Original Article Open Access

Archives Of Breast Cancer

DOI: 10.32768/abc.7394826150-371

Circulating MicroRNA Signatures as Prognostic Biomarkers in Breast Cancer: A Comprehensive Analysis of Expression Profiles, Functional Targets, and Survival Correlations

Adam Neama Al-Ebadi[a], Rasha Abd Ali Al-Khalidi*[b,c]

[a] Biotechnology Department, College of Science, University of Baghdad, Baghdad, Iraq
[b] Department of Biotechnology, College of Science, University of Baghdad, Baghdad, Iraq
[c] Biomedical Application Department, Artificial Intelligence, University of Baghdad, Baghdad, Iraq

* **Corresponding Author:** Rasha Abd Ali Al-Khalidi. Email: rasha.ali@sc.uobaghdad.edu.iq

ARTICLE INFO

Received:
17 April 2026
Revised:
01 July 2026
Accepted:
11 July 2026

Keywords:
liquid biopsy, epigenomics, transcriptome, survival analysis, gene expression profiling, computational biology

ABSTRACT

Background: Circulating microRNAs are promising noninvasive biomarkers in oncology. Notable gaps remain in relevant serum-focused research on early-stage breast cancer; therefore, this study integrated transcriptomic data to identify a signature with both diagnostic and prognostic value.

Methods: We retrieved serum microRNA data from the publicly available microarray data set GSE73002, which includes 1270 patients with early-stage breast cancer and 2686 healthy controls. We used limma's empirical Bayes method to analyze differential expression, predicted target genes using miRWalk, performed functional enrichment analysis, and assessed prognostic relevance using survival data from the TCGA-BRCA tissue cohort as a supportive, non-independent line of evidence.

Results: At an FDR threshold of 0.05, we identified 1450 upregulated microRNAs and 898 downregulated microRNAs. Core upregulated microRNAs, including hsa-miR-614, showed enrichment of predicted target genes, and high hsa-miR-614 expression was associated with shorter overall survival (HR, 2.81; 95% CI, 1.24-6.39; $P = .0098$).

Conclusion: Early breast tumorigenesis may be characterized by systemic dysregulation of circulating microRNAs that govern epigenetic and transcriptional networks. These specific signatures, notably hsa-miR-614, may serve as promising candidates for liquid biopsy-based diagnostic and prognostic biomarkers, pending further experimental and prospective clinical validation. The survival association reported here was derived from a tissue-based cohort (TCGA-BRCA) and should be interpreted as supportive evidence rather than independent validation in circulating samples; prospective studies using independent serum or plasma cohorts are needed to confirm prognostic utility in liquid biopsy settings.

© 2026 The Author(s). This open-access article is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License, which permits sharing and adapting the material for any purpose except commercial purposes, provided appropriate credit is given, a link to the license is provided, and any changes are indicated.

INTRODUCTION

Breast cancer is the most common cancer worldwide, and it is also the leading cause of cancer-related deaths among women.[1] At present, the core diagnostic methods used in clinical practice center on imaging and tissue biopsy, but these approaches face two key problems: they pose significant challenges for early diagnosis, and they lack sufficient accuracy for prognostic stratification.[1] In particular, early-stage lesions exhibit only molecular changes, so the academic community urgently needs minimally invasive biomarkers to address this gap.

MicroRNAs (miRNAs) are small noncoding RNAs approximately 22 nucleotides long that regulate gene expression at a posttranscriptional level and participate in the processes of tumorigenesis, including, but not limited to, proliferation, invasion, immune modulation, and angiogenesis.[2–4] Serum- and plasma-based circulating miRNAs display substantial stability and are, therefore, ideal candidates for use as liquid biopsy-based indicators of cancer diagnosis or prognosis.[4–6] Numerous studies have previously advocated for miRNA signatures for breast cancer; however, many reports are limited by small sample size, lack of independent validation,

and inadequate consideration of the integration of differential expression with functional inferences or clinical outcome evaluation.[7–11]

Accordingly, a significant gap remains in the implementation of an integrative end-to-end data-driven framework that links large-cohort differential expression profiling with target inference, pathway-level interpretation, interaction network context, and supportive survival analysis.[12] Therefore, the aim of this study was to use a bioinformatic workflow to discover differentially expressed circulating miRNAs in early-stage breast cancer, characterize their predicted target genes and the enriched biological pathways associated with them, and explore their prognostic relevance using TCGA-BRCA survival data.

METHODS

Data acquisition and study design

This study used publicly available serum miRNA expression profiles obtained from the Gene Expression Omnibus (GEO) database (accession number GSE73002) (View GEO Dataset GSE73002). The data set comprised microarray data generated using the highly sensitive 3D-Gene Human miRNA V20_1.0.0 platform (Toray Industries, Japan; GEO platform GPL18941). The initial dataset included 4684 serum samples. To ensure a case-control comparison, rigorous inclusion and exclusion criteria were applied based on the provided clinical metadata. The final analytical cohort consisted of 1,270 samples from patients with early-stage breast cancer and 2,686 samples from healthy controls. Samples were excluded from analysis based on the following explicit criteria: (i) patients with benign breast diseases (n = 53); (ii) patients with other malignancies (n = 514); and (iii) samples with incomplete or missing clinical annotations. The designation of early-stage breast cancer was adopted from the original GSE73002 study design and annotations, which focused on identifying serum miRNA biomarkers for early breast cancer detection. Breast cancer and healthy control samples were selected based on the disease-status labels in the GEO metadata (Shimomura et al., Cancer Sci, 2016). Molecular subtype, age, sex, and sample collection date were not used as exclusion criteria, as complete individual-level demographic metadata were not uniformly available in the GEO submission. Institutional review board approval and patient consent were not required, as this study entailed the secondary analysis of publicly available, deidentified data.

Preprocessing and normalization of microarray data

Gene expression microarray data were preprocessed using R (version 4.1.2). First, $\log_2$ transformation and quantile normalization were applied to reduce technical variation and stabilize variance across arrays, after which quality control was carried out to assess sample structure and identify technical outliers. Because all samples were generated using standardized experimental workflows on the same microarray platform, no batch effect correction was applied. Ultimately, all 3956 valid samples (1270 breast cancer cases and 2686 controls) were included in the downstream differential expression analysis.

Differential expression analysis

Differential miRNA expression between the breast cancer and healthy control groups was quantified using the empirical Bayes method implemented in the limma framework. Linear models were fitted to the normalized expression data using the lmFit function, and a specific contrast was designed to compare the ”Tumor” group with the ”Non-Cancer” control group. Moderated t-statistics and $\log_2$ fold changes (logFC) were calculated using the eBayes function. Differentially expressed miRNAs were identified using a Benjamini–Hochberg false discovery rate (FDR) threshold of 0.05. Upregulated miRNAs were defined as those with logFC > 0, whereas downregulated miRNAs were defined as those with logFC < 0.

Given the large cohort size (n = 3,956), no arbitrary fold-change threshold was applied to avoid excluding consistently deregulated miRNAs exhibiting modest but potentially biologically relevant expression changes. Significantly differentially expressed miRNAs were subsequently ranked according to the absolute magnitude of their logFC values. The top 10 upregulated and top 10 downregulated miRNAs were selected for downstream analyses, prioritizing the most strongly deregulated candidates while maintaining analytical feasibility for downstream functional analyses. This prioritization strategy was adopted to identify high-confidence candidate biomarkers for subsequent independent validation and detailed functional investigation. Fold-change values were calculated as $2^{\log_2 FC}$. To assess the robustness of this prioritization strategy, a sensitivity analysis was performed using expanded ranking thresholds (top 15 and top 20 differentially expressed miRNAs ranked by absolute logFC). The resulting candidate lists were compared with the original top 10 selection.

Target gene prediction and prioritization

To elucidate the potential biological roles of the top 10 significantly upregulated miRNAs, putative mRNA targets were predicted using the miRWalk v3.0 database. To minimize the inclusion of false-positive interactions, the target prediction was refined by intersecting the results with interactions recorded or predicted in at least one additional database (eg, miRTarBase or TargetScan) integrated within the miRWalk platform. The predicted target genes were then mapped to their respective binding regions, including the 3′-untranslated region (3′-UTR), 5′-untranslated region (5′-UTR), and coding sequence (CDS). Target genes were designated high-confidence candidate regulatory hubs if they satisfied at least one biological support criterion: (i) being targeted by ≥2 of the top 10 upregulated miRNAs, or (ii) having predicted interactions in ≥2 distinct binding regions (3′-

UTR, 5′-UTR, or CDS) within the same gene. In all cases, candidate targets were required to be supported by intersection of miRWalk predictions with at least one additional database (miRTarBase or TargetScan).

Functional enrichment

Functional annotation of the prioritized high-confidence target genes was performed using the Enrichr web server. The analysis included Gene Ontology (GO) terms (Biological Process, Molecular Function, and Cellular Component) from the GO 2025 database, as well as pathway enrichment analyses using the Kyoto Encyclopedia of Genes and Genomes (KEGG 2021), Reactome, and WikiPathways databases. Statistical significance for enrichment analysis was assessed using Benjamini–Hochberg adjusted *P* values (FDR-adjusted $P < .05$) as reported by the Enrichr platform. Results were further prioritized by the Enrichr combined score ($-\log_{10}$ *P*-value × z-score). The enriched functional categories and pathways were subsequently visualized using standardized bar charts and clustergrams to illustrate the relationship between the target genes and their associated biological themes.

Construction of the miRNA-mRNA interaction network

An miRNA–mRNA regulatory network was constructed using the top upregulated miRNAs and their high-confidence functional target genes. Comprising 164 miRNA–target pairs, the network was generated via miRNet 2.0, a comprehensive web-based platform that integrates data from miRTarBase, TarBase, and miRDB. This interaction network was exported and visualized in Cytoscape (version 3.9.1) to identify central nodes and regulatory hubs in the breast cancer-associated miRNA signature. To determine whether there are patterns of network regulation by candidate miRNAs, the miRNA-mRNA interaction network was constructed using hierarchical clustering and presented as a heatmap.

Survival analysis and prognostic validation

To evaluate the prognostic significance of the identified circulating miRNAs, we used the miRPower module of Kaplan Meier Plotter (View Kaplan Meier Plotter), which integrates transcriptomic and clinical outcome data from The Cancer Genome Atlas breast invasive carcinoma cohort (TCGA BRCA; View GDC Data Portal). Patients were stratified into high- and low-expression groups based on the median expression level of each specific miRNA. No optimal cut-point selection algorithm was employed to prevent statistical overfitting. Overall survival (OS) was assessed using Kaplan-Meier survival curves, and the statistical difference between the survival distributions was determined using the log-rank test. Hazard ratios (HR) and corresponding 95% confidence intervals (CIs) were calculated in a univariate framework without adjustment for clinicopathological covariates. A log-rank *P* value $< .05$ was considered statistically significant. It should be noted that this survival analysis is based on tissue-derived expression data from the TCGA-BRCA cohort and, therefore, constitutes a supportive line of evidence regarding the clinical relevance of hsa-miR-614, rather than an independent validation of its circulating expression pattern.

Statistical analysis and data visualization

All statistical analyses were conducted using R. The supplied R script for miRNA microarray preprocessing, differential-expression analysis and volcano plotting is included in the Supplementary Material. Data visualization, including volcano plots, heatmaps, and bar charts, was performed using the ggplot2, EnhancedVolcano, and pheatmap packages. The raw data and processed expression matrices from the GSE73002 dataset are publicly accessible via the GEO repository. All customized analytical scripts and derived data tables supporting the findings of this study will be deposited in a public GitHub repository and assigned a formal digital object identifier (DOI) via Zenodo to ensure complete computational reproducibility.

RESULTS

The GEO dataset GSE73002 was used, comprising 4684 serum samples tested on the 3D-Gene Human miRNA V20_1.0.0 platform. After applying the inclusion criteria, the 1270 breast cancer samples and the 2686 healthy control samples were retained. To ensure cohort homogeneity, samples associated with benign breast disease, other non-malignant conditions, or non breast cancer types were excluded. This filtering approach provided a very strong case-control comparison for patients with breast cancer and healthy controls (Figure 1).

Differential expression analysis using the limma package identified numerous miRNAs that were significantly deregulated between serum samples from breast cancer and healthy controls (FDR < 0.05). In total, 1450 miRNAs showed positive $\log_2$ fold changes and 898 showed negative $\log_2$ fold changes versus controls, while 192 were not statistically significant. The large sample size (n = 3,956) enabled the limma analysis to detect consistent and significant differences in circulating miRNA expression between groups. Most observed $\log_2$ fold changes were modest, suggesting coordinated systemic dysregulation rather than extreme expression alterations. Therefore, subsequent analyses focused on the most strongly deregulated and top-ranked miRNAs (Figure 2).

The volcano plot (Figure 2A) demonstrates a clear predominance of upregulated miRNAs. The top 10 most deregulated miRNAs are shown in Figure 2B, including hsa-miR-4532, hsa-miR-614, and hsa-miR-4734 among the upregulated miRNAs, and hsa-miR-3157-3p, hsa-miR-4776-5p, and hsa-miR-8067 among the downregulated miRNAs.

Sensitivity analysis demonstrated complete retention of all top 10 miRNAs within both top 15 and top 20 candidate sets, indicating that the principal candidate set were robust to alternative ranking thresholds.

