

RESEARCH

Open Access



A comprehensive assessment of VCAN transcriptional expression and evaluation as an effective prognostic biomarker against breast cancer: in silico study

Md. Hasan Jafre Shovon¹, Dhrubo Ahmed Khan¹, Md. Mohaimenul Islam Tareq¹, Md. Imtiaz¹,
Md Nazmul Hasan Zilani² and Md. Nazmul Hasan^{1*}

Abstract

Background The VCAN gene provides instructions for making a protein called versican which is a type of protein known as a proteoglycan. Versican is a key ingredient of the extracellular matrix, and due to its widespread expression in the body, versican is involved in cell adhesion, proliferation, and migration. Mutations or alterations of this protein could result in the disintegration of the fine-tuned molecular machinery which can lead to uncontrolled cell proliferation.

Results VCAN is a novel prognostic marker for multiple cancers, and it showed tremendous results on breast cancer prognosis based on the data available on multiple websites. So, we targeted VCAN to analyze the expression and the outcome of breast cancer. This is a server-based study, and the expression of VCAN shows upregulation in breast cancer subtypes as compared to the normal tissue. The promoter methylation analysis suggested that overexpression of VCAN may be due to hypomethylation. Mutation analysis showed a positive correlation with VCAN expression where missense-type mutation has the highest percentage (77.33%), truncating (17.33%), and splice (4%) and somatic mutation frequency is 1.8%. VCAN was closely related to ten different genes and coexpressed with five of the genes among them. Five distinct compounds are linked to the methylation and mutagenesis of VCAN, according to the gene–drug interaction.

Conclusions The upregulation of VCAN is closely correlated with promoter methylation and the clinical features of breast cancer patients. The whole study suggests that the breast cancer patient's survival rate gets lower when the VCAN expression level gets higher. We anticipated that these findings will lead to further improvements in breast cancer prognosis and the significance of VCAN as a biomarker for breast cancer prognosis.

Keywords Breast cancer, Cancer biomarker, VCAN, Gene expression, Bioinformatics

Background

Cancer occurs due to uncontrolled cell proliferation, and it can develop anywhere in the human body (Farrior n.d.). Cancer is a big concern for all human beings as it is one of the major reasons for human death (Sung *et al.* 2021). Breast cancer is one of the most common types of cancer. More than a million cases of breast cancer are diagnosed each year (Lin *et al.* 2022). VCAN can

*Correspondence:

Md. Nazmul Hasan
mn.hasan@just.edu.bd

¹ Laboratory of Pharmaceutical Biotechnology and Bioinformatics, Department of Genetic Engineering and Biotechnology, Jashore University of Science and Technology, Jashore 7408, Bangladesh

² Department of Pharmacy, Faculty of Biological Science and Technology, Jashore University of Science and Technology, Jashore 7408, Bangladesh



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

be a prolific biomarker for BC prognosis. The VCAN gene provides instructions for making a protein called versican. Versican is a type of protein known as a proteoglycan, which means it has several sugar molecules attached to it. Versican is a chondroitin sulfate proteoglycan found in the extracellular matrix that is important for changes in cell phenotype associated with development and disease (Jørgensen et al. 2022). Versican is found in the extracellular matrix of many different tissues and organs. The extracellular matrix is the intricate lattice of proteins and other molecules that forms in the spaces between cells (Antonius et al. 2022; Fadholly et al. 2020a, b, c; Yamaguchi 2000). Versican interacts with many proteins and molecules to facilitate the assembly of the extracellular matrix and ensure its stability (Wu et al. 2005). Versican is involved in cell adhesion, proliferation, and migration. Alteration in the versican function can result in serious trouble in cell adhesion, proliferation, and migration (Dituri et al. 2022). Overexpression of VCAN causes uncontrolled cell proliferation. VCAN mRNA expression analysis indicated that it can be a biomarker for breast cancer, brain cancer (Ricciardelli et al. 2009), bladder cancer, esophagus cancer (Thelin et al. 2012), lung cancer (Brunn et al. 2021), pancreas cancer (Skandalis et al. 2006) and stomach cancer (Setoguchi et al. 2011). Human BC is a heterogeneous disease with different neoplasms originating from the epithelial cells (Azam and Sounni 2022; Zhang et al. 2023). The molecular underpinning of the BC cells led to the identification of prognostic tools to estimate therapeutic reactions and predict gene expression signatures which can lead to the long-term survival of the patients (Munshi et al. 2022; Al Saber et al. 2022; Samad et al. 2020). Early prognosis of human BC can save many lives and a potent biomarker can uncover BC in its early stages which can be cured (Agostinetto et al. 2022; Biswas et al. 2021, 2023; Han et al. 2016; Kaddoura et al. 2023). VCAN can be a potential biomarker for human BC because it is correlated with BC-affected patients. BC-affected cases are estimated in 2022 which is 15% of overall cancers in 2022, and 43,250 deaths are recorded in 2022 which is 7.1% of total cancer deaths in 2022 by breast cancer (<https://seer.cancer.gov/>). The death rate of BC is getting higher day by day, and BC is one of the cancers which hardly produce any symptoms in the pre-stages (Azamjah et al. 2019; Biswas et al. 2022; Kumaratharan 2015). Numerous transcriptome studies have been conducted on VCAN, and the results indicate that VCAN expression is upregulated in a number of common solid malignancies, including ovarian, pancreatic, breast, lung, esophageal, and colorectal cancers (Hirani et al. 2021). However, no research have been conducted on VCAN expression that are related to VCAN overexpression (Fadholly et al. 2020a, b, c). So, the

novel aspect of this study is its comprehensive examination of VCAN transcriptional expression in breast cancer and its evaluation as a prognostic biomarker using in silico techniques, which may be able to reveal important information about the molecular mechanisms underlying the development of breast cancer and identify a new biomarker for forecasting patient outcomes. This study has the potential to significantly impact breast cancer research and clinical practice, guiding tailored treatment plans and enhancing patient outcomes.

Methods

Differential expression analysis

UALCAN (<http://ualcan.path.uab.edu/>) (Chandrashekar et al. 2017, 2022), TIMER2.0 (<http://timer.cistrome.org/>) (version 2.0) (Li et al. 2016, 2017, 2020; Ricciardelli et al. 2009), GEPIA2 (<http://gepia2.cancer-pku.cn/#index>) (Tang et al. 2019), and GENT2 (Park et al. 2019) (<http://gent2.appex.kr/gent2/>) databases were utilized to analyze mRNA expression of VCAN in a number of cancers and normal tissues. The TCGA dataset on the UALCAN server includes 738 normal tissues and 8609 tumors. There are 10,656 tumors and 790 datasets of normal tissue on the TIMER2.0 server. In GEPIA2, samples from the TCGA and GTEx projects make up 8587 normal samples and 9736 tumor samples. GENT2 delivers gene expression across 72 different tissues and more than 68,000 samples.

Expression analysis of targeted biomarker in normal tissues and human BC-affected tissues

The GEPIA2 (Tang et al. 2019) and UALCAN (Chandrashekar et al. 2017, 2022) servers were used to assess the differences in mRNA expression between healthy and cancer-affected tissues. The GEPIA2 server features 291 samples of normal tissues and 1085 samples of human BC tumor tissues matching TCGA and GTEx datasets where Log_2FC and q -value were considered 1 and 0.01 using the ANOVA differential method. There are 114 TCGA samples of normal tissues, and 1097 TCGA samples of human BC tumor tissues are used for the investigation on the UALCAN website. Both the servers were utilized to compare the expression between healthy and cancerous human BC tissues. Both servers are renowned for examining the differences in expression between healthy and cancer tissues.

Immunohistochemistry staining

Immunohistochemistry (IHC) is a technique used to detect and localize specific proteins in cells and tissues. The immunohistochemical staining of VCAN at the protein level was examined using the Human Protein Atlas server (HPA) (<https://www.proteinatlas.org/>) (Uhlén

et al. 2015). The HPA server was used to download the IHC image of the VCAN protein expression using GTEx dataset in normal tissues and human BC tissues in order to compare the differences at the protein level.

A relationship of VCAN expression and clinical features of BC patients

Based on the clinicopathological characteristics of the BC patients, the UALCAN server (Chandrashekar et al. 2017, 2022) was utilized to assess the mRNA expression of VCAN in patients' major subclass, age, gender, race, menopause status, and individual cancer stage using the TCGA dataset. Comparing normal, healthy tissues with tumor-affected tissues is allowed for study.

Promoter methylation of VCAN expression

Promoter methylation is a frequent epigenetic alteration that has been linked to several diseases, including cancer. By silencing genes essential for cell proliferation, differentiation, and repair, it is believed that it contributes to the growth of tumors. The connection between promoter methylation and overexpression of VCAN was examined using Mexpress (<https://mexpress.be/>) (Koch et al. 2015, 2019) and the UALCAN website. A *p* value of 0.05 or lower was regarded as significant.

Correlation of VCAN expression with mutation and copy number alteration

The somatic copy number changes as a result of unregulated cell division brought on by mutations. cBioPortal server (<https://www.cbioportal.org/>) (version 5.3.3) (Cerami et al. 2012; Gao et al. 2013) was used for the assessment of the mutation and copy number alteration of the VCAN gene in BC patients. To conduct the analysis, 3835 samples from four different TCGA datasets were employed, including TCGA, Cell 2015, TCGA, Firehose Legacy, TCGA, Nature 2012, and TCGA, PanCancer Atlas.

