Discover Oncology

REVIEW Open Access

Harnessing plasma transcriptomics for non-invasive cancer biomarker identification: a comprehensive review

Nur Mazidah Haji Noor Mohamed[1], Nurulisa Zulkifle[1,2] and Siti Razila Abdul Razak[1,2*]

*Correspondence:
Siti Razila Abdul Razak
sitirazila@usm.my
[1]Department of Biomedical Science, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Sains@Bertam, Kepala Batas 13200, Pulau Pinang, Malaysia
[2]Advanced Management of Liver Malignancies Research Program, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Sains@Bertam, Kepala Batas 13200, Pulau Pinang, Malaysia

Abstract

Plasma transcriptomic profiling has emerged as a promising non-invasive strategy for cancer biomarker discovery, offering a dynamic change in gene expression profiles associated with carcinogenesis. Plasma-based transcriptomic profiling offers a less-invasive alternative for detecting cancer biomarkers compared to tissue-based methods. Unlike traditional tissue biopsies, plasma-based analyses enable easier, repeatable sampling and real-time disease monitoring through liquid biopsy. Despite its clinical potential, plasma transcriptomics remain underutilized due to technical limitations such as RNA degradation, low yield, and lack of standardized protocols. This review presents a comprehensive overview of current transcriptomic technologies including high-throughput sequencing (RNA-seq), single cell RNA sequencing (scRNA-seq), and spatial transcriptomics, and evaluates their suitability for plasma-based applications. This review also discusses the analytical workflows used to process transcriptomic data, including differential expression analysis, pathway enrichment, and protein-protein interaction network mapping, as well as tool for assessing biomarker performance through ROC curves and survival analysis. Furthermore, this review highlights emerging trends in integrative analysis, such as meta-analysis and artificial intelligence (AI)-assisted multiomics integration, which significantly enhance biomarker discovery and therapeutic target identification. The review also outlines critical challenges in plasma transcriptomic profiling and proposes strategies for advancing the field toward clinical translation. Overall, this review aims to equip cancer researchers and clinicians with foundational knowledge and practical insights into plasma-based transcriptomics, reinforcing its potential in precision oncology and non-invasive cancer diagnostic.

Keywords Plasma transcriptomic profiling, Liquid biopsy, RNA-sequencing, cancer biomarkers, Non-invasive diagnostics, Bioinformatic

1 Introduction

Transcriptomics refers to the study of transcriptomes, encompassing the complete set of coding (messenger RNA) and non-coding RNA molecules, including transfer RNA, ribosomal RNA, and small RNAs such as microRNA, long non-coding RNA, small

interfering RNA, and small nuclear RNA [1]. In cancer, transcriptomes expression levels are often altered, and profiling these differentially expressed transcripts can aid in identifying cancer-specific genes that may serve as biomarkers or therapeutic targets [2]. Additionally, transcriptomic profiling plays a crucial role in predicting therapeutic resistance, monitoring disease progression, identifying molecular subtypes, detecting fusion transcripts from chromosomal rearrangements, and evaluating immune infiltration within tumors [3–5]. These capabilities highlight transcriptomics as a powerful tool for enhancing cancer diagnosis, guiding treatment strategies, and improving patient outcomes.

Liquid biopsy using plasma provides a less invasive alternative to traditional tissue biopsies in cancer research [6]. Plasma has emerged as a valuable source for transcriptomic studies aimed at identifying cancer biomarkers [7–10]. It is particularly advantageous due to its accessibility via minimally invasive venepuncture, unlike tissue biopsies that require surgical intervention [11, 12]. This facilitates frequent sampling for disease monitoring. Plasma-based biomarkers support early cancer detection and enable continuous tracking of disease progression and treatment response [13]. This review provides an overview of emerging methods in transcriptomic profiling using plasma for cancer research.

2 Types of transcriptomic profiling

2.1 High-Throughput sequencing (HTS)

RNA sequencing (RNA-seq) has emerged as a powerful and comprehensive technique for detecting gene expression profiles in plasma samples due to its sensitivity and ability to capture diverse RNA species [14, 15]. It enables detailed analysis of all transcribed genes within a biological sample and is particularly effective in identifying differentially expressed genes (DEGs) between disease and healthy states. This high-throughput approach supports biomarker discovery and the identification of therapeutic targets, enhancing disease diagnosis and monitoring including for various cancers [14, 16, 17].

However, plasma-derived RNA presents unique biological challenges that could influence RNA-seq performance. Circulating cell-free RNA (cfRNA) in plasma is often highly fragmented, present in low abundance, and easily degraded. These challenges increased the need for optimized sample processing and RNA extraction methods for plasma-based RNA-seq studies [18]. Beyond biological limitations, several quality issues in raw RNA-seq data, such as adapter contamination and low-quality reads, can affect downstream analysis and lead to erroneous conclusions. Therefore, implementing rigorous quality control procedures, such as read quality assessment and contaminant removal, is essential prior to data analysis to ensure reliability and reproducibility [19, 20].

Numerous studies have utilised RNA-seq for plasma-based cancer research. For example, DEGs such as COX1, COX2, COX3, ND1, ND2, ND4L, and ATP6 were identified in plasma from non-small cell lung cancer patients compared to healthy individuals [21]. Similarly, Jin et al. used RNA-seq to compare plasma gene expression in colorectal cancer patients before and after surgery, identifying prognostic biomarkers including HPGD, PACS1, and TDP2 [22]. Another study using small RNA deep sequencing identified differentially expressed microRNAs such as hsa-miR-182-5p in colorectal cancer patients versus normal controls [23]. Furthermore, RNA-seq of plasma from patients with various cancers (colorectal, gastric, liver, lung, and esophageal) were performed to

identify cancer-specific biomarkers [24]. Collectively, these studies highlight the potential of plasma RNA-seq for non-invasive cancer biomarker discovery.

2.2 Single-Cell RNA sequencing

Single cell RNA sequencing (scRNA-seq) enables the analysis of gene expression at the resolution of individual cells, making it particularly valuable for assessing intra-tumoral heterogeneity [25]. While scRNA-seq is commonly applied to tumor tissue sample [26–29], it has also been used to study extracellular vesicles (EVs) derived from tissue, peripheral blood mononuclear cells, and blood samples [30–32]. Its application to plasma-derived EVs offers the possibility of exploring RNA heterogeneity across different EV subpopulations in cancer patients.

Nonetheless, scRNA-seq presents technical challenges for plasma analysis due to low RNA abundance, the acellular nature of plasma, and potential contamination noise. Additionally, the high cost and methodological complexity of scRNA-seq limit its routine application in plasma transcriptomics [33].

2.3 Spatial transcriptomics

Spatial transcriptomics (ST) profiles, gene expression while preserving spatial context within intact tissue section [34]. This technique is valuable in investigating spatial variations in gene expression, tumor heterogeneity, prognostic factors, and tissue-specific interactions [35, 36]. Unlike scRNA-seq, which loses spatial context during single-cell preparation, ST retains positional information, enabling detailed analysis of tumor architecture and microenvironmental interactions [37]. However, spatial transcriptomics is not suitable for plasma-based transcriptomic profiling. Since ST requires structured tissue sections to maintain spatial gene expression maps, it is incompatible with plasma, a fluid matrix lacking spatial organization [35].

3 Computational methods in transcriptomic data analysis

Computational analysis of transcriptomic data typically follows a sequential workflow, as illustrated in Fig. 1. The process generally begins with quality control of raw sequencing reads, followed by alignment to a reference genome and quantification of gene expression. Differential expression analysis is then performed to identify genes with significant expression changes between conditions. These DEGs can subsequently be subjected to pathway enrichment analysis and protein–protein interaction (PPI) network construction to uncover biological functions and molecular interactions. In this section, we provide an overview of commonly used computational tools at each step, emphasizing their practical relevance for plasma transcriptomics.

3.1 Differential gene expression analysis

Galaxy and GEO2R are among the most widely used platforms for DEG analysis in plasma transcriptomics. Both are considered high-level, user-friendly platforms that enable researchers to perform analyses without requiring advanced computational expertise.

