Investigating diagnostic, Prognostic, and therapeutic role of Long Non-Coding RNA MALAT1 in Breast Carcinoma: in silico study

ARSLAN SHUJA1, MUHAMMAD NASIR SHAHBAZ2, SUMAIRA SHARIF3, AQSA IKRAM4, MAMOONA SHUJA4, NAVEED SHUJA4

1- Institute of Molecular biology and biotechnology (IMBB), CRIMM, The University of Lahore, Lahore, Pakistan.  
2- Institute of Molecular biology and biotechnology (IMBB), CRIMM, The University of Lahore, Lahore, Pakistan.  
3- Associate Professor, Institute of Molecular biology and biotechnology (IMBB), CRIMM, The University of Lahore, Lahore, Pakistan.  
4- Assistant Professor, Institute of Molecular biology and biotechnology (IMBB), CRIMM, The University of Lahore, Lahore, Pakistan.  
5- Ai-Aleem Medical college, Lahore, Pakistan.  
6- Dean Faculty of Biological Sciences, Lahore University of Biological & Applied Sciences (UBAS), Professor Lahore Medical and Dental College, Lahore, Pakistan.

Corresponding Author: Prof. Muhammad Nasir Shahbaz, Email: shahbaznasir985@gmail.com, Cell: +92 333 6479181

Abstract:
Background: Metastasis-Associated Lung Adenocarcinoma Transcript 1 (MALAT1) is a nuclear-enriched long non-coding RNA, implicated in the tumorigenesis of various cancers, including breast carcinoma (BC). It is associated with cancer proliferation, invasion, migration, and metastasis, and plays a role in immune system modulation.

Aims and Objectives: This in silico study investigates MALAT1’s diagnostic, prognostic, and therapeutic potential in BC. It examines MALAT1’s differential expression in Breast Cancer (BC) versus normal tissues, its impact on patient survival, and its interaction with miR-561 affecting TOP2A mRNA degradation.

Methodology: Using publicly available datasets from GEO and TCGA, we conducted differential expression, survival, and prognostic analyses, along with network and pathway studies and drug repurposing analyzes and Tools like Graph Pad Prism and GEPIA data set was used for demographics.

Results: Our analyzes indicate overexpression of MALAT1 in BC tissues, its correlation with poorer survival rates, and involvement in key oncogenic pathways. Additionally, drug repurposing analyzes have identified potential MALAT1-targeted therapeutic strategies.

Conclusion: MALAT1 serves as a significant biomarker for BC diagnosis and prognosis and is identified as a potential therapeutic target. This study lays the groundwork for future research into MALAT1-targeted therapies in BC.

Keywords: Long Non-Coding RNA, Breast Carcinoma, Diagnostic Biomarker, Drug Repurposing, Survival Analysis, Protein-Protein Interactions, Immune Surveillance, Drug Repurposing.
Introduction:
MALAT-1, also known as the metastatic-associated lung adenocarcinoma transcript-1, is a non-coding intergenic RNA that spans 12,820 nucleotides on chr11q13. RNase P enzymatically processes the precursor transcript to produce the mature long non-coding RNA[1]. The structure of MALAT-1 is stabilized, and the lack of a poly-A tail is compensated for by a triple helical helix at the 3′ end. It regulates gene transcription through a variety of means, including repressing target gene promoters, modulating RNA-binding proteins or activating mesenchymal transcription factors, modifying chromatin, and directing post-transcriptional processing. It also affects cell death and DNA repair. When MALAT-1 is targeted with chemotherapy, cancer becomes more sensitive and DNA is damaged[2].

MALAT-1 promotes the growth, invasion, migration, and distant metastases of cancer. A number of recent studies have brought attention to MALAT-1’s immunomodulatory function and demonstrated how it might help cancer cells evade immune surveillance by suppressing the immune system and controlling the expression of various molecules connected to the tumor microenvironment. The triple-negative breast cancer (TNBC) cell culture exhibits a notable upregulation of MHC class I chain-related protein A/B expression with MALAT-1 knockdown, along with the inhibition of checkpoint markers B7-H4 and PD-L1[3].

