Scientific Research Publishing

The Expression and Prognostic Significance of Major MicroRNA Genes in Breast Cancer Based on Bioinformatics Analysis

Qiongge Wang*, Jiuli Hu*, Junwei Liang, Lanfang Liu, Shuoyang Xiao, Rui Wang#, Chanchan Hu#

Hospital of Chengde Medical College, Chengde, China

Email: #chanchanhu@126.com

Abstract

Objective: Breast Cancer (BC) is characterized by high complexity and heterogeneity, and microRNA (miRNA) is bound up with the occurrence and development of BC. In this study, we evaluated the prognostic value of miRNA in BC. **Background:** Breast ductal and lobular cancers are the most common types of Breast Carcinomas (BC) and indicate the high complexity heterogeneity in this disease. Each BC patient has unique morphological and molecular features. MicroRNAs (miRNAs) play a critical role in human oncogenesis, progression, and prognosis. Our study aimed to identify potential prognostic biomarkers of breast ductal and lobular cancers to predict the overall survival outcome. **Methods:** All analyzed miRNA sequencing and clinical data were obtained from the Genomic Data Commons Data Porta. edgeR package in R software was used to analyze the differential miRNA expression profiles. Complete survival information and differentially expressed miRNA expression were obtained and the Caret package was used for random division of the samples along with their profiles into two groups (training group and test group). We performed univariate Cox regression analyses for miRNAs in the training group. We utilized three different web-based tools to identify the target genes of miRNAs and used the Perl language to evaluate the target genes for miRNA signature. STRING database was used to assess PPIs. **Results:** A total of 304 differentially expressed miRNAs were identified (213 were upregulated and 91 were downregulated). Among these, nine (hsa-miR-204-5p, hsa-miR-7706, hsa-miR-1247-3p, hsa-miR-20b-3p, hsa-miR-605-5p, hsa-miR-615-3p, hsa-miR-4652-5p, hsa-miR-9-5p, hsa-miR-2115-5p) miRNAs which remarkably correlated with overall survival rate of BC were performed

---

*Qiongge Wang and Jiuli Hu have contributed equally to this work.
#Chanchan Hu and Rui Wang have equally cooperated and should be considered the joint corresponding authors.

by Cox regression analysis and miRNA signature risk score built. And then we performed the model of BC patients for three years survival risk, the AUCs of ROC were 0.804, 0.667, and 0.739 in the training, test, and entire groups, respectively. miRNAs were differentially expressed in tumor-related biological processes and pathways by functional enrichment and bioinformatic analysis. **Conclusion:** The current study provided novel insights into the miRNA-based mRNA network in BC. The nine miRNA and ten hub genes may be independent prognostic signatures for survival prediction in BC patients.

Keywords

Breast Cancer, miRNA, Overall Survival, Biomarkers

1. Introduction

At present, BC of female patients has surpassed lung cancer as the leading cause of global cancer incidence. Globally, with an estimated 2.3 million new cases being reported, which corresponds to 11.7% of all cancer cases, it is the fifth leading cause of cancer-related mortality, contributing to 685,000 deaths [1]. There are several histologic subtypes of invasive breast carcinoma, including, but not limited to the following: invasive carcinoma of no special type (commonly known as invasive ductal carcinoma); invasive lobular carcinoma; tubular carcinoma; carcinoma with medullary features (sometimes called medullary carcinoma); and metaplastic carcinoma. Among these subtypes, Invasive (or Infiltrating) Ductal Carcinoma (IDC) accounts for the majority (approximately 75%) of BCs, followed by Invasive Lobular Carcinoma (ILC) which accounts for 5% - 15% of BCs [2]. BC is a heterogeneous and complex disease with unique morphological and molecular features, unlike a disease where only a few genes, proteins, and/or signaling pathways contribute to its progression in a simple, independent, and autonomous manner. Previous studies showed that patients with the same type of BC can show differential treatment responses, which further contributes to the high heterogeneity in this disease [3]. Therefore, histological features have certain limitations for diagnosis, prognosis, and prediction of clinical outcomes for individual patients [4]. Many prognosis factors are available for BC; therefore, it is necessary to develop suitable prognostic biomarkers for the diagnosis and treatment of BC.

MicroRNAs (miRNAs) are a group of small non-coding RNAs that can interrupt the expression of protein-coding genes by binding to their mRNAs and inhibiting the subsequent protein translation [5]. miRNAs can regulate the expression of genes associated with various biological phenomena including homeostasis, development, proliferation, differentiation, and apoptosis [6]. Deregulated signaling and expression of miRNAs have been well-studied in the pathogenesis of various cancers. Aberrant expression of miRNAs is vital for the initiation and

progression of human malignancies as they function as both tumor suppressors and oncogenes [7]. This study aimed at identifying the potential prognostic biomarkers to predict overall survival in BC. We also evaluated the clinical potential of miRNAs as biomarkers for early as well as differential diagnosis and prognosis of BC.

2. Methods

Data Source and Data Processing

We obtained the first dataset for our analyses from The Genomic Data Commons Data Portal (GDC, https://portal.gdc.cancer.gov/), a robust data-driven platform. MiRNA sequencing data for invasive carcinoma samples of the breast (103 normal vs 1057 tumor), mRNA sequencing data for invasive carcinoma samples of the breast (112 normal vs 1066 tumor), and the corresponding clinical information were obtained TCGA. Then we obtained the information on individual mature miRNA expression profiles through Perl language.

Survival Data and Identification of Differentially Expressed miRNAs and mRNAs

EdgeR package (R software v. 4.0.3) was used to compare the DEMs and DEGs in tumor group with normal group. The expression profiles of miRNAs and mRNAs were normalized, according to mean value > 1, p-value (FDR) < 0.05, and |log-2FC| > 1. The clinical information of patients with survival time ≥ 30 days was obtained, and then we correspondingly combined with differentially expression profiles information of miRNA and mRNA.

Evaluation of the Prognostic Model

The Caret package was used for random division of the samples into two groups (training group and test group) with complete survival information and differentially miRNA expression profile information. We also performed univariate Cox regression analyses of miRNAs in the training group.

To reduce the number of miRNAs with a similar expression, we subjected these to a stepwise multivariate Cox regression and constructed the miRNA-based prognostic model with a p-value < 0.005. Then, the risk scores for the prognostic miRNA signature comprising multiple miRNAs were established based on the summation of the product of each miRNA and its coefficient. Then we tested the Proportional Hazards Assumption of the Cox model. Further, this model was used to evaluate the survival prognosis for each patient in the training and test groups. Furthermore, we classified the patients into the high-risk group and the low-risk group on median value grouping of risk score performed by Kaplan-Meier curve and log-rank tests. The predictive power of the miRNA signature was assessed using the "survival ROC" package by calculating the AUC of 3 year-ROC curves [8].

Independent Prognostic Evaluation of miRNA Signature for Clinical Variables

After the classification based on the age, gender, size, and/or extent of main

cancer, clinical-stage, lymph nodes, distant metastasis, we used univariate Cox regression analysis for evaluating the prognostic value of miRNA signature and overall patient survival in the training group. Variables with p-value < 0.05 in univariate Cox regression were further used for multivariate Cox regression analysis which determined whether these variables could function as independent prognostic factors. Forest plots were used for comparison of the predictive power of this risk model with other clinical characteristics.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichment Analyses

We utilized three different web-based tools, miRTarBase (https://mirtarbase.cuhk.edu.cn/), targetScan (http://www.targetscan.org) and miRDB (http://www.mirdb.org/), to identify the target genes of the miRNAs and collate the target genes of miRNA signature which were covered in at least 2 databases. Venn diagram and Cytoscape 3.7.1 were used to map the relationship between the miRNA and these target genes. We performed functional gene enrichment analyses to identify the major biological attributes of these genes. Annotation by Gene Ontology (GO) and enrichment analyses by Kyoto Encyclopedia of Genes and Genomes (KEGG) were performed through "clusterProfiler" package [9] and the "org.Hs.eg.db" package in R software, with p adjust < 0.05 and q value < 0.05, set as the cut-off criteria.

Protein-Protein Interaction (PPI) Network and Module Analyses

STRING database (http://string-db.org/) was used for information on PPIs. The medium confidence interval of 0.4 was used as the threshold. Cytoscape_v3.7.1 software was used to visualize the resulting PPI network, and we identified the top 10 hub genes which formed the hub. The Kaplan-Meier method with log-rank test < 0.05 was used to verify the intersection gene correlation with overall survival.

Statistical Analysis

All statistical analyses were performed by R language 4.0.3 version and attached packages.

3. Results

### Identification of DEMs and DEGs

TCGA is an open-access database containing miRNA/mRNA profiles and their corresponding patient clinical information. Based on the screening criteria, 304 miRNAs were differentially expressed (DEmiRNAs) in 1057 tumor and 103 normal samples; 213 DEmiRNAs were upregulated and 91 were downregulated. 7364 mRNAs were differentially expressed between 1066 tumor and 112 normal samples; 4894 DEmRNAs were up-regulated while 2470 were downregulated (Figure 1).
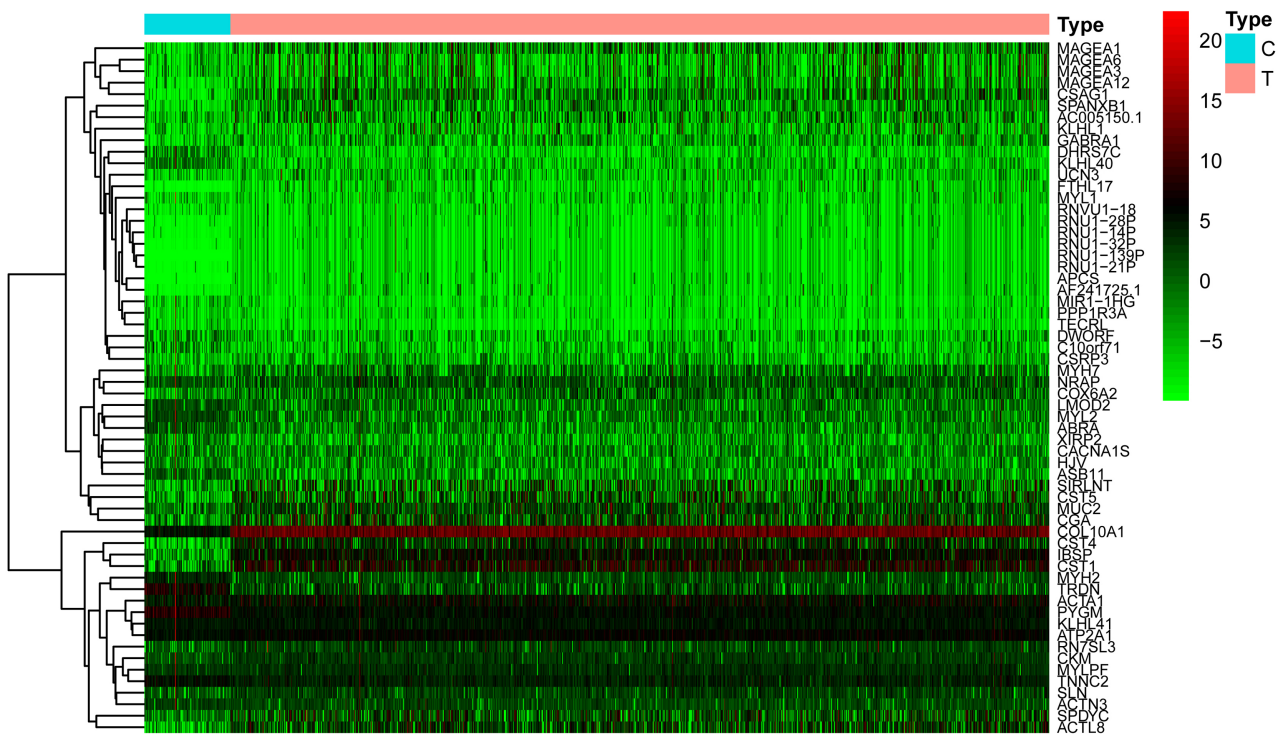
Prognostic Model Construction

