Heliyon 10 (2024) e27011

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Txnrd1 as a prognosticator for recurrence, metastasis and response to neoadjuvant chemotherapy and radiotherapy in breast cancer patients

Raghavendra S. Patwardhan [a], Archita Rai [a,b], Deepak Sharma [a,b], Santosh K. Sandur [a,b,**], Sejal Patwardhan [b,c,*]

[a] Radiation Biology & Health Sciences Division, Bhabha Atomic Research Centre, Trombay, Mumbai, 400085, India
[b] Homi Bhabha National Institute, Mumbai, 400094, India
[c] Patwardhan Lab, Advanced Centre for Treatment Research & Education in Cancer, (ACTREC), Tata Memorial Centre (TMC), Kharghar, Navi Mumbai, 410210, India

ARTICLE INFO

Keywords:
Thioredoxin reductase
Nrf2
ROS
Ionizing radiation
Therapy response

ABSTRACT

Thioredoxin reductase 1 (Txnrd1) is known to have prognostic significance in a subset of breast cancer patients. Despite the pivotal role of Txnrd1 in regulating several cellular and physiological processes in cancer progression and metastasis, its clinical significance is largely unrecognized. Here, we undertook a retrospective comprehensive meta-analysis of 13,322 breast cancer patients from 43 independent cohorts to assess prognostic and predictive roles of Txnrd1. We observed that Txnrd1 has a positive correlation with tumor grade and size and it is over-expressed in higher-grade and larger tumors. Further, hormone receptor-negative and HER2-positive tumors exhibit elevated Txnrd1 gene expression. Patients with elevated Txnrd1 expression exhibit significant hazards for shorter disease-specific and overall survival. While Txnrd1 has a positive correlation with tumor recurrence and metastasis, it has a negative correlation with time to recurrence and metastasis. Txnrd1^{High} patients exhibit 2.5 years early recurrence and 1.3 years early metastasis as compared to Txnrd1^{Low} cohort. Interestingly, patients with high Txnrd1 gene expression exhibit a pathologic complete response (pCR) to neoadjuvant chemotherapy, but they experience early recurrence after radiotherapy. Txnrd1^{High} MDA-MB-231 cells exhibit significant ROS generation and reduced viability after doxorubicin treatment compared to Txnrd1^{Low} MCF7 cells. Corroborating with findings from meta-analysis, Txnrd1 depletion leads to decreased survival, enhanced sensitivity to radiation induced killing, poor scratch-wound healing, and reduced

Abbreviations: Auranofin, AF; DFS, disease-free survival; DSS, disease-specific survival; DRFS, distant recurrence-free survival; DMFS, distant metastasis-free survival; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; HR, hazard ratio; LABC, locally advanced breast cancer; MFI, metastasis-free interval; OS, overall survival; pCR, pathologic complete response; PFS, progression-free survival; PR, progesterone receptor; RD, residual disease; RCB, residual cancer burden; RFS, recurrence-free survival; ROS, reactive oxygen species; RR, relative risk; Trx, thioredoxin; TrxR, thioredoxin reductase; Txnrd1, thioredoxin reductase 1; TNBC, triple-negative breast cancer; QoL, quality of life.

* Corresponding author. Homi Bhabha National Institute Principal Investigator, Patwardhan Lab, Advanced Centre for Treatment, Research & Education in Cancer (ACTREC), Tata Memorial Centre (TMC) Kharghar, Navi Mumbai, 410210, India.

** Corresponding author. Homi Bhabha National Institute Head, Free Radical Biology Section Radiation Biology & Health Sciences Division, Bhabha Atomic Research Centre, Trombay, Mumbai, 400085, India.

E-mail addresses: sskumar@barc.gov.in (S.K. Sandur), spatwardhan@actrec.gov.in (S. Patwardhan).

https://doi.org/10.1016/j.heliyon.2024.e27011
Received 17 August 2023; Received in revised form 17 January 2024; Accepted 22 February 2024
Available online 29 February 2024

invasion potential in MDA-MB-231 cells. Thus, Txnrd1 appears to be a potential predictor of recurrence, metastasis and therapy response in breast cancer patients.