Furthermore, MALAT1 regulates the suppressive function by lowering peripheral blood mononuclear cells (PBMCs) in cancer patients and adversely influencing myeloid-derived suppressor cells (MDSCs). In fact, in an immune-competent animal model, MALAT-1 antisense oligonucleotides (ASO) cause MALAT-1 knockdown, which reduces both immunosuppressive tumor-associated macrophages (TAM) and MDSC[4]. On the other hand, there was also a noticeable rise in cytotoxic CD8+ T cells, which provided fresh insights into the prominent function of MALAT-1 in controlling the development of cancer[5]. Using data from The Cancer Genome Atlas (TCGA), we investigated MALAT-1 gene expression in a variety of cancer types. Several studies have found MALAT-1 overexpression in a variety of malignancies, including esophageal squamous cell carcinoma (ESCC), gastric cancer (GC), non-small cell lung cancer (NSCLC), colorectal cancer (CRC), pancreatic cancer, breast cancer (BC), and hepatocellular carcinoma. MALAT-1 expression levels in samples from various cancers and healthy persons show that high MALAT-1 expression is widespread in a wide range of tumors, including melanoma, cholangiocarcinoma, cervical cancer (CC), ESCC, and HCC. However, in other malignancies such as stomach adenocarcinoma, thyroid carcinoma, and bladder carcinoma, possible differences in MALAT-1 expression between normal and malignant samples were observed[6].

There are samples with similar MALAT-1 expression levels, which contradicts with the commonly seen differential in MALAT-1 expression between normal and malignant samples. This ambiguity shows that MALAT-1 has a pleiotropic effect on cancer cells. Breast cancer is a big global problem. With 30% of newly reported cases, this disease is the most common cancer in women and one of the leading causes of cancer-related deaths. Age, menarche history, reproductive patterns, physical activity, breast features, and body habitus are all unique elements that contribute to breast cancer development[7].

Over the past 20 years, there have been enormous advancements in the early identification and treatment of breast cancer; nonetheless, the illness still poses a serious threat to public health. It is imperative that new therapy approaches for BC be developed in this setting. In order to better understand the underlying mechanism of BC development and create new intervention options, we implemented a long non-coding RNA (LncRNA)-based approach[8]. A sizable fraction of non-coding RNAs, known as LncRNA genes, are essential to both normal development and the process of cancer. More than 20,000 protein-coding genes, or 23,000 LncRNA genes, are thought to be present in the human genome. Their interactions with microRNAs, mRNAs, proteins, and genomic DNA control their physiological and pathological roles. Transcribing long noncoding RNAs (LncRNAs) from loci connected to cancer has been proposed as a major contributor to cancer risk, since deregulation of gene expression is a key event in carcinogenesis[9].

LncRNA dysregulation, deficiency, or mutations have been linked to a number of complicated disorders, including cancer of the breast. Tumor suppressors and oncogene classes can be distinguished by their roles and modes of expression. Furthermore, accumulating exploratory data lends credence to the concept that LncRNAs serve as competitive endogenous RNAs (ceRNAs), vying with one another for the binding of microRNA (miRNA) to their target genes, which are upregulated. The role of lncRNA may be identified and understood with the important assistance of the ceRNA hypothesis[10]. The LncRNA expression profile in BC is often dysregulated, and several LncRNAs, including LINC01133, ZEB1-AS1, and ABHD11-AS1, have been linked to BC carcinogenesis[11].

Determining the role of unique LncRNAs in a tumorigenic way is extremely important, as LncRNA-based treatment techniques have improved. Therefore, the goal of this work is to clarify how LncRNA
MALAT1, which functions as a miR-561 sponge in BC, contributes to cancer[12].

The nuclear-enriched long non-coding RNA Metastasis-Associated Lung Adenocarcinoma Transcript 1 (MALAT1) is frequently overexpressed in patient Tumors and metastases. MALAT1 overexpression is associated with tumor development and metastasis in a variety of tumor types, including breast cancer. MALAT1, which was first identified as an IncRNA, was shown to be more prevalent in early human non-small cell lung malignancies that metastasized. MALAT1 is overexpressed in a variety of different human malignancies, including lymphoma, multiple myeloma, hepatocellular carcinoma, esophageal squamous cell carcinoma, breast, lung, ovarian, prostate, cervical, endometrial, colorectal, gastric, pancreatic, sarcoma, bladder, and brain[13]. Being a ceRNA, MALAT1 may interact with miRNAs, including miR-205, miR-1297, miR-217, and miR-155, which can lead to modifications in cellular behaviours such inversion, migration, apoptosis, proliferation, and metastasis. Furthermore, MALAT1 may influence the carcinogenesis of cancer by stimulating the ERK/MAPK, PI3K/AKT, and Wnt/-catenin pathways; however, concurrent stimulation of oncogenic pathways may have potent carcinogenic consequences. Two subunits make up the gene Topoisomerase alpha 2 (TOP2A), which is found on 17q12–21. It is a marker of chemotherapy resistance and proliferation in a variety of cancer forms, including breast and adrenocortical carcinomas[14].

