Original Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2024 (4 Jun);38.63. https://doi.org/10.47176/mjiri.38.63

Integrated Analysis of Multi-Omic Data Reveals Regulatory Mechanisms and Network Characteristics in Breast Cancer

Zahra Hosseinpour[1], Mostafa Rezaei Tavirani[2]*, Mohammad Esmaeil Akbari[3]

Received: 12 Aug 2023 Published: 4 Jun 2024

Abstract

Background: Breast cancer is a complex and heterogeneous disease, and understanding its regulatory mechanisms and network characteristics is essential for identifying therapeutic targets and developing effective treatment strategies. This study aimed to unravel the intricate network of interactions involving differentially expressed genes, microribonucleic acid (miRNAs), and proteins in breast cancer through an integrative analysis of multi-omic data from Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA) dataset.

Methods: The TCGA-BRCA dataset was used for data acquisition, which included RNA sequencing data for gene expression, miRNA sequencing data for miRNA expression, and protein expression quantification data. Various R packages, such as TCGAbiolinks, limma, and RPPA, were employed for data preprocessing and integration. Differential expression analysis, network construction, miRNA regulation exploration, pathway enrichment analysis, and independent dataset validation were performed.

Results: Eight consistently upregulated hub genes—including ACTB, HSP90AA1, FN1, HSPA8, CDC42, CDH1, UBC, and EP300—were identified in breast cancer, indicating their potential significance in driving the disease. Pathway enrichment analysis revealed highly enriched pathways in breast cancer, including proteoglycans in cancer, PI3K-Akt, and mitogen-activated protein kinase signaling.

Conclusion: This integrated multi-omic data analysis provides valuable insights into the regulatory mechanisms, network characteristics, and functional roles of genes, miRNAs, and proteins in breast cancer. The findings contribute to our understanding of the molecular landscape of breast cancer, facilitate the identification of potential therapeutic targets, and inform strategies for effective treatment.

Keywords: Breast Cancer, Target Therapy, Multiomic, Hub-Genes

Conflicts of Interest: *None declared*
Funding: *The Cancer Research Center of Shahid Beheshti University of Medical Sciences supported the present study.*

Cite this article as Hosseinpour Z, Rezaei Tavirani M, Akbari ME. Integrated Analysis of Multi-Omic Data Reveals Regulatory Mechanisms and Network Characteristics in Breast Cancer. *Med J Islam Repub Iran.* 2024 (4 Jun);38:63. https://doi.org/10.47176/mjiri.38.63

Introduction

Breast cancer, recognized globally as a leading cause of cancer-related deaths among women, is a highly intricate and heterogeneous disease (1). Understanding its underlying regulatory mechanisms, gene expressions, and biological pathways is a critical focal point in oncological research, essential for developing effective therapeutic strategies and improving patient survival rates (2). The recognition of this need forms the foundation upon which this research study on breast cancer is built. Diverse molecular alterations characterize the pathology of breast cancer and contribute to the broad range of clinical outcomes observed in patients (3). These molecular alterations manifested

Corresponding author: *Dr Mostafa Rezaei Tavirani, tavirani@sbmu.ac.ir*

1. Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2. Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Surgical Oncology, Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

↑What is "already known" in this topic:
Breast cancer is a complex and heterogeneous disease, and previous research has emphasized the importance of understanding its regulatory mechanisms, network characteristics, and molecular alterations to develop effective treatment strategies.

→What this article adds:
This article adds insights into the regulatory mechanisms, network characteristics, and functional roles of genes, miRNAs, and proteins in breast cancer through an integrated analysis of multi-omic data, facilitating the identification of potential therapeutic targets and informing effective treatment strategies.

[DOI: 10.47176/mjiri.38.63]

through differential gene and protein expressions and microribonucleic acid (miRNA) behaviors, are the dynamic players driving the disease's progression. However, a detailed exploration and understanding of the intricate network interplay between these molecular elements is needed to truly make sense of these intricate interactions and workings (4). This brings us to the core objectives of this study, which aims to unravel the complex network of interactions involving differentially expressed genes, microRNAs (miRNAs), and proteins in breast cancer.

Through an integrated analysis of multi-omic data, the study sheds light on the various regulatory mechanisms, network characteristics, and functional roles of genes, miRNAs, and proteins in breast cancer (5). In seeking to map this complex landscape, we will harness the considerable potential of the expansive multi-omic data offered by The Cancer Genome Atlas Breast Cancer (TCGA-BRCA) dataset. By leveraging this large-scale dataset, we aim to perform an in-depth analysis highlighting the various molecular alterations present in breast cancer. The comprehensive scope of this dataset to provide vital insights into the possible therapeutic targets and inform the development of effective treatment strategies. The novelty of this approach lies in our focus on an integrative, multiomic analysis, moving beyond static, singular datasets to generate a detailed examination of the complex interwikiing of genes, miRNAs, and proteins in cancerous cells. We will utilize various R packages for data preprocessing and integration to facilitate this analytical process, such as TCGAbiolinks, limma, and reverse-phase protein array (RPPA). After acquiring the necessary dataset, a differential expression analysis will identify dysregulated genes, miRNAs, and proteins within the TCGA-BRCA dataset. Identifying these dysregulated elements will provide deeper insights into the molecular markers of breast cancer and could, ultimately, become potential therapeutic targets. Understanding the pathways these dysregulated genes, miRNAs, and protein markers navigate through will be the focus of the pathway enrichment analysis. This knowledge will not only enhance our understanding of the underlying molecular mechanisms driving breast cancer but also pave the way for the development of targeted therapeutic interventions. In addition, by validating our findings using independent datasets and conducting gene expression analysis, we aim to strengthen the reliability and significance of our results.