The miRNA mature expression profiles of the total group (N = 1033) were randomly divided into the training group (N = 517) and test group (N = 516). Based
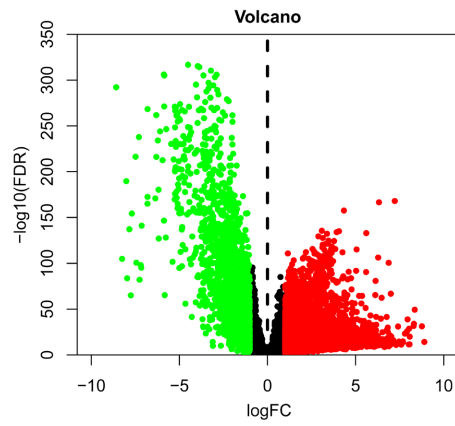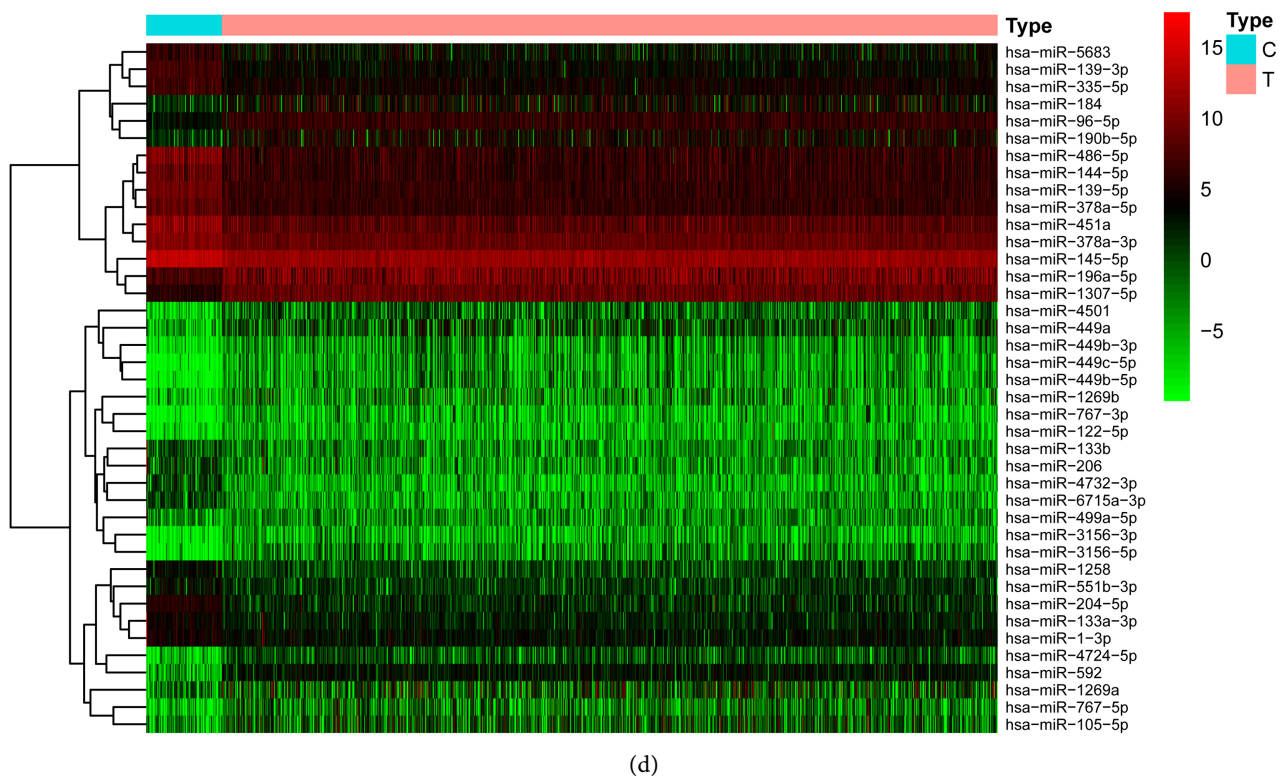
(a)
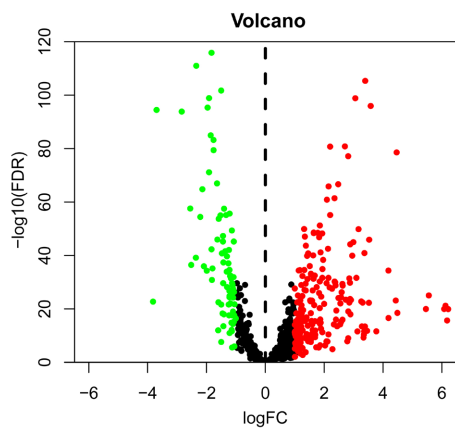
(b)

(c)

(d)

Figure 1. Differentially expressed mRNA and miRNA between Breast Cancer (BC) and normal tissues. (a) The volcano plot for mRNA from the TCGA data portal; (b) The volcano plot for miRNA from the TCGA data portal; (c) Hierarchical clustering of differentially expressed mRNA expression levels; (d) Hierarchical clustering of differentially expressed miRNA expression levels.

on univariate Cox regression results for the training group, a total of fifteen miRNAs were significantly associated with overall patient survival (p-value < 0.05) (Table 1).

In our study, nine (hsa-miR-204-5p, hsa-miR-7706, hsa-miR-1247-3p, hsa-miR-20b-3p, hsa-miR-605-5p, hsa-miR-615-3p, hsa-miR-4652-5p, hsa-miR-9-5p, hsa-miR-2115-5p) miRNAs were selected (Table 2) by stepwise multivariate Cox regression analysis. A predictive miRNA signature model was established by the summation of the product of each miRNA and its multivariate Cox regression coefficient as follows: miRNA signature risk score = ((−0.249) × expression of hsa-miR-204-5p) + (0.228 × expression of hsa-miR-7706) + ((−0.243) × expression of hsa-miR-1247-3p) + ((−0.306) × expression of hsa-miR-20b-3p) + (0.485 × expression of hsa-miR-605-5p) + (0.136 × expression of hsa-miR-615-3p) + (0.194 × expression of hsa-miR-4652-5p) + (0.151 × expression of hsa-miR-9-5p) + (0.339 × expression of hsa-miR-2115-5p). Kaplan-Meier method showed that these nine miRNAs were significantly associated with overall patient survival (p value < 0.05; Figure 2).
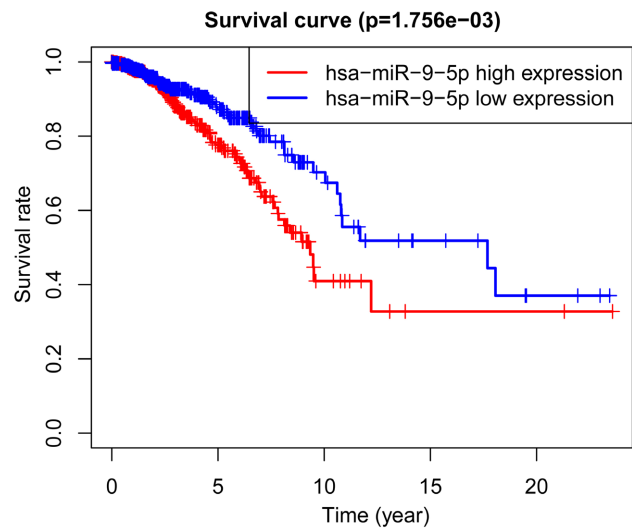
Evaluation of Overall Survival

According to median value grouping based on the risk score in the training, test, and entire groups, Kaplan-Meier curves showed that indeed the high-risk group had poorer overall survival as compared to the low-risk group (p = 1.474E−08, p = 8.708E−03, and p = 7.825E−09, respectively; Figures 3(a)-(c)). The 5-year

Table 1. Univariate and multivariate Cox regression of differentially expressed miRNAs.
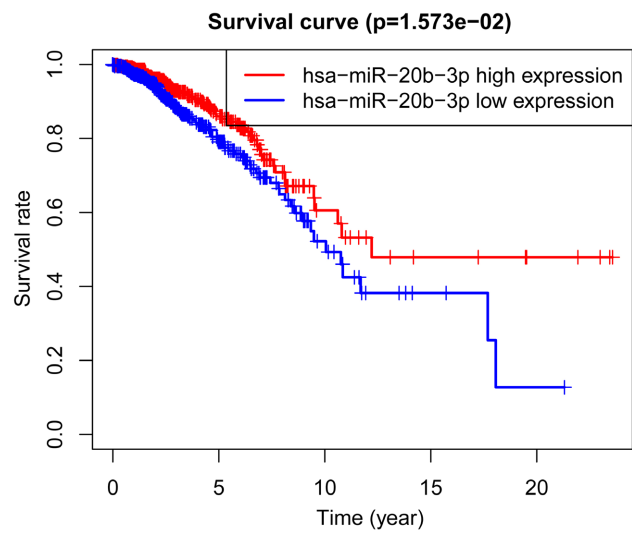
Id	Univariate Cox Regression				Multivariate Cox Regression				
	HR	HR.95L	HR.95H	p-value	Coef	HR	HR.95L	HR.95H	P-value
Hsa-MiR-204-5P	0.788	0.667	0.931	0.005	−0.249	0.780	0.642	0.946	0.012
Hsa-MiR-7705	1.354	1.082	1.695	0.008					
Hsa-MiR-148b-5P	1.401	1.079	1.818	0.011					
Hsa-MiR-454-3P	1.363	1.066	1.742	0.013					
Hsa-MiR-3677-3P	1.210	1.018	1.437	0.030					
Hsa-MiR-379-5P	0.800	0.664	0.964	0.019					
Hsa-MiR-301b-5P	1.346	1.119	1.617	0.002					
Hsa-MiR-7706	1.371	1.075	1.750	0.011	0.228	1.256	0.999	1.580	0.051
Hsa-MiR-1247-3P	0.793	0.701	0.898	≤ 0.001	−0.243	0.784	0.683	0.901	0.001
Hsa-MiR-20b-3P	0.838	0.709	0.990	0.038	−0.306	0.737	0.620	0.875	≤ 0.001
Hsa-MiR-605-5P	1.549	1.124	2.134	0.008	0.485	1.625	1.173	2.251	0.004
Hsa-MiR-615-3P	1.187	1.008	1.397	0.040	0.136	1.146	0.958	1.370	0.136
Hsa-MiR-4652-5P	1.222	1.010	1.478	0.039	0.194	1.214	0.991	1.486	0.061
Hsa-MiR-9-5P	1.167	1.055	1.291	0.003	0.151	1.162	1.039	1.300	0.009
Hsa-MiR-5691	0.728	0.548	0.968	0.029					
Hsa-MiR-9-3P	1.171	1.031	1.329	0.015					
Hsa-MiR-1229-3P	1.231	1.001	1.514	0.049					
Hsa-MiR-2115-5P	1.254	1.045	1.503	0.015	0.339	1.404	1.156	1.705	0.001

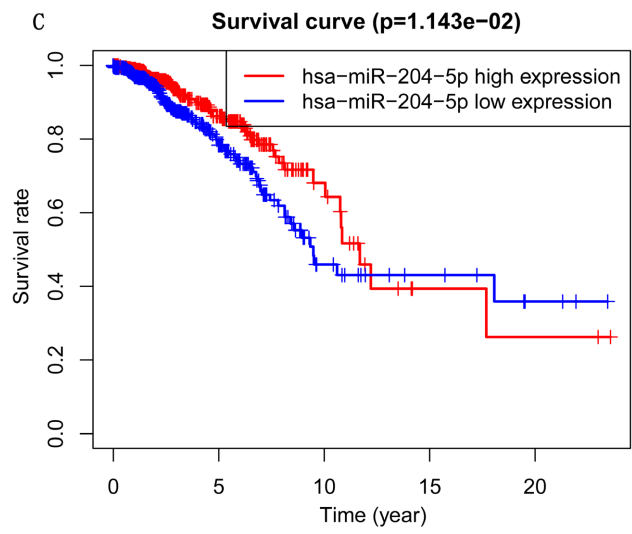**Table 2.** 53 positive expression of genes correlate with survival prognosis.

Oncogene		Anti-Oncogene	
ANO6	FAM155B	C1orf189	C2orf88
POP1	CENPI	CACNA1E	CCR5
PTK6	CLEC3A	CD5	CHST3
RACGAP1	B4GALT3	RYR1	RIC3
REEP1	KNSTRN	SELL	S100B
SOX11	MRPS23	SLA2	SH2D1A
ZNF367	POU3F2	SNX22	SLC6A2
LRRC59	FFAR2	TENM1	SPIB
DLG3	CHD5	TTC24	TIPARP
C2CD2		UNC13D	USP44
		PTPRZ1	FREM1
		PTX3	PRKCQ
		PAK3	CLEC10A
		NTRK2	MYH1
		INSM1	GLT1D1
		EDAR	DPF3
		DACT2	CX3CL1

Survival curve (p=1.756e−03)

(a)

Survival curve (p=1.573e−02)

(b)

Survival curve (p=1.143e−02)

(c)

Survival curve (p=7.417e−03)

(d)

Survival curve (p=7.278e−03)

(e)

