Original Article Open Access

Archives Of Breast Cancer

DOI: 10.32768/abc.2024114360-370

Identification of Potential Predictive Transcript Isoform-Biomarkers for the Early Diagnosis of Breast Cancer Using Bioinformatics Tools

Hulya Gundesli*[a], Medi Kori[b,c], Cihan Ergul[a]

[a]Gulhane Faculty of Medicine, University of Health Sciences, Department of Medical Biology, Ankara, Turkey

[b]Institute of Health Sciences, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

[c]Faculty of Health Sciences, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

ARTICLE INFO

Received:
22 May 2024
Revised:
11 September 2024
Accepted:
18 September 2024

Keywords:
Breast Cancer, early-stage, transcript isoform-specific biomarker, diagnostic biomarker

ABSTRACT

Background: Several studies have demonstrated that the expression status of isoforms is more informative as a biomarker than overall gene expression. This study aimed to determine highly but significantly expressed transcript isoforms and evaluate their prognostic and diagnostic impact in breast invasive carcinoma (BRCA) Stage I patients.

Methods: The differentially expressed genes and their transcript isoforms in BRCA Stage I were determined using the Cancer Differentially Expressed Isoform (Cancer DEIso) and gene platform based on The Cancer Genome Atlas (TCGA) data. The prognostic and diagnostic impact of significantly upregulated top 10 genes and their transcripts were determined using the Cancer DEIso tool, the Kaplan-Meier (KM) method, and the Receiver Operating Characteristic Curve (ROC) approach, respectively. Isoform-level protein-protein interactions (PPI) were constructed using the Domain Interaction Graph Guided ExploreR (DIGGER) database. ConsensusPathDB was used to perform pathway enrichment analysis based on the constructed interactions.

Results: The results revealed that NM_024037, NM_001143782, and NM_021619 transcript isoforms have significant diagnostic ability to distinguish stage I BRCA patients from normal with AUC values 93.2%, 77.1% and 75.3%, respectively. KM-plot analysis showed that these three isoforms have no prognostic significance in Stage I patients, but their upregulation was correlated with decreased survival in BRCA patients regardless of stage. Isoform-based pathway enrichment analyses indicated that these three isoforms were involved in chromatin organization, senescence, DNA damage and several signaling pathways which contributes to cancer when there is misregulation.

Conclusion: NM_024037, NM_001143782, and NM_021619 transcript isoforms are potential biomarkers for detecting early-stage BRCA. Thus, it is essential to find out how these three isoforms contribute to the development of breast carcinogenesis and develop a new approach for capturing breast tumors at an earlier stage of the clinical landscape.

***Address for correspondence:**
Hülya Gundesli, MSc, PhD,
Gulhane Faculty of Medicine, University of Health Sciences, Department of Medical Biology, Ankara, Turkey
Email: hulya.gundesli@sbu.edu.tr

INTRODUCTION

Despite the substantial advancements in cancer research, breast invasive carcinoma (BRCA) still holds the record as the most prevalent neoplastic malignancy, especially in women.[1] It accounts for 25% of cancer cases and about 17% of cancer deaths.[2] The GLOBOCAN Cancer Tomorrow prediction tool

estimates that the incidence of the disease will increase by over 55% by 2050.[3] Mangone *et al.* pointed out that the 5-year relative survival rate of BRCA patients at stages I, II, III, and IV was 100%, 89.7%, 71.4%, and 29.1%, respectively (with 95% confidence intervals P-value <0.05).[4] Therefore, detecting breast cancer at an early stage plays a pivotal role in reducing patient mortality and eradicating one of the biggest challenges in healthcare.

BRCA is a genetically and clinically heterogeneous disease with several subtypes. These subtypes have been classified based on the expression of the following hormone receptors: estrogen (ER+), human epidermal growth factor (HER2+), and progesterone (PR+).[5] In addition, there is also the triple-negative (TNBC) subtype, which is characterized by the lack of expression of these receptors.[5] The expression status of these three receptors, Ki67, a proliferative index, and p53, is used as molecular predictive and prognostic breast markers.[6] Furthermore, other markers have been reported such as several gene alterations, miRNAs, urokinase-type plasminogen activator (uPA) and plasminogen activator inhibitor1 (PAI-1), programmed death ligand 1 (PD-L1), and programmed cell death protein 1 (PD-1) receptors as well as microsatellite repeat instability (MSI) for BRCA.[6-11] However, as tumor cells are highly heterogeneous, these biomarkers are unfortunately insufficient to classify BRCA patients and detect tumor formation at earlier stages.

In the search for reliable early detection methods, biomarkers have played an increasingly important role in recent years. They offer a non-invasive, cost-effective, and efficient means of early detection compared to traditional methods such as mammography, ultrasonography, and biopsy. Although biopsy is a reliable tool for a specific diagnosis compared to other methods, it is inefficient as it requires multiple procedures, usually taking 7 to 10 days to obtain a result. Therefore, finding the most efficient biomarker(s) to enable the earliest diagnosis, increase the possibility of personalized treatment, and facilitate monitoring the response to therapy is crucial.

It has been shown that the expression status of specific isoforms is more relevant and informative as a biomarker than overall gene expression for disease diagnosis, subtyping, prognosis, and drug response prediction.[12-15] Additionally, Zhleh *et al.* indicated that expression-based features were significantly more strongly associated with sensitivity to most drugs than mutations and copy number alterations.[15] Several isoform-specific potential biomarkers have been reported as predictive outcome indicators in breast cancer not only for detecting tumor formation or progress but also for drug response.[16-19] Nevertheless, none of them are considered predictive biomarkers for early stage and are routinely used in clinics.

This study used web-based bioinformatics platforms to assess the transcript isoforms expressed in Stage I BRCA tissue samples. We further analyzed differentially expressed isoforms that are not functionally involved in BRCA and aimed to discover novel potential predictive isoform biomarkers that can be used for early diagnosis. As BRCA is a global health concern that requires the development of novel and highly efficient approaches for diagnosis and treatment, this study opens a pathway for the detection of potential early isoform biomarker(s) to improve diagnosis and personalized treatment.

METHODS

Analysis of differentially expressed transcript isoforms in stage I of BRCA

This study used the Cancer DEIso database to search for potential BRCA isoform biomarkers. This database collected different types of data from TCGA, including RNA sequencing and clinical data, and provided information on differential expression at both gene and isoform levels through differential analysis considering the stages of the samples. [20,21] In this database, differential expression analysis is performed using the average fragments per kilobase per million mapped fragments (FPKM) ratio between the selected conditions. Accordingly, differential expression analysis of each transcript in Stage I of BRCA was computed by the average FPKM ratio of Condition 2 (Stage I, n=182) and Condition 1 (Normal, n=113). The false discovery rate (FDR) was used to correct for multiple testing. The q-value and fold-change parameters were calculated to assess the statistical significance of differential expression. The database uses the Cufflinks tool for these calculations. Thus, the fold change of average FPKM was set as ≥ 2. The q-values were calculated by the one-tailed independent t-test to determine whether the average value of FPKM of Condition 2 was greater than Condition 1. The q-value cut-off was defined as 0.05.

