RNA Sequencing-Based Total RNA Profiling; The Oncogenic MiR-191 Identification as a Novel Biomarker for Breast Cancer

Sevan Omer Majed

Biology Department, Education College, Salahaddin University-Erbil, Kurdistan Region-Iraq

ARTICLE INFO

ABSTRACT

This study aims to screen the differential expression of total RNA transcripts in formalin-fixed paraffin-embedded tissues (FFPETs) in breast cancer (BRCA) and normal adjacent tissues (NATs) and identify miR-191 as a new biomarker for early diagnosing BRCA. Differentially expressed genes (DEGs) by MACE-Seq and differentially expressed ncRNAs (DEncRNAs) by the TrueQuant technique were examined in this study. The miR-191 expression level was measured by Real Time-qPCR. An average of 4,739 coding genes from 25,713 significantly down-regulated genes was identified, whereas 3,954 coding genes were significantly up-regulated in the BRCA against NAT. An average of 1450 ncRNAs, including up-regulated= 679 and down-regulated= 780, were differentially expressed in 7 paired samples of BRCA and NAT. Among the ncRNAs, 227 microRNAs, including unchanged= 152, down=53, and up=22, were differentially expressed. MiR-191 was one of the 22 significant up-regulation, with $p=0.0001$. RT-qPCR results confirmed that miR-191, $p=0.003$, was significantly over-expressed in 120 paired samples of BRCA and NAT. Furthermore, NextSeq 500 revealed that a single nucleotide polymorphism (C>T) newly occurred in the mature sequence of miR-191-5p seed region in BRCA samples. However, the putative target genes regulated by the miR-191-5p were recognized by the above ten computational programs for the prediction. MACE-Seq outcomes showed that the genes of *CDK6*($P=0.0001$), *DAPK1*($P=0.02$), *MTC7*($P=0.04$), *SETD1B*($P=0.005$), *CALN1*($P=0.01$), and *TMOD2*($P=0.001$) were significantly over-expressed in the BRCA against the NATs. The expression level of the targets was adversely related to the miR-191-5p.

Introduction

Breast cancer (BRCA) is a genomic illness wreaking havoc on women's health worldwide (1–3). According to the Global Cancer Project 2018, The most frequent cancer diagnosed in women is BRCA (11.7 percent of cases all), which is followed by lung (11.4 percent) colorectal (10.0 percent) cancers (4). In terms of incidence and mortality, breast cancer is the leading reason for cancer death in women, followed by colorectal and lung cancers (4,5). BRCA originates when the gene expression of individual changes, causing breast cells to grow and multiply abnormally. These alterations can occur in the breast's ducts, glands, and other tissues (6,7). MicroRNAs (miRNAs), as single-stranded RNAs, have a length of 19-22nt and a category of regulatory RNAs regulating the expression of a gene by targeting messenger RNAs (mRNAs) at the post-transcriptional stage (8,9). They are vital regulators of both physiologic and pathologic processes, controlling the expression of 10–30% of the human transcriptome (10). Their abnormal expression might have a role in human cancers, especially BRCA (11). Specific BRCA clinicopathologic characteristics, such as expression level of estrogen and progesterone receptor, vascular invasion, and tumor stage, are related to miRNA expression. Furthermore, BRCA has been found to have an abnormal expression of miRNAs between normal and cancerous tissues, as well as between breast cancer molecular subtypes, such as Her2+, luminal A, luminal B, and basal-like (3,12). A study revealed that miRNAs in BRCA have significantly contributed to the onset and development. MiR-21 upregulation in this carcinoma, for instance, has been related to progressive clinical-stage, poor patient prognosis, lymph node metastases, and targeted key tumor-suppressor genes including PDCD4, TPM1, and PTEN (13,14). Another overexpressed miRNA was miR-1228, which targets the SCAI protein, promoting BRCA cell proliferation and metastasis.

Moreover, miRNAs regulate oncogenes, such as the tyrosine kinase receptor HER-2 and HER-3, which are regulated, respectively, by miR-125b and miR-205 (15). In metastatic breast cancer, the family of miR-200, which targets ZEB transcription factor members and reduces cell motility and invasiveness, was reduced (16).

A study showed that miR-191 is strongly up-regulated in several human cancers, such as prostate, lung, colon, pancreatic, and stomach cancer (17,18). In BRCA, a substantial positive association was also found between miR-191 expression and ER levels (11). Another study indicated that p53 in BRCA causes miR-191-5p to be downregulated. miR-191-5p overexpression in BRCA cell lines (MCF7 and ZR-75) inhibited apoptosis, while downregulation of miR-191-5p had the reverse impact. SOX4 gene was also discovered to be a miR-191-5p direct target. In MCF7 cells, overexpression of SOX4 was found to enhance p53 protein level. Overexpression of miR-191-5p resulted in a decrease in the levels SOX4 and consequently p53 (19). Another earlier research revealed that miR-191 has previously been identified as an oncogenic miRNA that is overexpressed in breast cancer and aids in malignant transformation development; therefore, antisense miR-191 (anti-miR-191) suppression of miR-191 has enormous therapeutic potential. A cationic liposome-based on stearylamine (SA) to deliver anti-miR-191 inhibitor (anti-miR-191) was generated to BRCA cells (MCF-7 and ZR-75-1) in culture and tested its effectiveness (20). A function for miR-191 in carcinogenesis is supported by a number of data, including the fact that miR-191 enhances prostate cancer radiation resistance over interaction with RXRA (18). However, miR-191 enhances human gastric cancer carcinogenesis (21) and colorectal cancer carcinogenesis by targeting C/EBP (22). MiR-191 in liver cancer is triggered by a dioxin family carcinogen, which is overexpressed (23). Furthermore, miR-191 suppression decreased hepatocellular carcinoma cell proliferation and tumor development (24). Another recent study revealed that miRNA-191 in liver cancer increased cell proliferation through has_circ_0000204/miR-191/KLF6 axis (25).