PPI network of VCAN-enriched genes

A protein–protein interaction (PPI) network is a graph that depicts the physical interactions that take place between proteins in a cell. Protein interactions are represented by the edges and nodes of the graph, respectively.

PPI networks can be utilized to investigate the composition and operation of cells and to find new therapeutic targets. Additionally, they can be used to create new treatments and comprehend the disease's underlying causes. Using the STRING server (<https://string-db.org/>) (version 11.5) (von Mering et al. 2003) and Cytoscape software (version 3.9.1) (Paul Shannon et al. 1971), a PPI network comprising VCAN-enriched

genes was generated. Using the Cytoscape software, the STRING server was utilized to obtain the VCAN-enriched genes dataset.

Gene–drug interaction network of the VCAN

The study of gene–drug interaction networks has the potential to completely change how medications are created and used. Understanding the complicated connections between genes and medications can help create safer and more effective treatments for a variety of disorders. Utilizing the comparative Toxicogenomics Database (CTD) (<http://ctdbase.org/>) (Mattingly et al. 2003) and Cytoscape software (version 3.9.1) (Paul Shannon et al. 1971), the VCAN gene–drug interaction network was created. Datasets of several chemotherapeutic medicines that can affect VCAN were downloaded from the CTD database (last updated on March 31, 2023) which consists of 2,690,443 number of chemical–gene interaction data (curated), and the interaction network was visualized using the Cytoscape software.

Results

Differential expression analysis of VCAN on several types cancer

Four different web server-based platforms were used to examine the expression analysis of VCAN in various cancer types. According to research on the UALCAN server, VCAN is elevated in a variety of cancers, including those of the bladder, breast, bile duct, colon, esophagus, head and neck, liver, lung, and stomach (Fig. 1A). It is also upregulated in sarcomas of the skin, thyroid, and thymomas. Surprisingly, VCAN is downregulated in a variety of malignancies, including uterine, pancreatic, pancreaticobiliary, pheochromocytoma, and paraganglioma (Fig. 1A) and cervical, KICH, pancreatic, and prostate cancers. Three more servers, TIMER2.0, GEPIA2, and GENT2, were used to investigate VCAN expression in various cancer types; the results are shown in Fig. 1B–D, respectively. To conduct the analysis, the TCGA database was used by all four servers.

VCAN expression analysis on normal and breast cancer affected tissues

Using the TCGA dataset, UALCAN and GEPIA2 were used to perform VCAN expression analyses on normal and cancer-affected tissues. A total of 291 healthy and 1085 cancer-affected samples were used in the analysis by the GEPIA2 server (Fig. 2A). The investigation of the various BC subtypes expression of the VCAN was also carried out on the GEPIA2 server. The study discovered four distinct BC subtypes of expression. The expression of VCAN in normal and tumor-affected samples was carried out using the UALCAN server, which is shown in

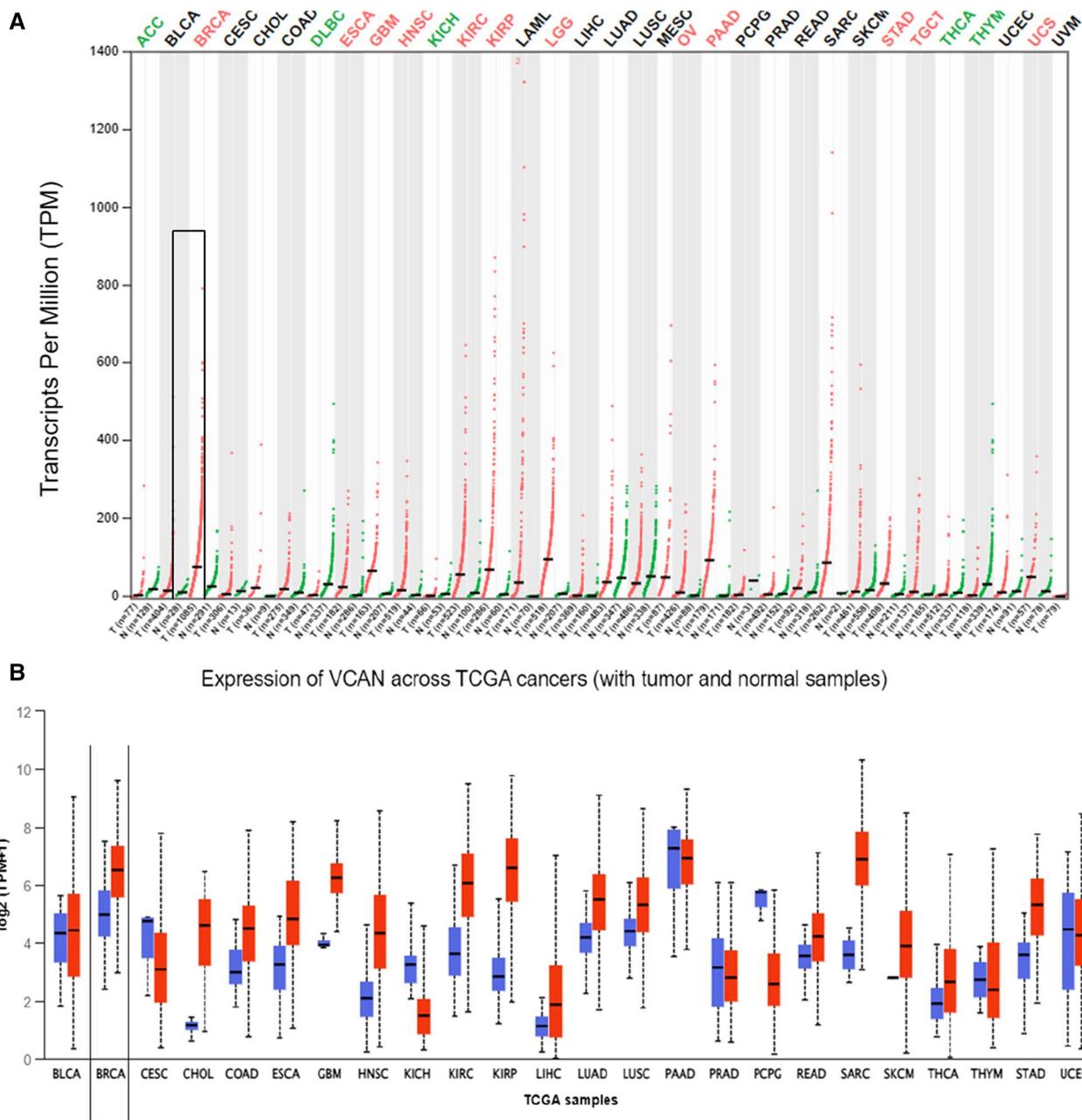


Fig. 1 VCAN gene expression pattern in several human cancers; **A** comparison of cancer and normal tissues; on the GEPIA2 server, high and low mRNA expression are denoted, respectively, by the colors red and green; **B** the box plot on the UALCAN server displays the gene expression profile of the VCAN gene in 24 different types of human malignancies, including tumors and matching samples of normal tissue. The average expression value in all tumor and normal tissues, respectively, is shown in red and blue boxes. **C** The comparison of cancerous and normal tissues, where high and low levels of mRNA expression are denoted on the TIMER2.0, respectively, by the colors red and blue; **D** a box-plot depicting the expression of VCAN mRNA in tumors and their corresponding normal tissues based on the GENT2 database, where boxes stand in for the median, dots for outliers, red boxes represent tumor tissues, and blue boxes represent normal tissues

Fig. 2B and contains 114 normal and 1097 tumor TCGA samples. In Fig. 2C, they are depicted as basal-like (291 samples), HER2 (66 samples), Luminal A (415 samples), and Luminal B (194 samples).

Immunohistochemistry (IHC) staining

VCAN protein-level immunohistochemistry (IHC) was carried out utilizing the HPA (Human Protein Atlas) server, and the dataset’s VCAN expression data