Diagnostic performance of differentially expressed transcript isoforms in Stage I BRCA

RNA sequencing by expectation maximization (RSEM) normalized isoform expression data of Stage I BRCA were collected by TCGA. [21] To evaluate the diagnostic performance of the transcript isoforms, the Receiver Operating Characteristic (ROC) curve method was used, which utilizes the parameters of sensitivity (indicating how well the isoform identifies true positives) and specificity (indicating how well

the isoform identifies true negatives) to predict diagnostic ability.

The area under the ROC curve (AUC) was calculated to determine the overall diagnostic accuracy. A transcript isoform with an AUC value ≥ 70% was considered statistically significant [22] and accepted as a diagnostic isoform.

Prognostic performance of differentially expressed transcript isoforms in Stage I BRCA

To evaluate the prognostic performance of isoforms with significant diagnostic ability, clinical information (i.e., days to patient death, patient vital status, and days to last follow-up) of stage I BRCA samples from TCGA was collected and used in the analysis in addition to RNA sequencing-RSEM-normalised isoform expression data. The prognostic abilities of the isoforms were assessed using the Kaplan-Meier (KM) plots and the logrank test. All analyses were performed using the survival package in R/Bioconductor (version 4.0.2). [23] Using this package, the samples were divided into two groups (Group 1 and Group 2) according to the calculated prognostic index, a linear component of the Cox model. The cox proportional hazard ratio [HR = (O1/E1)/(O2/E2)] was calculated based on the ratio between the relative mortality rates in Group 1 and Group 2, where O and E are the observed and expected number of deaths, respectively.

In addition, the prognostic capacity of the isoform biomarker candidates was evaluated for all BRCA samples without consideration of stage differences using the Cancer DEIso database. [20] The database performed a survival analysis using the FPKM values generated by Cufflinks. Accordingly, the database divided the samples into two groups, considering the percentile threshold (50%). One group consisted of patients with higher FPKM values considering the threshold, and the other group consisted of patients with lower FPKM values considering the threshold. Using the created groups, the database performed a survival analysis based on the KM method with Python and calculated the logrank P-value. A transcript isoform with a logrank P-value <0.05 was considered statistically significant in the survival analysis.

Analysis of isoform-specific protein-protein interactions

PPI at the isoform level was reconstructed using the DIGGER database. [24] Accordingly, the Ensemble and UCSC IDs of each differentially expressed transcript isoform were determined, and either protein or domain interactions were reconstructed for differentially expressed transcript isoforms. Using the proteins that interacted with the isoforms, pathway enrichment analyses were performed with ConsensusPathDB using only the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database as a source of the pathways. [25] Fisher's exact test and a FDRadjustment were applied to the P-values in the enrichment analyses. An adjusted P-value <0.01 was considered statistically significant.

RESULTS

Differentially Expressed Transcript Isoforms in Stage I of BRCA

In this study, differentially expressed 6465 transcript isoforms were determined using the Cancer DEIso database, 93 of which could not be directly and functionally linked to BRCA by searching the literature in Pubmed (Supplementary Table 1, Supplementary Table 2). Subsequently, 11 transcript isoforms were selected based on their significantly higher expression than normal tissues in Stage I BRCA. (Supplementary Table 2). Accordingly, NM_001033555 (one of the transcript isoforms of *SPECC1*), NM_001206916 and NM_001206917 (two of *CACNB3* transcript isoforms), NM_024037 (one of the two transcript isoforms of *AUNIP)*, NM_001301824 (one of the *AZIN2* transcript isoforms), NM_001143782 (one isoform of *FKBP11*), NM_080860 (one isoform of *RSPH1*), NM_001297721 (one isoform of *C1orf43)*, NM_001164638 (one isoform of *ENDOV*), NM_001085365 and NM_021619 (transcript isoforms of *MZT2A* and *PRDM12*, respectively) showed q-values between 3.202×10^{-24} and 3.564×10^{-9} as well as fold changes between 5.089 and 2.014 (Figure 1, Table 1).

Diagnostic Performance Analysis of Differentially Expressed Transcript Isoforms

In this study, the TCGA-normalized isoform expressions of stage I BRCA data were used for diagnostic analysis. Since expression data could only be obtained from TCGA for 3 of the 11 statistically significant isoforms, the diagnostic analyses were only performed for these three isoforms (*AUNIP:* NM_024037, *FKBP11*: NM_001143782, *PRDM12*: NM_021619). The ROC curve method was used for the diagnostic performance analyses. AUC values were calculated to understand the discriminatory ability of the isoforms between stage I BRCA patients and normal samples. Accordingly, the transcript isoform NM_024037 with the highest AUC value (0.93) had the most significant diagnostic impact compared to the other two (Figure 2A-C).

