Oncology Treatment Discovery, 2024, Volume 2, Issue 3
http://ojs.bbwpublisher.com/index.php/OTD
Online ISSN: 2981-8079

Long Non-coding RNA and Progression of Breast Cancer

Usama Ahmed[1]*, Muhammad Abubakar[2], Salma Saeed Khan[3], Baqa ur Rehman[2]

[1]Department of Medicine, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen 518060, China
[2]Department of Biosciences, COMSATS University Islamabad, Park Rd, Islamabad Capital Territory 45550, Pakistan
[3]Biomedical and Life Sciences, Kohsar University Murree, Murree 47150, Punjab, Pakistan

Corresponding author: Usama Ahmed, usama.med.19@gmail.com

Abstract: Breast cancer is a major problem in the health sector worldwide. The prevalence of breast cancer incidence is on the rise in many countries despite efforts to eradicate it. The substantial efforts for the treatment of breast cancer are quite successful but the main hindrance is early detection which in turn is dependent on various factors i.e. genetic history, environmental exposures, hormonal causes, and sedentary lifestyle. In previous studies, lncRNAs were known to modulate the progression of many cancers like gastric, lung, ovarian, and colorectal cancers. This study was designed to find these lncRNAs role in breast cancer development. Bioinformatics tools were used for in silco analysis of these lncRNAs LINC01356, LINC01357, ARF4-AS1, and FAM83B. These lncRNAs were evaluated in GEPIA2 for subtype analysis and survival probability, Phylo CSF UCSC genome browser for conserved sequences and protein-coding potential, BcGenExMiner for correlation analysis and UALCAN for protein-coding gene expression. These above-mentioned lncRNAs showed high expression in the basal subtype of TNBC as compared with other subtypes and they showed poor survival rates in the basal subtype. LINC01356 and LINC01357 showed positive correlation with TAFA3 and SLC16A1. ARF4-AS1 positively correlated with ASB14 and FAM83B positively correlated with HCRTR2. These lncRNAs showing positive correlation with different genes have a role in breast cancer development and progression.

Keywords:

Online publication: October 2, 2024

1. Introduction

Breast cancer is a complex and heterogeneous type of cancer and is a highly prevalent cancer among females exhibiting a high mortality rate whereas its frequency is low in males approximately 1%. It depends on multiple factors including genetic history, radiation exposure, environmental carcinogens, poor diet etc. [1]. Cancer is globally prevalent but differs in the incidence rate and mortality rate due to differences in genetic makeup, ethnicity and living standards [2,3]. Breast cancer involves multiple risk factors, such as some of which are modifiable risk factors that can be altered to reduce the incidence of breast cancer like dietary uptake, environmental exposure,

and sedentary lifestyle and some are non-modifiable risk factors which are unavoidable like genetic history and demographic variables which all lead towards progression of breast cancer [4,5].

Long non-coding RNAs (lncRNAs) are non-coding RNA molecules of length greater than 200 nucleotides. About 98% human genome is transcribed into noncoding RNAs and only 2% is transcribed into protein coding transcript [6]. These RNAs are involved in the regulation of genetic expression through different post-transcriptional and epigenetic mechanisms. Their functionality involves housekeeping, genomic imprinting, translation, RNA degradation and protein kinetics [7,8]. lncRNA is categorized into five categories depending upon its origin in the genome including sense lncRNA, antisense lncRNA, bidirectional lncRNA, intronic lncRNA and intergenic lncRNA. lncRNAs are comparatively high in concentration in the nucleus as compared to mRNA whereas in cytoplasm, they are in lesser concentration associated with ribosomes [9]. lncRNAs biogenesis involves DNA elements i.e., promoters, enhancers, and intergenic regions through which they are transcribed [10]. Many mechanisms are undergoing their synthesis such as the breakage of ribonuclease P (RNaseP) to form mature ends, synthesis of circular structures and formation of small nucleolar RNA (snoRNA) [11].

Initially, it was considered as nonfunctional component of RNA polymerase but later on, it has shown diversified functionality such as metabolic pathways, cellular developmental processes, imprinting, cell cycle regulation, transcription, translation, splicing, and mRNA decay. There are three levels of mRNA regulation. Epigenetic regulation, transcriptional and post-transcriptional regulation [12]. The genetic expression is usually regulated by various epigenetic factors like DNA methyl transferases and chromatin-modifying complexes and during tumor progression epigenetic silencing inactivates tumor suppressive genes [13]. So, lncRNA mainly targets this silencing to regulate processes like ANRIL, H19, and HOTAIR, achieve repress their transcriptional activity by chromatin-remodeling or histone-modifying proteins. lncRNAs can perform their function by interacting with RNA, DNA, protein and metal ions and forming tertiary structures [14]. This lncRNA action is categorized into five subtypes based on interacting molecules and effects i.e., lncRNA interacts with protein, DNA, and RNA and it can also act as a catalyst which results in its various biological roles [15].

lncRNAs are involved in various human diseases and their progressions and breast cancer is one of them. lncRNAs such as XIST, H19 and MALATI were most comprehensively studied for their role in breast cancer progression [16]. XIST is downregulated lncRNA found in many cancer cell lines including breast tumors by suppressing AKT signaling so it is a tumor suppressor. H19 is expressed in luminal A and luminal B subtypes of breast cancer. It works by promoting proliferation and hindering apoptosis in cancerous cells so acts as a tumor progressor. MALAT1 is upregulated in luminal A, luminal B and HER2 subtypes, where it promotes cancer cell invasion and migration and targeting it slows tumor growth and metastasis so it also promotes cancer progression [17]. In the present study, differentially expressed lncRNAs were selected including LINC01356, LINC01357, ARF4-AS1, and FAM83B. The study is focused on the expression of lncRNAs and their role in triple negative breast cancer. This study explored their effect on breast cancer progression through correlation with protein-coding genes to elucidate their role in breast cancer.

2. Experimental plan

This study was designed to investigate the genetic expression of long non-coding RNA expression in breast cancer. It was divided into two parts of wet lab and dry lab. For this purpose, post-surgical breast cancer samples were collected for RNA isolation. Then due to COVID-19, the study moved toward in silico work. For in silico work, different bioinformatics tools were used to explore different features of LncRNA to predict their role in the progression of breast cancer.

2.1. Wet lab

The present study was first approved by the Ethical Review Committee of COMSATS University Islamabad. The samples that were collected for this study from various tertiary care hospitals involved blood samples of breast cancer patient and post-surgical normal tissue and tumor tissue. Before sample collection, the consent form was signed. After surgery, fresh tumor sample of size almost 5–6 mm was excised and a control sample of about 2 cm was removed from the surrounding region of the same patient. Tissue and blood was taken from the patient ignoring age, gender, and stage. Post-surgery samples were immediately shifted to RNA later® and were stored in a refrigerator at 2–8 ℃. The sample was transferred from the hospital in an ice box along with ice packs and maintained at a temperature of 2–8 ℃. Samples were still in RNA later® before RNA extraction so to preserve RNA integrity. Samples were removed from RNA later® and transferred to a petri dish having 1 mL chilled PBS DEPC. Trizol reagent method was used to extract RNA from samples. Surgical blades separate fats from tissue by cutting tissue into small pieces. These pieces were transferred to a glass homogenizer with 1mL PBS. After some time as crushing proceeds, 10 µL of Betamerceptaethanol was added until the tissue softened. Then, 100 µL of the mixture was taken into Eppendorf tubes for extraction of DNA. After this, 1 mL triazole reagent was added to the remaining sample and further crushed for RNA extraction. When the sample was homogenized and it gave a pinkish milky appearance it was transferred in the Eppendorf tube then 200 µL chloroform was added for protein degradation and incubated at room temperature for 5 minutes after inversion so to separate nuclear protein. After incubation centrifugation of labeled Eppendorf containing crushed tissue was done at 12,000 rpm for 15 minutes at 4 ℃. After centrifugation, three layers were formed. After which, the upper aqueous layer was transferred into newly labeled Eppendorf tubes for further RNA treatment.

After this step, 2 µL of glycogen and then an equal volume of isopropanol to separate the aqueous layer was added in a labeled Eppendorf tube. Then, sodium acetate which is equal to the volume of $1/10^{th}$ aqueous layer was added so that the pellet became visible and tubes were stored at -20 ℃ overnight. Next day, the study thawed the mixture and then centrifugation was done at 12,000 rpm at 4 ℃ for 10–15 minutes. Now in the Eppendorf tube, the RNA pellet was carefully observed at the bottom of tube and the supernatant was removed by placing tip opposite to the pipette. 200 µL of 70% ethanol was added in each Eppendorf tube for washing the pellet. Again, the tubes were spined at 4 ℃ for 5 minutes at 12,000 rpm. After spinning, the supernatant was carefully removed with a pipette without touching the pellet and the leftover pellet was dried at a dry heater (thermoblock) at 70 ℃ for 1 minute along with DEPC water. The dried pellet was dissolved in nuclease-free water (DEPC). After that, 35–60 µL DEPC was used for pellet resuspension according to the size of the pellet, then vortex Eppendorf tubes by moving them on a rack. A short spin was done for 15 seconds, then again dry heat for 45 seconds at 70 ℃. Then, they were shortly spun for 15 seconds and were stored at 20 ℃ for cold shock.

RNA integrity was confirmed by the technique of denaturing gel electrophoresis. 1.2% agarose was weighed and dissolved in 100 mL 1X TBE buffer by heating in a microwave until it was completely dissolved. Then, it was allowed to cool down and 172 µL bleach for denaturing was added. In the next step, 7 µL ethidium bromide was added and moved slowly to mix and poured into a gel tray to solidify. It was prepared by mixing 2 µL of loading dye with 6 µL of formamide and 5 µL RNA sample was mixed with 8 µL master mix. Then, it was short spun and heat shock at 70 ℃ for 1 minute was given followed by a short spin and then placed in a freezer for 1 minute. Samples were then loaded into gel wells and horizontal gel was run first at 300 mA for 10 minutes at 90 V and then run for 35 minutes at 100 V.

Another method utilized for RNA quantification was through a spectrophotometer where optical density was measured and results were given by OD value. Firstly, dilutions of 1:100 were prepared by using 198 µL DEPC water and 2 µL RNA. Then, the machine was turned on for 10 minutes before measuring. The blank

was set to zero after baseline correction. Then, the samples were filled in a cuvette and test sample was run and finally OD values were taken as 260/280 ratio.

2.2. Dry lab / in silico

To find a highly differentially expressed lncRNA, the study utilized multiple online software, such as Ensemble genome browser, Tanric and Lncar, as they are common databases giving phenotype-specific lncRNA. Expression of lncRNAs was screened out and further validated using data from The Cancer Genome Atlas (TCGA) database. Data was taken from the Atlas of lncRNA in cancer (TANRIC) database which was quantified based on $\log_2$RPKM value. Out of multiple candidates, 5 highly expressed candidates were selected, that is, DDIT4-AS1, ASHIL-AS1, LINC01356, LINC01357 and ARF4-AS1. Tissue-specific gene expression and regulation of selected lncRNAs were determined using the Genotype-Tissue Expression database (GTex). This database was utilized to find out what changes occur at the genomic level that led to the conversion of healthy cells into faulty cells that cause various diseases. So, the database was targeted to find out the root cause of the disease and to develop treatment accordingly for improvement in diagnostic and therapeutic hallmarks. By applying filters of the chosen organ, the study checked tissue-specific expression analysis of targeted 5 lncRNAs in breast tissue and whole blood. Overall survival analysis of breast cancer patients was calculated by comparing them with healthy individuals using GEPIA 2 (Gene expression profiling interactive index). GEPIA was previously used but later it was upgraded to GEPIA2 to provide insights with more functioning. It provides a survival analysis graph in comparison between normal individuals and tumor patients and is presented in the form of curves. For breast cancer, BRCA subtype was selected to present survival curves of 5 genes in breast cancer.

GEPIA2 (gene expression profiling interactive analysis) was used to analyze the cancer genome atlas (TCGA). GEPIA is web-based is a web-based tool that provides quick and targeted functions based on TCGA and Gtex data. Expression DIY was targeted to analyze and compare the expression of normal individuals and tumor patients with BRCA cancer type. GEPIA 2 was used for lncRNAs subtype analysis. This study allows choosing a subtype and finding its expression analysis or survival analysis. Then for the breast cancer subtype, subtype filter BRCA was applied and box plot was plotted for normal and tumor differential expression of selected lncRNAs. PhyloCSF was used for comparative differentiation to analyze the coding potential of lncRNA by selecting phyloCSF tracks. PhyloCSF was used for the alignment of multispecies genomes to recognize coding regions in the genome. Similarly, selected genes (lncRNAs) were aligned with 100 vertebrate genome to recognize the protein-coding potential of genes.

lncATLAS is a bioinformatics tool that was used to find lncRNAs subcellular localization because lncRNAs functionality is dependent on their location. For this purpose, it uses various RNA sequencing datasets. Relative concentration index (RCI) was used to find the location of 6768 lncRNAs which were represented in different compartments of 15 cell lines. lncATLAS was approachable through an informative web server, where lncRNA of interest was retrieved through names. Localization was given across cell types and organelles. BcGenExMiner is a recently developed web-based application to evaluate prognostic information of breast cancer genes using a prognostic module known as a correlation module. This module involves three kinds of gene expression correlation analysis. In the first type of analysis, 10 targeted lncRNA correlation coefficient can be computed. In the second type, targeted lncRNA can be compared with 2 sets of genes for correlation coefficient. In the third type, targeted lncRNA can be compared with genes in close vicinity for correlation coefficient. The correlation analysis was performed in molecular subtypes of TNBC (HER2+, luminal A, luminal B and normal). In a gene correlation targeted analysis, 3 to 16 different genes were selected and analyzed in 4 molecular subtypes of BC. Results were displayed in the form of a heat map where the dark

red coefficient is represented as (-1) and dark blue is represented as (+1). The correlation of normal and basal subtypes was compared. The selected LncRNA was correlated with cis and trans-protein-coding genes.