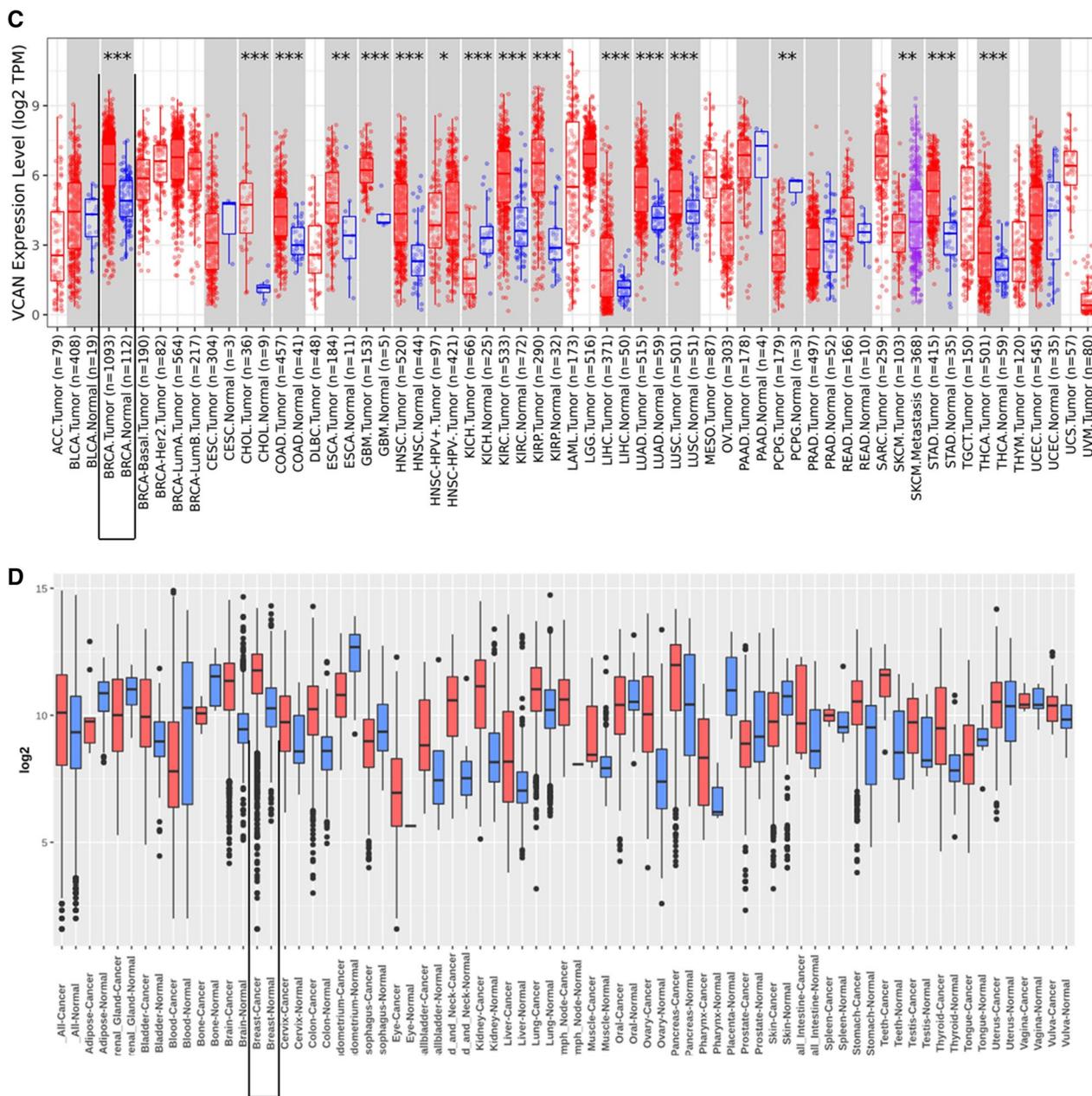


Fig. 1 continued

were compared with the data acquired from the TCGA database. The findings of the analysis revealed that the VCAN expression had a strong IHC staining on the tissues damaged by BC, compared to moderate staining on normal tissues. This is shown in Fig. 3.

A relationship of VCAN expression and clinical features of BC patients

UALCAN server was used to investigate the clinical characteristics of BC-afflicted patients and healthy

individuals in connection to the expression of VCAN. The findings pointed to increased expression of VCAN in patients' major subclass, age, gender, race, menopause status, and individual cancer stage (Fig. 4A–F). In this study, 1065 TCGA samples with information on different stages of cancer, 967 samples of menopause status, 1087 samples of different ages, 1089 samples of gender, 988 samples of different races, and 833 samples of patient's major subclass with

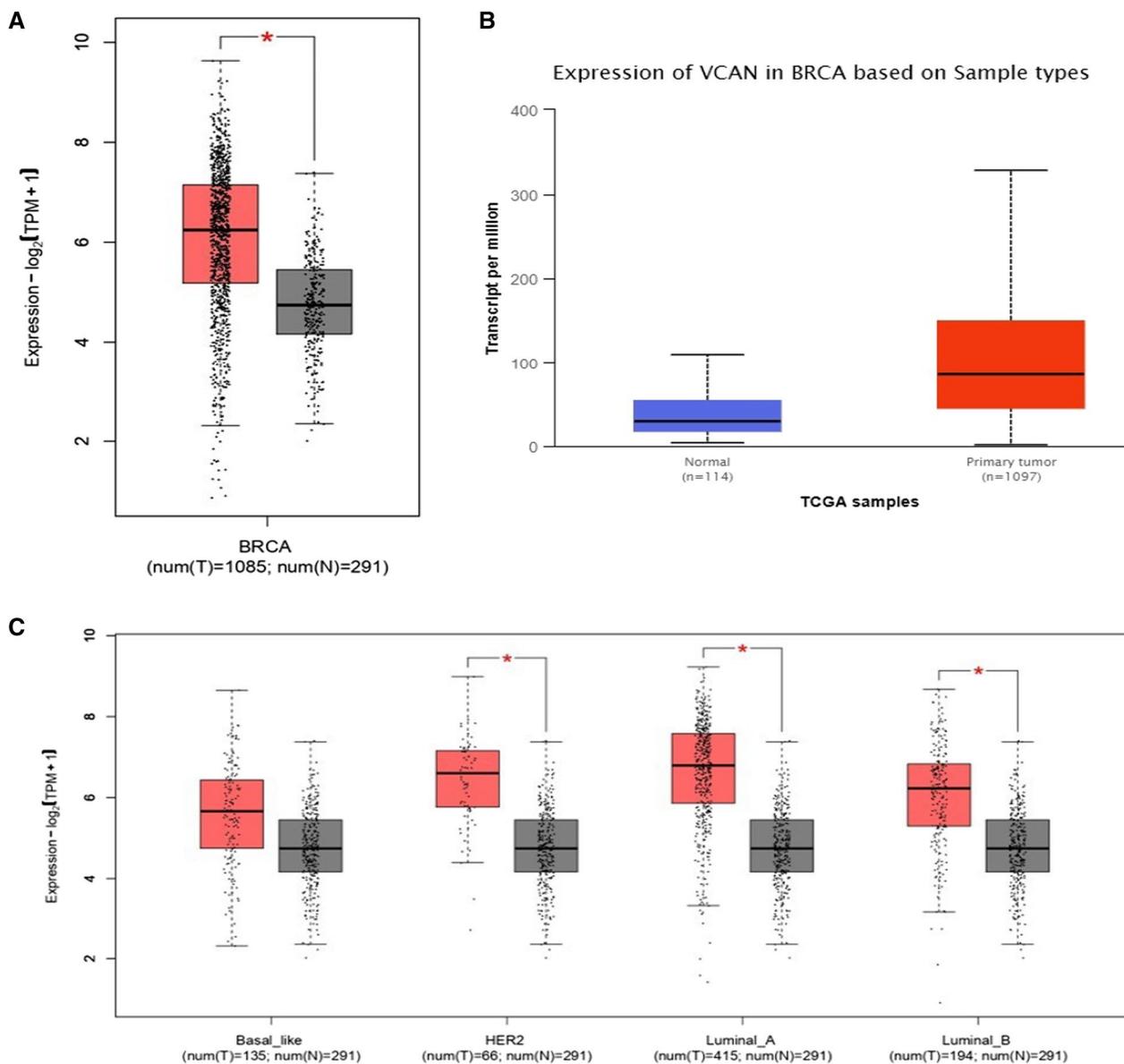


Fig. 2 VCAN expression analysis on normal and breast cancer-affected tissues in which **A, B** box-plots displaying the VCAN mRNA expression in breast cancer and normal tissues using the GEPIA2 and UALCAN servers; **C** tissue-wise expression of a gene signature in different breast cancer sub-types using box plot where gray color denotes normal and red color denotes tumor-affected tissues on the GEPIA2 server

information on histologic subtypes were used. We used 114 TCGA samples of healthy individuals that were utilized as controls. Despite the fact that there have only been a few cases of BC in male patients where the VCAN expression was substantially higher in men compared to women. In the initial stage of cancer, VCAN expression is higher; subsequently in the fourth stage, it seems a little bit lower. Patients aged between 21 and 40 had higher levels of VCAN expression than the other age groups.