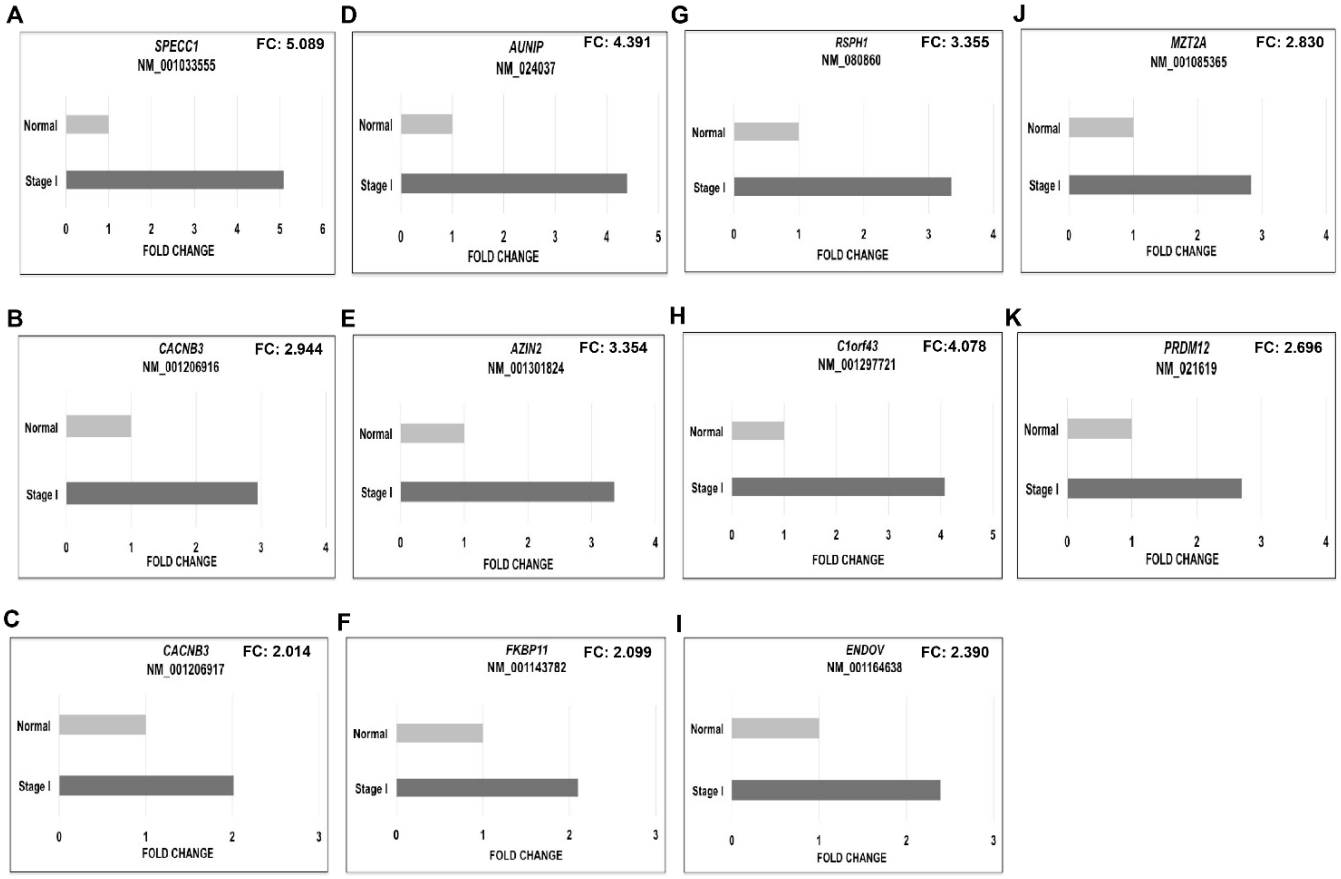

Figure 1. Significantly upregulated 11 transcript isoforms in BRCA stage I. The expression level of A) NM_001033555 transcript isoform of *SPECC1* gene, B) NM_001206916 transcript isoform of *CACNB3* gene, C) NM_001206917 transcript isoform of *CACNB3* gene, D) NM_024037 transcript isoform of *AUNIP* gene E) NM_001301824 transcript isoform of *AZIN2* gene, F) NM_001143782 transcript isoform of *FKBP11* gene, G) NM_080860 transcript isoform of *RSPH1* gene, H) NM_001297721 transcript isoform of *C1orf43* gene, I) NM_001164638 transcript isoform of *ENDOV* gene, J) NM_001085365 transcript isoform of *MZT2A* gene, K) NM_021619 transcript isoform of *PRDM12* gene were found to be increased in BRCA samples compared to normal in Stage I. Differential expression was defined as FC ≥ 2. q ≤ 0.05 was considered statistically significant.

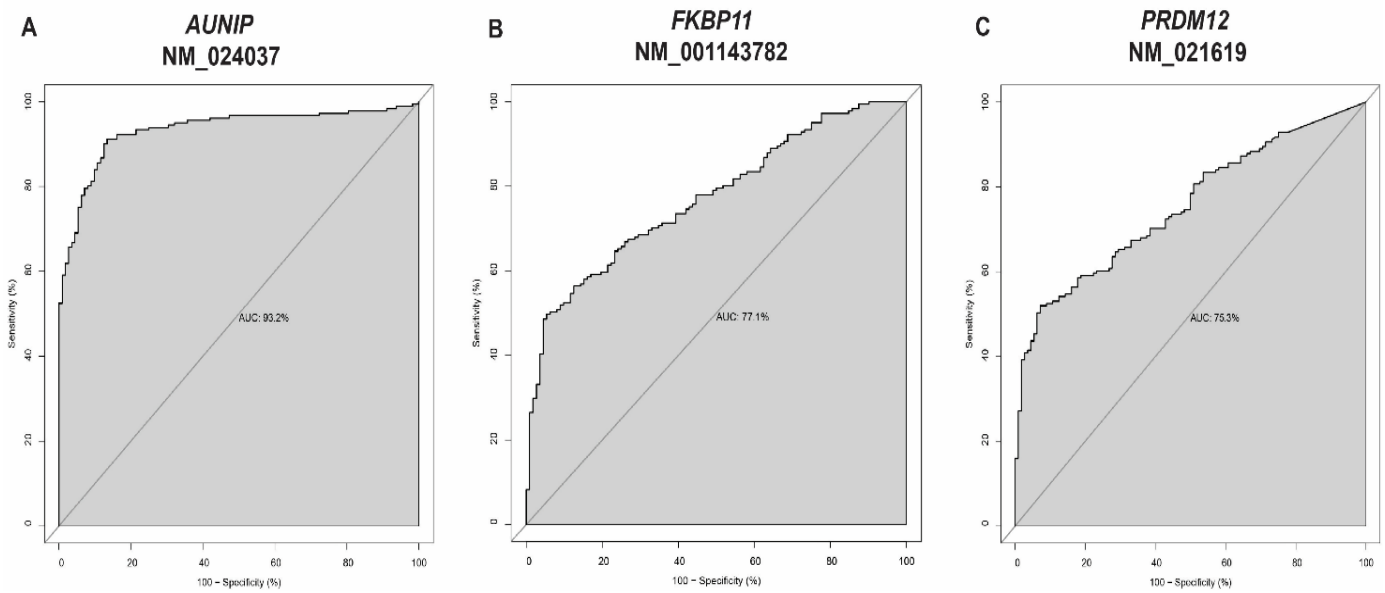

Figure 2. The ROC curve analyses of three upregulated transcript isoforms in BRCA Stage I. Three transcript isoforms: A) NM_024037 (AUNIP), B) NM_001143782 (FKBP11), and C) NM_021619 (PRDM12), with AUC values of 93.2%, 77.1%, and 75.3%, respectively.

In addition, increased expression levels of NM_001143782 and NM_021619 isoforms displayed appropriate AUC values (0.771 for NM_001143782; 0.753 for NM_021619) in Stage I BRCA. Therefore, all these transcript isoforms demonstrated a statistically significant diagnostic capacity (AUC value ≥ 70%).

