Iranian Journal of Basic Medical Sciences
ijbms.mums.ac.ir

Differential expression of the lncRNA MIR17HG and miR-18a/miR-20a as candidate ncRNA signatures in stage II/III breast cancer and bioinformatic analysis of ncRNA–mRNA regulatory axes

Reyhane Alidousti Shahraki [1#], Sara Tavallaei [2#], Afsaneh Malekzadeh Shafaroudi [3], Saba Abbassioun [1], Nahid Nafissi [4], Seyed-Javad Mowla [1*]

[1] Molecular Genetics Department, Biological Sciences Faculty, Tarbiat Modares University, Tehran, Iran
[2] Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, Isfahan, Iran
[3] Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad, Iran
[4] Rasoul Akram Hospital Clinical Research and Development Center (RCRDC), Iran University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Article type:
Original

Article history:
Received: Apr 21, 2026
Accepted: Jul 21, 2026

Keywords:
Breast cancer
Endocytosis
Hippo signaling pathway
Long non-coding RNA
MIR17HG
MicroRNAs
Noncoding RNA
RNA degradation

ABSTRACT

Objective(s): Breast cancer has remained a significant cause of mortality associated with cancer, necessitating a deeper molecular understanding to improve diagnostics and targeted therapies. Noncoding RNAs (ncRNAs) and their regulatory ncRNA–mRNA axes have recently emerged as important modulators of cancer progression and potential tools for earlier detection and more precise stratification of breast cancer. Among them, MIR17HG variants and the precise regulatory axes of miR-18a/miR-20a, and their mechanisms in stage II/III BC, remain elusive.

Materials and Methods: This study addresses this critical gap by integrating data from TCGA miRNAseq and RNAseq datasets with experimental qRT-PCR validation on patient tissues. We investigate MIR17HG transcript variants and their associated miRNAs, especially miR-18a and miR-20a, as candidate stage-associated molecules in stage II/III breast cancer. Co-expressed mRNAs were identified through gene co-expression networks (WGCNA).

Results: Potential MIR17HG–miR-18a/20a-mRNA regulatory axes were computationally predicted and functionally annotated to elucidate their involvement in cancer-related signaling pathways. Our findings showed that both MIR17HG splice variants were down-regulated, while their associated miRNAs (miR-18a-5p and miR-20a-5p) were overexpressed. Based on *in silico* prediction and functional annotation, we propose three candidate MIR17HG-centred axes: MIR17HG–miR-20a–BTG3 (linked to RNA degradation), MIR17HG–miR-18a/miR-20a–PARD6B (Hippo signaling and cell polarity), and MIR17HG–miR-20a–ZNF367 (endocytosis-related processes).

Conclusion: Together, these data support MIR17HG and its associated miRNAs as candidate stage II/III breast cancer-associated ncRNA signatures and provide a framework for future functional and clinical validation studies.

▶Please cite this article as:
Alidousti Shahraki R, Tavallaei S, Malekzadeh Shafaroudi A, Abbassioun S, Nafissi N, Mowla SJ. Differential expression of the lncRNA MIR17HG and miR-18a/miR-20a as candidate ncRNA signatures in stage II/III breast cancer and bioinformatic analysis of ncRNA–mRNA regulatory axes. Iran J Basic Med Sci 2026; 29:

Introduction

Breast cancer (BC), the predominant malignancy affecting women worldwide, has been a major contributor to cancer-related mortality over the past two decades (1, 2). The urgent need for early cancer detection and improved therapeutic targeting has driven extensive research to identify novel molecular biomarkers and develop more personalized treatment approaches (3). Among these, regulatory biomolecules and molecular interaction networks- particularly non-coding RNA-mediated gene regulation involving lncRNA–miRNA–mRNA axes- have gained considerable attention in driving oncogenesis and tumor suppression (4–6). Long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) regulate gene expression at transcriptional and post-transcriptional levels (7). Their tissue-specific expression patterns and stability in bodily fluids have established miRNAs as promising noninvasive biomarkers for cancer diagnosis, prognosis, and therapeutic monitoring (8–10). The MIR17-92 polycistronic transcript, located within the third intron of C13orf25, spans over 800 nucleotides and comprises six evolutionarily conserved miRNAs (miR-17, miR-18a, miR-19a, miR-20a-5p, miR-19b-1, and miR-92-1). Two primary transcript variants have been identified: variant 1 (uc010tie), a longer two-exon transcript, and variant 2 (uc001vlq), a shorter four-exon isoform (11, 12). However, their regulatory

*Corresponding author: Seyed Javad Mowla. Molecular Genetics Department, Biological Sciences Faculty, Tarbiat Modares University, Tehran, Iran. Tel: +98-21-82883464, Fax: +98-21-82884717, Email: sjmowla@modares.ac.ir
#These authors contributed eqully to this work

capacities and potential stage-specific clinical relevance remain poorly characterized. The MIR17-92 cluster is overexpressed in several malignancies, including gastric and lung cancers(13, 14), and both its host gene (MIR17HG) and encoded miRNAs exhibit dual oncogenic and tumor-suppressive roles in cancer progression. Nevertheless, their specific functions in BC remain insufficiently defined. Differences in the expression patterns of its variants or among populations may reflect distinct roles, serving as novel candidate stage-associated molecules at the isoform level and promising biomarkers for future translational studies, including functional validation and personalized therapeutic strategies in larger patient cohorts.

While several studies have focused on MIR17HG and miR-18a/20a in BC, no previous study has simultaneously (i) characterized the isoform-specific expression of MIR17HG transcript variants, (ii) integrated miRNA-seq and total RNA-seq data with WGCNA to define MIR17HG-associated modules in stage II/III BC, and (iii) proposed MIR17HG–miR-18a/20a–BTG3/PARD6B/ZNF367 regulatory axes linked to RNA degradation, Hippo signaling, and endocytosis. Here, we address this gap by performing a comprehensive analysis of TCGA miRNA-seq and bulk RNA-seq datasets utilizing differential gene expression (DGE) analysis to elucidate expression patterns and quantify levels of miR-18a-5p, miR-20a-5p, and MIR17HG. Subsequently, we validated these expression patterns experimentally through qPCR on stage II–III BC tissues and adjacent normal tissues. To identify co-expressed mRNAs with MIR17HG, co-expression network analysis was performed. Based on these data, we predicted MIR17HG-miRNA-mRNA regulatory axes involving these markers to provide new insights into the molecular regulation of BC and identify candidate stage-associated molecules. Our findings unravel the transcriptional dysregulation of MIR17HG and its encoded miRNAs in BC stage II-III, validate MIR17HG variants and miR-18a/20a expression in human tissue samples, and predict potential regulatory axes involving these markers. The diagram illustrating the process of this study is presented in Figure 1.

Materials and Methods

MiRNA-seq data acquisition and preprocessing

MiRNA-seq datasets for stage II–III BC and matched normal samples were obtained from The Cancer Genome Atlas (TCGA) via the GDC portal. Female patients with primary BC were selected, and datasets were merged for analysis in R (v4.4.1). Preprocessing included low-count gene filtering (genes with<10 reads across all samples removed), variance-stabilizing transformation (VST) for normalization, and outlier removal via hierarchical clustering.

Differential miRNA expression analysis

Differentially expressed miRNAs (DE-miRNAs) between tumor and normal tissues were identified using the DESeq2 package (v1.46.0). Raw counts were modeled using a negative binomial distribution, with dispersions estimated logarithmically. Wald tests were applied to assess significance, followed by Benjamini-Hochberg correction for multiple testing. DE-miRNAs were filtered using thresholds of |log2(fold change) |>0.58 and *P*-value<0.05. Additionally, adjusted *P*-values (FDR) were calculated. All final candidate genes were checked for their FDR values and were significant (adjusted *P*<0.05), demonstrating the robustness of the findings.

Bulk RNA-seq data processing

Total RNA-seq datasets for stage II–III BC tumors (TCGA) and normal mammary tissues (GTEx) were merged and processed in R. Preprocessing steps mirrored miRNA-seq analysis, including low-count filtering, VST normalization, and outlier detection via principal component analysis (PCA) and hierarchical clustering.