Methods

Data Acquisition

The study's selection of breast cancer patients was based on specific inclusion and exclusion criteria to ensure the relevance and quality of the data used in the analysis. Patients with various tumor stages were included to encompass the heterogeneity of breast cancer. Moreover, the focus was on invasive breast carcinoma cases with different histological subtypes to explore the regulatory mechanisms and network characteristics across diverse subtypes. The patient selection process also considered the availability of multiomic data, including gene expression data obtained through RNA sequencing, miRNA data obtained through miRNA-Seq, and protein expression data obtained through protein expression quantification. The data preprocessing steps involved the utilization of several R packages. The TCGAbiolinks package filters the mRNA data based on the project, data category, type, and experimental strategy. MiRNA data were extracted using miRNA-Seq as the experimental strategy, while protein data utilized protein expression quantification. The mRNA and miRNA data were further processed using the limma package, which included normalization, log transformation, and batch correction. The protein data were normalized and log-transformed using the RPPA package. Subsequently, sample barcodes that had data available across all 3 data types (mRNA, miRNA, and protein) were obtained for further integrated analysis. (6). These steps ensured our study's comprehensive investigation of interactions and regulatory relationships between different molecular layers.

Data Analysis

Differential expression analysis was performed using the R package limma to identify abnormally expressed mRNA and miRNAs in the TCGA-BRCA dataset. In addition, the RPPA R package was utilized to identify abnormally expressed proteins in the same dataset. This analysis specifically focused on comparing breast cancer samples in late stages (stages I and II) with breast cancer samples in advanced stages (stages III and IV). Identifying differentially expressed genes involved applying initial selection criteria, including a log2 fold change of ≥ 1 and a P value < 0.05. Subsequently, the top-upregulated proteins and downregulated miRNAs were identified. To investigate interactions between the differentially expressed genes, a protein-protein interaction network was constructed using the STRINGdb and igraph R packages (7). For constructing the protein-protein interaction network, a new instance of the STRING database was instantiated using Version 11.0. The organism specified was Homo sapiens, and a minimum score threshold of 200 was applied. The "get interactions" function was used to retrieve known and predicted interactions between the upregulated genes, resulting in an edge list. Based on this edge list, an igraph object representing the network was created (8).

Exploring Graph Characteristics

Various methods were utilized to analyze the graph and explore its characteristics, including its nodes representing entities and edges representing their relationships. The objective was to gain insights into the connectivity patterns within the graph. First, the degree distribution of the graph was calculated using the degree-distribution function, providing information about the frequency of nodes with different degrees and illuminating the distribution of connectivity. Next, a power-law model was fitted to the degree distribution using the power-law-fit function. This model describes the relationship where the frequency of nodes with a given degree follows a power-law distribution. Fitting this model allowed the estimation of the power-law exponent, which characterizes the distribution's shape (9). To assess the goodness-of-fit of the power-law model, the P value was obtained from the Kolmogorov-Smirnov test us-

[DOI: 10.47176/mjiri.38.63]

ing the KS.p attribute of the fitted model. The P value represents the probability of observing a test statistic as extreme as the one obtained, assuming that the data follow a power-law distribution (10). These analyses provided valuable insights into the characteristics of the graph and the distribution of connectivity among its nodes. The top 1% of high-degree hub nodes were found using quantile and saved as "hubs." Global clustering coefficient and betweenness centrality were calculated, and the result showed that the hubs were highly connected with high betweenness, implying important roles (11). In summary, network analysis provided insights into potential interactions among differentially expressed breast cancer genes and identified hub genes that may play critical functional roles in the disease.

Exploring miRNA Regulation of Hub Genes

The multiMiR R package investigated regulatory relationships between hub genes and downregulated miRNAs. This package integrates miRNA-target interactions from multiple databases and predicts new interactions based on sequence complementarity. Using the getMiRNATargets function, both predicted and validated interactions for the downregulated miRNAs were retrieved. Moreover, the multiMiR package offers valuable insights into disease-drug associations. Utilizing its extensive database, we can explore connections between hub genes associated with breast cancer and specific drugs targeting those genes or related pathways. This analysis enhances our understanding of the regulatory landscape and potential therapeutic interventions for breast cancer. This integrated approach, combining miRNA-target interactions and disease-drug associations, provides a comprehensive view of the regulatory mechanisms underlying breast cancer. These findings contribute to unraveling the complex network of interactions involving hub genes, dysregulated miRNAs, and potential therapeutic strategies in breast cancer research (12).

Pathway Enrichment and Visualization

Pathway enrichment analysis was performed using the KEGGREST R package to understand the functional roles of the identified hub genes. The KEGG database contains information on molecular pathways and networks in various organisms. The keggGet function obtained KEGG pathway annotations for all human genes and the keggEnricher function analyzed pathway enrichment for the hub genes.

A hypergeometric test was used to determine if the hub genes were overrepresented in particular KEGG pathways.