1. Introduction

Breast cancer is a heterogeneous disease both at the clinical and molecular level with 4 major molecular subtypes and 21 distinct histological subtypes. Almost 81% of breast cancers are invasive, about 83% are hormone receptor-positive, whereas approximately 15% are HER2-positive [1]. Age-standardized incidence and mortality rate for breast cancer is 47.8 and 13.6 per 100,000 cases, respectively [2]. Mortality due to breast cancer has significantly declined over the past few decades due to improved treatments and early detection [3]. However, there is an overall increase in the incidence of local-stage disease [4–6]. The use of chemo, hormonal or targeted molecular therapy shows clinical partial/complete regression and better pathological complete response (pCR) for locally advanced breast cancers (LABC), triple-negative breast cancer (TNBC) and metastatic disease when combined with surgery [7,8]. However, there is still a risk of loco-regional or distant recurrence and second primary tumors among these patients that reduces disease-free survival (DFS), distant metastasis-free survival (DMFS), progression-free survival (PFS), leading to poor quality of life (QoL) among cancer survivors [9,10]. Local, regional or distant breast cancer recurrence may occur due to several factors that sustain or promote the survival of a few cancer cells, which were intended to be eliminated by therapy [9,11]. Tumor size, lymph node involvement, close tumor margins, age, lack of hormone, radiotherapy, and inflammatory breast cancer are some contributory factors [9,11,12].

Moreover, other than these traditional histopathological prognostic factors, numerous factors related to patient and tumor characteristics can influence breast cancer progression and survival. Hence, gene expression profiling is clinically essential for prognosis and treatment planning [13,14]. Similarly, assessing the response rate to standard preoperative chemotherapy requires tools for evaluating the risk of recurrence. While pCR and residual cancer burden (RCB) are significant predictors of recurrence-free survival (RFS) for all breast cancer patients [15], predicting breast cancer pCR to neoadjuvant chemotherapy demands the identification of candidate genes/pathways associated with poor response. Radiotherapy is often delivered post-breast-conserving surgery with systemic adjuvant treatment to destroy remnant cancer cells in the breast, chest wall or underarm area to reduce the risk of recurrence [16–18]. However, in many cases, local or ipsilateral breast tumor recurrence is observed even after postoperative radiotherapy [19]. Tumors exhibit intrinsic radio-resistance and may also acquire a radioresistant phenotype after exposure to fractionated irradiation during radiotherapy [20]. This limits the outcome of radiotherapy, leading to therapy failure or disease return. Hence, tools or methodologies that may aid in educated decisions regarding treatment choice hold great translational potential in enhancing therapeutic gain. In this regard, gene expression profiling may function as a guiding principle in predicting tumor behaviour, risk of recurrence, therapy response and overall clinical outcome of the patient [21–23]. Hence, there is a need to emphasize the basis of treatment by evaluation of tumor gene profiling in the current standard of care to enhance real-world outcomes and global health status of breast cancer patients.

We and others have reported that over-expression of Nrf2 and its dependent oxidative stress response pathway regulates intrinsic and acquired radio-resistance in cancer cells [24–26]. Thioredoxin reductase (TrxR), under transcriptional control of Nrf2, is a pivotal intracellular redox-regulatory protein. TrxR is a dithiol-disulfide oxido-reductase, that reduces critical cysteine residues over several proteins involved in antioxidant defense (thioredoxin, peroxiredoxin, methionine sulfoxide reductase), cell survival (PTEN), apoptosis (ASK1), transcription factors (Nrf2, NF-κB, p53, Hif1α), and DNA synthesis (ribonucleotide reductase), redox factor-1 (Ref-1) etc. [27,28]. TrxR is gaining wider attention as a pharmacologic target for cancer therapy [29–31], and we have recently reviewed the literature highlighting the emergence of TrxR as an attractive target for tumor radio-sensitization [32]. Thioredoxin reductase 1 (Txnrd1) is over-expressed in cholangiocarcinoma, colon adenocarcinoma, esophageal carcinoma, head & neck squamous cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous carcinoma and stomach adenocarcinoma [33,34]. Txnrd1 over-expression is associated with dysplastic transformation of human breast epithelial cells with a certain degree of malignancy [35]. Thioredoxin system is reported to be involved in breast cancer cell invasion, metastasis and poor prognosis of breast cancer patients [36,37].