Correlation of promoter methylation with VCAN expression

A key factor in regulating gene expression is promoter methylation. It is an important epigenetic process, and any changes or irregularities in the promoter methylation might result in genetic flaws like cancer. MEXPRESS and the UALCAN server were used to evaluate the relationship between promoter methylation and the overexpression of VCAN. The MEXPRESS server’s findings reveal a negative association between promoter methylation and VCAN upregulation (a *p* value less

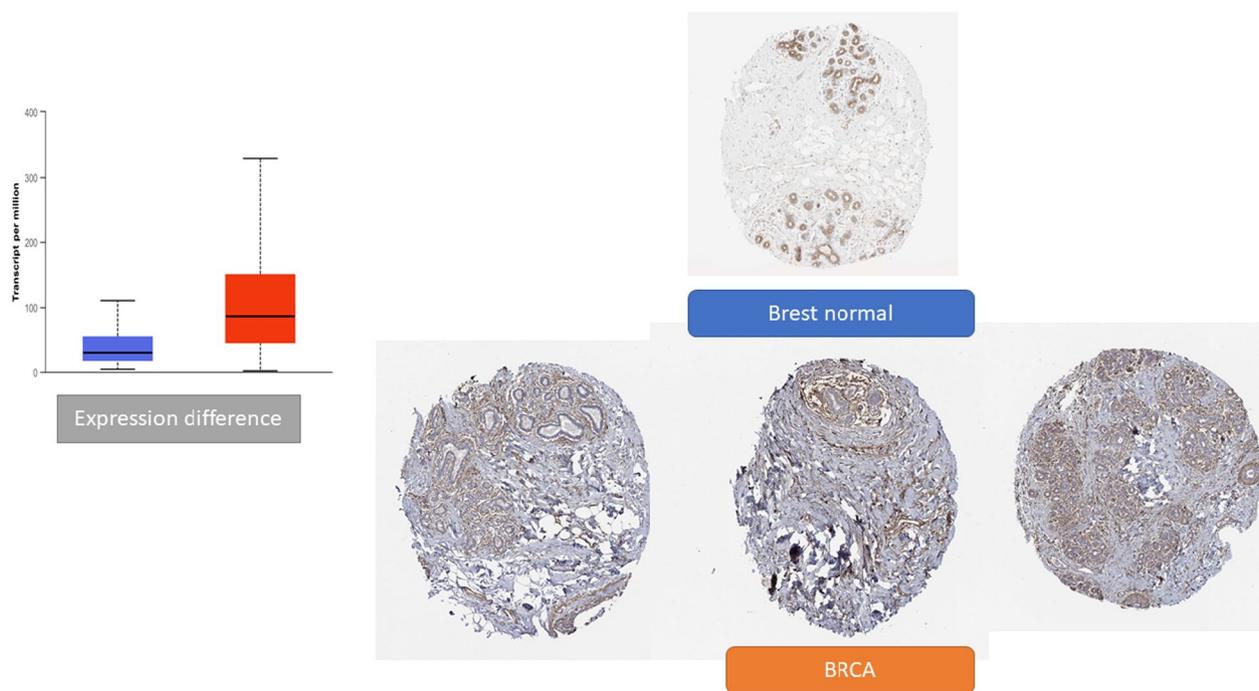


Fig. 3 Comparison of VCAN gene expression and immunohistochemistry images in normal and tumor tissues. Breast invasive carcinoma (BRCA) tissues had significantly greater VCAN protein expression levels than normal tissues

than 0.05 is considered as significant) (Fig. 5A). The MEXPRESS server hypothesized that the overexpression of VCAN might be caused by hypomethylation. The promoter methylation in the TCGA samples was examined by the UALCAN server using the beta value (Fig. 5B). The Beta number represents the level of DNA methylation, which ranges from 0 (unmethylated) to 1 (Fully methylated). Different beta value cut-offs have been considered to signify hyper-methylation [Beta value: 0.7–0.5] or hypo-methylation [Beta-value: 0.3–0.25]. The UALCAN server employed 97 normal and 793 tumor-affected TCGA samples in its study.

Copy number alteration and mutation

The cBioPortal server was used to assess mutations and copy number alterations. Four TCGA datasets of BC investigations yielded a total of 3835 TCGA samples (Fig. 6A). The findings revealed that VCAN was altered in a total of 75 samples, including 48 duplication mutations in patients with multiple samples; among the samples, 58 were missense mutations, 13 were nonsense mutations, and four splice mutations occurred with a 1.8% somatic mutation rate frequency (Fig. 6B, C and Table 1). With a 3.06% alteration frequency in 817 samples, the breast invasive carcinoma (TCGA, Cell 2015) studies had the highest alteration frequency. This included 2.2% mutation (18 cases), 0.73% deep deletion (six cases), and numerous

alterations (0.12%) (one case). A correlation between the overexpression of VCAN in BC and the mutation and copy number alteration may exist, according to studies from the cBioPortal database.

PPI network construction and visualization

The VCAN-enriched genes were found using the STRING database and the Cytoscape program to analyze the protein–protein interaction. According to functional interaction network research, VCAN was closely related to ten different genes shown in Fig. 7 and Table 2. With regard to coexpression rates, VCAN with CD44, FBLN1, FBN1, FN1, and GPC1 are, in that order, 0.062, 0.081, 0.102, 0.097, and 0.062. In contrast, there is no coexpression exhibited with SDC1, SDC4, SELL, TLR2, or TLR6 during protein–protein interaction.

Gene–Drug interaction network analysis of the VCAN

The gene–drug interaction network was built using the CTD database and visualized using Cytoscape software to determine the association between VCAN expression and the numerous medicinal drugs that are currently on the market. It was found that certain medications may be able to regulate the expression of VCAN as shown in Fig. 8 and Table 3. Aflatoxin B1 and bisphenol A increase the methylation, and sodium arsenite affects the methylation.

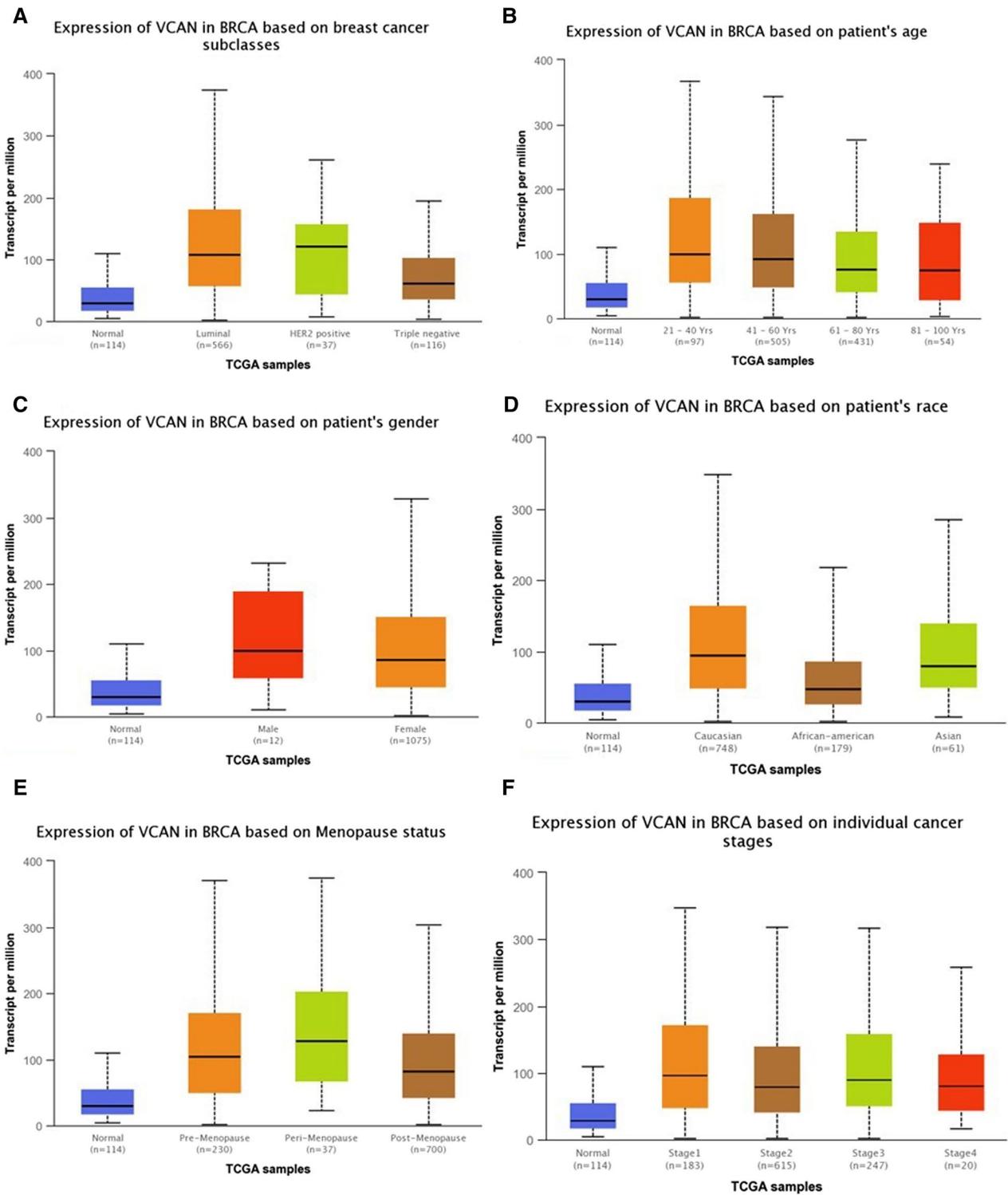


Fig. 4 VCAN expression analysis in relation to clinical characteristics of BC patients. The VCAN mRNA expression in BRCA shows **A** breast cancer major subclasses, **B** patient's age, **C** patient's gender, **D** patient's race, **E** menopause status, and **F** individual cancer stages. In order to create these graphs, significant differences between normal variables and other variables were compared