UALCAN is an interactive web portal that was utilized to check the expression of positively correlated cis or trans-protein-coding genes with LncRNA. It allowed users to perform and analyze the relative expression of a protein-coding gene across basal and normal subtypes, as well as in other subtypes like HER2, luminalA and luminal B so to interpret over and under-expressed (up and down-regulated) genes in breast cancer. In silico validation of targeted genes was done and further identification of tumor subgroup specific candidate biomarkers.

3. Results

3.1. RNA extraction

Blood and tissue samples i.e. normal tissue and tumor from breast cancer patients were taken and processed for RNA extraction using TRIZOL reagent. After this, agarose gel was used to check the quality of extracted RNA by visualizing band on the gel. After that, formamide was used to remove the secondary structure of RNA by giving initial denaturation at 70 degrees centigrade. Figure 1 shows intact bands of rRNA of the size of 28SrRNA, 18SrRNA and 5.8SrRNA and it is indicating RNA is not denatured and of good quality.

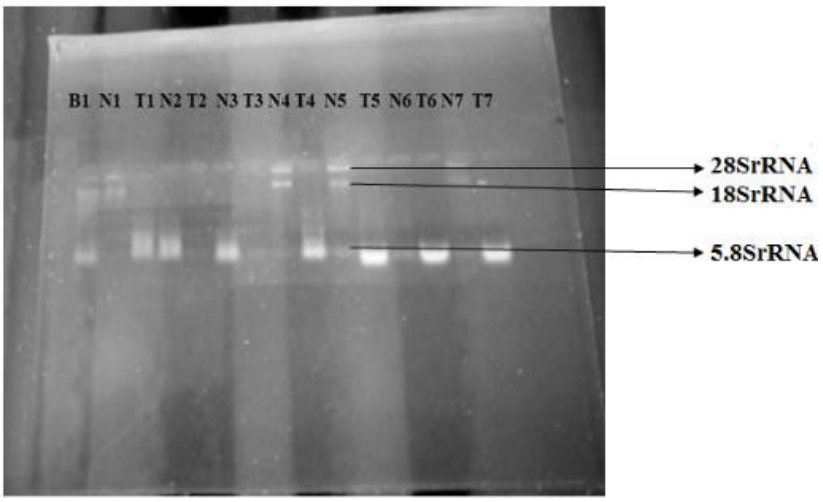


Figure 1. Extraction of RNA from tissue and blood samples from breast cancer patients.

lncRNA having differential expression was selected and shortlisted based on tumor and normal samples database from TCGA and GTEx. They were analyzed by comparing their values between tumor and normal samples. Six potential candidate lncRNA genes were selected based on their fold difference between the two datasets.

Table 1. List of selected lncRNA that differentially expressed in breast cancer

Rank	Transcript ID	Name	F_{pkm} value subset 1 breast cancer/ TNB C	F_{pkm} value subset 2 Breast cancer/ normal TNB C
1	ENSG00000269926	DDIT4-AS1	13.10	2.49
2	ENSG00000235919	ASHIL-AS1	2.83	1.39
3	ENSG00000215866	LINC01356		
4	ENSG00000224167	LINC01357		
5	ENSG00000272146	ARF4-AS1		

3.2. DDIT4-AS1

3.2.1. Subtype analysis of DDIT4-AS1 using GEPIA-2

DDIT4-AS1 is located at chromosome 10q22.1 having 2 exons. DDIT4-AS1 is anti-sense, non-coding RNA with 1 transcript only. Expression analysis among four subtypes of BC was performed using GEPIA2. Results showed that DDIT4-AS1 is significantly upregulated in all four subtypes of breast cancer and can play a role in the progression of all these types.

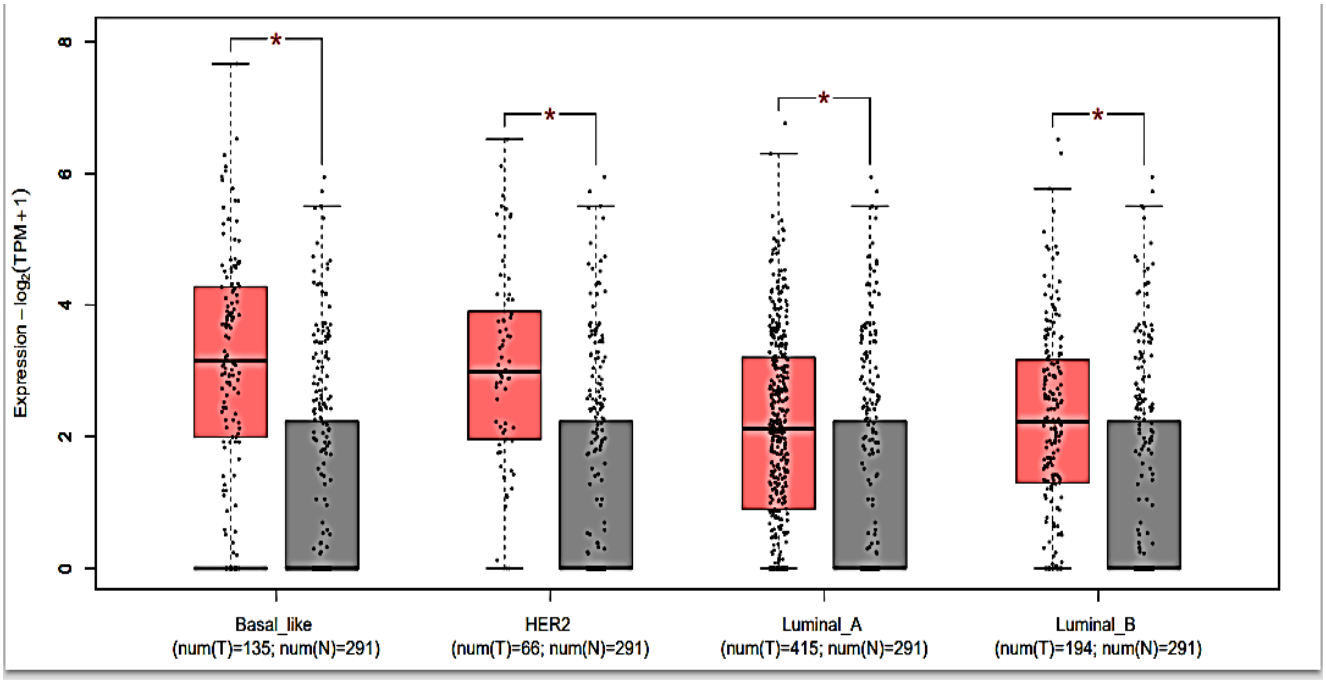


Figure 2. Subtype analysis of DDIT4-AS1 based on TCGA data. The box plot in red represents tumor while in grey represents normal. DDIT4-AS1 expression is upregulated in all four subtypes of BC.

3.2.2. Survival analysis of DDIT4-AS1

The impact of DDIT4-AS1 on basal breast cancer patient survival probability was evaluated using GEPIA2 and showed less difference between the survival probability of the low and high expression group. However, the high expression group showed less survival probability after 250 months showing it is not oncogenic.

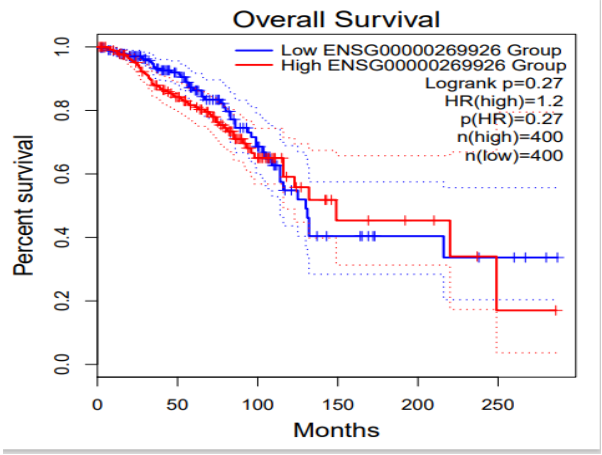


Figure 3. Survival analysis of DDIT4-AS1 using GEPIA-2. The above graph showed less difference between the survival probability of low and high expression groups in the basal subtype of breast cancer. The survival probability is shown on the y-axis whereas the time duration in months is shown on the x-axis. HR represents the hazard ratio.

3.4. ASHIL-AS1

3.4.1. Subtype analysis ASHIL-AS1

ASHIL-ASI is located at chromosome 1q22 having 2 exons. This gene encodes for a non-coding RNA with 2 different transcript isoforms. Tissue specific expression analysis using GTex showed high expression in normal breast-mammary tissue.

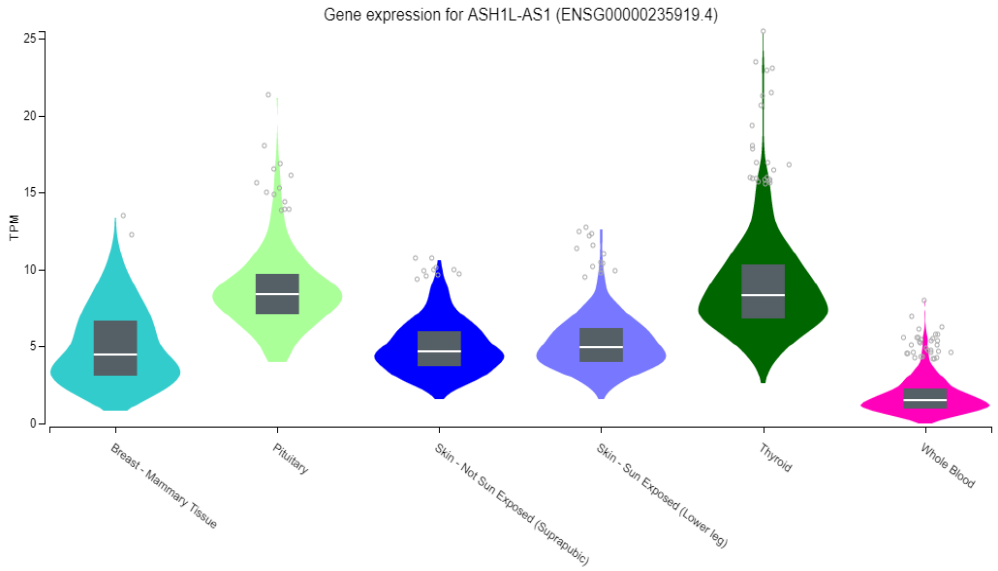
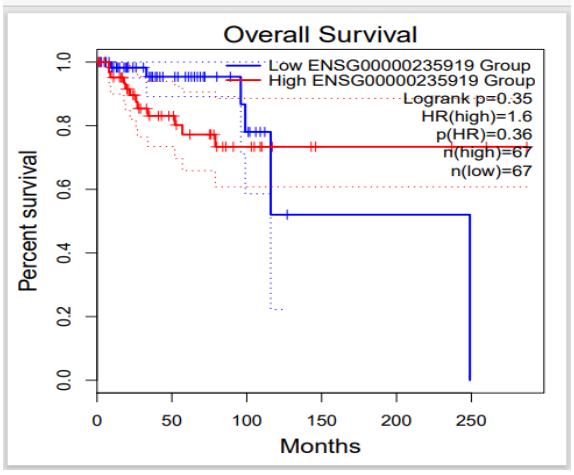


Figure 4. Tissue-specific expression analysis of ASHIL-AS1 Expression analysis showed around 13TPM expression of ASHIL-AS1 in normal breast mammary tissue. Expression values are in transcript per million (TPM). Light Blue color represents breast mammary tissue, mustard represents pituitary, dark blue represents non-exposed skin, light purple represents exposed skin, dark green represents thyroid and pink represents whole blood.

Figure 5. Survival analysis of ASHIL-AS1 using GEPIA-2. The above graph showed a decrease in the survival possibility of the low survival group and indicated a constant level of survival for the high group at 0.8 so it is not oncogenic. The survival probability is shown on the y-axis whereas the time duration in months is shown on the x-axis. HR represents the hazard ratio.



3.4.2. Survival analysis of ASHIL-AS1

The impact of ASHIL-AS1 on basal breast cancer patient survival probability was evaluated using GEPIA2 and showed a decrease in the survival percentage of the low group whereas the high group became constant at 0.8 indicating not oncogenic nature.

3.5. LINC01356

3.5.1. Tissue-based expression analysis of LINC01356

LINC01356 is located at chromosome 1p13.2 having 7 exons. This gene is long intergenic non protein

coding RNA with 5 different transcript isoforms. Tissue specific expression analysis using GTex showed low expression in normal breast mammary tissue.

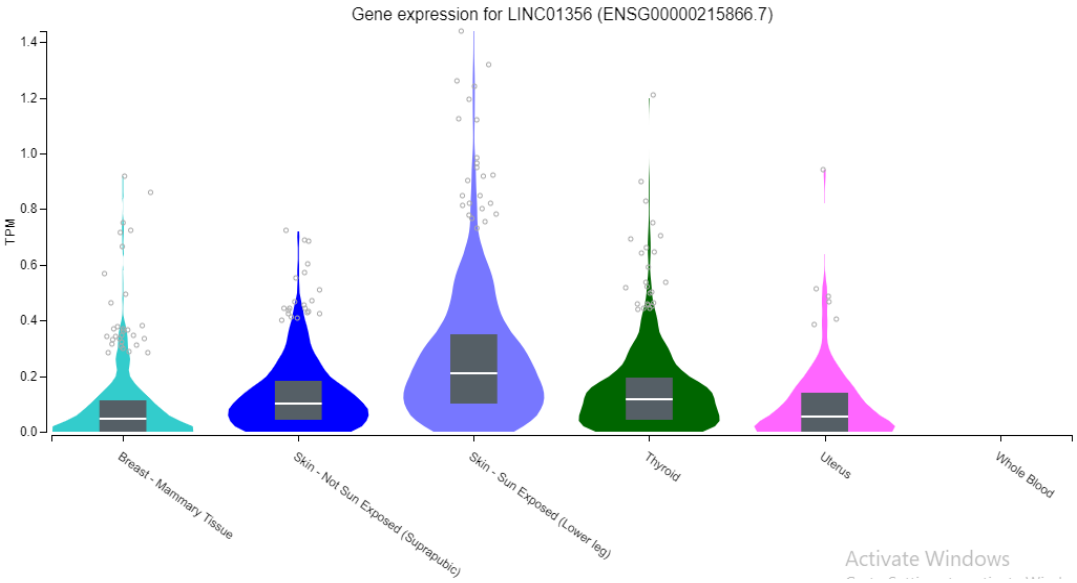


Figure 6. Tissue specific expression analysis of LINC01356 Expression analysis showed insignificant expression of around 0.8 TPM expression of LINC01356 in normal breast mammary tissue. Expression values are in transcript per million (TPM). Light blue shows breast mammary tissue, dark blue shows unexposed skin, light purple shows exposed skin, dark green shows the thyroid, magenta shows the uterus and pink shows whole blood.