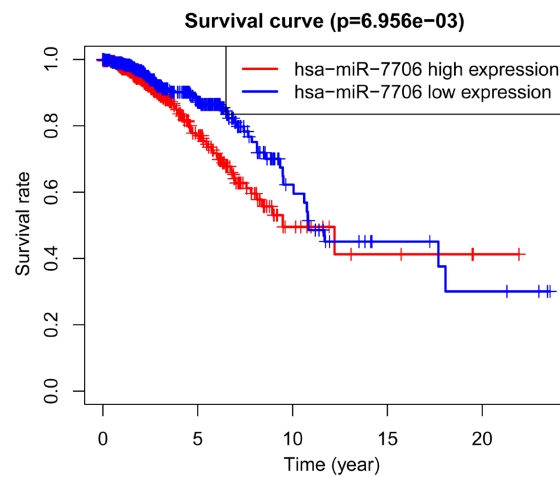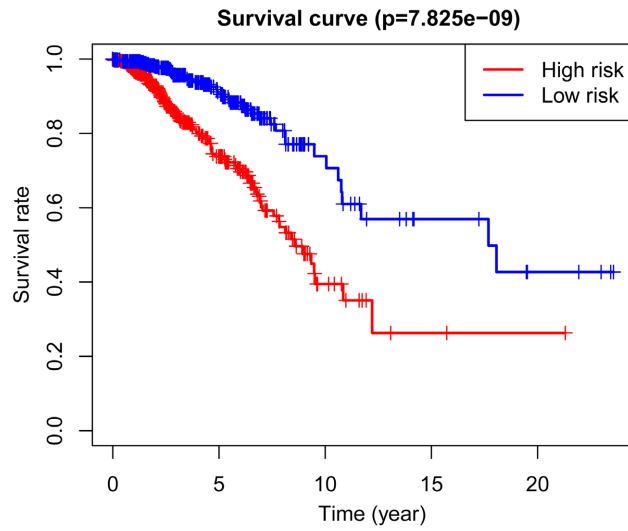**Survival curve (p=2.152e−02)**
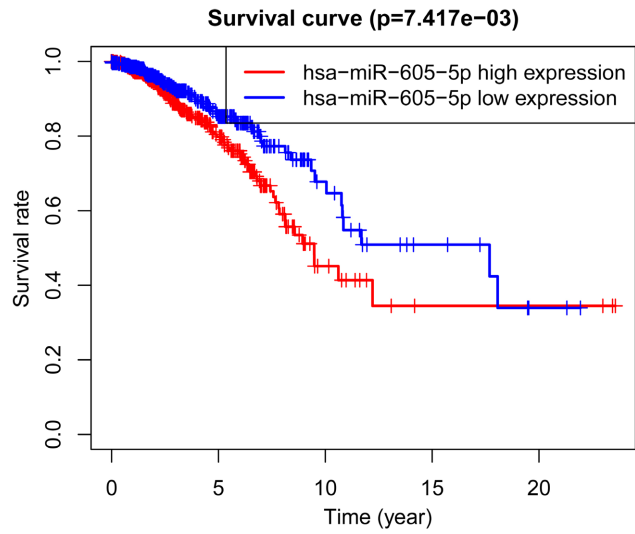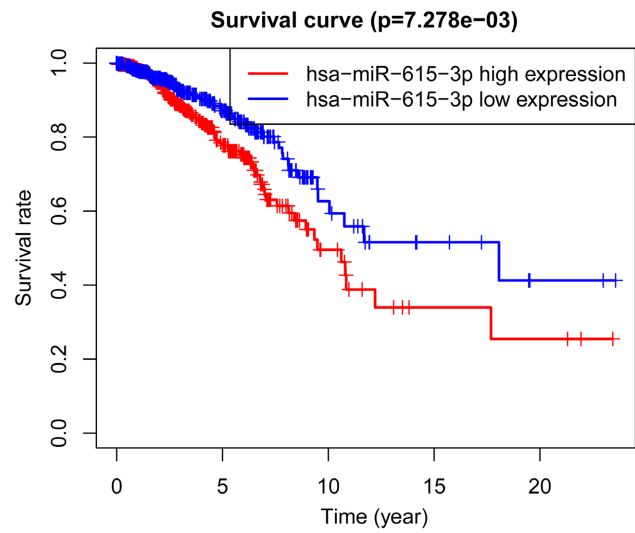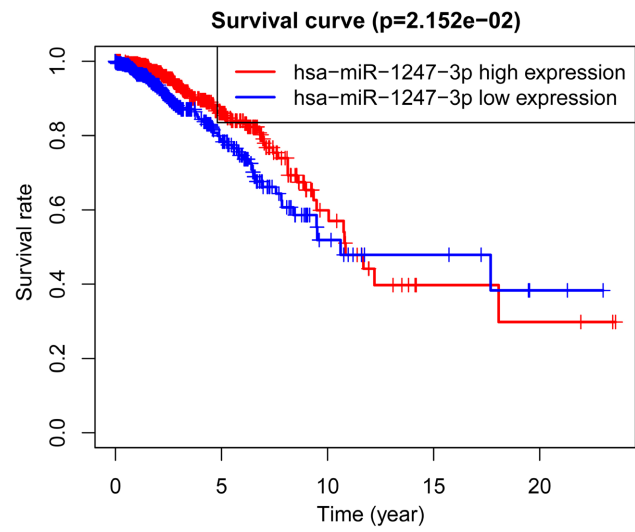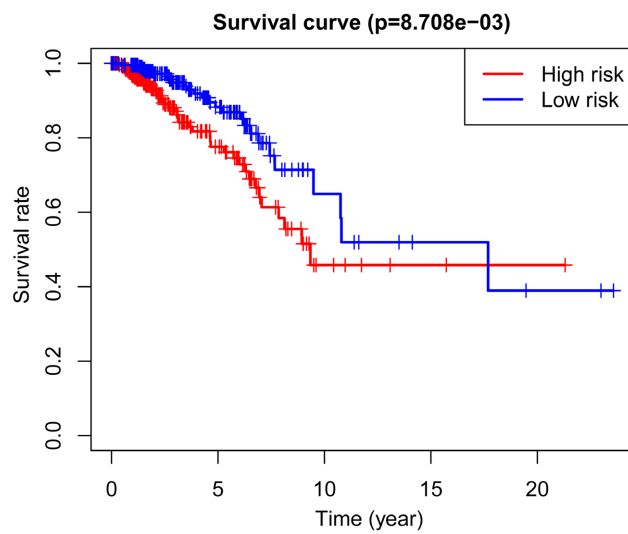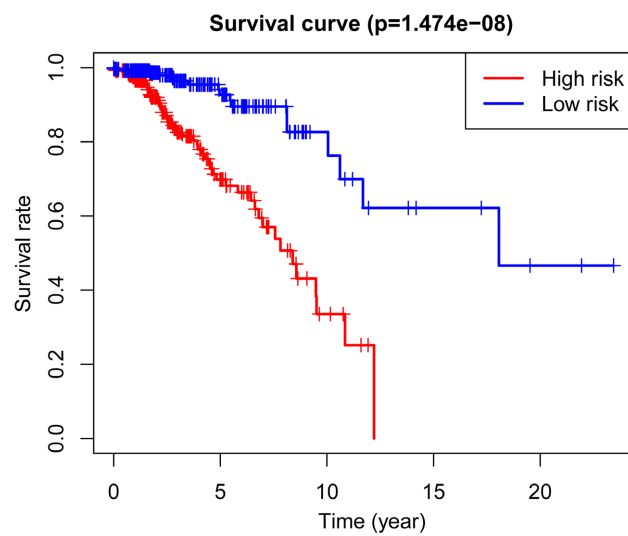

(f)
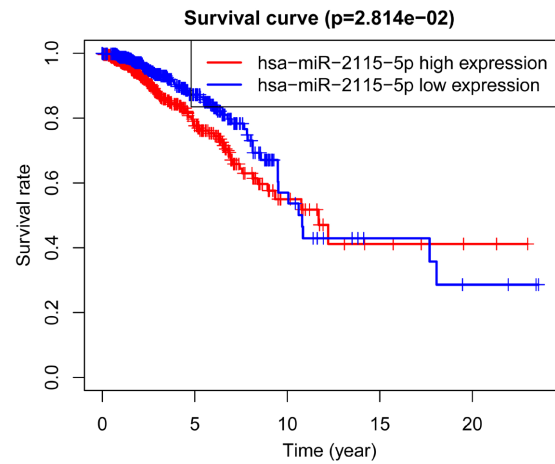
(g)
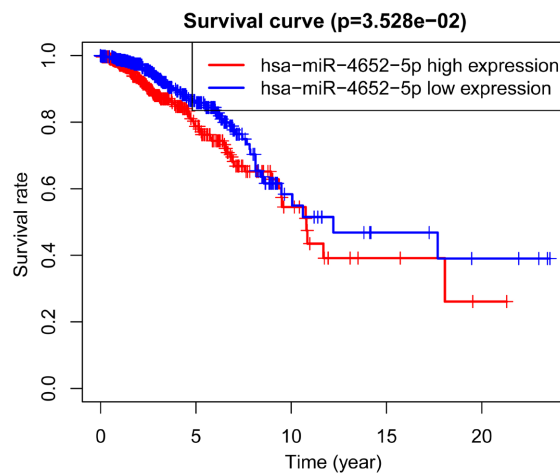
(h)

(i)

Figure 2. Nine miRNAs associated with overall survival in BC patients using Kaplan–Meier curves and log-rank tests. The patients were stratifified into high and low expression groups according to the median expression of each miRNA. (a) Hsa-miR-9-5p. (b) Hsa-miR-20b-3p; (c) hsa-miR-204-5p; (d) Hsa-miR-605-5p; (e) Hsa-miR-615-3p; (f) Hsa-miR-1247-3p; (g) Hsa-miR-2115-5p; (h) Hsa-miR-4652-5p; (i) Hsa-miR-7706.
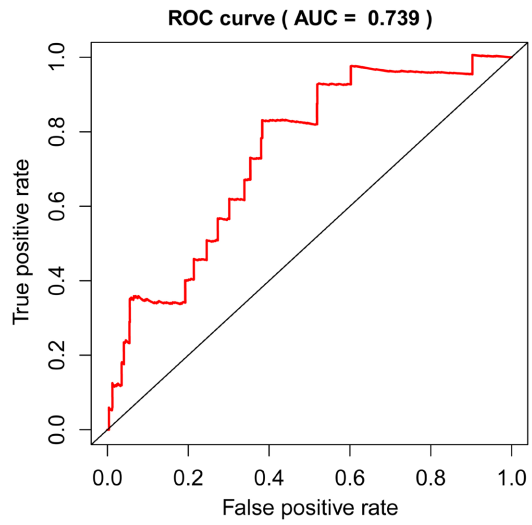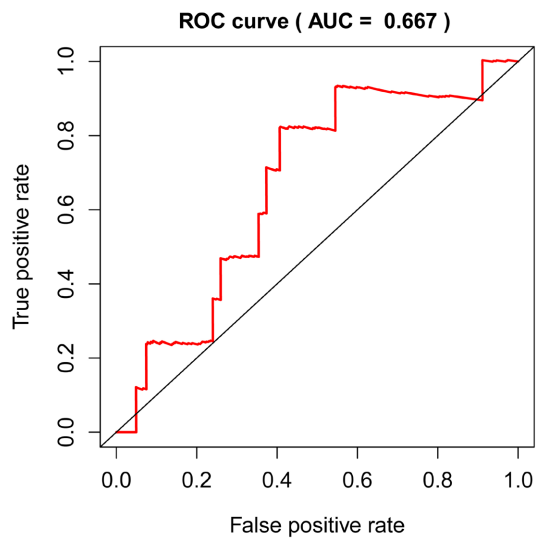
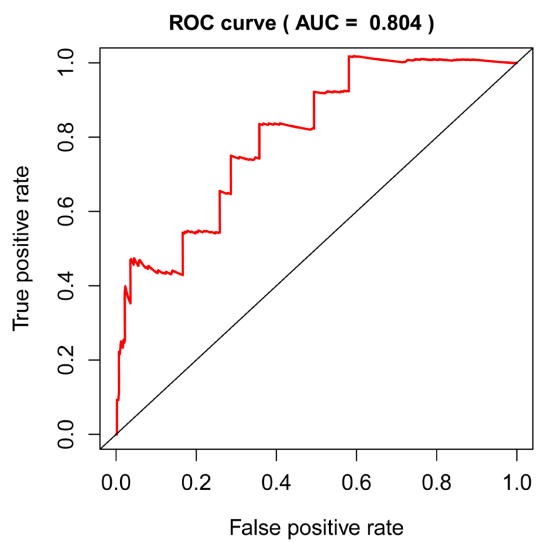**Survival curve (p=7.825e−09)**

(a)

Survival curve (p=8.708e−03)

(b)

Survival curve (p=1.474e−08)

(c)

ROC curve (AUC = 0.739)

(d)

ROC curve (AUC = 0.667)
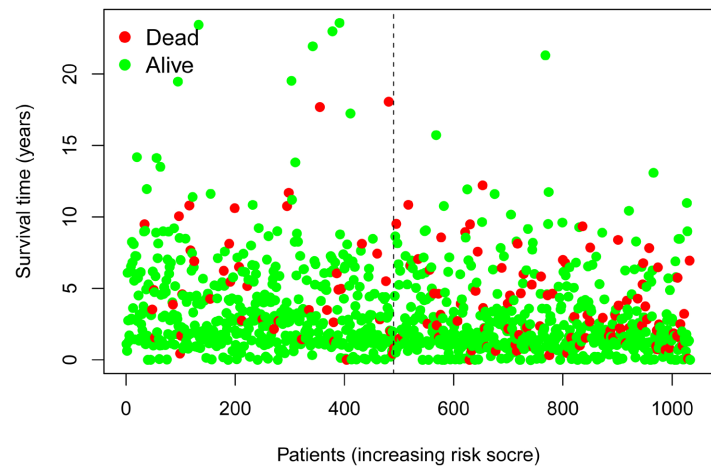

(e)

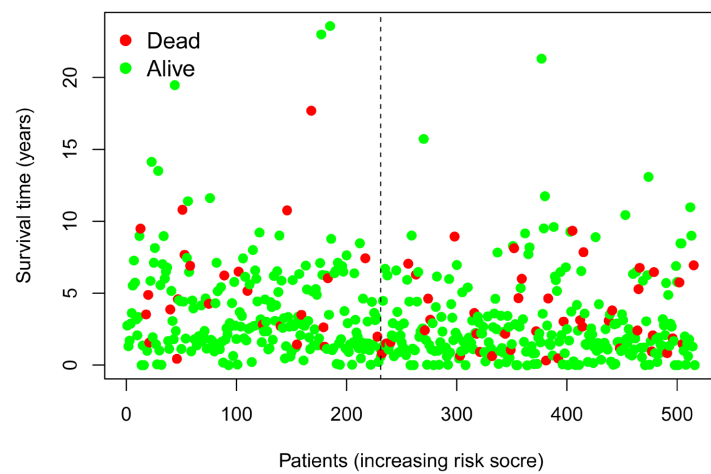**ROC curve (AUC = 0.804)**

(f)

(g)

(h)

(i)

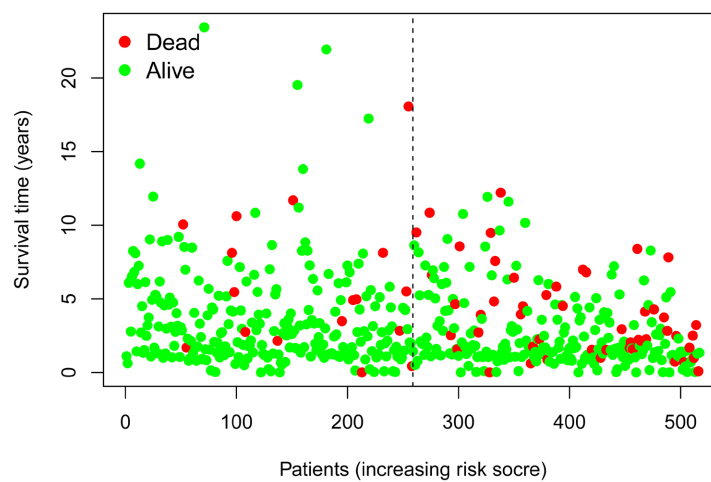**Figure 3.** Validation and evaluation of the predictive nine-miRNA signature. Kaplan-Meier curves in the entire group (a), test group (b), training group (c); The AUC of three years dependent curve in the entire group (d), test group (e), training group (f), Survival status in high and low risk patients for entire group (g), test group (h), training group (i), red dots represent death, green dots represent alive.

overall survival for training-group patients was 69.8% and 92.7% for the high and low-risk groups, respectively. The 5-year overall survival for test group patients was 76.1% and 86.9% in the high and low-risk groups, respectively. The entire group showed a 5-year overall survival of 73.8% and 90.5% in the high and low-risk groups, respectively.