The present study aimed to screen differentially expressed non-coding RNAs (DEncRNAs) and miRNAs (DEmiRNAs) associated with BRCA; the research design is mapped in Fig. 1. In this work, the miR-191 expression level in formalin-fixed paraffin-embedded tissues (FFPETs) of BRCA compared to normal adjacent tissues (NAT) was identified as a biomarker for BRCA diagnostics. Another aim was to analyze of precursor sequence of miR-191 to detect any single nucleotide polymorphism (SNP), If present. Then, identification of the potential targeted genes of miR-191 was made. The biological role of mRNA-miR-191 interaction was determined in BRCA cells by Gene Ontology (GO) enrichment analysis. The survival curve of Kaplan Meier was considered to find out the prognostic significance relationship between patient overall survival (OS) and the expression levels of miR-191 and identified target genes. Our results may offer pivotal information helping to explain the DEncRNAs and DEmiRNAs from breast cancer and can offer a biomarker for the treatment of BRCA patients.
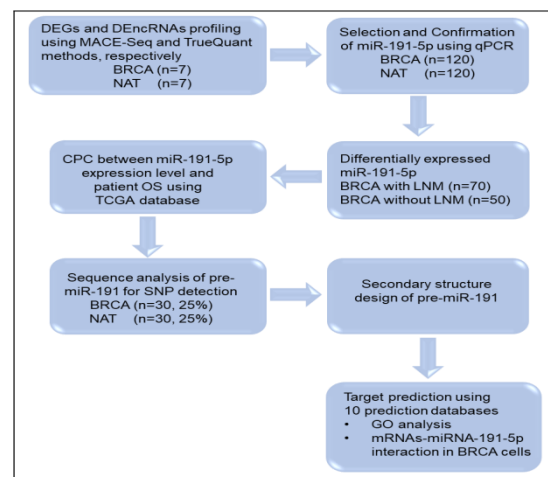
Figure 1. Workflow of the design study. DEGs; are differentially expressed genes. DEncRNAs; are differentially expressed non-coding RNAs. MACE-Seq; Massive Analysis of cDNA End Sequencing. BRCA; breast cancer. NAT; Normal adjacent tissues. LNM; Lymph node metastasis. CPC; clinicopatlogical correlation. OS; overall survival. pre; precursor.

Materials and methods
Ethics approval and consent to join
The Human Research Ethics Committee (HREC) at Salahuddin University, Erbil, had previously approved all procedures used during the study (Reference No. 4d/132). By signing the verified consent of all study patients, permission for publication was gained.

Throughout the study, all processes were carried out in compliance with the Helsinki Declaration of 1964, and all participants obtained written informed consent and approval for publication.

Sample collection of FFPET

At Al-Mufti and Luay's clinicopathological laboratories, 120 paired samples of FFPE blocks of Kurdish patients with BRCA were obtained. None of these patients had hormone, chemotherapy and radiotherapy. A pair of paraffin blocks (1 cancerous tissue and 1 normal adjacent tissue (NAT)) were obtained for each patient. Stains of Eosin and Hematoxylin were used to identify the cancerous and non-cancerous sites before beginning the total RNA extraction. from the malignant zone, NATs were taken 2cm away. biopathological characteristics of the samples were determined by employing a questionnaire. Table 1 demonstrates a summary of clinicopathological features of the BRCA cases.

RNA extraction

Malignant and non-malignant regions of FFPE of seven paired samples were identified using stains of Eosin and Hematoxylin prior to total RNA extraction in a histopathological laboratory (in GenXPro GmbH, Germany). Then, 10μm of each sample was deparaffinized twice with absolute methanol (each 1ml) for 5 minutes and three times with ethanol (70%, 90%, and 95%) for 3 minutes each. The total RNA extraction was achieved by the Quick-RNA FFPE Kit (Cat: R1008, Zymo Research, USA). Then, DNA contamination was removed using DNAse I digestion (Baseline-ZERO kit, Epicentre, USA). In order to determine the total RNA concentration, the RNA Pico Sensitivity Assay on a LabChip GXII Touch (PerkinElmer, USA) was applied.

Small RNA-sequencing library construction for analysis of DEncRNAs

TrueQuant-libraries were constructed using a TrueQuant kit containing all reagents by GenXPro GmbH, as previously described (26). This kit also contains TrueQuant barcodes/unique molecule identifiers (UMIs) that are used to barcode each transcript individually before PCR reactions. Briefly, from 1μg of total RNA, poly-adenylated mRNA was extracted. Fourteen TrueQuant-libraries were prepared.

After that, the small RNA (sRNA) libraries were also made by applying small RNA molecules. These sRNA libraries were built by the GenXPro GmbH to the rest of the non-polyadenylated RNAs using the flashPAGE™ Fractionator System (Life Technologies). The small RNAs were attached directly to TrueQuant Adapters, as previously described by Hafner and his coworkers (27). A Hiseq2000 (Illumina, US) was applied to order the nucleotide populations of sRNA molecules comprising p5 and p7 adapters. The ordered reads were cut out for quality sequence to exclude flanking adapters. The data was available in (the OmiRas database under c0cf26993013).

Table 1. Clinicopathological feature of the 120 samples of BRCA

Pathological feature	No. of Cases (%)
Age	
>50	67 (55.83)
≥50	53 (44.17)
BMI	
≤20	22 (18.33)
20-25	37 (25.83)
≥25	61 (55.83)
Tumor size	
≤2cm	52 (43.33)
2-5cm	60 (50)
≥5cm	8 (6.67)
Tumor grade	
Grade I	13 (10.83)
Grade II	78 (65)
Grade III	29 (24.17)
Lymph node metastasis	
Yes	70 (58.33)
No	50 (41.67)
Lymphatic invasion	
Yes	83 (69.17)
No	37 (30.83)
Venous invasion	
Yes	(74) (61.67)
No	(46) (38.33)
Biomarkers	
Triple negative	60 (50)
ER+ & PR+	30 (25)
ER- & PR-	10 (8.33)
ER+ & HER2+	20 (16.67)
Ki67 status	
Yes	112 (93.33)
No	0
Not available	8 (6.67)

BMI; Body mass index. ER; Estrogen. PR; Progesterone HER; human epidermal growth factor receptor 2

RNA-sequencing library construction for analysis of DEmRNAs

To understand the differentially expressed mRNA (DEmRNA) transcripts, RNA sequencing was carried out from the seven paired samples that were used to

identify DEncRNA molecules by applying Massively Analysis of cDNA End sequencing (MACE-Seq). The materials and methods for extraction of total RNA were previously described and also found in (3). Biotinylation of the cDNA cap structures was done precisely. Streptavidin beads attracted full-length cDNAs with biotinylated caps, which randomly were trimmed to be around 300 bps. The purified mRNA transcripts were sequenced using Illumina Hiseq2000 with 1 ×100 bps. The RNA libraries were then made by MACE-kit (GenXPro GmbH, Germany). The mapped reads were trimmed for the quality sequence to eliminate flanking adapters.

Statistics and bioinformatics analysis

Later, the Bowtie 2 tool was utilized to match the arranged reads of DEncRNA and DEmRNA transcripts to the human reference genome (GRCh38.p13) and annotate them with the appropriate properties, as described in (28). Mapped reads to each gene were quantified using the HT-seq tool (29). Differential expression processing and normalized counts were carried out with the DESeq2 tool, as previously defined in Love et al. (30). Negative binomial generalized linear models were used in this program. Analysis of differential transcripts was then performed with edgeR, Glimma and Limma packages (31). DEncRNA and DEmRNA transcripts were log2-fold change (Log2FC) >2 or <−2 and by an adjusted *P*-value < 0.05, and FDR. These parameters were grouped into a final Excel table. For more analysis, the outputs of DESeq2 are imported into RStudio (V1.2.1335, R V 3.5.3). A scatter plot for DEncRNAs, a volcano plot for differentially expressed microRNAs, and a scatter plot for DEmRNAs were designed to show the significance of differentially expressed reads in different libraries.