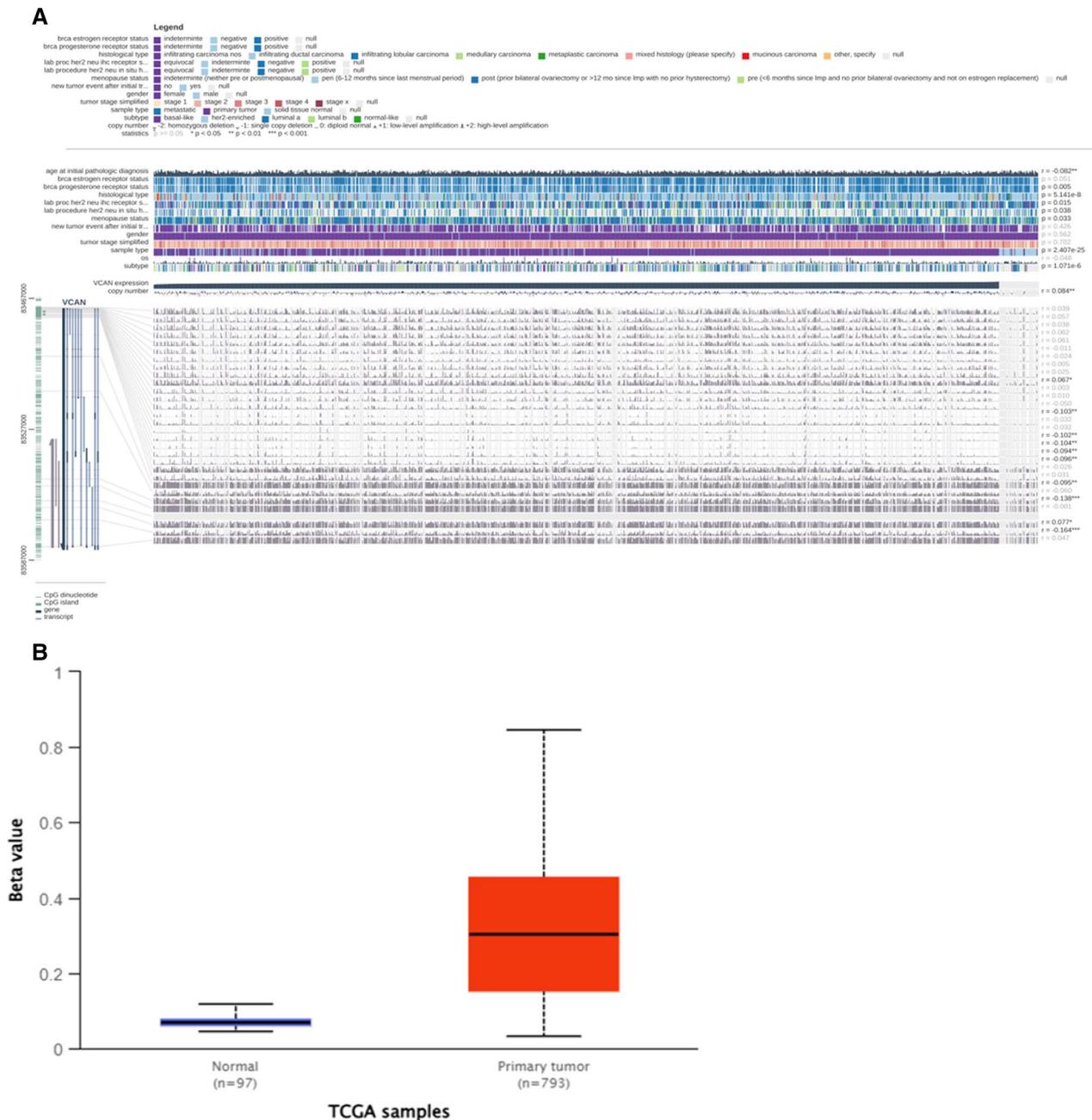


Fig. 5 Promoter methylation and VCAN expression correlation were utilized using the MEXPRESS AND UALCAN server. **A** Analysis of the relationship between the expression of VCAN and the methylation of its promoter using MEXPRESS server. A negative sign shows a negative association between the expression of VCAN and its promoter methylation when a certain probe is used at a particular CpG island, while a positive sign suggests a positive correlation between the two ($p < 0.05$). **B** UALCAN server was used to identify the beta value of normal and primary tumors. The Beta number represents the level of DNA methylation, which ranges from 0 (unmethylated) to 1 (fully methylated)

Discussion

The study provided evidence of a bioinformatics analysis, which was carried out to look at the VCAN expression pattern, assess its link to the prognosis of breast cancer, and forecast the relationship between VCAN and other

proteins as well as the interaction of VCAN with drugs. VCAN is a member of the versican/aggrecan family of proteoglycans. Versican is a proteoglycan protein, which means that several sugar molecules are attached to it (Dibha et al. 2022). Versican is found in the extracellular

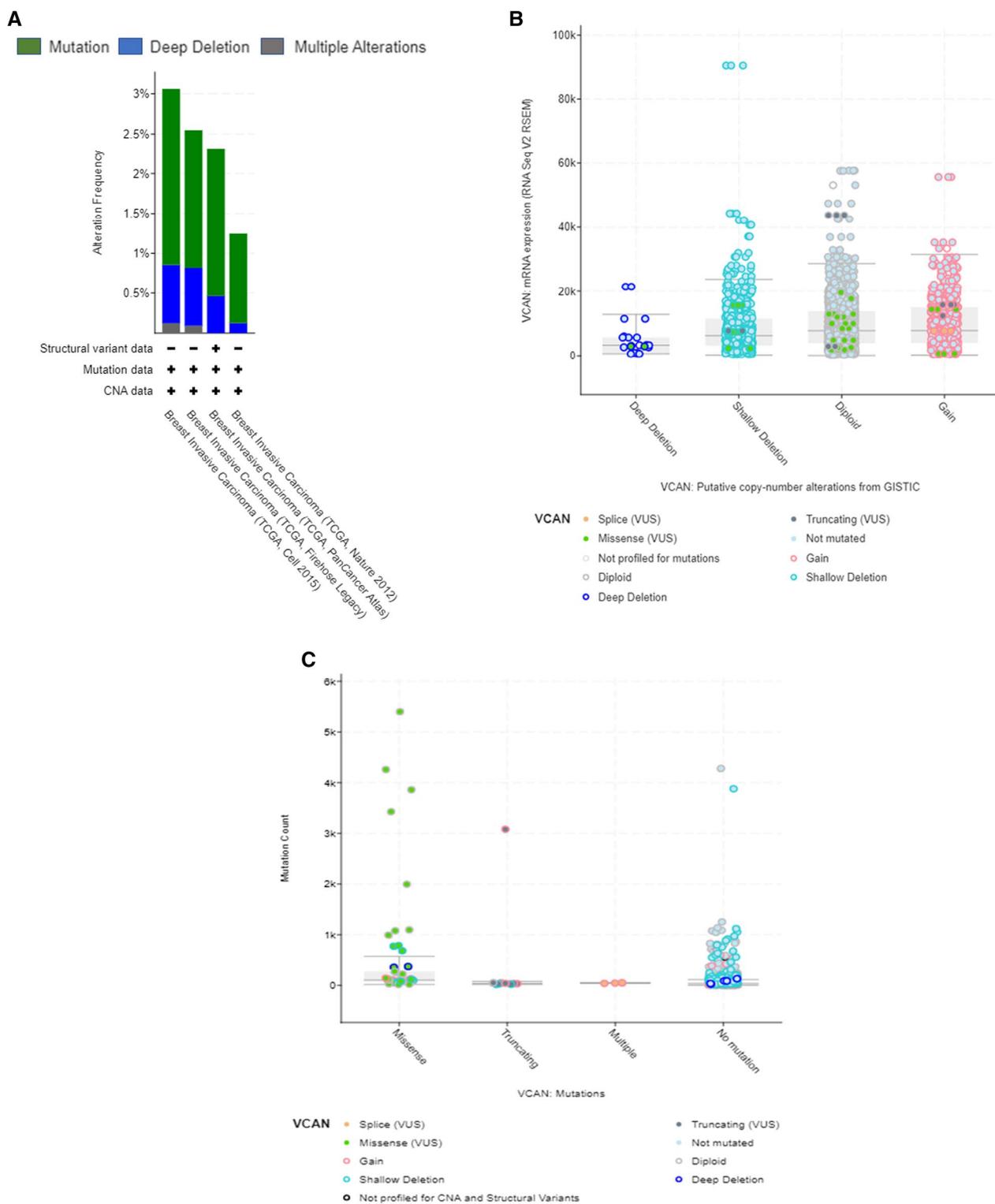


Fig. 6 VCAN mutations and genetic changes in BC tissue. **A** The mutation rates and genomic changes in the VCAN gene are depicted in a bar diagram. **B** VCAN putative copy number alteration from GISTIC and **C** mutation types and no mutation occurrence in different mutation count using TCGA dataset