Galaxy (https://usegalaxy.org/) is an open-access platform for scientific computing that integrates a suite of tools for RNA-seq analysis [38]. These include quality assessment (e.g., FastQC, Trimmomatic, Cutadapt, Trim Galore), read alignment to a

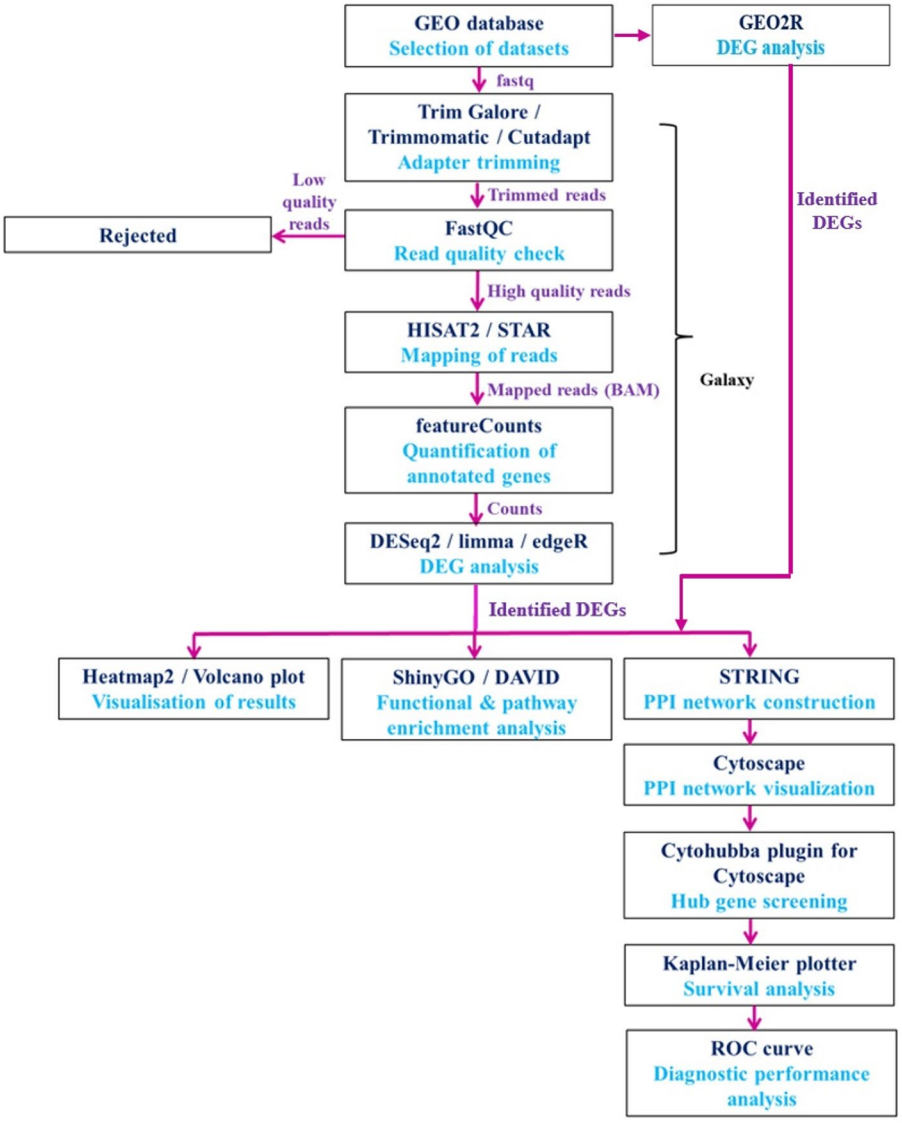

Fig. 1 The overview of workflow involving Galaxy or GEO2R, used in cancer-related transcriptomic profiling studies

reference genome (e.g., HISAT2, STAR), gene quantification (e.g., featureCounts), DEG analysis (e.g., DESeq2, limma, edgeR), and data visualization (e.g., Heatmap2, Volcano plots) [39, 40]. Its modular design and graphical interface make it accessible for researchers with limited programming background, and it has been widely adopted in cancer-related transcriptomic studies [41–44].

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) is another web-based tool that allows direct comparison of gene expression across two or more groups from the GEO database. It generates DEG lists and visualization outputs such as volcano plots, Venn diagrams, and box plots [45]. GEO2R is particularly useful for quick, exploratory analyses and has been extensively applied in cancer biomarker discovery studies [46–49]. Figure 1 illustrates the typical analysis pipelines using Galaxy and GEO2R.

While Galaxy and GEO2R simplify DEG analysis through their high-level, graphical interfaces, tools such as DESeq2 [50], edgeR [51], and limma [52] represent lower-level

statistical packages that provide greater flexibility and precision for researchers with computational expertise.

Comparative evaluations have shown that DESeq2 often identifies more DEGs than limma, with approximately 90% overlap between the two tools, reflecting high consistency [53]. In a broader analysis, DESeq2 detected the most significant DEGs [760], followed by edgeR [597] and limma [278], with 273 DEGs overlapping across all three methods [54]. Despite identifying fewer DEGs overall, limma remains valuable for detecting subtle but biologically important expression changes. Notably, edgeR and limma require normalized RNA-seq count data, whereas DESeq2 can be applied directly to raw counts [55]. In addition, DESeq2 is particularly suitable for studies with at least six samples per group [56]. Importantly, all three methods have been successfully applied in plasma transcriptomic studies, supporting their practical utility in analyzing circulating cell-free RNA [57–59].

3.2 Pathway enrichment analysis

DAVID and ShinyGO are among the most widely used tools for pathway enrichment and functional annotation of DEGs in cancer research. DAVID (Database for Annotation, Visualization, and Integrated Discovery) (https://davidbioinformatics.nih.gov/tools.jsp) provides detailed pathway names and statistical data, which can be used with visualization tools like SR Plot to generate enrichment diagrams [60].

ShinyGO (https://bioinformatics.sdstate.edu/go77/) offers a more user-friendly interface that directly produces graphical outputs, enhancing usability and efficiency [61]. Figure 2 shows Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway outputs