Furthermore, it has been documented that a number of miRNAs directly inhibit TOP2A, a target of cancer, to provide a regulatory function. MALAT1 overexpression was found in BC tissues and cells in this investigation. Experiments using loss-of-function indicate that BC invasion and proliferation are suppressed when MALAT1 is silenced. Furthermore, the main way in which MALAT1 and miR-561 interact has been determined [15]. As a target of miR-561, we discovered that MALAT1 can induce miR-561 to "sponge" information and hinder TOP2A mRNA degradation by blocking its connection with TOP2A. Through thorough in silico analysis, our goal was to clarify the diagnostic, prognostic, and therapeutic functions of long non-coding RNA MALAT1 in breast cancer [16].

Our methodology was created to examine the expression and functional consequences of MALAT1 throughout different stages of breast cancer by utilizing cutting-edge computational tools and publically available information. The potential of MALAT1 as a diagnostic and prognostic biomarker is shown by its differential expression in breast cancer, and addressing it may present new treatment opportunities.

Materials and Methods:

Collection and selection of Data:
The gene expression omnibus (GEO) and genome atlas (TCGA) dataset of breast cancer was detected and gather by current study. Current data was collected through breast cancer tissues gene expression profiles by comparing with normal tissues and considered treatment responses, clinical outcomes and survival rates by applying MALAT1 expression analysis. The correlation between breast cancer patients and clinical features was measured. GEPIA 2 was used for the analysis

Analysis of Differential Expression:

In present study bioinformatics tools such as DESeq2 and edge R were used to find out the difference between MALAT1 expression, breast cancer cells and normal breast tissues and this stage was so helpful and crucial in biomarker for breast cancer detection.

Survival and Prognosis Analysis:

Kaplan-Meier survival analysis and Cox proportional hazards regression models were used for prediction of MALAT1 expression levels in breast cancer patients. In case of breast cancer patients MALAT1 expression may serve as a valuable prognostic biomarker.

Pathway and Network Analysis:

In Breast cancer cases STRING and Cystoscope were used for potential pathways in MALAT1 which is molecular interactions between protein synthesis in key signalling pathways.

Examining Drug Repurposing:

In this study Drug Bank and Connectivity Map (CMap) was used from medication repurposing research

Drug Repurposing Analysis

We conducted a medication repurposing research, utilizing resources such as Drug Bank and the Connectivity Map (CMap), to identify medications and compounds that are already on the market and may be able to target MALAT1 or associated pathways in breast cancer. This strategy provides a tool for identifying unique therapeutic options that might be promptly implemented in clinical trials.

Statistical Analysis

We applied stringent statistical approaches throughout the examination to ensure the dependability of our findings. Differential expression research employed statistical tests appropriate for high-dimensional data,
with several testing changes made to lower the false discovery rate. Survival curves were compared using log-rank tests, and any confounding factors were included while revising Cox models for survival analysis. GraphPad prism and Roc Package Bioconductor was used for the analysis considering p value<0.05. Our method was designed to provide a comprehensive understanding of MALAT1’s roles in breast cancer, including its potential utility as a biomarker for diagnosis and prognosis, as well as the implications for therapeutic intervention. Our goal was to make a significant contribution to breast cancer research and pave the path for innovative novel targeted therapeutic approaches.

Results:

![Figure 1: Prognosis and Therapy Implications of Long Non-Coding RNA MALAT1 Expression Analysis in Breast Carcinoma](image)

**Quantitative Analysis of MALAT1 Expression in Breast Carcinoma and Corresponding Normal Tissue Using GEPIA Datasets**

**Analysis of MALAT1 Expression Profiles:**

The quantification of MALAT1 transcripts in breast cancer and normal breast tissue is shown in the fig-1 (a). The central mark of the box plot on the left violin plot indicates a limited expression range of MALAT1 in normal mammary tissue, which is typified by a reduced median expression level and a constricted distribution. On the other hand, an increased expression range with a noticeably higher median is seen in the right violin plot for breast cancer tissue, indicating a considerable elevation of MALAT1 under neoplastic circumstances. Individual dots that represent outlier data points that deviate from the top and lower quartiles are displayed. The bar chart provides a detailed breakdown of MALAT1 expression across various breast carcinoma subtypes, classified as "T" (Tumor) and "N" (Normal), against a gradient color scale reflective of expression intensity. Each bar, presumably representing normalized expression values like TPM, conveys the relative abundance of MALAT1 transcripts within a specific subtype, with a discernible trend of higher expression in tumor samples compared to their normal counterparts.