Table 1 TCGA dataset's list of BC's VCAN mutational sites and types

Sample ID	Cancer type detailed	Protein change	Mutation type	No. of samples
TCGA-BH-A0EE-01	Breast invasive ductal carcinoma	M2799T	Missense_Mutation	123
TCGA-C8-A12P-01	Breast invasive ductal carcinoma	S1547C	Missense_Mutation	141
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	X348_splice	Splice_Site	52
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	D1008G	Missense_Mutation	52
TCGA-AR-A1AV-01	Breast invasive ductal carcinoma	E1272*	Nonsense_Mutation	41
TCGA-D8-A27V-01	Breast invasive lobular carcinoma	Q2845Sfs*8	Frame_Shift_Del	3080
TCGA-A2-A0YG-01	Invasive breast carcinoma	X348_splice	Splice_Site	50
TCGA-A2-A0YG-01	Invasive breast carcinoma	D1008G	Missense_Mutation	50
TCGA-AR-A1AV-01	Invasive breast carcinoma	E1272*	Nonsense_Mutation	38
TCGA-BH-A0EE-01	Invasive breast carcinoma	M2799T	Missense_Mutation	79
TCGA-C8-A12P-01	Breast invasive ductal carcinoma	S1547C	Missense_Mutation	123
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	X348_splice	Splice_Site	45
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	D1008G	Missense_Mutation	45
TCGA-AR-A1AV-01	Breast invasive ductal carcinoma	E1272*	Nonsense_Mutation	40
TCGA-BH-A0EE-01	Breast invasive ductal carcinoma	M2799T	Missense_Mutation	59
TCGA-C8-A12P-01	Breast invasive ductal carcinoma	S1547C	Missense_Mutation	109
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	X348_splice	Splice_Site	38
TCGA-A2-A0YG-01	Breast invasive ductal carcinoma	D1008G	Missense_Mutation	38
TCGA-AR-A1AV-01	Breast invasive ductal carcinoma	E1272*	Nonsense_Mutation	33
TCGA-GM-A2DB-01	Breast invasive ductal carcinoma	Q2031H	Missense_Mutation	125
TCGA-A8-A09Z-01	Breast invasive lobular carcinoma	T1822I	Missense_Mutation	1095
TCGA-A7-A0CE-01	Breast invasive ductal carcinoma	N3147T	Missense_Mutation	94
TCGA-A7-A3RF-01	Breast invasive mixed mucinous carcinoma	P365H	Missense_Mutation	32
TCGA-AC-A23H-01	Breast invasive ductal carcinoma	S2826L	Missense_Mutation	4259
TCGA-AR-A0TX-01	Breast invasive ductal carcinoma	P1204A	Missense_Mutation	276
TCGA-E2-A2P6-01	Breast invasive lobular carcinoma	S3260F	Missense_Mutation	104
TCGA-A8-A07L-01	Breast invasive ductal carcinoma	M1368L	Missense_Mutation	81
TCGA-E9-A1NF-01	Breast invasive carcinoma (NOS)	E1739*	Nonsense_Mutation	51
TCGA-AN-A046-01	Breast invasive ductal carcinoma	K2228T	Missense_Mutation	5400
TCGA-D8-A27T-01	Breast invasive lobular carcinoma	V949L	Missense_Mutation	21
TCGA-S3-AA14-01	Breast invasive ductal carcinoma	E2310Q	Missense_Mutation	16
TCGA-BH-A0EE-01	Invasive breast carcinoma	M2799T	Missense_Mutation	67
TCGA-C8-A12P-01	Invasive breast carcinoma	S1547C	Missense_Mutation	120
TCGA-A7-A0CE-01	Invasive breast carcinoma	N3147T	Missense_Mutation	77
TCGA-AR-A0TX-01	Invasive breast carcinoma	P1204A	Missense_Mutation	223
TCGA-A8-A07L-01	Invasive breast carcinoma	M1368L	Missense_Mutation	77
TCGA-GM-A2DB-01	Breast invasive ductal carcinoma	Q2031H	Missense_Mutation	103
TCGA-A8-A09Z-01	Breast invasive lobular carcinoma	T1822I	Missense_Mutation	1078
TCGA-A7-A0CE-01	Breast invasive ductal carcinoma	N3147T	Missense_Mutation	96
TCGA-A7-A3RF-01	Invasive breast carcinoma	P365H	Missense_Mutation	46
TCGA-AC-A23H-01	Breast invasive ductal carcinoma	S2826L	Missense_Mutation	3860
TCGA-E2-A105-01	Breast invasive ductal carcinoma	P2900Afs*5	Frame_Shift_Ins	53
TCGA-AR-A0TX-01	Invasive breast carcinoma	P1204A	Missense_Mutation	258
TCGA-E2-A2P6-01	Breast invasive lobular carcinoma	S3260F	Missense_Mutation	106
TCGA-A8-A07L-01	Breast mixed ductal and lobular carcinoma	M1368L	Missense_Mutation	89
TCGA-E9-A1NF-01	Breast invasive ductal carcinoma	E1739*	Nonsense_Mutation	47
TCGA-GM-A2DB-01	Breast invasive ductal carcinoma	Q2031H	Missense_Mutation	81
TCGA-A8-A09Z-01	Breast invasive lobular carcinoma	T1822I	Missense_Mutation	990
TCGA-A7-A0CE-01	Breast invasive ductal carcinoma	N3147T	Missense_Mutation	90

Table 1 (continued)

Sample ID	Cancer type detailed	Protein change	Mutation type	No. of samples
TCGA-A7-A3RF-01	Breast invasive mixed mucinous carcinoma	P365H	Missense_Mutation	44
TCGA-AC-A23H-01	Breast invasive ductal carcinoma	S2826L	Missense_Mutation	3427
TCGA-E2-A105-01	Breast invasive ductal carcinoma	P2900Afs*5	Frame_Shift_Ins	45
TCGA-A8-A0A6-01	Breast invasive lobular carcinoma	H3333P	Missense_Mutation	1995
TCGA-AR-A0TX-01	Breast invasive ductal carcinoma	P1204A	Missense_Mutation	216
TCGA-E2-A2P6-01	Breast invasive lobular carcinoma	S3260F	Missense_Mutation	93
TCGA-A8-A07L-01	Breast invasive ductal carcinoma	M1368L	Missense_Mutation	76
TCGA-E9-A1NF-01	Breast mixed ductal and lobular carcinoma	E1739*	Nonsense_Mutation	36
TCGA-AN-A0AT-01	Breast invasive ductal carcinoma	T570S	Missense_Mutation	117
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	S1945F	Missense_Mutation	785
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	E2413K	Missense_Mutation	785
TCGA-D8-A1XJ-01	Breast invasive carcinoma (NOS)	G1144C	Missense_Mutation	80
TCGA-AN-A0AT-01	Invasive breast carcinoma	T570S	Missense_Mutation	100
TCGA-B6-A0IE-01	Invasive breast carcinoma	V1296Wfs*42	Frame_Shift_Del	26
TCGA-A8-A09C-01	Breast mixed ductal and lobular carcinoma	T2115Hfs*44	Frame_Shift_Del	20
TCGA-AN-A0AT-01	Breast invasive ductal carcinoma	T570S	Missense_Mutation	98
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	S1945F	Missense_Mutation	770
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	E2413K	Missense_Mutation	770
TCGA-D8-A1XJ-01	Invasive breast carcinoma	G1144C	Missense_Mutation	70
TCGA-A8-A09C-01	Breast invasive ductal carcinoma	T2115Hfs*44	Frame_Shift_Del	17
TCGA-AN-A0AT-01	Breast invasive ductal carcinoma	T570S	Missense_Mutation	87
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	S1945F	Missense_Mutation	682
TCGA-D8-A1JA-01	Breast invasive ductal carcinoma	E2413K	Missense_Mutation	682
TCGA-D8-A1XJ-01	Breast invasive ductal carcinoma	G1144C	Missense_Mutation	56
TCGA-A8-A07R-01	Breast invasive ductal carcinoma	D1177H	Missense_Mutation	372
TCGA-A8-A07R-01	Breast invasive ductal carcinoma	D1177H	Missense_Mutation	357

matrix of numerous organs and tissues (Yamaguchi 2000). The extracellular matrix is the complex web of substances, including proteins, that form in the gaps between cells (Fadholly et al. 2020a, b, c; Kharisma et al. 2020; Widyananda et al. 2021). Versican accelerates extracellular matrix synthesis and ensures its stability by interacting with a wide variety of proteins and substances. Cell growth is correlated with VCAN expression as well as it has been labeled as a cancer-causing gene after some studies suggested that it might be carcinogenic (Yu et al. 2003). Breast, ovarian, and lung cancers are only a few of the tumors for which VCAN expression levels have been connected (Du et al. 2013). VCAN is another factor that could be a predictor of breast cancer. We still do not know what part VCAN might play in tumor immunology and tumor growth. We studied the expression levels of VCAN in breast cancer using UALCAN, GENT2, TIMER, and GEPIA databases. VCAN expression was shown to be highly expressed in breast tumors, which is consistent. We evaluated the expression of VCAN in samples from normal and various subtypes of breast cancer and discovered a significant connection suggesting that

VCAN expression may be a viable diagnostic biomarker for breast cancer. Using TCGA samples, we also examined the clinical characteristics by the expression of VCAN in BRCA depending on distinct cancer stages, menopause status, and patient's age, gender and race. Although only a few cases of males have been diagnosed, the results unexpectedly demonstrated that the VCAN expression is significantly higher in males compared to females. Another study has found that while normal tissues do not contain any IHC stain, the expression of VCAN exhibits rough IHC staining in tumor-affected tissues, suggesting that VCAN may be a significant biomarker for breast cancer. Gains or losses in the genomic area can have either suppressive or cancer-promoting effects (Qiu et al. 2008). Investigations on VCAN mutations, copy number changes, and mutant mRNA expressions were conducted using the cBioPortal webserver (Gao et al. 2013). Seventy-five altered individuals made up the sample under analysis, with a somatic mutation frequency of 1.8%; subsequently, we examined the close connection between DNA methylation and VCAN expression. When different DNA methylation clusters are compared, it becomes clear that VCAN expression