3.5.2. Subtype analysis of LINC01356

Expression analysis was performed among four subtypes of BC using GEPIA2. It was calculated by mean value of log2 (TPM+1) in each subtype of breast cancer. Results showed that LINC01356 is upregulated in basal-like breast cancer as compared to the other three subtypes. However, the upregulation of linc01356 is not significant.

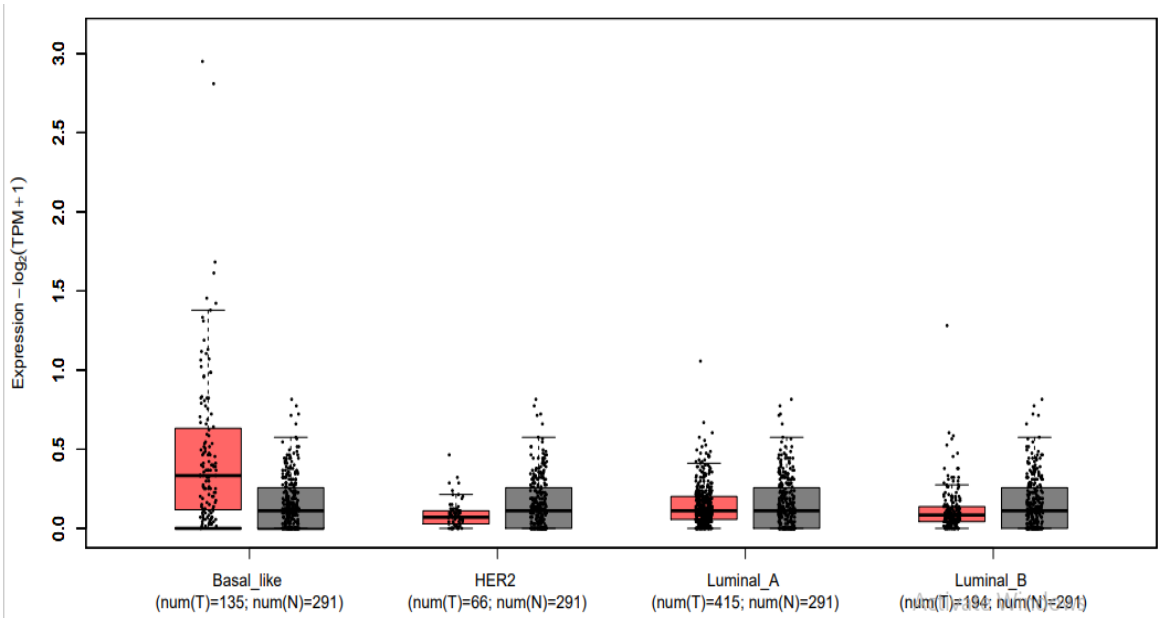
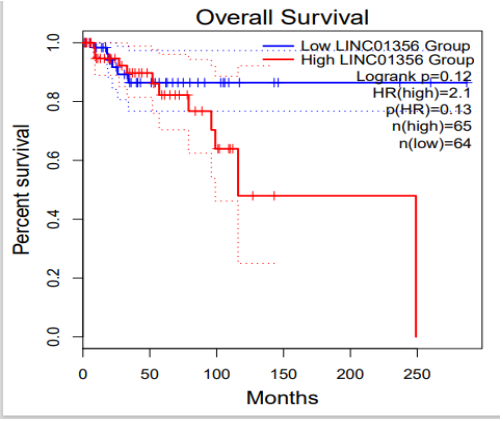


Figure 7. Subtype analysis of LINC01356 based on TCGA data. The box plot in red represents the tumor while in grey represents normal. LINC01356 expression is upregulated in the basal subtype of BC as compared to the other three subtypes.

3.5.3. Survival analysis of LINC01356

LINC01356 is upregulated in the basal subtype survival graph as we evaluated it through GEPIA2. There are two groups i.e. red represents the high expression group whereas blue represents the low expression group. HR represents the hazard ratio which is an indication of the association between different treatments (radiotherapy and chemotherapy) and survival time. The graph shows a decrease in survival rate for 250 months for high survival groups.

Figure 8. Graph showing overall survival analysis of LINC01356. The higher expression group showed low survival probability as it decreases to 0 during 250 months so it is oncogenic. The percent survival is shown on the y-axis and the time duration in months is shown on the x-axis.



3.5.4. Gene structure LINC01356

LINC01356 is a long intergenic non-coding RNA having a size of 1819 bp. It has 5 transcript isoforms. Table 2 shows multiple transcripts of LINC01356 along with their transcript length.

Table 2. List of transcripts of LINC01356

Transcript ID	Name	Bp	Protein	Bio type
ENST00000401018.5	LINC01356-201	1819	No protein	lncRNA
ENST00000661136.1	LINC01356-204	1430	No protein	lncRNA
ENST00000667852.1	LINC01356-205	1222	No protein	lncRNA
ENST00000433505.5	LINC01356-202	767	No protein	lncRNA
ENST00000449572.2	LINC01356-203	702	No protein	lncRNA

3.5.5. Protein coding potential using PhyloCSF

The coding potential of LINC01356 was further evaluated by using PhyloCSF tool uses a calculate the phylogenetic conservation score multi-species nucleotide sequence alignment, which represents a probable protein-coding region. LINC01356 shows all the peaks towards negative axis indicating that it does not code for any protein or shorter peptide.

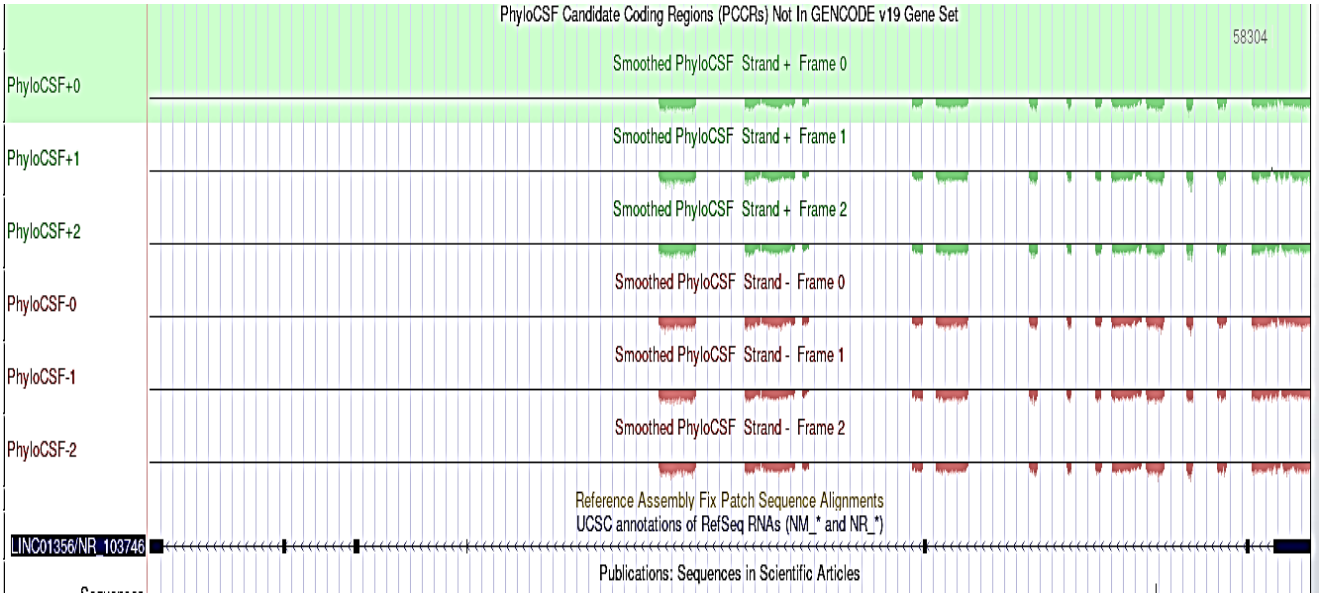


Figure 9. Protein coding potential of LINC01356. The coding potential was evaluated and PhyloCSF. The coding potential was evaluated in three frames indicated as 0, 1and 2. All the peaks lie in the negative axis indicating that LINC01356 has no coding potential.

3.5.6. Conservation of sequence and structure

Few alignments of 100 vertebrates showed little conservation of LINC01356 in rhesus and elephant.

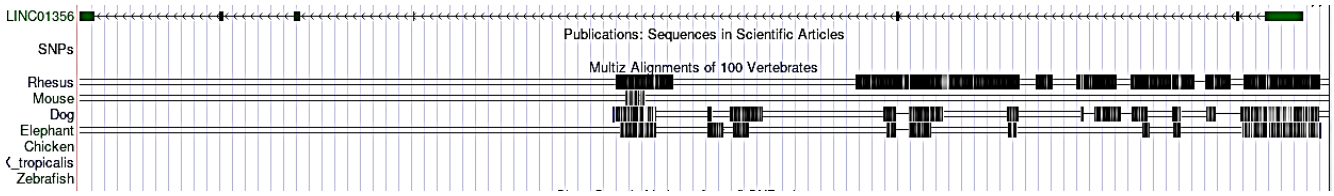


Figure 10. Multiz alignment of 100 vertebrates showing little bit conservation of LINC01356 in rhesus and elephant.

3.5.7. Nucleo-cytoplasmic localization of LINC01356

The nucleo-cytoplasmic localization of LINC01356 was determined by using lnc-ATLAS. MALATI was used as a reference nuclear gene whereas DANCR was used as a reference cytoplasmic gene. LINC01356 showed cytoplasmic localization in different cell lines i.e. H1. HESC and HUVEC and indicating their role in regulating gene expression via RNA-RNA or RNA-protein interaction.

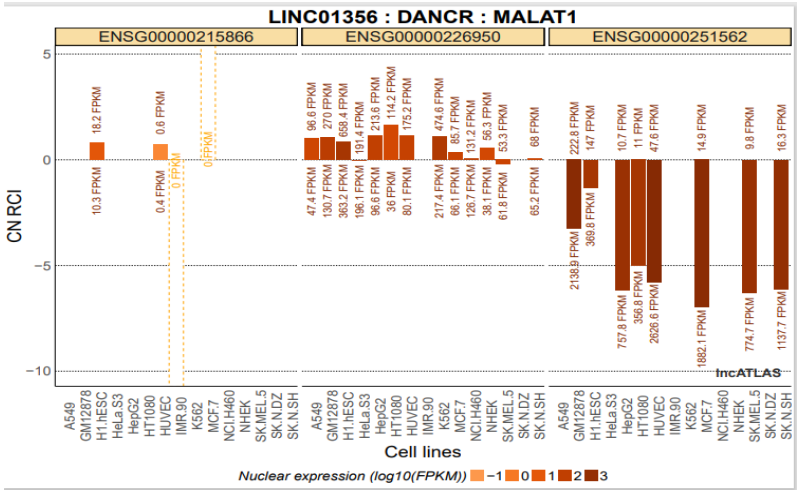


Figure 11. Nucleo-cytoplasmic localization of LINC01356. LINC01356 showed few cytoplasmic localizations in multiple cell lines. MALATI was used as a reference gene for the nucleus and DANCR was used as a reference gene for the cytoplasm. RCI represents the relative concentration index based on the comparison of gene concentration per unit RNA mass between two cellular compartments.

3.5.8. Macromolecular interactions involving LINC01356

LINC01356 can interact with neighboring genes or genes on other chromosomes in a cis or trans manner. Two protein-coding genes (TAFA3 and SLC16A1) which are in close vicinity were selected for correlation expression with LINC01356 (non-coding gene). BcGenExMinor4.5 tool was used to evaluate the correlation expression of coding and non-coding genes between four breast cancer subtypes.

Figure 12. Correlation of Cis interactors TAFA3 and SLC16A1 with LINC01356 predicted by BcGenExMiner4.5 in the normal subtype. TAFA3 showed a strong correlation as compared to SLC16A1 in a normal subtype of breast cancer. Blue color represents a negative correlation whereas red represents a strong positive correlation.