Confirmation of miR-191 expression using qRT-PCR
Total RNA extraction for miR-191 expression determination

MiR-191 among the DEncRNA results was up-regulated significantly (*P*=0.0001) in BRCA, compared to NAT; therefore, it was chosen to approve its differential expression in 120 Kurdish cases using the quantitative Real-Time polymerase chain reaction (qRT-PCR) technique. One hundred twenty paired FFPE blocks of BRCA tissues and NATs were

analyzed. All procedures were done in Salahuddin University Research Center (SURC).

Deparaffinization

For each test, 20 mg of the FFPET was used to isolate total RNA transcripts. The tissue was deparaffinized by its incubation in an Eppendorf tube, containing 1ml of xylene for at 50℃ 5 minutes. Then, the Eppendorf tube was centrifuged at 14,000 rpm for 2 minutes to generate pellets. Next, rehydration of the deparaffinized pellet was performed three times with 1ml of 100%, 95% and 90% ethanol, respectively. Then, the pellet was air-dried for more than 10 minutes at ambient temperature. Then, PowerGen 125 Tissue Homogenizer (Fisher Scientific, USA) was used to homogenize the pellet.

Total RNA extraction

The overall RNA transcripts were isolated by the FFPE RNA/DNA Purification Plus kit (Cat. No. 54300, NORGEN BIOTEK CORP, Canada). As previously defined in the kit procedure. The overall RNA concentration was measured using the Nanodrop technique (Fisher Scientific, USA). Then, all of the qRT-PCR products were ordered from the Abmgood company in the United States via the Iraq Biotech company. In the present study, 5μl of the total RNA was used to make a sufficient amount of complementary DNA (cDNA) utilizing the miRNA All-In-One cDNA Synthesis Kit (Cat. No. G898, Abmgood company, US). In order to run a qPCR, the total volume for each well was 20 μl, containing 10μl of BrightGreen miRNA qPCR MasterMix-ROX (Cat. No. MasterMix-mR), 0.5μl for each reverse and forward primers (Cat. No. MPH01555), 1.5μl of template cDNA, and 7.5μl of SNORD44 primers (Cat. No. MPH0003) and U6-2 primers (Cat. No. MPH0001) were used as housekeeping primers.

Determination of miR-191 differential expression utilizing qRT-PCR

The following three-step cycling program was used to set up the qRT-PCR reaction. After 10 minutes of enzyme activation at 95°C, 40 cycles of denaturing for 10 seconds at 95°C, annealing at 65°C for 20 secs, and extension at 72°C for 20 sec were set up.

Selected targeted genes regulated by miR-191-5p

To explore the chosen target genes regulated by miR-191-5p in BC, above ten computational programs of prediction were used to search the binding spots in the 3'UTR (3'untranslated region). Table 2 was shown the details of target predicted sites. The identified target genes are the intersection of all database results.

Differentially expressed genes (DEGs)

The seven paired samples that were used to identify DEncRNA molecules, were also used to identify differential expression of the identified target genes by applying Massively Analysis of cDNA End sequencing (MACE-Seq). The procedure of these techniques and bioinformatic analysis were described in (3). Then, the result of the identified target genes was statistically analyzed by GraphPad Prism (V 8.0.1).

Gene Ontology (GO) enrichment analysis

In this study, the biological part of the identified target genes of miR-191-5p in BRCA was investigated using the GenXPro tool (http://tools.genxpro.net/modules/GO Enrichment Tool v2). The gene enrichment analysis was performed to show Gene Ontology (GO) terms, including Molecular Function (MF) and Biology Process (BP).

Kaplan-Meier survival analysis

Kaplan-Meier estimator was utilized to measure a prognostic relationship between expression levels of miR-191 and selected target genes with patient overall survival (OS). The BRCA data was gained from the cBioPortal databases (http://www.cbioportal.org/) and OncoLnc (http://www.oncolnc.org/). On Aug 2021, overall survival (OS) per month was then analyzed using RStudio (R 3.0.1) software. The data and R scripts are kept for future research.

Table 2. The use of computational prediction tools to find potential targets binding with miR-191-5p seed sites were discovered.

| Target predicted Sites	Species	Tool properties	links
miRmap	Chimpanzee, Human, Rat, Mouse, Zebrafish, Cow, Chicken, Opossum	seed match, Conservation, free energy	https://mirmap.ezlab.org/
miRWalk	Human, Rat, Mouse	seed match Conservation, free energy	http://mirwalk.uni-hd.de/
TargetMiner	Human, Rat, Mouse, Fly	seed location ,Conservation	https://www.targetMiner.html
MirTar2	Human, rat, Mouse, Dog, Chicken	seed location, Conservation	http://www.mirdb.org/
DIANA	Any	seed match, Conservation, free energy	http://www.microrna.gr/micro T-CDS
Target scan	Human, Fly, Mouse, Fish, Worm	seed location Conservation,	http://www.targetscan.org/
Microrna. org	Human, Fruit Fly, mouse, rat	seed match, Conservation, free energy	http://www.microrna.org/
RNA22	Human, Fruit Fly, Worm Mouse,	Seed match, free energy	https://cm.jefferson.edu/rna22/
MirPath	Human, Mouse, D. melanogaster, C. elegans, R. norvegicus, D. rerio and G. gallus	Conservation, seed match free energy	http://snf-515788.vm.okeanos.grnet.gr/index.php?r=mirpath/geneList
miRTarBase	Human, Rat Mouse,	Conservation, seed location	http://mirtarbase.mbc.nctu.edu.tw/php/index.php
PicTar - Tools4miRs	Human, Rat, Mouse, Fly	seed location, Conservation,	https://tools4mirs.org/software/target_prediction/pictar/

Results and discussion

Transcriptional landscape of DEncRNA in FFPET of BRCA

Kurdish women ncRNA transcript profiles of 7 BRCA samples and their matched NATs were analyzed to identify differential expression. In this result, Software Asset Management (SAM) tool was used to filter out 1450 ncRNA molecules. The raw data was then normalized by using, *P*-value less than 0.05 and log2 fold-changes (log2FCs), as seen in the scatter plot (Fig. 2A), indicating that 670 ncRNA transcripts were over-expressed; whereas 780 ncRNAs were down-expressed in BRCA, compared with NAT. Each blue dot represents only one ncRNA molecule. The x-axis signifies the NAT data, and the y-axis signifies the BC data. Among the 1450 DEncRNAs, 227 microRNAs in the BRCA were differentially expressed, compared to the NAT. among 227 microRNAs, 152 microRNAs were down-regulated significantly, but 22 microRNAs were up-regulated significantly. The others were unchanged (Fig. 2B). In the present study, miR-191 was found to have a high expression value (*P*=0.0001) in the BRCA cells; therefore, it was focused on confirming the expression value using the RT-qPCR instrument. Then, the precursor miRNA-191 sequence in 30 (25%) of BRCA samples was compared to the sequence in 30 (25%) NAT samples to see if any polymorphism was present.