Table 1. All transcript isoforms with their differential expression and prognostic performance

Gene	RefSeq ID	Avg FPKM Stage I	Avg FPKM Normal	q-Value (Cancer DEIso)	Survival (Logrank p-Value)
SPECC1	NM_001033554	0.767	0.351	6.757e-15	0.216
	NM_001033553	1.434	2.506	1.000e+00	0.581
	NM_001033555	4.824	0.948	3.202e-24	0.023
	NM_001243438	0.218	0.233	7.033e-01	0.656
	NM_001243439	0.217	0.397	9.985e-01	0.903
	NM_152904	0.324	0.931	1.000e+00	0.151
CACNB3	NM_001206916	9.570	3.251	1.051e-23	0.04
	NM_000725	1.955	1.955	8.998e-19	0.183
	NM_001206915	0.011	0.008	3.415e-01	0.808
	NM_001206917	4.950	2.458	3.564e-09	0.042
AUNIP	NM_024037	1.919	0.437	3.521e-22	0.03
AZIN2	NM_001287490	0.182	0.052	4.352e-08	0.779
	NM_001301824	1.261	0.376	1.418e-19	0.043
	NM_001293562	2.984	1.733	1.705e-18	0.775
	NM_001301823	0.247	0.104	2.585e-09	0.637
	NM_001301825	0.115	0.056	1.504e-02	0.576
	NM_001301826	0.442	0.440	4.805e-01	0.941
FKBP11	NM_052998	0.311	0.176	1.743e-10	0.016
	NM_001143782	5.340	2.544	3.357e-19	0.046
	NM_001143781	1.519	0.598	5.958e-04	0.062
RSPH1	NM_016594	28.013	18.501	3.64e-07	0.062
	NM_080860	3.489	1.040	2.277e-18	0.003
C1orf43	NM_001286506	3.907	0.789	5.146e-17	0.065
	NM_138740	18.740	5.048	6.796e-29	0.066
	NM_001098616	74.031	62.210	2.673e-03	0.305
	NM_001297717	1.018	0.168	4.985e-11	0.453
	NM_001297718	7.396	3.909	8.779e-04	0.129
	NM_001297720	12.750	6.839	7.257e-16	0.447
	NM_001297721	1.203	0.295	3.634e-18	0.045
	NM_001297723	1.281	0.683	1.046e-12	0.883
ENDOV	NM_015449	70.457	35.981	2.304e-38	0.562
	NM_001164638	3.086	1.291	3.706e-18	0.019
	NM_001164637	2.377	2.694	9.350e-01	0.484
	NM_173627	1.899	1.427	4.848e-04	0.016
MZT2A	NM_001085365	60.856	21.503	6.990e-18	0.017
PRDM12	NM_021619	0.062	0.023	2.124e-17	0.019

Survival Analysis of Differentially Expressed Transcript Isoforms

To understand the prognostic performance of the isoforms, TCGA-normalized isoform expressions from Stage I BRCA data and clinical data were used and integrated, and survival analyses were performed. In the survival analyses, isoforms NM_024037, NM_001143782, and NM_021619, which have high diagnostic performance, were evaluated using the KM method based on their logrank P-values and hazard ratios. As a result, no statistically significant P-values were observed (for isoform NM_024037 HR: 1.556 and p-value: 0.375; for isoform NM_001143782 HR: 1.749 and P-value: 0.262 and for isoform

NM_021619 HR: 1.646 and P-value: 0.314). This showed that these isoforms were not significantly correlated with overall survival in Stage I BRCA (Figure 3A-C).

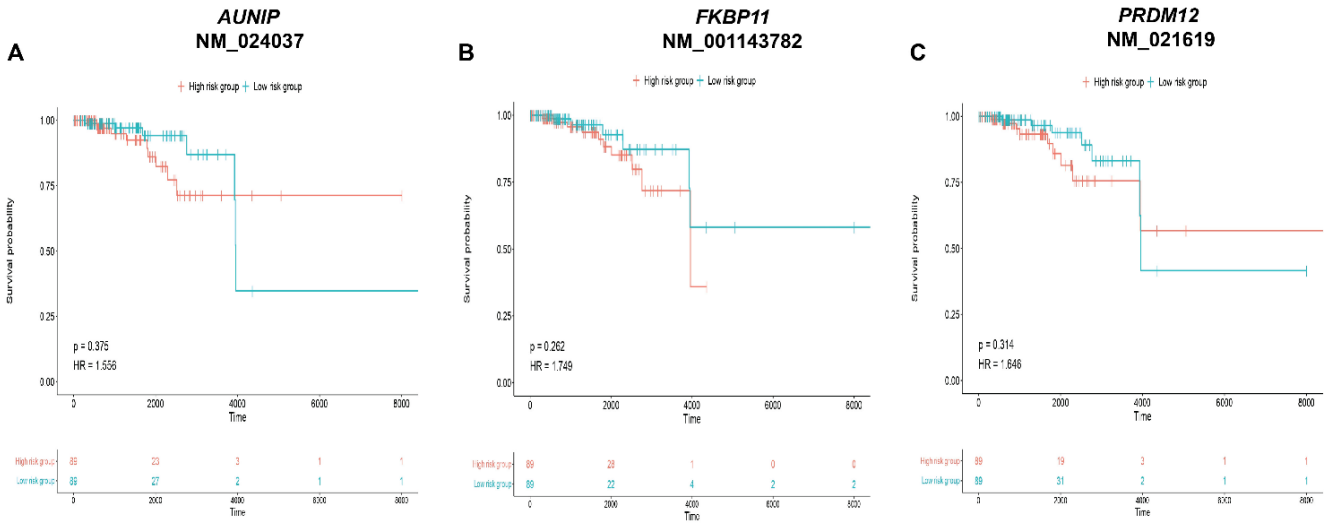

Figure 3. KM-plot analyses of three diagnostically significant transcript isoforms in Stage I of BRCA. The upregulation of A) NM_024037 (*AUNIP*), B) NM_001143782 (*FKBP11*), and C) NM_021619 (*PRDM12*) is not correlated with poor survival in the high-risk group of patients compared with the low-risk group in stage I.

In addition, the prognostic capacity of these three transcripts was evaluated without considering stage differences, and all cancer samples were analyzed simultaneously. The cancer DEIso database was used for these analyses, and again, the results were evaluated using the KM method and logrank P-values. All three transcript isoforms had statistically significant P-values (logrank P-value <0.05). Therefore, these three isoforms were associated with reduced overall survival of BRCA patients (Figure 4A-C).