The AUC of ROC based on the nine-miRNA signature for three years were 0.804, 0.667, 0.739 in the training, test, and entire groups, respectively (Figures 3(d)-(f)), which showed that the model had better performance for prediction of BC patient survival risk. In addition, in all the three groups, the high-risk score patients had higher mortality rates as compared to their low-risk score counterparts (Figures 3(g)–(i)).

Independence of Clinical Factors

The nine-miRNA signature was evidently associated with overall patient survival in univariate Cox regression analysis (hazard ratio HR = 1.051, confidence interval 95% CI = 1.025 - 1.078, p = 1.148E−04). Multivariate Cox regression analysis showed that the nine-miRNA signature was also independent of the overall survival for conventional clinical factors (HR = 1.039, 95% CI = 1.012 - 1.066, p = 3.794E−03), such as T stage, lymph node status, and distant metastasis. This makes it a potential prognostic marker for BC. In addition, clinical stage and age were also found to be independent prognostic factors (HR = 1.738, 1.032, 95% CI = 1.036 - 2.916, 1.016 - 1.048, p = 0.036, 8.56E−05) (Figure 4).

Prediction of Target Genes

We predicted the target genes regulated by the nine miRNAs from at least 2 databases. The overlapping target genes were identified to further enhance the reliability of the bioinformatic analyses. Using the above three databases, the results indicated that 1314, 431, 1297, 609, 151, 381, 627, 331 and 82 genes overlapped for hsa-miR-9-5p, hsa-miR-20b-3p, hsa-miR-204-5p, hsa-miR-605-5p, hsa-miR-615-3p, hsa-miR-1247-3p, hsa-miR-2115-5p, hsa-miR-4652-5p, hsa-miR-7706, respectively, as shown using the Venn diagram (Figure 5) and the network map of miRNA target genes. A total of 15404 target genes were predicted for the nine miRNAs. To assess whether the target genes of these miRNAs participated in the progression of BC, the 7364 DEmRNAs (upregulated 4894, downregulated 2470) described above, were used for further analyses. The intersection of target mRNAs for upregulated miRNAs (hsa-miR-9-5p, hsa-miR-20b-3p, hsa-miR-2115-5p, hsa-miR-4652-5p, hsa-miR-7706, hsa-miR-615-3p) and downregulated mRNAs and target mRNAs for down-regulated miRNAs (hsa-miR-204-5p, hsa-miR-605-5p, hsa-miR-1247-3p) and upregulated mRNAs were taken. The results were obtained for a total of 555 genes including 261 upregulated genes, 294 downregulated genes. The sub-network of the nine miRNAs and their 555 target genes are shown in Figure 6.

GO Annotation and KEGG Enrichment Analyses

We performed GO and KEGG analyses for a detailed understanding of the selected 522 target genes associated with BC. The top thirty terms from the GO annotations for Biological Process (BP), Cellular Component (CC), and Molecular
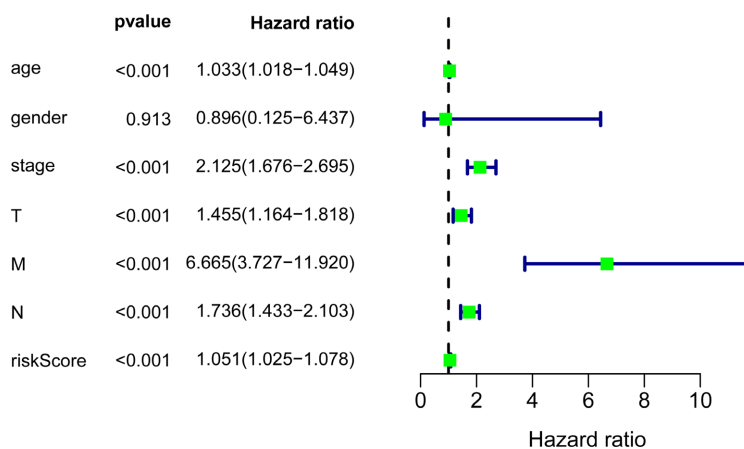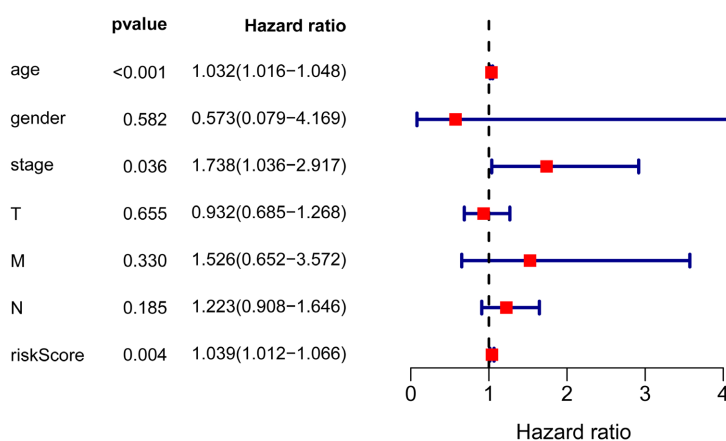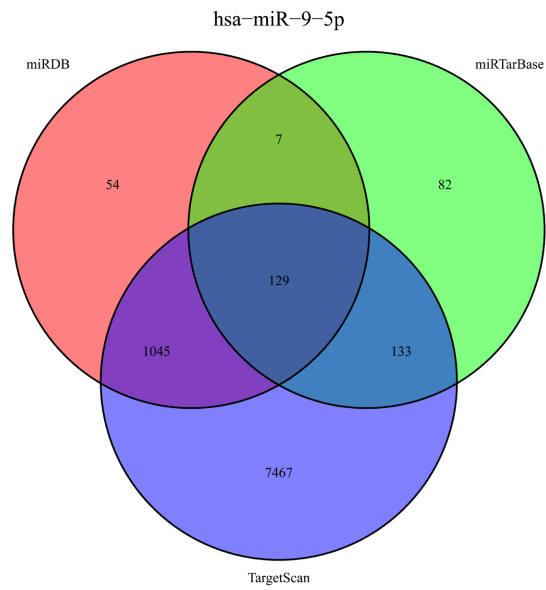
Figure 4. Comparison of risk score and clinical features in predicting the accuracy of patients' survival prognosis. (a) Forest plot of univariate Cox regression analysis; (b) Forest plot of multivariate Cox regression analysis.
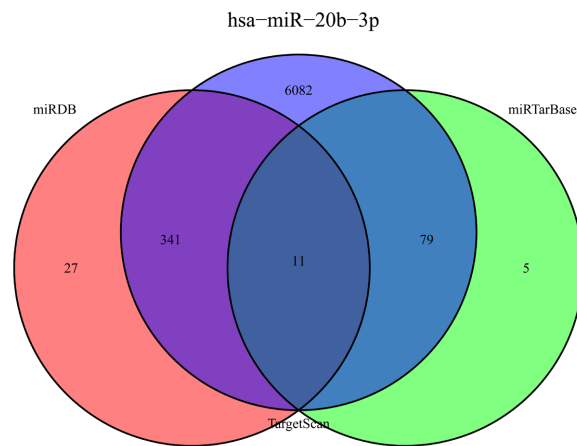
Function (MF) are shown in dotplot (**Figures 7(a)-(c)**). BP included 447 annotations such as axonogenesis, embryonic organ development, and sensory system development. CC contained 50 annotations including synaptic membrane, collagen-containing, and extracellular matrix transcription regulator complex. There were 25 MF annotations, including metal ion transmembrane transporter activity, ion channel activity, and actin-binding. A total of 433 KEGG pathways were enriched, of which 5 signaling pathways were significantly enriched including calcium signaling pathway, cytokine-cytokine receptor interaction, MAPK signaling pathway, proteoglycans in cancer, and cAMP signaling pathway. In addition, a readable graphic representation of the complex relationship between target genes and relative KEGG pathway using the "pathway-gene network" is shown in **Figures 7(d)-(e)**.

PPI Network and Module Selection

Using the STRING database and Cytoscape software, a total of 546 target genes were incorporated into the PPI network consisting of 178 nodes and 327

hsa−miR−9−5p

(a)

hsa−miR−20b−3p

(b)

hsa−miR−204−5p

(c)

hsa−miR−605−5p

(d)

hsa−miR−615−3p

(e)

hsa−miR−1247−3p

(f)
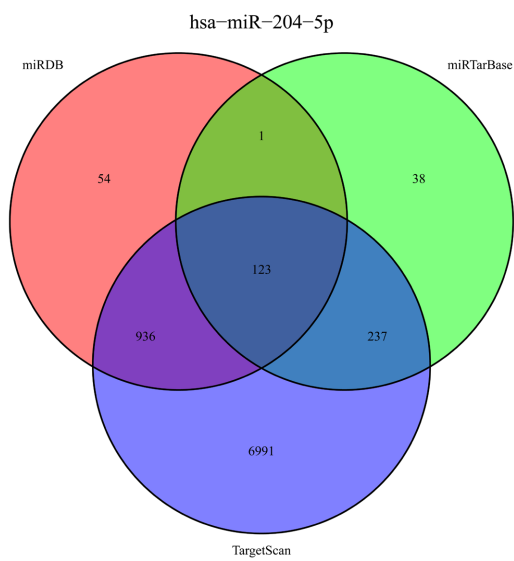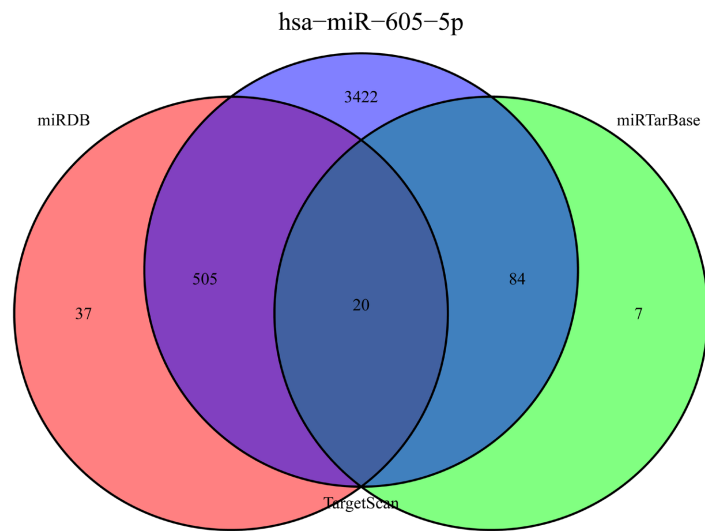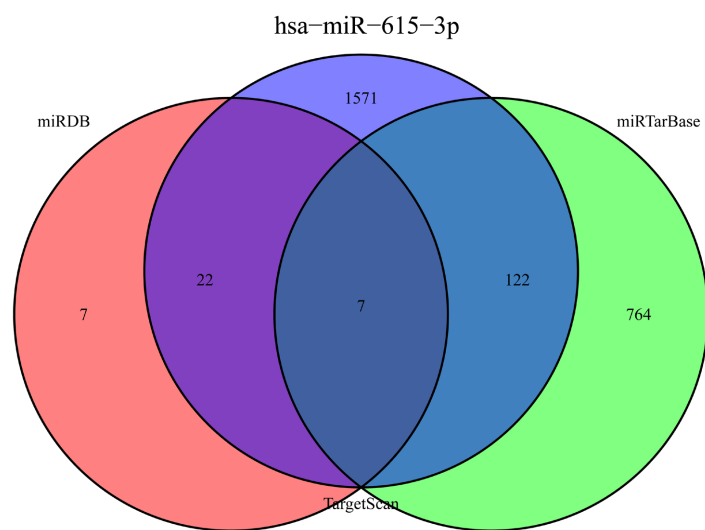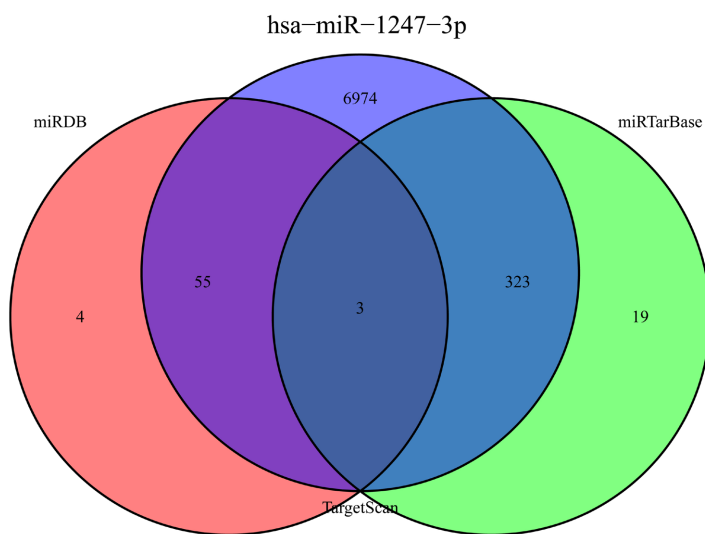
Figure 5. Venn diagram of target genes for nine miRNAs. (a) Hsa-miR-9-5p; (b) Hsa-miR-20b-3p; (c) Hsa-miR-204-5p; (d) Hsa-miR-605-5p; (e) Hsa-miR-615-3p; (f) Hsa-miR-1247-5p; (g) Hsa-miR-2115-5p; (h) Hsa-miR-4652-5p; (i) Hsa-miR-7706.