**Prognostic Implications:**

A prognostic relevance is inferred if MALAT1 expression levels demonstrably associate with patient outcomes, such as survival rates and disease-free intervals. The increased MALAT1 expression in tumorous tissues could suggest a negative prognostic indicator, warranting further validation through survival analysis that correlates these expression levels with patient follow-up data.
Diagnostic Potential:

The stark differential expression of MALAT1 depicted in the violin plots suggests its potential utility as a diagnostic biomarker for breast carcinoma. The consistent overexpression in tumour tissues relative to normal tissues validates MALAT1’s potential role in diagnostic stratification.

Therapeutic Considerations:

The potential for MALAT1 as a therapeutic target hinge on its functional impact on tumour progression. Should the modulation of MALAT1, possibly through targeted antisense oligonucleotide therapies, demonstrate efficacy in altering tumour behaviour, it would substantiate MALAT1’s candidacy as a therapeutic target. While the graphical representations illuminate the overexpression of MALAT1 in breast carcinoma and hint at its differential expression across subtypes, a comprehensive interpretation demands exact numerical expression values and standard statistical measures. Furthermore, a thorough exploration into the dataset metadata is required to contextualize the subtype labels and to ascertain the robustness of sample sizes across categories. These details are pivotal for a nuanced understanding of MALAT1’s diagnostic, prognostic, and therapeutic potential in breast carcinoma.

Fig-2: Prognostic Significance of MALAT1 Expression in Breast Carcinoma: (a) Expression of MALAT1 in Breast Carcinoma Compared to Normal Tissue (b) MALAT1 Expression Distribution by Stages of Breast Carcinoma (c) Overall Survival Analysis Stratified by MALAT1 Expression Levels in Breast Carcinoma (d) Disease-Free Survival in Breast Carcinoma Patients Categorized by Levels of MALAT1 Expression
Expression of MALAT1 in Breast Carcinoma Compared to Normal Tissue

The fi-2 (a) boxplot shows the expression levels of MALAT1 in BRCA tissues compared to normal breast tissues. The y-axis represents the expression level of MALAT1, quantified as transcripts per million (TPM). The BRCA samples are denoted in red, while the normal samples are in black. There's a clear indication that MALAT1 expression is significantly higher in BRCA samples than in normal breast tissue. The red boxplot (cancer tissue) shows a higher median expression level and more variability than the black boxplot (normal tissue), which is relatively lower and more consistent. This suggests that MALAT1 could serve as a potential diagnostic marker, where higher levels might indicate the presence of breast carcinoma.

MALAT1 Expression Distribution by Stages of Breast Carcinoma

The violin plot in Figure 2 (b) compares the distribution of MALAT1 expression across different stages of BRCA. The shape of each violin plot represents the density of the data points at different expression levels, with the white dot indicating the median expression level. There seems to be a slightly increasing trend in the median expression level of MALAT1 from Stage I to Stage IV, although Stage X (possibly a control or unknown stage group) has a lower median expression level. This could imply that MALAT1 expression may correlate with disease progression and therefore may have prognostic value.

Overall Survival Analysis Stratified by MALAT1 Expression Levels in Breast Carcinoma

Protein-Protein Interactions:

Fig-3: Protein-Protein Interactions among MALAT1 in BRCA

This Kaplan-Meier curve fig2-(c)(d) illustrates the overall survival of BRCA patients stratified by the expression level of MALAT1. The patients are divided into high (red line) and low (blue line) MALAT1 expression groups. The x-axis represents the time in months, and the y-axis shows the percentage of survival. The lower expression of MALAT1 correlates with a higher overall survival rate, while high expression correlates with lower survival. This indicates that high MALAT1 expression could be associated with poorer prognosis in BRCA patients.

Disease-Free Survival in Breast Carcinoma Patients Categorized by Levels of MALAT1 Expression

This Kaplan-Meier curve, like Figure (c), displays disease-free survival (DFS) in BRCA patients. The groups are again split according to MALAT1 expression levels. According to the overall survival statistics, patients with lower MALAT1 expression levels have a larger proportion of disease-free survival than those with higher expression levels. This supports MALAT1's potential prognostic use, as greater levels may signal a lower chance of disease-free survival or a higher risk of recurrence in BRCA patients.