Adjusted P values were calculated using the Benjamini-Hochberg method for multiple testing correction (13). The Pathview R package was used to visualize the enriched pathways by creating pathway diagrams that highlighted the differentially expressed hub genes within each pathway. This integrated approach, combining the KEGGREST and Pathview packages with network analysis, provided a comprehensive understanding of the functional roles of the hub genes within specific pathways in primary breast cancer (14).

Validation

To validate the upregulated hub genes identified from TCGA-BRCA, the independent gene expression dataset GSE139038 was utilized. This dataset contained transcriptomic data from early-stage and late-stage breast cancer samples (15). Differential expression analysis between late- and early-stage samples in GSE139038 was performed using GEO2R to identify significantly upregulated genes (log2FC > 1 and adjusted $P < 0.05$) (16).

Expression Analysis by qPCR

In this study, gene expression analysis was performed using a specific set of locked nucleic acid primers obtained from Exiqon. The 2-ΔΔCT method was utilized to normalize the relative expression levels of the target gene. Normalization was performed using the endogenous reference gene BET. Real-time polymerase chain reaction (PCR) reactions were conducted on the Chromo4 system, a Bio-Rad instrument. The real-time PCR kit provided (Takara Bio) was used for the PCR reactions. To ensure accuracy and reproducibility, all samples were run in technical triplicates, and each sample underwent 40 cycles of PCR amplification. The thermal cycling conditions for the quantitative PCR (qPCR) reactions were as follows: an initial predenaturation step at 95°C for 10 minutes, followed by denaturation at 95°C for 20 seconds, annealing at 60°C for 30 seconds, and extension at 72°C for 30 seconds. By following these experimental procedures, we aimed to accurately quantify the expression levels of the target gene using qPCR. This approach allowed us to assess the relative expression changes and investigate the gene's involvement in the studied context.

Results

Data Acquisition

The analysis of the TCGA-BRCA dataset resulted in the following outcomes. A total of 878 samples had available data for mRNA, miRNA, and protein expression (Table 1). In the protein expression analysis, 3 proteins were significantly upregulated ($P < 0.05$). 39 miRNAs showed significant downregulation for miRNA expression with a fold change of <–1 and a P of < 0.05. In the mRNA analysis, 2905 genes were identified as significantly upregulated (fold change >1; $P < 0.05$). These findings reveal specific molecular alterations in breast cancer, including dysregulated proteins, downregulated miRNAs, and upregulated mRNA genes. The integrated analysis of multiomic data from TCGA-BRCA provides valuable insights into the molecular landscape of breast cancer and enhances our under-

Table 1. Demographic and Clinical Information of Patients

n	835
Pathologic stage	(%)
Normal	9 (1.1)
Stage_I	127 (15.2)
Stage_II	479 (57.4)
Stage_III	202 (24.2)
Stage_IV	18 (2.2)
Age, years, mean (SD)	58.34 (13.28)

[DOI: 10.47176/mjiri.38.63]

standing of the underlying mechanisms involved in the disease.

Network Analysis and Degree Distribution

Various analyses were conducted to understand the connectivity patterns within the graph. Initially, the degree distribution of the graph was calculated to determine the frequency of nodes with different degrees, providing valuable information about connectivity distribution. Subsequently, a power-law model was fitted to the degree distribution, yielding a power-law exponent of 6.497. This exponent suggests the presence of highly connected nodes potentially following a power-law pattern. The goodness-of-fit of the power-law model was assessed using a Kolmogorov-Smirnov test, resulting in a P value of 0.943, indicating a good fit. The Pearson correlation coefficient between the in-degree and out-degree of the graph was 0.0467, suggesting a very weak positive correlation. These analyses offered insights into the graph's characteristics and node connectivity distribution. The clustering coefficient was calculated to measure the tendency of nodes to cluster together, yielding a value of 0.222, indicating a moderate level of clustering and local interconnectedness. A log-log plot (Figure 1) was generated to visualize the degree distribution, illustrating the frequency of nodes with different degrees and aiding in examining distribution patterns.

The plot provides insights into the degree distribution, highlighting the presence of highly connected nodes and potential patterns within the distribution. In summary, the analyses revealed a potential power-law distribution with a power-law exponent of 6.497, supported by the goodness-of-fit test. A very weak positive correlation was observed between the in-degree and out-degree of the nodes. In addition, a moderate level of clustering was found within the network. The log-log plot enhances the understanding of the degree distribution and provides valuable information about the frequency of nodes with different degrees.

Multimer Results

This study investigated potential regulatory relationships between the hub genes identified in the network analysis and downregulated miRNAs. A total of 20 hub genes—namely, AKT1, MAPK3, ALB, ACTB, HSP90AA1, MAPK1, HSPA4, FN1, HSPA8, ERBB2, ESR1, CDC42, RHOA, CDH1, HSP90AB1, EEF2, SOD1, UBC, EP300, and DECR1—were analyzed. The interactions between these hub genes and downregulated miRNAs were examined using the mulitmiR R package, which integrates miRNA-target interactions from multiple databases. The findings revealed specific miRNAs that were either predicted or experimentally validated to target the hub genes of interest. The analysis presented in Table 2 summarizes the downregulated miRNAs and their corresponding target genes (hub genes) identified in this study. It was observed that several miRNAs exhibited downregulation and were predicted or experimentally validated to target specific hub genes. Among them, hsa-miR-107 was found to target genes—including CDC42, ACTB, HSPA4, HSP90AA1, HSP90AB1, MAPK3, ALB, ESR1, HSPA8, and DECR1. These findings suggest potential regulatory interactions between the identified hub genes and the downregulated miRNAs, indicating their potential involvement in the observed gene expression changes.