Figure 6. Sub network map of miRNA regulating mRNA. The hexagon represents miRNA; The circle stands for mRNA; Red means upregulated; Green means downregulated.

edges. 10 hub genes (GNG7, CDK1, EGF, BIRC5, STAT1, ESR1, FZD4, DSN1, APOB, GNAI1) were identified as hubs (Figure 8). In addition, the Kaplan-Meier method showed 53 (Table 2) of the 546 gene expressions were positively correlated with survival prognosis; a high expression of ANO6, B4GALT3, C2CD2, CENPI, CHD5, CLEC3A, DLG3, FAM155B, FFAR2, KNSTRN, LRRC59, MRPS23, 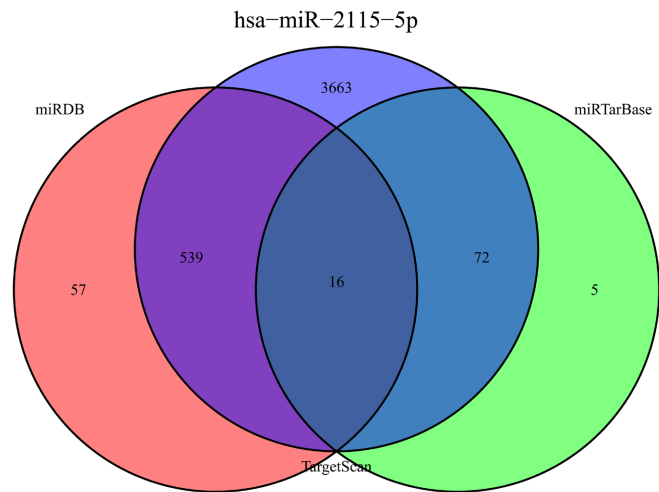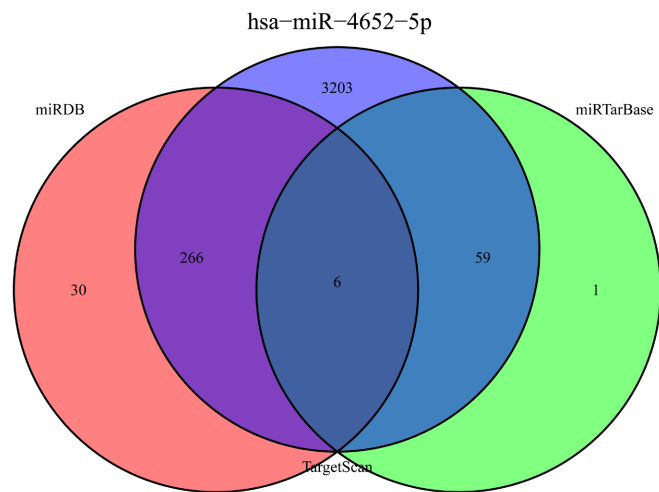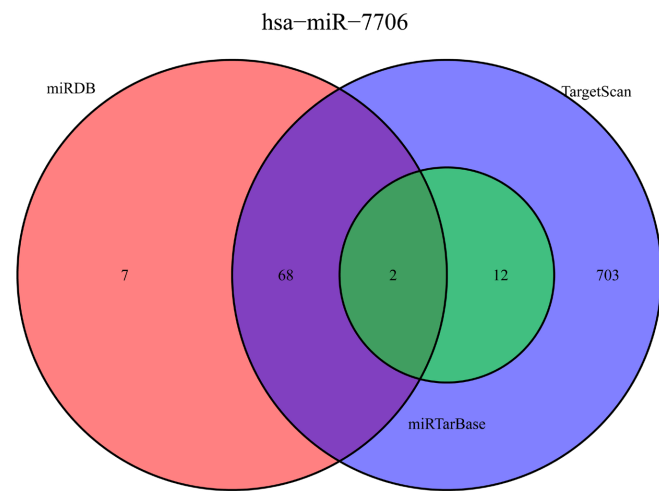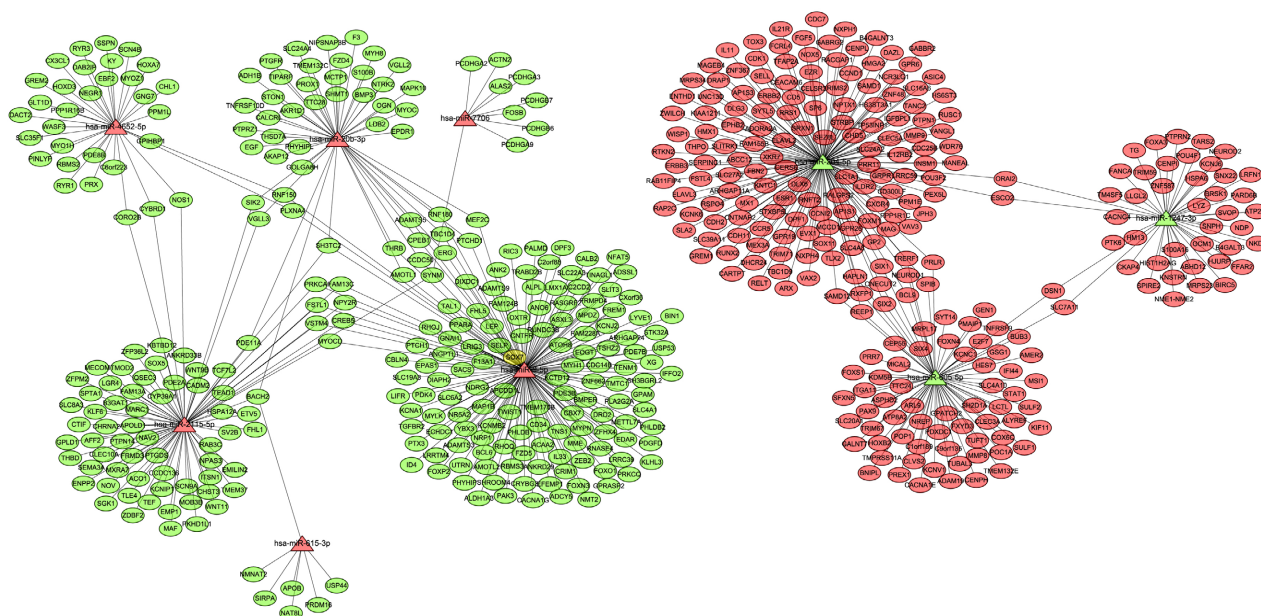POP1, POU3F2, PTK6, RACGAP1, REEP1, SOX11, ZNF367 was associated with poorer overall survival in BC patients (Figure 9).

4. Discussion

Among women, the three most frequent malignancies are breast, lung, and colorectal cancers. These account for 50% of all new diagnoses; breast cancer alone accounts for 30% of female tumors [10]. Recently, with the advancement in early diagnoses and comprehensive treatment strategies, the mortality rate of BC has gradually decreased. However, distant metastasis is a major cause of death among these patients [11] [12]. Therefore, the identification of highly specific and sensitive prognostic markers for BC is important. Recently, miRNAs have been widely demonstrated to be associated with a variety of diseases including breast cancer [13]. A vast and growing body of evidence firmly supports the involvement of miRNAs in tumors including breast cancer [14]. However, most of these have studied signal genes. To capture the genes associated with DEMs in breast cancer, we screened the DEGs and DEMs and identified the key prognostic DEMs, all of which may be potential therapeutic targets. In the present study, from the TCGA database, the miRNA and mRNA expression profile data on invasive carcinoma of the breast were analyzed. DEMs and DEGs were analyzed using the edgeR package, and their targets were predicted. All the patients were randomly

(a)

(b)

(c)

(d)

(e)
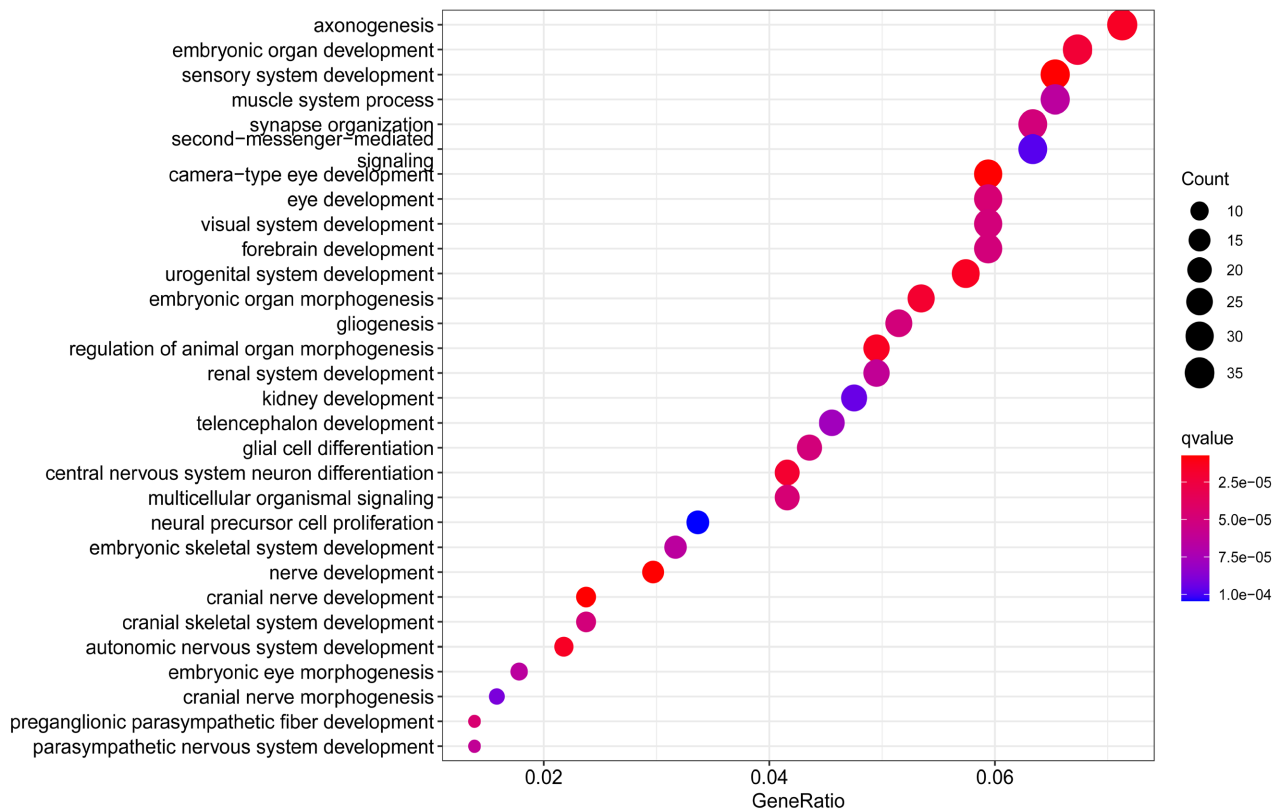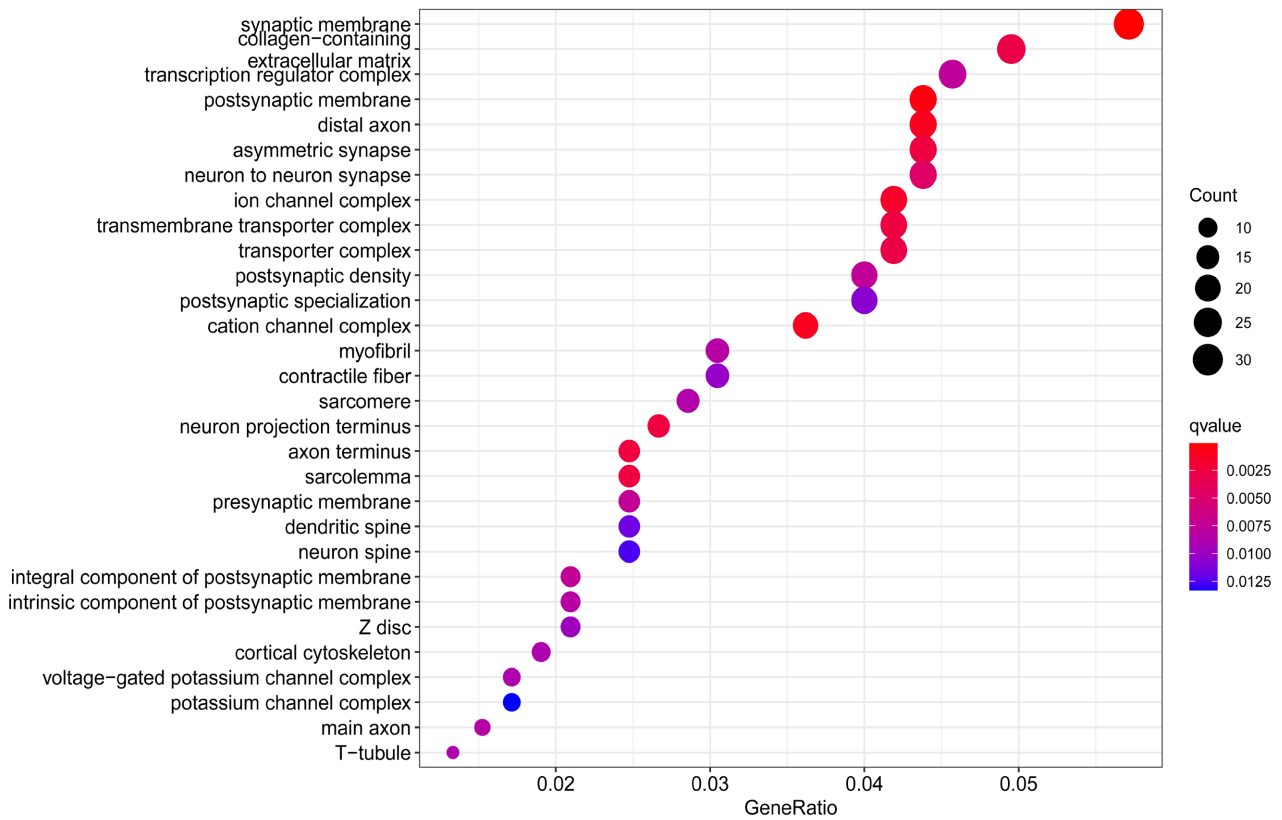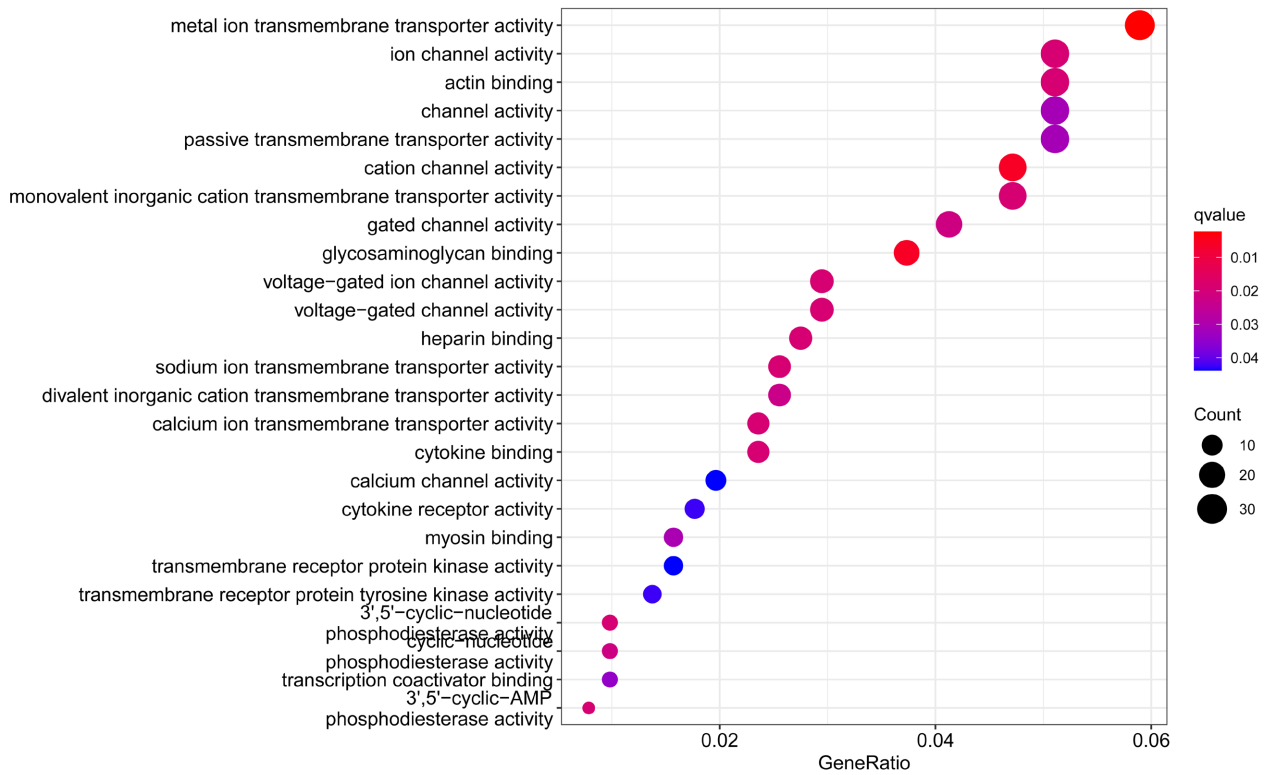
Figure 7. Functional enrichment analysis of target genes associated BC. (a) BP; (b) CC; (c) MF; (d) Bubble of KEGG signal pathway shown the counts of genes; (e) Circos of KEGG signal pathway shown the "pathway-gene" network.