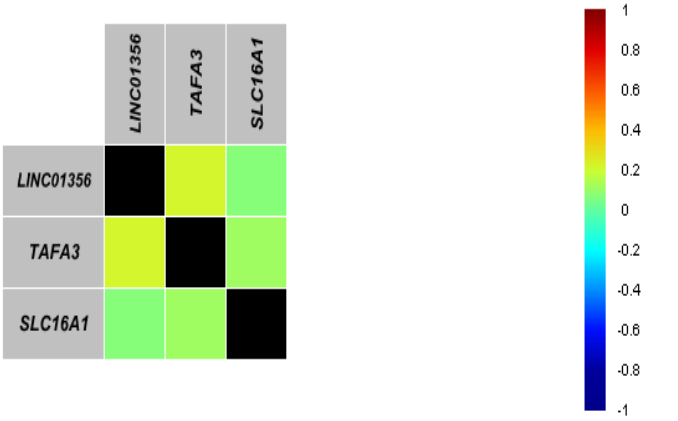
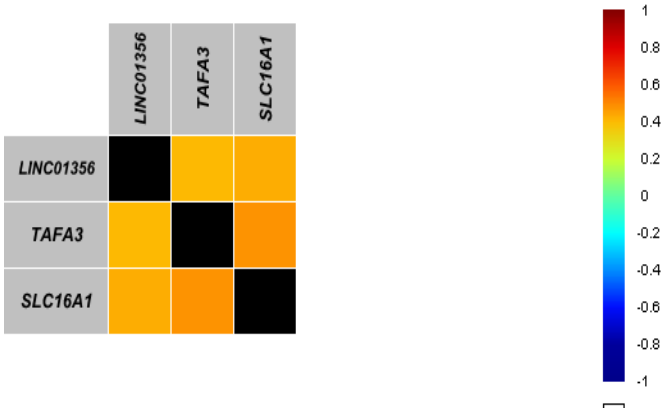
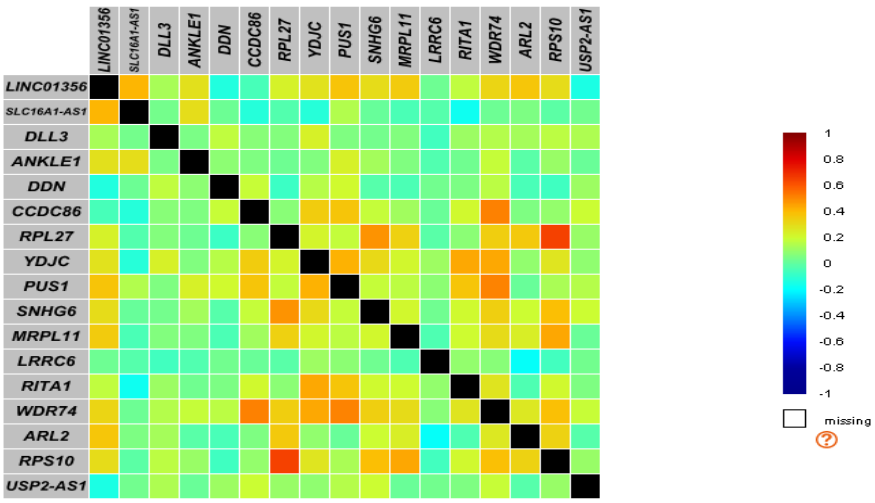


Figure 13. Correlation of Cis interactors TAFA3 and SLC16A1 with LINC01356 predicted by BcGenExMiner4.5 in basal (TNBC) subtype. Both TAFA3 and SLC16A1 showed a strong correlation with LINC01356.



Linc01356 has cytoplasmic localization. Therefore, different trans-acting genes were explored using BcGenExMiner4.5. The following heat map exhibits different trans-interacting genes with positive expression correlation with LINC01356.

Figure 14. Heat map showing the correlation of LINC01356 with different genes in normal breast cancer tissue. SLC16A1-AS1 showed a positive correlation with LINC01356 based on TCGA RNA seq data as predicted by bcExGenMiner4.5. Blue color represents a negative correlation whereas red represents a strong positive correlation.



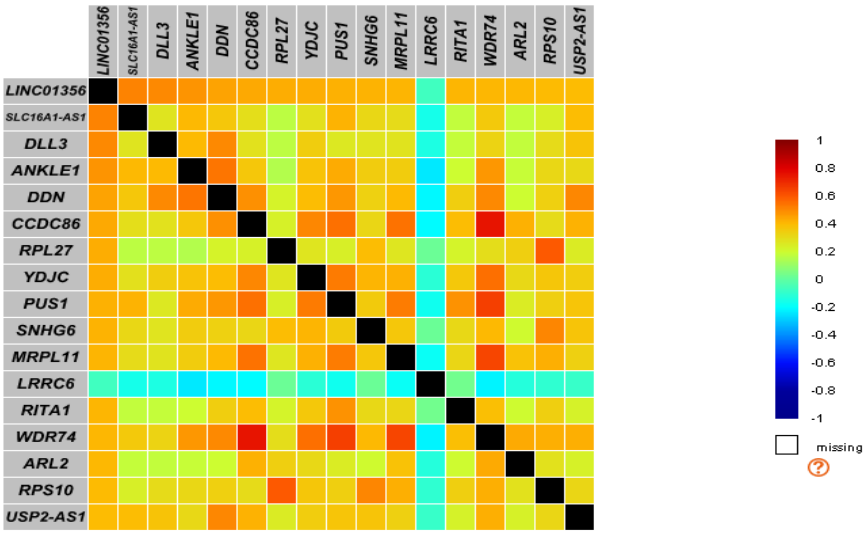


Figure 15. Heat map showing the correlation of LINC01356 with different genes in the basal subtype. SLC16A1-AS1, DLL3, ANKLE1, and DDN showed a more positive correlation with LINC01356 as compared to other genes based on TCGA RNA seq data as predicted by bcExGenMiner4.5. Blue color represents a negative correlation whereas red represents a strong positive correlation.

3.6. LINC01357

3.6.1. Tissue-based expression analysis of LINC01357

LINC01357 is located at chromosome 1p13.2 having 3 exons. This gene is long intergenic non-protein coding RNA with 3 different transcript isoforms. Tissue-specific expression analysis using GTex showed insignificant expression in normal breast mammary tissue.

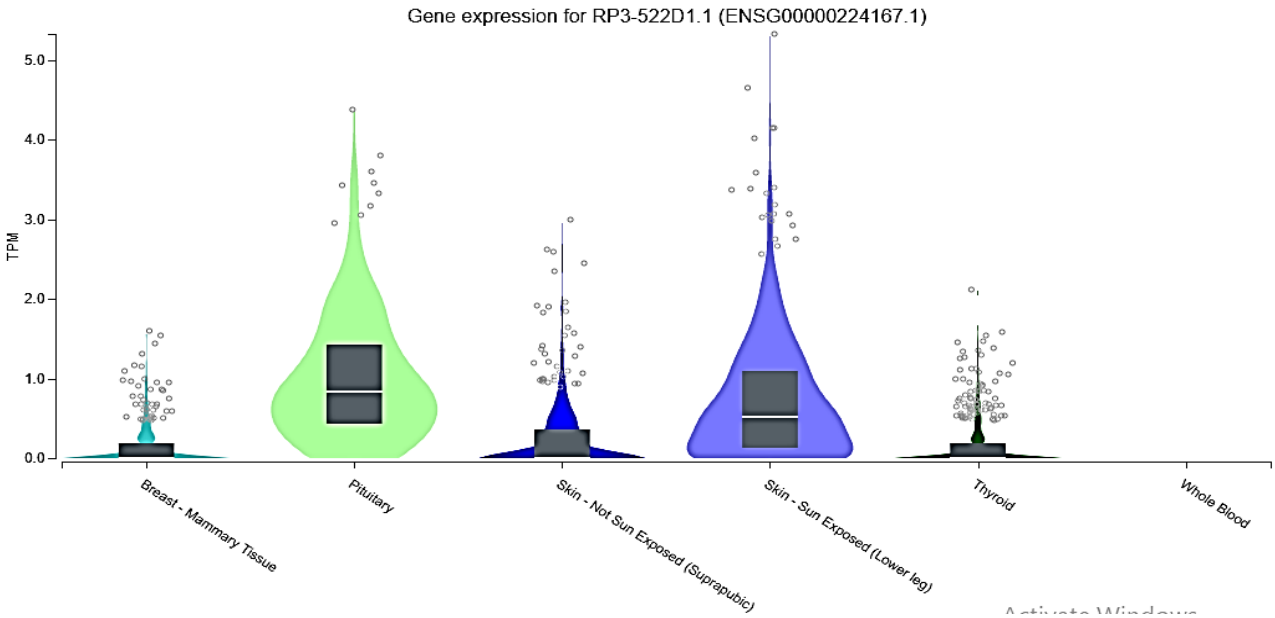


Figure 16. Tissue-specific expression analysis of LINC01357 Expression analysis showed insignificant expression around 1TPM expression of LINC01357 in normal breast mammary tissue. Expression values are in transcript per million (TPM). Light blue shows breast mammary tissue, mustard shows the pituitary gland, DARK blue shows unexposed skin, sky blue shows exposed skin, dark green represents the thyroid and pink represents whole blood.

3.6.2. Subtype analysis of LINC01357

Expression analysis was performed among four subtypes of BC using GEPIA2. It was calculated by mean value of log2 (TPM+1) in each subtype of breast cancer. Results showed that LINC01357 is upregulated in basal-like breast cancer as compared to the other three subtypes.

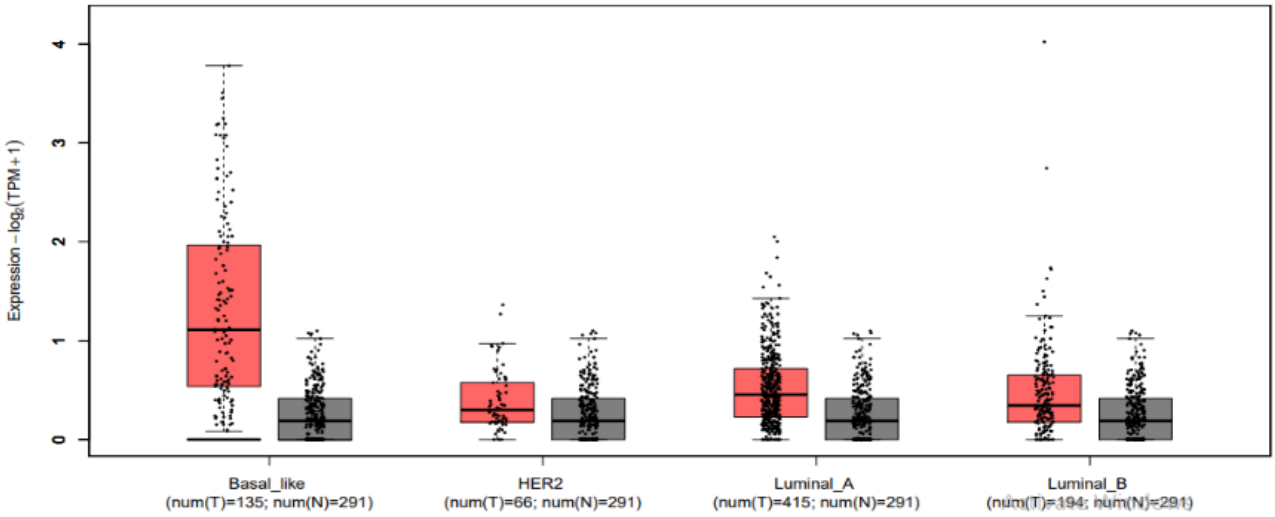


Figure 17. Subtype analysis of LINC01357 based on TCGA data. The box plot in red represents the tumor while in grey represents normal. LINC01357 expression is upregulated in a basal subtype of BC than the other three subtypes.

3.6.3. Survival analysis of LINC01357

LINC01357 is highly upregulated in the basal subtype survival graph as we evaluated it through GEPIA2. There are two groups i.e. red represents the high expression group whereas blue represents the low expression group. HR represents the hazard ratio which is an indication of the association between different treatments (radiotherapy and chemotherapy) and survival time.

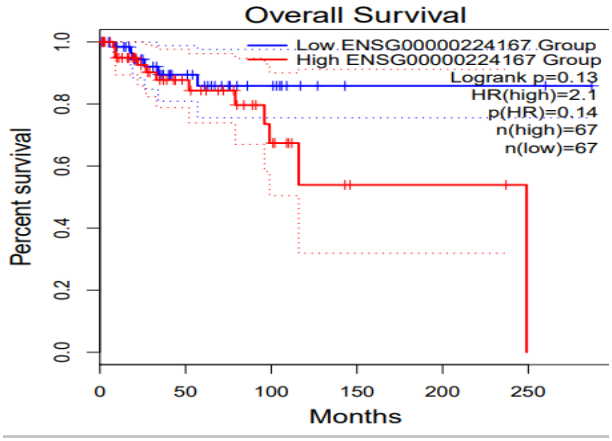


Figure 18. Graph showing overall survival analysis of LINC01357. The higher expression group showed low survival probability. Blue color represents low expression; red represents high expression and n represents the number of patients. The percent survival is shown on the y-axis and the time duration in months is shown on the x-axis.

3.6.4. Gene structure

LINC01357 is a long intergenic non-coding RNA having a size of 1819 bp. It has 3 transcript isoforms. Table 3 shows multiple transcripts of LINC01357 along with their transcript length.

Table 3. List of transcripts of LINC01357

Transcript ID	Name	Bp	Protein	Bio type
ENST00000658577.1	LINC01357-203	749	No protein	lncRNA
ENST00000456651.1	LINC01357-202	587	No protein	lncRNA
ENST00000422022.1	LINC01357-201	490	No protein	lncRNA

3.6.5. Protein coding potential using PhyloCSF

The coding potential of LINC01357 was further evaluated by using the PhyloCSF tool. PhyloCSF uses a calculate the phylogenetic conservation score multi-species nucleotide sequence alignment, which represents a probable protein-coding region LINC01357 shows all the peaks towards the negative axis and does not code for any protein.

Figure 19. Protein coding potential of LINC01357. The coding potential was evaluated and PhyloCSF. The coding potential was evaluated in three frames indicated as 0, 1 and 2. All the peaks lie in the negative axis indicating that LINC01357 does not code for a protein.

3.6.6. Conservation of sequence and structure

Few alignments of 100 vertebrates showed conservation of LINC01357 in rhesus. However, less conservation is observed in elephants and mice.

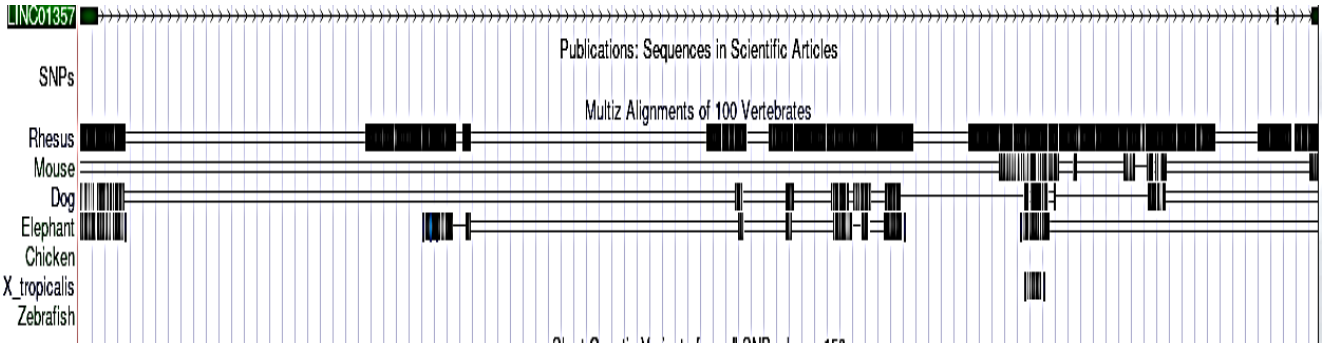
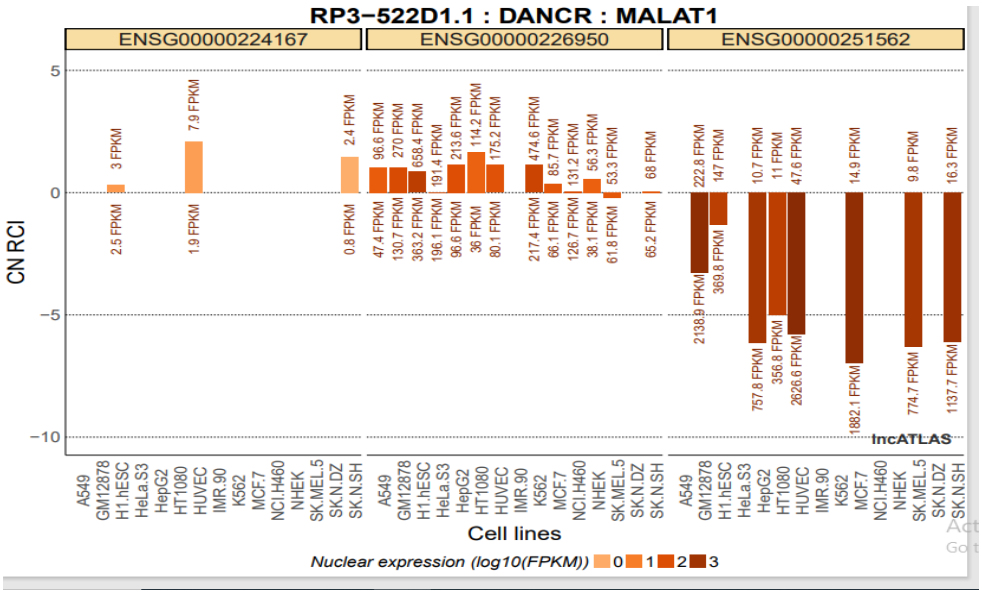


Figure 20. Multiz alignment of 100 vertebrates showing little bit of conservation of LINC01357.

3.6.7. Nucleo-cytoplasmic localization of LINC01357

The nucleo-cytoplasmic localization of LINC01357 was determined by using lnc-ATLAS. MALATI was used as a reference nuclear gene whereas DANCR was used as a reference cytoplasmic gene. LINC01357 showed few cytoplasmic localization in different cell lines indicating its role in regulating gene expression via RNA-RNA or RNA-protein interaction.

Figure 21. Nucleo-cytoplasmic localization of LINC01357 LINC01357 showed few cytoplasmic localizations in multiple cell lines. MALATI was used as a reference gene for the nucleus and DANCR was used as the reference gene for the cytoplasm. RCI represents the relative concentration index based on the comparison of gene concentration per unit RNA mass between two cellular compartments.



3.6.8. Macromolecular interactions involving LINC01357

LINC01357 can interact with neighboring genes or genes on other chromosomes in a cis or trans manner. Two protein-coding genes (TAFA3 and SLC16A1) which are in close vicinity were selected for correlation expression with LINC01357 (non-coding gene). BcGenExMinor 4.5 tool was used to evaluate the correlation expression of coding and non-coding genes between four breast cancer subtypes.

Figure 22. Correlation of Cis interactors TAFA3 and SLC16A1 with LINC01357 predicted by BcGenExMiner 4.5 in normal subtype. TAFA3 showed a negative correlation whereas SLC16A1 is negatively correlated in normal subtype of breast cancer. Blue color represents a negative correlation whereas red represents a strong positive correlation.

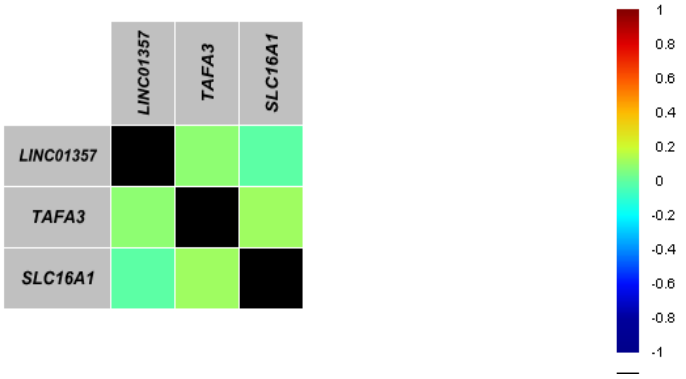


Figure 23. Correlation of Cis interactors TAFA3 and SLC16A1 with LINC01357 predicted by BcGenExMiner4.5 in basal (TNBC) subtype. SLC16A1 and TAFA3 showed strong correlation with LINC01357.

3.6.9. UALCAN database analysis

SLC16A1 is a protein coding gene which interact with LINC01357. It is highly up regulated in TNBC subtype as compared to other subtypes which is calculated through UALCAN database.

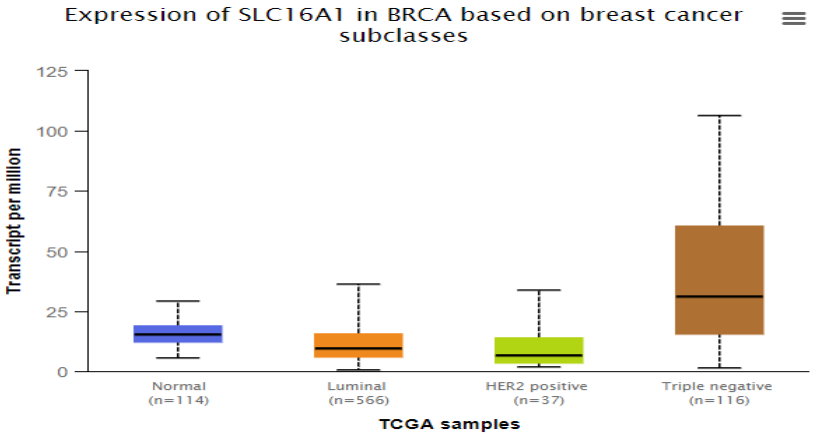


Figure 24. Box whisker plot showing SLC16A1 expression in breast cancer subtypes. SLC16A1 is significantly upregulated in TNBC while showing nominal expression in other subtypes as illustrated by the box plot.

Table 4. Statistical Significance of SLC16A expression in BRCA

Comparison	Statistical significance
Normal-vs-Luminal	3.863800E-02
Normal-vs-HER2 Positive	9.416100E-02
Normal-vs-TNBC	7.597200E-03
Luminal-vs-HER2 Positive	3.771600E-01
Luminal-vs-TNBC	2.843500E-02
HER2 Positive-vs-TNBC	2.218000E-01

3.7. ARF4-AS1

3.7.1. Tissue-based expression analysis of ARF4-AS1

ARF4-AS1 is located at chromosome 3p14.3 having 3 exons. This gene is long intergenic non-protein coding RNA with 4 different transcript isoforms. Tissue-specific expression analysis using GTex showed significant expression in normal breast mammary tissue.

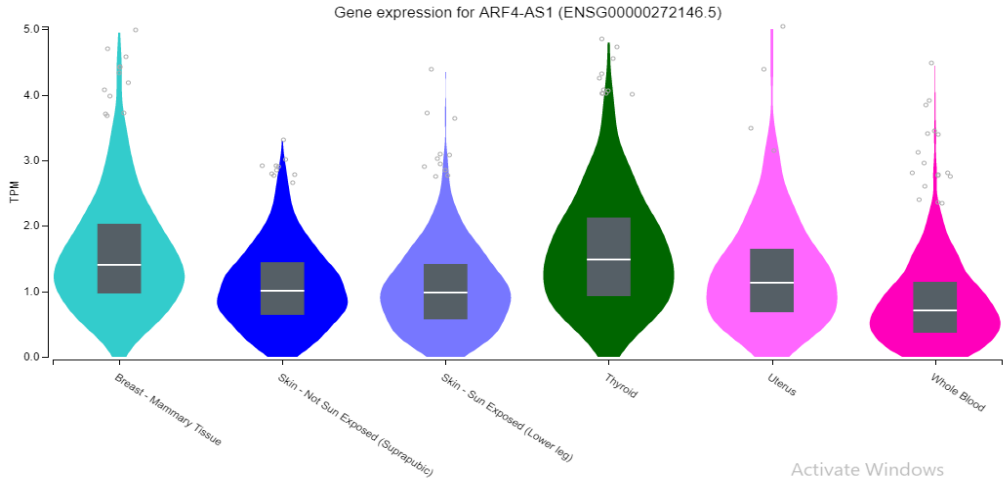
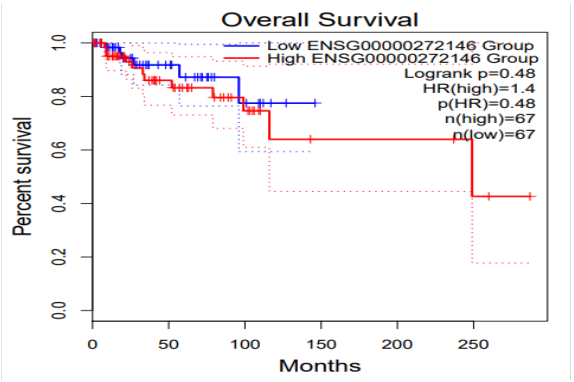


Figure 25. Tissue-specific expression analysis of ARF4-AS1 Expression analysis showed insignificant expression around 5TPM expression of ARF4-AS1 in normal breast mammary tissue. Expression values are in transcript per million (TPM). Light blue represents breast mammary tissue, dark blue represents unexposed skin, sky blue represents exposed skin, dark green represents the thyroid, light purple represents the uterus and pink represents whole blood.

3.7.2. Survival analysis of ARF4-AS1

ARF4-AS1 is highly upregulated in the basal subtype survival graph as the study evaluated it through GEPIA2. There are two groups i.e. red represents the high expression group whereas blue represents the low expression group.HR represents the hazard ratio which is an indication of the association between different treatments (radiotherapy and chemotherapy) and survival time. Figure 26 represents a decrease in the survival rate of the high expression group so it is oncogenic.

Figure 26. Graph showing overall survival analysis of ARF4-AS1. The higher expression group showed low survival probability. Blue color represents low expression; red represents high expression and n represents a number of patients. The percent survival is shown on the y-axis and the time duration in months is shown on the x-axis.



3.7.3. Gene structure

ARF4-AS1 is a long intergenic non-coding RNA having a size of 327bp. It has 4 transcript isoforms. Table 5 shows multiple transcripts of ARF4-AS1 along with their transcript length.

Table 5. List of transcripts of ARF4-AS1

Transcript ID	Name	Bp	Protein	Bio type
ENST00000606192.5	ARF4-AS1-201	327	No protein	lncRNA
ENST00000658756.1	ARF4-AS1-204	588	No protein	lncRNA
ENST00000607782.1	ARF4-AS1-203	552	No protein	lncRNA
ENST00000607297.1	ARF4-AS1-202	437	No protein	lncRNA

3.7.4. Protein coding potential using PhyloCSF

The coding potential of ARF4-AS1 was further evaluated by using PhyloCSF tool. PhyloCSF uses a calculate the phylogenetic conservation score multi-species nucleotide sequence alignment, which represents a probable protein-coding region. All peaks are towards negative axis showing it does not code for protein or small peptides.

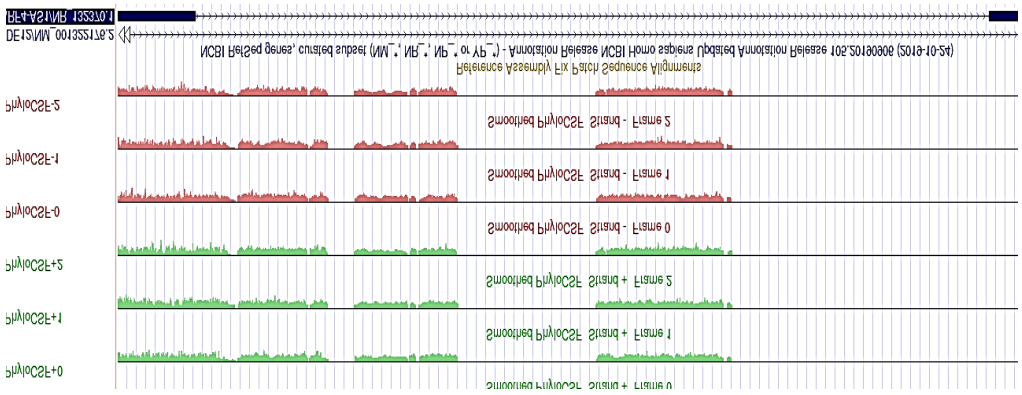


Figure 27. Protein coding potential of ARF4-AS1. The coding potential was evaluated and PhyloCSF. The coding potential was evaluated in three frames indicated as 0, 1and 2.

3.7.5. Conservation of sequence and structure

Few alignments of 100 vertebrates showed little conservation of ARF4-AS1 in rhesus, mouse and dog.

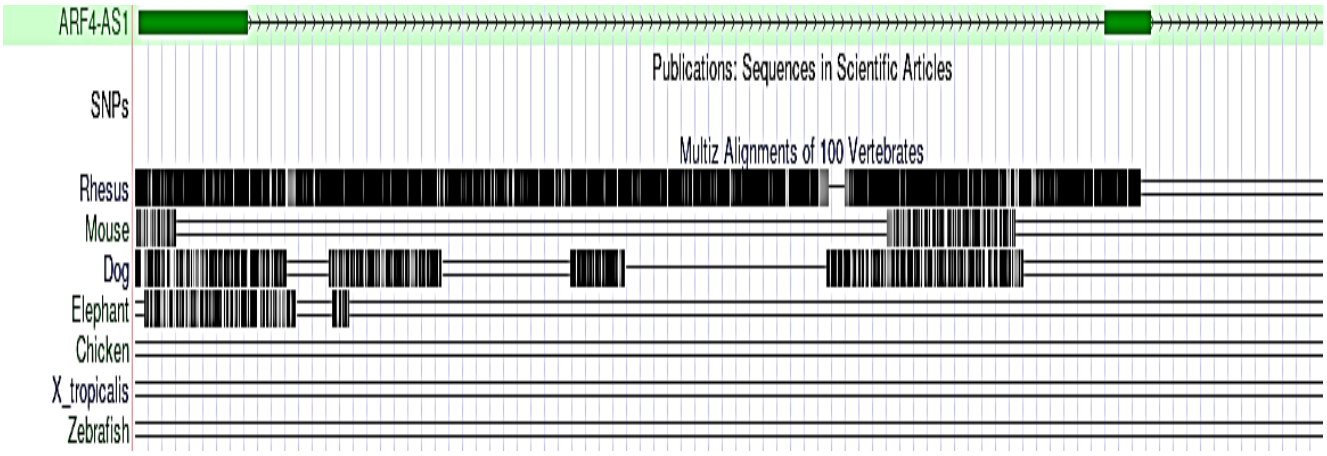


Figure 28. Multiz alignment of 100 vertebrates showing little bit of conservation of LINC01357.

3.7.6. Nucleo-cytoplasmic localization of ARF4-AS1

The nucleo-cytoplasmic localization of ARF4-AS1 was determined by using lnc-ATLAS. MALATI was used as a reference nuclear gene whereas DANCR was used as a reference cytoplasmic gene. LINC01357 showed various cytoplasmic localization in different cell lines i.e., HT1080. A549 indicates its role in regulating gene expression via RNA-RNA or RNA-protein interaction whereas a single line shows its nuclear localization.

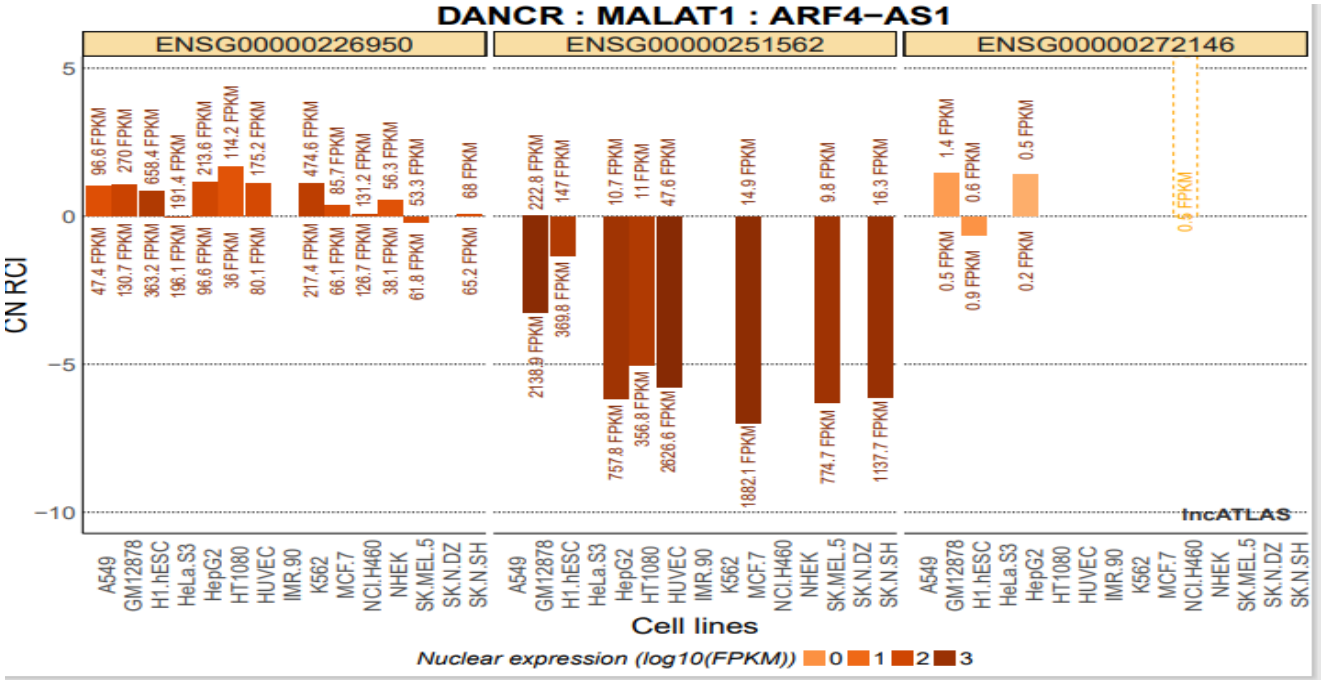


Figure 29. Nucleo-cytoplasmic localization of ARF4-AS1 ARF4-AS1 showed few cytoplasmic localizations in multiple cell lines. MALATI was used as the reference gene for the nucleus and DANCR was used as the reference gene for the cytoplasm. RCI represents the relative concentration index based on the comparison of gene concentration per unit RNA mass between two cellular compartments.

3.7.7. Macromolecular interactions involving ARF4-AS1

ARF4-AS1 can interact with neighboring genes or genes on other chromosomes in a cis or trans manner.

Two protein coding genes (ASB14 AND SPATA 12) which are in close vicinity were selected for correlation expression with ARF4-AS1 (non-coding gene). BcGenExMinor4.5 tool was used to evaluate the correlation expression of coding and non-coding genes between four breast cancer subtypes.

Figure 30. Correlation of cis interactors ASB14 and SPATA12 with ARF4-AS1 predicted by BcGenExMiner4.5 in normal subtype. ASB14 showed a strong correlation as compared to SPATA12. Blue color represents a negative correlation whereas red represents a strong positive correlation.

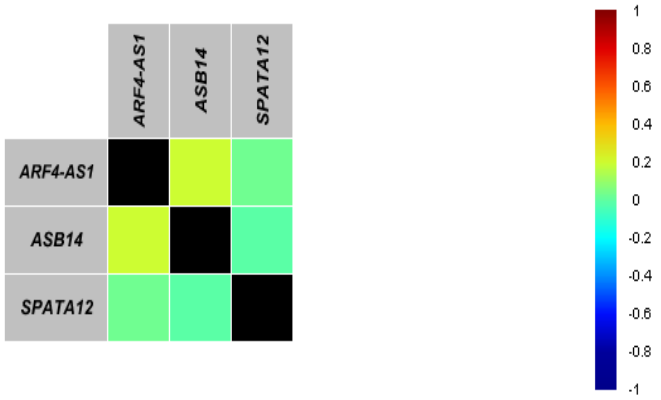
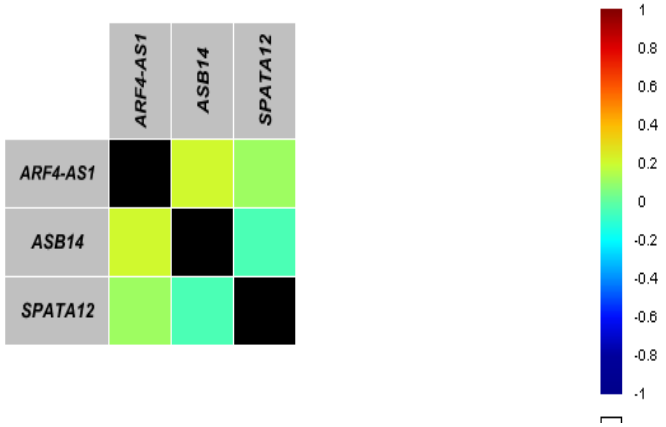


Figure 31. Correlation of Cis interactors ASB14 and SPATA12 with ARF4-AS1 predicted by BcGenExMiner4.5 in basal (TNBC) subtype. ASB14 showed more stronger correlation than SPATA12.



3.7.8. UALCAN database analysis

SPATA12 is a protein coding gene that interacts with ARF4-AS1. It is highly upregulated in the TNBC subtype as compared to other subtypes. It is calculated through the UALACAN database.

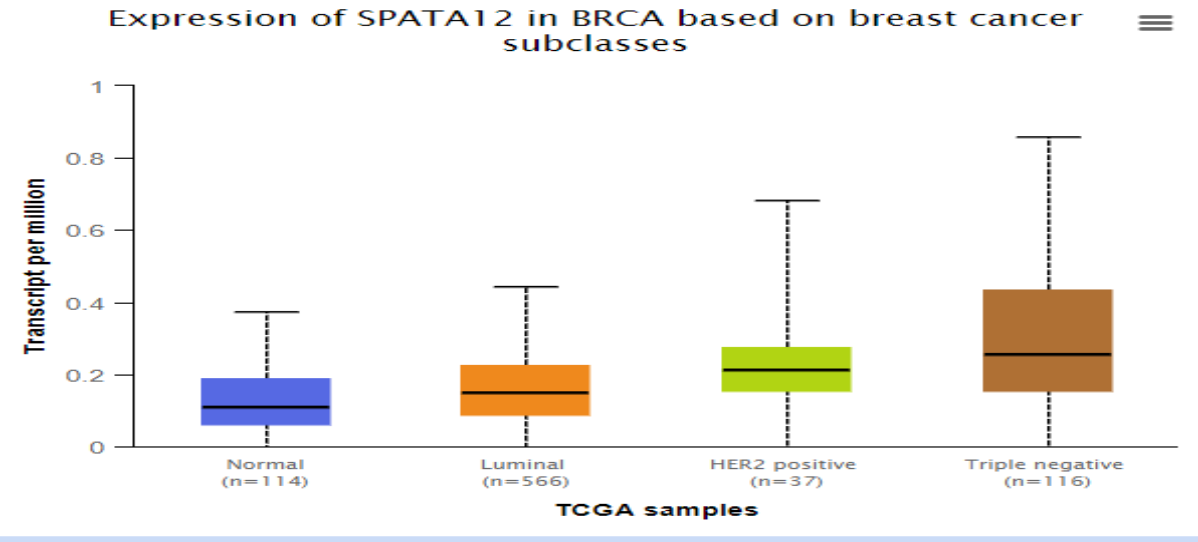


Figure 32. Box whisker plot showing SAPATA12 expression in breast cancer subtypes. SPATA12 is significantly upregulated in TNBC while showing nominal expression in other subtypes as illustrated by the box plot.

Table 6. Statistical significance of SPATA12 expression in BRCA

Comparison	Statistical significance
Normal-vs-Luminal	1.849390E-04
Normal-vs-HER2 Positive	5.232200E-04
Normal-vs-TNBC	3.89670518075036E-11
Luminal-vs-HER2 Positive	8.264600E-03
Luminal-vs-TNBC	4.59779999628651E-08
HER2 Positive-vs-TNBC	5.185400E-01

3.8. FAM83B

3.8.1. Subtype analysis of FAM83B

FAM83B is located at chromosome 6p12.1 having 5 exons. Expression analysis was performed among four subtypes of BC using GEPIA2. It was calculated by mean value of log2 (TPM+1) in each subtype of breast cancer. Results showed that FAM83B is upregulated in basal like breast cancer as compared to the other three subtypes.

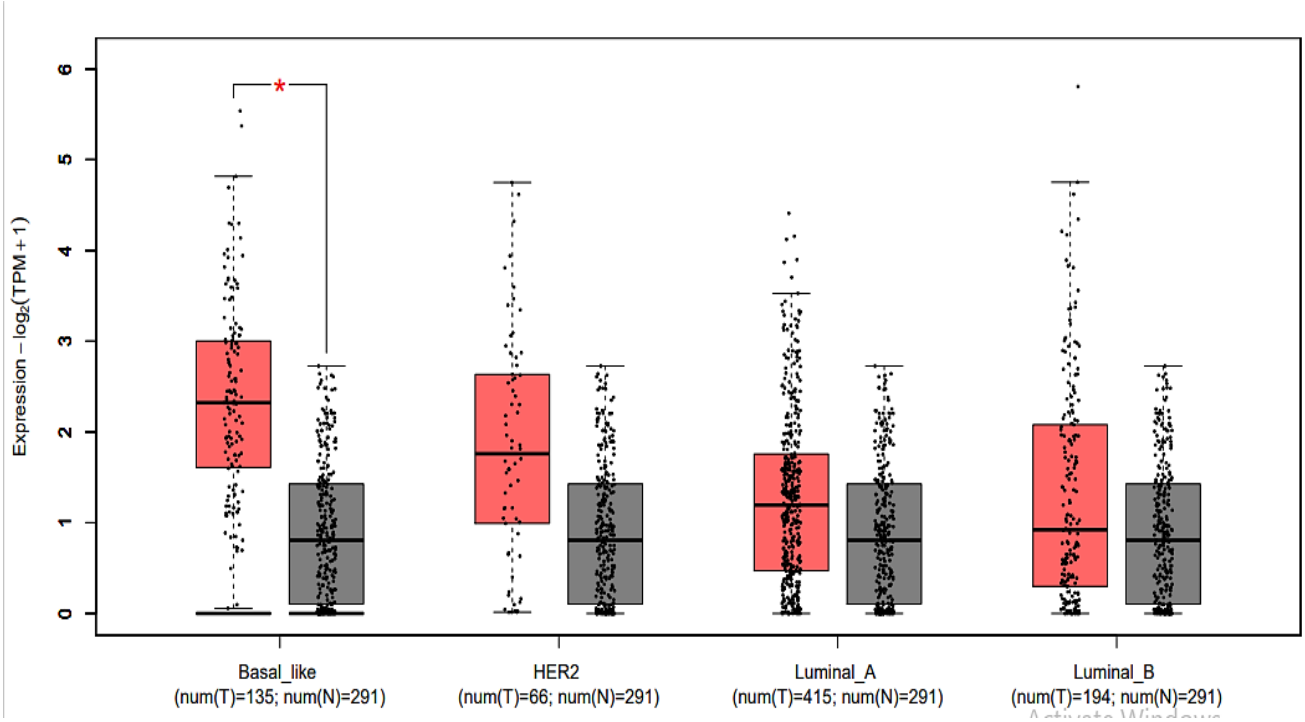
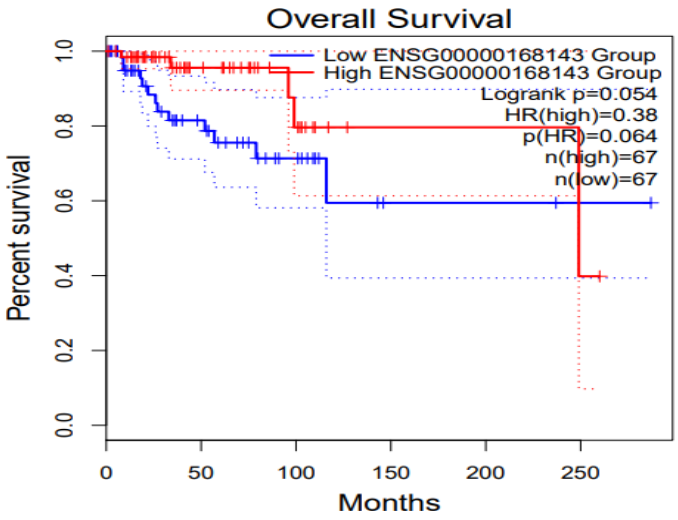


Figure 33. Subtype analysis of FAM83B based on TCGA data. The box plot in red represents the tumor while in grey represents normal. FAM83B expression is upregulated in the basal subtype of BC than the other three subtypes.

3.8.2. Survival analysis of FAM83B

FAM83B is highly upregulated in the basal subtype survival graph as we evaluated it through GEPIA2. There are two groups i.e. red represents the high expression group whereas blue represents the low expression group. HR represents the hazard ratio which is an indication of the association between different treatments (radiotherapy and chemotherapy) and survival time. Figure 34 represents there is difference between the two groups. The survival rate of breast cancer is decreasing over 250 months as presented.

Figure 34. Graph showing overall survival analysis of FAM83B. The higher expression group showed low survival probability. Blue color represents low expression; red represents high expression and n represents the number of patients. The percent survival is shown on the y-axis and the time duration in months is shown on the x-axis.



3.8.3. Gene structure

FAM83B is a long intergenic non-coding RNA having the size of 6244bp. It has 1 transcript isoform. Table 7 shows the transcript of FAM83B along with its transcript length.

Table 7. List of transcripts of FAM83B

Transcript ID	Name	Bp	Protein	Bio type
ENST00000306858.8	FAM83B-201	6244	No protein	lncRNA

3.8.4. Protein coding potential using PhyloCSF

The coding potential of FAM83B was further evaluated by using PhyloCSF tool. PhyloCSF uses a calculate the phylogenetic conservation score multi-species nucleotide sequence alignment, which represents a probable protein-coding region. All peaks are towards negative axis showing that it does not code for any proteins or small peptide.

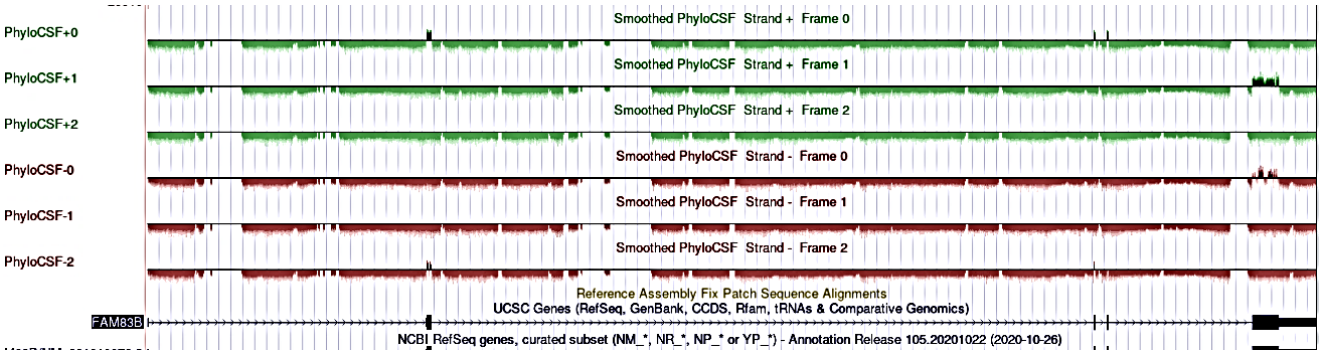


Figure 35. Protein coding potential of FAM83B. The coding potential was evaluated and PhyloCSF. The coding potential was evaluated in three frames indicated as 0, 1and 2.

3.8.5. Conservation of sequence and structure

Few alignments of 100 vertebrates showed high conservation of FAM83B in rhesus and few alignments in mouse, dog and elephant.

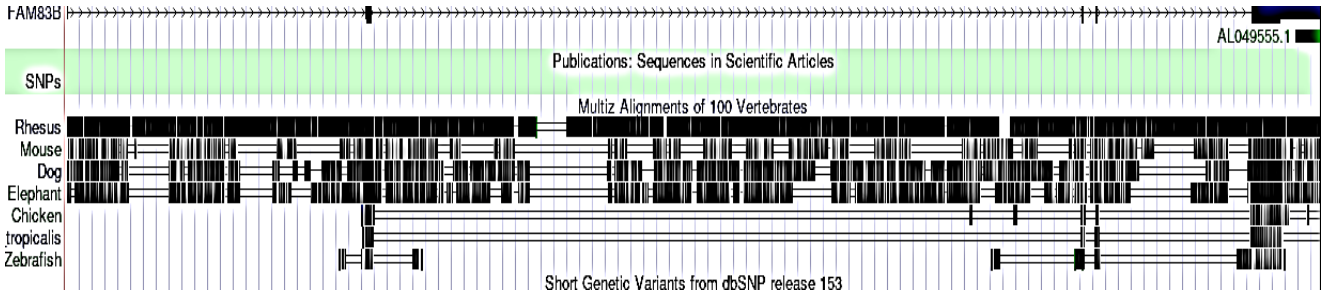


Figure 36. Multiz alignment of 100 vertebrates showing little bit conservation of FAM83B.

3.8.6. Nucleo-cytoplasmic localization of FAM83B

The nucleo-cytoplasmic localization of FAM83B was determined by using lnc-ATLAS. MALATI was used as a reference nuclear gene whereas DANCR was used as a reference cytoplasmic gene. FAM83B showed few cytoplasmic localizations in different cell lines i.e. H1 Hesc and Hela. S3 indicates its role in regulating gene expression via RNA-RNA or RNA-protein interaction whereas a single line shows its nuclear localization.

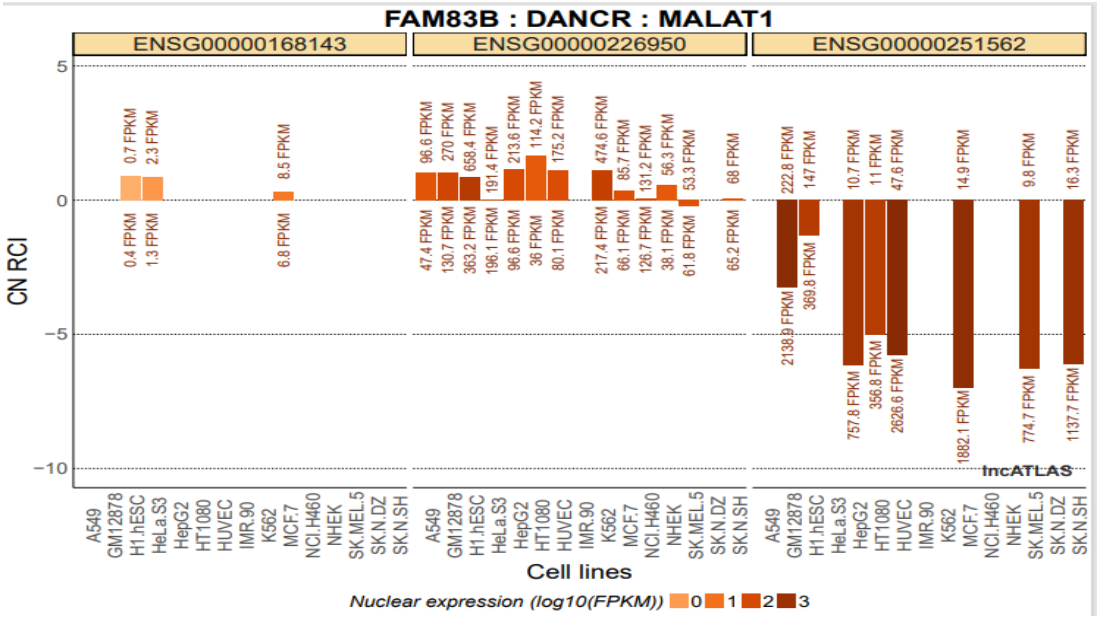


Figure 37. Nucleo-cytoplasmic localization of FAM83B. FAM83B showed few cytoplasmic localization in multiple cell lines. MALATI was used as a reference gene for the nucleus and DANCR was used as a reference gene for the cytoplasm. RCI represents the relative concentration index based on the comparison of gene concentration per unit RNA mass between two cellular compartments.

3.8.7. Macromolecular interactions involving FAM83B

FAM83B can interact with neighboring genes or genes on other chromosomes in a cis or trans manner. Two protein coding genes (HCRTR2 and GFRAL) which are in close vicinity were selected for correlation expression with FAM83B (non-coding gene). BcGenExMinor4.5 tool was used to evaluate the correlation expression of coding and non-coding genes between four breast cancer subtypes.

Figure 38. Correlation of Cis interactors HCRTR2 and GFRAL with FAM83B predicted by BcGenExMiner4.5 in the normal subtype. HCRTR2 showed a positive correlation whereas GFRAL showed a negative correlation with FAM83B.

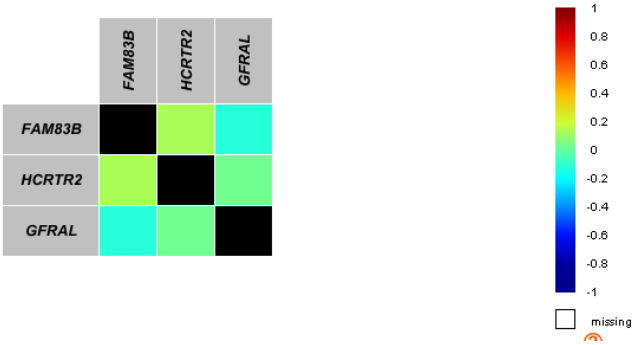
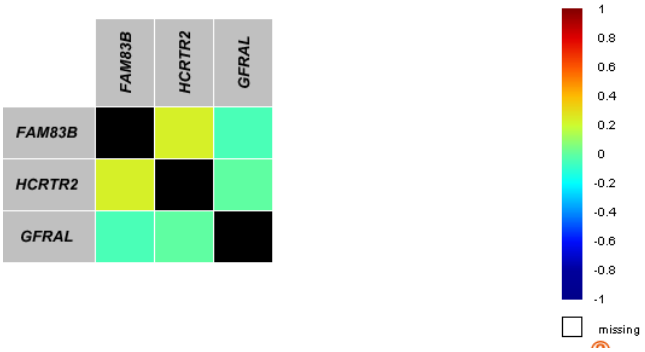


Figure 39. Correlation of Cis interactors HCRTR2 and GFRAL with FAM83B predicted by BcGenExMiner4.5 in basal (TNBC) subtype. HCRTR2 showed a strong correlation whereas GFRAL showed a negative correlation.



3.8.8. Trans interaction

Trans interaction is the interaction of genes on same chromosomes or on different chromosomes. Different trans-interacting genes were correlated with FAM83B.

Figure 40. Heat map showing the correlation of FAM83B with different genes in normal breast cancer tissue. PERP, DSC2, and DSG2 showed strong positive correlation whereas XPO5, SERPINB5, GRHL1 and ATL2 also positively correlated with FAM83B based on TCGA RNA seq data as predicted by bcExGenMiner4.5. Blue color represents a negative correlation whereas red represents a strong positive correlation.

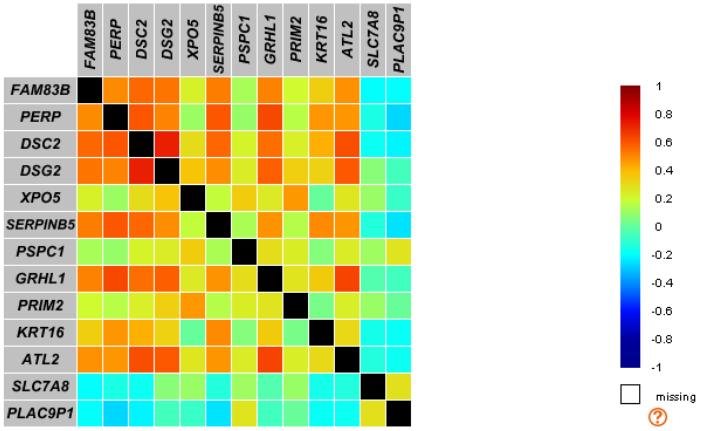
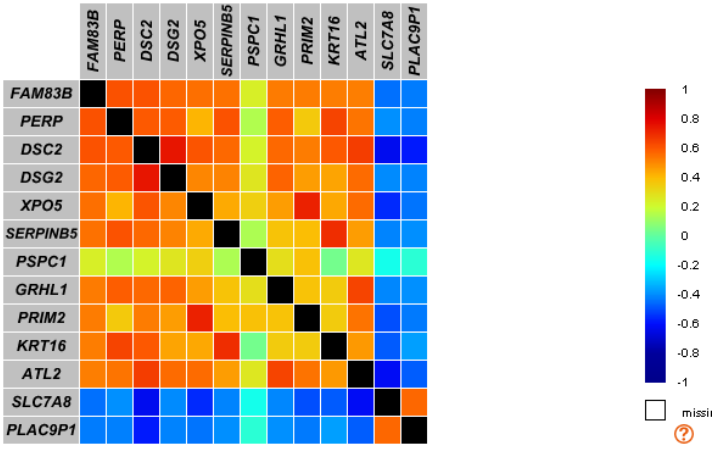


Figure 41. Heat map showing the correlation of FAM83B with different genes in the basal subtype. PERP, DSC2, and DSG2 showed a more positive correlation with FAM83B as compared to other genes based on TCGA RNA seq data as predicted by bcExGenMiner4.5. Blue color represents a negative correlation whereas red represents a strong positive correlation.