Given the pivotal role of Txnrd1 in regulating invasion, metastasis, chemoresistance and radioresistance of breast cancer cells, it warrants a detailed investigation in a larger cohort, preferably across the available global datasets, to make it more robust. To this end, collating and analyzing data from public gene expression datasets can be highly rigorous and resourceful. However, to our knowledge, there is no systematic meta-analysis assessing Txnrd1 gene expression across different clinicopathological parameters of breast cancer, probing its possible predictive value for recurrence, metastasis and response to neoadjuvant chemotherapy and radiotherapy. The objective of this retrospective analysis is to perform a comprehensive scrutiny to investigate prognostic and/or predictive role of Txnrd1 in breast cancer patients. Here, we performed a systematic analysis of RNAseq and microarray data of 13,322 breast cancer patients from 43 independent gene expression datasets to examine the association of Txnrd1 expression with tumor grade, size, stage, histology, menopause status, pathologic T & N stage, PAM50 molecular subtypes, overall survival (OS), disease-specific survival (DSS), recurrence and time to recurrence, metastasis and time to metastasis, response to neoadjuvant chemotherapy and radiotherapy. Further, we have performed experiments in relevant breast cancer cell lines to validate the interpretations of this study. Our findings project Txnrd1 as a valuable predictor for disease recurrence and therapy response in breast cancer patients.

Table 1
List of datasets used in the study.