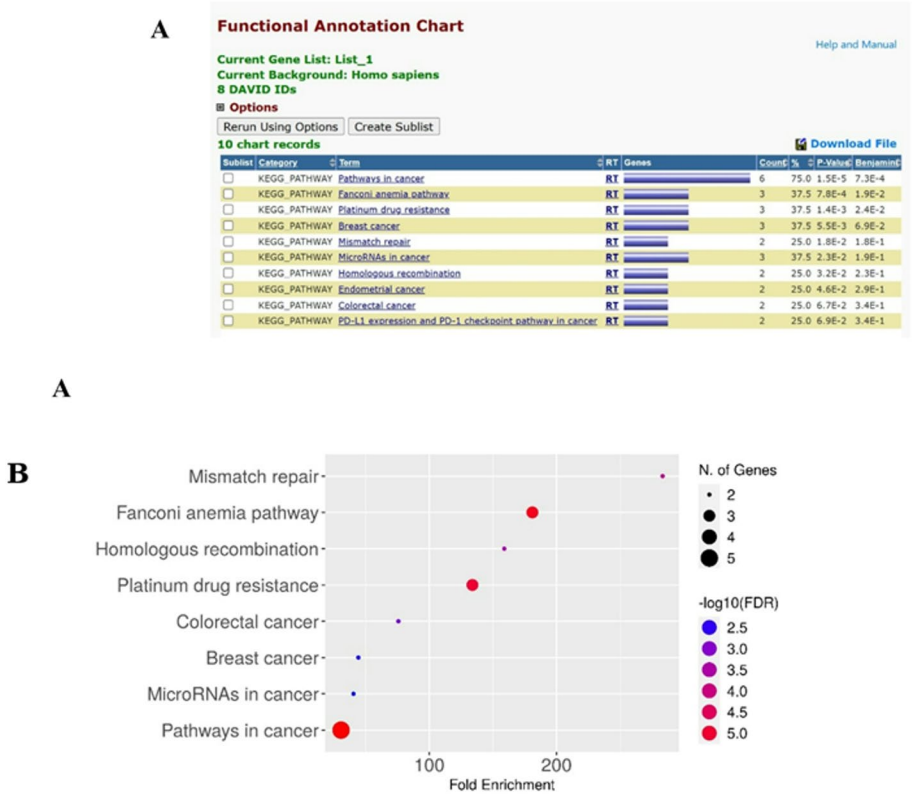

Fig. 2 Example outputs for KEGG pathway analysis using (A) DAVID and (B) ShinyGO

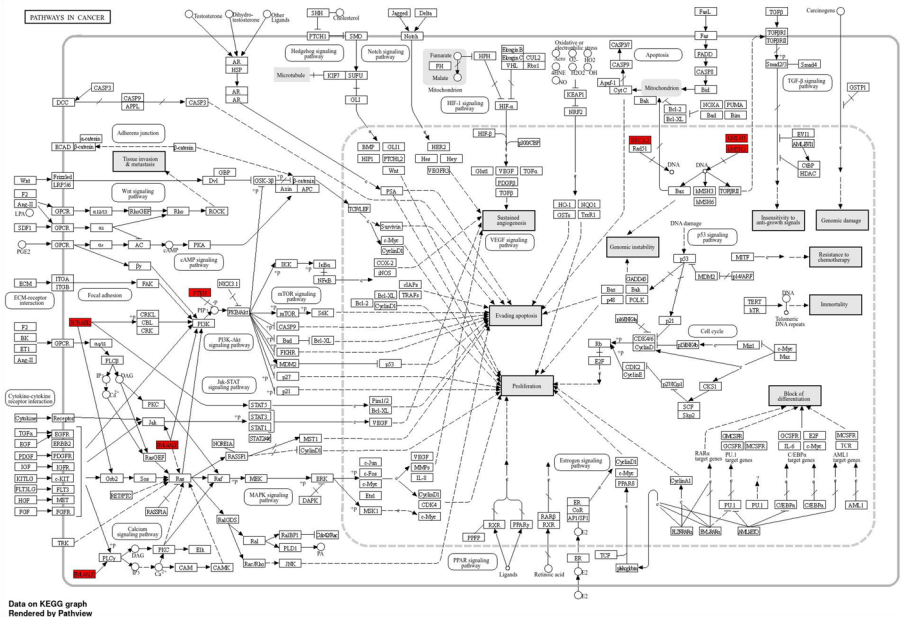

Fig. 3 Example of KEGG pathway diagrams generated by KEGG. Red indicates the involvement of query genes in the pathway

Table 1 Available tools for functional and pathway enrichment analysis

Tools	Description	Usage in plasma transcriptomics for cancer research	References
DAVID	Web server for gene functional annotation and pathway enrichment	Used in cancer-related plasma transcriptomic studies	[22, 57, 60]
ShinyGO	Web server for gene functional annotation and pathway enrichment	Widely adopted in cancer-related studies, including tissue-, plasma-, and serum-derived exosomal transcriptomic studies	[61–64]
IPA (Ingenuity Pathway Analysis)	Commercial pathway analysis tool	Used in cancer-related plasma transcriptomic studies	[65, 66]
Reactome	Curated, expert-authored, and peer-reviewed pathway database	Very limited use in cancer-related plasma transcriptomic studies	[67, 68]
WikiPathways	Open-source biological pathway database	No reported use in cancer-related plasma transcriptomic studies	[69]
EnrichR	Comprehensive resource for curated gene sets	Used in cancer-related plasma transcriptomic studies	[70, 71]

produced by both DAVID and ShinyGO for selected genes (e.g., PALB2, BRCA1, BRCA2, PTEN, MLH1, MSH2, ABL1, ALK, and BMPR1A). ShinyGO also produces KEGG pathway diagrams highlighting the specific involvement of query genes [61–64], as illustrated in Fig. 3. While ShinyGO has more limited species coverage compared to DAVID, both tools are well-suited for enrichment analysis in human transcriptomic studies.

In addition to DAVID and ShinyGO, several other resources are available for pathway enrichment, each with distinct strengths and limitations (Table 1). Ingenuity Pathway Analysis (IPA) is a commercial software that offers advanced visualization and curation, and it has been used in cancer-related plasma transcriptomic studies [65, 66]. Reactome is a peer-reviewed, expert-curated pathway database, but its application in plasma transcriptomic studies remains limited [67–68]. WikiPathways is an open-source

and community-driven pathway database that is increasingly popular in bioinformatics, although its use in plasma transcriptomic profiling has not yet been reported [69]. EnrichR is a comprehensive platform that integrates multiple gene set libraries, making it useful for functional annotation across diverse datasets; it has also been applied in plasma transcriptomic studies [70–71].

Table 1 provides a summary of these enrichment tools, outlining their main features and reported applications in cancer-related plasma transcriptomic research.

3.3 Protein-protein interaction network construction

The STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) (https://string-db.org) is one of the most widely used resources for constructing PPI networks based on DEGs from transcriptomic profiling. STRING integrates computational predictions, experimental data, and curated knowledge to provide both functional and physical associations among proteins [72, 73]. These networks can help reveal biological connections among DEGs at the protein level, offering insights into potential pathways and mechanisms of disease.

PPI networks generated in STRING can be further visualized and analyzed in Cytoscape using the StringApp plugin [74, 75]. In these networks, nodes represent genes or proteins, while edges denote molecular interactions [76]. Figure 4 illustrates an example of a PPI network constructed with STRING and visualized in Cytoscape. Plugins such as CytoHubba extend this analysis by identifying hub genes through ranking algorithms, including degree, density of maximum neighborhood component (DMNC), maximal clique centrality (MCC), and betweenness [77–79]. Genes consistently identified across multiple ranking methods are typically considered core hub genes and may represent strong biomarker candidates.

In addition to STRING and Cytoscape, several other resources are available for PPI network analysis, including PiSITE, ComPPI, IntAct, and BioGRID. However, our literature search indicates that their adoption in plasma transcriptomic studies remains limited. Table 2 provides a summary of these additional tools and their potential relevance to plasma transcriptomic research.

3.4 Diagnostic performance test and survival analysis

Evaluating the clinical potential of candidate biomarkers often involves testing their diagnostic performance and prognostic value. Two commonly applied approaches are receiver operating characteristic (ROC) curve analysis and survival analysis.

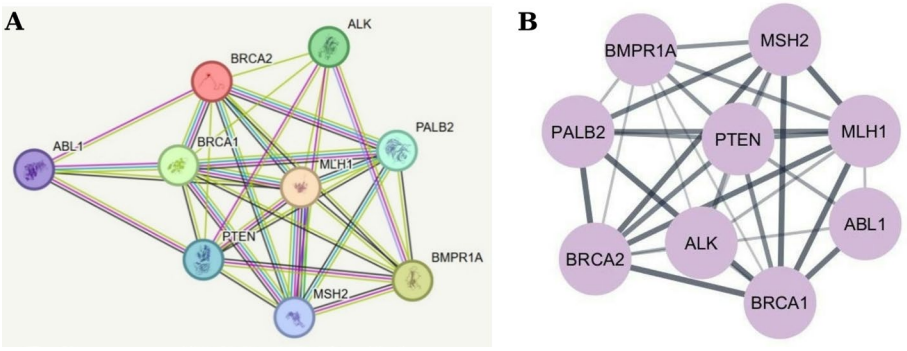

Fig. 4 The example of PPI network constructed using (A) STRING database, and (B) visualised using Cytoscape

Table 2 Other available tools for PPI network construction

Tools	Description	Use in plasma transcriptomics for cancer research	References
PiSITE	Structural PPI site prediction based on 3D data For structural bioinformatics and protein interface prediction	No reported use in cancer-related plasma transcriptomic studies	[80]
ComPPI	Cellular compartment-specific PPI database	No reported use in cancer-related plasma transcriptomic studies	[81]
IntAct	Freely accessible database for experimentally validated molecular interactions including PPI	Used in cancer-related studies, including plasma transcriptomics	[82–87]
BioGRID	Comprehensive repository of protein, genetic, and chemical interactions curated from literature	Used in cancer research, including tissue- and plasma-based transcriptomic studies	[86–90]