Differential mRNA expression analysis

Differentially expressed genes (DEGs) between tumor and normal tissues were identified using DESeq2 (as in Section 2), with identical significance thresholds (|log2FC| > 0.58, *P*<0.05). Additionally, adjusted p-values (FDR) were calculated. All final candidate genes that passed multi-layer analyses—including WGCNA—were checked for their FDR values and were significant (adjusted *P*<0.05), demonstrating the robustness of the findings.

Co-expression network construction and module-trait analysis

A signed hybrid co-expression network was built using the WGCNA package (v1.73) (15). First, pairwise gene correlations were calculated using the Pearson method. Soft thresholding power was selected to achieve a scale-free topology (R2>0.8). After thresholding gene correlations to the selected power, the gene-gene adjacency matrix and the topological overlap matrix (TOM) were created. Using the hclust function, the hierarchical clustering of modules was examined, and close modules were merged (modules were identified via dynamic tree cutting with deepSplit =

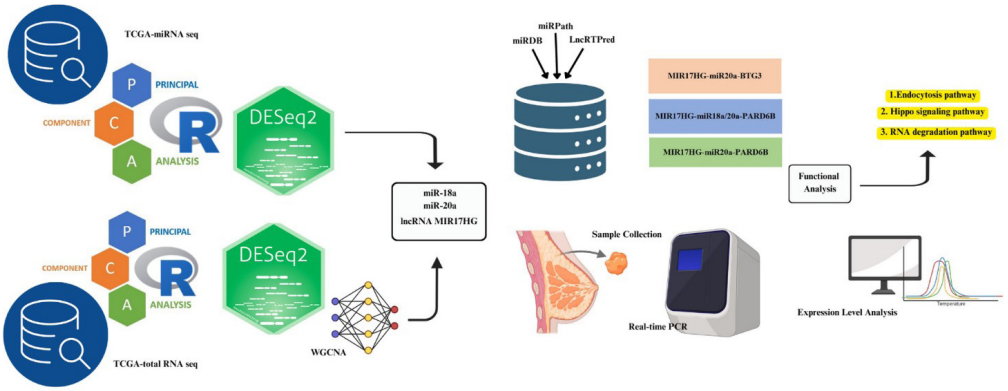

Figure 1. A schematic overview of the Study design and Analytical Framework (Created using Microsoft PowerPoint)

2 and minClusterSize=30, and the merging cut-off was 0.1). Module-trait relationships were calculated to assess correlations with BC status. Module eigengenes (MEs) and membership (MM) values were used to prioritize modules for downstream analysis.

MiRNA-lncRNA-mRNA axes identification

Target mRNAs for miR-18a-5p and miR-20a-5p were predicted using the miRDB database (16) (applying the MirTarget algorithm (17) and target score >80). Common mRNA targets of miR-18a-5p and miR-20a-5p were identified and validated by retaining only those present in significant modules based on Module-Trait analysis. LncRTPred tool (18), integrating Decision Tree, KNN, Random Forest, and LightGBM machine learning algorithms, assessed binding probabilities between MIR17HG and target mRNAs based on their sequences. Axes were retained if binding probability exceeded 80% across all models.

Functional enrichment analysis

Pathway enrichment for miRNA targets was performed using the DIANA-miRPath v4.0 web platform (19). KEGG pathways with $P<0.05$ were retained, and results were filtered to include only pathways involving mRNAs in the identified axes.

Clinical sample collection

To validate our computational findings, an independent cohort consisting of 36 matched pairs of tumor and tumor margin tissues from stage II–III patients undergoing surgery at Khatam-ol-Anbia and Rasule-Akram Hospitals (Tehran, Iran) was collected from January 2020 to July 2020 and utilized for experimental validation. Patients' written informed consent was collected before participation. This study was conducted with the approval of the Research Ethics Committee of Ferdowsi University of Mashhad (Code: IR.UM.REC.1399.104). Females with histopathologically confirmed stage II/III BC were enrolled and None of the patients had been treated with preoperative radiotherapy, chemotherapy, or other relevant modalities. Detailed clinicopathological characteristics of the patient tissue samples are provided in Table 1. For long-term storage, tissues were snap-frozen in liquid nitrogen and stored at −80 °C until RNA extraction.

Table 1. Clinicopathological characteristics of the patient tissue samples

Characteristic	Value
Number of patient tissue samples	36 paired
Sex	Female
Age (years)	52.09 ± 11.37
Tumor Location	Left / Right
Histological Type	IDC / DCIS / LVI
Tumor Size	1 -10 (cm)
Clinical Stage	
Stage II	20 (55.6%)
Stage III	16 (44.4%)
Molecular Subtype	n (%)
Luminal A/B	22 (61.1 %)
HER2-enriched	5 (13.8 %)
TNBC	9 (25 %)

RNA Extraction and qRT-PCR

Total RNA was isolated using RiboEx reagent (GeneAll, South Korea, Cat. No.301-001). RNA purity (A260/A280 ratio >1.8) was verified via NanoDrop spectrophotometry, and integrity was confirmed by agarose gel electrophoresis. DNA contamination was removed using DNase I (NEB, UK, Cat. No. EN0521). For miRNA analysis, the poly-A method was employed for cDNA synthesis. In the initial step, total RNA was polyadenylated using a Poly(A) Polymerase Tailing Kit (NEB, England, Cat. No. M0276S). Following the manufacturer's instructions, we incubated 1 µg of total RNA with 1 µl of 10× reaction buffer, 1 µl of 10 mM ATP, and 1 unit of poly(A) polymerase at 37 °C for 45 min, followed by enzyme inactivation through incubation of the sample at 65 °C for 5 min. Complementary DNA (cDNA) was synthesized using a Fermentase cDNA Synthesis Kit (Cat. No. K1622) and specific miRNA sequences were synthesized with oligo dT primers (RevertAid M-MuLV RT, Thermo Fisher Scientific, Cat. No. SO131). In brief, the cDNA synthesis was conducted in a total volume of 20 µl, which included 10 µl of the polyadenylation reaction product, 1 µl of Anchored Oligo(dT), 1 µl of RiboLock RNase Inhibitor (20 U/µl), 4 µl of 5X Reaction Buffer, 2 µl of 10 mM dNTP Mix, and 1 µl of RevertAid M-MuLV RT (200 U/µl) (Thermo Fisher Scientific, UK, Cat. No. EP0441). The reaction was incubated at 42 °C for 70 min and then terminated by heating at 70 °C for 5 min, following the manufacturer's recommended protocol.

Primers for miRNAs (poly-A method) and MIR17HG (exon-exon junction) were designed using Primer3 and OligoAnalyzer (IDT) to ensure absence of secondary structures. All primers (Table 2) were synthesized by TAG Copenhagen (Denmark).

Table 2. Primer sequences used for quantitative real-time PCR (qRT-PCR) analysis (5′–3′)

Gene	Sequence	Type of primer	Primer length
miR-18a-5p	CGTAAGGTGCATCTAGTGCAGA	Forward	22
mir-20a-5p	GTGCGTAAAGTGCTTATAGTGC	Forward	22
17HG-v1-F	GGCTACTCCTCCTGTCATACA	Forward	22
17HG-v1-R	CCGCACACCTACTTTTACTCTC	Reverse	22
17HG-v2-F	ATTTCCTCTAAAAGGCAGACCTGT	Forward	24
17HG-v2-R	ATGTGACACTGAGTTGTTCTCC	Reverse	22
U6	GAACGATACAGAGAAGATT	Forward	19