Transcriptional landscape of DEGs in FFPET of BRCA

The protein-coding genes at the same samples used for ncRNA profiling were analyzed. It was detected that 25,713 genes were differentially expressed in the BRCA ($\geq$ 2-fold and FDR $\leq$ 0.001, with *P*-value less than 0.05). Figure 3A displayed the differential expression of 8,693 genes. Four thousand seven hundred thirty-nine genes from 8,693 genes were down-regulated significantly, whereas 3,954 were up-regulated

significantly in the BRCA compared with the NAT. Others were unchanged by 572.
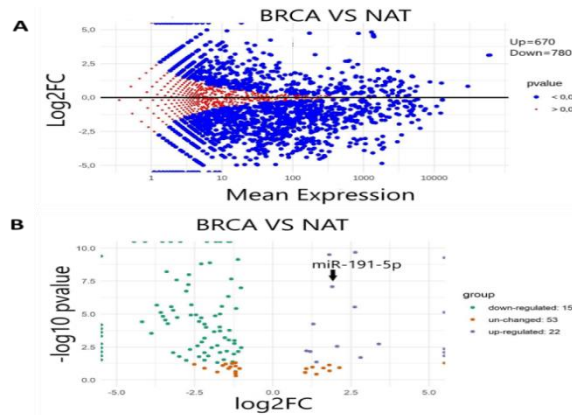
Figure 2. Analysis of small RNAs ' differential expression in BRCA against adjacent normal tissues; **A)** Comparison of total small non-coding RNAs' differential expression, with *P*-value < 0.05. **B)** Comparison of the microRNAs' differential expression in the breast cancer and normal adjacent tissues, with *P*-value < 0.05. MiR-191 was pointed out in the volcano plot.
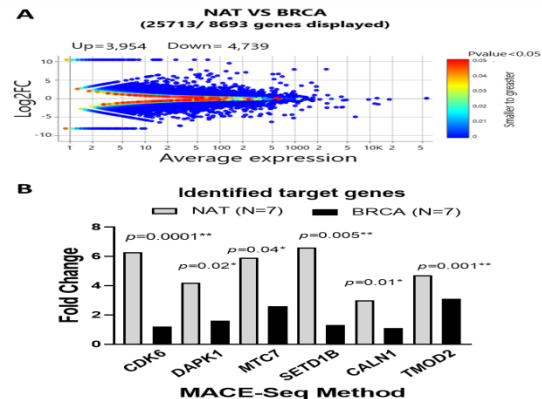
Figure 3. Analysis of differential expression of protein-coding genes in breast cancer against normal tissues. **A)** differential expression of 8,693 genes was displayed from 25,713 genes. Among the 8,693 genes, 4,739 genes were down-regulated significantly, and 3,954 genes were up-regulated significantly in BRCA, as compared to the normal adjacent tissues, with *P*-value < 0.05. **B)** Comparison of differential expression of target genes regulated by miRNA-191-5p. All the targets were down-expressed in the BRCA, with *P*value< 0.05. **N** denotes the number of samples.

Based on the eleven computational prediction targets, six target genes, such as *CDK6, DAPK1, MTC7, SETD1B, CALN1,* and *TMOD2*, were chosen to have a complementarity of base pairing

with miR-191-5p. According to the MACE-Seq result, the expression levels of these genes between BRCA and NAT were significantly different. The expression levels of *CDK6*(*P*=0.0001), *DAPK1*(*P*=0.02), *MTC7*(*P*=0.04), *SETD1B*(*P*=0.005), *CALN1*(*P*=0.01), and *TMOD2*(*P*=0.001), were down-regulated significantly in the BRCA, as compared to NAT (Fig. 3B). Table 3 provided brief information about the gene name, gene ID, chromosome location, and binding site of the seed site of miR-191-5p with 3'UTR (untranslated region) of the target genes.

Table 3. The information about the identified target genes that were regulated by miR-191-5p

Gene	Gene ID	Gene Name	Chromosome location	miRNA-191-5p seed site binds the 3'UTR of target genes by base pairing
CDK6	ENST00000265734	Cyclin-dependent kinase 6	7q21.2	5'..GCAUAAUUAUUCUUAUUCCGUUU. ||||||| 3'..GUCGACGAAAACCCUAAGGCAAC.
DAPK1	ENST00000196730	Death associated protein kinase 1	9q21.33	5'.UCUAGGAAGACAUUUUUCCGUUU. ||||||| 3'.GUCGACGAAAACCCUAAGGCAAC.
TMC7	ENST00000421369	Transmembrane channel-like 7	16p12.3	5'.AAUAUCUCUUAAGAUUUCCGUUA. ||||||| 3' GUCGACGAAAACCCUAAGGCAAC.
SETD1B	ENST00000267197	SET domain containing 1B	12q24.31	5'.UUUGAAGAGAAGUGUUUCCGUUG. ||||||| 3'.GUCGACGAAAACCCUAAGGCAAC.
CALN1	ENST00000329008	Calneuron 1	7q11.22	5'.UGCACCCCUCACCGCUUCCGUUC. ||||||| 3'.GUCGACGAAAACCCUAAGGCAAC.
TMOD2	ENST00000249700	Tropomodulin 2	15q21.2	5'.UCAAAGUUCGUUCAUUUCCGUUA. ||||||| 3' GUCGACGAAAACCCUAAGGCAAC.

MiR-191 expression confirmation

At the present study, TrueQuant-based ncRNA outcomes showed miR-191 to have a high expression value in 7 paired samples of BRCA and NAT, with the *P*value= 0.0001; therefore, it was focused to confirm expression value using the qRT-PCR technique in 120 cases with the BRCA (Fig. 4I). The qRT-PCR results confirmed the miR-191 expression value in the BRCA cells to be higher significantly, compared with NAT, with the *P*value=0.003(Fig. 4II). In addition, a comparison of the miR-191 expression levels between the BRCA with and without lymph node metastasis was made. According to the results, the miR-191 expression levels in the BRCA with lymph node metastasis were significantly higher compared to those without, with Mann-Whitney test *P*=0.0012 (Fig. 4III). Moreover, the Kaplan-Meier curve

was designed to show the prognostic relationship between the miR-191 expression level and overall patient survival (OS). It was found that there was a significant correlation between them, with *P*value=0.001 (Fig. 4IV).
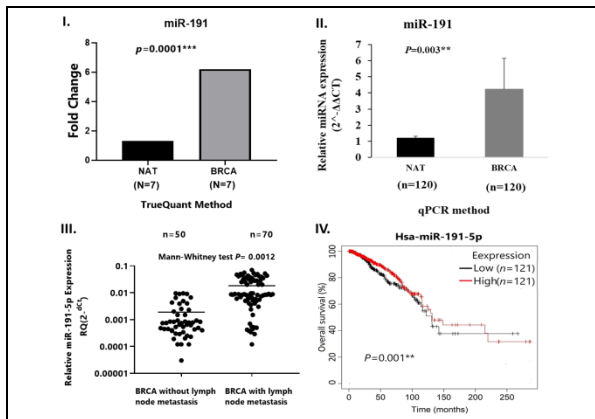
Analysis of the precursor-sequence of miR-191