For the top 10 miRNAs significantly upregulated

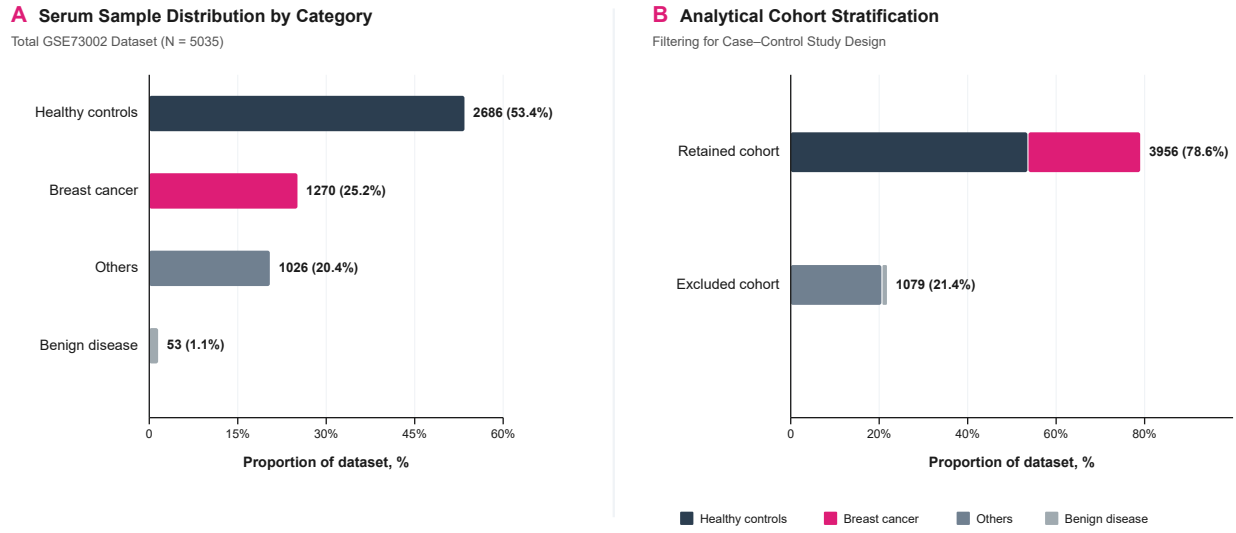


Figure 1. Study Design and Sample Filtering Flowchart for Dataset GSE73002.
Distribution of serum samples from the GSE73002 dataset following quality control and eligibility filtering. A, Overall sample classification across 4 primary categories: healthy controls (n = 2686 [53.4%]), breast cancer (n = 1270 [25.2%]), other non-breast conditions (n = 1026 [20.4%]), and benign breast disease (n = 53 [1.1%]). B, Cohort stratification retaining early-stage breast cancer cases and healthy controls (n = 3956 [78.6%]) for case-control differential expression analysis while excluding non-eligible disease categories (n = 1079 [21.4%]).

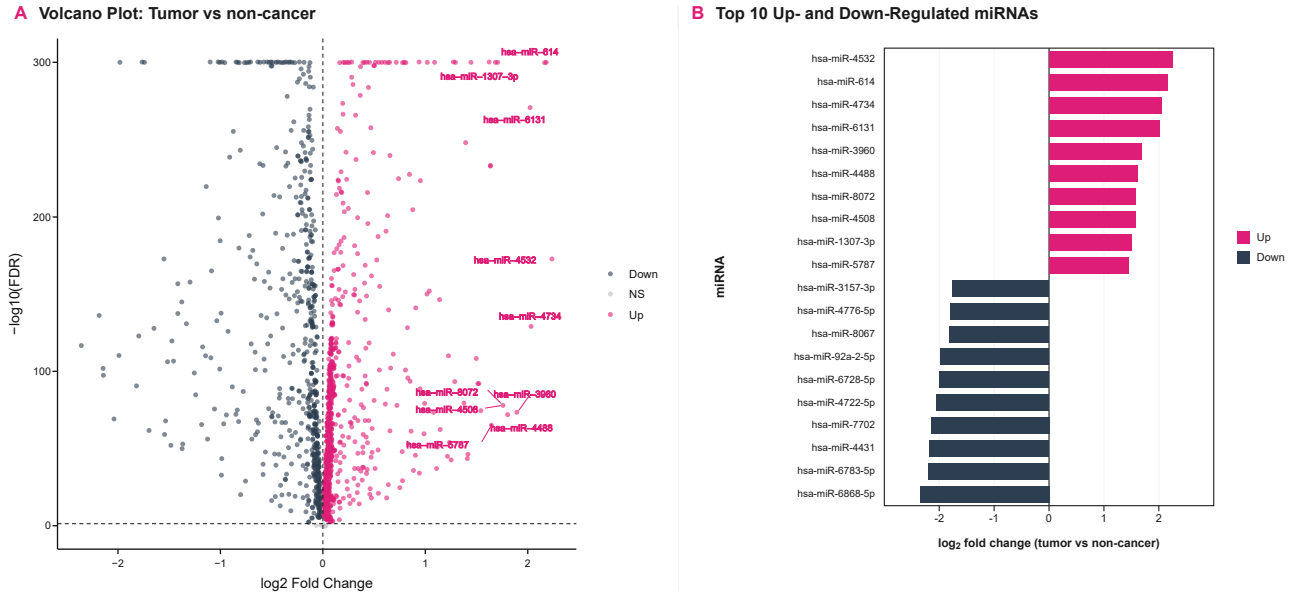


Figure 2. Differential Expression of Circulating miRNAs in Serum Samples From Patients With Breast Cancer vs Healthy Controls.
A, Volcano plot of differentially expressed serum miRNAs. The x-axis indicates $\log_2$ fold change and the y-axis indicates $-\log_{10}$(FDR). Upregulated miRNAs ($\log_2$ fold change > 0, FDR < .05) are shown in red and downregulated miRNAs ($\log_2$ fold change < 0, FDR < .05) are shown in blue. B, Horizontal bar chart of the top 10 upregulated and top 10 downregulated miRNAs ranked by absolute $\log_2$ fold change. Top upregulated candidates include hsa-miR-4532, hsa-miR-614, and hsa-miR-4734 (magenta bars, $\log_2$ fold change > 1.4); top downregulated candidates include hsa-miR-6868-5p, hsa-miR-6783-5p, and hsa-miR-4431 (dark slate bars, $\log_2$ fold change < −1.7).

in breast cancer vs healthy control samples, we performed target prediction using miRWalk and filtered the results by retaining only overlapping predictions from two independent databases to minimize false-positive interactions. The predicted targets were distributed across mRNA binding regions, with the majority mapping to the 3′-UTR and CDS, whereas relatively few targets were assigned to the 5′-UTR (Figure 3A). Notably, a limited set of genes appeared repeatedly across regions, indicating convergent post-transcriptional regulation by multiple miRNAs. To prioritize putative regulatory hubs, target genes were ranked by the number of top-ranking miRNAs predicted to regulate them, revealing a small