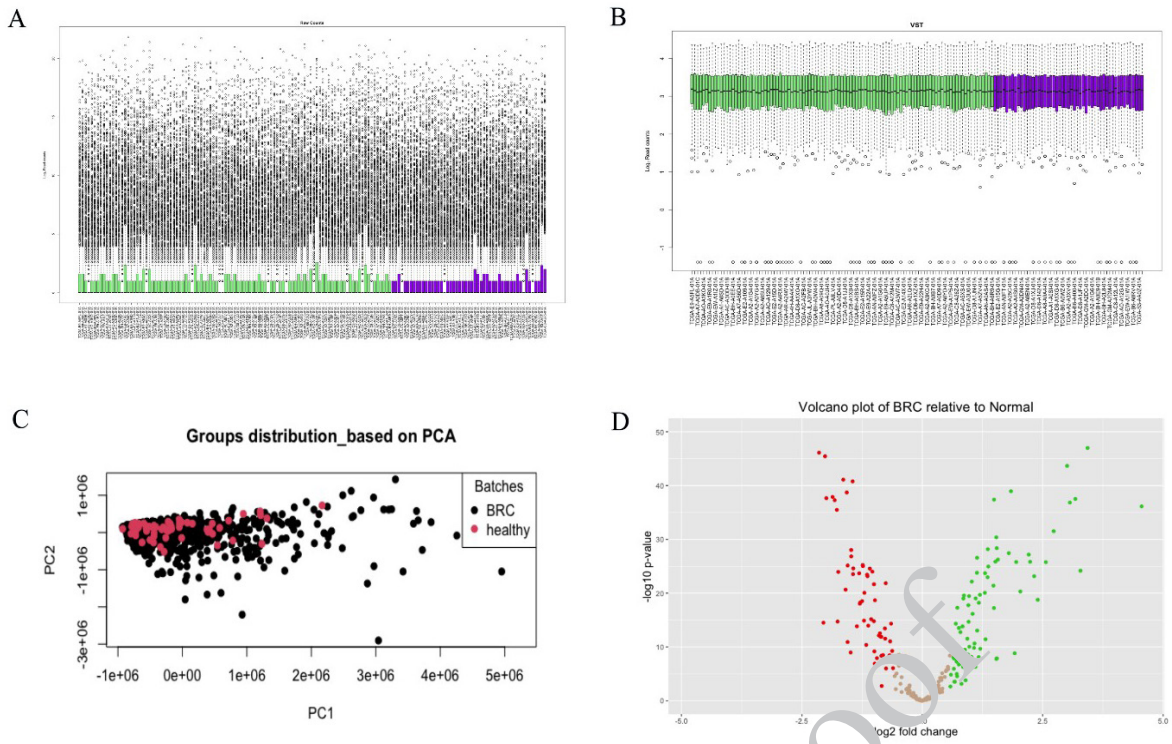

Figure 2. MiRNA-seq data Analysis and differential gene expression (DGE)
(A) The raw miRNA-seq expression matrix, which encompasses both tumor and normal cases, displays a considerable number of low-count miRNAs, as depicted in the boxplot. This observation underscores the necessity for gene filtering and normalization. Green bars correspond to stages II-III of primary breast cancer, while purple bars represent normal cases. (B) As illustrated by the expression median in the boxplot, low-count miRNAs were subsequently removed, and the data underwent appropriate normalization. (C) Principal component analysis (PCA) revealed several outlier samples in both the tumor and normal groups, which were removed in the next step. (D) In the volcano plot of filtered differentially expressed genes (DEGs), miRNAs exhibiting significant changes in expression are illustrated in green for up-regulation and red for down-regulation. In contrast, those with non-significant changes are represented in a creamy color. Specifically, up-regulated miRNAs are defined by a Log2FC > 0.58 and a P-value<0.05, signifying an expression increase of at least 1.5-fold compared to the normal state. In contrast, down-regulated miRNAs are characterized by a Log2FC<0.58 and a P-value<0.05, indicating a decrease in expression by at least 1.5-fold compared to the normal condition (created using R Software).

Expression of miRNAs (U6 reference) and MIR17HG (B2M reference) was quantified using StepOne Plus (Applied Biosystems) and RealQ Plus 2x Master Mix (Ampliqon Co, Denmark, Cat. No. A25742). Cycling conditions: an initial denaturation step at 95 °C (5 min), 40 cycles of 95 °C (25 sec), 59 °C (25 sec) for extension, and gene-specific annealing at 72 °C (25 sec). All reactions were performed in duplicate for each sample to minimize experimental error, and Melt curve analysis confirmed specificity. qRT-PCR raw data were analyzed using GraphPad Prism software (v.10). Data were expressed as mean±SD. Differences between groups were evaluated using Student's t-test, and a *P*-value>0.05 was considered statistically significant. Relative expression was calculated via 2−ΔΔCt.

Results

MiR-18a-5p and miR-20a-5p are overexpressed in stage II–III breast cancer

Analysis of miRNA-seq data from 728 stage II–III BC cases and 60 normal controls identified 265 differentially expressed miRNAs (DE-miRNAs). Among these, miR-18a-5p (log2FC=2.29, *P*<0.05) and miR-20a-5p (log2FC=1.69, *P*<0.05) were significantly overexpressed in tumor tissues, suggesting their potential roles in BC progression (Figure 2).

MIR17HG Is down-regulated and co-expressed with tumor-suppressive markers

Total RNA-seq analysis of 566 stage II–III BC tumors and 168 adjacent normal tissues revealed significant down-regulation of the lncRNA MIR17HG (log2FC=−1.63, *P*<0.05). Co-expression network analysis placed MIR17HG within the turquoise module, which exhibited the strongest negative correlation with BC status (*cor*=−0.95, *P*<0.05) (Figure 3G). This module's enrichment for tumor-suppressor genes suggests MIR17HG's potential regulatory role, potentially mediated through its encoded miRNAs, miR-18a-5p and miR-20a-5p.

BTG3, PARD6B, and ZNF367 are identified as mRNAs involved in the MIR17HG-miRNA-mRNA regulatory axes

Using the miRDB database, seven high-confidence mRNA targets (target score>80) shared by miR-18a-5p and miR-20a-5p were identified: *BTG3*, *ZBTB47*, *SUCO*, *ZBTB20*, *RORA*, *PARD6B*, and *ZNF367*. Binding probability analysis (>80% across four machine learning models in the LncRTPred web tool(16)) prioritized *BTG3*, *PARD6B*, and *ZNF367* as core mRNAs interacting with MIR17HG (Figures 4-6). These findings suggest three functional axes:
miR-20a-5p–*BTG3*
miR-18a-5p/miR-20a-5p–*PARD6B*
miR-20a-5p–*ZNF367*

Functional enrichment implicates key pathways in bc progression

DIANA-miRPath analysis linked these axes to three KEGG pathways (*P*<0.05):
RNA degradation pathway: Mediated by the miR-20a-5p–*BTG3* axis.

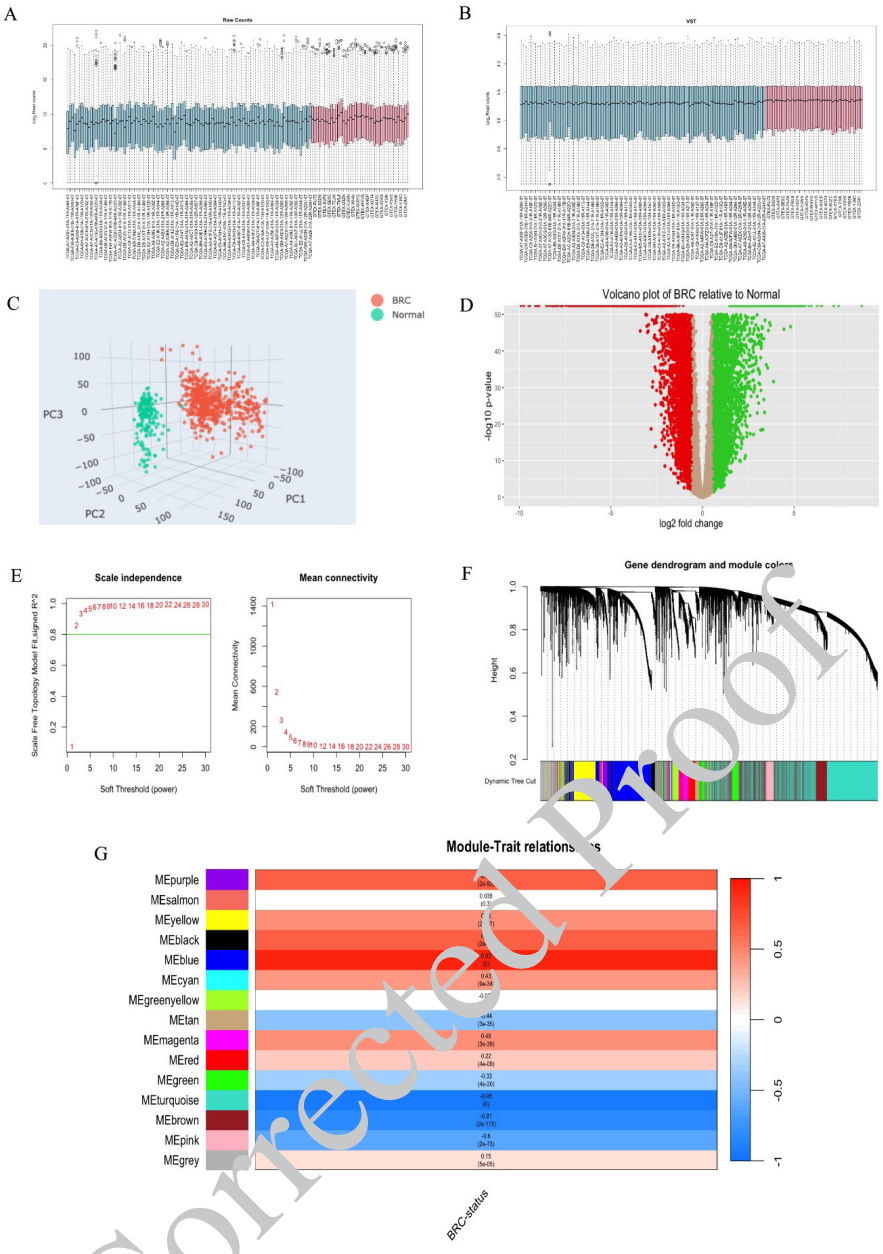

Figure 3. DGE, co-expression network construction, and module-trait analysis on bulk total RNA-seq data

(A, B) The median expression status before and after normalization is presented. Observation A highlights the importance of normalization. The blue bars represent stages II-III of primary breast cancer, while the pink bars correspond to normal cases. Part B boxplot demonstrates that the data underwent appropriate normalization. (C) The 3D principal component analysis (PCA) revealed several outlier samples within both the tumor and normal groups, which were removed in the next step. (D) In the volcano plot of filtered differentially expressed genes (DEGs), genes exhibiting significant changes in expression are illustrated in green for up-regulation and red for down-regulation. In contrast, those with non-significant changes are represented in a creamy color. Specifically, up-regulated genes are defined by a Log2FC>0.58 and a *P*-value<0.05, signifying an expression increase of at least 1.5-fold compared to the normal state. In contrast, down-regulated genes are characterized by a Log2FC<0.58 and a *P*-value<0.05, indicating a decrease in expression by at least 1.5-fold compared to the normal condition. (E) The analysis of the scale-free index and mean connectivity for various soft-threshold powers revealed that a power of 6 provided the best fit for the scale-free index and aligned with the data status. (F) A module dendrogram was created to visualize clustered genes after merging closely related modules, based on their dissimilarity measurements (1-TOM). (G) Module-trait analysis plots indicate that the most related modules for stages II and III of primary breast cancer are turquoise (correlation -0.95) and blue (correlation +0.93) (created using R Software).

Hippo signaling pathway: Associated with the miR-18a-5p/miR-20a-5p–*PARD6B* axis.

Endocytosis pathway: Linked to the miR-20a-5p–*PARD6B* axis.

While MIR17HG itself lacks direct pathway annotations, its co-expression with these mRNAs and miRNAs suggests indirect regulatory roles in these processes.

Experimental validation of MIR17HG and miRNA expression

qRT-PCR was performed on 36 paired stage II–III BC and adjacent normal tissue samples to validate the *in silico* expression pattern. Both MIR17HG transcript variants showed significant down-regulation in tumors: Variant 1 exhibited a 3.2-fold decrease (*P*=0.004) and Variant 2 showed a 2.8-fold decrease (*P*=0.007) (Figure 7A, B). Conversely, miR-18a-5p (4.1-fold, *P*=0.002) and miR-20a-5p (3.6-fold, *P*=0.003) were significantly overexpressed in tumor tissues compared to adjacent normal tissues (Figure 8A, B).

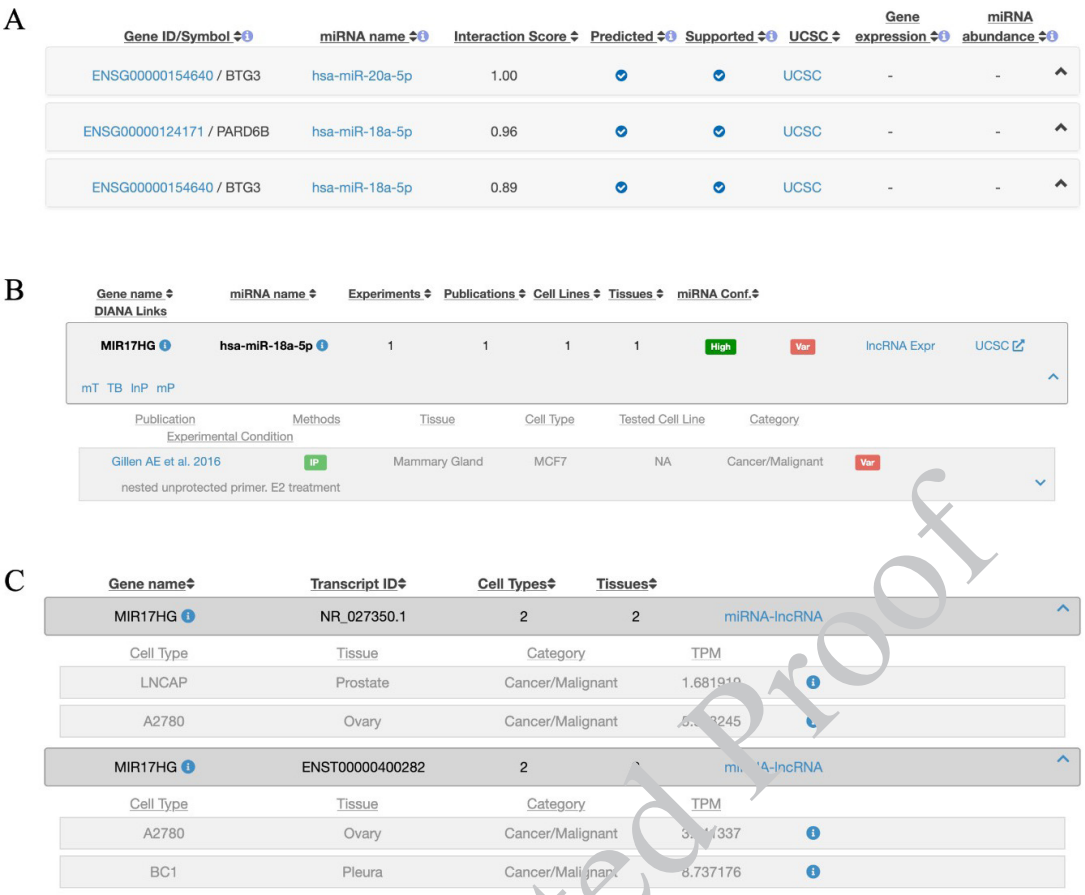

Figure 4. Assessment of investigated miRNA-lncRNA-mRNA interactions based on some databases
(A) Evaluation of miRNA-mRNA interactions using the microT web server (20): The miRNA-mRNA relationships identified in our analysis were evaluated with microT, which provides results from other studies. (B) Evaluation of miRNA-lncRNA interactions using the LncBase web platform (21, 22). This platform indicates whether any studies have identified a miRNA target for the specified lncRNA. It indicates that miR-18a has been reported as a target for MIR17HG in a study related to cancerous mammary glands, showing a strong correlation. (C) According to LncBase, MIR17HG has been implicated in cancers of the prostate, ovary, and pleura (Created using microT web server, LncBase web platform, Microsoft PowerPoint).

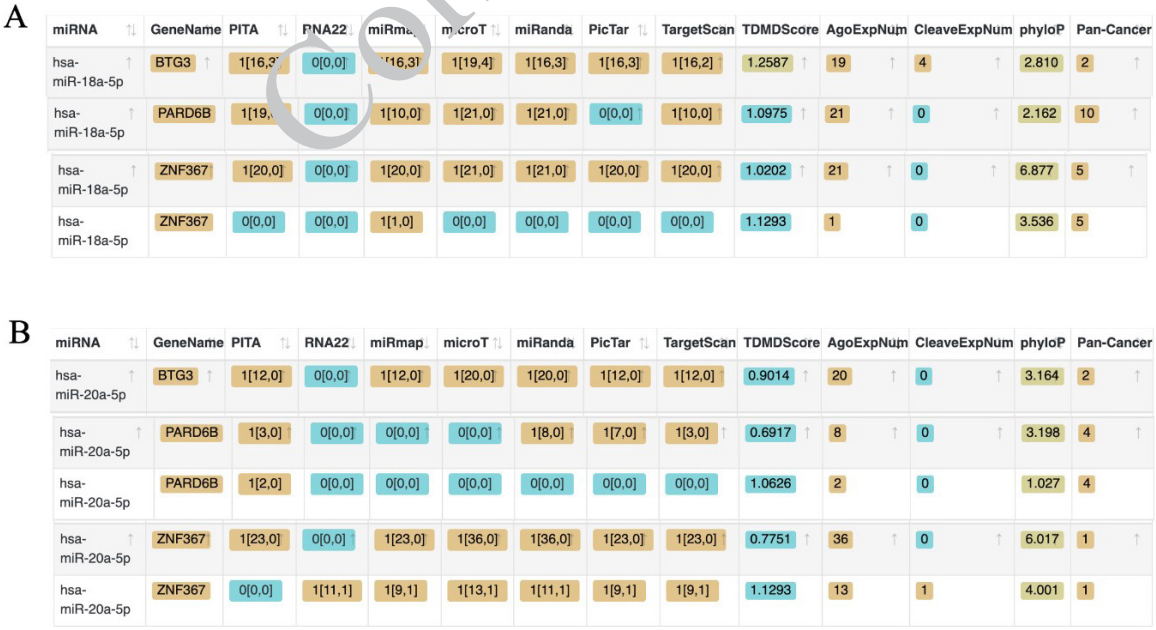

Figure 5. Evaluation of miRNA-mRNA target score based on various databases by the starBase web platform (23)
This section presents the target scores for all three mRNAs associated with mir-18a-5p (A) and mir-20a-5p (B). The results indicate that the selected mRNAs demonstrate strong scores and high confidence across multiple databases (Created using starBase, Microsoft PowerPoint)

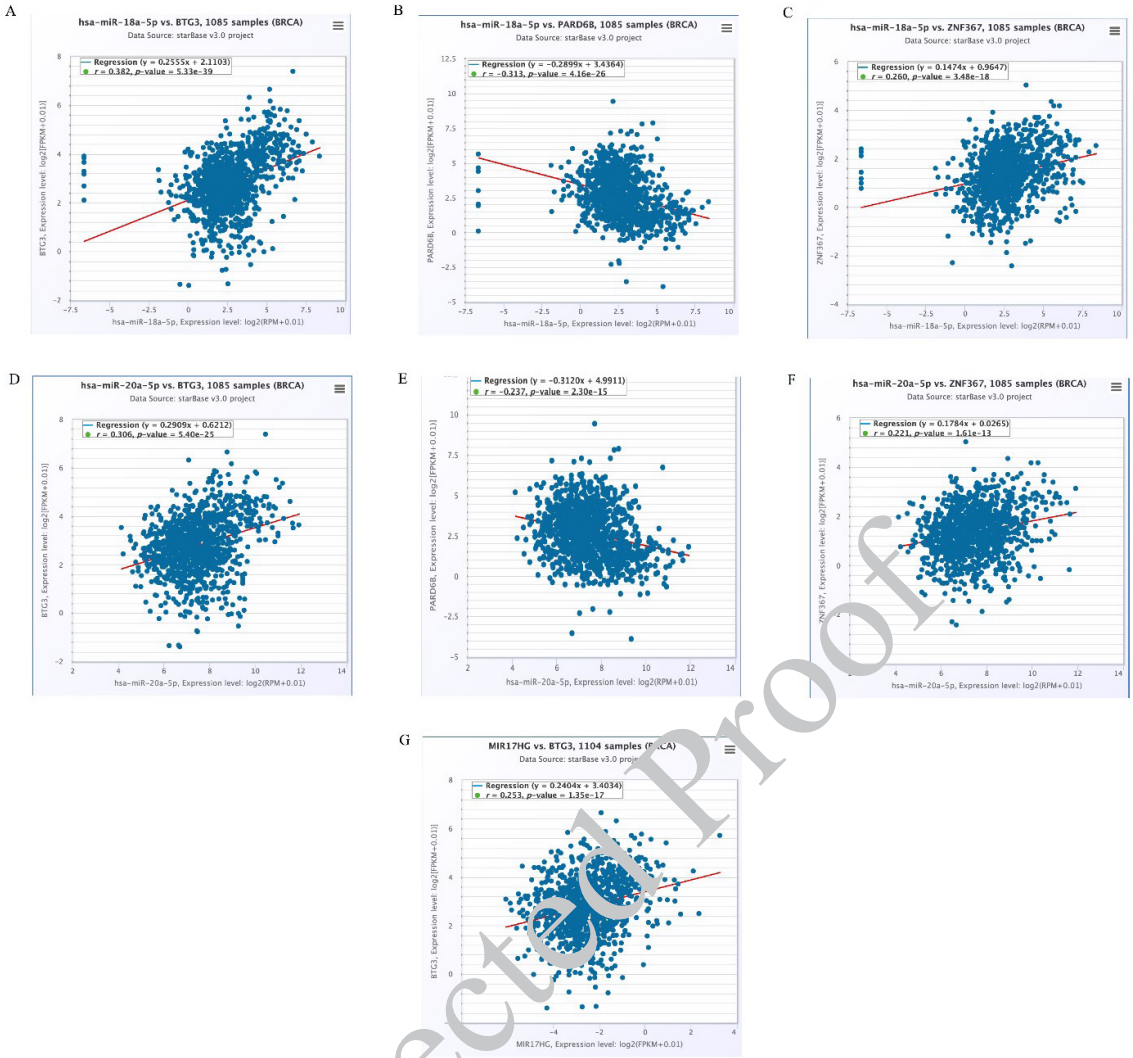

Figure 6. Expression relationship scatter plot between miRNAs and their targets based on the starBase web platform
Within the PanCancer section of starBase, the analysis identified significant associations between miR-18a and its mRNA targets in cases of invasive breast carcinoma (A, B, C). Furthermore, analogous relationships have been documented for mir-20a-5p and its mRNA targets (D, E, F), in addition to the interaction between MIR17HG and BTG3 (G) (Created using starBase, Microsoft PowerPoint)

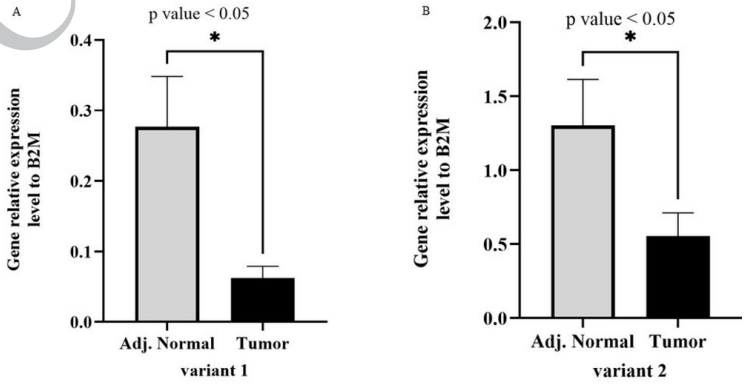

Figure 7. Results of *in vitro* analyses of MIR17HG variants (relative to B2M) in tumor tissue samples in stage II/III compared with adjacent normal tissues.
(A, B) The results of qRT-PCR analysis of uc010tie and uc001vlq, which demonstrated a significant down-regulation of both variants in breast tumors compared to adjacent normal tissues (*P*-value<0.05) (Created using GraphPad Prism 10)

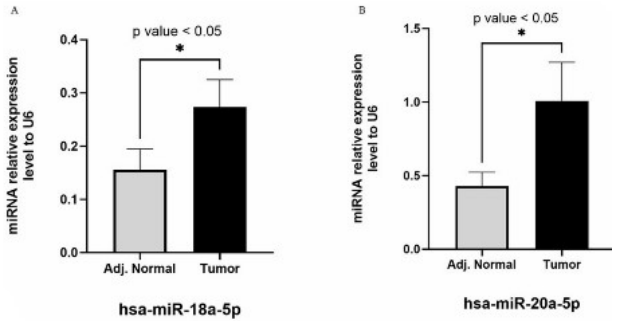

Figure 8. qRT-PCR analysis results of miR-18a-5p and mir-20a-5p (relative to U6) in tumor tissue samples compared to adjacent normal tissue
(A, B) The expression level of miR-18a-5p and mir-20a-5p were significantly up-regulated in breast tumors compared to adj-normal (*P*-value < 0.05) (Created using GraphPad Prism 10)