To detect single nucleotide polymorphism (SNP) in the miRNA-191 gene, 25% of paired samples of BRCA and NAT were sequenced. It was found that the location of the miR-191 gene is sited on chromosome 3 at 3p21.31, NC_000003.12 based on the Entrez gene in the human genome (Fig. 5A). MiR-191 is found in a cluster with miR-425, separated by 384 nucleotides (nt). MiR-191 and -425 are found in intron 1 of the DALRD3 and NDUFAF3 genes. The cluster microRNAs are located as follows: mir-191: chr3 49058051-49058142 [-] and mir-

425 chr3: 49057581-49057667The miR-191 gene included 92nt and was highly conserved in BRCA and NAT, as shown with star marks. Compared to the human reference sequence (RefSeq), a single SNP, which Cytosine (C) in the BRCA sample was changed into Thymine (T), was discovered in the precursor (pre-) sequence of miR-191at the location of 49,020,689 as compared to the NAT and RefSeq (Fig. 5B). Then, the miR-191 gene was translated into the pre-microRNA. Where mature miR-191-5P (22nt in length) and -3P (22nt in length) were determined in the pre-miRNA-191. Seed region withing mature miR-191-5P and-3p was bolded and colored with a yellow background. The SNP (A) was squared within of the seed region (Fig. 5C).

Figure 4. miR-191-5p expression level comparison in BRCA and NAT. **I. The** miR-191-5p expression levels in BRCA were compared with NAT using the TrueQuant method. **II.** the overexpression (P =0.003) of miR-191 was approved using RT-qPCR in 120 paired samples of BRCA and NAT. **III.** MiR-191-5p expression level in BRCA tissues with lymph as compared with that without. The miR-191-5p expression values were higher significantly in BRCA with lymph node metastasis than that without ($P = 0.0012$). RQ denotes the relative expression level of miR-191-5p to U6 (internal reference gene). **IV.** Prognostic correlation between the miR-191-5p expression levels and overall patient survival (OS). The miR-191-5p expression levels were associated with overall patient survival significantly (P =0.001).

Stem-loop structure of pre-miR-191

Within the nucleus, the primary (pri-) miRNA-191 was processed by a microprocessor complex (Drosha/DGCR8) to generate 92nt stem-loop structures labeled pre-miRNA. Fig. 6 showed a stem-loop structure for miR-191 in BRCA and NAT. It also defined the sites of two possible mature miR-191 transcripts, miR-191-5p in yellow and miR-191-3p in red. The SNP location was also shown in BRCA mature-miR-191-5p.
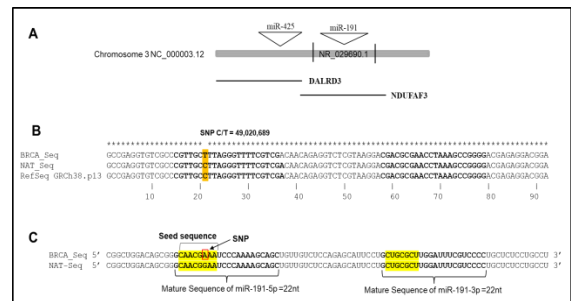
Gene Ontology (GO) enrichment analysis

To determine the biological role of the *CDK6, DAPK1, MTC7, SETD1B, CALN1, and TMOD2* regulated by miR-191-5p in the BRCA, Gene Ontology (GO) analysis was performed using the GenXpro databases. The GO terms of the identified genes were assigned according to biological process (BP), and molecular function (MF). In the BP, the identified target genes played a key role in the cellular process (GO:0009987), biological regulation (GO: 0065007), and developmental process (GO:0032502), as shown in Figure 7A. The minimum number of the target genes were involved in the multicellular organismal process (GO:0032501), metabolic process (GO:0008152), reaction to stimuli (GO:0050896), and signaling (GO:0023052).
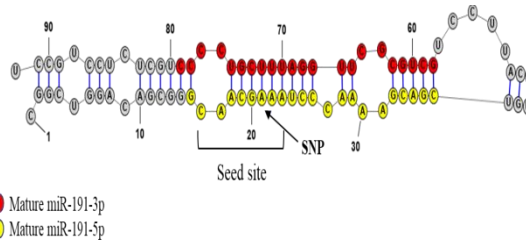

Figure 5. Schematic representation of the miR-191 genes. **A** Genomic location, **B** Comparison of miRNA-191 gene in BRCA and NAT with RefSeq. **C** Precursor-miRNA-191 (92 nt), mature miR-191-5p and 3p in NAT and BRCA.

Figure 6. The pre-miRNA-191 stem-loop structure is depicted in the BRCA. The pre-miR-191 stem's paired strands also include nonhomology areas that interfere with a proper base pairing. An SNP was pointed out within the mature sequence of the miR-191-5p seed sequence colored in the yellow background, and the mature sequence of miR-191-3p is shown in the red background.

However, the maximum number of the target genes in the MF had a great role in transporter activity (GO:0005215), but the minimum number of them had a role in binding (GO:0005488) catalytic activity (GO:0003824) (Fig. 7B).

mRNA-miR-191 interaction

A growing number of evidences has revealed that miR-191 has emerged as an important contributor in biological processes, because it plays a vital role in cancer growth and development, especially lung, prostate, and leukemia. A single nucleotide polymorphism of adenine (A) was found in the seed region of the miR-191-5p in the BRCA samples, it may impact the interaction between miR-191-5p seed site and 3'untranslated region (3'UTR) of *CDK6, DAPK1, MTC7, SETD1B, CALN1, and TMOD2* (Fig. 8A). To study the miR-191-5p effect on the BRCA cells, the biological role of the six genes was recognized as direct targets. The GeneXPro database revealed that miRNA-191-5p promotes several biological processes by targeting these genes. BRCA cell survival, cell growth, cell proliferation and migration. It was found that the genes of *CDK6, DAPK1, SETD1B, CALN1, and TMOD2* can directly influence the BRCA cell growth; whereas, the genes of *DAPK1, CALN1, and MTC7* can directly influence the apoptotic process (Fig. 8B).
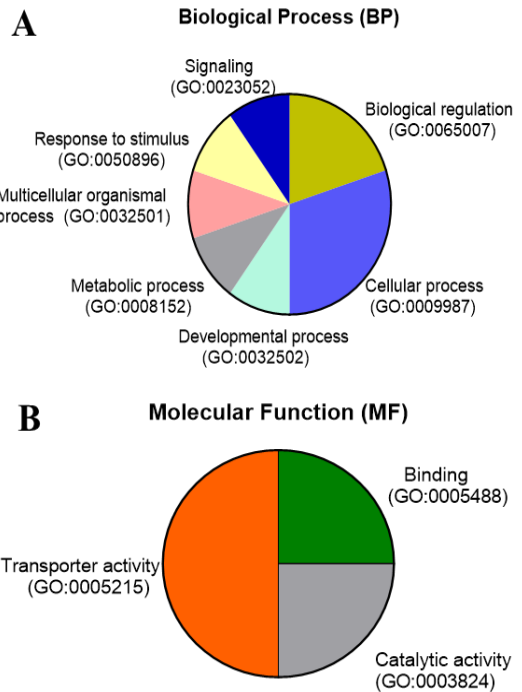