Sr. No.	Dataset	Technology	Parameters	No. of patients	Ref.
1	TCGA	RNA-seq	ER, PR, HER2, PAM50, N, T, histology, menopause, OS, DSS, DMI, DFS, PFS	1218	[59]
2	Metabric	RNA-seq	Grade, ER, PR, HER2, PAM50, N, T, histology, menopause, recurrence, RFS, OS, DSS, DMI, DFS, radiotherapy	1904	[60]
3	GSE11121	Affymetrix HGU	Grade, size, metastasis, DMFS	200	[61]
4	GSE2034	Affymetrix HGU	Bone relapse	286	[62]
5	GSE9893	MLRG Human 21K V12.0	Size, Grade, T, N, histologic subtype, recurrence, OS, metastasis	132	[63]
6	GSE10886	Agilent Human 1A	ER, PR, HER2, node, grade, size, PAM50, recurrence, RFS, OS	189	[64]
7	GSE18229	Agilent Human 1A	Grade, node, size, ER, PR, HER2, PAM50, recurrence, OS, RFS	337	[65]
8	GSE21653	Affymetrix HGU	Grade, ER, PR, PAM50, T, N, histologic subtype, DFS, DRFS	266	[66]
9	GSE22219	Illumina humanRef-8 v1.0 expression beadchip	Grade, node, size, ER, recurrence, Metastasis, DMFS	216	[67]
10	GSE22226	Agilent Whole Human Genome Microarray	ER, PR, HER2, PAM50, grade, T, size, recurrence, RFS, OS	221	[68]
11	GSE30682	Illumina HumanWG-6	ER, PR, HER2, PAM50, T, N, histologic subtypes, recurrence, RFS	165	[69]
12	GSE31448	Affymetrix HGU	ER, PR, T, N, histologic subtypes	353	[66]
13	GSE31863	SWEGENE Human	Histologic subtypes, ER, PR	143	[70]
14	GSE76275	Affymetrix HGU	Histologic subtypes, menopause status	265	[71]
15	GSE80999	Agilent-028,004 SurePrint G3 Human GE	Grade, ER, subtype,	381	[72]
16	GSE103744	Illumina HumanHT-12	ER	172	[73]
17	GSE119295	Affymetrix Human Exon 1.0 ST Array	ER, PR, HER2, grade	765	[74]
18	GSE180962	Agilent_human_DiscoverPrint	HR, HER2, pCR	163	[75]
19	GSE6861	Affymetrix Human X3P Array	pCR	161	[76]
20	GSE22093	Affymetrix HGU	Grade, T, N, pCR	103	[77]
21	GSE41998	Affymetrix HGU	ER, PR, HER2, menopause, pCR, RCB	279	[78]
22	GSE32603	UCSF/HAQQLAB_Human_40,986_ISPY	HR, HER2, RCB, histology, menopause, recurrence, RFS	221	[79]
23	GSE25055	Affymetrix HGU	Grade, ER, PR, HER2, PAM50, N, T, pCR, recurrence, RFS	310	[80]
24	GSE25065	Affymetrix HGU		198	
25	GSE25066 (25,055/ 25,065)	Affymetrix HGU		508	
26	GSE34138	Illumina HumanWG-6	ER, PR, PAM50, pCR	178	[81]
27	GSE130788	Agilent-014,850 Whole Human Genome Microarray	pCR	199	[82]
28	GSE4056	DKFZ/Operon Human Oligo	Grade, pCR	100	[83]
29	GSE158309	Affymetrix HGU	Grade, Size, N, Subtype, Metastasis, DMFS	461	[84]
30	GSE3494	Affymetrix HGU	Size, DSS, ER, PR	251	[85]
31	GSE2990	Affymetrix HGU	ER	189	[86]
32	GSE39004	Affymetrix Human Gene 1.0 ST Array	OS	72	[87]
33	Caldas 2007		Grade, size, DFS, recurrence, OS,	242	[88]
34	Chin 2006	Affymetrix HGU	Size, N, T, DSS, recurrence,	173	[89]
35	Hess 2006	Affymetrix HGU	Grade, size, N, T, histology	133	[90]
36	Miller 2005	Affymetrix HGU	ER, PR, size, DSS	251	[91]
37	VantVeer 2002	Affymetrix HGU	Grade, size, metastasis,	117	[92]
38	Vijver 2002	Affymetrix HGU	OS, metastasis, DMFS	295	[93]
39	Yau 2010 (GSE2034/ 5327/7390)	Affymetrix HGU	ER, PAM50, metastasis, DMFS	683	[94]
40	LN Wang 2005	Affymetrix HGU	ER, recurrence, RFS	286	[62]
41	Desmedt 2007	Affymetrix HGU	Size, DSS, DMFS, OS	198	[95]
42	GSE124647	Affymetrix HGU	PFS, OS, PR, stage	140	[96]
43	GSE7390	Affymetrix HGU	Size, ER, RFS, DMFS, OS	198	[97]

2. Materials & methods

2.1. Cells and reagents

The human breast cancer cell lines MCF-7, MDA-MB-468 and MDA-MB-231 were procured from the National Centre for Cell Science (NCCS, Pune, India). Cells were cultured in DMEM medium containing 10% (v/v) fetal bovine serum (FBS), 200 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin (HiMedia, India). Thioredoxin reductase assay kit was purchased from Cayman Chemicals (IN, USA). The RNA isolation kit was purchased from HiMedia, India. cDNA synthesis and RT-PCR kit were procured from Roche (MA, USA). Doxorubicin, auranofin (AF), MTT, formaldehyde, crystal violet and H_2DCFDA were procured from

Merck (NJ, USA). TrxR1 CRISPR plasmids, which consist of a pool of three plasmids, each encoding the Cas9 nuclease and a TrxR1-specific 20 nt guide RNA (gRNA) derived from the GeCKO (v2) library designed for maximum knockout efficiency was purchased from Santacruz Biotechnology (TX, USA). Lipofectamine 3000 transfection reagent was purchased from Invitrogen. All other chemicals used were of analytical grade.