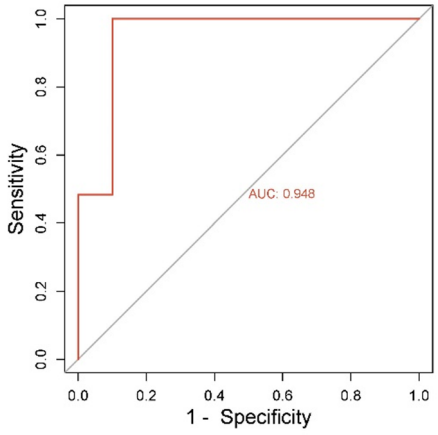

Fig. 5 Example of a ROC curve generated with AUC value 0.948 for gene RPS7 based on normalized expression data from the GEO dataset (GSE142987)

The ROC curve plots sensitivity (true positive rate) against specificity (true negative rate), providing a graphical representation of diagnostic accuracy. The area under the curve (AUC) serves as a key performance metric, with values closer to 1.0 indicating high discriminative ability [91, 92]. Online tools such as easyROC (http://biosoft.erciyes.edu.tr/app/easyROC/) and SR plot (https://www.bioinformatics.com.cn/plot_basic_one_or_multi_ROC_curve_plot_106_en) are frequently used to generate ROC curves [93]. An AUC value close to 1 indicates high diagnostic accuracy. Biomarkers with AUC values above 0.8 are generally considered clinically relevant [94, 95]. Figure 5 presents an example of a ROC curve generated using publicly available plasma transcriptomic data, where an AUC of 0.94 demonstrates strong diagnostic ability. Several studies have also validated plasma-based cancer biomarkers using ROC analysis, further supporting its clinical utility [96–99]. In addition to online platforms, offline alternatives such as the pROC R package or Python-based implementations (e.g., scikit-learn) provide advanced customization and greater control over data handling, making them suitable for studies involving sensitive or confidential datasets [100, 101].

Survival analysis is equally important for assessing prognostic value. Tools such as the KM Plotter (kmplot.com/analysis/), allow researchers to evaluate associations between gene expression levels and patient survival outcomes [102, 103]. Kaplan–Meier plots can indicate whether high or low gene expression correlates with favorable or poor

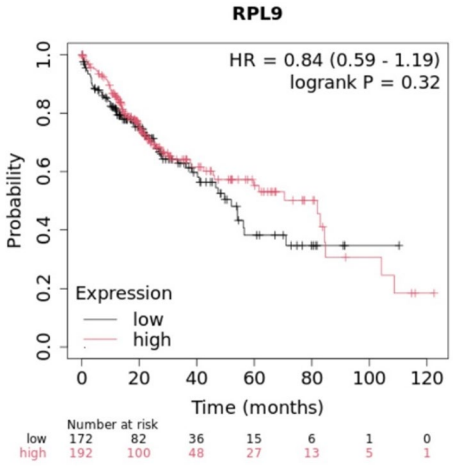

Fig. 6 Example of a KM plot

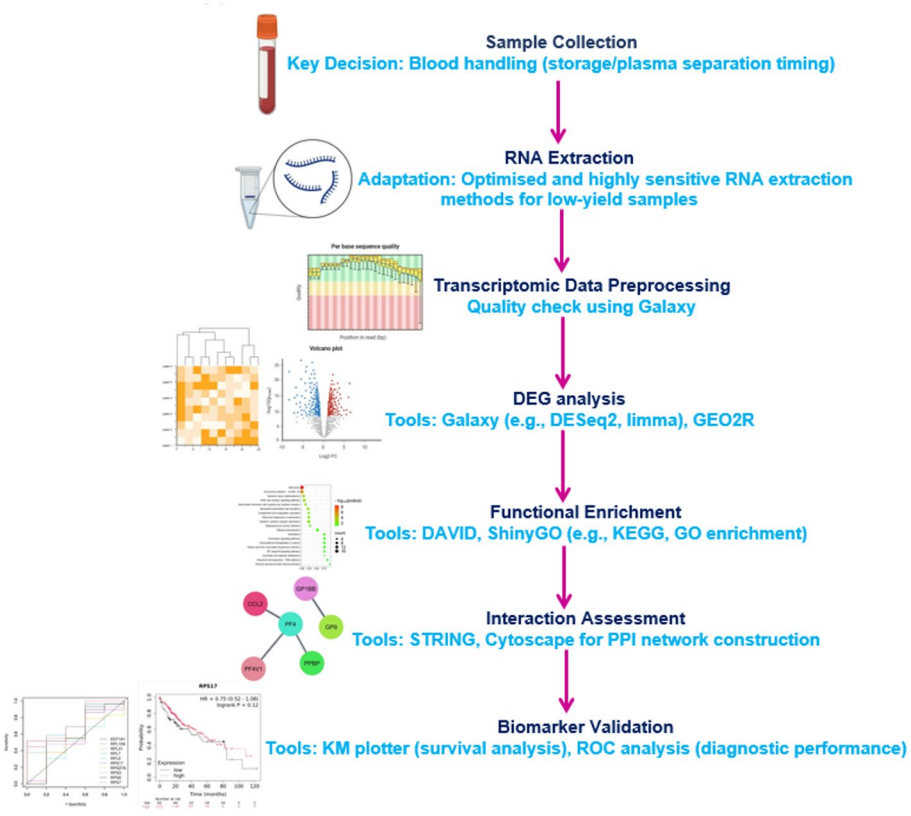

Fig. 7 Plasma transcriptomic analysis workflow for less-invasive cancer biomarker discovery (created with BioRender)

prognosis [104]. More advanced strategies, including multi-gene risk score models, further improve prognostic stratification [105, 106]. Figure 6 shows an example Kaplan–Meier plot, while Fig. 7 provides an overview of a complete plasma transcriptomic analysis workflow.

4 Integrated analysis as the emerging trend

4.1 Meta-analysis

Meta-analysis combines data from multiple studies with similar objectives to produce a single, more precise estimate of the overall effect size [107]. In the context of plasma-based transcriptomic profiling, studies often vary in terms of patient populations, sample handling procedures, and analytical techniques. Meta-analyses help standardise these differences, reducing heterogeneity and providing a more balanced and reliable synthesis of findings [108].

By integrating dataset, meta-analyses can improve the diagnostic accuracy of cancer biomarkers, enhance statistical power, and increase the robustness of research outcomes [109–112]. This is particularly beneficial for plasma-based studies with small sample sizes, where pooling data from multiple sources can overcome limitations in statistical significance. Furthermore, meta-analysis is a cost-effective strategy, as it relies on existing datasets rather that requiring new and often expensive sequencing experiments [112].

In recent years, there has been a notable increase in meta-analyses involving differential gene expression data from plasma, serum, and blood for cancer biomarker [113–116]. These efforts demonstrate that meta-analysis is a powerful tool for integrating transcriptomic data to advance our understanding of cancer biomarkers and their clinical applications.

4.2 Artificial intelligence

Artificial intelligence (AI) has become an essential tool in transcriptomic data analysis, particularly in managing and interpreting large, complex datasets that are difficult to handle using conventional methods [117, 118]. For instance, machine learning (ML) algorithms such as XGBoost and Boruta have been applied to identify prognostic biomarkers in kidney cancer using transcriptomic data from 531 kidney renal clear cell carcinoma and 72 normal control samples [119]. Moreover, a range of ML methods including Extra-Trees, GBDT, XGBoost, AdaBoost, KNN, MLP, RFECV-LR, RFECV-SVM, RF, LR-L1, and SVC-L1, has been employed to detect diagnostic and prognostic biomarkers in breast cancer by analyzing transcriptomic profiles from 356 cancerous and 345 normal breast tissue samples [120].

Beyond transcriptomics alone, AI techniques such as deep learning have been used to integrate multi-omics data (e.g., transcriptomics, proteomics, genomics, epigenomics, and metabolomics) to enhance cancer classification, diagnosis, prognosis prediction, and treatment response modeling [121–123]. For example, a concatenation autoencoder, which is an AI-based deep learning tool, was used to predict the survival rate of breast cancer patients by integrating multi-omics data from 60 000 samples [124].