**Figure 7.** Gene Ontology (GO) enrichment analysis; A) Biological Process (PB). B) Molecular Function (MF).

Kaplan–Meier (K-M) survival analysis

The expression level of each *CDK6, DAPK1, MTC7, SETD1B, CALN1*, and *TMOD2* in the BRCA development was examined to demonstrate the probability of a BRCA at a given time surviving. Using the K-M curve, their prognostic level for BRCA cases in a large TCGA dataset was calculated. The expression level of these genes was classified into two groups; such as over- and down-expression. The impact of these genes' high or low expression levels on overall survival (OS) was then measured. According to the results, a prognostic association was observed between the expression level of *CDK6* ($P=0.0001$), *DAPK1*($P=0.03$), *MTC7*($P=0.0004$), *SETD1B*($P=0.001$), *CALN1*($P=0.002$), *and TMOD2*($P=0.05$) (Fig. 9). Our findings suggested that these genes can be used as robust biomarkers for prognosis in patients with the BRCA.

In this study, the techniques of MACE-Seq and TrueQuant for seven paired samples of FFPETs of BRCA and NAT were utilized to profile the differential expression of DEGs and small non-coding RNAs (ncRNAs), respectively. From the MACE-seq workflow, it was found that the total number of DEGs in the BRCA was 25,713 genes, out of which 4,739 were observed to be significantly down-regulated and 3,954 were up-regulated. A study found that the total number of DEGs found in the BRCA was 18,498, with 4114 being up-regulated and 3475 being down-regulated, according to the RNA-seq process (32). However, it was observed that the total number of ncRNA transcripts in the BRCA was 1,450 out of which 670 were up-regulated, and 780 were down-regulated.
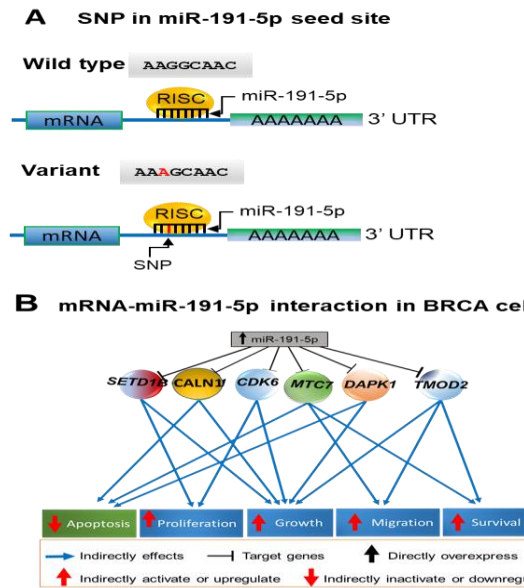
Figure 8. MiR-191-5p and mRNA interaction. **A)** A complementarity of miR-191-5p seed sequence with the target gene was shown in both NAT and BRCA. An SNP (A) may affect the base pairing of miR-191-5p with its target in the BRCA. **B)** The biological function of target genes regulated by miR-191-5p was computationally determined in the BRCA cell.
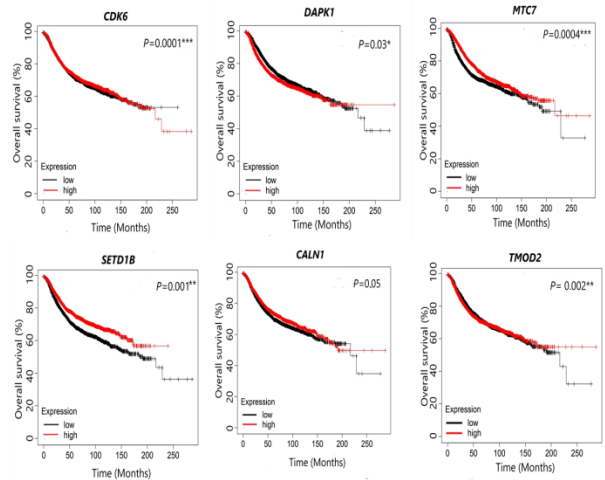
Figure 9. Kaplan Meier curve analysis for the target genes regulated by miR-191-5p. A correlation of the expression of *CDK6, DAPK1, MTC7, SETD18, TMOD2,* and *CALN1* and overall patient survival was shown, with *P*-value < 0.05. The expression levels of these genes were significantly correlated with overall patient survival significantly.

Among the 1,450 ncRNAs, 227 microRNAs were differentially expressed; 22 microRNAs were up-regulated, and 53 were down-regulated significantly. According to recent research, small non-coding RNAs, such as microRNAs and transfer RNA fragments (tRFs), are differentially expressed in the BRCA and may function as a regulatory player in the biological processes (33,34). For example, microRNAs have recently been discovered to have a significant role in various molecular mechanisms that drive the start and development of many cancer types. The genesis of abnormal microRNA expression in human malignancies and their involvement was examined in tumor metastasis, as well as whether they function as oncogenes or tumor suppressors. They may have a great impact on the cancer cell biological processes, including cell growth, migration, proliferation, invasion, and metastasis (35).

Among the 22 microRNAs, the miR-191 expression level in BRCA tissues was higher than the NATs significantly; thus, this microRNA was focused on confirming using RT-qPCR in 120

paired samples of BRCA and NAT to identify the miR-191 as a biomarker for early diagnosing the BRCA. More than 20 distinct cancers have been linked to miR-191, making it a well-known miRNA like miR-155 or miR-21. miR-191 overexpression in 16 cancer types [breast cancer (female), lung, colon, stomach, bladder, prostate, pancreas, liver, ovarian cancer, oral squamous carcinoma, anaplastic large cell lymphoma, esophageal squamous carcinoma, pituitary adenoma, osteosarcoma, B-ALL, and acute myeloid leukemia (AML)] was recorded; but, Its levels are reported to be downregulated in 5 additional cancers (severe medulloblastomas, male breast cancer, retinoblastoma, thyroid follicular tumor, and melanoma) (36). Another study reported that microRNAs, such as miR-191, -21 and -152b promote breast cancer biological processes by regulating their targets genes (37). In the hypoxic milieu, miR-191 promotes breast cancer migration through complicated modulation of TGF signaling, according to a study (38).

Within the seed sequence of the mature miR-191-5p of the BRCA, A single nucleotide polymorphism (SNP) (C>T) was newly identified. SNPs in miRNA genes and miRNA-associated pathways (miR-SNPs) were reported to impair miRNA biogenesis and modulate miRNA-mRNA target interactions, resulting in substantial impacts on gene expression and biological processes of a cell. A study showed that A SNP in MDM4's 3′UTR resulted in the emergence of an illegitimate miR-191 target location, lowering MDM4 expression and dramatically slowing the development of ovarian cancer and tumor-related mortality (39). In this study, six putative target genes; including *CDK6, DAPK1, MTC7, SETD1B, CALN1, and TMOD2* were also identified for miR-191-5p using more than ten computational prediction programs. These genes were negatively associated with that of miR-191-5p. A study showed that the genes of *BDNF*, *CEBPB*, *SATB1*, *SOX4*, and *PLCD1* were

experimentally identified as direct miR-191-5p targets. It was found that these genes were differentially expressed (40). The MACE-Seq findings showed that the expression level of *CDK6, MTC7, DAPK1, SETD1B, CALN1, and TMOD2* were down-regulated in the BRCA cells significantly when compared with the NATs. Some studies showed that miR-29c, miR-1299, and Let-7 enhanced BRCA progression via targeting the *CDK6* gene (41–43). Another study found that miR-141 and miR-26-5p regulated *DAPK1* to stop apoptosis in intestinal mucosal cells and ovarian cancer (44,45).

Precision medicine has recently received too much attention, and there is an increasing need for prognostic biomarkers to be identified. The miR-191-5p was characterized as an oncogenic player in cancer formation and progression and is expressed differentially in cancerous and non-cancerous tissues (37,38). Kaplan-Meier (KM) survival showed a clinicopathological relationship between the miR-191 expression and patients' overall survival. A recent study showed that microRNA-191 overexpression might be applied as an independent prognostic marker for poor prognosis in prostate cancer (46). Another previous study revealed that a poor prognosis in colorectal cancer was clinicopathologically correlated with miRNA-191 expression by controlling the tissue inhibitor of metalloprotease 3.

Gene ontology categories of select target transcripts for the miR-191-5p showed that the genes of *CDK6*, *DAPK1*, *MTC7*, *SETD1B*, *CALN1*, and *TMOD2* played a key role in enhancing BRCA formation and are negatively correlated with that of miR-191-5p in cancerous and non-cancerous tissues because these target genes are involved in several biological processes, and molecular functions. An additional work disclosed the *CDK6* gene to be regulated by miR-191-5p to inhibit apoptosis in BRCA (19). Another study reported that the *DAPK1* downregulation could promote cancer

cell growth and metastasis in several types of cancer (47,48). Another research found that the *SETD1B* down expression was recognized to play a key role in the development of several cancer types (49,50). Moreover, *CALN1* was found to be a miR-675 target gene, and its expression was shown to be inversely associated with miR-675 expression in human gastric cancer. The *CALN1* gene and its related miR-191 serve as oncogenes, encouraging cell proliferation and malignant transformation (51).

Conclusion

To the best of the author's understanding, the present research is the first to show that miR-191 is overexpressed in BRCA. It can negatively and positively interact with several genes that are key in tumorigenesis and breast cancer progression. The direct targeting of *CDK6, DAPK1, MTC7, SETD1B, CALN1, and TMOD2* by miR-191-5p resulted in some of the significant impact of miR-191 on breast cancer cells. The understanding of miRs' molecular mechanisms in breast cancer has grown, suggesting that miR-191 can be a potential biomarker for early detection of breast cancer in the future.

Acknowledgments

Funding

Interest Conflict

The author declares no conflict of interest. MS is a teacher of Salahaddin University, Erbil, a subsidiary of the ministry of higher education in the Kurdistan Region Government (KRG). The researcher provided the fund for this study.

Authors contribution

SM conducted the project.

References

1. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast cancer. Lancet 2005;365:1727–41.

2. Loh H-Y, Norman BP, Lai K-S, Rahman NMANA, Alitheen NB, Osman MA. The Regulatory Role of MicroRNAs in Breast Cancer. Int J Mol Sci 2019;20.

3. Majed SO, Mustafa SA. MACE-Seq-based coding RNA and TrueQuant-based small RNA profile in breast cancer: tumor-suppressive miRNA-1275 identified as a novel marker. BMC Cancer 2021;21:473.

4. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209–49.

5. Cao W, Chen H-D, Yu Y-W, Li N, Chen W-Q. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chin Med J (Engl) 2021;134:783–91.

6. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424.

7. Toda H, Seki N, Kurozumi S, Shinden Y, Yamada Y, Nohata N, et al. RNA-sequence-based microRNA expression signature in breast cancer: tumor-suppressive miR-101-5p regulates molecular pathogenesis. Mol Oncol 2020;14:426–46.

8. Bhaskaran M, Mohan M. MicroRNAs: History, Biogenesis, and Their Evolving Role in Animal Development and Disease. Vet Pathol 2013;51:759–74.

9. Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. Nat Rev Mol Cell Biol 2019;20:21–37.

10. Mendell JT, Olson EN. MicroRNAs in Stress Signaling and Human Disease. Cell 2012;148:1172–87.

11. Iorio M V, Ferracin M, Liu C-G, Veronese A, Spizzo R, Sabbioni S, et al. MicroRNA gene expression deregulation in human breast cancer. Cancer Res 2005;65:7065–70.

12. Di Leva G, Piovan C, Gasparini P, Ngankeu A, Taccioli C, Briskin D, et al. Estrogen Mediated-Activation of miR-191/425 Cluster Modulates Tumorigenicity of Breast Cancer Cells Depending on Estrogen Receptor Status. PLOS Genet 2013;9:e1003311.

13. Qian B, Katsaros D, Lu L, Preti M, Durando A, Arisio R, et al. High miR-21 expression in breast cancer associated with poor disease-free survival in early stage disease and high TGF-β1. Breast Cancer Res Treat 2009;117:131–40.

14. Wang H, Tan Z, Hu H, Liu H, Wu T, Zheng C, et al. microRNA-21 promotes breast cancer proliferation and metastasis by targeting LZTFL1. BMC Cancer 2019;19:1–13.

15. Iorio M V, Casalini P, Piovan C, Di Leva G, Merlo A, Triulzi T, et al. microRNA-205 regulates HER3 in human breast cancer. Cancer Res 2009;69:2195–200.

16. Korpal M, Ell BJ, Buffa FM, Ibrahim T, Blanco MA, Celià-Terrassa T, et al. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. Nat Med 2011;17:1101–8.

17. Volinia S, Calin GA, Liu C-G, Ambs S, Cimmino A, Petrocca F, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci 2006;103:2257 LP – 2261.

18. Ray J, Haughey C, Hoey C, Jeon J, Murphy R, Dura-Perez L, et al. miR-191 promotes radiation resistance of prostate cancer through interaction with RXRA. Cancer Lett 2020;473:107–17.

19. Sharma S, Nagpal N, Ghosh PC, Kulshreshtha R. P53-miR-191-SOX4 regulatory loop affects apoptosis in breast cancer. Rna 2017;23:1237–46.

20. Sharma S, Rajendran V, Kulshreshtha R, Ghosh PC. Enhanced efficacy of anti-miR-191 delivery through stearylamine liposome formulation for the treatment of breast cancer cells. Int J Pharm 2017;530:387–400.

21. Peng W-Z, Ma R, Wang F, Yu J, Liu Z-B. Role of miR-191/425 cluster in tumorigenesis and diagnosis of gastric cancer. Int J Mol Sci 2014;15:4031–48.

22. Zhang X-F, Li K, Gao L, Li S-Z, Chen K, Zhang J-B, et al. miR-191 promotes tumorigenesis of human colorectal cancer through targeting C/EBPβ. Oncotarget 2015;6:4144.

23. Elyakim E, Sitbon E, Faerman A, Tabak S, Montia E, Belanis L, et al. hsa-miR-191 is a candidate oncogene target for hepatocellular carcinoma therapy. Cancer Res 2010;70:8077–87.

24. He Y, Cui Y, Wang W, Gu J, Guo S, Ma K, et al. Hypomethylation of the hsa-miR-191 locus causes high expression of hsa-mir-191 and promotes the epithelial-to-mesenchymal transition in hepatocellular carcinoma. Neoplasia 2011;13:841-IN23.

25. Tian F, Yu C, Wu M, Wu X, Wan L, Zhu X. MicroRNA-191 promotes hepatocellular carcinoma cell proliferation by has_circ_0000204/miR-191/KLF6 axis. Cell Prolif 2019;52:e12635.

26. Zajac BK, Amendt J, Horres R, Verhoff MA, Zehner R. De novo transcriptome analysis and highly sensitive digital gene expression profiling of Calliphora vicina (Diptera: Calliphoridae) pupae using MACE (Massive Analysis of cDNA Ends). Forensic Sci Int Genet 2015;15:137–46.

27. Hafner M, Landgraf P, Ludwig J, Rice A, Ojo T, Lin C, et al. Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. Methods 2008;44:3–12.

28. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.

29. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166–9.

30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:1–21.

31. Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. F1000Research 2016;5.

32. Tripathi R, Soni A, Varadwaj PK. Integrated analysis of dysregulated lncRNA expression in breast cancer cell identified by RNA-seq study. Non-Coding RNA Res 2016;1:35–42.

33. Dhahbi JM, Spindler SR, Atamna H, Boffelli D, Martin DIK. Deep sequencing of serum small RNAs identifies patterns of 5′ tRNA half and YRNA fragment expression associated with breast cancer. Biomark Cancer 2014;6:BIC-S20764.

34. Koi Y, Tsutani Y, Nishiyama Y, Ueda D, Ibuki Y, Sasada S, et al. Predicting the presence of breast cancer using circulating small RNAs, including those in the extracellular vesicles. Cancer Sci 2020;111:2104–15.

35. Di Leva G, Croce CM. Roles of small RNAs in tumor formation. Trends Mol Med 2010;16:257–67.

36. Nagpal N, Kulshreshtha R. miR-191: an emerging player in disease biology. Front Genet 2014;5:99.

37. MAR-AGUILAR F, LUNA-AGUIRRE CM, MORENO-ROCHA JC, ARAIZA-CHÁVEZ J, Trevino V, RODRÍGUEZ-PADILLA C, et al. Differential expression of miR-21, miR-125b and miR-191 in breast cancer tissue. Asia-Pacific J Clin Oncol 2013;9:53–9.

38. Nagpal N, Ahmad HM, Chameettachal S, Sundar D, Ghosh S, Kulshreshtha R. HIF-inducible miR-191 promotes migration in breast cancer through complex regulation of TGFβ-signaling in hypoxic microenvironment. Sci Rep 2015;5:1–14.

39. Lin J, Zandi R, Shao R, Gu J, Ye Y, Wang J, et al. A miR-SNP biomarker linked to an increased lung cancer survival by miRNA-mediated down-regulation of FZD4 expression and Wnt signaling. Sci Rep 2017;7:1–11.

40. Yang L, Li Q, Wang Q, Jiang Z, Zhang L. Silencing of miRNA-218 promotes migration and invasion of breast cancer via Slit2-Robo1 pathway. Biomed Pharmacother 2012;66:535–40.

41. Liu L, Tian Q, Liu J, Zhou Y, Yong H. Upregulation of hsa_circ_0136666 contributes to breast cancer progression by sponging miR-1299 and targeting CDK6. J Cell Biochem 2019;120:12684–93.

42. Zhang Y, Zhang H, Kang H, Huo W, Zhou Y, Zhang Y. Knockdown of long non-coding RNA HOST2 inhibits the proliferation of triple negative breast cancer via regulation of the let-7b/CDK6 axis. Int J Mol Med 2019;43:1049–57.

43. Zhang H, Zhao B, Wang X, Zhang F, Yu W. LINC00511 knockdown enhances paclitaxel cytotoxicity in breast cancer via regulating miR-29c/CDK6 axis. Life Sci 2019;228:135–44.

44. Zhou B, Zhang W, Yan Z, Zhao B, Zhao J, Feng W, et al. MicroRNA-26b-5p Targets DAPK1 to Reduce Intestinal Ischemia/Reperfusion Injury via Inhibition of Intestinal Mucosal Cell Apoptosis. Dig Dis Sci 2021:1–12.

45. Li D, Xu D, Xu Y, Chen L, Li C, Dai X, et al. MicroRNA-141-3p targets DAPK1 and inhibits apoptosis in rat ovarian granulosa cells. Cell Biochem Funct 2017;35:197–201.

46. Liu J, Yan Y, Shi J, Wu Y, Li Y, Dai L, et al. Upregulation of microRNA-191 can serve as an independent prognostic marker for poor survival in prostate cancer. Medicine (Baltimore) 2019;98.

47. Zhao J, Zhao D, Poage GM, Mazumdar A, Zhang Y, Hill JL, et al. Death-associated protein

kinase 1 promotes growth of p53-mutant cancers. J Clin Invest 2015;125:2707–20.

48. Park G-B, Jeong J-Y, Kim D. Gliotoxin enhances autophagic cell death via the DAPK1-TAp63 signaling pathway in paclitaxel-resistant ovarian cancer cells. Mar Drugs 2019;17:412.

49. Neto JNM, do Amaral Rabello D, Pittella-Silva F. Expression profiling of SETD family genes in breast cancer patients 2016.

50. Lindner P, Paul S, Eckstein M, Hampel C, Muenzner JK, Erlenbach-Wuensch K, et al. EMT transcription factor ZEB1 alters the epigenetic landscape of colorectal cancer cells. Cell Death Dis 2020;11:1–13.

51. Li H, Yu B, Li J, Su L, Yan M, Zhu Z, et al. Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. Oncotarget 2014; 5:2318.