Pathway Enrichment and Visualization

To identify proteins associated with breast cancer progression, omic data analysis compared late-stage and early-stage samples. Three proteins of interest, WIPI1, IGFBP2, and DUSP6, were identified in this analysis. WIPI1 is involved in autophagy processes (17), IGFBP2 plays a role

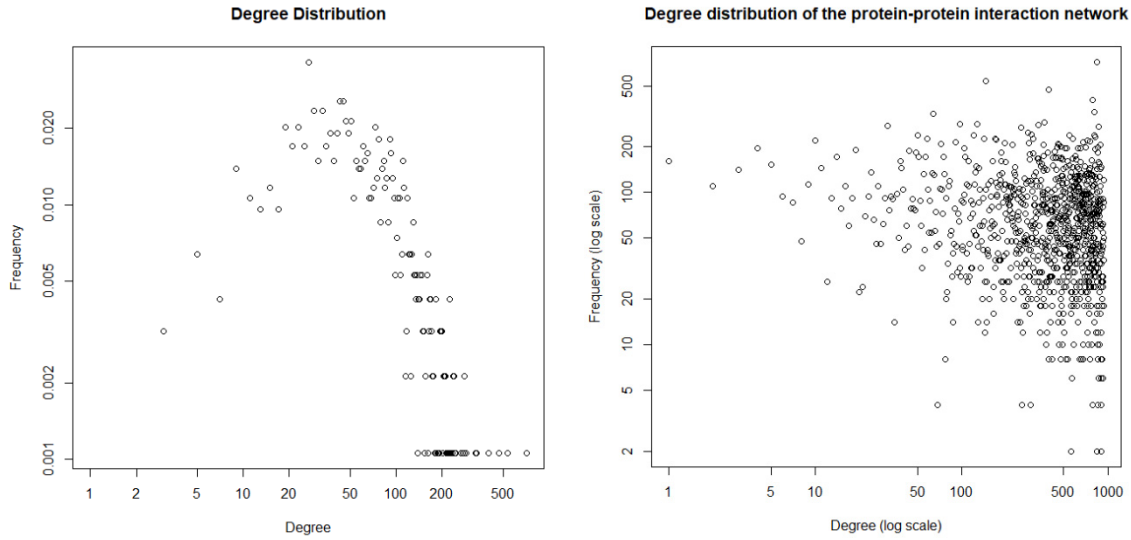


Figure 1. Log-Log Plot of Degree Distribution.
The plot provides insights into the degree distribution, highlighting the presence of highly connected nodes and potential patterns within the distribution. In summary, the analyses revealed a potential power-law distribution with a power-law exponent of 6.497, supported by the goodness-of-fit test. A very weak positive correlation was observed between the in-degree and out-degree of the nodes. In addition, a moderate level of clustering was found within the network. The log-log plot enhances the understanding of the degree distribution and provides valuable information about the frequency of nodes with different degrees.

[DOI: 10.47176/mjiri.38.63]

Table 2. Downregulated miRNAs and Their Corresponding Target Genes (Hub Genes)

Downregulated miRNA	Target Gene (Hub Gene)
hsa-miR-375, hsa-miR-940, hsa-miR-484, hsa-miR-326	ERBB2
hsa-miR-107, hsa-miR-577, hsa-miR-375, hsa-miR-484	CDC42
hsa-miR-375, hsa-miR-592, hsa-miR-577, hsa-miR-147b, hsa-miR-484, hsa-miR-940, hsa-miR-429, hsa-miR-1270	RHOA
hsa-miR-451a, hsa-miR-326, hsa-miR-1976	AKT1
hsa-miR-484, hsa-miR-326, hsa-miR-107	ACTB
hsa-miR-429, hsa-miR-484, hsa-miR-940	EP300
hsa-miR-107, hsa-miR-375	HSPA4
hsa-miR-107, hsa-miR-375, hsa-miR-484, hsa-miR-935	HSP90AA1
hsa-miR-107, hsa-miR-375 , hsa-miR-484	HSP90AB1
hsa-miR-1976, hsa-miR-484, hsa-miR-1269a, hsa-miR-326, hsa-miR-940, hsa-mi R-451a	MAPK1
hsa-miR-107, hsa-miR-375	MAPK3
hsa-miR-1269a, hsa-miR-326	UBC
hsa-miR-107	ALB
hsa-miR-429, hsa-miR-577, hsa-miR-1270	FN1
hsa-miR-934, hsa-miR-107	ESR1
hsa-miR-375, hsa-miR-107	HSPA8
hsa-miR-484	CDH1
hsa-miR-107, hsa-miR-429	DECR1

in MAPK signaling regulation (18), and DUSP6 is associated with the MAPK pathway (19). Enrichment analysis revealed several significantly enriched pathways in breast cancer. The Proteoglycans in the cancer pathway showed the highest enrichment (P = 4.9E-10), followed by the PI3K-Akt pathway (P = 1.48E-05) and the mitogen-activated protein kinase (MAPK) signaling pathway (P = 4.65E-06). The selection of WIPI1, IGFBP2, and DUSP6 as proteins of interest aligns with their involvement in the PI3K-Akt and MAPK pathways and their potential significance in breast cancer progression. These proteins likely mediate the effects of proteoglycans on cancer progression, given their associations with key pathways. Visualizing these pathways using Pathview provides a comprehensive understanding of the molecular mechanisms and potential crosstalk between these proteins and other molecules within the pathways (see Supplementary Figure 1 for proteoglycans in the cancer pathway and Supplementary Figure 2 for PI3K-Akt signaling pathway). The complete list of enriched pathways in breast cancer is provided in Table 3.

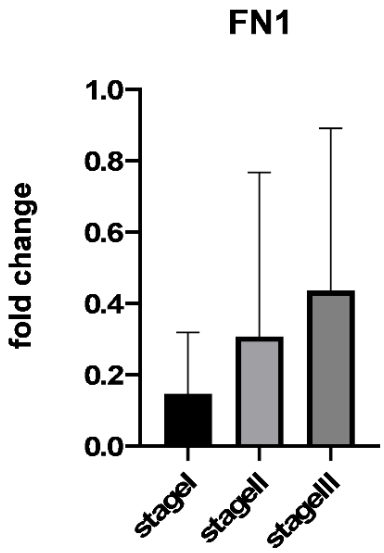


Figure 2. The comparative expression analysis of FN1 in different stages of BC, *P = 0.6743 (nonsignificant).

Validation

To validate the upregulated hub genes from TCGA, the independent gene expression dataset GSE139038 was used. This dataset included transcriptomic data from early and late-stage breast cancer samples. Differential expression analysis using GEO2R compared late- and early-stage samples in GSE139038, with criteria set as a log2 fold change (log2FC) >1 and $P < 0.05$. The upregulated hub genes from TCGA were cross-referenced with the upregulated genes from the late- versus early-stage comparison in GSE139038. This validation process identified 8 consistently upregulated hub genes—ACTB, HSP90AA1, FN1, HSPA8, CDC42, CDH1, UBC, EP300—in both datasets (Table 4). These hub genes show reproducibility across independent cohorts of early- and late-stage breast cancer samples, indicating their potential as biomarkers of disease progression. The validation analysis using GEO2R confirms the consistent upregulation of the hub genes identified from TCGA during breast cancer progression in the external dataset (GSE139038). These validated hub genes can be further investigated as potential prognostic biomarkers or therapeutic targets in breast cancer research. Table 4 provides a summary of the validated hub genes.

FN1 Expression

In this study, the expression of FN1 was investigated. Pathway enrichment analysis revealed the significance of the proteoglycan pathway in cancer. Specifically, the mRNA expression of FN1, which serves as a hub gene in the proteoglycans pathway, was found to be significantly higher in tumor tissues from women with late-stage breast cancer compared to early-stage breast samples. To further

[DOI: 10.47176/mjiri.38.63]

Table 3. Enriched Pathways

ID	Description	Gene Ratio	*P* Value
hsa05205	Proteoglycans in cancer	9/21	<0.001
hsa05417	Lipid and atherosclerosis	9/21	<0.001
hsa04520	Adherens junction	8/21	<0.001
hsa04510	Focal adhesion	8/21	<0.001
hsa05131	Shigellosis	8/21	<0.001
hsa05132	Salmonella infection	8/21	<0.001
hsa05215	Prostate cancer	7/21	<0.001
hsa04915	Estrogen signaling pathway	7/21	<0.001
hsa05135	Yersinia infection	7/21	<0.001
hsa04015	Rap1 signaling pathway	7/21	<0.001
hsa04810	Regulation of actin cytoskeleton	7/21	<0.001
hsa04010	MAPK signaling pathway	7/21	<0.001
hsa04151	PI3K-Akt signaling pathway	7/21	<0.001

Table 4. Validated Upregulated Hub Genes in Breast Cancer Progression

Gene	*P* Value	logFC	Gene	*P* Value	logFC
AKT1	0.037	0.983348	ESR1	0.031	0.359
MAPK3	0.967	0.013248	CDC42	0.002	1.131
ALB	NA	NA	RHOA	0.004	0.752
ACTB	0.091	1.036123	CDH1	0.001	1.182
HSP90AA1	0.036	1.333578	HSP90AB1	0.011	0.782
MAPK1	0.039	0.983738	EEF2	0.012	0.775
HSPA4	0.070	0.938271	SOD1	0.094	-0.027
FN1	0.014	1.724786	UBC	0.000	1.194
HSPA8	0.010	1.15904	EP300	0.000	1.144
ERBB2	0.169	0.711806	DECR1	0.021	0.600

investigate this observation, blood samples were collected from women with stage I, II, and III breast cancer (n = 3), and their mRNA expression profiles were analyzed using RT-qPCR. The results demonstrated a stage-dependent increase in FN1 mRNA expression (Figure 2). This suggests that FN1 can potentially serve as a candidate blood mRNA biomarker for breast cancer. The statistical significance of the differences between the normal and cancer groups was assessed using an F test, and the graphs were generated using GraphPad Prism 9. The primer sequences used in this study are available and specified as follows:

20	GTGTGTTGGGAATGGTCGTG	f-FN1
19	CGCTTGTGGAATGTGTCGT	r-FN1

Discussion

The integrated analysis of multi-omic data in this study aimed to unravel the regulatory mechanisms and network characteristics associated with breast cancer. By analyzing the TCGA-BRCA dataset, several key findings were obtained, including dysregulated genes, miRNAs, and proteins and their network characteristics and regulatory mechanisms. These findings contribute to our understanding of the molecular landscape of breast cancer and have implications for identifying potential therapeutic targets and strategies. One of this study's most important findings was identifying dysregulated genes, miRNAs, and proteins in breast cancer. The differential expression analysis revealed a set of genes that were abnormally expressed in breast cancer samples compared to normal samples. Similarly, dysregulated miRNAs and proteins were also identified. These findings highlight the molecular alterations that are associated with breast cancer and provide insights into potential biomarkers for the disease. The construction of a protein-protein interaction network allowed for exploring interactions between differentially expressed genes. The network analysis provided valuable insights into the connectivity patterns and identified hub genes with important functional roles in breast cancer. It was observed that these hub genes were highly connected and had high betweenness centrality, indicating their significance in the disease. This finding suggests that these hub genes may serve as potential therapeutic targets for breast cancer treatment. Furthermore, the study investigated the potential regulatory relationships between hub genes and downregulated miRNAs. By utilizing the multiMiR R package, known and predicted miRNA-target interactions were explored. This analysis provided additional insights into the regulatory landscape of breast cancer and expanded our understanding of potential therapeutic interventions. Integrating miRNA-target interactions and disease-drug associations further enhanced the comprehensive view of the regulatory mechanisms underlying breast cancer. Several miRNAs were found to be downregulated and predicted or experimentally validated to target specific hub genes. Notably, hsa-miR-107 was identified as a regulator of genes—including CDC42, ACTB, HSPA4, HSP90AA1, HSP90AB1, MAPK3, ALB, ESR1, HSPA8, and DECR1. These findings suggest potential regulatory interactions between the downregulated miRNAs and the identified hub genes, indicating their involvement in the observed gene expression changes. Takahashi et al have highlighted miRNA-107 as a potential therapeutic target for disease treatment, particularly in human lung cancers. In the context of lung cancers, miRNA-107 and miR-185 have been implicated in regulating the cell cycle, emphasizing their potential contribution to the regulation of cell cycle processes in malignant tumors (20). Furthermore, epige-

[DOI: 10.47176/mjiri.38.63]

netic silencing of miR-107 has been shown to modulate Cyclin-Dependent Kinase 6 expression in pancreatic cancer. These findings suggest that targeting miRNA-107 holds promise for potential therapeutic interventions in various diseases (21). Pathway enrichment analysis was performed to gain insights into the functional roles of the hub genes within specific molecular pathways. The analysis revealed an association between the proteoglycans in the cancer pathway (22) and several other pathways, including the PI3K-Akt signaling pathway (23), MAPK signaling pathway (24), and cell cycle pathway (25). This finding suggests that the dysregulated hub genes may be involved in multiple pathways that play crucial roles in breast cancer development and progression. The interplay between these pathways may contribute to the complex nature of the disease. The identification of the link between the proteoglycans in cancer pathways and other pathways provides valuable data for comprehending the molecular mechanisms underlying breast cancer. It suggests potential crosstalk and interactions between these pathways, which may contribute to the dysregulation of key genes and signaling pathways involved in breast cancer. Targeting these interconnected pathways and hub genes may represent a promising approach to developing effective therapeutic strategies for breast cancer. Despite the valuable insights gained from this integrated analysis, certain limitations should be considered. First, the analysis was based on publicly available datasets, such as TCGA-BRCA. While these datasets are valuable, they have inherent limitations, such as sample heterogeneity and potential batch effects. Furthermore, the analysis focused on specific molecular alterations, and other important factors, such as epigenetic modifications and noncoding RNAs, were not considered. Future studies should aim to incorporate a broader range of omics data and utilize more comprehensive datasets to enhance our understanding of breast cancer further. Another limitation of this study is the small sample size used for the analysis of FN1 mRNA expression in blood samples from women with stage I, II, and III breast cancer. The study included only 3 patients, which may not represent the broader breast cancer population. As a result, the statistical power to detect significant differences between the normal and cancer groups was limited. The F test used to assess the statistical significance of the differences did not reach significance, possibly due to the small sample size. Therefore, caution should be exercised when interpreting the results, and further studies with larger sample sizes are necessary to validate the potential utility of FN1 as a blood mRNA biomarker for breast cancer. Increasing the sample size would enhance the statistical power and provide more robust and reliable results. In conclusion, the integrated analysis of multiomic data in breast cancer has provided valuable insights into the regulatory mechanisms, network characteristics, and functional roles of dysregulated genes, miRNAs, and proteins. Identifying hub genes and pathways in our study offers promising prospects for developing targeted therapies and precision medicine approaches in breast cancer treatment. By targeting these specific molecular drivers, we can aim for more effective and personalized treatment strategies to potentially improve patient outcomes. These findings contribute to our understanding of the molecular mechanisms underlying breast cancer and may have implications for identifying potential therapeutic targets and strategies. However, it is important to consider the study's limitations and further investigate the role of additional molecular factors in breast cancer pathogenesis.