Figure 8. Hub genes of PPI network. The darker the color, the bigger the degrees.

Survival curve (p=8.193e−03)

Survival curve (p=1.556e−02)

Survival curve (p=4.581e−03)
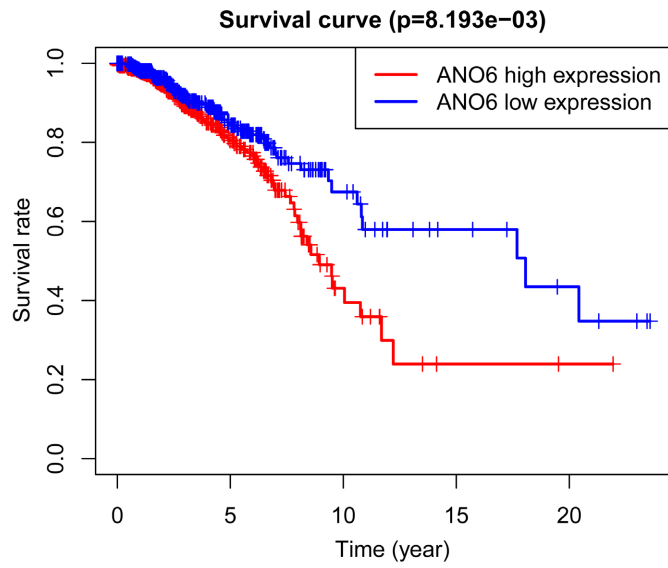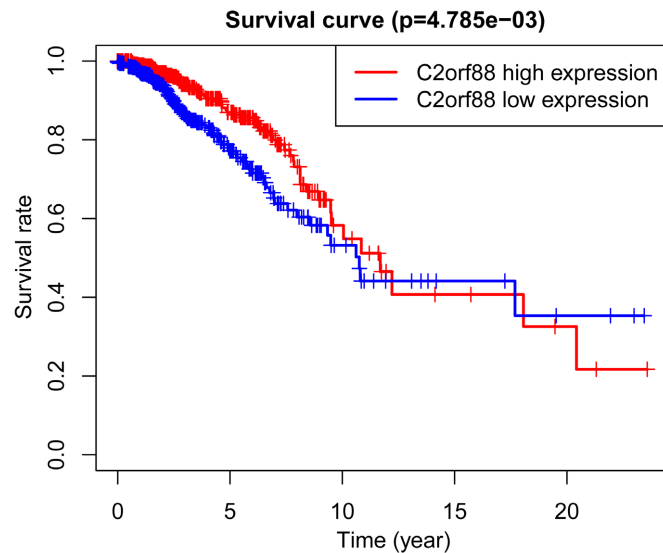
Figure 9. Target genes associated with over survival.