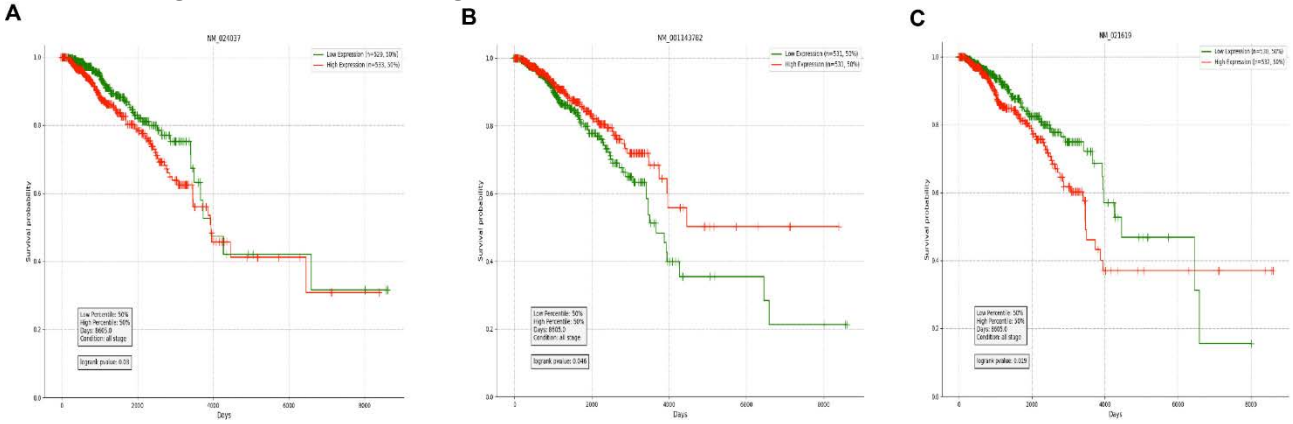

Figure 4. The Survival analyses of three diagnostically significant transcript isoforms in BRCA patients at all stages. The three isoforms A) NM_024037 (longrank P-value: 0.03), B) NM_001143782 (logrank P-value: 0.046), C) NM_021619 (logrank P-value: 0.019) were found to be associated with decreased survival in BRCA patients for all stages. Logrank P-value <0.05 was considered significant.

Isoform Level Protein Interactions

DIGGER database was used to visualize the interaction patterns of three transcript isoforms with significant diagnostic impact. For this purpose, the ensemble IDs of the individual transcripts were obtained and used to determine the isoform-based interactions. Isoform-level interaction analysis showed that AUNIP-201 (ENST00000374298.4) and FKBP11-203 (ENST00000453172.2) interact with seven proteins (Figure 5A-B). In addition, PRDM12-201 (ENST00000253008.3) was the isoform that only interacted with EZH2 (Enhancer of zest homolog 2) (Figure 5C). We could also visualize the domain interaction of PRDM12-201 and EZH2 (Figure 5C). Moreover, a pathway enrichment analysis was performed to understand whether these PPIs might play a role in BRCA initiation/progression, particularly for the isoforms of FKBP11 and AUNIP. Thus, the interacting protein partners of these isoforms were found to be associated with several signalling pathways related to tumor formation or cancer progression (Supplementary Table 3 and Supplementary Table 4).

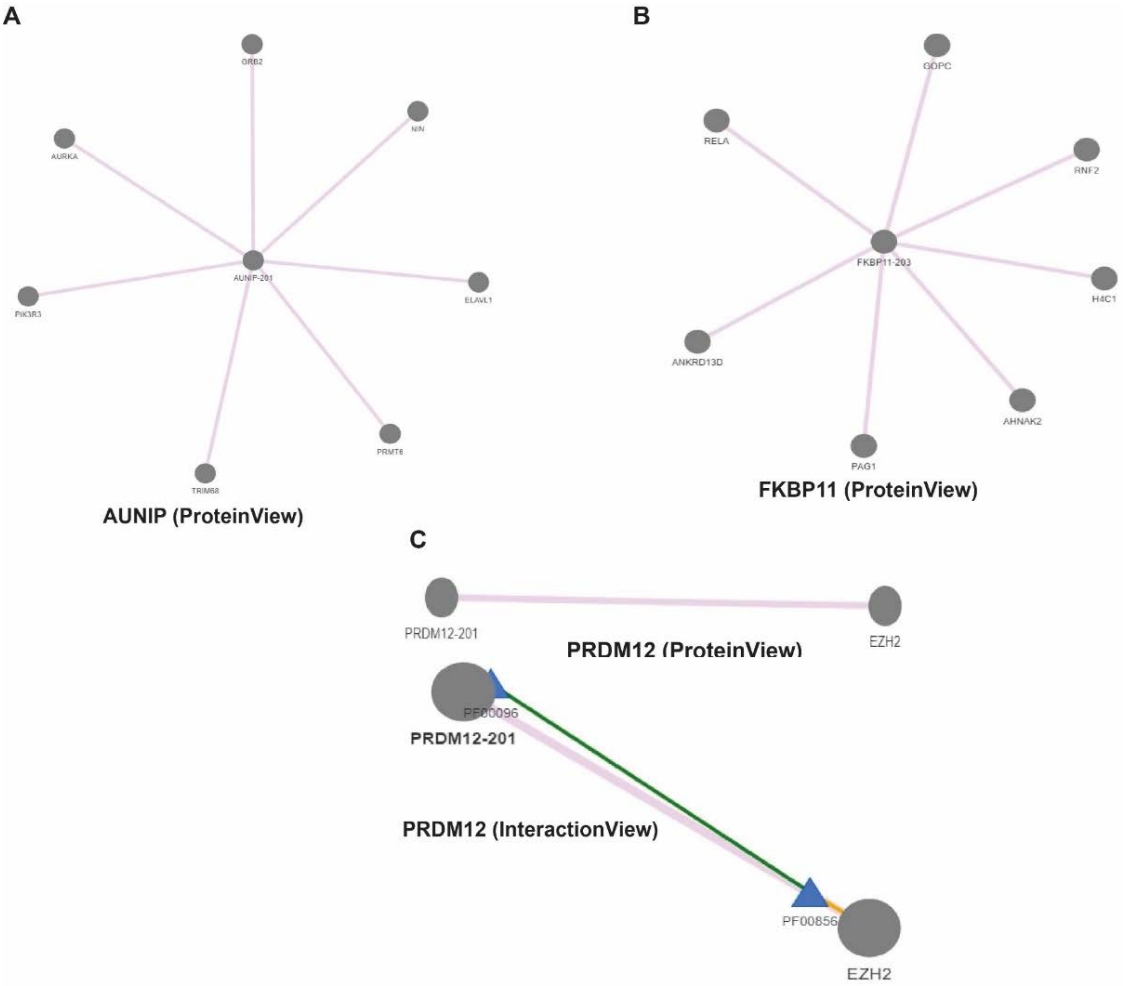

Figure 5. PPI analysis of three transcript isoforms with significant AUC values. The visualizing network of A) AUNIP, B) FKBP11, and C) PRDM12 isoforms interacted with proteins based on the DIGGER database. Transcript isoforms were represented as Ensembl Names, AUNIP-201 (ENST00000374298.4), FKBP11-203 (ENST00000453172.2), and PRDM12-201 (ENST00000253008.3), respectively.

DISCUSSION

Multiple transcript and protein isoforms are generated by alternative splicing events that significantly increase transcriptomic and proteomic diversity. Several studies revealed that cancer exhibits extensive RNA dysregulation, so in addition to gene signature differences, isoform signature differences can be used to distinguish normal tissue from cancer. [26-29] In line with this idea, this study uncovered the various differentially expressed transcript isoforms and investigated whether these spliced isoforms are relevant to the survival and diagnosis of stage I BRCA patients. As a result, our data demonstrated that elevated expression of NM_024037 (belongs to *AUNIP* gene), NM_001143782 (belongs to *FKBP11* gene), and NM_021619 (belongs to *PRDM12* gene) have a high diagnostic performance and could be considered as potential biomarkers for early-stage BRCA diagnosis.