2.2. Study design and gene expression data

Literature was reviewed and studies to be included in the meta-analysis were selected and assessed for suitability. To avoid the introduction of bias, we did not set any exclusion criteria. This retrospective study used data from 43 different public gene expression datasets generated by microarray/RNAseq analysis of tumor specimens from 13,322 patients with breast cancers. The details about the dataset, platform used, cohort size, parameters studied and reference to the original study are outlined in Table 1. NCBI Gene Expression Omnibus, GEO (http://www.ncbi.nlm.nih.gov/geo/) was searched for gene expression studies. GEO was chosen as it is a comprehensive and MIAME (Minimum Information About a Microarray Experiment) compliant database, studies are published in peer-reviewed journals, and clinical outcome is published and available for analysis. Raw as well as normalized gene expression files are deposited in GEO and available for download and downstream processing. We downloaded pre-processed, normalized data from GSE series matrix files associated with each study accessible at GEO. Data from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database were obtained via the cBioPortal (http://www.cbioportal.org/). Breast cancer data for Caldas 2007, Chin 2006, Hess 2006, Miller 2005, VantVeer 2002, Vijver 2002, Yau 2010 (GSE2034/5327/7390), LN Wang 2005, Desmedt 2007 and data for differential gene expression of Txnrd1 in breast cancer versus normal tissue were available from (https://xena.ucsc.edu/). TCGA RNA-seq data was retrieved using UCSC Xena platform. The normalized gene expression data was accessed from UCSC Xena repository. Details on the data generation are available in the original TCGA study [38]. Details about the data normalization and pre-processing are available on UCSC Xena webpage. Protein expression data from Clinical Proteomic Tumor Analysis Consortium (CPTAC) for breast cancer and data for human protein atlas was obtained from (http://ualcan.path.uab.edu/analysis-prot.html) and https://www.proteinatlas.org/ENSG00000198431-TXNRD1/pathology/breast+cancer respectively. The rest of the datasets were available from (https://www.ncbi.nlm.nih.gov/geo/). Data for Txnrd1 expression in different breast cancer cell lines by Neve 2006 & Heiser 2012 was available from (https://xena.ucsc.edu/).

2.3. Optimal cut-point selection

Numerous studies discuss diverse methodologies to determine the optimal cut-off for continuous variables, such as gene expression data [39,40]. There is an emphasis on a data-driven approach to determine the optimal cut-off for maximum significance between the chosen arms (high vs low). However, outcome-oriented methods are expected to have better statistical indicators than data-oriented methods [41]. Given the challenges associated with optimal cut-points in evaluating quantitative prognostic factors, we used cut-off finder, an easy-to-use web application for determining cut-off points in molecular data [42]. Using Txnrd1 as a biomarker and recurrence/metastasis/DFI/DSS/DMFS/OS/PFS/RFS/tumor grade/tumor size as outcomes, multiple methods were employed for cut-off determination viz. Fisher's exact test, ROC curve (Euclidean distance), ROC curve (Manhattan distance), minimum sensitivity, minimum specificity, log-rank test, etc. We observed that by selecting the correct outcome classification (in ~75% population) and sensitivity ~25%, specificity ~75% as threshold criteria, the study population need to be divided into quartiles according to Txnrd1 gene expression. Hence, across datasets, patients with Txnrd1 gene expression $<$ Q1 (25th percentile) were categorized as $Txnrd1^{Low}$ while those with Txnrd1 gene expression $>$ Q3 (75th percentile) were stratified as $Txnrd1^{High}$ subgroups respectively in the present study [43].

2.4. Study endpoints

Study endpoints were divided into three main categories. Cross-study comparison involved analyzing Txnrd1 gene expression across various clinicopathological parameters in individual datasets. Data used in this analysis was not normalized to avoid masking of underlying associations due to the normalization process. Survival analysis was conducted to determine the cumulative mean time to outcome following data normalization to derive meaningful conclusions using pooled data. Finally, meta-analysis was performed to combine effect size (hazard ratio, relative risk) across datasets.

2.5. Cross-study comparison

The mean expression of Txnrd1 was calculated across various clinical parameters, including tumor grades, pathologic T & N stage, ER, PR, and HER2 expression status, molecular subtypes, recurrence and metastasis, and pathologic complete response. Spearman's correlation coefficients were computed for tumor size, grade, recurrence, metastasis, and their intervals using IBM SPSS version 21. To assess the significance of differences, t-tests were employed for two-group comparisons, while one-way ANOVA was utilized for scenarios involving more than two groups, with a significance level set at $p < 0.05$. The expression heatmap representing Txnrd1 across different parameters was generated using the "pheatmap" package, and a forest plot was created using the "ggplot2" and "ggrepel" packages in R software. Additional plots were generated using GraphPad Prism.

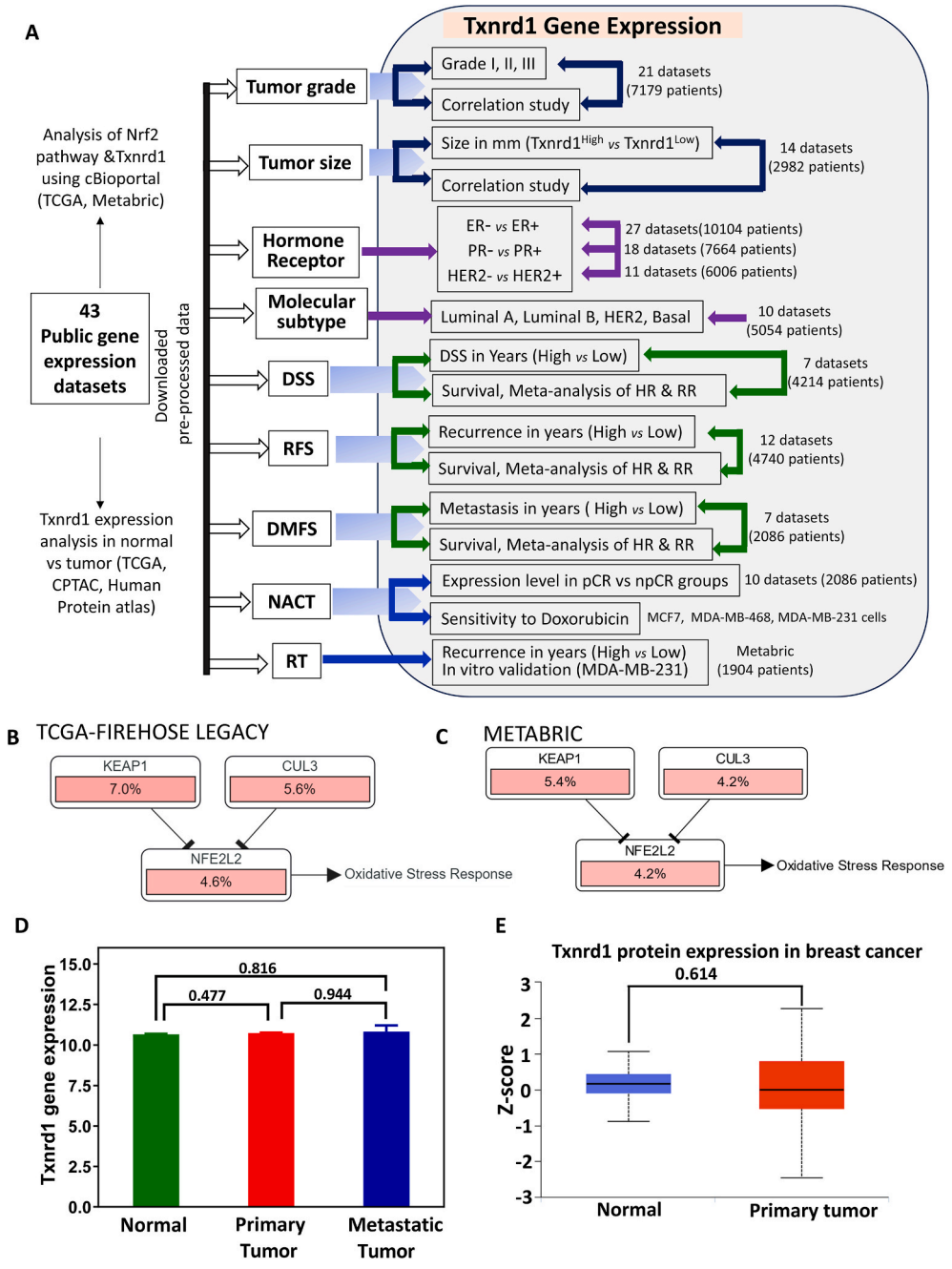


Fig. 1. Txnrd1 is not over-expressed in breast tumor tissue

Workflow outlining the strategy employed for analysis of public gene expression data with respect to Txnrd1 gene expression is shown in (A). Pathway analysis was conducted using cBioPortal, and Txnrd1 gene expression between normal and tumor tissue was assessed using TCGA data. Txnrd1 protein expression and prognostic value were estimated using CPTAC and human protein Atlas, respectively. Pre-processed, normalized data was downloaded and used for downstream analysis. Briefly, Txnrd1 gene expression was assessed across tumor grade, size, hormone receptor expression status and molecular subtypes. Survival analysis was conducted for DSS, OS, RFS & DMFS between Txnrd1 high vs low cohort. Meta-analysis for hazard ratio (HR) and relative risks (RR) were performed. The number of datasets used and cohort size are indicated at respective places. Response to neoadjuvant chemotherapy (NACT) and radiotherapy (RT) was performed with respect to Txnrd1 gene expression status. Validation experiments were performed in MCF7, MDA-MB-468 & MDA-MB-231 cells (A). Pathway level alteration in Nrf2, Keap1 and Cul3 gene expression as per TCGA-Firehose Legacy (B) and METABRIC (C) obtained from cBioportal. Bar graph shows Txnrd1 gene expression from TCGA-BRCA database for normal tissue (n = 114), primary tumor (n = 1097) and metastatic tumor (n = 7). Data is presented as mean ± SEM (D). Box plot shows Z-values for protein expression of Txnrd1 in normal (n = 18) and primary tumor tissue (n = 125) (E). Interactive survival scatter plots and survival analysis of Txnrd1 in breast cancer from human protein atlas is shown in (F). mRNA expression z-scores relative to all samples

from TCGA-Firehose Legacy database with expression heatmap obtained from cBioportal is shown in (G). Statistical significance is calculated using One-way ANOVA for (D) and Student's t-test for (E).

2.6. Survival analysis

Txnrd1 gene expression data was quantile normalized for individual datasets. The data was pooled, and stratification into $Txnrd1^{High}$ and $Txnrd1^{Low}$ subgroups was performed by a quartile-based method, as explained in 2.3. Quantile normalization is the most preferred data normalization method for microarray and RNA-seq data. It transforms the statistical distributions across samples to be the same and assumes global differences in the distribution are induced by only technical variation. It was selected for data transformation to adjust the distribution of values in a dataset to make them conform to a common distribution. It helps remove systematic biases and ensures that the distribution of expression values is consistent across samples, making it easier to compare and analyze the data [44,45]. Survival analysis was conducted on normalized data for individual datasets and respective Kaplan-Meier survival curves are shown in supplementary figures. Combined survival analysis of all the datasets was performed on pooled data from respective datasets. Kaplan-Meier estimator, along with the Log-Rank (Mantel-Cox) test for DSS or OS or RFS or MFS, was plotted using the IBM SPSS version 21 software package. The selection criteria for datasets and the corresponding cohort sizes are outlined in Fig. 1A, providing a visual representation of the dataset selection process.