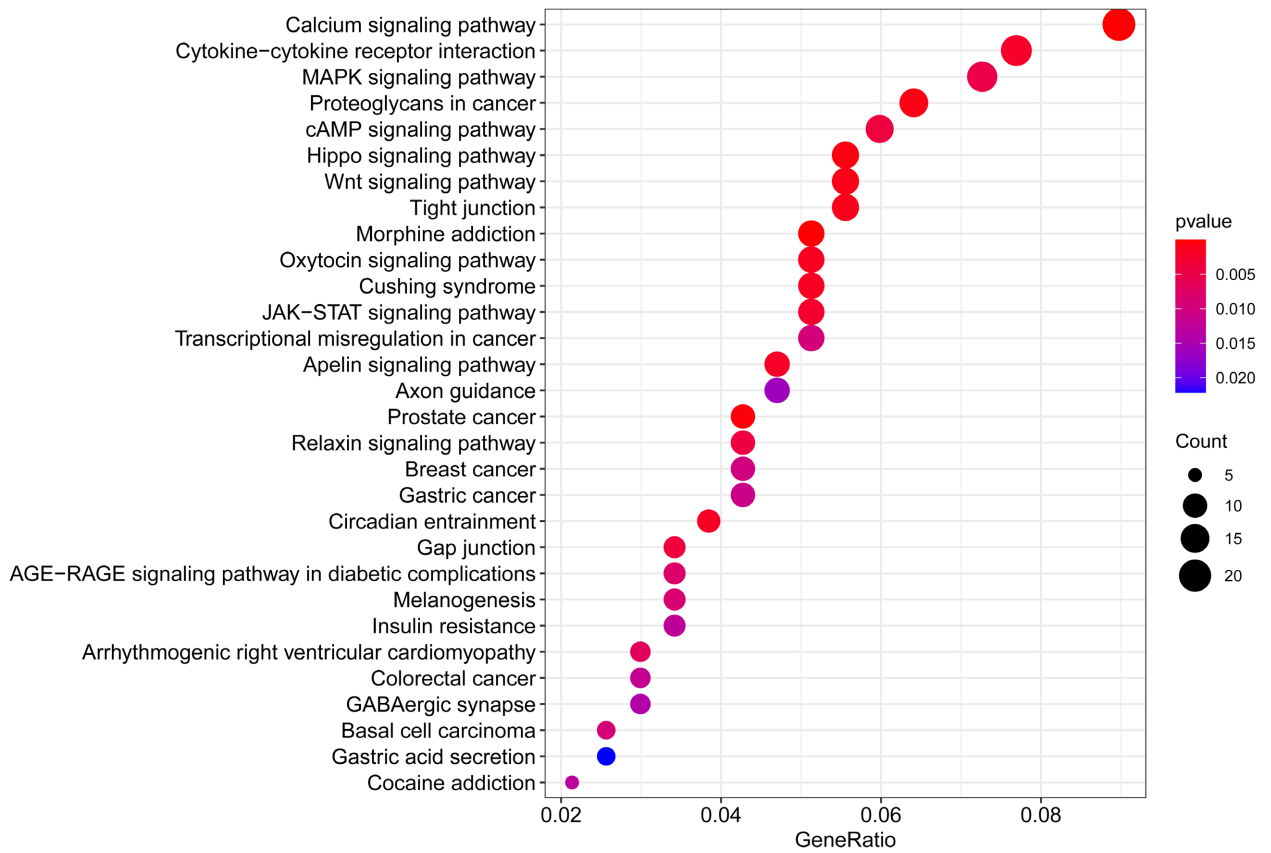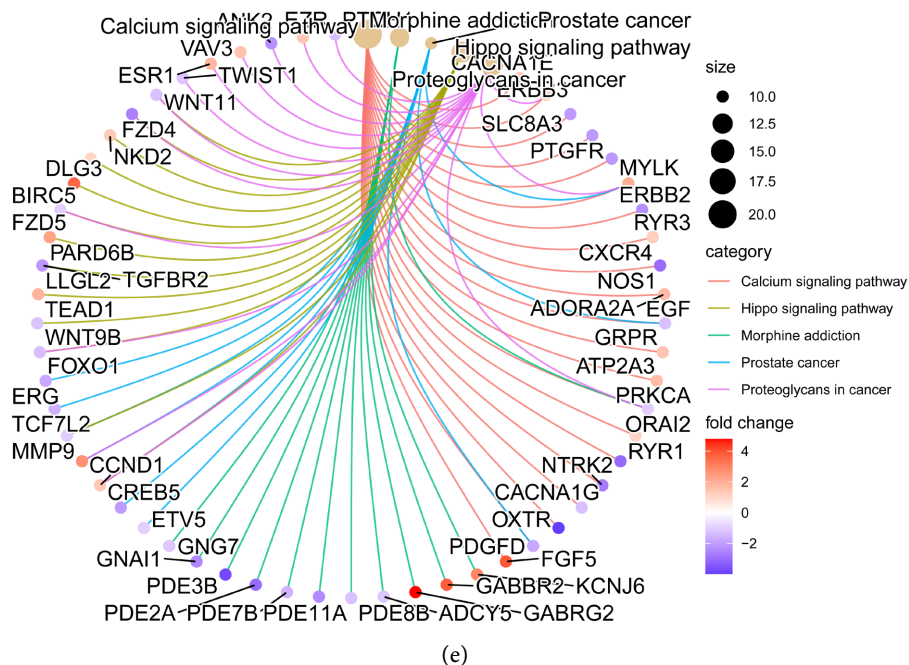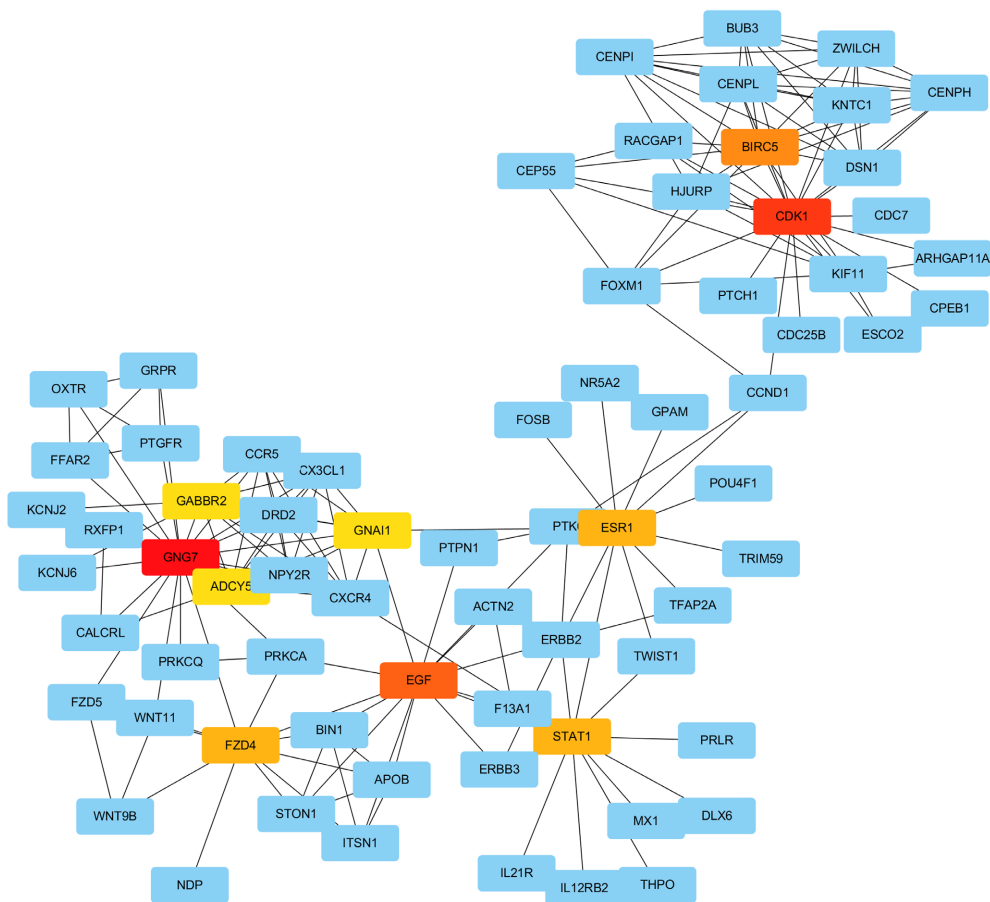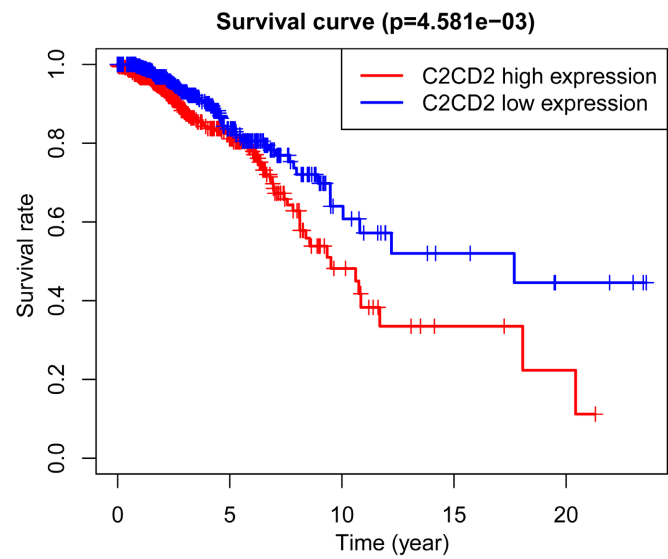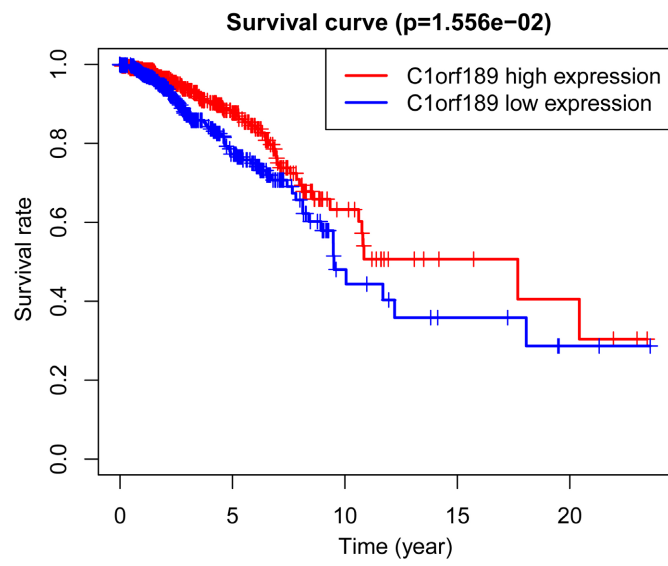divided into training and test groups. We found a nine miRNA-based signature model (hsa-miR-204-5p, hsa-miR-7706, hsa-miR-1247-3p, hsa-miR-20b-3p, hsa-miR-605-5p, hsa-miR-615-3p, hsa-miR-4652-5p, hsa-miR-9-5p, hsa-miR-2115-5p), using univariate Cox regression and stepwise multivariate Cox regression in training group. We also validated the utility of the nine miRNA signature model in the test group and the entire group. The median value grouping based on the risk score, and the results from K-M analysis indicated that the high-risk group indeed had poorer overall survival as compared to the low-risk groups across all three groups. We used the ROC curve to evaluate the predictive power of the nine-miRNA signature for overall survival in the three groups. Our results indicated that our model had a reliable performance. The results of univariate Cox regression and multivariate Cox regression revealed that the nine-miRNA signature was independent of overall survival for the conventional clinical factors in breast cancer patients. Several studies on the nine miRNAs show their involvement in the initiation, progression, and prognosis of various cancers. miR-204-5p is involved in the regulation of multiple human malignant tumor progression and is also a tumor suppressor for some human cancers such as malignant melanoma, hepatocellular, and esophageal squamous cell cancer [15] [16] [17]. Some studies show that expression levels of miRNA-7706 significantly decrease in patients with lymph node metastasis. Expression levels of miRNA-299 and miRNA-7706 in tumor tissues were significantly lower than those in adjacent tissues. These data may provide new parameters for clinical staging and prediction of lymph node metastasis and these two miRNAs could effectively inhibit the proliferation and invasion of cancer cells [18]. Previous studies show that miR-1247 is aberrantly expressed in several cancers [9] [19] [20] [21]. miR-125b and miR-20b display context-dependent properties in regulating the tumorigenicity of various cancers. They are key mediators of Wnt signaling activity. Together, these three

form a regulatory circuit [22] [23] [24] [25]. miR-605-5p promotes invasion and proliferation by targeting TNFAIP3 in NSCLC [26]. It also inhibits melanoma progression and glutamine catabolism by targeting GLS [27]. In a recent study, Lei *et al.* found that ectopic miR-615-3p expression in breast cancer cells promotes EMT-like traits. The tumor-promoting role of miR-615-3p is directly linked to its ability to target 3'-UTR of PICK1 [28]. MiR-4652, miR-9-5p, miR-2115 are expressed in several cancers. However, the current research mechanism of these three miRNA in tumors has not been reported yet. Thus, experiments in the future need to focus on the potential roles of these miRNA in BC.

To further understand the regulatory mechanism of the nine miRNA-based signatures in breast cancer, the target genes of nine miRNAs were predicted using three prediction databases. For breast cancer, we obtained the intersection of the target genes of these miRNAs and the differentially expressed genes from the TCGA database, and performed functional enrichment analyses for the overlapping genes. The GO annotations of the target genes were primarily associated with axonogenesis, embryonic organ development, sensory system development, synaptic membrane, collagen-containing, extracellular matrix transcription regulator complex, metal ion transmembrane transporter activity, ion channel activity, and actin-binding. KEGG analysis enriched signaling pathways included calcium signaling pathway, cytokine-cytokine receptor interaction, MAPK signaling pathway, proteoglycans in cancer and cAMP signaling pathway. Calcium signaling pathways are involved in proliferation, angiogenesis, invasion, immune evasion, disruption of cell death pathways, ECM remodelling, Epithelial-Mesenchymal Transition (EMT) and drug resistance [29]. Proteoglycans (PGs) and Glycosaminoglycans (GAGs) have important roles in cancer development, including regulation of metabolic patterns, angiogenesis and distant metastases, and treatment resistance [30]. Mitogen-Activated Protein Kinase (MAPK) pathways are among the most conserved signal transduction pathways with critical functions in different aspects of tumorigenesis such as tumor growth, apoptosis, angiogenesis, invasion, metastasis, and drug resistance. MAPK signaling plays critical functions in the various aspects of human malignancies [31]. These signaling pathways exert their effects on tumors to varying degrees. These five signaling pathways are only the tip of an iceberg of the target genes involved in signaling pathways. Thus, the construction of the miRNA prognosis model may have implycations in understanding the regulation of tumor signaling pathways.

To find the key regulatory nodes in the miRNA signature model breaast cancer, 10 hub genes (GNG7, CDK1, EGF, BIRC5, STAT1, ESR1, FZD4, DSN1, APOB, GNAI1) were identified with Cytoscape 3.7.1 and its plug-in (degree ranking of cytoHubba). Xu *et al.* show that GNG7 expression negatively correlates with patient grade and overall survival in clear cell renal cell carcinoma (ccRCC). Reduced expression of GNG7 was related to mTOR1, E2F, G2M, and MYC signaling pathways. In addition, ccRCC cells with lower GNG7 expression show an increased proportion of cells in of G2/M cell-cycle phase [32]. Currently, CDK1, EGF

receptor expression are enhanced in several cancer types including gastric cancer, esophageal adenocarcinoma, oral squamous cell carcinoma, ovarian cancer, liver cancer, colorectal cancer, and breast cancer. CDK1-dysregulation leads to robust tumor growth, chromosomal instability, and a higher proliferation rate in cancer cells. Overexpression of EGF receptors can augment cell growth by increased active ligand: receptor complex formation [33] [34] [35] [36]. Survivin (also called BIRC5) is a part of several cancer-cell signaling networks. The current cancer therapeutic strategies use survivin as a target [37]. Previous studies show that Stat3 and ESR1 have their respective mechanisms in breast cancer [38] [39]. FZD4 (frizzled 4) is a transmembrane protein belonging to the frizzled receptor family. It is involved in the signaling processes in a variety of malignant tumors and participates in cell proliferation, trans-differentiation, and carcinogenesis [40]. DSN1 is involved in cell cycle transition, cell growth. Depletion of the DSN1 gene induced G2/M arrest and significantly decreased migration and invasion of colorectal cancer and hepato-cellular carcinoma [41] [42] [43]. Heterotrimeric G proteins are crucial in multiple cellular signal transduction pathways [44]. The guanine nucleotide-binding protein G(i), α-1 subunit (GNAI1), belongs to the Gαi family [45]. These proteins primarily function as inhibitors of adenylyl cyclase. Gαi signaling can regulate cell proliferation and differentiation. It is taxol resistance in human ovarian carcinoma cells and hydrogen peroxide-induced apoptosis in human lung cancer cells [46] [47] [48] [49] [50]. The current mechanistic basis of APOB function in tumors has not yet been reported. Thus, the miRNA signature may affect the survival prognosis and progression in breast cancer patients by regulation through CDK1, EGF, BIRC5, STAT3, ESR1, FZD4. In addition, the Kaplan-Meier method showed that a positively associated expression of 53 genes with survival prognosis; however the high expression of ANO6, B4GALT3, C2CD2, CENPI, CHD5, CLEC3A, DLG3, FAM155B, FFAR2, KNSTRN, LRRC59, MRPS23, POP1, POU3F2, PTK6, RACGAP1, REEP1, SOX11, ZNF367 were associated with poorer overall survival outcomes.

5. Conclusion

In summary, we constructed a new predictive model based on miRNA signature through miRNA mature expression profiling for the prognosis of BC. By grouping, we also verified and evaluated the predictive ability of the model. The most important finding was its utility as an independent prognostic factor in BC. Additionally, the target genes of the model provide a new direction for investigations on the pathogenesis and diagnosis, and development of targeted inhibitors for the breast cancer. However, this is a bioinformatic study and extensive clinical validations are required to confirm our findings of a reliable predictor and therapeutic target signature for BC.