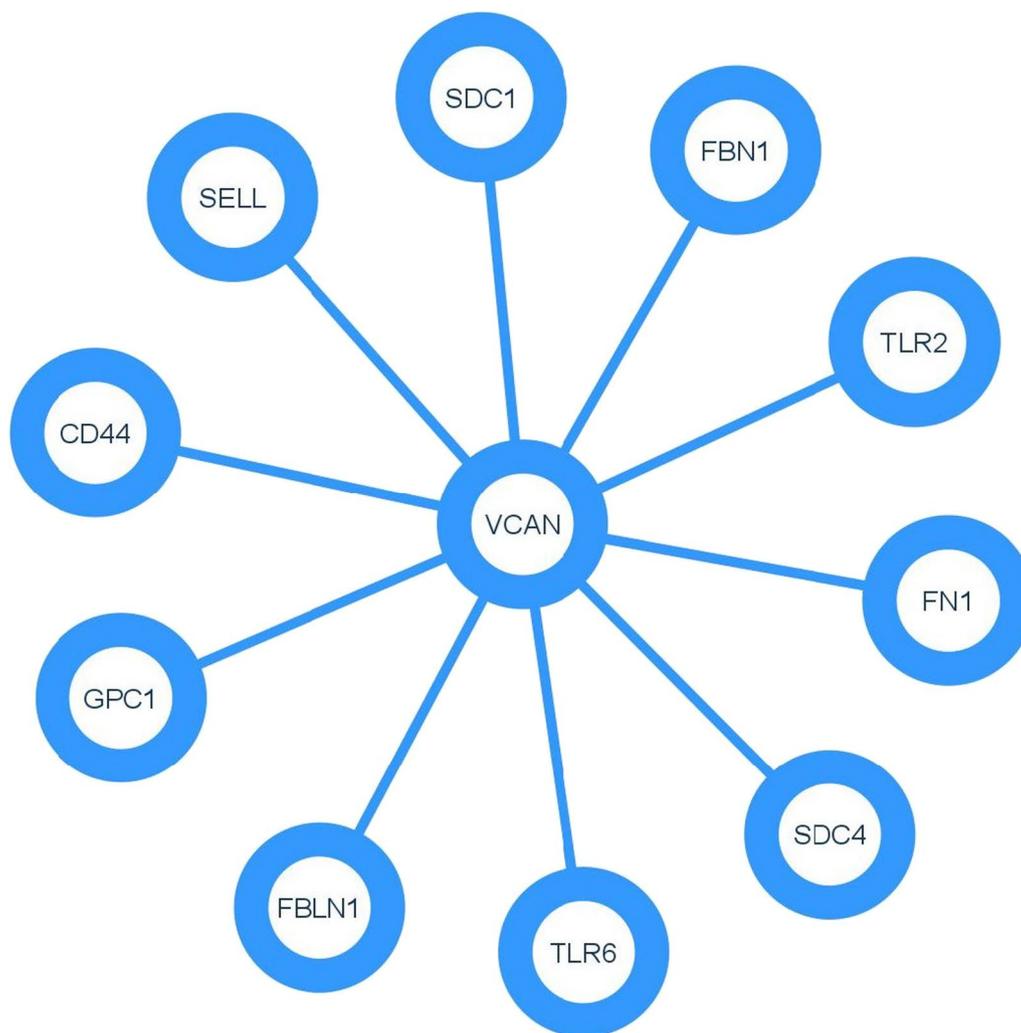


Fig. 7 PPI network of genes with VCAN enrichment

Table 2 Coexpression of VCAN-enriched genes

Node1	Node2	Coexpression	Experimentally determined interaction	Combined score
CD44	VCAN	0.062	0.409	0.993
FBLN1	VCAN	0.081	0.213	0.969
FBN1	VCAN	0.102	0.27	0.959
FN1	VCAN	0.097	0	0.972
GPC1	VCAN	0.062	0	0.955
SDC1	VCAN	0	0	0.952
SDC4	VCAN	0	0	0.953
SELL	VCAN	0	0.316	0.965
TLR2	VCAN	0	0.07	0.998
TLR6	VCAN	0	0.07	0.965

increases as DNA methylation levels decrease. The state of VCAN DNA methylation in human cancer has also been shown to be decreased. We looked further at the protein–protein interaction network of VCAN using the STRING service (von Mering et al. 2003) and Cytoscape tools (Paul Shannon et al. 1971) to identify the VCAN-enriched genes. VCAN and ten additional genes have strong interactions, according to functional interaction network studies. Using the CTD database (Mattingly et al. 2003) and the Cytoscape tool to identify the VCAN gene and drug interactions, it was further discovered that a number of medications may be able to impact the expression of VCAN. For instance, sodium arsenite affects methylation, while aflatoxin B1 and bisphenol A increase it. In conclusion, the considerable prognosis linked to high VCAN expression in breast cancer raises the idea that VCAN could be employed as a biomarker for BC prediction and diagnosis.

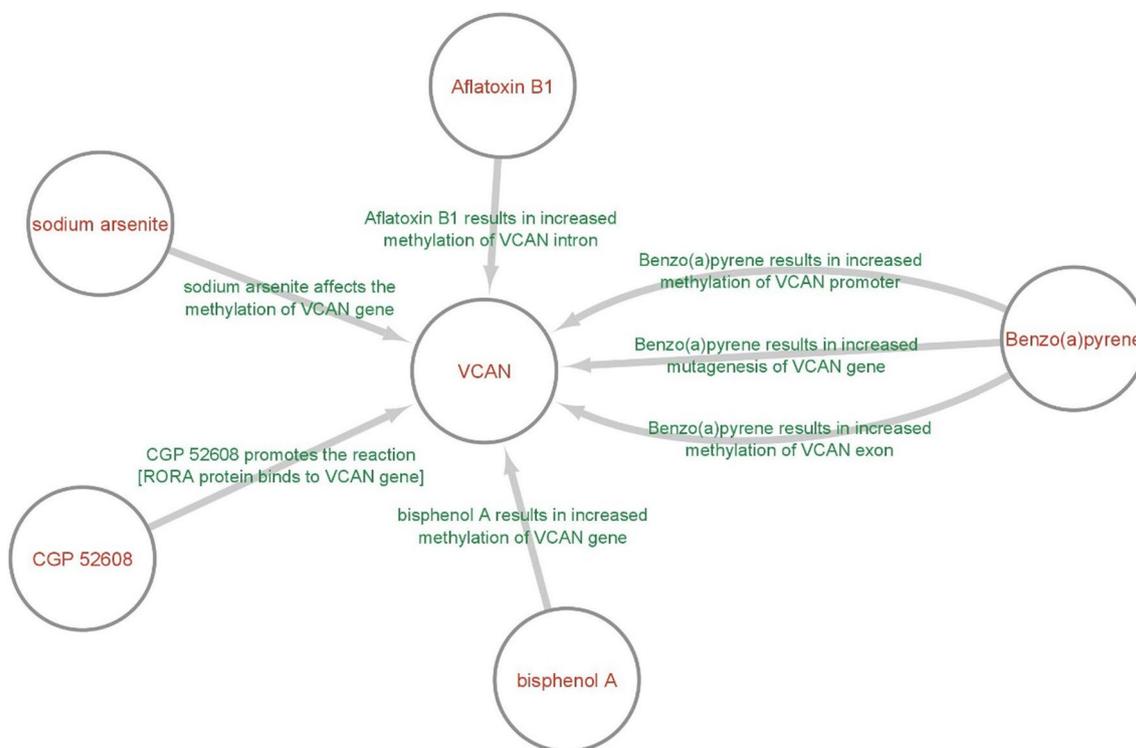


Fig. 8 VCAN-targeted drugs from CTD’s gene–drug interaction network. The numbers of arrows between each interaction show the number of investigations that have supported that interaction

Table 3 Detail of VCAN-targeted drugs from CTD’s gene–drug interaction network

Chemical name	Chemical ID	Interaction
bisphenol A	C006780	Bisphenol A results in increased methylation of VCAN gene
Benzo(a)pyrene	D001564	Benzo(a)pyrene results in increased methylation of VCAN exon
Benzo(a)pyrene	D001564	Benzo(a)pyrene results in increased methylation of VCAN promoter
Benzo(a)pyrene	D001564	Benzo(a)pyrene results in increased mutagenesis of VCAN gene
CGP 52608	C092451	CGP 52608 promotes the reaction [RORA protein binds to VCAN gene]
Aflatoxin B1	D016604	Aflatoxin B1 results in increased methylation of VCAN intron
Sodium arsenite	C017947	Sodium arsenite affects the methylation of VCAN gene
Vehicle emissions	D001335	Vehicle emissions results in increased methylation of VCAN gene

Conclusions

All the analyses revealed that VCAN is overexpressed in the majority of human malignancies. The oncogenesis of these tumors is implicated here, suggesting a role for VCAN. We investigated the expression of mRNA, DNA methylation, mutations and CNAs, associated genes, and prognostic features to assess the viability of VCAN as a potential prognostic marker in BC development. Our significant analysis reveals a clear statement on the overexpression of VCAN that is responsible for the development of breast cancer. According to all of

these multi-omics analyses, VCAN represents a novel diagnostic and prognostic biomarker for breast cancer. Our analysis suggested that further research is required for any therapeutic consequences to treat breast cancer.

Abbreviations

- Bc Breast cancer
- HPA Human Protein Atlas
- IHC Immunohistochemistry staining
- TCGA The Cancer Genome Atlas
- PPI Protein–protein interaction
- CTD Comparative Toxicogenomics Database

Acknowledgements

The authors are immensely grateful to each other for their intellectual contributions to this project.

Author contributions

HJS involved in writing—original draft preparation, resources, data curation, visualization, and investigation. DAK involved in investigation and data curation. MIT involved in investigation and visualization. MI involved in investigation and resources. MNHZ involved in conceptualization, validation, and supervision. MNH involved in conceptualization, validation, and project administration. All authors have read and agreed to the published version of the manuscript.

Funding

No external or internal funding was received for this project.

Availability of data and materials

Not applicable.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 17 March 2023 Accepted: 5 June 2023

Published online: 09 June 2023

References

- Agostinetto E, Gligorov J, Piccart M (2022) Systemic therapy for early-stage breast cancer: learning from the past to build the future. *Nat Rev Clin Oncol* 19(12):763–774
- Al Saber M et al (2022) A comprehensive review of recent advancements in cancer immunotherapy and generation of CAR T cell by CRISPR-Cas9. *Processes* 10(1):16
- Antonius Y et al (2022) Prediction of aflatoxin-B1 (AFB1) molecular mechanism network and interaction to oncoproteins growth factor in hepatocellular carcinoma. *J Pure Appl Microbiol* 16(3):1844–1854
- Azam A, Sounni NE (2022) Lipid metabolism heterogeneity and crosstalk with mitochondria functions drive breast cancer progression and drug resistance. *Cancers* 14(24):1–18
- Azamjah N, Soltan-zadeh Y, Zayeri F (2019) Global trend of breast cancer mortality rate: a 25-year study. *Asian Pac J Cancer Prev* 20:2015–2020
- Biswas P et al (2021) Analysis of SYK gene as a prognostic biomarker and suggested potential bioactive phytochemicals as an alternative therapeutic option for colorectal cancer: an in-silico pharmacoinformatics investigation. *J Personal Med* 11(9):888
- Biswas P et al (2022) A comprehensive analysis and anti-cancer activities of quercetin in ROS-mediated cancer and cancer stem cells. *Int J Mol Sci* 23(19):11746
- Biswas P et al (2023) Advanced implications of nanotechnology in disease control and environmental perspectives. *Biomed Pharmacother* 158:114172. <https://doi.org/10.1016/j.biopha.2022.114172>
- Brunn D et al (2021) Interferon regulatory factor 9 promotes lung cancer progression via regulation of versican. *Cancers* 13(2):1–19
- Cerami E et al (2012) The CBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2(5):401–404
- Chandrashekar DS et al (2017) UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* (United States) 19(8):649–658. <https://doi.org/10.1016/j.neo.2017.05.002>
- Chandrashekar DS et al (2022) UALCAN: an update to the integrated cancer data analysis platform. *Neoplasia* (United States) 25(C):18–27. <https://doi.org/10.1016/j.neo.2022.01.001>
- Dibha AF et al (2022) Utilization of secondary metabolites in algae *Kappaphycus alvarezii* as a breast cancer drug with a computational method. *Pharmacogn J* 14(3):536–543
- Dituri F et al (2022) Proteoglycans in cancer: friends or enemies? A special focus on hepatocellular carcinoma. *Cancers* 14(8):1–21
- Du WW, Yang W, Yee AJ (2013) Roles of versican in cancer biology—tumorigenesis, progression and metastasis. *Histol Histopathol* 28(6):701–713
- Fadholly A et al (2020a) Apoptosis of hela cells via caspase-3 expression induced by chitosan-based nanoparticles of annona squamosa leaf extract. In vitro study. *Indian J Pharm Educ Res* 54(2):416–421
- Fadholly A, Ansori ANM, Sucipto TH (2020b) An overview of naringin: potential anticancer compound of citrus fruits. *Res J Pharm Technol* 13(11):5613–5619
- Fadholly A, Muhammad AAN, Nugraha AP (2020c) Anticancer potential of naringenin: an overview. *Biochem Cell Arch* 20:2971–2977
- Farrior JH “Frank E. Adair, M.D., George T. Pack, M.D., J. H. Farrior, M.D. (From the Mixed Service of the Memorial Hospital).” 1104–20
- Gao J et al (2013) Integrative analysis of complex cancer genomics and clinical profiles using the CBioPortal. *Sci Signal* 6(269):p11
- Han Lu, Ma P, Liu SM, Zhou X (2016) Circulating long noncoding RNA GASS as a potential biomarker in breast cancer for assessing the surgical effects. *Tumor Biol* 37(5):6847–6854
- Hirani P et al (2021) Targeting versican as a potential immunotherapeutic strategy in the treatment of cancer. *Front Oncol* 11(August):1–12
- Jørgensen SM, Lorentzen LG, Chuang CY, Davies MJ (2022) Peroxynitrous acid-modified extracellular matrix alters gene and protein expression in human coronary artery smooth muscle cells and induces a pro-inflammatory phenotype. *Free Radic Biol Med* 186(March):43–52
- Kaddoura R, Alqutami F, Asbaita M, Hachim M (2023) In silico analysis of publicly available transcriptomic data for the identification of triple-negative breast cancer-specific biomarkers. *Life* 13(2):422
- Kharisma VD et al (2020) Molecular simulation: the potency of conserved region on E6 HPV-16 as a binding target of black tea compounds against cervical cancer. *Biochem Cell Arch* 20(August):2795–2802
- Koch A, De Meyer T, Jeschke J, Van Criekeing W (2015) MEXPRESS: visualizing expression, DNA methylation and clinical TCGA data. *BMC Genom* 16(1):1–6. <https://doi.org/10.1186/s12864-015-1847-z>
- Koch A et al (2019) MEXPRESS update 2019. *Nucleic Acids Res* 47(W1):W561–W565
- Kumaratharan N (2015) A review of medical image classification and evaluation methodology for breast cancer diagnosis with computer aided mammography (June 2016)
- Li Bo et al (2016) Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol* 17(1):1–16
- Li T et al (2017) TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Can Res* 77(21):e108–e110
- Li T et al (2020) TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res* 48(W1):W509–W514
- Lin RH et al (2022) Breast cancer survival analysis model. *Appl Sci (switzerland)* 12(4):1–19
- Mattingly CJ, Colby GT, Forrest JN, Boyer JL (2003) The comparative toxicogenomics database (CTD). *Environ Health Perspect* 6:2002–2004
- Munshi M et al (2022) Novel compounds from endophytic fungi of ceriops decandra inhibit breast cancer cell growth through estrogen receptor alpha in in-silico study. *Inform Med Unlocked* 32(March):101046. <https://doi.org/10.1016/j.imu.2022.101046>
- Park SJ, Yoon BH, Kim SK, Kim SY (2019) GENT2: an updated gene expression database for normal and tumor tissues. *BMC Med Genom* 12(Suppl 5):1–8
- Qiu W et al (2008) No evidence of clonal somatic genetic alterations in cancer-associated fibroblasts from human breast and ovarian carcinomas. *Nat Genet* 40(5):650–655
- Ricciardelli C et al (2009) The biological role and regulation of versican levels in cancer. *Cancer Metastasis Rev* 28(1–2):233–245
- Samad A et al (2020) Computational assessment of MCM2 transcriptional expression and identification of the prognostic biomarker for human

- breast cancer. *Heliyon* 6(10):e05087. <https://doi.org/10.1016/j.heliyon.2020.e05087>
- Setoguchi T et al (2011) Microarray analysis identifies versican and CD9 as potent prognostic markers in gastric gastrointestinal stromal tumors. *Cancer Sci* 102(4):883–889
- Shannon P et al (1971) Cytoscape: a software environment for integrated models. *Genome Res* 13(22):426
- Skandalis SS et al (2006) The greatly increased amounts of accumulated versican and decorin with specific post-translational modifications may be closely associated with the malignant phenotype of pancreatic cancer. *Biochim Biophys Acta Gen Subj* 1760(8):1217–1225
- Sung H et al (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 71(3):209–249
- Tang Z et al (2019) GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 47(W1):W556–W560
- Thelin MA et al (2012) Dermatan sulfate is involved in the tumorigenic properties of esophagus squamous cell carcinoma. *Can Res* 72(8):1943–1952
- Uhlén M et al (2015) Tissue-based map of the human proteome. *Science* 347(6220):1260419
- von Mering C et al (2003) STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 31(1):258–261
- Widyananda MH et al (2021) Molecular docking study of sea urchin (*Arbacia lixula*) peptides as multi-target inhibitor for non-small cell lung cancer (NSCLC) associated proteins. *J Pharm Pharmacogn Res* 9(4):484–496
- Wu YJ et al (2005) The interaction of versican with its binding partners. *Cell Res* 15(7):483–494
- Yamaguchi Y (2000) Lecticans: organizers of the brain extracellular matrix. *Cell Mol Life Sci* 57(2):276–289
- Yu XJ et al (2003) Analysis of gene expression profiles in pancreatic carcinoma by using CDNA microarray. *Hepatobiliary Pancreat Dis Int* 2(3):467–470
- Zhang H, Wang M, Xu Y (2023) Understanding the mechanisms underlying obesity in remodeling the breast tumor immune microenvironment: from the perspective of inflammation. *Cancer Biol Med* 20:1–19

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