This study showed a significant association between BRCA and the transcript isoform NM_024037, which belongs to the *AUNIP* gene. AUNIP (Aurora Kinase A and Ninein Interacting Protein), also known as AIBp, plays a crucial role in mitotic spindle assembly, the maintenance of centrosomal structure during the cell cycle process, and the repair of DNA double-strand break. [30,31] It has been shown that there is an interaction between AUNIP and Aurora-A, co-expressed in brain tumors. [30] Moreover, one of the studies reported that *AUNIP* expression was upregulated in oral squamous cell carcinoma (OSCC), and depletion of *AUNIP* expression inhibited cell proliferation in OSCC. [32] AUNIP has also been associated with tumor infiltration and might serve as a prognostic and diagnostic biomarker for hepatocellular carcinoma (HCC) and lung adenocarcinoma (LUAD). [33] In addition, a recent study indicated that AUNIP was one of the critical potential diagnostic and therapeutic targets for pancreatic cancer.[34]

To our knowledge, no association between AUNIP and BRCA development has been reported. In this study, ROC curve analysis revealed that one of the two transcript isoforms of the gene, NM_024037, had

the most significant diagnostic ability among the other isoforms with considerable AUC value (93%) and might be an excellent early predictor to detect tumor tissue. On the other hand, increased expression of NM_024037 did not correlate with survival in Stage I since the early-stage survival ratio was very high up to 5 years. Moreover, the enrichment analysis showed that the isoform might be involved in pathways associated with the occurrence and progression of human malignant tumors. Overall, this data suggests that NM_024037 isoform may play a crucial role as a promoter of breast tumor formation and can be used as a diagnostic isoform biomarker. However, functional and more detailed molecular studies are required to understand the function of NM_024037 isoform in BRCA development.

The transcript isoform NM_001143782, which belongs to the *FKBP11* gene, showed statistically significant results for predicting the diagnosis of Stage I BRCA patients. *FKBP11* (FKBP Prolyl Isomerase 11) is a member of FK506-binding proteins (FKBPs), which belong to the family of conservative intracellular immunophilins. [35] The *FKBP* gene family comprises 16 members that play critical roles in various biological activities such as cellular homeostasis, metabolism, T-cell activation, and carcinogenesis. [36,37] However, several studies have demonstrated that increased expression of *FKBP11* promotes tumorigenesis in OSCC, clear cell renal cell carcinoma (ccRCC), LUAD, and osteosarcoma. [38-41] Therefore, these findings suggest that FKBP11 acts on the development of diverse cancer types; however, there is currently no data on FKBP11 to demonstrate its involvement with BRCA development.

Our results put the importance of FKBP11 at the isoform level (NM_001143782) for predicting the prognosis and diagnosis of BRCA. The increased expression of NM_001143782 transcript isoform was associated with decreased survival when differences in the patient stage were not accounted for as expected. In addition, survival analysis of this isoform showed no significance when Stage I BRCA patients were considered. The enrichment analysis also supports the idea that this isoform could contribute to the pathways involved in regulating gene expression, activating pro-tumorigenic signalling, and transforming healthy cells into cancer. Consequently, in vitro and in vivo studies should be performed to clarify the exact role of the NM_001143782 transcript isoform in the development of breast cancer.

Our study also revealed that the transcript isoform NM_021619, which belongs to the *PRDM12*, has a high diagnostic capacity in Stage I BRCA patients. *PRDM12* belongs to the PRDM [PRDI-BF1 (positive regulatory domain I-binding factor 1) and RIZ1 (Retinoblastoma Protein-Interacting Zinc Finger Gene 1) homologous domain containing] gene family that encode for Kruppel-like zinc finger proteins. [42] Histone-modifying enzymes interact with PRDM proteins directly or indirectly. [42,43] Thus, these proteins were proposed to regulate gene expression. [44] Sorrentino *et al.* demonstrated that mRNA expression of *PRDM12* was increased in various cancers, such as breast, ovary, lung, colon, kidney, liver, prostate, and thyroid, compared to normal adult tissues. [45] However, several studies have indicated that derivative chromosome 9 deletion or rearrangements within a region containing the *PRDM12* gene prompt the aggressive, chronic myeloid leukemia (CML) phenotype and correlate with poor survival of patients. [46,47] A recent study has reported that the combined expression of some genes, including *PRDM12*, is associated with poor progression-free survival in BRCA patients with lymph node invasion. [48] Our results showed that high-level expression of *PRDM12* transcript isoform, NM_021619, is linked to poorer survival for all BRCA stages, and that it has a significant ability to distinguish normal from Stage I cancer tissue. Besides, no significant prognostic impact of this isoform has been detected for Stage I. Notably, the interaction between PRDM12 and EZH2, which is also involved in BRCA growth and metastasis, demonstrates the functional relevance of PRDM12 in the development and progression of breast tumors. [49,50] Therefore, accurate functional data are essential to elucidate the possible role of NM_021619 in the early phase of breast tumorigenesis.

Besides these findings, the main limitation of the study is that the data were obtained from bioinformatics analyses. For this reason, these potential early predictive BRCA isoform biomarkers need to be functionally analyzed exhaustively by performing both in vitro and in vivo assays to determine how their translation from bench to bedside will be impactful.

CONCLUSION

This study proposes the three isoforms NM_024037, NM_001143782, and NM_021619 as diagnostic biomarkers that can be used mainly for early-stage BRCA patients (i.e., Stage I). As far as we know, this study is the first to report the analysis of TCGA data based on transcript isoforms at the level of expression, diagnosis, and prognosis in BRCA Stage I. As previously emphasized, analysis at the isoform level is critical for discovering precious biomarkers to impede the misdiagnosis of cancers. Because of this, the present study provides valuable insight into the discovery of valid and reliable BRCA markers and the development of novel personalized

therapeutic strategies for this intractable healthcare problem.

ETHICAL CONSIDERATIONS

Ethical approval is not required as all data used in this study was obtained from a publicly accessible database.

FUNDING

This research received no financial support from any agencies.

CONFLICT OF INTEREST

The authors declare that there is no competing interest.

ACKNOWLEDGEMENTS

The authors would like to declare that there are no additional acknowledgments.

DATA AVAILABILITY

The authors confirm that all data used in this study was publicly accessible.