Discussion

The present study offers a significant contribution to understanding the complex regulatory roles of noncoding RNAs (ncRNAs) in stage II-III BC, particularly focusing on MIR17HG, miR-18a/20a gene expression, and MIR17HG-miR-18a/20a–mRNA axes. By integrating bioinformatic predictions with experimental validation, this research identifies novel candidates for biomarker development and therapeutic strategies. In mammals, MIR17HG has two paralogs—miR-106b-25 (chromosome 7) and miR-106a-363 (chromosome X)—arising from gene duplications (24). It is known as a polycistronic host gene for the miR-17-92 cluster. Our findings revealed significant down-regulation of MIR17HG splice variants in stage II–III BC, aligning with Xu *et al.* (25) but extending their findings by distinguishing between isoforms. However, its derived miRNAs, miR-18a-5p and miR-20a-5p, are overexpressed in stage II/III BC, presenting an intriguing paradox. Previous studies in colon cancer have shown that MIR17HG positively regulates miR-18a expression, where suppression of MIR17HG would be expected to reduce this miRNA level. This discrepancy highlights a critical area for further mechanistic investigation, suggesting that the regulatory interplay between MIR17HG and its embedded miRNAs may be highly context-dependent and involve additional, as-yet-unidentified post-transcriptional mechanisms in BC that allow for inverse expression patterns (26, 27). For instance, loss of heterozygosity at 13q31.3 (harboring MIR17HG) occurs in ~20% of BC cases (28–31), while MYC—a key transcriptional activator of MIR17HG—was co-expressed with MIR17HG in the tumor-suppressive turquoise module (log2FC=−1.25). This paradox (MYC/MIR17HG down-regulation vs. oncogenic roles) may reflect subtype-specific regulation, as MIR17HG is elevated in triple-negative BC but reduced in ER+ tumors (32, 33). These findings likely arise from methodological or subtype variability, underscoring the need for isoform-specific analyses. Furthermore, the miR-17-92 cluster itself exhibits functional diversity; its increased expression is associated with longer overall survival in luminal A BC but shorter overall survival in HER2-enriched and TNBC subtypes.

MiR-18a-5p and miR-20a-5p display dual oncogenic and tumor-suppressive functions depending on cellular context and pathway-specific effects. MiR-18a-5p has been shown to promote tumor progression in lung cancer (via IRF2 targeting) and cervical cancer (via the OCT4–miR-18a axis (34, 35), while acting as a tumor suppressor in HER2-positive BC by targeting HER2 and inhibiting the PI3K/AKT pathway. Additionally, it suppresses tumor growth and invasion in cisplatin-resistant ovarian cancer through targeting MMP-3. These findings underscore the context-dependent nature of this miRNA, influenced by tumor type, microenvironment, and target gene networks (34–37). Similarly, miR-20a-5p exhibits context-dependent roles. While it appears oncogenic in BC, it functions as a tumor suppressor in endometrial cancer, where it inhibits proliferation and invasion via targeting Janus kinase 1 (Jak1) (38). A particularly noteworthy aspect of miR-20a-5p's function, which represents an area of reduction in many studies, including the present one, is its subcellular localization. Research indicates that high miR-20a-5p expression in stromal fibroblasts, TNBC cells, non-small cell lung cancer (NSCLC), colorectal cancer (CRC), and the cytoplasm of cancer cells correlates with increased Ki67 expression (a proliferation marker) and higher odds of relapse, suggesting an oncogenic role in these compartments (39). Conversely, high nuclear miR-20a-5p expression in cancer cells is associated with smaller tumor size and lower odds of lymph node metastasis, potentially indicating a tumor-suppressive role or a mechanism of sequestration within the nucleus. This spatial compartmentalization of miRNA function adds a critical layer of complexity that warrants deeper investigation beyond overall expression levels (40). Previous studies have demonstrated that miR-20a-5p targets key cell cycle regulators, including E2F, c-Myc, and Rb (26, 41–43), while miR-18a-5p modulates hypoxia responses via HIF1A in BC (44).

One of the key regulatory axes identified in our study is the MIR17HG-miR-20a-BTG3 pathway. This axis plays a crucial role in the RNA degradation pathway. This is a fundamental quality-control mechanism for gene expression. Previous research indicates that this regulatory mechanism modulates cellular functions by governing post-transcriptional responses, which are often dysregulated in various cancers. Within this context, we focused on BTG3, a member of the TOB/BTG family located at 17q21, which acts as a tumor suppressor by inhibiting the G1-S cell cycle transition. Studies have demonstrated that TOB family proteins promote mRNA de-adenylation and decay by interacting with the PABPC1 complex and poly(A) nucleases. Consistent with these roles, our data revealed that BTG3 is co-expressed within the turquoise module, which is a highly negatively correlated gene set with BC. Furthermore, the observed up-regulation of miR-20a-5p in our samples suggests that it may function as a post-transcriptional suppressor of BTG3 (45-51). Given the established anti-proliferative role of BTG3, it can be hypothesized that miR-20a-5p binds to the BTG3 transcript, thereby silencing its expression and promoting tumorigenesis. This regulatory loop aligns with previous evidence showing that the suppression of BTG3, whether through miRNA-mediated mechanisms or other processes like antisense-mediated duplex formation, contributes significantly to BC progression.

The second regulatory axis involves the MIR17HG-miR-18a/20a-PARD6B pathway, which appears to be functionally integrated into the Hippo signaling cascade. PARD6B, a key member of the PAR complex (alongside PAR3 and aPKC), is essential for maintaining apical-basal cell polarity. Emerging evidence highlights the crosstalk between cell polarity complexes and the Hippo pathway;

specifically, the PAR complex can inhibit Hippo signaling, thereby promoting YAP/TAZ nuclear localization and cell proliferation. In our study, PARD6B was identified within the blue module, showing the strongest positive correlation with BC and a significant 3.32-fold up-regulation. This aligns with previous studies reporting PARD6B amplification in BC, particularly in ER-positive subtypes. While miRNAs typically induce mRNA degradation, the observed concurrent up-regulation of both PARD6B mRNA and its predicted regulators (miR-18a/20a-5p) suggests a complex regulatory landscape. It can be hypothesized that these miRNAs might be failing to suppress PARD6B translation in the tumor microenvironment effectively, or that the massive transcriptional up-regulation of PARD6B overcomes miRNA-mediated silencing. This accumulation of PARD6B likely contributes to the dysregulation of cell polarity and the subsequent inhibition of the Hippo pathway, facilitating malignant progression (52– 57)

The third identified axis is the MIR17HG-miR-20a-PARD6B pathway, which highlights the role of endocytosis in cancer progression. Endocytosis is a critical regulator of various cellular processes, including immune surveillance, tumor immune evasion, and epithelial-to-mesenchymal transition (EMT). PARD6B is known to facilitate cell migration and EMT by coordinating endocytic adaptors such as Numb. Loss of cellular polarity, often driven by aberrant PARD6B expression, is a hallmark of malignancy, leading to the disruption of tissue architecture. Our results showed a significant 3.23-fold up-regulation of PARD6B in the blue module, which is strongly associated with BC. While miR-20a-5p is predicted to target PARD6B, the robust up-regulation of PARD6B mRNA suggests that the regulatory interplay in the tumor cells is complex. We propose that the overexpression of PARD6B facilitates malignant progression by enhancing endocytic trafficking and promoting EMT. The apparent decoupling of PARD6B mRNA and its predicted miRNA regulation may reflect a failure of the miRNA-mediated silencing machinery in the face of strong transcriptional activation, allowing PARD6B to exert its oncogenic effects on cell polarity and metastasis (58- 60).

Taken together, our qRT-PCR analyses show the potential of candidate genes in biomarker studies. Moreover, integration of the findings with transcriptomic analysis suggests that the proposed MIR17HG–miR-18a/miR-20a–BTG3/PARD6B/ZNF367 axes may participate in RNA degradation, Hippo signaling, and endocytosis in BC. The findings highlight the complex and often paradoxical roles of ncRNAs in cancer. However, these data should be interpreted as providing a framework for future mechanistic and clinical studies rather than definitive mechanistic proof. Future research should prioritize elucidating the precise molecular mechanisms governing the observed inverse expression of MIR17HG and its miRNAs. Although we integrated TCGA miRNA-seq and bulk RNA-seq data with qRT-PCR validation in a well-defined cohort of stage II/III BC patients, our analyses are primarily correlative and based on transcript-level measurements. While it identifies mRNA targets for the miRNAs, comprehensive protein-level validation of these interactions in BC tissues is not presented. The functional consequences of modulating these axes *in vivo* in BC models, beyond the initial expression profiling in patient samples, remain to be fully elucidated. Future studies should include functional assays to directly validate the predicted MIR17HG–miR-18a/miR-20a–BTG3/PARD6B/ZNF367 axes and their biological roles, as well as protein-level validation of BTG3, PARD6B, and ZNF367, to link RNA dysregulation to downstream proteomic changes. Furthermore, while the study focuses on stage II/III BC, this stage encompasses considerable heterogeneity. A more detailed exploration of how these identified axes might differ across specific molecular subtypes within stage II/III BC (e.g., luminal B, HER2-enriched, TNBC) could provide more targeted insights. The precise molecular mechanisms that dictate the dual roles of miR-18a-5p and miR-20a-5p, or the factors influencing their subcellular localization, are also areas that warrant deeper investigation. Additionally, a systematic evaluation of the diagnostic or prognostic performance of these candidates is suggested to address the limitations of this study. Such efforts will be crucial for translating these promising discoveries into effective diagnostic and therapeutic strategies for personalized BC treatment.

Conclusion

This study provides a comprehensive investigation of the expression patterns of MIR17HG splice variants and their associated miRNAs (miR-18a-5p and miR-20a-5p) in stage II/III BC. By integrating large-scale TCGA datasets with WGCNA-based co-expression analysis, *in silico* prediction of ncRNA–mRNA interactions, pathway enrichment, and qRT-PCR validation in patient tissues, we identify MIR17HG and its encoded miRNAs as candidate stage II/III breast cancer–associated ncRNA signatures. The proposed MIR17HG–miR-18a/miR-20a–BTG3/PARD6B/ZNF367 axes point to potential involvement of RNA degradation, Hippo signaling, and endocytosis in BC progression and highlight putative therapeutic targets.

Acknowledgment

This work was financially supported by Tarbiat Modares University of Tehran, Iran. The results presented in this paper were part of a student thesis. We would like to express our deepest thanks to Dr Sara Parsa and Dr. Akram Gholipour for their experimental support, comments, and suggestions that helped improve the project's performance and results.

Furture Dirctions

Given the correlative and transcript-level nature of our analyses, these findings should be regarded as hypothesis-generating rather than definitive. Future studies should (i) delineate isoform-specific functions of mir17hg transcript variants using gain- and loss-of-function assays, (ii) experimentally validate mirna–mrna interactions and protein-level changes for btg3, pard6b, and znf367, and (iii) evaluate the diagnostic and prognostic performance of these ncrnas, as well as their therapeutic potential, in larger and independent bc cohorts and preclinical models.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Data Availability Statement

The datasets used in this study are publicly available from the TCGA project of the GDC portal (https://portal.gdc.

cancer.gov/projects/TCGA-BRCA). Further inquiries can be directed to the corresponding author.

Ethics Statement

The Ethics Committee of Ferdowsi University of Mashhad, Iran, approved the studies involving humans. The studies were conducted in accordance with the local legislation and institutional requirements.

Patient Consent Statement

Written informed consent was obtained from all individual participants included in the study.

Authors' Contributions

R ASh conceived the study and designed the experiments, performed the experiments; R ASh and S T performed data processing, analysis, and interpretation of results, prepared the manuscript, created visualizations, and critically revised the article; A MSh, N N, and S A collected resources; SJ M supervised, directed, and managed the study and acquired funding; R ASh, S T, A MSh, N N, S A, and SJ M approved the final version for publication.

Conflicts of interest

The authors have declared that no competing interests exist.

Declaration

The authors acknowledge the use of artificial intelligence-based tools to assist with writing this manuscript. Specifically, Grammarly (https://www.grammarly.com, Grammarly Inc., San Francisco, CA, USA) was used for English language editing to improve grammar, spelling, and readability of the manuscript. The authors take full responsibility for the content of this publication.

References

1. Booker R. A brief report on the 2020 Canadian Global Oncology Workshop. Can Oncol Nurs J 2021;31:345–346.
2. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, *et al.* Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med 2025;1–9.
3. McGuire A. Factors influencing diagnosis and outcome in HER2 receptor. Cancer Metastasis Rev 2015;34:145–155.
4. Amaral PP, Mattick JS. Noncoding RNA in development. Mamm Genome 2008;19:454–92.
5. Brosnan CA, Voinnet O. The long and the short of noncoding RNAs. Curr Opin Cell Biol 2009;21:416–425.
6. Beermann J, Piccoli MT, Viereck J, Thum T. Non-coding RNAs in development and disease: Background, mechanisms, and therapeutic approaches. Physiol Rev 2016; 96:1297-1325.
7. Shi J, Zhou T, Chen Q. Exploring the expanding universe of small RNAs. Nat Cell Biol 2022;24:415–423.
8. Romano G, Veneziano D, Acunzo M, Croce CM. Small non-coding RNA and cancer. Carcinogenesis 2017;38:485–491.
9. Esteller M. Non-coding RNAs in human disease. Nat Rev Genet 2011;12:861–874.
10. Gholipour A, Shakerian F, Zahedmehr A, Irani S, Malakootian M, Mowla SJ. Bioinformatics analysis to find novel biomarkers for coronary heart disease. Iran J Public Health 2022;51:1152.
11. Olive V, Li Q, He L. mir-17-92: A polycistronic oncomir with pleiotropic functions. Immunol Rev 2013;253:158–166.
12. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. Cell 2008;133:217–222.
13. Li H, Wu Q, Li T, Liu C, Xue L, Ding J, *et al.* The miR-17-92 cluster as a potential biomarker for the early diagnosis of gastric cancer: Evidence and literature review. Oncotarget 2017;8:45060.
14. Zhang X, Li Y, Qi P, Ma Z. Biology of MiR-17-92 cluster and its progress in lung cancer. Int J Med Sci 2018;15:1443.
15. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:1–13.
16. Chen Y, Wang X. miRDB: An online database for prediction of functional microRNA targets. Nucleic Acids Res 2020;48:D127–D-131.
17. Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol 2019;20:1–10.
18. Das G, Das T, Parida S, Ghosh Z. LncRTPred: Predicting RNA–RNA mode of interaction mediated by lncRNA. IUBMB Life 2024;76:53–68.
19. Tastsoglou S, Skoufos G, Miliotis M, Karagkouni D, Koutsoukos I, Karavangeli A, *et al.* DIANA-miRPath v4. 0: Expanding target-based miRNA functional analysis in cell-type and tissue contexts. Nucleic Acids Res 2023;51:W154–W159.
20. Tastsoglou S, Alexiou A, Karagkouni D, Skoufos G, Zacharopoulou E, Hatzigeorgiou AG. DIANA-microT 2023: Including predicted targets of virally encoded miRNAs. Nucleic Acids Res 2023;51:W148–W153.
21. Karagkouni D, Paraskevopoulou MD, Tastsoglou S, Skoufos G, Karavangeli A, Pierros V, *et al.* DIANA-LncBase v3: Indexing experimentally supported miRNA targets on non-coding transcripts. Nucleic Acids Res 2020;48:D101–D110.
22. Paraskevopoulou MD, Karagkouni D, Vlachos IS, Tastsoglou S, Hatzigeorgiou AG. MicroCLIP super learning framework uncovers functional transcriptome-wide miRNA interactions. Nat Commun 2018;9:3601.
23. Li JH, Liu S, Zhou H, Qu LH, Yang JH. StarBase v2. 0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014;42:D92–D97.
24. Lu Y, Thomson JM, Wong HYF, Hammond SM, Hogan BLM. Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. Dev Biol 2007;310:442–453.
25. Xu J, Hu M, Gao Y, Wang Y, Yuan X, Yang Y, *et al.* LncRNA MIR17HG suppresses breast cancer proliferation and migration as ceRNA to target FAM135A by sponging miR-454-3p. Mol Biotechnol 2023;65:2071–2085.
26. Hossain MM, Sultana A, Barua D, Islam MN, Gupta A, Gupta S. Differential expression, function and prognostic value of miR-17–92 cluster in ER-positive and triple-negative breast cancer. Cancer Treat Res Commun 2020;25:100224.
27. Yuan G, Liu B, Han W, Zhao D. LncRNA-MIR17HG mediated upregulation of miR-17 and miR-18a promotes colon cancer progression via activating Wnt/β-catenin signaling. Transl Cancer Res 2019 Aug;8:1097–1108.
28. Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: A comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death Differ 2013;20:1603–1614.
29. Mihailovich M, Bremang M, Spadotto V, Musiani D, Vitale E, Varano G, *et al.* MiR-17-92 fine-tunes MYC expression and function to ensure optimal B cell lymphoma growth. Nat Commun 2015;6:8725.
30. Ota A, Tagawa H, Karnan S, Tsuzuki S, Karpas A, Kira S, *et al.* Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res 2004;64:3087–3095.
31. Calin GA, Croce CM. Chromosomal rearrangements and microRNAs: a new cancer link with clinical implications. J Clin Invest 2007;117:2059–2066.
32. Chacon-Cortes D, Smith RA, Lea RA, Youl PH, Griffiths LR. Association of microRNA 17–92 cluster host gene (MIR17HG) polymorphisms with breast cancer. Tumor Biol 2015;36:5369–5376.
33. Moi L, Braaten T, Al-Shibli K, Lund E, Busund LTR. Differential

expression of the miR-17-92 cluster and miR-17 family in breast cancer according to tumor type; results from the Norwegian Women and Cancer (NOWAC) study. J Transl Med 2019;17:1–20.
34. Kolenda T, Guglas K, Kopczyńska M, Sobocińska J, Teresiak A, Bliźniak R, *et al.* Good or not good: Role of miR-18a in cancer biology. Reports Pract Oncol Radiother 2020;25:808–819.
35. Shen K, Cao Z, Zhu R, You L, Zhang T. The dual functional role of MicroRNA-18a (miR-18a) in cancer development. Clin Transl Med 2019;8:32.
36. Calvano Filho CMC, Calvano-Mendes DC, Carvalho KC, Maciel GA, Ricci MD, Torres AP, *et al.* Triple-negative and luminal A breast tumors: Differential expression of miR-18a-5p, miR-17-5p, and miR-20a-5p. Tumor Biol 2014;35:7733–7741.
37. Enblad G, Karlsson H, Gammelgård G, Wenthe J, Lövgren T, Amini RM, *et al.* A phase I/IIa trial using CD19-targeted third-generation car T cells for lymphoma and leukemia. Clin Cancer Res 2018;24:6185–6194.
38. He Y, Ma H, Wang J, Kang Y, Xue Q. miR-20a-5p inhibits endometrial cancer progression by targeting janus kinase 1. Oncol Lett 2021;21:427.
39. Guo L, Zhu Y, Li L, Zhou S, Yin G, Yu G, *et al.* Breast cancer cell-derived exosomal miR-20a-5p promotes the proliferation and differentiation of osteoclasts by targeting SRCIN1. Cancer Med 2019;8:5687–5701.
40. Stoen MJ, Andersen S, Rakaee M, Pedersen MI, Ingebriktsen LM, Donnem T, *et al.* Overexpression of miR-20a-5p in Tumor Epithelium Is an Independent Negative Prognostic Indicator in Prostate Cancer-A Multi-Institutional Study. Cancers (Basel) 2021;13:4096.
41. Pickering MT, Stadler BM, Kowalik TF. MiR-17 and miR-20a temper an E2F1-induced G1 checkpoint to regulate cell cycle progression. Oncogene 2009;28:140–145.
42. Dews M, Homayouni A, Yu D, Murphy D, Sevignani C, Wentzel E, *et al.* Augmentation of tumor angiogenesis by a Myc activated microRNA cluster. Nat Genet 2006;38:1060–1065.
43. Li JY, Zhang Y, Zhang WH, Jia S, Kang Y, Zhu XY. Differential distribution of miR-20a and miR-20b may underlie metastatic heterogeneity of breast cancers. Asian Pacific J Cancer Prev 2012;13:1901–1906.
44. Krutilina R, Sun W, Sethuraman A, Brown M, Seagroves TN, Pfeffer LM, *et al.* MicroRNA-18a inhibits hypoxia-inducible factor 1α activity and lung metastasis in basal breast cancers. Breast Cancer Res 2014;16:R78.
45. Deng M, Wang X, Xiong Z, Tang P. Control of RNA degradation in cell fate decision. Front cell Dev Biol 2023;11:1164546.
46. Nogueira G, Fernandes R, García-Moreno JF, Romão L. Nonsense-mediated RNA decay and its bipolar function in cancer. Mol Cancer 2021;20:72.
47. O'Malley S, Su H, Zhang T, Ng C, Ge H, Tang CK. TOB suppresses breast cancer tumorigenesis. Int J Cancer 2009;125:1805–1813.
48. Ezzeddine N, Chang TC, Zhu W, Yamashita A, Chen CYA, Zhong Z, *et al.* Human TOB, an antiproliferative transcription factor, is a poly (A)-binding protein-dependent positive regulator of cytoplasmic mRNA deadenylation. Mol Cell Biol 2007;27:7791–7801.
49. Ezzeddine N, Chen CYA, Shyu AB. Evidence providing new insights into TOB-promoted deadenylation and supporting a link between TOB's deadenylation-enhancing and antiproliferative activities. Mol Cell Biol 2012;32:1089–1098.
50. Gou W feng, Yang X feng, Shen D fu, Zhao S, Liu Y peng, Sun H zhi, *et al.* The roles of BTG3 expression in gastric cancer: A potential marker for carcinogenesis and a target molecule for gene therapy. Oncotarget 2015;6:19841–19867.
51. Yanagida S, Taniue K, Sugimasa H, Nasu E, Takeda Y, Kobayashi M, *et al.* ASBEL, an ANA/BTG3 antisense transcript required for tumorigenicity of ovarian carcinoma. Sci Rep 2013;3:1305.
52. Karasek C, Ashry M, Driscoll S, Knott JG. A tale of two cell-fates: Role of the hippo signaling pathway and transcription factors in early lineage formation in mouse preimplantation embryos. Mol Hum Reprod 2020;26:653–664.
53. Ahmed SM, Maitra IG. Mechanisms of polarity protein expression control. Curr Opin Cell Biol 2016;42:38–45.
54. Yang H, Yang J, Zheng X, Chen T, Zhang R, Chen R, *et al.* The hippo pathway in breast cancer: The extracellular matrix and hypoxia. Int J Mol Sci 2024;25:12868.
55. Nolan ME, Aranda V, Lee S, Lakshmi B, Basu S, Allred DC, *et al.* The polarity protein Par6 induces cell proliferation and is overexpressed in breast cancer. Cancer Res 2008;68:8201–8209.
56. Muthuswamy SK, Xue B. Cell polarity as a regulator of cancer cell behavior plasticity. Annu Rev Cell Dev Biol 2012;28:599–625.
57. Gupta SRR, Nagar G, Mittal P, Rana S, Singh H, Singh R, *et al.* Breast cancer therapeutics and hippo signaling pathway: Novel microRNA-gene-protein interaction networks. Omi A J Integr Biol 2023;27:273–280.
58. Sadri F, fatemeh Hosseini S, Rezaei Z, Fereidouni M. Hippo-YAP/TAZ signaling in breast cancer: Reciprocal regulation of microRNAs and implications in precision medicine. Genes Dis 2024;11:760–771.
59. Cunliffe HE, Jiang Y, Fornace KM, Yang F, Meltzer PS. PAR6B is required for tight junction formation and activated PKCζ localization in breast cancer. Am J Cancer Res 2012;2:478.
60. Marques E, Englund JI, Tervonen TA, Virkunen E, Laakso M, Myllynen M, *et al.* Par6G suppresses cell proliferation and is targeted by loss-of-function mutations in multiple cancers. Oncogene 2016;35:1386–1398.