Li et al. (2025) integrated plasma-based transcriptomic and lipidomic profiles from hepatocellular carcinoma (HCC) patients and normal controls using deep learning and ML tools and identified 12 lipid markers as HCC diagnostic markers [125]. Meanwhile, Wei et al. (2024) established plasma miR-15a/16 as a potential early-stage HCC diagnostic biomarker using ML [126]. The AI-facilitated integration of plasma-based multi-omics data from cancer patients, particularly when utilising large sample sizes, holds significant promise for uncovering plasma-based less-invasive novel biomarkers for cancer diagnosis.

While AI offers significant advantages in terms of accuracy, efficiency, and reduction of human error, it also require advanced technical expertise for effective implementation [127]. Compared to traditional methods, AI-assisted approaches present a steep learning curve, making them less accessible to researchers without specialized computational backgrounds.

4.3 Multi-omics integration for biomarker discovery

Integrating multi-omics data provides a more comprehensive view of cancer biology, enabling the identification of novel biomarkers that may be missed through single-omics analyses [128]. While tissue-based studies have led to demonstrating the power of combining transcriptomic and proteomic analyses, plasma-based multi-omics remains underexplored. For example, one study integrated transcriptomic profiling of buffy coat with plasma proteomics in HCC patients, and established SH3PXD2B and CD70 as potential blood-based biomarkers for HCC [129]. Similarly, transcriptomic profiling of tissue, combined with blood and plasma proteomic profiling, has played a critical role in identifying biomarkers at different stages of breast cancer, aiding early diagnosis and treatment planning [130]. However, despite the growing interest in circulating cell-free RNA (cfRNA), integrated transcriptomic studies focused on discovering cancer biomarkers sourced solely from plasma samples remain limited. This gap likely stems from the technical and biological challenges associated with cfRNA in plasma, including its low abundance, fragmentation, and potential contamination from lysed blood cells. Nevertheless, future studies on integrated plasma transcriptomic profiling could open new avenues for minimally invasive biomarker discovery in cancer.

5 Challenges and future directions

5.1 Challenges in transcriptomic profiling

Despite its promises, plasma-based transcriptomic profiling faces several challenges. RNA is highly susceptible to degradation during sample collection and handling, where improper storage and delays in processing can significantly compromise RNA integrity [131]. Degraded RNA fragments can lead to inaccurate RNA-seq reads [132]. Moreover, the inherently low abundance of circulating RNA in plasma samples can impair data quality. Low input RNA increases the risk of adapter dimer formation and the inclusion of non-target reads, which may contaminate sequencing results [133]. Proper handling protocols and highly sensitive RNA extraction methods are essential to address these issues.

Another major challenge is the lack of standardized protocols for plasma transcriptomic analysis. Variability in blood collection storage conditions, and data processing pipelines can significantly affect RNA profiles and compromise reproducibility [134]. Differences in bioinformatics tools used for read alignment, quantification, and DEGs analysis further contribute to inconsistent results. Establishing standardized protocol would help minimize variability across studies.

Integrating transcriptomic with other omics also presents computational challenges requiring advanced statistical and analytical tools for effective interpretation [128, 135]. Researchers in low- resource settings may face difficulties accessing the technical infrastructure and expertise required to adopt these tools. Additionally, the high cost of sequencing remains a barrier to widespread implementation [136].

5.2 Future directions

The following outlines some key suggestions for future research to further advance plasma-based transcriptomic profiling:

- Increasing the application of AI-based techniques to integrate transcriptomic data with other omics layer, thereby improving model accuracy for cancer diagnosis, prognosis, and treatment planning.
- Advancing the clinical translation of plasma transcriptomic by developing reproducible biomarkers, validating them in large and diverse patient cohorts, and standardizing RNA extraction and analytical protocols.

6 Conclusion

Transcriptomic profiling is a powerful tool in cancer research, offering insights into disease mechanisms, biomarker discovery, and therapeutic targets. However, its application in plasma-based studies introduces specific challenges, including issues of RNA degradation, variability in data processing pipeline, and the complexity of integrating diverse data types.

Despite these challenges, integrated transcriptomic profiling holds great potential for enhancing diagnostic precision and enabling personalized treatment approaches. Ongoing efforts to improve standardization, adopt advanced analytical tools, and validate findings in clinical settings will be essential to realize the full potential of plasma transcriptomics in cancer research.

Acknowledgements
This work was funded by grant from Fundamental Research Grant Scheme, Ministry of Higher Education Malaysia (FRGS/1/2023/SKK10/USM/02/7) and Advanced Management of Liver Malignancies Research Program, Advanced Medical and Dental Institute, Universiti Sains Malaysia.

Dual publication
The authors declare that this study have not been previously published or are under consideration for publication elsewhere.

Authorship clarification
All authors agreed with the content, approved the version to be published and agree to be accountable for all aspects of the work.

Author contributions
NMHNM, NZ and SRAR conceptualized and designed the study. NMHNM and SRAR contributed to data interpretation and manuscript drafting. All authors reviewed, edited, and approved the final manuscript.

Funding
This work was funded by a grant from Fundamental Research Grant Scheme, Ministry of Higher Education Malaysia (FRGS/1/2023/SKK10/USM/02/7) and Advanced Management of Liver Malignancies Research Program, Advanced Medical and Dental Institute, Universiti Sains Malaysia.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
The study was based on publicly available transcriptomic datasets from the GEO database and did not involve human subjects or animal experiments requiring ethical approval. This study did not involve the recruitment or participation of human subjects.

Consent for publication
Not applicable. This study does not include any individual person's data in any form (including individual details, images, or videos).

Competing interests
The authors declare no competing interests.

Received: 1 July 2025 / Accepted: 2 December 2025

Published online: 07 December 2025

References

1. Assis AF, Oliveira EH, Donate PB, Giuliatti S, Nguyen C, Passos GA. What is the transcriptome and how it is evaluated. Transcriptomics in health and disease. Springer; 2022. pp. 3–50.
2. Supplitt S, Karpinski P, Sasiadek M, Laczmanska I. Current achievements and applications of transcriptomics in personalized cancer medicine. Int J Mol Sci. 2021;22(3):1422.
3. Mark Reverter. The importance of transcriptomics in cancer immunology. J Data Min Genomics Proteom. 2023;14(2).
4. Li K, Huang J, Tan Y, Sun J, Zhou M. Single-cell and bulk transcriptome analysis reveals tumor cell heterogeneity and underlying molecular program in uveal melanoma. J Translational Med. 2024;22(1):1020.
5. Balakrishnan K, Xiao Y, Chen Y, Dong J. Elevated expression of cell adhesion, metabolic, and mucus secretion gene clusters associated with tumorigenesis, metastasis, and poor survival in pancreatic ductal adenocarcinoma. Cancers. 2024;16(23):4049.
6. Law EW, Settell ML, Kurani SS, Eckert EC, Liu MC, Greenberg-Worisek AJ. Liquid biopsy: emergence of an alternative cancer detection method. Clin Transl Sci. 2020;13(5):845.
7. Reggiardo RE, Maroli SV, Peddu V, Davidson AE, Hill A, LaMontagne E, et al. Profiling of repetitive RNA sequences in the blood plasma of patients with cancer. Nat Biomedical Eng. 2023;7(12):1627–35.
8. Nishiwada S, Cui Y, Sho M, Jun E, Akahori T, Nakamura K, et al. Transcriptomic profiling identifies an Exosomal MicroRNA signature for predicting recurrence following surgery in patients with pancreatic ductal adenocarcinoma. Ann Surg. 2022;276(6):e876–85.
9. Bahrambeigi V, Lee JJ, Branchi V, Rajapakshe KI, Xu Z, Kui N, et al. Transcriptomic profiling of plasma extracellular vesicles enables reliable annotation of the Cancer-specific transcriptome and molecular subtype. Cancer Res. 2024;84(10):1719–32.
10. Rajkumar T, Amritha S, Sridevi V, Gopal G, Sabitha K, Shirley S, et al. Identification and validation of plasma biomarkers for diagnosis of breast cancer in South Asian women. Sci Rep. 2022;12(1):100.
11. Kumar V, Ray S, Ghantasala S, Srivastava S. An integrated quantitative proteomics workflow for cancer biomarker discovery and validation in plasma. Front Oncol. 2020;10:543997.
12. Metatla I, Roger K, Chhuon C, Ceccacci S, Chapelle M, Schmit PO, et al. Neat plasma proteomics: getting the best out of the worst. Clin Proteomics. 2024;21(1):22.
13. Zhang H, Chan DW. Cancer biomarker discovery in plasma using a tissue-targeted proteomic approach. Cancer Epidemiol Biomarkers Prev. 2007;16(10):1915–7.
14. Geraci F, Saha I, Bianchini M, Editorial. RNA-Seq Analysis: Methods, Applications and Challenges. Front Genet 2020; 11. 2020.
15. Danielson KM, Rubio R, Abderazzaq F, Das S, Wang YE. High throughput sequencing of extracellular RNA from human plasma. PLoS ONE. 2017;12(1):e0164644.
16. Rapin N, Bagger FO, Jendholm J, Mora-Jensen H, Krogh A, Kohlmann A, et al. Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. Blood. J Am Soc Hematol. 2014;123(6):894–904.
17. Arasu A, Balakrishnan P, Velusamy T. RNA sequencing analyses reveal differentially expressed genes and pathways as Notch2 targets in B-cell lymphoma. Oncotarget. 2020;11(48):4527.
18. Zhong P, Bai L, Hong M, Ouyang J, Wang R, Zhang X, et al. A comprehensive review on Circulating CfRNA in plasma: implications for disease diagnosis and beyond. Diagnostics. 2024;14(10):1045.
19. Sheng Q, Vickers K, Zhao S, Wang J, Samuels DC, Koues O, et al. Multi-perspective quality control of illumina RNA sequencing data analysis. Brief Funct Genomics. 2017;16(4):194–204.
20. Zhou Q, Su X, Jing G, Chen S, Ning K. RNA-QC-chain: comprehensive and fast quality control for RNA-Seq data. BMC Genomics. 2018;19:1–10.
21. Wang L, Wang J, Jia E, Liu Z, Ge Q, Zhao X. Plasma RNA sequencing of extracellular RNAs reveals potential biomarkers for non-small cell lung cancer. Clin Biochem. 2020;83:65–73.
22. Jin N, Kan CM, Pei XM, Cheung WL, Ng SSM, Wong HT, et al. Cell-free Circulating tumor RNAs in plasma as the potential prognostic biomarkers in colorectal cancer. Front Oncol. 2023;13:1134445.
23. Al-Sheikh YA, Ghneim HK, Alharbi KK, Aboul-Soud MA. Screening for differentially-expressed MicroRNA biomarkers in Saudi colorectal cancer patients by small RNA deep sequencing. Int J Mol Med. 2019;44(6):2027–36.
24. Chen S, Jin Y, Wang S, Xing S, Wu Y, Tao Y, et al. Cancer type classification using plasma cell-free RNAs derived from human and microbes. Elife. 2022;11:e75181.
25. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Med. 2017;9:1–12.
26. Pang L, Xiang F, Yang H, Shen X, Fang M, Li R, et al. Single-cell integrative analysis reveals consensus cancer cell States and clinical relevance in breast cancer. Sci Data. 2024;11(1):289.
27. Xu L, Saunders K, Huang SP, Knutsdottir H, Martinez-Algarin K, Terrazas I et al. A comprehensive single-cell breast tumor atlas defines epithelial and immune heterogeneity and interactions predicting anti-PD-1 therapy response. Cell Rep Med. 2024;5(5).
28. Prazanowska KH, Lim SB. An integrated single-cell transcriptomic dataset for non-small cell lung cancer. Sci Data. 2023;10(1):167.
29. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. Nat Commun. 2021;12(1):2540.
30. Zhou B, Guo W, Guo L, Li Y, Zheng Z, Huai Q, et al. Single-cell RNA-sequencing data reveals the genetic source of extracellular vesicles in esophageal squamous cell carcinoma. Pharmacol Res. 2023;192:106800.
31. Lin S, Zhou S, Han X, Yang Y, Zhou H, Chang X, et al. Single-cell analysis reveals exosome-associated biomarkers for prognostic prediction and immunotherapy in lung adenocarcinoma. Aging. 2023;15(20):11508.

32. Li Y, Yue L, Zhang S, Wang X, Zhu Y, nan, Liu J, et al. Proteomic, single-cell and bulk transcriptomic analysis of plasma and tumor tissues unveil core proteins in response to anti-PD-L1 immunotherapy in triple negative breast cancer. Comput Biol Med. 2024;176:108537.
33. Adil A, Kumar V, Jan AT, Asger M. Single-cell transcriptomics: current methods and challenges in data acquisition and analysis. Front NeuroSci. 2021;15:591122.
34. Maciejewski K, Czerwinska P. Scoping review: methods and applications of Spatial transcriptomics in tumor research. Cancers. 2024;16(17):3100.
35. Williams CG, Lee HJ, Asatsuma T, Vento-Tormo R, Haque A. An introduction to Spatial transcriptomics for biomedical research. Genome Med. 2022;14(1):68.
36. Yu Q, Jiang M, Wu L. Spatial transcriptomics technology in cancer research. Front Oncol. 2022;12:1019111.
37. Huang S, Ouyang L, Tang J, Qian K, Chen X, Xu Z, et al. Spatial transcriptomics: a new frontier in cancer research. Clin Cancer Bull. 2024;3(1):13.
38. The Galaxy platform. For accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 2022;50(W1):W345–51.
39. Batut B, van den Beek M, Doyle MA, Soranzo N. RNA-seq data analysis in galaxy. RNA Bioinf. 2021;367–92.
40. Jiang G, Zheng JY, Ren SN, Yin W, Xia X, Li Y, et al. A comprehensive workflow for optimizing RNA-seq data analysis. BMC Genomics. 2024;25(1):631.
41. Kader Z. Identifying potential biomarkers for colorectal cancer diagnosis using an RNA-seq analysis workflow. 2021.
42. Soni V, Dawson T, Lin L, Crandall K, Sherman J, Keidar M. High-throughput RNA-Seq and In-silico analysis of glioblastoma cells treated with cold atmospheric plasma and temozolomide. 2024.
43. Bhardwaj M, Begum F, Singh D, Krupanidhi S, Yadav VK, Sahoo DK, et al. Identification of biomarkers associated with paget's disease of bone and bone metastasis from breast cancer patients. Cancer Rep. 2024;7(9):e70003.
44. Saddiqa A, Zakir M, Sheikh M, Muneer Z, Hassan A, Ali I, et al. On discovery of novel hub genes for ER + and TN breast cancer types through RNA seq data analyses and classification models. Sci Rep. 2024;14(1):20840.
45. Clough E, Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. Nucleic Acids Res. 2024;52(D1):D138–44.
46. Kalaki NS, Razizadeh MH, Tameshkel FS, Marzidare AA, Babaei M, Sayad S, et al. The breast cancer biomarkers associated with the development of the disease; an in-silico-based study. Immunopathol Persa. 2024;11(1):e41722–41722.
47. Asadzadeh A, Behboodian B, Bernoos P, Shojaei Barjouei M. Identification of breast cancer biomarkers by constructing Protein-Protein interaction and miRNAs-mRNAs networks. Middle East J Cancer. 2025;16(1):39–48.
48. Öztan G. Gene expression profiling with transcriptomic data analysis in small cell lung cancer. Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi. 2024;11(2):276–84.
49. Akhtar A, Hameed Y, Ejaz S, Abdullah I. Identification of gastric cancer biomarkers through in-silico analysis of microarray based datasets. Biochem Biophys Rep. 2024;40:101880.
50. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
51. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.
52. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
53. Tong Y. The comparison of limma and DESeq2 in gene analysis. In EDP Sciences; 2021. p. 03058.
54. Qu H, Qu M, Wang S, Yu L, Jia Q, Wang X et al. Differential expression analysis: simple pair, interaction, time-series. Bio-101. 2022;e4455.
55. Liu S, Wang Z, Zhu R, Wang F, Cheng Y, Liu Y. Three differential expression analysis methods for RNA sequencing: limma, EdgeR, DESeq2. J Visualized Experiments (JoVE). 2021;(175):e62528.
56. Li D, Zand MS, Dye TD, Goniewicz ML, Rahman I, Xie Z. An evaluation of RNA-seq differential analysis methods. PLoS ONE. 2022;17(9):e0264246.
57. Chen M, Mithraprabhu S, Ramachandran M, Choi K, Khong T, Spencer A. Utility of Circulating cell-free RNA analysis for the characterization of global transcriptome profiles of multiple myeloma patients. Cancers. 2019;11(6):887.
58. Metzenmacher M, Váraljai R, Hegedüs B, Cima I, Forster J, Schramm A, et al. Plasma next generation sequencing and droplet digital-qPCR-based quantification of Circulating cell-free RNA for noninvasive early detection of cancer. Cancers. 2020;12(2):353.
59. Chen Q, Guo X, Wang H, Sun S, Jiang H, Zhang P, et al. Plasma-free blood as a potential alternative to whole blood for transcriptomic analysis. Phenomics. 2024;4(2):109–24.
60. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50(W1):W216–21.
61. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics. 2020;36(8):2628–9.
62. Yeh C, Lai HC, Grabbe N, Willett X, Lin ST. Revolutionizing detection of minimal residual disease in breast cancer using Patient-. Derived Gene Signature; 2025.
63. Hannan S, Hami I, Dey RK, Gupta SD. A systematic exploration of key candidate genes and pathways in the biogenesis of human gastric cancer: a comprehensive bioinformatics investigation. J Translational Gastroenterol. 2024;2(1):9–20.
64. Lim J, Kang M, Ahn YH, Cho MS, Lee JH, Kang JL, et al. Comprehensive profiling of serum exosomes by a Multi-Omics approach reveals potential diagnostic markers for brain metastasis in lung cancer. Cancers. 2025;17(12):1929.
65. Hannan NJ, Cohen PA, Beard S, Bilic S, Zhang B, Tong S, et al. Transcriptomic analysis of patient plasma reveals Circulating miR200c as a potential biomarker for high-grade serous ovarian cancer. Gynecologic Oncol Rep. 2022;39:100894.
66. Northrop-Albrecht EJ, Wu CW, Berger CK, Taylor WR, Foote PH, Doering KA, et al. An investigation of plasma cell-free RNA for the detection of colorectal cancer: from transcriptome marker selection to targeted validation. PLoS ONE. 2024;19(8):e0308711.
67. Rothfels K, Milacic M, Matthews L, Haw R, Sevilla C, Gillespie M, et al. Using the reactome database. Curr Protocols. 2023;3(4):e722.

68. Decruyenaere P, Daneels W, Morlion A, Verniers K, Anckaert J, Tavernier J, et al. Characterizing the Cell-Free transcriptome in a humanized diffuse large B-Cell lymphoma Patient-Derived tumor xenograft model for RNA-Based liquid biopsy in a preclinical setting. Int J Mol Sci. 2024;25(18):9982.
69. Agrawal A, Balcı H, Hanspers K, Coort SL, Martens M, Slenter DN, et al. WikiPathways 2024: next generation pathway database. Nucleic Acids Res. 2024;52(D1):D679–89.
70. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 2016;44(W1):W90–7.
71. Tosevska A, Morselli M, Basak SK, Avila L, Mehta P, Wang MB, et al. Cell-free RNA as a novel biomarker for response to therapy in head & neck cancer. Front Oncol. 2022;12:869108.
72. Szklarczyk D, Nastou K, Koutrouli M, Kirsch R, Mehryary F, Hachilif R, et al. The STRING database in 2025: protein networks with directionality of regulation. Nucleic Acids Res. 2025;53(D1):D730–7.
73. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46.
74. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape stringapp: network analysis and visualization of proteomics data. J Proteome Res. 2018;18(2):623–32.
75. Majeed A, Mukhtar S. Protein–Protein interaction network exploration using cytoscape. Protein-Protein interactions: methods and protocols. Springer; 2023. pp. 419–27.
76. Miryala SK, Anbarasu A, Ramaiah S. Discerning molecular interactions: a comprehensive review on biomolecular interaction databases and network analysis tools. Gene. 2018;642:84–94.
77. Hamdy H, Yang Y, Cheng C, Liu Q. Identification of potential hub genes related to aflatoxin B1, liver fibrosis and hepatocellular carcinoma via integrated bioinformatics analysis. Biology. 2023;12(2):205.
78. Ma Y, Guo J, Li D, Cai X. Identification of potential key genes and functional role of CENPF in osteosarcoma using bioinformatics and experimental analysis. Experimental Therapeutic Med. 2022;23(1):1–12.
79. Yang Z, Li J, Song H, Mei Z, Jia X, Tian X, et al. Unraveling the molecular links between benzopyrene exposure, NASH, and HCC: an integrated bioinformatics and experimental study. Sci Rep. 2023;13(1):20520.
80. Higurashi M, Ishida T, Kinoshita K. PiSite: a database of protein interaction sites using multiple binding States in the PDB. Nucleic Acids Res. 2009;37(suppl1):D360–4.
81. Veres DV, Gyurkó DM, Thaler B, Szalay KZ, Fazekas D, Korcsmáros T, et al. ComPPI: a cellular compartment-specific database for protein–protein interaction network analysis. Nucleic Acids Res. 2015;43(D1):D485–93.
82. Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, et al. The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. 2014;42(D1):D358–63.
83. Del Toro N, Shrivastava A, Ragueneau E, Meldal B, Combe C, Barrera E, et al. The intact database: efficient access to fine-grained molecular interaction data. Nucleic Acids Res. 2022;50(D1):D648–53.
84. Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. Genome Biol. 2010;11(5):R53.
85. Meng X, Wang J, Yuan C, Li X, Zhou Y, Hofestädt R, et al. CancerNet: a database for decoding multilevel molecular interactions across diverse cancer types. Oncogenesis. 2015;4(12):e177–177.
86. Zhang J, Pei J, Durham J, Bos T, Cong Q. Computed cancer interactome explains the effects of somatic mutations in cancers. Protein Sci. 2022;31(12):e4479.
87. Zgoda VG, Kopylov AT, Tikhonova OV, Moisa AA, Pyndyk NV, Farafonova TE, et al. Chromosome 18 transcriptome profiling and targeted proteome mapping in depleted plasma, liver tissue and HepG2 cells. J Proteome Res. 2013;12(1):123–34.
88. Oughtred R, Rust J, Chang C, Breitkreutz B, Stark C, Willems A, et al. The biogrid database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci. 2021;30(1):187–200.
89. Dovrou A, Bei E, Sfakianakis S, Marias K, Papanikolaou N, Zervakis M. Synergies of radiomics and transcriptomics in lung cancer diagnosis: a pilot study. Diagnostics. 2023;13(4):738.
90. Xing A, Tong HH, Liu S, Zhai X, Yu L, Li K. The causal association between obesity and gastric cancer and shared molecular signatures: a large-scale Mendelian randomization and multi-omics analysis. Front Oncol. 2023;13:1091958.
91. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochemist Reviews. 2008;29(Suppl 1):S83.
92. Li J. Area under the ROC curve has the most consistent evaluation for binary classification. PLoS ONE. 2024;19(12):e0316019.
93. Çorbacıoğlu ŞK, Aksel G. Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. Turkish J Emerg Med. 2023;23(4):195.
94. Hassanzad M, Hajian-Tilaki K. Methods of determining optimal cut-point of diagnostic biomarkers with application of clinical data in ROC analysis: an update review. BMC Med Res Methodol. 2024;24(1):84.
95. Franco-Pereira AM, Nakas CT, Pardo MC. Biomarker assessment in ROC curve analysis using the length of the curve as an index of diagnostic accuracy: the binormal model framework. AStA Adv Stat Anal. 2020;104:625–47.
96. Mohammadi D, Zafari Y, Estaki Z, Mehrabi M, Moghbelinejad S. Evaluation of plasma circ_0006282 as a novel diagnostic biomarker in colorectal cancer. J Clin Lab Anal. 2022;36(1):e24147.
97. Song Y, Cao P, Li J. Plasma circular RNA hsa_circ_0001821 acts as a novel diagnostic biomarker for malignant tumors. J Clin Lab Anal. 2021;35(11):e24009.
98. Jiang HG, Dai CH, Xu YP, Jiang Q, Xia XB, Shu Y, et al. Four plasma MiRNAs act as biomarkers for diagnosis and prognosis of non–small cell lung cancer. Oncol Lett. 2021;22(5):1–12.
99. Tao Y, Liu J, Qiu W, Li Y. LncRNA MANCR is downregulated in non-small cell lung cancer and predicts poor survival. Discover Oncol. 2025;16(1):40.
100. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res. 2011;12:2825–30.
101. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+to analyze and compare ROC curves. BMC Bioinformatics. 2011;12(1):77.
102. Šutić M, Dmitrović B, Jakovčević A, Džubur F, Oršolić N, Debeljak Ž, et al. Transcriptomic profiling for prognostic biomarkers in Early-Stage squamous cell lung cancer (SqCLC). Cancers. 2024;16(4):720.

103. Shi Y, Li Y, Yan C, Su H, Ying K. Identification of key genes and evaluation of clinical outcomes in lung squamous cell carcinoma using integrated bioinformatics analysis. Oncol Lett. 2019;18(6):5859–70.
104. Lánczky A, Győrffy B. Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. J Med Internet Res. 2021;23(7):e27633.
105. Zhao X, Yan H, Yan X, Chen Z, Zhuo R. A novel prognostic Four-Gene signature of breast cancer identified by integrated bioinformatics analysis. Dis Markers. 2022;2022(1):5925982.
106. Oshi M, Angarita FA, Tokumaru Y, Yan L, Matsuyama R, Endo I, et al. A novel three-gene score as a predictive biomarker for pathologically complete response after neoadjuvant chemotherapy in triple-negative breast cancer. Cancers. 2021;13(10):2401.
107. Abbas A, Hefnawy MT, Negida A. Meta-analysis accelerator: a comprehensive tool for statistical data conversion in systematic reviews with meta-analysis. BMC Med Res Methodol. 2024;24(1):243.
108. Dekkers OM. Meta-analysis: key features, potentials and misunderstandings. Res Pract Thromb Haemostasis. 2018;2(4):658–63.
109. Kane LE, Mellotte GS, Mylod E, O'Brien RM, O'Connell F, Buckley CE, et al. Diagnostic accuracy of blood-based biomarkers for pancreatic cancer: a systematic review and meta-analysis. Cancer Res Commun. 2022;2(10):1229–43.
110. Quintana DS. A guide for calculating study-level statistical power for meta-analyses. Adv Methods Practices Psychol Sci. 2023;6(1):25152459221147260.
111. Papakostidis C, Giannoudis PV. Meta-analysis. What Have We learned? Injury. 2023;54:S30–4.
112. Keel BN, Lindholm-Perry AK. Recent developments and future directions in meta-analysis of differential gene expression in livestock RNA-Seq. Front Genet. 2022;13:983043.
113. Padroni L, De Marco L, Dansero L, Fiano V, Milani L, Vasapolli P, et al. An epidemiological systematic review with Meta-Analysis on biomarker role of Circulating MicroRNAs in breast cancer incidence. Int J Mol Sci. 2023;24(4):3910.
114. Jasemi SV, Zia S, Mirbahari SG, Sadeghi M. A systematic review and meta-analysis to evaluate blood levels of interleukin-6 in lung cancer patients. Kardiochirurgia I Torakochirurgia Polska/Polish J Thorac Cardiovasc Surg. 2023;20(4):240–50.
115. Pal A, Dhar A, Shamim MA, Rani I, Negi RR, Sharma A et al. Selenium levels in colorectal cancer: A systematic review and Meta-Analysis of serum, plasma, and colorectal specimens. J Trace Elem Med Biol. 2024;127429.
116. Ssedyabane F, Obuku EA, Namisango E, Ngonzi J, Castro CM, Lee H et al. The diagnostic accuracy of serum and plasma MicroRNAs in detection of cervical intraepithelial neoplasia and cervical cancer: A systematic review and meta-analysis. Gynecologic Oncol Rep. 2024;101424.
117. Saeed S, Ahmed S, joseph shalom. Machine Learning in the Big Data Age: Advancements, Challenges, and Future Prospects. 2024.
118. Gui Y, He X, Yu J, Jing J. Artificial intelligence-assisted transcriptomic analysis to advance cancer immunotherapy. J Clin Med. 2023;12(4):1279.
119. Zhong S, Chen S, Lin H, Luo Y, He J. Selection of M7G-related LncRNAs in kidney renal clear cell carcinoma and their putative diagnostic and prognostic role. BMC Urol. 2023;23(1):186.
120. Mirza Z, Ansari MS, Iqbal MS, Ahmad N, Alganmi N, Banjar H, et al. Identification of novel diagnostic and prognostic gene signature biomarkers for breast cancer using artificial intelligence and machine learning assisted transcriptomics analysis. Cancers. 2023;15(12):3237.
121. Benkirane H, Pradat Y, Michiels S, Cournède PH, CustOmics. A versatile deep-learning based strategy for multi-omics integration. PLoS Comput Biol. 2023;19(3):e1010921.
122. Ballard JL, Wang Z, Li W, Shen L, Long Q. Deep learning-based approaches for multi-omics data integration and analysis. BioData Min. 2024;17(1):38.
123. Ozaki Y, Broughton P, Abdollahi H, Valafar H, Blenda AV. Integrating omics data and AI for cancer diagnosis and prognosis. Cancers. 2024;16(13):2448.
124. Tong L, Mitchel J, Chatlin K, Wang MD. Deep learning based feature-level integration of multi-omics data for breast cancer patients survival analysis. BMC Med Inf Decis Mak. 2020;20:1–12.
125. Li S, Zhang Y, Guan L, Dong Y, Zhang M, Zhang Q et al. Integrating explainable deep learning with multi-omics for screening progressive diagnostic biomarkers of hepatocellular carcinoma covering the inflammation-cancer transformation. J Pharm Anal. 2025;101253.
126. Wei H, Luo S, Bi Y, Liao C, Lian Y, Zhang J, et al. Plasma microRNA-15a/16-1-based machine learning for early detection of hepatitis B virus-related hepatocellular carcinoma. Liver Res. 2024;8(2):105–17.
127. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023;23(1):689.
128. Sanches PHG, de Melo NC, Porcari AM, de Carvalho LM. Integrating molecular perspectives: strategies for comprehensive Multi-Omics integrative data analysis and machine learning applications in Transcriptomics, Proteomics, and metabolomics. Biology. 2024;13(11):848.
129. Hong J, Eun JW, Baek GO, Cheong JY, Park S, Kim SS, et al. Multiomics profiling of Buffy coat and plasma unveils etiology-specific signatures in hepatocellular carcinoma. Clin Mol Hepatol. 2024;30(3):360.
130. Yao F, Yan C, Zhang Y, Shen L, Zhou D, Ni J. Identification of blood protein biomarkers for breast cancer staging by integrative transcriptome and proteome analyses. J Proteom. 2021;230:103991.
131. Wang C, Liu H. Factors influencing degradation kinetics of mRNAs and half-lives of microRNAs, circRNAs, LncRNAs in blood in vitro using quantitative PCR. Sci Rep. 2022;12(1):7259.
132. Shi H, Zhou Y, Jia E, Pan M, Bai Y, Ge Q. Bias in RNA-seq library preparation: current challenges and solutions. Biomed Res Int. 2021;2021(1):6647597.
133. Wong RK, MacMahon M, Woodside JV, Simpson DA. A comparison of RNA extraction and sequencing protocols for detection of small RNAs in plasma. BMC Genomics. 2019;20:1–12.
134. Sun J, Yang X, Wang T, Xing Y, Chen H, Zhu S et al. Impact of blood storage conditions on the transcript profile of plasma cell-free RNA. bioRxiv. 2021;2021–03.
135. Mohr AE, Ortega-Santos CP, Whisner CM, Klein-Seetharaman J, Jasbi P. Navigating challenges and opportunities in multi-omics integration for personalized healthcare. Biomedicines. 2024;12(7):1496.

136. Vidanagamachchi SM, Waidyarathna K. Opportunities, challenges and future perspectives of using bioinformatics and artificial intelligence techniques on tropical disease identification using omics data. Front Digit Health. 2024;6:1471200.

Publisher's note