with RNA-seq results, then used proteomic data to validate whether the observed transcriptional patterns had functional relevance at the protein level.

The first major theme that emerged from the data was the gap between transcript abundance and protein abundance. Participants repeatedly described this gap as one of the most important challenges in systems biology-based modeling. In several cases, genes with high mRNA expression did not show similar protein abundance, while some proteins appeared more biologically relevant despite limited transcript-level signals. One laboratory practitioner explained, "Sometimes the transcript looks promising, but when we check the protein data, the pattern is not strong. That is where interpretation becomes difficult." This finding shows that researchers experience RNA-protein inconsistency not only as a technical issue, but also as an interpretive problem. The inconsistency forces researchers to reconsider whether a biological signal should be treated as meaningful, uncertain, or dependent on additional regulatory mechanisms.

The second theme was the role of computational tools as interpretive instruments. Participants emphasized that software, algorithms, and statistical models were useful for organizing large omics datasets, but they could not replace biological reasoning. Several informants described the use of correlation analysis, pathway enrichment, network modeling, and machine learning as common strategies in integrating transcriptomic and proteomic data. However, they also stressed that computational output must be interpreted carefully. One participant stated, "The software can produce clusters and networks, but the researcher still needs to decide whether the result makes biological sense." This finding indicates that modeling gene expression regulation requires both computational competence and biological judgment. The model is not treated as a purely technical product, but as a scientific interpretation shaped by data quality, analytical assumptions, and domain knowledge.

The third theme was the importance of data quality, database availability, and contextual limitations. Participants reported that transcriptomic datasets were often easier to access than proteomic datasets, especially in research environments with limited laboratory infrastructure. Proteomic analysis was viewed as more expensive, technically demanding, and dependent on specialized equipment such as mass spectrometry. One postgraduate student noted, "We can obtain RNA-seq data from public repositories, but matched proteomic data are harder to find, especially for local samples." This finding is relevant to the Indonesian research context, where the development of omics research still depends on access to facilities, funding, training, and international databases. Documentation analysis also showed that many available datasets were produced outside Indonesia, which limits the direct relevance of some models for local biological populations, biodiversity, or disease contexts.

The fourth theme was interdisciplinary collaboration. Participants stated that integration of transcriptomics and proteomics requires collaboration among molecular biologists, bioinformaticians, statisticians, clinicians, and domain experts. The complexity of multi-omics

data made it difficult for one researcher to master all analytical and interpretive stages. One lecturer explained, "A good model needs people who understand the biology, the computation, and the disease or organism being studied." This statement shows that systems biology is not only a methodological approach, but also a collaborative research culture. Observational notes from online discussions supported this finding, as researchers often negotiated the meaning of results through discussion, comparison with prior studies, and repeated checking of pathway databases.

The fifth theme concerned the construction of biological meaning from integrated data. Participants did not define successful integration only through strong statistical correlation between RNA and protein. Instead, they viewed integration as successful when it helped explain a biological process more clearly. For example, integrated data were considered meaningful when they clarified why certain genes were active, how proteins supported a specific pathway, or why a disease-related mechanism appeared at a particular molecular stage. One informant stated, "The important thing is not only whether the numbers match, but whether the model can explain the biological process." This finding shows that qualitative meaning-making plays an important role in systems biology-based modeling. Researchers must move from numerical patterns to conceptual explanations that can be communicated to other scientists.

Overall, the findings indicate that the integration of transcriptomics and proteomics is experienced as both a technical and interpretive process. The main patterns that emerged from the data include RNA-protein inconsistency, the need for computational-biological interpretation, limitations of data infrastructure, interdisciplinary collaboration, and the construction of biological meaning. These patterns show that systems biology-based modeling of gene expression regulation cannot rely only on algorithmic procedures. It also requires contextual judgment, methodological transparency, and critical reflection on how molecular data are transformed into biological explanations.

Discussion

The findings support the view that gene expression regulation is a multilayered process that cannot be fully explained through transcriptomic data alone. Participants consistently viewed transcriptomics as an entry point, while proteomics was seen as an important layer for functional confirmation. This finding is in line with Veenstra (2021), who explains that multi-omics approaches strengthen systems biology because they connect molecular changes with biological functions. The result also supports Sanches et al. (2024), who argue that the integration of transcriptomics, proteomics, and other omics layers can reveal biological mechanisms that remain hidden in single-layer analysis. In this study, participants showed that RNA data were useful, but they needed protein-level interpretation to build a more credible model of gene expression regulation.

The finding on RNA-protein inconsistency is also consistent with previous empirical studies. Zhou et al. (2020) found that cell surface protein abundance cannot always be predicted directly from transcriptomic data. This aligns with participants' experiences that high mRNA expression does not always produce a similar protein-level pattern. Lakkis et al. (2022) responded to this problem by developing a deep learning approach to improve cell surface protein prediction from single-cell transcriptomic data. Chen et al. (2025) also advanced this field through transformer-based modeling for joint analysis of single-cell transcriptomics and proteomics. The present study adds a qualitative perspective by showing that RNA-protein inconsistency is not only a computational problem, but also a problem of scientific interpretation.

The findings also show that computational tools play an important role in data integration, but they require biological interpretation. This result supports Argelaguet et al. (2020), who introduced MOFA+ as a framework for integrating multimodal single-cell data. However, this study shows that the value of such tools depends on how researchers interpret the output in relation to biological theory, data quality, and research objectives. Hernández-Lemus and Ochoa (2024) emphasize that multi-omic integration faces challenges related to scale, dimensionality,

and heterogeneity. The participants in this study experienced these challenges directly when they had to decide whether a computational pattern represented a valid biological mechanism or only a statistical signal.

The Indonesian context adds an important dimension to the discussion. Participants highlighted that data availability, research infrastructure, and access to proteomic technology remain practical challenges. This finding is relevant because Indonesian biological and genetic diversity requires stronger local omics research. Natri et al. (2022) showed that gene regulation in Indonesian populations is shaped by global and local ancestry variation. Ibeh et al. (2024) also demonstrated that alternative splicing patterns in Indonesian populations reveal important post-transcriptional variation. These studies support the argument that local biological context matters in gene regulation research. The present study extends this argument by showing that local context also affects how researchers access data, choose methods, and interpret integrated omics results.

The findings on interdisciplinary collaboration support the core principle of systems biology. Participants viewed gene expression modeling as a collaborative process that requires expertise from molecular biology, bioinformatics, statistics, and applied biological fields. This finding is consistent with Patel et al. (2024), who showed that transcriptome-proteome relationships can be understood through modularized biological organization. Such modular interpretation requires both computational modeling and biological explanation. Sidhaye et al. (2023) also showed that integrated transcriptome and proteome analysis can reveal post-transcriptional regulation in human brain organoids. However, this study adds that the development of such interpretation depends on collaborative dialogue among researchers, not only on the availability of advanced datasets.

The results further suggest that successful integration should not be measured only by strong correlation between RNA and protein. Participants emphasized that integration becomes meaningful when it helps explain a biological process. This finding offers a perspective that complements quantitative multi-omics studies. Li et al. (2022) showed that integrated transcriptomic and proteomic analysis can explain flooding stress response in *Myricaria laxiflora*. Zhao et al. (2023) demonstrated that transcriptomics-proteomics integration can reveal inflammation-related protein translation in diabetic nephropathy. Liu et al. (2024) also identified molecular phenotypic alterations in colorectal cancer through integrated analysis. These studies show the biological value of integration, while the present study explains how researchers interpret that value during the modeling process.

Theoretically, this study contributes to the understanding of gene expression regulation as a layered and interpretive biological process. It supports the systems biology view that biological function emerges from interactions among molecular components rather than from isolated genes. The findings also strengthen the argument that qualitative research can contribute to bioinformatics and molecular biology by examining how researchers make decisions, interpret uncertainty, and construct meaning from complex data. Sibilio et al. (2025) noted that multi-omics integration in complex diseases faces major methodological challenges due to high-dimensional and heterogeneous data. This study adds that these challenges also require interpretive awareness, methodological reflexivity, and transparent decision-making.

Practically, the findings indicate that researchers and institutions need to strengthen training in both computational analysis and biological interpretation. Bioinformatics education should not only teach software operation, but also help researchers understand the assumptions, limits, and interpretive consequences of analytical models. Research institutions in Indonesia also need to improve access to proteomic facilities, develop local omics databases, and encourage collaboration across disciplines. These steps are important because multi-omics integration will become more relevant for biomedical research, biodiversity studies, agriculture, and precision health.

Future studies can expand this research by involving more informants from different institutional backgrounds, including clinical researchers, agricultural scientists, and computational biologists. Further research can also compare qualitative experiences across countries or research environments to understand how infrastructure and scientific culture affect multi-omics integration. In addition, future studies may combine qualitative interviews with bibliometric analysis, workflow analysis, or case comparison of specific transcriptomics-proteomics projects. Such studies can deepen understanding of how integrated omics models are developed, validated, and translated into biological knowledge.

4. Conclusions

This study concludes that the integration of transcriptomics and proteomics plays an important role in strengthening systems biology-based modeling of gene expression regulation. The findings show that transcriptomic data cannot stand alone as a complete explanation of biological activity, because mRNA expression does not always align with protein abundance. Through qualitative analysis, this study identified several key themes, including RNA-protein inconsistency, the role of computational tools, data infrastructure limitations, interdisciplinary collaboration, and the construction of biological meaning from integrated omics data. These findings contribute to the literature by showing that multi-omics integration is not only a technical process, but also an interpretive process that requires biological reasoning, methodological transparency, and contextual understanding.

Theoretically, this study reinforces the view that gene expression regulation is a multilayered process involving transcriptional, post-transcriptional, translational, and post-translational mechanisms. Practically, the findings suggest the need to strengthen bioinformatics training, improve access to proteomic infrastructure, and develop local omics databases that are relevant to the Indonesian research context. From a policy perspective, research institutions and funding bodies should support collaborative multi-omics research involving molecular biologists, bioinformaticians, statisticians, and applied science experts. Future studies should involve broader research settings, compare experiences across institutions or countries, and combine qualitative inquiry with bibliometric or workflow analysis to deepen understanding of how integrated omics models are developed and validated.

References

Ahmed, S. K. (2024). The pillars of trustworthiness in qualitative research. *Journal of Medicine, Surgery, and Public Health, 2*, 100051. https://doi.org/10.1016/j.glmedi.2024.100051

Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., Ameen, B. M. M., & Khdhir, R. M. (2025). Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health, 6*, 100198. https://doi.org/10.1016/j.glmedi.2025.100198

Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J. C., & Stegle, O. (2020). MOFA+: A statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biology, 21*, 111. https://doi.org/10.1186/s13059-020-02015-1

Braun, V., & Clarke, V. (2022). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health, 24*(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Chen, Y., Fan, X., Shi, C., Shi, Z., & Wang, C. (2025). A joint analysis of single cell transcriptomics and proteomics using transformer. *npj Systems Biology and Applications, 11*, 1. https://doi.org/10.1038/s41540-024-00484-9

Dahal, N. (2025). Qualitative data analysis: Reflections, procedures, and some points for consideration. *Frontiers in Research Metrics and Analytics, 10*, 1669578. https://doi.org/10.3389/frma.2025.1669578

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine, 292*, 114523. https://doi.org/10.1016/j.socscimed.2021.114523

Hernández-Lemus, E., & Ochoa, S. (2024). Methods for multi-omic data integration in cancer research. *Frontiers in Genetics, 15*, 1425456. https://doi.org/10.3389/fgene.2024.1425456

Ibeh, N., Kusuma, P., Darusallam, C. C., Malik, S. G., Sudoyo, H., McCarthy, D. J., & Gallego Romero, I. (2024). Profiling genetically driven alternative splicing across the Indonesian archipelago. *The American Journal of Human Genetics, 111*(11), 2458–2477. https://doi.org/10.1016/j.ajhg.2024.09.004

Lakkis, J., Schroeder, A., Su, K., Lee, M. Y. Y., Bashore, A. C., Reilly, M. P., & Li, M. (2022). A multi-use deep learning method for CITE-seq and single-cell RNA-seq data integration with cell surface protein prediction and imputation. *Nature Machine Intelligence, 4*(11), 940–952. https://doi.org/10.1038/s42256-022-00545-w

Li, L., Huang, G., Xiang, W., Zhu, H., Zhang, H., Zhang, J., Ding, Z., Liu, J., & Wu, D. (2022). Integrated transcriptomic and proteomic analyses uncover the regulatory mechanisms of *Myricaria laxiflora* under flooding stress. *Frontiers in Plant Science, 13*, 924490. https://doi.org/10.3389/fpls.2022.924490