The results illustrated from the datasets shows that the differential expression of MALAT1 in malignant vs normal tissues may provide chances for usage as a diagnostic marker. Its expression correlates with patient survival and illness progression, indicating its potential value as a prognostic marker. Regarding therapeutic implications, although not explicitly depicted in the graphs, the noteworthy influence of MALAT1 on survival outcomes suggests that it may be a viable target for treatment, particularly if measures could mitigate its expression or obstruct its activity in BRCA.

Protein-Protein Interactions:

In Fig-3 Protein-Protein interactions MALAT1 plays a variety of roles in the T-cell receptor signalling cascade, including modulating BCL10's nuclear export, promoting TRAF6 activation by oligomerization, and amplifying BCL10-induced NF-kappa-B activation. Th17 cell differentiation and T-cell adhesion are two of the downstream immunological responses triggered by its ubiquitin ligase activity. As a lengthy non-coding RNA, MALAT1 may be able to control the expression of
important elements like MALT1, BCL10, or TRAF6, as well as the expression of ZC3H12A and RC3H1, which MALAT1 cleaves in response to TCR stimulation. MALAT1's regulatory function might help maintain the delicate equilibrium between immune cell responses and functions. Highlighting the increasing intricacy of immune process gene regulation mediated by non-coding RNA.

Discussion
Our research has clarified MALAT1's complex function in BC, enhancing its usefulness as a predictive and diagnostic biomarker. MALAT1's significant overexpression in BC tissues indicates that the disease's molecular pathophysiology involves it [17, 18, 19]. Moreover, the prognostic analyses demonstrate a relationship between elevated MALAT1 levels and reduced patient survival, hence designating MALAT1 as a putative prognostic indicator[20, 21]. Our network and pathway analysis provide mechanistic insights that indicate MALAT1 is implicated in important carcinogenic pathways such as PI3K/AKT, ERK/MAPK, and Wnt/β-catenin. Its role as a miRNA sponge, particularly in sequestering miR-561, affects the stability of TOP2A mRNA, a marker of proliferation and chemotherapy resistance[22]. Notably, the drug repurposing analysis suggests several existing drugs that could potentially target MALAT1 or its associated pathways, providing a promising avenue for BC therapy. This highlights the therapeutic potential of MALAT1, where modulating its expression or function could enhance treatment efficacy and potentially sensitize tumors to chemotherapy[23, 24].

Conclusion
The present study confirms the diagnostic and prognostic relevance of MALAT1 in BC and proposes its potential as a therapeutic target. By establishing the overexpression of MALAT1 in BC and its association with poor survival outcomes, we advocate for the incorporation of MALAT1 in BC management strategies. Our in-silico analysis offers valuable insights into the role of MALAT1 in BC and sets the stage for translational research to develop MALAT1-targeted therapies, marking a step forward in the battle against this pervasive disease.

Acknowledgment:
We would like to acknowledge our fellow colleagues and our departmental staff for helping and motivating us.

Authors Contributions:
All authors contributed equally and sincerely in present research.

Funding:
No external funding was received for the Research.

Conflict of Interest:
No conflict of Interest was declared by any author during the current study.

Abbreviations:
LncRNA: Long Non-Coding RNA
MALAT1: Metastasis-Associated Lung Adenocarcinoma Transcript 1
BC: Breast Carcinoma
GEO: Gene Expression Omnibus
TCGA: The Cancer Genome Atlas
TPM: Transcripts Per Million
ASO: Antisense Oligonucleotides
MDSCs: Myeloid-Derived Suppressor Cells
TAM: Tumor-Associated Macrophages
miRNA: MicroRNA
cRNA: Competitive Endogenous RNA
PD-L1: Programmed Death-Ligand 1
TNBC: Triple-Negative Breast Cancer
ESCC: Esophageal Squamous Cell Carcinoma
GC: Gastric Cancer
NSCLC: Non-Small Cell Lung Cancer
CRC: Colorectal Cancer
CC: Cervical Cancer
HCC: Hepatocellular Carcinoma
ERK/MAPK: Extracellular Signal-Regulated Kinases/Mitogen-Activated Protein Kinases
PI3K/AKT: Phosphoinositide 3-Kinases/Protein Kinase B
TOP2A: Topoisomerase II Alpha

References:


