Gene Expression **2023** vol. 000(000) | 000–000 Epub
DOI: 10.14218/GE.2023.00094

Original Article

Exploring the Role of Seven Key Genes in Breast Cancer: Insights from *In Silico* and *In Vitro* Analyses

Devashish Mehta[1], Jignesh Prajapati[2], Vinal Upadhyay[1], Krupa Shah[1], Vijay Devanhalli[3] and Rakesh M. Rawal[1,2*]

[1]*Department of Life Science, University School of Sciences, Gujarat University, Ahmedabad, India;* [2]*Department of Biochemistry & Forensic Science, University School of Sciences, Gujarat University, Ahmedabad, India;* [3]*Department of Surgery, HCG Cancer Centre, Science City Road, Sola, Ahmedabad, India.*

Received: August 16, 2023 | **Revised:** September 20, 2023 | **Accepted:** October 26, 2023 | **Published online:** January 02, 2024

Abstract

Background and objectives: Breast cancer remains a significant global health concern, warranting further exploration into its genetic basis and potential therapeutic targets. This study aimed to elucidate the genetic associations of seven pivotal genes with breast cancer and discern their potential role in disease prognosis.

Methods: The genes *VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4* were curated from over 30 articles. Their association with breast cancer was analyzed using both *in silico* and *in vitro* techniques. The *in silico* assessment involved constructing a protein-protein interaction network, accompanied by Gene Ontology and pathway enrichment analysis. Further, survival and expression analysis were conducted using Kaplan-Meier Plotter and the UALCAN database respectively. At the protein level, expression was observed using the Human Protein Atlas database. The *in vitro* validation involved analyzing mRNA expression levels in 10 breast cancer tissue samples.

Results: The study revealed that all seven genes are significantly upregulated in breast cancer tissues compared to normal tissues, highlighting their critical role in tumor development and progression. Protein-protein interaction analysis confirmed their central involvement in vital biological processes related to breast cancer. Survival analysis showed that high expressions of these genes are associated with poorer patient prognosis, with hazard ratios indicating their potential as prognostic markers. *In vitro* validation further supported their overexpression in breast cancer, suggesting their importance in molecular landscape of the disease and their value as targets for therapeutic intervention.

Conclusions: This study provides a profound understanding of the genetic landscape of breast cancer, emphasizing the significance of the selected seven genes in the pathology of the disease. These findings suggest potential avenues for therapeutic targeting and the advancement of personalized medicine in breast cancer treatment.

Keywords: Breast cancer; Protein-protein interaction; Gene ontology; Kaplan-Meier plotter; UALCAN; Gene expression.

Abbreviations: BC, breast cancer; cDNA, complementary DNA; GO, gene ontology; HER2, human epidermal growth factor receptor 2; HPA, Human Protein Atlas; HR, hazard ratio; KEGG, Kyoto Encyclopedia of Genes and Genomes; PCR, polymerase chain reaction; PPI, protein-protein interaction; qRT-PCR, quantitative reverse transcriptase PCR; SEM, standard error of the mean; STRING, Search Tool for the Retrieval of Interacting Genes; TCGA, The Cancer Genome Atlas; UALCAN, The University of ALabama at Birmingham CANcer data analysis Portal.

***Correspondence to:** Rakesh M. Rawal, Department of Life Science, University School of Sciences, Gujarat University, Ahmedabad 380009, Gujarat, India. ORCID: https://orcid.org/0000-0002-7985-1187. Tel: +91 99252 44855, E-mail: rakeshrawal@gujaratuniversity.ac.in; rakeshrawal.gu@gmail.com

How to cite this article: Mehta D, Prajapati J, Upadhyay V, Shah K, Devanhalli V, Rawal RM. Exploring the Role of Seven Key Genes in Breast Cancer: Insights from *In Silico* and *In Vitro* Analyses. *Gene Expr* 2023;000(000):000–000. doi: 10.14218/GE.2023.00094.

Introduction

Globally, breast cancer (BC) is the leading type of cancer in women, making up roughly one-fourth of all cancer diagnoses and contributing to 15% of cancer-related fatalities.[1] Even with technological progress in early detection and treatment options, BC continues to pose a substantial healthcare issue. There are marked variations in survival outcomes depending on geographic locations and demographic groups.[2] This highlights the pressing necessity to further decode the molecular intricacies that drive the onset and advancement of BC, and to pinpoint new diagnostic markers and treatment avenues. One of the hallmarks of BC is its genetic heterogeneity, reflecting the diversity of mutations, gene expression patterns, and epigenetic alterations that occur in different subtypes and stages of the disease.[3] Several genes have been implicated in

BC pathogenesis, either as drivers of oncogenesis or as modifiers of tumor behavior and response to therapy.[4] However, the functional roles and interactions of these genes in BC remain incompletely understood. The advent of computational biology and bioinformatics techniques has revolutionized our understanding of complex diseases like BC. These techniques allow for the analysis of large-scale genomic data to identify genetic alterations associated with disease and to understand their functional implications.[5] In particular, network analysis can reveal the complex interactions among genes and their involvement in key biological processes and pathways.[6]

In this study, we leveraged these advanced techniques to focus on seven key genes, *VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4,* which have been implicated in various aspects of BC biology. These genes were selected based on a comprehensive literature review, which revealed their involvement in various aspects of BC biology, such as angiogenesis, DNA repair, cell cycle regulation, and apoptosis.[7–10] For instance, mutations in the *BRCA1* and *BRCA2* genes are well-known risk factors for hereditary BC.[11] However, these mutations account for only a small fraction of all BCs, suggesting that other genes may also play significant roles.

A combination of *in silico* and *in vitro* analysis was used to gain insights into the roles of these genes in BC. The *in silico* analysis, which included protein-protein interaction (PPI) network analysis, gene ontology (GO) enrichment, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, allowed us to examine the interactions among these genes and their involvement in key biological processes and pathways. Kaplan-Meier survival analysis and The University of ALabama at Birmingham CANcer data analysis Portal (UALCAN) database expression analysis provided insights into the prognostic value of these genes in BC. The *in vitro* analysis enabled us to validate these findings in actual BC tissues. These methods provided a comprehensive view of the roles of these genes in BC and their potential as therapeutic targets. Our results provide new insights into the genetic landscape of BC and highlight potential targets for therapeutic intervention.

Materials and methods

Manual curation of 246 genes related to BC was done from 36 research papers that were impactful and relevant to the field. The 246 genes and their references are listed in Table S1. Seven genes (*VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4*) were selected for analysis, based on how often they appeared in the literature, their known biological role, and their potential as therapeutic targets. This allowed for a deeper investigation of their function in BC in the *in silico* and *in vitro* analyses. To provide comprehensive information, a compiled list of collected research articles and the list of 246 genes studied is shown in Table S1.

PPI network analysis

The first step in the *in silico* analysis was the construction of a PPI network of the seven selected genes. It was achieved using the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/), a widely recognized tool for evaluating gene interactions. The PPI network was constructed using a confidence score threshold of >0.5 to ensure the reliability of the interactions. This score represents the likelihood of the interaction being genuine, with higher scores indicating a higher level of confidence.[12]

GO and pathway enrichment analysis

GO and pathway enrichment analyses were performed using the online platform https://www.bioinformatics.com.cn, which is designed for data analysis and visualization. GO enrichment and KEGG pathway analysis were carried out using this platform.[13,14] GO analysis included three categories, biological processes, molecular functions, and cellular components. This analysis provided insights into the biological functions and protein lists associated with the seven selected genes. Statistical significance was determined by a p-value of < 0.05. This analysis offered a deeper understanding of the functional roles and pathways these genes are involved in, further elucidating their significance in BC.

Survival analysis and expression of selected genes

To assess the relationship between the selected genes and BC patient survival time, a survival analysis was performed using the Kaplan-Meier Plotter (http://kmplot.com/analysis/). Gene expression levels were segmented into two groups, those above the median were high, and those at or below the median were low. This method of categorization allowed us to distinguish survival outcomes between the two groups.[15] To evaluate the relationship between gene expression and survival rates, a hazard ratio (HR) was calculated. A lower HR value implies a more favorable prognosis, whereas a higher HR points to an increased risk and less favorable outcome.

Additionally, the correlation between the expression of the key genes was evaluated using UALCAN (http://ualcan.path.uab.edu/). This tool was also used to examine the relationship between gene expression and other clinical parameters, including tumor stages, subclasses, and ethnicity.[16,17] Statistical significance was determined by a p-value of < 0.05. This comprehensive analysis provided valuable insights into the prognostic value of these genes in BC.

Protein levels of selected genes in the HPA database

In addition to the gene expression and survival analysis, the protein levels of the selected genes in tumors and normal tissues were evaluated. This was achieved using the Human Protein Atlas (HPA) database (https://www.proteinatlas.org/), which contains immunohistochemistry-based expression data for specific human tissues.[18] This analysis provided a comprehensive view of the protein expression levels of the seven genes, offering additional insights into their roles in BC. Comparing the protein levels in tumor tissues with those in normal tissues helped to further elucidate the potential impact of these genes on BC pathology.

Patient inclusion and tissue sampling process

The current investigation included a cohort of 10 patients, each undergoing tumor resection procedures at the HCG Cancer Centre located in Ahmedabad, Gujarat. A distinctive feature of our research design was the procurement of two types of tissue samples from each patient, BC tumors and adjacent normal tissue. The patient selection criteria ensured that none of the participants had been treated with radiation therapy or chemotherapy prior to surgery. This was to ensure that the gene expression patterns were not affected by their treatment, allowing for a better understanding of the native tumor and adjacent tissue gene profiles.

For each of the 10 patients, BC tumor tissue was collected together with a sample of adjacent normal tissue, ensuring a minimum separation distance of 5 cm from the tumor margin. This was done to lower the risk of including potentially cancer-affected cells within the normal tissue sample. On collection, the tissues were rapidly frozen in liquid nitrogen, a process that instantaneously halts biological activity and preserves the integrity of cellular

Table 1. Primer sequences for the selected genes

Gene	Primer sequence	No of bases
VEGFA	Forward: 5′-CTTGCCTTGCTGCTCTACC-3′	19
	Reverse: 5′-CACACAGGATGGCTTGAAG-3′	19
BRCA1	Forward: 5′-ACAGCTGTGTGGTGCTTCTGTG-3′	22
	Reverse: 5′-CATTGTCCTCTGTCCAGGCATC-3′	22
CHEK1	Forward: 5′-CCCGCACAGGTCTTTCCTT-3′	19
	Reverse: 5′-GGCTGGGAAAAGCTGATCC-3′	19
CDK1	Forward: 5′-TGGATCTGAAGAAATACTTGGATTCTA-3′	27
	Reverse: 5′-CAATCCCCTGTAGGATTTGG-3′	20
CCNB1	Forward: 5′-CTTAGACAAATTCTGAACTAGTGTACA-3′	27
	Reverse: 5′-ATTCTTGACAACGGTGAAT-3′	19
RAD51	Forward: 5′-CCAGACCCAGCTCCTTTACC-3′	20
	Reverse: 5′-CACTGCGACACCAAACTCATC-3′	21
XRCC4	Forward: 5′-CTGAGGAGGATGGGCTTTATGAT-3′	23
	Reverse: 5′-CAAGATTTGTCTGCATTCGGTGT-3′	23
18S rRNA	Forward: 5′-GGAGTATGGTTGCAAAGCTGA-3′	21
	Reverse: 5′-ATCTGTCAATCCTGTCCGTGT-3′	21

components. They were then stored at an ultra-low temperature of −80 °C, which is optimal for long-term preservation of biological samples, until they were retrieved for further examination. The participating patient ages varied at the time of diagnosis, ranging from 35 to 55 years. All participants provided written informed consent affirming their voluntary participation in the study. This commitment ensured ethical transparency and protected patient rights. Furthermore, the study adhered to the Declaration of Helsinki (revised in 2013) and underwent review and approval by the Ethics Committee and Institutional Review Board of Gujarat University (approval number GU/IEC/08/2017), providing additional ethical assurance for this research undertaking.

RNA isolation and quantitative reverse transcriptase (qRT)-PCR

Total RNA was isolated from 25 mg samples of tumor tissue and adjacent normal tissue. The extraction process was carried out using Trizol reagent (Thermo Fisher Scientific, Invitrogen Bioservices India Pvt Ltd.), as per the manufacturer's instructions. Briefly, the tissue samples were homogenized in the Trizol reagent and subsequently subjected to chloroform extraction. Following this, the RNA was precipitated from the aqueous phase by isopropanol, washed with ethanol, and resuspended in RNAse-free water. The isolated RNA was then converted to complementary DNA (cDNA) to facilitate downstream applications. The iScript cDNA Synthesis kit (Bio-Rad, Hercules, CA, USA) was employed for this purpose. This process involved combining the extracted RNA with the reaction mix and primers, followed by incubation at a specific temperature sequence to enable reverse transcription.

Following cDNA synthesis, the gene expression levels were measured by qRT-PCR using a Quant Studio 5 platform (Applied Biosystems Inc., Waltham, MA, USA). The qRT-PCR mix included cDNA, Brilliant III SYBR-Green dye (Agilent), and primers specific to the genes of interest (*VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4*) and the internal control (18S rRNA). The primer sequences are listed in Table 1. The qRT-PCR protocol commenced with an initial denaturation phase (5 m at 95°C), followed by a 40-cycle amplification and quantification program. Each cycle of the program included a 15 s denaturation phase at 95°C and a 1 m annealing/extension phase at 60°C. To ensure specificity and avoid amplification of nontarget sequences, each qRT-PCR run included a melting curve analysis, a process that identifies the unique thermal profile of the amplified products. The relative gene expression levels were computed using the $2^{-\Delta\Delta Ct}$ method, which considers the efficiency of PCR amplification and enables the comparison of gene expression levels across different samples. In this calculation, the expression levels of the genes of interest were normalized to the expression level of the internal *18S rRNA* control.

Results

PPI network analysis

The initial step in the *in silico* analysis involved constructing a PPI network for the seven selected genes using STRING. The network, shown in Figure 1, revealed significant enrichment of interactions among the seven genes, as indicated by a PPI enrichment *p*-value of 3.11 e^{-05}. This suggests that these genes were biologically interconnected as a group in the context of BC. The network consisted of seven nodes and 14 edges, with an average node degree of 4 and an average local clustering coefficient of 0.876, indicating a high degree of interconnectivity among the genes. These results underscore the potential significance of these seven genes in BC, providing a foundation for further analysis.

GO enrichment and KEGG analysis

GO Enrichment and KEGG analysis of the seven selected genes, *VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4*, revealed their significant involvement in various biological processes, cellular components, molecular functions, and key biologi-

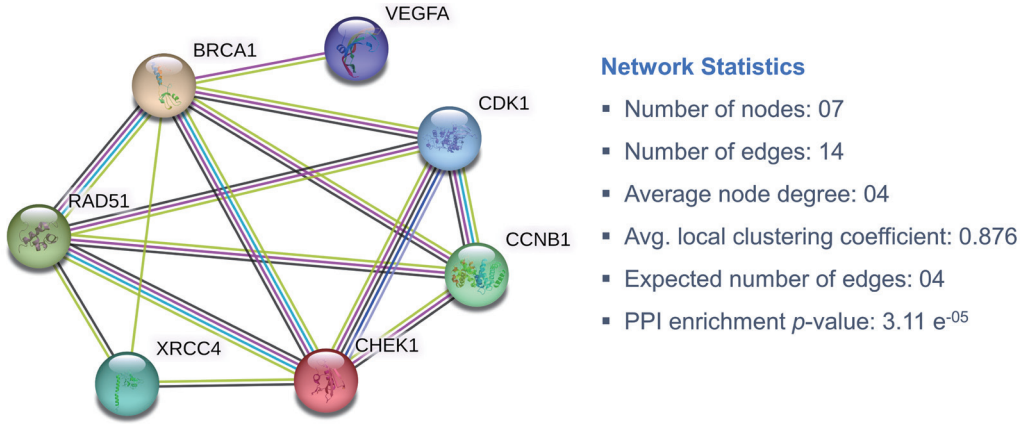

Fig. 1. PPI network of seven selected genes with statistics using STRING. PPI, protein-protein interaction; STRING, Search Tool for the Retrieval of Interacting Genes.

cal pathways, all of which are critical in the context of BC biology. The GO enrichment and KEGG analysis are shown in Figure 2 and in Supplementary Tables S2–S5. The results demonstrated that these genes participate in critical biological processes such as responses to DNA damage, regulation of the cell cycle, and chromatin modification. Notably, all of the genes except *XRCC4* were involved in these processes, suggesting their significant roles in the development and progression of BC. In addition, the

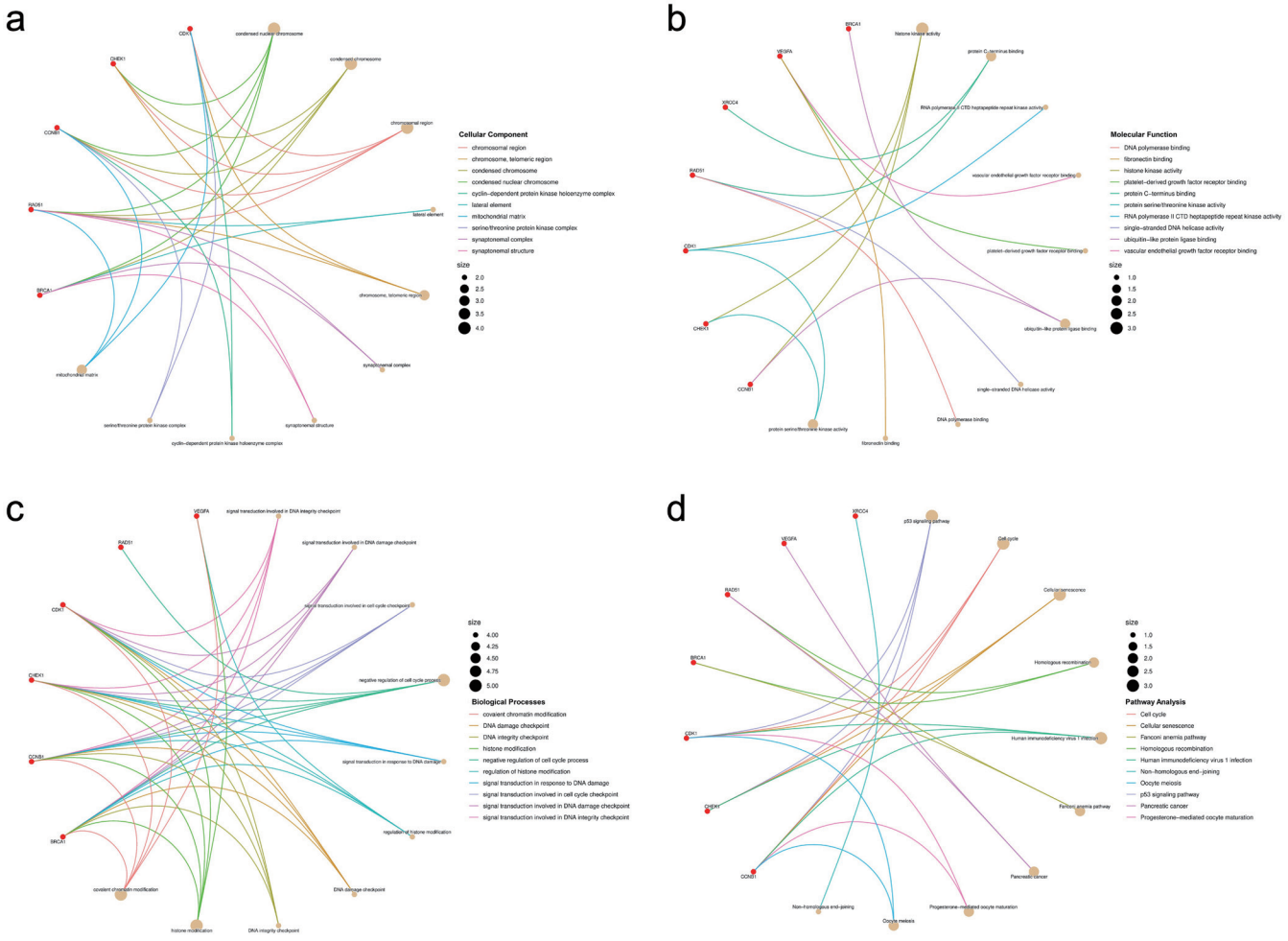

Fig. 2. Functional enrichment analysis of selected genes. (a) Cellular component. (b) Molecular function. (c) Biological process. (d) KEGG pathways. KEGG, Kyoto Encyclopedia of Genes and Genomes.

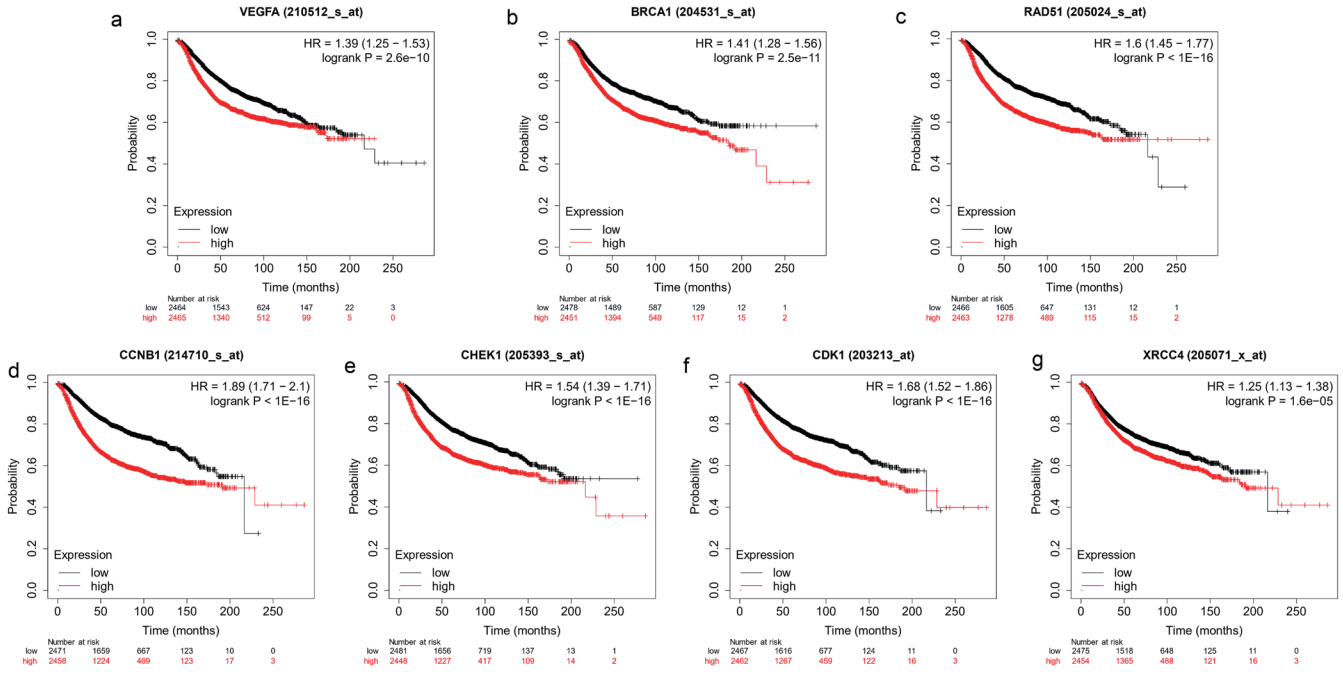

Fig. 3. Kaplan-Meier survival curves illustrating the prognostic values of seven key genes in breast cancer. This figure demonstrates the survival impact of high versus low expression levels of (a) *VEGFA*, showing its influence on angiogenesis and tumor progression; (b) *BRCA1*, highlighting its role in DNA repair and impact on survival; (c) *RAD51*, emphasizing its function in homologous recombination repair; (d) *CCNB1*, indicating its involvement in cell cycle regulation; (e) *CHEK1*, illustrating its importance in DNA damage response; (f) *CDK1*, demonstrating its critical role in cell cycle progression; and (g) *XRCC4*, showcasing its part in DNA repair mechanisms. Each plot corresponds to a different gene, with HR and *p*-value indicating the statistical significance of its association with patient survival. Higher HR values suggest a greater risk and poorer prognosis with high gene expression. BC, breast cancer; HR, hazard ratio.

genes were found to be associated with key cellular components such as chromosomal regions, telomeric regions of chromosomes, condensed chromosomes, condensed nuclear chromosomes, and others. These associations indicate the crucial roles of genes in various cellular processes such as cell cycle regulation, DNA replication and repair, protein phosphorylation, and signal transduction. Molecular function analysis revealed the association of genes with several key molecular functions, including DNA polymerase binding, fibronectin binding, histone kinase activity, and others. These functions suggest the critical roles of genes in various molecular processes such as DNA replication and repair, protein binding and interaction, kinase activity, and signal transduction.

KEGG pathway analysis further revealed the significant involvement of these seven genes in several key biological pathways. Three genes, *CCNB1*, *CHEK1*, and *CDK1*, were implicated in the p53 signaling pathway, cell cycle, cellular senescence, and human immunodeficiency virus 1 infection. These pathways are crucial for cell growth regulation and DNA damage response, both of which are critical processes in cancer development and progression. The genes *BRCA1* and *RAD51*, both involved in DNA repair, were found to be involved in the homologous recombination and Fanconi anemia pathways. Dysregulation of these pathways can lead to genomic instability, a hallmark of cancer. *VEGFA*, a gene involved in angiogenesis, is implicated in the pancreatic cancer pathway. While that pathway is not directly related to BC, angiogenesis is a common process in many types of cancer by promoting tumor growth and metastasis. The *XRCC4*, gene also involved in DNA repair, was highlighted in the non-homologous end-joining pathway. Like *BRCA1* and *RAD51*, defects in *XRCC4* contribute to genomic instability and cancer development. In addition to these pathways, the analysis identified several other pathways where these genes have significant roles, including viral carcinogenesis, microRNAs in cancer, the phosphoinositide-3-kinase-Akt signaling pathway, bladder cancer, and the vascular endothelial growth factor signaling pathway.

Survival analysis

Kaplan-Meier survival analysis revealed significant associations between the expression of the seven selected genes and overall survival in BC patients. HRs ranged from 1.25 (*XRCC4*) to 1.68 (*CDK1*), indicating varying degrees of risk associated with high expression of these genes. High expression of *CDK1*, with an HR of 1.68, was associated with the highest risk of shorter overall survival, while high expression of *XRCC4*, with an HR of 1.25, was associated with the lowest risk within this gene set. Importantly, all these associations were statistically significant, with p-values < 0.05, underscoring the potential of these genes as prognostic biomarkers in BC, with varying degrees of impact on patient survival. These findings are shown in Figure 3 along with the HRs and p-values for each gene.

Expression analysis

Assessment of the expression of the seven pivotal genes was conducted with the UALCAN platform. This evaluation scrutinized how gene expression correlated with multiple clinical factors, such as stage of tumor development, cancer subclass, and ethnic background. In the comparison of normal and primary tumor tissues, all seven genes had significant differences in expression (all $p < 0.001$; Fig. 4). This suggests that these genes were upregulated in primary tumor tissues. When examining individual

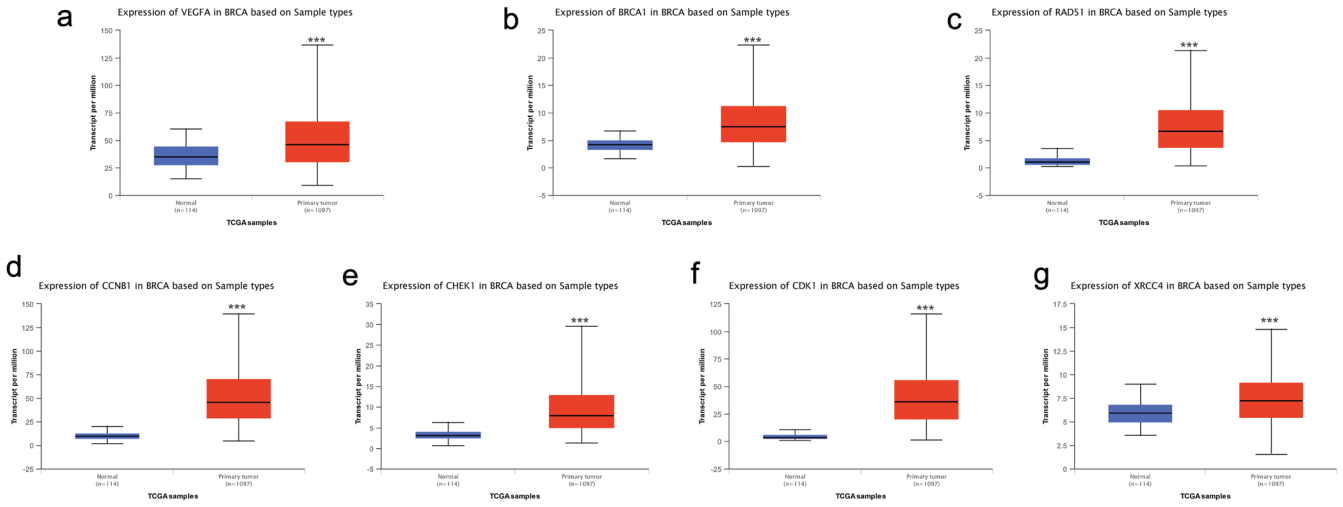

Fig. 4. Expression validation in UALCAN database for seven genes (data from TCGA database). (a) *VEGFA*. (b) *BRCA1*. (c) *RAD51*. (d) *CCNB1*. (e) *CHEK1*. (f) *CDK1*. (g) *XRCC4*. Data are means ± SEMs. ***$p < 0.001$. CGA, The Cancer Genome Atlas; SEM, standard error of mean; UALCAN, The University of ALabama at Birmingham CANcer data analysis Portal.

cancer stages, there were significant differences in gene expression across different stages. For instance, high expression of all seven genes was associated with stage 2 compared with normal tissues. However, the expression of these genes varied across different stages, with some having higher expression in later stages of cancer (Fig. 5).

In terms of subclasses of BC, all seven genes had significantly higher expression in luminal, human epidermal growth factor receptor 2 (HER2) positive, and triple negative BC subclasses compared with normal tissues (Fig. 6). Lastly, when comparing gene expression based on ethnicity, all seven genes had significantly higher expression in Caucasian, African American, and Asian populations compared with normal tissues (Fig. 7). The detailed results of this analysis are shown in Tables S6–S9. These results suggest that the expression of these seven key genes is significantly associated with BC and varies across different clinical parameters. These genes could potentially serve as valuable prognostic markers for BC.

Validation of the genes within the HPA database

The expression levels of the identified hub genes were confirmed to be significantly upregulated through immunostaining data obtained from the HPA database. In BC tissues, the genes *VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CDK1*, and *XRCC4* were overexpressed compared with normal tissues (Fig. 8). However, it should be noted that data for *CHEK1* was not available in the HPA database.

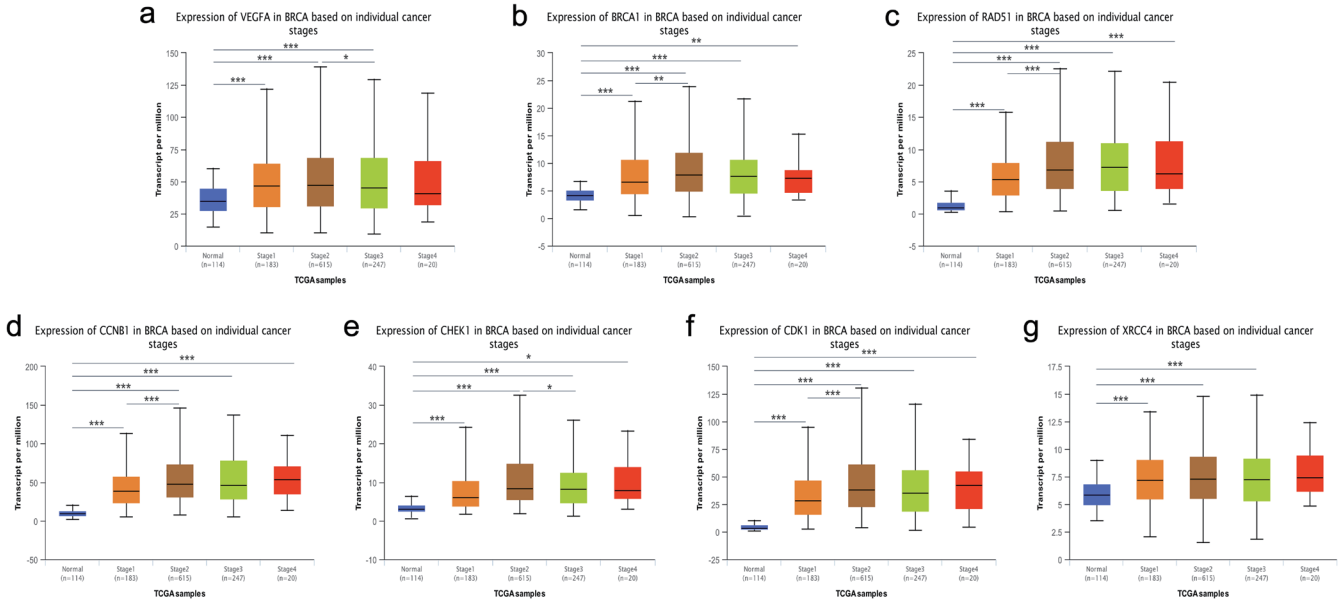

Fig. 5. Relative expression of seven genes in normal tissues and BC tissues with different stages. (a) *VEGFA*. (b) *BRCA1*. (c) *RAD51*. (d) *CCNB1*. (e) *CHEK1*. (f) *CDK1*. (g) *XRCC4*. Data are means ± SEMs. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. BC, breast cancer; SEM, standard error of mean.

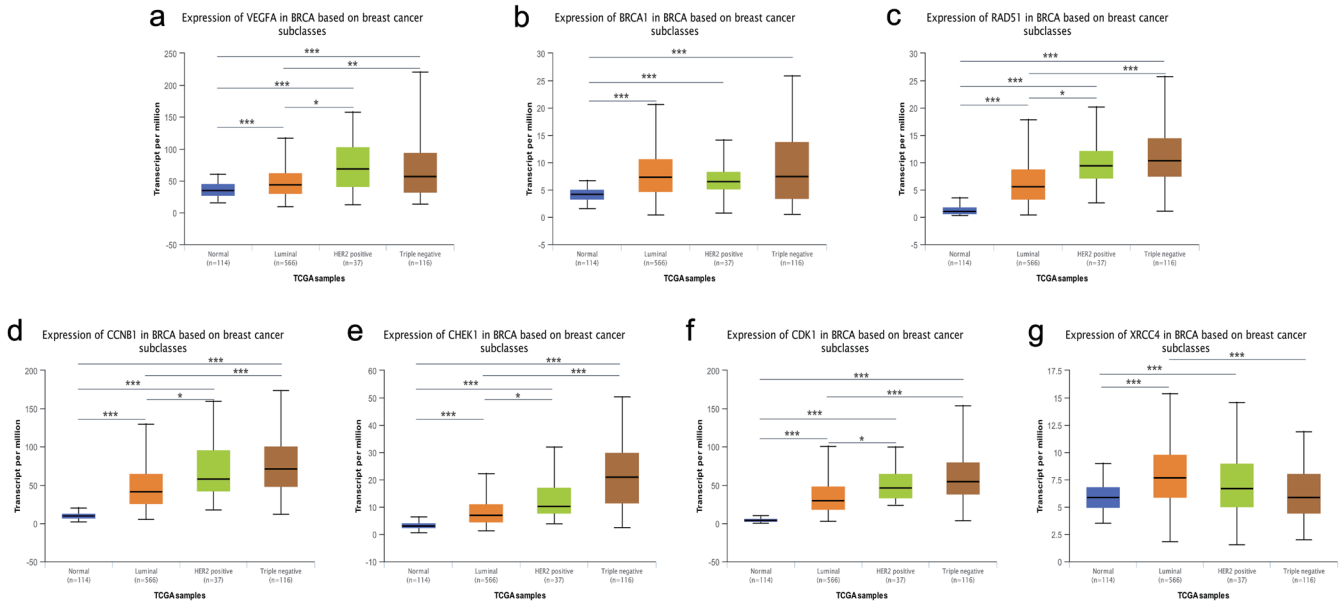

Fig. 6. Relative expression of seven genes in normal tissues and BC tissues of different subclasses. (a) *VEGFA*. (b). *BRCA1*. (c) *RAD51*. (d) *CCNB1*. (e) *CHEK1*. (f) *CDK1*. (g) *XRCC4*. Data are means ± SEMs. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. BC, breast cancer; SEM, standard error of mean.

Despite this, the validation of the other genes underscores their potential role as pivotal contributors to the development and progression of BC.

In vitro validation of gene expression

In this study, a cohort of 10 patients undergoing tumor resection procedures at HCG Cancer Center in Ahmedabad, Gujarat, was examined. Two types of tissue samples were procured from each patient, BC tumors and adjacent normal tissue. The selection criteria ensured that none of the participants had undergone radiation therapy or chemotherapy prior to their surgical intervention, preserving the native gene expression patterns in the tumor and adjacent tissue. Expression levels of seven genes of interest (*VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, *XRCC4*) and the internal control (18S rRNA) were measured by qRT-PCR. The relative gene expression levels were computed using the $2^{-\Delta\Delta Ct}$ method, with the expression levels of the genes of interest normalized to the expression level of the internal control, 18S rRNA.

The results revealed significant overexpression of all seven genes in the BC tumor tissues compared with adjacent normal

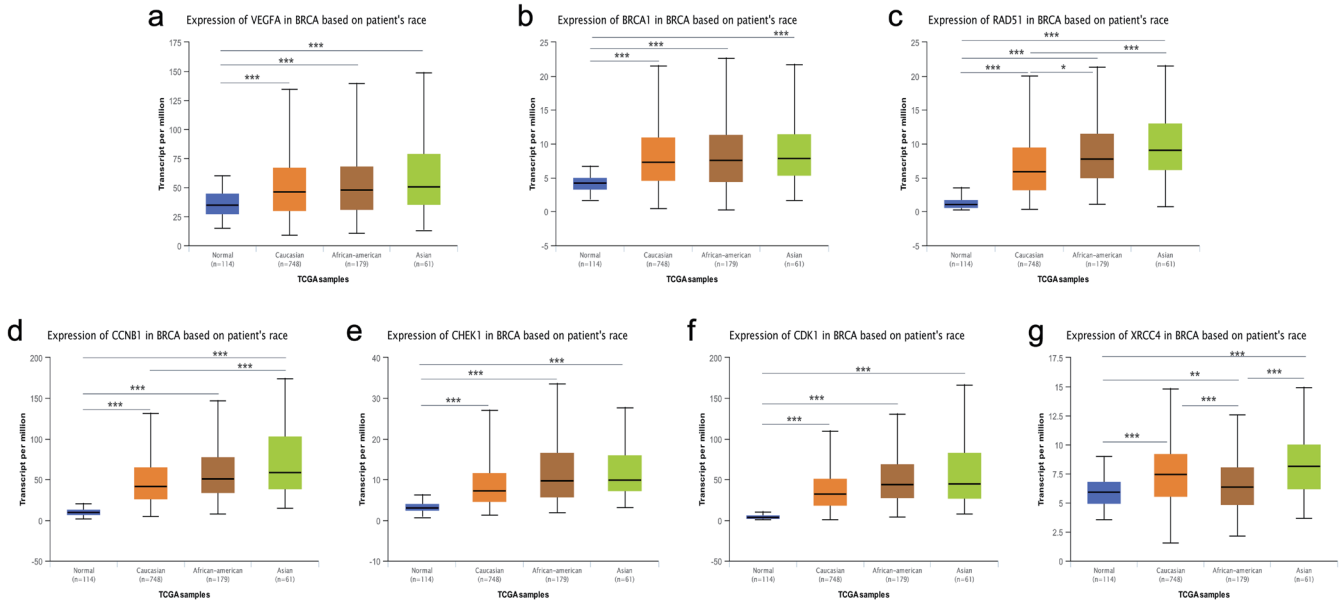

Fig. 7. Relative expression of seven genes in normal tissues and BC tissues from different races. (a) *VEGFA*. (b) *BRCA1*. (c) *RAD51*. (d) *CCNB1*. (e) *CHEK1*. (f) *CDK1*. (g) *XRCC4*. Data are means ± SEMs. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. BC, breast cancer; SEM, standard error of mean.

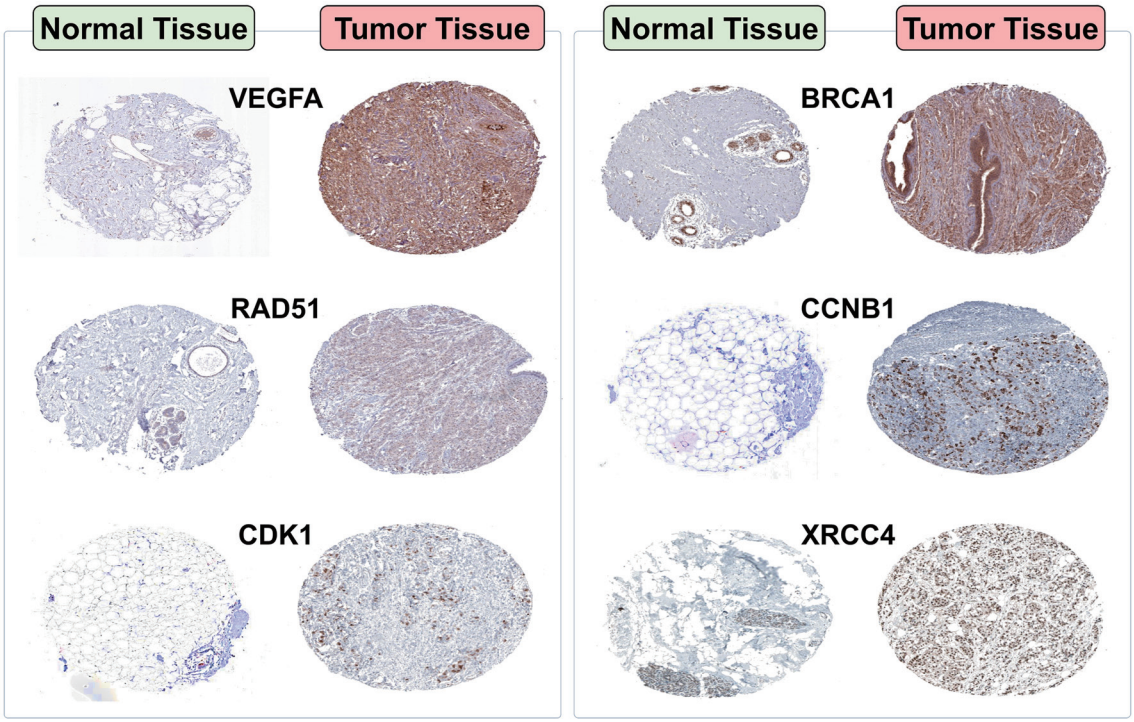

Fig. 8. Protein level analysis of genes with differential expression in the Human Protein Atlas database. (Magnification: 200×).

tissues. The fold change in expression ($2^{-\Delta\Delta Ct}$ values) and their standard deviations are shown in Figure 9. These results were significant for all genes except *VEGFA*, as indicated by the *p*-values calculated by *t*-tests. The overexpression of these genes in BC tumor tissues compared with adjacent normal tissue provides valuable insights into the molecular mechanisms underlying cancer development and progression, revealing potential targets for therapeutic intervention.

Discussion

BC remains a significant global health concern, with high incidence and mortality rates among women worldwide.[19] The complex genetic basis of BC necessitates comprehensive studies to understand its genetic underpinnings and identify potential therapeutic targets. This study contributes to this understanding by focusing on seven key genes, *VEGFA*, *BRCA1*, *RAD51*, *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4*. The seven genes selected for this study have been implicated in various aspects of BC biology. For instance, *VEGFA* is known to have a crucial role in angiogenesis, a process that promotes tumor growth and metastasis.[20] Similarly, *BRCA1* and *RAD51* are involved in DNA repair, and defects in these genes can lead to genomic instability, a hallmark of cancer.[21] Moreover, *CCNB1*, *CHEK1*, and *CDK1* are implicated in cell cycle regulation and DNA damage response, processes critical in cancer development and progression.[22] Finally, *XRCC4*, like *BRCA1* and *RAD51*, is involved in DNA repair, and its defects can contribute to genomic instability and cancer development.[23]

The approach taken in this investigation used both *in silico* and *in vitro* methods to offer an exhaustive scrutiny of the selected genes. Additionally, part of the research design included building a PPI network, facilitated by using the STRING database and Cytoscape software for analysis. This approach allowed us to visualize and analyze the interactions among the seven genes of interest and their potential partners in the context of BC. The PPI network analysis also enabled us to identify the most significant clusters and hub genes within the network, which may represent key regulators or mediators of BC biology. Furthermore, we performed GO enrichment and KEGG pathway analysis to explore the functional roles and pathways associated with the genes in the network. The *in silico* methodology is a useful tool to complement the experimental data and generate new hypotheses or insights into the molecular mechanisms of BC. By integrating multiple sources of information, such as gene expression, protein interactions, functional annotations and pathways, the *in silico* methodology provides a comprehensive and holistic view of complex biological systems. Moreover, the *in silico* methodology can help prioritize genes or pathways for further validation or intervention, as well as identify potential biomarkers or therapeutic targets.[24,25]

Several previous studies have used similarly *in silico* methodologies to investigate BC at different levels, such as gene expression, copy number variation, methylation, microRNA or long noncoding RNA. For example, Dashti *et al*.[26] used a system biology strategy to identify key genes and signaling pathways in BC based on gene expression data from two microarray datasets. They also performed survival analysis to evaluate the prognostic value of the identified genes. Gur *et al*.[21] used real-time PCR-based expression profiling of *BRCA1*-induced genes in primary breast tumors and performed functional interaction and gene enrichment analysis using Cytoscape. They also assessed the clinical importance of the obtained genes using Kaplan-Meier survival analysis. Kour *et al*.[27] used karyotypic analysis to detect chromosomal aberrations in BC patients and performed *in-silico* analysis using Cytoscape to identify genes and pathways associated with chromosomal instability. These studies demonstrate the applicability and utility of the *in*

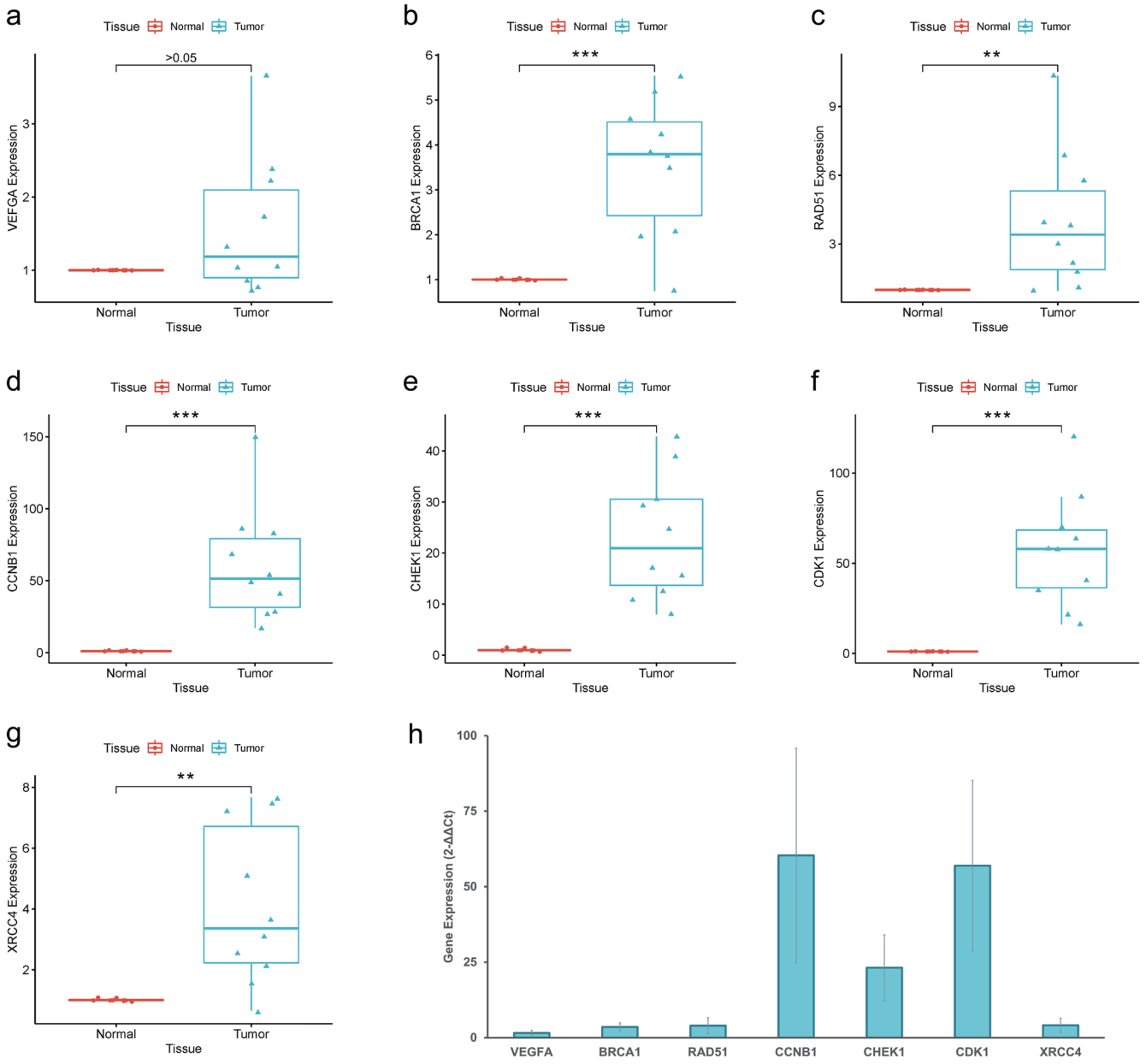

Fig. 9. Quantitative real-time PCR results for the selected seven genes. (a) *VEGFA*, (b) *BRCA1*, (c) *RAD51*, (d) *CCNB1*, (e) *CHEK1*, (f) *CDK1*, and (g) *XRCC4*, showing significant upregulation in tumor tissues compared to normal tissues, with *p*-values indicating strong statistical significance. The ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$ denote statistical significance levels in the t-test for the individual genes. (h) Bar graph representing the collective overexpression profile of all genes, normalized against 18S rRNA expression. Data are means ± standard deviations. PCR, polymerase chain reaction.

silico methodology in BC research.

In vitro validation confirmed the overexpression of all seven genes in BC tumor tissues compared with adjacent normal tissues. This finding aligns with existing literature, further substantiating the roles of these genes in BC. For instance, the role of *VEGFA* in promoting angiogenesis in BC has been well-documented.[28,29] Likewise, the involvement of *BRCA1* and *RAD51* in DNA repair and their association with BC have been previously reported. Interestingly, a previous study using qRT-PCR showed that *BRCA1*-induced genes were differentially expressed in primary breast tumors compared with normal tissues, indicating that *BRCA1* may modulate the expression of downstream genes involved in various cellular processes.[30] Furthermore, the involvement of *CCNB1*, *CHEK1*, and *CDK1* in cell cycle progression and DNA damage response in BC cells has been reported in several studies. For instance, a study by Wang *et al*.[31] found that *CCNB1*, along with *CDK1*, had a complex interrelation with other genes, promoting cancer progression through multiple pathways. Another study by Fang *et al*.[32] highlighted *CCNB1* and *CDK1* as potential antiBC drug targets and prognostic markers, as their overexpression was strongly associated with a low survival rate in BC patients. Similarly, the role of *XRCC4* in DNA repair and genomic stability in BC cells has been documented. A study by

Cheng *et al*.[33] found that overexpression of *XRCC4* and other genes was associated with resistance to CDK4/6 inhibitors, a class of drugs used in BC treatment. These studies provide further evidence supporting the findings of this study, underscoring the significant roles of *CCNB1*, *CHEK1*, *CDK1*, and *XRCC4* in BC.

In addition to the broader landscape of BC genetics, it is important to note specialized subtypes like invasive micropapillary carcinoma. Recent studies have highlighted unique molecular pathways associated with this subtype.[34] Although our study did not specifically focus on invasive micropapillary carcinoma, some of the key genes we analyzed, such as *BRCA1* and *RAD51*, have been implicated in complex DNA repair mechanisms that might be relevant in this context.[35] The clinical significance of *HER2* expression in ductal carcinoma *in situ* has recently garnered attention. *HER2* overexpression has been associated with a more aggressive phenotype and could potentially alter therapeutic strategies for ductal carcinoma *in situ*.[36] While our study did not focus on *HER2*, the genes we examined may interact with HER2 signaling pathways, warranting future investigation. Adipokines are another critical component in the multifaceted landscape of BC. Recent evidence suggests that adipokines may interact with key molecular pathways involved in BC progression.[37] Given that some of the genes we analyzed, like *VEGFA*, are involved in angiogenesis, a process that could be influenced by adipokines, there may be significant interplay worth exploring in future studies.

Overexpression of these genes in BC tumor tissues, as confirmed by the *in vitro* validation, provides valuable insights into the molecular mechanisms underlying cancer development and progression, revealing potential targets for therapeutic intervention. Despite the promising findings, the present study is not without its limitations. The relatively small sample size may limit the generalizability of the results. Furthermore, the experimental design did not include functional assays to assess the effects of gene knockdown or overexpression on cell viability, apoptosis, migration, or invasion. These limitations highlight the need for further validation with larger cohorts and more robust methods to confirm the findings.

Conclusions

This study contributes to the understanding of the genetic landscape of BC and underscores the potential of seven genes studied as targets for therapeutic intervention. The findings emphasize the importance of these genes in the pathology of BC and pave the way for personalized medicine. Future research should focus on exploring the therapeutic potential of these genes and developing targeted therapies for BC. This could involve functional studies to elucidate the precise roles of these genes in BC and preclinical studies to evaluate the efficacy of targeted therapies. Ultimately, these efforts could lead to improved treatment strategies and outcomes for patients with BC.

Acknowledgments

JP acknowledges the University Grants Commission, New Delhi, India for providing fellowship for the award of CSIR-NET Junior Research Fellowship. The authors acknowledge the Department of Biochemistry and Forensic Science, School of Sciences, Gujarat University, and HCG Cancer Centre, Ahmedabad for providing samples to perform experiments.

Funding

This work was supported by the Research Support Scheme (RSS) of the Gujarat State Biotechnology Mission [GSBTM/JDR&D/608/2020/446-447].

Conflict of interest

All authors declare no conflict of interest.

Author contributions

Conceptualized the experiments (DM, VD, RR), curated the data (DM), performed the initial gene data analysis (VU, JP), performed the *in silico* experiments (JP), analyzed the data and wrote the manuscript (JP, DM, KS), supervised the research and revised the manuscript (RR). All authors discussed the manuscript critically for important intellectual content and approved the final version of the manuscript.

Ethics statement

All participants provided written informed consent affirming their voluntary participation in the study. This commitment ensured ethical transparency and protected patient rights. Furthermore, the study adhered to the Declaration of Helsinki (revised in 2013) and underwent review and approval by the Ethics Committee and Institutional Review Board of Gujarat University (approval number GU/IEC/08/2017), providing additional ethical assurance for this research undertaking.

Data sharing statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] 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(6):394–424. doi:10.3322/caac.21492, PMID:30207593.

[2] DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, *et al*. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69(6):438–451. doi:10.3322/caac.21583, PMID:31577379.

[3] Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, *et al*. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534(7605):47–54. doi:10.1038/nature17676, PMID:27135926.

[4] Stratton MR, Rahman N. The emerging landscape of breast cancer susceptibility. Nat Genet 2008;40(1):17–22. doi:10.1038/ng.2007.53, PMID:18163131.

[5] Beerenwinkel N, Greenman CD, Lagergren J. Computational cancer biology: an evolutionary perspective. PLoS Comput Biol 2016;12(2):e1004717. doi:10.1371/journal.pcbi.1004717, PMID:26845763.

[6] Creixell P, Reimand J, Haider S, Wu G, Shibata T, Vazquez M, *et al*. Pathway and network analysis of cancer genomes. Nat Methods 2015;12(7):615–621. doi:10.1038/nmeth.3440, PMID:26125594.

[7] Ferrara N, Gerber HP, LeCouter J. The biology of VEGF and its receptors. Nat Med 2003;9(6):669–676. doi:10.1038/nm0603-669, PMID:12778165.

[8] Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat Rev Cancer 2011;12(1):68–78. doi:10.1038/nrc3181, PMID:22193408.

[9] Thorslund T, McIlwraith MJ, Compton SA, Lekomtsev S, Petronczki M, Griffith JD, *et al*. The breast cancer tumor suppressor BRCA2 promotes the specific targeting of RAD51 to single-stranded DNA. Nat

Struct Mol Biol 2010;17(10):1263–1265. doi:10.1038/nsmb.1905, PMID:20729858.
[10] Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 2010;79:181–211. doi:10.1146/annurev.biochem.052308.093131, PMID:20192759.
[11] Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, *et al*. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;266(5182):66–71. doi:10.1126/science.7545954, PMID:7545954.
[12] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, *et al*. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47(D1):D607–D613. doi:10.1093/nar/gky1131, PMID:30476243.
[13] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284–287. doi:10.1089/omi.2011.0118, PMID:22455463.
[14] Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 2013;29(14):1830–1831. doi:10.1093/bioinformatics/btt285, PMID:23740750.
[15] Győrffy B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. Comput Struct Biotechnol J 2021;19:4101–4109. doi:10.1016/j.csbj.2021.07.014, PMID:34527184.
[16] Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, *et al*. UALCAN: An update to the integrated cancer data analysis platform. Neoplasia 2022;25:18–27. doi:10.1016/j.neo.2022.01.001, PMID:35078134.
[17] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, *et al*. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017;19(8):649–658. doi:10.1016/j.neo.2017.05.002, PMID:28732212.
[18] Asplund A, Edqvist PH, Schwenk JM, Pontén F. Antibodies for profiling the human proteome-The Human Protein Atlas as a resource for cancer research. Proteomics 2012;12(13):2067–2077. doi:10.1002/pmic.201100504, PMID:22623277.
[19] Afaya A, Ramazanu S, Bolarinwa OA, Yakong VN, Afaya RA, Aboagye RG, *et al*. Health system barriers influencing timely breast cancer diagnosis and treatment among women in low and middle-income Asian countries: evidence from a mixed-methods systematic review. BMC Health Serv Res 2022;22(1):1601. doi:10.1186/s12913-022-08927-x, PMID:36587198.
[20] Lyng MB, Laenkholm AV, Pallisgaard N, Ditzel HJ. Identification of genes for normalization of real-time RT-PCR data in breast carcinomas. BMC Cancer 2008;8:20. doi:10.1186/1471-2407-8-20, PMID:18211679.
[21] Gur B, Bozkurt B, Konu O, Seckin S, Yulug I. Real-time PCR-based expression profiling of BRCA1-induced genes in primary breast tumors. Breast Cancer Res 2005;7:P4.19. doi:10.1186/bcr1149.
[22] Bièche I, Laurendeau I, Tozlu S, Olivi M, Vidaud D, Lidereau R, *et al*. Quantitation of MYC gene expression in sporadic breast tumors with a real-time reverse transcription-PCR assay. Cancer Res 1999;59(12):2759–2765. PMID:10383126.
[23] Kelm JM, Samarbakhsh A, Pillai A, VanderVere-Carozza PS, Aruri H, Pandey DS, *et al*. Recent advances in the development of Non-PIKKs targeting small molecule inhibitors of DNA double-strand break repair. Front Oncol 2022;12:850883. doi:10.3389/fonc.2022.850883, PMID:35463312.
[24] Vaghasia H, Sakaria S, Prajapati J, Saraf M, Rawal RM. Interactive bioinformatics analysis for the screening of hub genes and molecular docking of phytochemicals present in kitchen spices to inhibit CDK1 in cervical cancer. Comput Biol Med 2022;149:105994. doi:10.1016/j.compbiomed.2022.105994, PMID:36103746.
[25] Upadhyay V, Raval A, Shah K, Shah FD, Rawal R. A Prognostic and Predictive Study of BCR-ABL Expression Based on Characterization of Fusion Transcripts. Indian J Clin Biochem 2020;35(1):88–94. doi:10.1007/s12291-018-0779-1, PMID:32071500.
[26] Dashti S, Taheri M, Ghafouri-Fard S. An in-silico method leads to recognition of hub genes and crucial pathways in survival of patients with breast cancer. Sci Rep 2020;10(1):18770. doi:10.1038/s41598-020-76024-2, PMID:33128008.
[27] Kour A, Sambyal V, Guleria K, Singh NR, Uppal MS, Manjari M, *et al*. In silico pathway analysis based on chromosomal instability in breast cancer patients. BMC Med Genomics 2020;13(1):168. doi:10.1186/s12920-020-00811-z, PMID:33167967.
[28] Shakouri A, Kahroba H, Hamishekar H, Abdolalizadeh J. Nanoencapsulation of Hirudo medicinalis proteins in liposomes as a nanocarrier for inhibiting angiogenesis through targeting VEGFA in the Breast cancer cell line (MCF-7). Bioimpacts 2022;12(2):115–126. doi:10.34172/bi.2021.39, PMID:35411300.
[29] Farooq M, Bhat GM, Besina S, Thakur N, Zahoor S, Rather RA, *et al*. Expression of HIF-1α and markers of angiogenesis and metabolic adaptation in molecular subtypes of breast cancer. Transl Med Commun 2023;8:2. doi:10.1186/s41231-023-00135-x.
[30] Kneubil MC, Goulart KOB, Brollo J, Coelho GP, Mandelli J, Orlandin BC, *et al*. Predictive value of DNA repair gene expression for response to neoadjuvant chemotherapy in breast cancer. Braz J Med Biol Res 2022;55:e11857. doi:10.1590/1414-431X2021e11857, PMID:35293552.
[31] Wang C, Xie X, Li W, Jiang D. Expression of KIF2A, NDC80, CDK1, and CCNB1 in breast cancer patients: Their interaction and linkage with tumor features and prognosis. J Clin Lab Anal 2022;36(9):e24647. doi:10.1002/jcla.24647, PMID:35949045.
[32] Fang L, Liu Q, Cui H, Zheng Y, Wu C. Bioinformatics analysis highlight differentially expressed CCNB1 and PLK1 genes as potential anti-breast cancer drug targets and prognostic markers. Genes (Basel) 2022;13(4):654. doi:10.3390/genes13040654, PMID:35456460.
[33] Cheng Q, Ma Z, Shi Y, Parris AB, Kong L, Yang X. FGFR1 overexpression induces cancer cell stemness and enhanced Akt/Erk-ER signaling to promote palbociclib resistance in luminal a breast cancer cells. Cells 2021;10(11):3008. doi:10.3390/cells10113008, PMID:34831231.
[34] Verras GI, Tchabashvili L, Mulita F, Grypari IM, Sourouni S, Panagodimou E, *et al*. Micropapillary breast carcinoma: from molecular pathogenesis to prognosis. Breast Cancer (Dove Med Press) 2022;14:41–61. doi:10.2147/BCTT.S346301, PMID:35310681.
[35] Verras GI, Mulita F, Tchabashvili L, Grypari IM, Sourouni S, Panagodimou E, *et al*. A rare case of invasive micropapillary carcinoma of the breast. Prz Menopauzalny 2022;21(1):73–80. doi:10.5114/pm.2022.113834, PMID:35388282.
[36] Akrida I, Mulita F. The clinical significance of HER2 expression in DCIS. Med Oncol 2022;40(1):16. doi:10.1007/s12032-022-01876-9, PMID:36352293.
[37] Verras GI, Tchabashvili L, Chlorogiannis DD, Mulita F, Argentou MI. Updated clinical evidence on the role of adipokines and breast cancer: a review. Cancers (Basel) 2023;15(5):1572. doi:10.3390/cancers15051572, PMID:36900364.