Lim, W. M. (2025). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal, 33*(2), 199–229. https://doi.org/10.1177/14413582241264619

Liu, J., Jin, X., Qiu, C., Han, P., Wang, Y., Zhao, J., Wu, J., Yan, N., & Song, X. (2024). Integrated transcriptomics-proteomics analysis identifies molecular phenotypic alterations associated with colorectal cancer. *Journal of Proteome Research, 23*(1), 175–184. https://doi.org/10.1021/acs.jproteome.3c00526

Miller, E. M., Porter, J. E., & Barbagallo, M. S. (2023). Simplifying qualitative case study research methodology: A step-by-step guide using a palliative care example. *The Qualitative Report, 28*(8), 2363–2379. https://doi.org/10.46743/2160-3715/2023.6478

Moilanen, T., Sivonen, M., Hipp, K., Kallio, H., Papinaho, O., Stolt, M., Turjamaa, R., Häggman-Laitila, A., & Kangasniemi, M. (2022). Developing a feasible and credible method for analyzing healthcare documents as written data. *Global Qualitative Nursing Research, 9*, 1–14. https://doi.org/10.1177/23333936221108706

Morgan, H. (2022). Conducting a qualitative document analysis. *The Qualitative Report, 27*(1), 64–77. https://doi.org/10.46743/2160-3715/2022.5044

Natri, H. M., Hudjashov, G., Jacobs, G. S., Kusuma, P., Saag, L., Darusallam, C. C., Metspalu, M., Sudoyo, H., Cox, M. P., Gallego Romero, I., & Banovich, N. E. (2022). Genetic architecture of gene regulation in Indonesian populations identifies QTLs associated with global and local ancestries. *The American Journal of Human Genetics, 109*(1), 50–65. https://doi.org/10.1016/j.ajhg.2021.11.017

Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2022). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher, 45*(3), 241–251. https://doi.org/10.1080/0142159X.2022.2057287

Patel, A., McGrosso, D., Hefner, Y., Campeau, A., Sastry, A. V., Maurya, S., Rychel, K., Gonzalez, D. J., & Palsson, B. O. (2024). Proteome allocation is linked to transcriptional regulation through a modularized transcriptome. *Nature Communications, 15*, 5234. https://doi.org/10.1038/s41467-024-49231-y

Sanches, P. H. G., de Melo, N. C., Porcari, A. M., & de Carvalho, L. M. (2024). Integrating molecular perspectives: Strategies for comprehensive multi-omics integrative data analysis and machine learning applications in transcriptomics, proteomics, and metabolomics. *Biology, 13*(11), 848. https://doi.org/10.3390/biology13110848

Sibilio, P., De Smaele, E., Paci, P., & Conte, F. (2025). Integrating multi-omics data: Methods and applications in human complex diseases. *Biotechnology Reports, 48*, e00938. https://doi.org/10.1016/j.btre.2025.e00938

Sidhaye, J., Trepte, P., Sepke, N., Novatchkova, M., Schutzbier, M., Dürnberger, G., Mechtler, K., & Knoblich, J. A. (2023). Integrated transcriptome and proteome analysis reveals posttranscriptional regulation of ribosomal genes in human brain organoids. *eLife, 12*, e85135. https://doi.org/10.7554/eLife.85135

Veenstra, T. D. (2021). Systems biology and multi-omics. *Proteomics, 21*(3–4), e2000306. https://doi.org/10.1002/pmic.202000306

Zhao, T., Zhan, D., Qu, S., Jiang, S., Gan, W., Qin, W., Zheng, C., Cheng, F., Lu, Y., Liu, M., Shi, J., Liang, H., Wang, Y., Qin, J., Zen, K., & Liu, Z. (2023). Transcriptomics-proteomics integration reveals alternative polyadenylation driving inflammation-related protein translation in patients with diabetic nephropathy. *Journal of Translational Medicine, 21*, 86. https://doi.org/10.1186/s12967-023-03934-w

Zhou, Z., Ye, C., Wang, J., & Zhang, N. R. (2020). Surface protein imputation from single cell transcriptomes by deep neural networks. *Nature Communications, 11*, 651. https://doi.org/10.1038/s41467-020-14391-0