Acknowledgements

Data Availability

Data underlying this study are provided in Supplementary Materials (Supplemental 1-4) and are available on TCGA (https://gdc.cancer.gov/), miRTarBase (https://mirtarbase.cuhk.edu.cn), targetScan (http://www.targetscan.org) and miRDB (http://www.mirdb.org/). The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F. (2021) Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA*: *A Cancer Journal for Clinicians*, **71**, 209-249. https://doi.org/10.3322/caac.21660

[2] Vachhani, H., Shah, P.A., Robila, V. and Idowu, M.O. (2018) Invasive Carcinomas, Diagnosis and Management of Breast Tumors. 191-232. https://doi.org/10.1007/978-3-319-57726-5_9

[3] Luque-Bolivar, A., Perez-Mora, E., Villegas, V.E. and Rondón-Lagos, M. (2020) Resistance and Overcoming Resistance in Breast Cancer. *Breast Cancer*, **12**, 211-229. https://doi.org/10.2147/BCTT.S270799

[4] Schnitt, S.J. (2010) Classification and Prognosis of Invasive Breast Cancer: From Morphology to Molecular Taxonomy. *Modern Pathology*, **23**, S60-S64. https://doi.org/10.1038/modpathol.2010.33

[5] Ha, M. and Kim, V.N. (2014) Regulation of microRNA Biogenesis. *Nature Reviews Molecular Cell Biology*, **15**, 509-524. https://doi.org/10.1038/nrm3838

[6] Fernandez-Valdivia, R., Takeuchi, H., Samarghandi, A., Lopez, M., Leonardi, J., Haltiwanger, R.S. and Jafar-Nejad, H. (2011) Regulation of Mammalian Notch Signaling and Embryonic Development by the Protein O-Glucosyltransferase Rumi. *Development*, **138**, 1925-1934. https://doi.org/10.1242/dev.060020

[7] Calin, G.A. and Croce, C.M. (2006) MicroRNA Signatures in Human Cancers. *Nature Reviews Cancer*, **6**, 857-866. https://doi.org/10.1038/nrc1997

[8] Heagerty, P.J., Lumley, T. and Pepe, M.S. (2000) Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics*, **56**, 337-344. https://doi.org/10.1111/j.0006-341X.2000.00337.x

[9] Shi, S., Lu, Y., Qin, Y., Li, W., Cheng, H., Xu, Y., Xu, J., Long, J., Liu, L., Liu, C. and Yu, X. (2014) miR-1247 Is Correlated with Prognosis of Pancreatic Cancer and Inhibits Cell Proliferation by Targeting Neuropilins. *Current Molecular Medicine*, **14**, 316-327. https://doi.org/10.2174/1566524014666140228120014

[10] Siegel, R.L., Miller, K.D. and Jemal, A. (2020) Cancer Statistics, 2020. *CA*: *A Cancer Journal for Clinicians*, **70**, 7-30. https://doi.org/10.3322/caac.21590

[11] Desantis, C.E., Fedewa, S.A., Sauer, A.G., Kramer, J.L., Smith, R.A. and Jemal, A. (2016) Breast Cancer Statistics, 2015: Convergence of Incidence Rates between Black and White Women. *CA*: *A Cancer Journal for Clinicians*, **66**, 31-42. https://doi.org/10.3322/caac.21320

[12] Li, Z. and Kang, Y. (2016) Emerging Therapeutic Targets in Metastatic Progression: A Focus on Breast Cancer. *Pharmacology & Therapeutics*, **161**, 79-96. https://doi.org/10.1016/j.pharmthera.2016.03.003

[13] Temian, D.C., Pop, L.A., Irimie, A.I. and Berindan-Neagoe, I. (2018) The Epigenetics of Triple-Negative and Basal-Like Breast Cancer: Current Knowledge. *Journal of Breast Cancer*, **21**, 233-243. https://doi.org/10.4048/jbc.2018.21.e41

[14] Loh, H.Y., Norman, B.P., Lai, K.S., Rahman, N.M.A.N.A., Alitheen, N.B.M. and Osman, M.A. (2019) The Regulatory Role of MicroRNAs in Breast Cancer. *International Journal of Molecular Sciences*, **20**, E4940. https://doi.org/10.3390/ijms20194940

[15] Shen, Z., Chai, T., Luo, F., Liu, Z., Xu, H., Zhang, P., Kang, M. and Chen, S. (2020) Loss of miR-204-5p Promotes Tumor Proliferation, Migration, and Invasion through Targeting YWHAZ/PI3K/AKT Pathway in Esophageal Squamous Cell Carcinoma. *OncoTargets and Therapy*, **13**, 4679-4690. https://doi.org/10.2147/OTT.S243215

[16] Luan, W., Qian, Y., Ni, X., Bu, X.F., Xia, Y., Wang, J.L., Ruan, H.R., Ma, S.J. and Xu, B. (2017) miR-204-5p Acts as a Tumor Suppressor by Targeting Matrix Metalloproteinases-9 and B-Cell Lymphoma-2 in Malignant Melanoma. *OncoTargets and Therapy*, **10**, 1237-1246. https://doi.org/10.2147/OTT.S128819

[17] Chu, Y., Jiang, M., Du, F., Chen, D., Ye, T., Xu, B., Li, X., Wang, W., Qiu, Z., Liu, H., Nie, Y., Liang, J. and Fan, D. (2018) miR-204-5p Suppresses Hepatocellular Cancer Proliferation by Regulating Homeoprotein SIX1 Expression. *FEBS Open Bio*, **8**, 189-200. https://doi.org/10.1002/2211-5463.12363

[18] Wang, F., Dai, M., Chen, H., *et al.* (2018) Prognostic Value of Hsa-mir-299 and Hsa-mir-7706 in Hepatocellular Carcinoma. *Oncology Letters*, **16**, 815-820. https://doi.org/10.3892/ol.2018.8710

[19] Scaravilli, M., Porkka, K.P., Brofeldt, A., *et al.* (2015) MiR-1247-5p Is Overexpressed in Castration Resistant Prostate Cancer and Targets MYCBP2. *Prostate*, **75**, 798-805. https://doi.org/10.1002/pros.22961

[20] Zhao, F., Lv, J., Gan, H., *et al.* (2015) MiRNA Profile of Osteosarcoma with CD117 and Stro-1 Expression: miR-1247 Functions as an onco-miRNA by Targeting MAP3K9. *International Journal of Clinical and Experimental Pathology*, **8**, 1451-1458.

[21] Fang, T., Lv, H., Lv, G., *et al.* (2018) Tumor-Derived Exosomal miR-1247-3p Induces Cancer-Associated Fibroblast Activation to Foster Lung Metastasis of Liver Cancer. *Nature Communications*, **9**, 191. https://doi.org/10.1038/s41467-017-02583-0

[22] Li, D., Ilnytskyy, Y., Kovalchuk, A., *et al.* (2013) Crucial Role for Early Growth Response-1 in the Transcriptional Regulation of miR-20b in Breast Cancer. *Oncotarget*, **4**, 1373-1387. https://doi.org/10.18632/oncotarget.1165

[23] Cascio, S., D'andrea, A., Ferla, R., *et al.* (2010) miR-20b Modulates VEGF Expression by Targeting HIF-1 Alpha and STAT3 in MCF-7 Breast Cancer Cells. *Journal of Cellular Physiology*, **224**, 242-249. https://doi.org/10.1002/jcp.22126

[24] Huang, T., Alvarez, A.A., Pangeni, R.P., *et al.* (2016) A Regulatory Circuit of miR-125b/miR-20b and Wnt Signalling Controls Glioblastoma Phenotypes through FZD6-Modulated Pathways. *Nature Communications*, **7**, 12885. https://doi.org/10.1038/ncomms12885

[25] So, A.Y., Zhao, J.L. and Baltimore, D. (2013) The Yin and Yang of microRNAs: Leukemia and Immunity. *Immunological Reviews*, **253**, 129-145. https://doi.org/10.1111/imr.12043

[26] Liao, Y., Cao, L., Wang, F., *et al.* (2020) miR-605-5p Promotes Invasion and Proliferation by Targeting TNFAIP3 in Non-Small-Cell Lung Cancer. *Journal of Cellular Biochemistry*, **121**, 779-787. https://doi.org/10.1002/jcb.29323

[27] Lu, R., Zhang, X., Li, X., *et al.* (2020) Circ_0016418 Promotes Melanoma Development and Glutamine Catabolism by Regulating the miR-605-5p/GLS Axis. *International Journal of Clinical and Experimental Pathology*, **13**, 1791-1801.

[28] Lei, B., Wang, D., Zhang, M., *et al.* (2020) miR-615-3p Promotes the Epithelial-Mesenchymal Transition and Metastasis of Breast Cancer by Targeting PICK1/TGFBRI Axis. *Journal of Experimental & Clinical Cancer Research*, **39**, 71. https://doi.org/10.1186/s13046-020-01571-5

[29] Santoni, G., Morelli, M.B., Marinelli, O., *et al.* (2020) Calcium Signaling and the Regulation of Chemosensitivity in Cancer Cells: Role of the Transient Receptor Potential Channels. *Advances in Experimental Medicine and Biology*, **1131**, 505-517. https://doi.org/10.1007/978-3-030-12457-1_20

[30] Wei, J., Hu, M., Huang, K., *et al.* (2020) Roles of Proteoglycans and Glycosaminoglycans in Cancer Development and Progression. *International Journal of Molecular Sciences*, **21**, 5983. https://doi.org/10.3390/ijms21175983

[31] Asl, E.R., Amini, M., Najafi, S., *et al.* (2021) Interplay between MAPK/ERK Signaling Pathway and MicroRNAs: A Crucial Mechanism Regulating Cancer Cell Metabolism and Tumor Progression. *Life Sciences*, **278**, Article ID: 119499. https://doi.org/10.1016/j.lfs.2021.119499

[32] Xu, S., Zhang, H., Liu, T., *et al.* (2019) G Protein Gamma Subunit 7 Loss Contributes to Progression of Clear Cell Renal Cell Carcinoma. *Journal of Cellular Physiology*, **234**, 20002-20012. https://doi.org/10.1002/jcp.28597

[33] Santamaria, D., Barriere, C., Cerqueira, A., *et al.* (2007) Cdk1 Is Sufficient to Drive the Mammalian Cell Cycle. *Nature*, **448**, 811-815. https://doi.org/10.1038/nature06046

[34] Lim, S. and Kaldis, P. (2013) Cdks, Cyclins and CKIs: Roles beyond Cell Cycle Regulation. *Development*, **140**, 3079-3093. https://doi.org/10.1242/dev.091744

[35] Malumbres, M. and Barbacid, M. (2009) Cell Cycle, CDKs and Cancer: A Changing Paradigm. *Nature Reviews Cancer*, **9**, 153-166. https://doi.org/10.1038/nrc2602

[36] Thompson, D.M. and Gill, G.N. (1985) The EGF Receptor: Structure, Regulation and Potential Role in Malignancy. *Journal of Cancer Survivorship*, **4**, 767-788.

[37] Li, F., Aljahdali, I. and Ling, X. (2019) Cancer Therapeutics Using Survivin BIRC5 as a Target: What Can We Do after over Two Decades of Study? *Journal of Experimental & Clinical Cancer Research*, **38**, 368. https://doi.org/10.1186/s13046-019-1362-1

[38] Dustin, D., Gu, G. and Fuqua, S.W. (2019) ESR1 Mutations in Breast Cancer. *Cancer*, **125**, 3714-3728. https://doi.org/10.1002/cncr.32345

[39] Han, W., Carpenter, R.L., Cao, X., *et al.* (2013) STAT1 Gene Expression Is Enhanced by Nuclear EGFR and HER2 via Cooperation with STAT3. *Molecular Carcinogenesis*, **52**, 959-969. https://doi.org/10.1002/mc.21936

[40] Wang, Y., Chang, H., Rattner, A., *et al.* (2016) Frizzled Receptors in Development and Disease. *Current Topics in Developmental Biology*, **117**, 113-139. https://doi.org/10.1016/bs.ctdb.2015.11.028

[41] Wei, L., Diao, Y., Qi, J., *et al.* (2013) Effect of Change in Spindle Structure on Proliferation Inhibition of Osteosarcoma Cells and Osteoblast under Simulated Microgravity during Incubation in Rotating Bioreactor. *PLoS ONE*, **8**, e76710. https://doi.org/10.1371/journal.pone.0076710

[42] Sun, C., Huang, S., Ju, W., *et al.* (2018) Elevated DSN1 Expression Is Associated with Poor Survival in Patients with Hepatocellular Carcinoma. *Human Pathology*, **81**, 113-120. https://doi.org/10.1016/j.humpath.2018.06.032

[43] Kline, S.L., Cheeseman, I.M., Hori, T., *et al.* (2006) The Human Mis12 Complex Is

Required for Kinetochore Assembly and Proper Chromosome Segregation. *Journal of Cell Biology*, **173**, 9-17. https://doi.org/10.1083/jcb.200509158

[44] Gilman, A.G. (1987) G Proteins: Transducers of Receptor-Generated Signals. *Annual Review of Biochemistry*, **56**, 615-649. https://doi.org/10.1146/annurev.bi.56.070187.003151

[45] Simon, M.I., Strathmann, M.P. and Gautam, N. (1991) Diversity of G Proteins in Signal Transduction. *Science*, **252**, 802-808. https://doi.org/10.1126/science.1902986

[46] Parekh, H.K., Adikari, M. and Vennapusa, B. (2006) Differential Partitioning of Gal-phai1 with the Cellular Microtubules: A Possible Mechanism of Development of Taxol Resistance in Human Ovarian Carcinoma Cells. *Journal of Molecular Signaling*, **1**, 3. https://doi.org/10.1186/1750-2187-1-3

[47] Ram, P.T. and Iyengar, R. (2001) G Protein Coupled Receptor Signaling through the Src and Stat3 Pathway: Role in Proliferation and Transformation. *Oncogene*, **20**, 1601-1606. https://doi.org/10.1038/sj.onc.1204186

[48] Taussig, R., Iñiguez-Lluhi, J.A. and Gilman, A.G. (1993) Inhibition of Adenylyl Cyclase by Gi Alpha. *Science*, **261**, 218-221. https://doi.org/10.1126/science.8327893

[49] Seo, M., Nam, H.J., Kim, S.Y., *et al.* (2009) Inhibitory Heterotrimeric GTP-Binding Proteins Inhibit Hydrogen Peroxide-Induced Apoptosis by Up-Regulation of Bcl-2 via NF-kappaB in H1299 Human Lung Cancer Cells. *Biochemical and Biophysical Research Communications*, **381**, 153-158. https://doi.org/10.1016/j.bbrc.2009.01.188

[50] Embry, A.C., Glick, J.L., Linder, M.E., *et al.* (2004) Reciprocal Signaling between the Transcriptional Co-Factor Eya2 and Specific Members of the Galphai Family. *Molecular Pharmacology*, **66**, 1325-1331. https://doi.org/10.1124/mol.104.004093