Conclusion

Our integrated analysis of multiomic data from the TCGA-BRCA dataset has provided valuable insights into the regulatory mechanisms, network characteristics, and functional roles of genes, microRNAs (miRNAs), and proteins in breast cancer. Through this comprehensive analysis, we have identified 8 consistently upregulated hub genes (ACTB, HSP90AA1, FN1, HSPA8, CDC42, CDH1, UBC, and EP300) that play potentially significant roles in driving the disease. These hub genes serve as potential therapeutic targets for breast cancer treatment. Furthermore, pathway enrichment analysis has revealed highly enriched pathways in breast cancer, including proteoglycans in cancer, PI3K-Akt, and MAPK signaling. The dysregulation of these pathways highlights their importance in the development and progression of breast cancer. Understanding the functional roles of the hub genes within these pathways provides crucial information for developing effective treatment strategies.

Authors' Contributions

Z.H. wrote the first draft of the manuscript, and this report is a portion of her PhD research. Each author reviewed the submitted version of the essay and made contributions to it.

Ethical Considerations

For this article, public databases were used and there are no ethical considerations.

Acknowledgment

None declared.

Conflict of Interests

The authors declare that they have no competing interests.

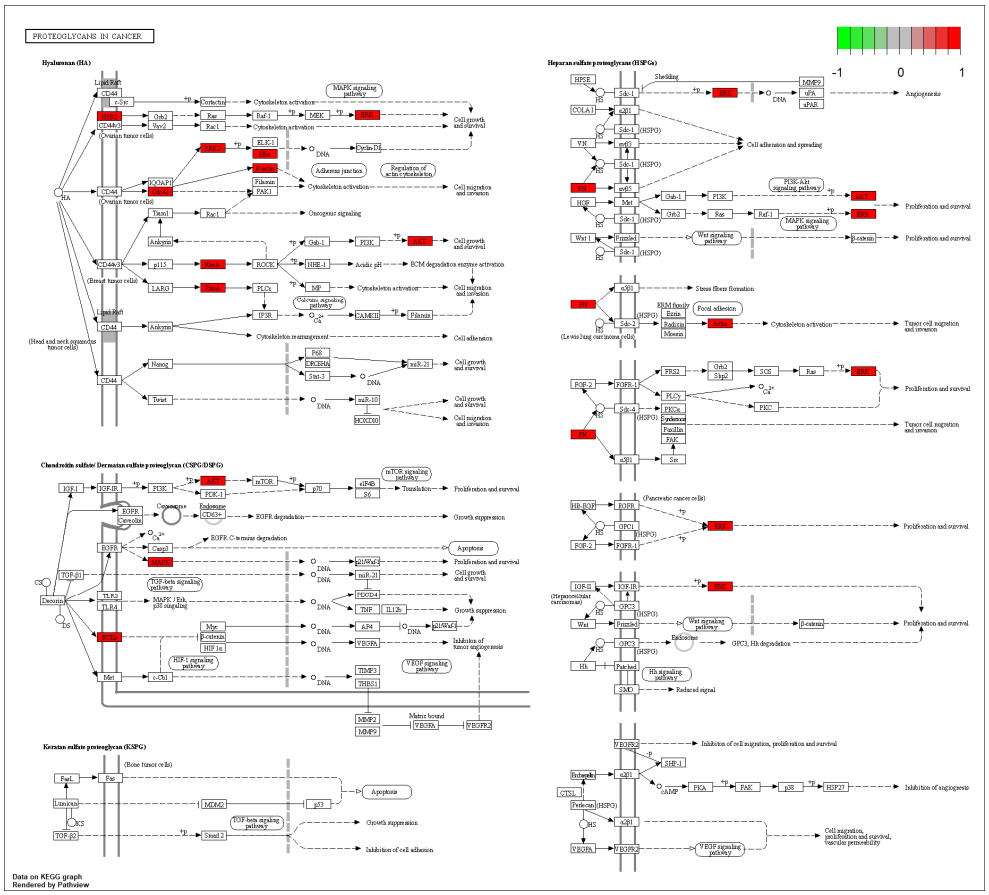
References

1. Guo L, Kong D, Liu J, Zhan L, Luo L, Zheng W, et al. Breast cancer heterogeneity and its implication in personalized precision therapy. Exp Hematol Oncol. 2023;12(1):3.
2. Malone ER, Oliva M, Sabatini PJB, Stockley TL, Siu LL. Molecular profiling for precision cancer therapies. Genome Medicine. Genome Med. 2020;12(1):8.
3. Russnes HG, Lingjærde OC, Børresen-Dale A-L, Caldas C. Breast Cancer Molecular Stratification: From Intrinsic Subtypes to Integrative Clusters. The American Journal of Pathology. AJP. 2017;187(10):2152-62.
4. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. Cell. Cell. 2011;144(6):986-98.
5. Netanely D, Stern N, Laufer I, Shamir R. PROMO: an interactive tool for analyzing clinically-labeled multi-omic cancer datasets. BMC Bioinformatics. BMC Bioinform. 2019;20(1):732.
6. Shehwana H, Kumar SV, Melott JM, Rohrdanz MA, Wakefield C, Ju Z, et al. RPPA SPACE: an R package for normalization and quantitation of Reverse-Phase Protein Array data. Bioinformatics. J Bioinform.