REFERENCES

1. De Miglio MR and Mello-Thoms C. Editorial: Reviews in breast cancer. Front Oncol. 2023; 13:1161583. doi:10.3389/fonc.2023.1161583.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021; 71:209–49. doi:10.3322/CAAC.21660.
3. GLOBOCAN Cancer Tomorrow Prediction Tool. Available from: https://gco.iarc.fr/tomorrow/en/dataviz/tables?cancers=20&years=2050&types=0.
4. Mangone L, Marinelli F, Bisceglia I, Braghiroli MB, Damato A, Pinto C. Five-year relative survival by stage of breast and colon cancers in northern Italy. Front Oncol. 2022; 12:982461. doi:10.3389/fonc.2022.982461.
5. Orrantia-Borunda E, Anchondo-Nuñez P, Acuña-Aguilar LE, Gómez-Valles FO, Ramírez-Valdespino CA. Subtypes of Breast Cancer. In: Mayrovitz HN. editor. Breast Cancer. Brisbane (AU): Exon Publications. Online first 22 Jun 2022. doi:10.36255/exon-publications-breast-cancer-subtypes.
6. Neves Rebello Alves L, Dummer Meira D, Poppe Merigueti L, Correia Casotti M, do Prado Ventorim D, Ferreira Figueiredo Almeida J, et al. Biomarkers in Breast Cancer: An Old Story with a New End. Genes. 2023; 14: 1364. doi:10.3390/genes14071364.
7. Nair MG, Somashekaraiah VM, Ramamurthy V, Prabhu JS, Sridhar TS. miRNAs: Critical mediators of breast cancer metastatic programming. Exp Cell Res. 2021; 401: 112518. doi: 10.1016/j.yexcr.2021.112518.
8. Uhl B, Mittmann LA, Dominik J, Hennel R, Smiljanov B, Haring F, et al. uPA-PAI-1 heteromerization promotes breast cancer progression by attracting tumorigenic neutrophils. EMBO Mol. Med. 2021; 13: e13110. doi: 10.15252/emmm.202013110.
9. Tu X, Qin B, Zhang Y, Zhang C, Kahila M, Nowsheen S, et al. PD-L1 (B7-H1) Competes with the RNA Exosome to Regulate the DNA Damage Response and Can Be Targeted to Sensitize to Radiation or Chemotherapy. Mol Cell. 2019; 74: 1215–26. doi: 10.1016/j.molcel.2019.04.005.
10. Zhang R, Yang Y, Dong W, Lin M, He J, Zhang X, et al. D-mannose facilitates immunotherapy and radiotherapy of triple-negative breast cancer via degradation of PD-L1. Proc Natl Acad Sci USA. 2022; 119: e2114851119. doi: 10.1073/pnas.2114851119.
11. Klouch KZ, Stern MH, Trabelsi-Grati O, Kiavue N, Cabel L, Silveira AB, et al. Microsatellite instability detection in breast cancer using drop-off droplet digital PCR. Oncogene. 2022; 41: 5289–97. doi: 10.1038/s41388-022-02504-6.
12. Trincado JL, Sebestyén E, Pagés A, Eyras E. The prognostic potential of alternative transcript isoforms across human tumors, Genome Med. 2016; 8: 85. doi:10.1186/s13073-016-0339-3.
13. Pal S, Bi Y, MacYszyn L, Showe LC, O'Rourke DM, Davuluri RV. Isoform-level gene signature improves prognostic stratification and accurately classifies glioblastoma subtypes, Nucleic Acids Res. 2014; 201442: 1–11. doi:10.1093/nar/gku121.
14. Zhang Z, Pal S, Bi Y, Tchou J, Davuluri RV. Isoform level expression profiles provide better cancer signatures than gene level expression profiles, Genome Med. 2013; 5, 33. doi:10.1186/gm437.
15. Safikhani Z, Thu KL, Silvester J, Smirnov P, Lupien M, Mak TW, et al. Gene isoforms as expression-based biomarkers predictive of drug response in vitro. Nat Commun. 2017; 160937. doi:10.1101/160937.
16. Avery-Kiejda, KA, Morten B, Wong-Brown MW, Mathe A, Scott RJ. The relative mRNA expression of p53 isoforms in breast cancer is associated with clinical features and outcome. Carcinogenesis. 2014; 35, 586–96. doi: 10.1093/carcin/bgt411.
17. Safikhani Z, Smirnov P, Thu KL, Silvester J, El-Hachem N, et al. Gene isoforms as expression-based biomarkers predictive of drug response in vitro. Nat Commun. 2017; 24,8(1), 1126. doi: 10.1038/s41467-017-01153-8.
18. Lin T, Qiu Y, Peng W, Peng L. Heat Shock Protein 90 Family Isoforms as Prognostic Biomarkers and Their Correlations with Immune Infiltration in Breast Cancer. Biomed Res Int. 2020: 2020:2148253. doi: 10.1155/2020/2148253.
19. Erdem M, Ozgul I, Dioken DN, Gurcuoglu I, Guntekin-Ergun S, Cetin-Atalay R, et al. Identification of an mRNA isoform switch for HNRNPA1 in breast

cancers. Sci Rep. 2021; 11:24444. doi:10.1038/s41598-021-04007-y.
20. Tzu-Hsien Y, Yu-Hsuan C, Sheng-Cian S, Po-Heng L, Ya-Chiao Y, Kai-Chi T, et al. Cancer DEIso: An integrative analysis platform for investigating differentially expressed gene-level and isoform-level human cancer markers. Comput Struct Biotechnol J. 2021; 19, 5149-59. doi: 10.1016/j.csbj.2021.09.005.
21. Tomczak K, Czerwin´ska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. Contemp Oncol (Pozn). 2015; 19(1A), A68–A77. doi: 10.5114/wo.2014.47136
22. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010; 5(9), 1315–16. doi: 10.1097/JTO.0b013e3181ec173d.
23. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating highthroughput genomic analysis with Bioconductor. Nat Methods. 2015; 12(2), 115–121. doi: 10.1038/nmeth.3252.
24. Louadi Z, Yuan K, Gress A, Tsoy O, Kalinina OV, Baumbach J, et al. DIGGER: exploring the functional role of alternative splicing in protein interactions. Nucleic Acids Research. 2021; 49, 309–18. doi: 10.1093/nar/gkaa768.
25. Kamburov A, Stelzl U, Lehrach H, Herwing R. The ConsensusPathDB interaction database: 2013 update. Nucleic Acids Res. 2013; 41, 793–800. doi:10.1093/nar/gks1055.
26. Zhang Z, Pal S, Yingtao B, Tchou J, Davuluri RV. Isoform level expression profiles provide better cancer signatures than gene level expression profiles. Genome Med. 2013; 5(4), 33. doi: 10.1186/gm437.
27. Mei J, Liu Y, Xu R, Hao L, Qin A, Chu C, Zhu Y, Liu X. Characterization of the expression and prognostic value of 14-3-3 isoforms in breast cancer. Aging (Albany NY). 2020; 12(19), 19597-19617. doi: 10.18632/aging.103919.
28. Gundesli H, Kori M, Arga KY. The Versatility of Plectin in Cancer: A Pan-Cancer Analysis on Potential Diagnostic and Prognostic Impacts of Plectin Isoforms. OMICS, 2023; 27(6):281-296. doi: 10.1089/omi.2023.0053.
29. Yang L, Gilbertsen A, Jacobson B, Pham J, Fujioka N, Henke CA, et al. SFPQ and Its Isoform as Potential Biomarker for Non-Small-Cell Lung Cancer. Int J Mol Sci. 2023; 24(15), 12500. doi: 10.3390/ijms241512500.
30. Lieu AS, Cheng TS, Chou CH, Wu CH, Hsu CY, Huang CYF, et al. Functional characterization of AIBp, a novel Aurora-a binding protein in centrosome structure and spindle formation. Int J Oncol. 2010; 37(2), 429–436. doi: 10.3892/ijo_0000069.
31. Lou J, Chen H, Han J, He H, Huen MSY, Feng XH et al. AUNIP/C1orf135 directs DNA double-strand breaks towards the homologous recombination repair pathway. Nat Commun. 2017; 8(1), 985. doi: 10.1038/s41467-017-01151-w.
32. Yang Z, Liang X, Fu Y, Liu Y, Zheng L, Liu F, et al. Identification of AUNIP as a candidate diagnostic and prognostic biomarker for oral squamous cell carcinoma. EBioMedicine. 2019; 47, 44-57. doi: 10.1016/j.ebiom.2019.08.013.
33. Ma C, Kang W, Yu L, Yang Z and Ding T. AUNIP expression is correlated with immune infiltration and is a candidate diagnostic and prognostic biomarker for hepatocellular carcinoma and lung adenocarcinoma. Front. Oncol. 2020; 10, 590006. doi: 10.3389/fonc.2020.590006.
34. Qu Y, Lu J, Mei W, Jia Y, Bian C, Ding Y et al. Prognostic biomarkers of pancreatic cancer identified based on a competing endogenous RNA regulatory network. Transl Cancer Res. 2022; 11(11), 4019-36. doi: 10.21037/tcr-22-709.
35. Harding MW, Galat A, Uehling DE, Schreiber SL. A Receptor for the Immunosuppressant FK506 Is a Cis-Trans Peptidyl-Prolyl Isomerase. Nature, 1989; 341, 758–60. doi: 10.1038/341758a0.
36. Solassol J, Mange A, Maudelonde T. FKBP Family Proteins as Promising New Biomarkers for Cancer. Curr Opin Pharmacol. 2011; 11, 320–25. doi: 10.1016/j.coph.2011.03.012.
37. Ge Y, Xu A, Zhang M, Xiong H, Fang L, Zhang X, et al. FK506 Binding Protein 10 Is Overexpressed and Promotes Renal Cell Carcinoma. Urol Int. 2017; 98, 169–76. doi: 10.1159/000448338.
38. Qiu L, Liu H, Wang S, Dai XH, Shang JW, Lian XL, et al. FKBP11 promotes cell proliferation and tumorigenesis via p53-related pathways in oral squamous cell carcinoma. Biochem Biophys Res Commun. 2021; 25, 559, 183-90. doi: 10.1016/j.bbrc.2021.04.096.
39. Sun Z, Qin X, Fang J, Tang Y and Fan Y. Multi-Omics Analysis of the Expression and Prognosis for FKBP Gene Family in Renal Cancer. Front Oncol. 2021; 11, 697534. doi: 10.3389/fonc.2021.697534.
40. Wang CC, Shen WJ, Anuraga G, Hsieh YH, Khoa Ta HD, Xuan DTM, et al. Penetrating Exploration of Prognostic Correlations of the FKBP Gene Family with Lung Adenocarcinoma. J Pers Med. 2023; 13, 49. doi:10.3390/jpm13010049.
41. Zeng D, Li J, Yuan X, Cai F, Yu B, Liu L, et al. FKBP11 improves the malignant property of osteosarcoma cells and acts as a prognostic factor of osteosarcoma. Aging (Albany NY). 2023; 15(7), 2450-59. doi: 10.18632/aging.204523.
42. Casamassimi A, Rienzo M, Di Zazzo E, Sorrentino A, Fiore D, Proto MC, et al. Multifaceted Role of PRDM Proteins in Human Cancer. Int J Mol Sci. 2020; 21, 2648. doi: 10.3390/ijms21072648.
43. Rienzo M, Di Zazzo E, Casamassimi A, Gazzerro P, Perini G, Bifulco M, et al. PRDM12 in Health and Diseases. Int. J. Mol. Sci. 2021; 22, 12030. doi:10.3390/ijms222112030.
44. Huang S, Shao G, Liu L. The PR Domain of the Rb-binding Zinc Finger Protein RIZ1 Is a Protein Binding Interface and Is Related to the SET Domain Functioning in Chromatin-mediated Gene Expression. J. Biol. Chem. 1998; 273, 15933–15939. doi: 10.3390/ijms19103250.
45. Sorrentino A, Federico A, Rienzo M, Gazzerro P, Bifulco M, Ciccodicola A, et al. PR/SET Domain Family and Cancer: Novel Insights from the Cancer Genome Atlas. Int. J. Mol. Sci. 2018; 19, 3250. doi: 10.3390/ijms19103250.

46. Reid AG, Nacheva EP. A potential role for PRDM12 in the pathogenesis of chronic myeloid leukaemia with derivative chromosome 9 deletion. Leukemia. 2003; 18, 178–80. doi: 10.1038/sj.leu.2403162.
47. Huet S, Dulucq S, Chauveau A, Ménard A, Chomel JC, Maisonneuve H, et al. Molecular characterization and follow-up of five CML patients with new BCR-ABL1 fusion transcripts. Genes Chromosom. Cancer. 2015; 54, 595–605. doi: 10.1002/gcc.22263.
48. Kuo CY, Moi SH, Hou MF, Luo CW, Pan MR. Chromatin Remodeling Enzyme Cluster Predicts Prognosis and Clinical Benefit of Therapeutic Strategy in Breast Cancer. Int. J. Mol. Sci. 2023; 24, 5583. doi:10.3390/ijms24065583.
49. Zhang L, Qu J, Qi Y, Duan Y, Huang YW, Zhou Z, et al. EZH2 engages TGFβ signaling to promote breast cancer bone metastasis via integrin β1-FAK activation. Nat Commun. 2022; 13(1), 2543. doi: 10.1038/s41467-022-30105-0.
50. Zhao Y, Hu Z, Li J, Hu T. EZH2 Exacerbates Breast Cancer by Methylating and Activating STAT3 Directly. J. Cancer. 2021; 12(17), 5220-30. doi: 10.7150/jca.50675. eCollection 2021.

How to Cite This Article

Gundesli H , Kori M, Ergul C. Identification of Potential Predictive Transcript Isoform-Biomarkers for the Early Diagnosis of Breast Cancer Using Bioinformatics Tools. Arch Breast Cancer. 2024; 11(4):360-70.
Available from: https://www.archbreastcancer.com/index.php/abc/article/view/950