3.8.9. UALCAN database analysis

BMP5 is a protein-coding gene that interacts with FAM83B. It is upregulated in the TNBC subtype as compared to other subtypes. It is calculated through the UALACAN database.

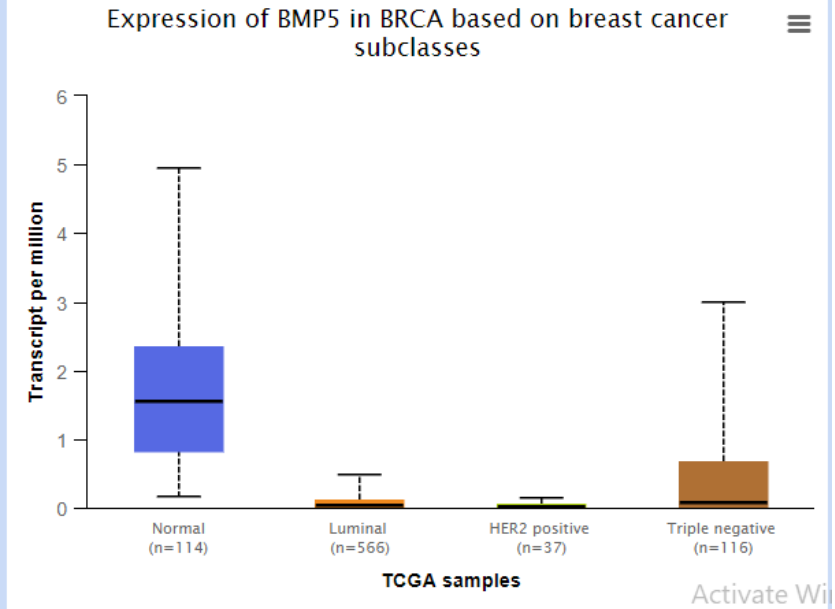


Figure 42. Box whisker plot showing expression in breast cancer subtypes. BMP5 is upregulated in TNBC while showing nominal expression in other subtypes as illustrated by the box plot.

Table 8. Statistical significance of BMP5 expression in BRCA

Comparison	Statistical significance
Normal-vs-Luminal	2.93720048283319E-10
Normal-vs-HER2 Positive	1.80211401357155E-12
Normal-vs-TNBC	2.239800E-02
Luminal-vs-HER2 Positive	8.882000E-02
Luminal-vs-TNBC	4.309800E-02
HER2 Positive-vs-TNBC	1.896320E-03

4. Discussion

Breast cancer is a health dilemma around the globe as it has a high mortality rate. About one million breast cancer cases were reported in the world so it comprises about 18% of all cancer among women. It is predicted that by 2021 incidence increase to 85 per 100,00 women. Breast cancer has two mechanisms that lead the patient towards death i.e. distant metastasis and invasion [18]. Long non-coding RNA plays a major role in these mechanisms by epithelial-mesenchymal transition, migration, invasion, and metastasis, and leads towards the progression of BC Related lncRNAs of promoting BC development [19].

This study shortlisted four genes on the basis of their expression analysis on online software. The first two are rejected because of their survival probability. Those genes that are selected for study are LINC01356, LINC01357, ARF4-AS1, and FAM83B. First, lncRNA LINC01356 was evaluated to find its role in the progression of breast cancer. Multiple online software are utilized to conclude the results. Subtype analysis showed high expression in the TNBC subtype of breast cancer. The survival probability was evaluated which showed a decrease in survival probability. By PhyloCSF, its protein-coding potential was evaluated which

showed it is a negative strand so do not code for any protein. Further, it was analyzed for correlation expression. LINC01356 lies in the cytoplasm. In cis interaction, TAFA3 and SLC16A showed a positive correlation in the basal subtype of TNBC. In trans interaction, SLC16A-AS1, DLL3, ANKLE1, and DDN were positively correlated in the basal subtype of TNBC. Previous studies on LINC01356 showed that it was upregulated in bladder cancer and promoted bone metastasis and can be utilized as an important diagnostic biomarker and also serve as a therapeutic target for bladder cancer and to control its metastasis for bone. Another study was also coherent with our study in the way that LINC01356 is significantly upregulated for tumor progression by significantly correlating with main immune checkpoints and they serve as major risk factors for bladder cancer.

Second, lncRNA LINC01357 was evaluated for breast cancer development. Results showed it was significantly upregulated in subtype analysis in the TNBC subtype of breast cancer. The survival probability curve showed that LINC01357 was an oncogene as it decreased the survival ratio of the patient over 250 months. LINC01357 has cytoplasmic localization. Two protein coding genes TAFA3 and SLC16A1 were correlated in a cis manner with LINC01357. SLC16A1 is a protein called MCT1 is transporter of monocarboxylic acid that transports short-chain fatty acids, lactate and pyruvate. It modifies the metabolic programming of cells to promote the development and progression of tumors [20]. It also serves as a target for MYC oncoproteins as SLC16A1 elevated level promotes malignancies targeting MYC or MYC involvement. So, it is also a potential target for therapeutic purposes. Another study also showed the oncogenic role of SLC16A1 as it is upregulated in many human malignancies [21,22]. This study demonstrated the effect of stress on cancer cell survival. It showed that suppression of SLC16A1 by specific siRNA blockage or any other inhibitor leads to suppression of glycolysis and lactate homeostasis of breast tumors. Another effect was observed known as the Warburg effect which showed that by suppressing SLC16A1 cells were metabolically shifted from anaerobic glycolysis towards oxidative phosphorylation so they serve as the major factor for tumor suppression.

Thirdly, lncRNA ARF4-AS1 was evaluated for breast cancer progression. Results showed significant expression in tissue specific expression and further survival analysis showed high expression in the basal subtype of TNBC so it is also oncogenic. It is highly conserved in rhesus. ASB14 was a protein coding gene that was correlated in cis manner ARF4-AS1. SPATA12 is a novel spermatogenesis-associated gene and plays a role in the inhibition of spermatogenesis and tumor genesis through the regulation of various genes of the cell cycle [23,24].

Fourth, lncRNA FAM83B was evaluated for its role in breast cancer development. Results concluded that it was highly upregulated in subtype analysis of TNBC subtype using GEPIA 2. It has a low survival possibility as shown by the survival probability curve found highly conserved in rhesus and has cytoplasmic localization. Two protein-coding genes HCRTR2 and GFRAL were correlated in a cis manner in which HCRTR2 showed a positive correlation in the basal subtype of TNBC [25], which was the protein coding gene for FAM83B. BMP5 (bone morphogenetic protein 5), a member of the TGF-β family was known to have a role in cancer progression and metastasis and previous studies showed it as an oncogene involved in various cancers i.e. lung, bladder, breast, ovarian and colon [26].

5. Conclusion

In the current study, lncRNAs were investigated for progression in breast cancer through different online bioinformatics tools and exploring their different features i.e. structure, multiple transcript isoforms, subtype analysis, survival probability and correlation expression with cis and trans interacting protein and RNA coding genes. This study analyzed four lncRNAs (LINC01356, LINC01357, ARF4-AS1, and FAM83B). LINC01356

showed a positive correlation with these protein coding genes i.e. TAFA3 and SLC16A1. LINC01357 showed a positive correlation with cis acting genes TAFA3 and SLC16A1 in a basal subtype of TNBC. ARF4-AS1 showed a strong correlation with ASB14 and SPATA12 which are protein-coding genes. FAM83B showed a good correlation with BMP5. Therefore, all these findings paved the way to consider lncRNAs as therapeutic targets for the development of breast cancer.

Authors contribution

Usama Ahmed – Literature review, Sampling, Experiment design, Experimental work, Result analyses
Muhammad Abubakar – Literature review, Methodology, Resources
Salma Saeed Khan – Experiment design, Methodology,
Baqa ur Rehman – Primer design, Methodology, Resources

Disclosure statement

The author declares that they have no conflict of interest.

References

[1] Zubair M, Wang S, Ali N, 2021, Advanced Approaches to Breast Cancer Classification and Diagnosis. Frontiers in Pharmacology, 11: 632079.

[2] Zavala VA, Bracci PM, Carethers JM, et al., 2021, Cancer Health Disparities in Racial/Ethnic Minorities in the United States. British Journal of Cancer, 124(2): 315–332.

[3] Martini R, Newman L, Davis M, 2022, Breast Cancer Disparities in Outcomes; Unmasking Biological Determinants Associated with Racial and Genetic Diversity. Clinical & Experimental Metastasis, 39(1): 7–14.

[4] Roheel A, Khan A, Anwar F, et al., 2023, Global Epidemiology of Breast Cancer Based on Risk Factors: A Systematic Review. Frontiers in Oncology, 13: 1240098.

[5] Kazemi A, Barati-Boldaji R, Soltani S, et al., 2021, Intake of Various Food Groups and Risk of Breast Cancer: A Systematic Review and Dose-Response Meta-Analysis of Prospective Studies. Advances in Nutrition, 12(3): 809–849.

[6] Mattick JS, Amaral PP, Carninci P, et al., 2023, Long Non-Coding RNAs: Definitions, Functions, Challenges and Recommendations. Nature Reviews Molecular Cell Biology, 24(6): 430–447.

[7] Mangiavacchi A, Morelli G, Orlando V, 2023, Behind the Scenes: How RNA Orchestrates the Epigenetic Regulation of Gene Expression. Frontiers in Cell and Developmental Biology, 11: 1123975.

[8] Kazimierczyk M, Wrzesinski J, 2021, Long Non-Coding RNA Epigenetics. International Journal of Molecular Sciences, 22(11): 6166.

[9] Yu W, et al., 2021, Identification of Immune-Related lncRNA Prognostic Signature and Molecular Subtypes for Glioblastoma. Frontiers in Immunology, 12: 706936.

[10] Constanty F, Shkumatava A, 2021, lncRNAs in Development and Differentiation: From Sequence Motifs to Functional Characterization. Development, 148(1): dev182741.

[11] Deogharia M, Gurha P, 2022, The “Guiding” Principles of Noncoding RNA Function. Wiley Interdisciplinary Reviews: RNA, 13(4): e1704.

[12] Kan RL, Chen J, Sallam T, 2022, Crosstalk Between Epitranscriptomic and Epigenetic Mechanisms in Gene Regulation. Trends in Genetics, 38(2): 182–193.

[13] Lee JE, Kim M-Y, 2022, Cancer Epigenetics: Past, Present and Future. Seminars in Cancer Biology, 83: 4–14.

[14] De Martino M, Esposito F, Pallante P, 2021, Long Non-Coding RNAs Regulating Multiple Proliferative Pathways in Cancer Cells. Translational Cancer Research, 10(6): 3140.

[15] Gandhi P, Wang Y, Li G, Wang S, 2024, The Role of Long Noncoding RNAs in Ocular Angiogenesis and Vascular Oculopathy. Cell & Bioscience, 14(1): 39.

[16] Wang M-Q, Zhu W-J, Gao P, 2021, New Insights into Long Non-Coding RNAs in Breast Cancer: Biological Functions and Therapeutic Prospects. Experimental and Molecular Pathology, 120: 104640.

[17] Wang W, Min L, Qiu X, et al., 2021, Biological Function of Long Non-Coding RNA (LncRNA) Xist. Frontiers in Cell and Developmental Biology, 9: 645647.

[18] Burstein HJ, 2022, Unmet Challenges in Systemic Therapy for Early-Stage Breast Cancer. The Breast, 62: S67–S69.

[19] Yi Y, et al., 2021, Tumor-Derived Exosomal Non-Coding RNAs: The Emerging Mechanisms and Potential Clinical Applications in Breast Cancer. Frontiers in Oncology, 11: 738945.

[20] Nguyen YT, Ha HT, Nguyen TH, et al., 2022, The Role of SLC Transporters for Brain Health and Disease. Cellular and Molecular Life Sciences, 79: 1–21.

[21] Gaballah A, Bartosch B, 2022, An Update on the Metabolic Landscape of Oncogenic Viruses. Cancers, 14(23): 5742.

[22] Chan KI, Zhang S, Li G, et al., 2024, MYC Oncogene: A Druggable Target for Treating Cancers with Natural Products. Aging Disease, 15(2): 640–697.

[23] Cheung S, Ng L, Xie P, et al., 2024, Genetic Profiling of Azoospermic Men to Identify the Etiology and Predict Reproductive Potential. Journal of Assisted Reproduction and Genetics, 41(4): 1111–1124.

[24] Xie D, Dai L, Yang X, et al., 2023, A Survival Model Based on the ASB Genes and Used to Predict the Prognosis of Kidney Renal Clear Cell Carcinoma. Genetics Research, 2023: e22.

[25] Krebs LC, Santos MMM, Siqueira MC, et al., 2024, Candidate Genes for Height Measurements in Campolina Horses. Animal Production Science, 64(1): 23071.

[26] Shao Y, Zhao C, Pan J, et al., 2021, BMP5 Silencing Inhibits Chondrocyte Senescence and Apoptosis as Well as Osteoarthritis Progression in Mice. Aging (Albany NY), 13(7): 9646–9664.