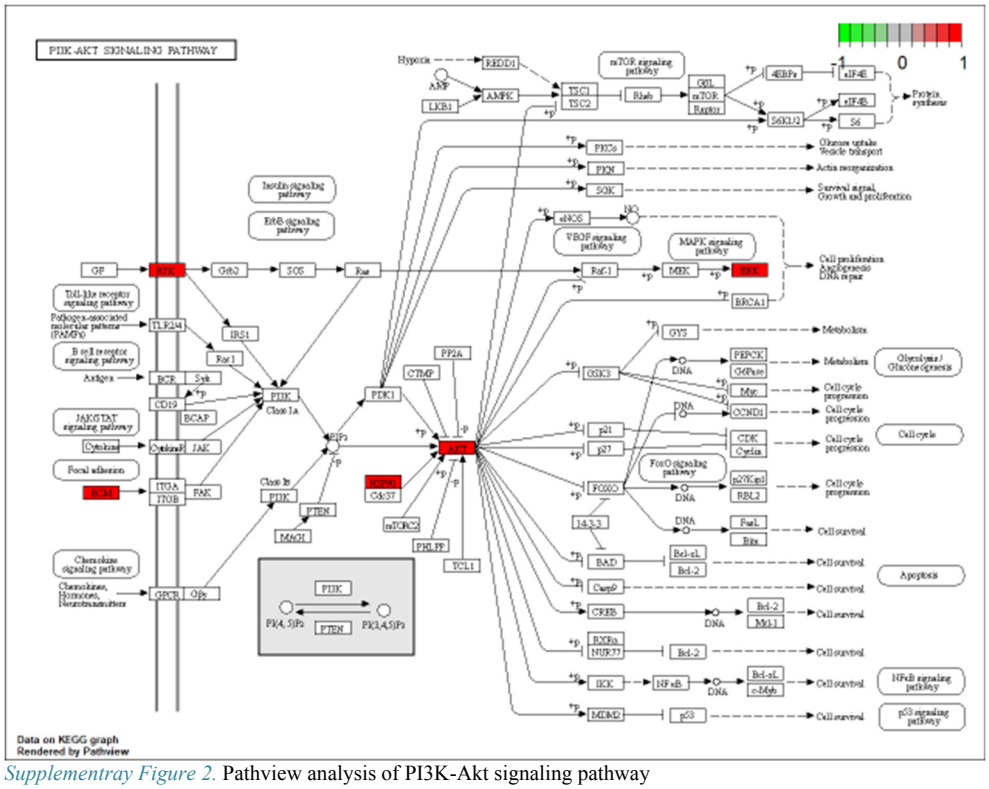
[DOI: 10.47176/mjiri.38.63]

2022;38(22):5131-3.
7. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, complex systems. 2006;1695(5):1-9.
8. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362-d8.
9. Clauset A, Shalizi CR, Newman ME. Power-law distributions in empirical data. SIAM Rev. 2009;51(4):661-703.
10. Shiryayev AN. Selected Works of AN Kolmogorov: Volume III Information Theory and the Theory of Algorithms: Springer Publishing Company, Incorporated; 2010.
11. Fatima U, Hina S, Wasif M. A novel global clustering coefficient-dependent degree centrality (GCCDC) metric for large network analysis using real-world datasets. Journal of Computational Science. J Comput Sci. 2023;70:102008.
12. Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. Nucleic Acids Res. 2014;42(17):e133.
13. Tenenbaum D. KEGGREST: Client-side REST access to KEGG. R package version. 2016;1(1).
14. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. J Bioinform. 2013;29(14):1830-1.
15. 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.
16. Clough E, Barrett T. The gene expression omnibus database. Statistical Genomics: Methods and Protocols. 2016:93-110.
17. Tsuyuki S, Takabayashi M, Kawazu M, Kudo K, Watanabe A, Nagata Y, et al. Detection of WIPI1 mRNA as an indicator of autophagosome formation. Autophagy. 2014;10(3):497-513.
18. Boughanem H, Yubero-Serrano EM, López-Miranda J, Tinahones FJ, Macias-Gonzalez M. Potential Role of Insulin Growth-Factor-Binding Protein 2 as Therapeutic Target for Obesity-Related Insulin Resistance. Int J Mol Sci. 2021;22(3):1133.
19. Ahmad MK, Abdollah NA, Shafie NH, Yusof NM, Razak SRA. Dual-specificity phosphatase 6 (DUSP6): a review of its molecular characteristics and clinical relevance in cancer. Cancer Biol Med. 2018;15(1):14-28.
20. Takahashi Y, Forrest AR, Maeno E, Hashimoto T, Daub CO, Yasuda J. MiR-107 and MiR-185 can induce cell cycle arrest in human non small cell lung cancer cell lines. PLoS One. 2009;4(8):e6677.
21. Lee KH, Lotterman C, Karikari C, Omura N, Feldmann G, Habbe N, et al. Epigenetic silencing of MicroRNA miR-107 regulates cyclin-dependent kinase 6 expression in pancreatic cancer. Pancreatology. 2009;9(3):293-301.
22. Iozzo RV, Sanderson RD. Proteoglycans in cancer biology, tumour microenvironment, and angiogenesis. J Cell Mol Med. 2011;15(5):1013-31.
23. Martini M, De Santis MC, Braccini L, Gulluni F, Hirsch E. PI3K/AKT signaling pathway and cancer: an updated review. Ann Med. 2014;46(6):372-83.
24. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 2002;12(1):9-18.
25. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. Nature. 2004;432(7015):316-23.

[DOI: 10.47176/mjiri.38.63]



Supplementray Figure 1. Pathview analysis of Proteoglycans in cancer pathway



Supplementray Figure 2. Pathview analysis of PI3K-Akt signaling pathway

[DOI: 10.47176/mjiri.38.63]