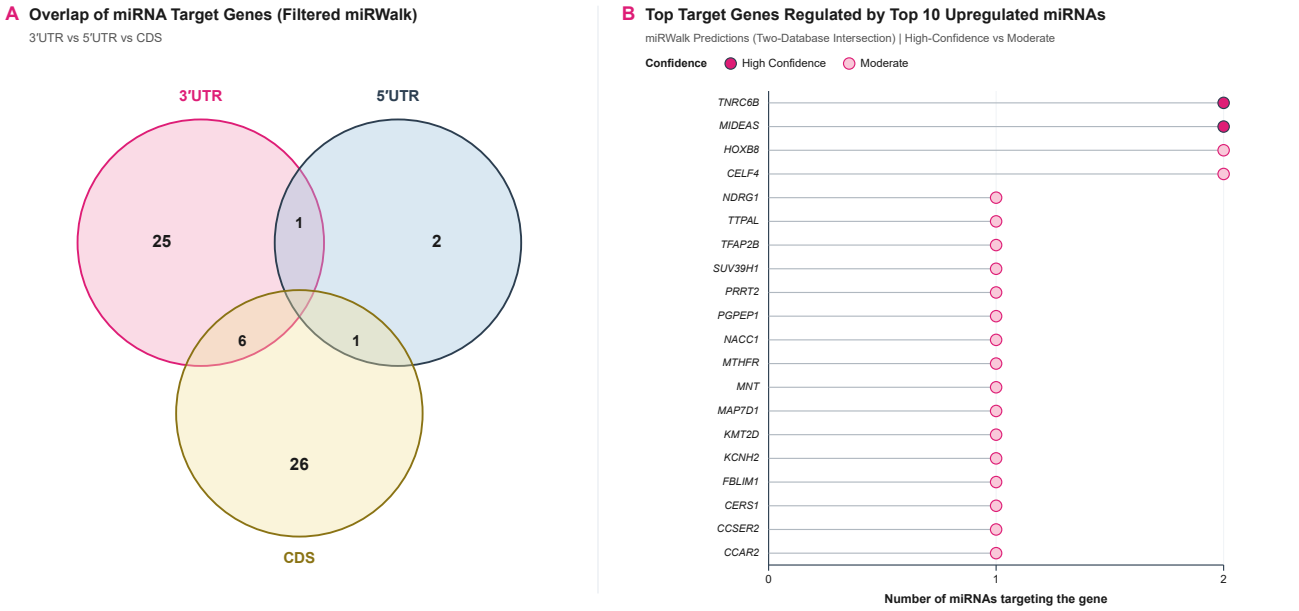


Figure 3. Predicted Target Genes of the Top 10 Upregulated miRNAs and Prioritization of Regulatory Hubs. A, Venn diagram illustrating the distribution and overlap of predicted mRNA target genes across binding regions (3′-UTR, 5′-UTR, and CDS) for the top 10 upregulated miRNAs based on miRWalk predictions filtered by two-database intersection. B, Lollipop plot showing prioritization of recurrent target genes ranked by the number of regulating miRNAs. Transcripts regulated by ≥2 miRNAs and supported across distinct binding regions were categorized as high-confidence candidate regulatory hubs (dark pink points; eg, *TNRC6B*, *MIDEAS*, *HOXB8*, *CELF4*).

subset of genes that were recurrently targeted (Figure 3B). Genes classified as high-confidence targets (i.e., those targeted by ≥2 miRNAs or supported across multiple binding regions) were designated as high-confidence candidate regulatory hubs based on recurrent targeting patterns.

The miRNA–mRNA interaction heatmap illustrates the presence or absence of predicted regulatory interactions between the top 10 upregulated miRNAs and their prioritized target genes, based on filtered miRWalk predictions. Rows represent miRNAs and columns represent target genes, with colored cells indicating predicted interactions. The heat map does not show random miRNA targeting of genes, with one class of miRNAs specifically targeting certain genes rather than others. Importantly, several miRNAs have been shown to target multiple genes, suggesting that these miRNAs are regulatory hubs. The similarities in targeting patterns are further accentuated by hierarchical clustering of targeting patterns, indicating common tendencies in targeting regulation within subsets of the most upregulated miRNAs (Figure 4).

Gene Ontology (GO) enrichment analysis was conducted using Enrichr to functionally characterize high- and moderate-confidence predicted mRNA targets of the top upregulated miRNAs (Figure 5). The analysis revealed significant enrichment of GO biological process terms related to transcriptional regulation and cellular stress responses, including negative regulation of RNA biosynthesis and responses to hypoxia and xenobiotic stimuli.

GO molecular function analysis indicated predominant enrichment of transcription-related activities, particularly cis-regulatory region binding, sequence-specific DNA binding, and chromatin-associated enzymatic functions, such as histone methyltransferase activity.

GO cellular component analysis showed enrichment of nuclear and chromatin-associated structures, as well as cytoskeletal and membrane-related complexes, suggesting a broad subcellular distribution of predicted miRNA targets.

Overall, the enrichment results indicate that the upregulated miRNAs converge on the posttranscriptional regulation of genes involved in transcriptional control, chromatin organization, and stress-related cellular processes. Clustergrams illustrating the relationships between enriched GO terms and predicted target genes are provided in Supplementary Figure S1.

Pathway enrichment analysis was performed using Enrichr to further characterize the functional pathways associated with high- and moderate-confidence predicted mRNA targets of the upregulated miRNAs (Figure 6). Reactome pathway analysis revealed significant enrichment of pathways related to RNA metabolism and post-transcriptional regulation, including processing of capped intron-containing pre-mRNA, post-transcriptional silencing by small RNAs, and regulation of RNA stability. Additional enriched Reactome pathways were involved in transcriptional regulation and cell communication.

Consistently, pathways related to histone modification, RNA polymerase II–mediated transcription, folate metabolism, sphingolipid metabolism, and cellular stress-related signaling were significantly enriched across all analyses performed using the WikiPathways platform. Pathway enrichment clustergrams are pro-

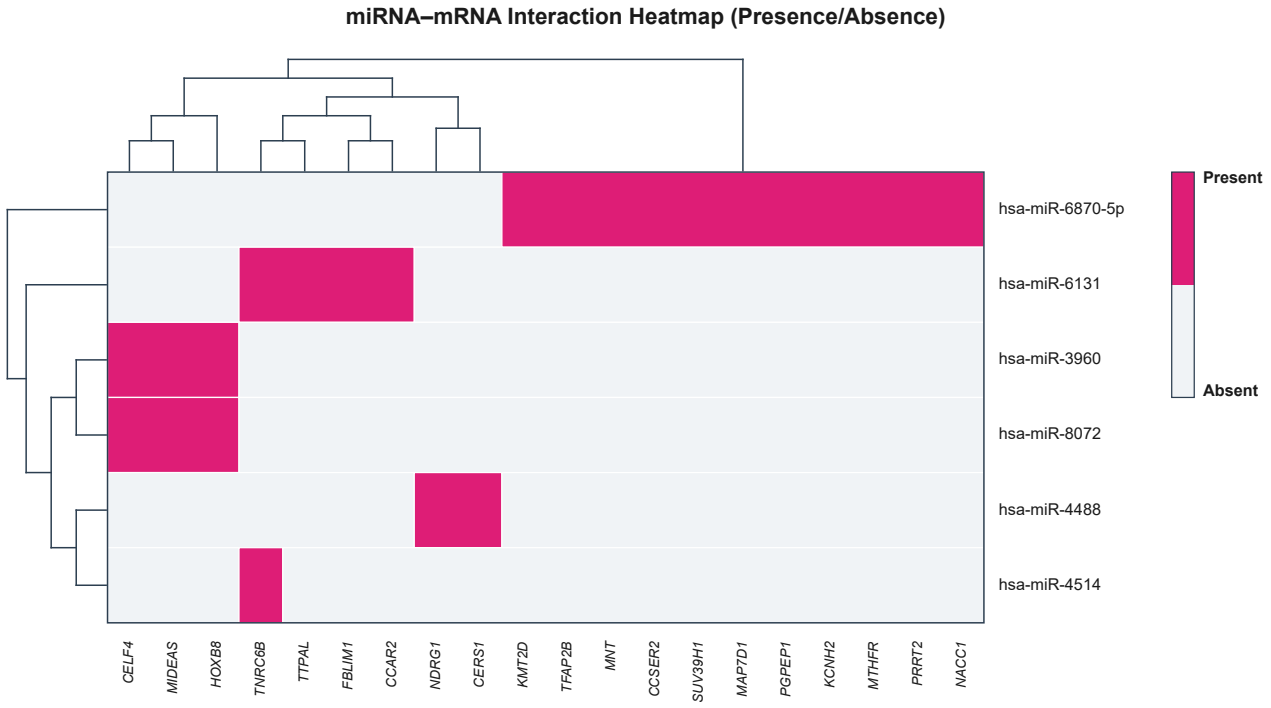


Figure 4. Predicted miRNA–mRNA Interaction Heatmap and Hierarchical Clustering.
Binary interaction heatmap displaying predicted regulatory interactions between the top upregulated circulating miRNAs (rows) and prioritized target genes (columns). Magenta filled cells indicate the presence of a predicted interaction, and light gray cells indicate absence. Hierarchical clustering dendrograms highlight co-regulatory miRNA clusters and target gene modules with convergent miRNA regulation (eg, *TNRC6B*, *KMT2D*, and *SUV39H1*).

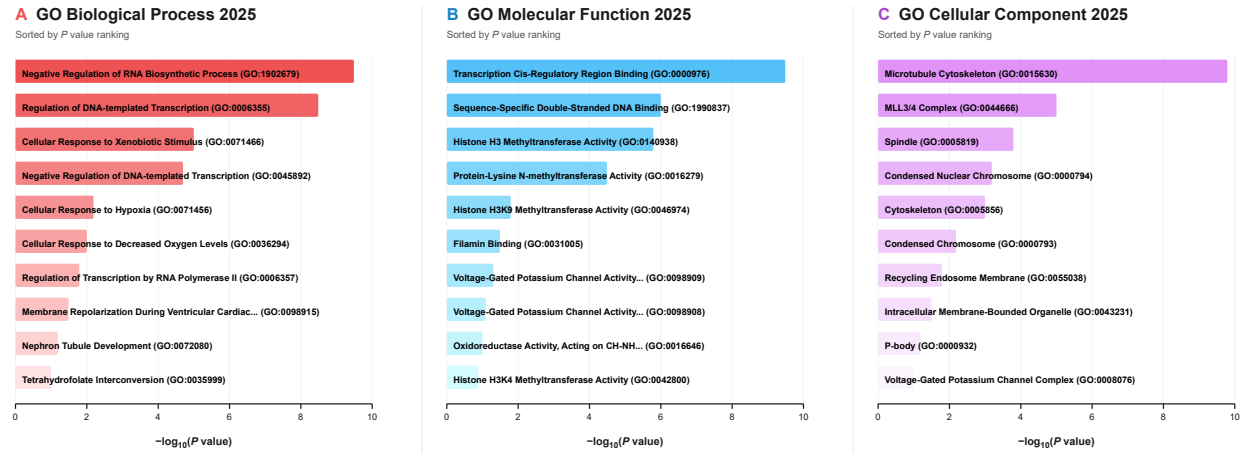


Figure 5. Gene Ontology (GO) Functional Enrichment Analysis of Predicted miRNA Target Genes.
Horizontal bar plots showing top enriched terms across Gene Ontology categories identified by Enrichr ($P < .05$): A, GO Biological Process terms, highlighting negative regulation of RNA biosynthesis, transcriptional control, and cellular stress responses; B, GO Molecular Function terms, emphasizing cis-regulatory region binding, sequence-specific DNA binding, and histone methyltransferase activity; and C, GO Cellular Component terms, showing enrichment of nuclear, chromatin-associated, and cytoskeletal structures. Bars are ordered by statistical significance ($-\log_{10}$ *P* value).

vided in Supplementary Figure S2. Taken together, the pathway-level enrichments provide evidence for predicted miRNA targets to modulate transcription, epigenetic regulation, RNA processing, and metabolic pathways, suggesting converging post-transcriptional control points regulated by upregulated miRNAs.

Kaplan–Meier survival analysis demonstrated that high hsa-miR-614 expression was significantly

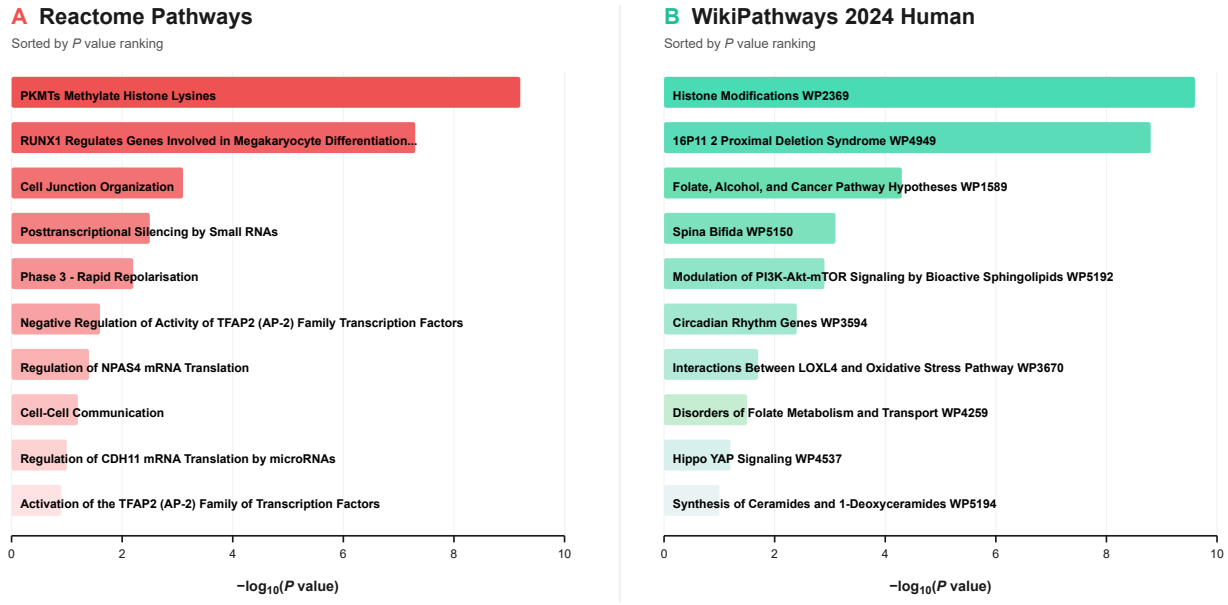


Figure 6. Pathway Enrichment Analysis of Predicted miRNA Target Genes.
Horizontal bar plots showing significantly enriched pathways identified by Enrichr ($P < .05$) ranked by $-\log_{10}$ P value: A, Reactome pathways, highlighting pre-mRNA processing, posttranscriptional silencing by small RNAs, and cell-cell communication; and B, WikiPathways (2024 Human), emphasizing histone modifications, 16p11.2 deletion syndrome, folate/alcohol pathways, and bioactive sphingolipid signaling.

associated with poorer overall survival in patients with breast cancer (HR, 2.81; 95% CI, 1.24-6.39; log-rank P = .0098). Patients were stratified according to median miRNA expression levels. This analysis was performed in a univariate framework using the Kaplan–Meier Plotter (miRPower) platform and did not include adjustment for clinicopathological covariates such as age or tumor stage. Elevated miR-614 expression was therefore associated with unfavorable clinical outcomes (Figure 7).

DISCUSSION

In the present study, a comprehensive bioinformatics framework was employed to systematically characterize the dysregulation of circulating miRNAs in early-stage breast cancer using large-scale serum microarray profiling. Differential expression analysis revealed a distinct bias toward the upregulation of circulating miRNAs, whereas the functional enrichment of their predicted targets converged on the transcriptional regulation, chromatin remodeling, and RNA metabolism pathways. Furthermore, survival modeling using a tissue-based TCGA-BRCA cohort identified hsa-miR-614 as a significant predictor of adverse prognosis. These results together suggest that the changed miRNA expression profile in the early stages of breast cancer development could be part of a larger system of regulation than a simple genetic aberration, though further validation experiments are necessary to prove this.

The observed upregulation of serum miRNAs is consistent with previous liquid biopsy studies, which have demonstrated active miRNA secretion during early-stage cancer development.[13,14] This process

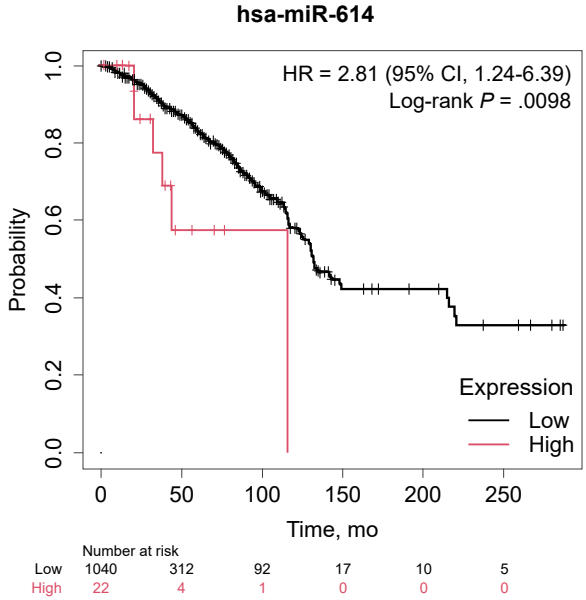


Figure 7. Kaplan–Meier Overall Survival Analysis of High vs Low hsa-miR-614 Expression in Breast Cancer (TCGA Cohort).
Overall survival curves generated via Kaplan–Meier Plotter for patients with breast invasive carcinoma in the TCGA cohort stratified by median hsa-miR-614 expression level. High hsa-miR-614 expression (red curve) was significantly associated with poorer overall survival compared with low expression (black curve) (univariate hazard ratio, 2.81 [95% CI, 1.24-6.39]; log-rank P = .0098). The table below the plot indicates the number of patients at risk at 50-month follow-up intervals.

occurs partly due to the packaging of miRNAs into EVs or their binding to proteins such as Ago2, which protect them from RNase cleavage and helps maintain their integrity in blood plasma.[15] The secretion of oncogenic miRNAs (oncomiRs) into the blood circulation is increasingly considered a strategy for intercellular communication, microenvironmental remodeling, and stress responses in many types of cancers.[8,16] Herein, we demonstrate active molecular regulatory reprogramming associated with serum miRNA levels in patients with early-stage breast cancer, thereby emphasizing the potential of miRNAs as diagnostic biomarkers of this disease.[5]

The predictive map of target enrichment for the highly upregulated microRNAs showed a strong correlation with biological pathways related to epigenetics and RNA processing, rather than with conventional proliferation pathways such as PI3K/AKT and MAPK. Notably, the computationally predicted target set included key chromatin modifiers, such as *KMT2D* and *SUV39H1*, which are established contributors to breast cancer progression, transcriptional plasticity, and therapy resistance.[17–19] However, the specific miRNA–mRNA interactions involving *KMT2D* and *SUV39H1* identified in the present study are based on in silico prediction using miRWalk with multi-database filtering and have not been experimentally validated. Therefore, these observations should be interpreted as hypothesis-generating rather than confirmed regulatory relationships. Alterations in histone methyltransferase activity and chromatin structure are increasingly recognized as early drivers of tumorigenesis that facilitate phenotypic adaptation.[20] Additionally, the enrichment of RNA metabolism pathways, particularly small RNA-mediated silencing, suggests feedback amplification loops within RNA regulatory networks, potentially involving core components of the silencing complex, such as *TNRC6*/*GW182*.[21] These findings suggest that circulating miRNA dysregulation may be associated with broader transcriptomic and epigenetic alterations.

Among the highly deregulated miRNAs, hsa-miR-614 demonstrated a significant association with poor overall survival in the TCGA-BRCA tissue cohort, providing supportive evidence rather than independent validation of its circulating serum expression. This finding is supported by experimental evidence demonstrating that elevated miR-614 expression enhances breast cancer cell motility and promotes epithelial-mesenchymal transition (EMT)-driven phenotypes.[22] Furthermore, miR-614 upregulation has been linked to increased cell proliferation and anti-apoptotic signaling in other malignancies, including ovarian and gastric cancers.[23] The observed association between early serum upregulation and adverse survival outcomes supports the biological plausibility of miR-614 as a marker of aggressive tumor behavior. This supportive survival analysis used TCGA-BRCA tissue-derived transcriptomic data.[11,24] An important consideration in interpreting these findings is the biological distinction between circulating serum miRNAs and tissue-derived miRNA expression profiles. While there have been reports in several investigations indicating the existence of partial correlation between serum miRNAs and tumor-associated miRNAs, it is not always a straightforward relationship. Serum miRNAs could have several sources, including tumor cells, stromal cells, immune cells, and non-cancerous host cells. Furthermore, serum miRNAs may be selectively released from various cell types either via extracellular vesicles or as protein-bound complexes (eg, with AGO2). Therefore, the expression patterns observed in tissues may not reflect the actual serum miRNA levels. Survival analysis done with TCGA-BRCA tissues should thus only provide a relevant biological context and not confirm the activity of serum miRNAs. An additional limitation of the survival analysis is the markedly unbalanced distribution of miR-614 expression within the TCGA-BRCA cohort, which resulted in a relatively small high-expression subgroup following the default Kaplan–Meier Plotter stratification. Although the association remained statistically significant, the limited number of patients in the high-expression category may reduce the stability of subgroup-specific estimates. Consequently, these findings should be interpreted as exploratory prognostic evidence and warrant confirmation in larger prospective cohorts with matched serum and tissue miRNA profiling.

In addition to their potential prognostic utility, the findings of this study hold considerable significance for the translation of hsa-miR-614 into a useful, clinically relevant biomarker. Notably, the nature of the prognostic signal observed, detected in serum from individuals suffering from early-stage disease and supported by the TCGA-BRCA tissue survival analysis, suggests that hsa-miR-614 is more suited to be considered as a promising prognostic biomarker than a diagnostic one. Clinically speaking, its applicability in this regard would involve its usefulness in the post-diagnosis process, that is, stratifying those individuals who are diagnosed with early-stage breast cancer but whose biological characteristics suggest increased risk of adverse outcomes. The growing body of evidence on the potential value of circulating miRNA signatures following diagnosis in assessing risk beyond clinicopathological variables (tumor size, nodal involvement, and subtype, among others) lends additional support to their use in this regard.[25]

Moreover, the serum-based detection of the current signal enhances its utility, given the inherent advantages of this method compared with tissue-based assessment via biopsy. In this context, hsa-miR-614 is a candidate for dynamic monitoring of treatment response and disease trajectory, as well as for early detection of disease recurrence or progression. The utility of circulating miRNAs as serial liquid biopsy analytes in breast cancer has been substantiated by recent clinical studies demonstrating their capacity to reflect real-time tumor behavior

and predict therapy resistance before radiographic progression.[26] Importantly, the integration of hsa-miR-614 into a multi-analyte liquid biopsy panel, combining established circulating markers such as CA 15-3, circulating tumor DNA (ctDNA), and other co-dysregulated miRNA candidates identified in the present study, may further enhance overall prognostic resolution and diagnostic sensitivity relative to any single biomarker alone.[27] It must be emphasized, however, that all translational projections presented here are hypothesis-generating. Prospective validation in independent, adequately powered cohorts with complete clinicopathological annotation, matched serum and tissue samples, and standardized pre-analytical protocols will be essential before hsa-miR-614 can be considered for integration into clinical practice.

Several miRNAs consistently upregulated in this study, including hsa-miR-4532, hsa-miR-4734, hsa-miR-6131, and hsa-miR-5787, remain largely undercharacterized in breast cancer. For instance, although circulating miR-4532 has been proposed as a potential biomarker for cholangiocarcinoma[28], its specific role in early breast tumorigenesis requires further investigation. The identification of these understudied miRNAs underscores the value of unbiased reanalysis of large public datasets, which expands the candidate biomarker landscape beyond well-known oncomiRs and may reveal early or subtype-specific regulatory changes. Nevertheless, we caution that our conclusions are exploratory and should be considered hypothesis generating, not confirmatory. The list of microRNAs obtained can be considered potential targets whose functions must be thoroughly validated through in vitro cell culture experiments and animal studies.[29] More importantly, this study shows that just because certain microRNAs exhibit high expression does not mean they are independently predictive of clinical outcomes.[30]

Strengths and limitations

One key advantage of this study is the combination of high-throughput, highly sensitive serum microarray data with extensive bioinformatics analyses, thereby providing sufficient statistical power to detect early changes in the regulation of circulating miRNAs. The TCGA-BRCA tissue survival analysis provides supportive evidence across biological compartments for the potential clinical relevance of these biomarkers, especially hsa-miR-614.

To begin with, the differential expression analysis in this study used microarrays, which, although quite standardized, have a narrower dynamic range for transcript detection than the more contemporary RNA-seq technology. Secondly, most biological functions and miRNA-target gene relationships were predicted solely using bioinformatics tools. Even though separate validation experiments will be conducted using techniques such as qRT-PCR, Luciferase reporter assay, and other gain/loss-of-function studies, the results obtained should be treated as preliminary for further hypothesis testing. Detailed technical batch metadata were not uniformly available in the public dataset; therefore, residual technical variation cannot be completely excluded. Another limitation is that the primary differential expression analysis was performed as a case–control comparison without adjustment for potential demographic and clinical confounders. Variables such as age, menopausal status, body mass index, and other patient-level characteristics were not uniformly available in the GEO-deposited metadata and therefore could not be incorporated into the statistical model. Consequently, residual confounding cannot be excluded. Future prospective studies should employ matched case–control designs and multivariable adjustment strategies to further evaluate the robustness of the identified miRNA signatures. Finally, the supportive survival analysis was performed using the Kaplan–Meier Plotter (miRPower) platform, which provides univariate survival estimates but does not readily support multivariable adjustment for potential clinical confounders. Consequently, the observed prognostic associations should be interpreted as exploratory rather than independent prognostic effects. Future studies should validate these findings using multivariable Cox proportional hazards models incorporating established clinicopathological variables, including age, tumor stage, molecular subtype, and receptor status. Although cross-compartment validation has been used in previous biomarker-discovery studies when matched serum cohorts with survival data are unavailable, such an approach should be regarded as supportive rather than definitive. Consequently, the TCGA-BRCA analysis presented here provides biological plausibility for the prognostic relevance of candidate miRNAs, but does not constitute direct clinical validation of their circulating serum expression patterns.

CONCLUSION

This integrative bioinformatic analysis systematically delineates the dysregulation of circulating miRNAs in early-stage breast cancer, a pattern consistent with potential systemic regulatory reprogramming that may be detectable in the peripheral blood. The pronounced upregulation of specific miRNAs, most notably hsa-miR-614, alongside undercharacterized candidates such as hsa-miR-4532 and hsa-miR-4734, may be consistent with an early disruption of global transcriptomic homeostasis, particularly involving chromatin remodeling and RNA metabolism pathways. Importantly, the association of elevated hsa-miR-614 expression with adverse overall survival in an independent TCGA-BRCA tissue cohort provides supportive evidence for its potential biological and prognostic relevance, although direct validation in matched serum cohorts remains necessary. Ultimately, these findings provide preliminary computational evidence supporting the potential of circulating miRNA signatures as minimally invasive, candidate liquid biopsy-based biomarkers for the early detection

and risk stratification of breast cancer, establishing a computational foundation for future experimental, mechanistic, and prospective clinical validation studies.

ETHICAL CONSIDERATIONS

The present study is based exclusively on the secondary analysis of publicly available, deidentified datasets obtained from the Gene Expression Omnibus (GEO; GSE73002) and The Cancer Genome Atlas Breast Cancer cohort (TCGA-BRCA). As no human participants were directly recruited and no identifiable personal information was accessed, institutional review board (IRB) approval and informed consent were not required for the current analysis. The companion prospective validation study conducted by the same research group was reviewed and approved by the Ethics Committee of the Department of Biotechnology, College of Science, University of Baghdad, Baghdad, Iraq, and was performed in accordance with the ethical principles of the Declaration of Helsinki.

DATA AVAILABILITY

The microRNA expression dataset analyzed in this study is publicly available in the Gene Expression Omnibus (GEO) repository under accession number GSE73002. Survival data were obtained from The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) cohort, accessible through the GDC Data Portal and the Kaplan–Meier Plotter platform. No new primary datasets were generated in this study. Processed analytical files and derived results generated during the study are available from the corresponding author upon reasonable request.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests.

FUNDING

None.

ACKNOWLEDGMENTS

The authors would like to thank the investigators who produced and uploaded the GSE73002 dataset into the Gene Expression Omnibus (GEO) repository, enabling a secondary computational analysis. We are grateful for the institutional support of the College of Science at the University of Baghdad.

AI DISCLOSURE

Artificial intelligence tools were used to improve sentence structure and clarity.

AUTHOR CONTRIBUTION

ANA: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. RAA: Conceptualization, Data curation, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing.

REFERENCES

1. Sun KX, Liang X, Zhu Q, Wu HL, Zhang GY, Yao YF, et al. Global patterns and trends in cancer-related premature death and their impact on life expectancy across 185 countries: A population-based analysis. *Mil Med Res*. 2025;12(1):56. doi: 10.1186/s40779-025-00645-9
2. Sharma D, Al-Khalidi R, Edgar S, An Q, Wang Y, Young C, et al. Co-delivery of indoleamine 2,3-dioxygenase prevents loss of expression of an antigenic transgene in dystrophic mouse muscles. *Gene Ther*. 2017;24(2):113-9. doi: 10.1038/gt.2016.82
3. Lánczky A, Nagy Á, Bottai G, Munkácsy G, Szabó A, Santarpia L, et al. miRpower: A web-tool to validate survival-associated miRNAs utilizing expression data from 2178 breast cancer patients. *Breast Cancer Res Treat*. 2016;160(3):439-46. doi: 10.1007/s10549-016-4013-7
4. Mou T, Deng W, Gu F, Pawitan Y, Vu TN. Reproducibility of methods to detect differentially expressed genes from single-cell RNA sequencing. *Front Genet*. 2019;10:1331. doi: 10.3389/fgene.2019.01331
5. Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol*. 2014;11(3):145-56. doi: 10.1038/nrclinonc.2014.5
6. Chen S, Wang L, Yao B, Liu Q, Guo C. miR-1307-3p promotes tumor growth and metastasis of hepatocellular carcinoma by repressing DAB2 interacting protein. *Biomed Pharmacother*. 2019;117:109055. doi: 10.1016/j.biopha.2019.109055
7. Shang C, Chen Q, Zu F, Ren W. Integrated analysis identified prognostic microRNAs in breast cancer. *BMC Cancer*. 2022;22(1):1170. doi: 10.1186/s12885-022-10242-x
8. He B, Zhao Z, Cai Q, Zhang Y, Zhang P, Shi S, et al. miRNA-based biomarkers, therapies, and resistance in cancer. *Int J Biol Sci*. 2020;16(14):2628-47. doi: 10.7150/ijbs.47203
9. Muniategui A, Nogales-Cadenas R, Vázquez M, Aranguren XL, Agirre X, Luttun A, et al. Quantification of miRNA-mRNA interactions. *PLoS One*. 2012;7(2):e30766. doi: 10.1371/journal.pone.0030766
10. Stempor PA, Cauchi M, Wilson P. MMpred: Functional miRNA-mRNA interaction analyses by miRNA expression prediction. *BMC Genomics*. 2012;13(1):620. doi: 10.1186/1471-2164-13-620
11. Volinia S, Croce CM. Prognostic microRNA/mRNA signature from the integrated analysis of patients with invasive breast cancer. *Proc Natl Acad Sci U S A*. 2013;110(18):7413-7. doi: 10.1073/pnas.1304977110
12. Ma L, Guo H, Zhao Y, Liu Z, Wang C, Bu J, et al. Liquid biopsy in cancer: Current status, challenges and future prospects. *Signal Transduct Target Ther*. 2024;9(1):336. doi: 10.1038/s41392-024-02021-w
13. Shimomura A, Shiino S, Kawauchi J, Takizawa S, Sakamoto H, Matsuzaki J, et al. Novel combination of serum microRNA for detecting breast cancer in the early stage. *Cancer Sci*. 2016;107(3):326-34. doi: 10.1111/cas.12880
14. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A*. 2008;105(30):10513-8. doi: 10.1073/pnas.0804549105

15. Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF. Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci U S A*. 2011;108(12):5003-8. doi: 10.1073/pnas.1019055108
16. Condrat CE, Thompson DC, Barbu MG, Bugnar OL, Boboc A, Cretoiu D, et al. miRNAs as biomarkers in disease: Latest findings regarding their role in diagnosis and prognosis. *Cells*. 2020;9(2):276. doi: 10.3390/cells9020276
17. Huang WL, Luo CW, Lin HS, Hung CM, Chen FM, Moi SH, et al. SUV39H1 expression as a guideline for omitting radiotherapy in lymph node-positive triple-negative breast cancer patients. *Cancer Genomics Proteomics*. 2023;20(6):582-91. doi: 10.21873/cgp.20407
18. Tinsley E, Bredin P, Toomey S, Hennessy BT, Furney SJ. KMT2C and KMT2D aberrations in breast cancer. *Trends Cancer*. 2024;10(6):519-30. doi: 10.1016/j.trecan.2024.02.003
19. Blawski R, Vokshi BH, Guo X, Kittane S, Sallaku M, Chen W, et al. Methylation of the chromatin modifier KMT2D by SMYD2 contributes to therapeutic response in hormone-dependent breast cancer. *Cell Rep*. 2024;43(5):114174. doi: 10.1016/j.celrep.2024.114174
20. Michalak EM, Visvader JE. Dysregulation of histone methyltransferases in breast cancer: Opportunities for new targeted therapies?. *Mol Oncol*. 2016;10(10):1497-515. doi: 10.1016/j.molonc.2016.09.003
21. Welte T, Goulois A, Stadler MB, Hess D, Soneson C, Neagu A, et al. Convergence of multiple RNA-silencing pathways on GW182/TNRC6. *Mol Cell*. 2023;83(14):2478-2492.e8. doi: 10.1016/j.molcel.2023.06.001
22. Dang TT, McIntosh AT, Morales JC, Pearson GW. miR614 expression enhances breast cancer cell motility. *Int J Mol Sci*. 2020;22(1):112. doi: 10.3390/ijms22010112
23. Zhang J, Gao D, Zhang H. Upregulation of miR-614 promotes proliferation and inhibits apoptosis in ovarian cancer by suppressing PPP2R2A expression. *Mol Med Rep*. 2018;17(5):6285-92. doi: 10.3892/mmr.2018.8714
24. Tian B, Hou M, Zhou K, Qiu X, Du Y, Gu Y, et al. A novel TCGA-validated, miRNA-based signature for prediction of breast cancer prognosis and survival. *Front Cell Dev Biol*. 2021;9:717462. doi: 10.3389/fcell.2021.717462
25. Wang G, Chang Y, Wu X, Li X, Li L, Zhang M. Prognostic nomogram for overall survival in upper aerodigestive tract extranodal natural killer/T-cell lymphoma, nasal type, stages IE and IIE: A SEER-based study. *Oncol Lett*. 2019;18(4):3767-76. doi: 10.3892/ol.2019.10719
26. Kubeczko M, Tudrej P, Tyszkiewicz T. Liquid biopsy utilizing miRNA in patients with advanced breast cancer treated with cyclin-dependent kinase 4/6 inhibitors. *Oncol Lett*. 2024;27(4):181. doi: 10.3892/ol.2024.14314
27. Simancas-Racines D, Román-Galeano NM, Vásquez JP. Liquid biopsy and multi-omic biomarkers in breast cancer: Innovations in early detection, therapy guidance, and disease monitoring. *Biomedicines*. 2025;13(12):3073. doi: 10.3390/biomedicines13123073
28. Supradit K, Prasopdee S, Phanaksri T, Tangphatsornruang S, Pholhelm M, Yusuk S, et al. Differential circulating miRNA profiles identified miR-423-5p, miR-93-5p, and miR-4532 as potential biomarkers for cholangiocarcinoma diagnosis. *PeerJ*. 2024;12:e18367. doi: 10.7717/peerj.18367
29. Sharifi Z, Talkhabi M, Taleahmad S. Identification of potential microRNA diagnostic panels and uncovering regulatory mechanisms in breast cancer pathogenesis. *Sci Rep*. 2022;12(1):20135. doi: 10.1038/s41598-022-24347-7
30. Wang L, Wang G, Song J, Yao D, Wang Y, Chen T. A comprehensive analysis of the prognostic characteristics of microRNAs in breast cancer. *Front Genet*. 2024;15:1293824. doi: 10.3389/fgene.2024.1293824

How to Cite This Article

Al-Ebadi AN, Al-Khalidi RAA. Circulating MicroRNA Signatures as Prognostic Biomarkers in Breast Cancer: A Comprehensive Analysis of Expression Profiles, Functional Targets, and Survival Correlations. Arch Breast Cancer. 2026; 13(4):521-531. doi: 10.32768/abc.7394826150-371
Available from: https://www.archbreastcancer.com/index.php/abc/article/view/1316