2.7. Meta-analysis

Multivariate Cox-regression analysis was conducted, incorporating Txnrd1 expression and other histo-clinical parameters as potential risk factors. This analysis yielded regression coefficients and Hazard Ratios (HR) with 95% confidence intervals (CI) for time-to-event outcomes (DSS, OS, RFS, MFS). To combine and analyze the hazard ratios from multiple independent studies, meta-analysis was performed using the "meta" package in R software employing the random effects model. Weight assignment was performed by the inverse-variance method. Pooled hazard ratio, z-value and associated significance are reported to quantify the hazard associated with Txnrd1 gene expression. These results are summarized in Table 5.

Between study and within study variance (Tau^2) was calculated using the DerSimonian-Laird estimator to quantify the uncertainty around the estimates. The inverse variance method with the DerSimonian-Laird estimator combines effect sizes, while the Jackson method helps calculate confidence intervals for measures of heterogeneity (τ^2 and τ). Jackson Method for Confidence Interval of Tau^2 and Tau, along with I^2 statistic derived from Cochran's Q statistic, assessed heterogeneity. The I^2 statistic generated from Cochran's Q statistic and degree of freedom in the meta-analysis represents the proportion of total variability in the HR due to heterogeneity rather than sampling error. These values for Tau^2, Tau, I^2, and Q along with associated statistical significance are summarized in Supplementary Table 7. Additionally, Egger's test was performed to assess the extent of publication bias (Supplementary Table 13). Sensitivity analysis was performed for meta-analysis with moderate heterogeneity by excluding the studies with higher weight assignment to assess the impact on pooled effect size, confidence interval and overall statistical significance.

The number of events in $Txnrd1^{High}$ and $Txnrd1^{Low}$ groups in the study cohort (DSS, OS, RFS, MFS) were enumerated, total number of subjects in high and low groups were noted and used as input values for calculating relative risk. Meta-analysis for log-transformed relative risk (lnRR) for the event was performed using the "meta" package in R software and the above-stated input values. Results of pooled log-transformed relative risks along with statistical tests for assessment of heterogeneity were performed as outlined above. The results are presented in Table 5 and Supplementary Table 7, respectively.

2.8. MTT assay

Overnight seeded cells (3000/well in a 96-well plate) treated with vehicle or doxorubicin were cultured for 72 h, followed by the addition of MTT solution (10 mg/ml), incubation for 3 h and lysis by addition of 100 μl DMSO to each well by proper mixing. Absorbance at $\lambda = 570$ nm was recorded, and a cell viability plot was generated.

2.9. ROS measurement

Overnight incubated cells were stained with H_2DCFDA (Ex. 485 nm/Em. 535 nm) followed by treatment with vehicle or doxorubicin (10 μM, 4 h). Cells were washed with 1X PBS and acquired using a Partec Cyflow flowcytometer (GmBH, Germany) or using a Synergy H1 multimode microplate reader (BioTek OR, USA).

2.10. RNA isolation, cDNA synthesis and qRT-PCR

Cells were processed for RNA isolation using the kit from Himedia (Mumbai, India) following the manufacturer's protocol [25], followed by two-step cDNA synthesis and quantitative real-time PCR. The primer list is provided in Supplementary Table 11.

2.11. Transfection

Overnight grown MDA-MB-231 cells (10^5 cells/2 mL resuspended in 24-well plate) were transfected with TrxR1 CRISPR/Cas9 KO

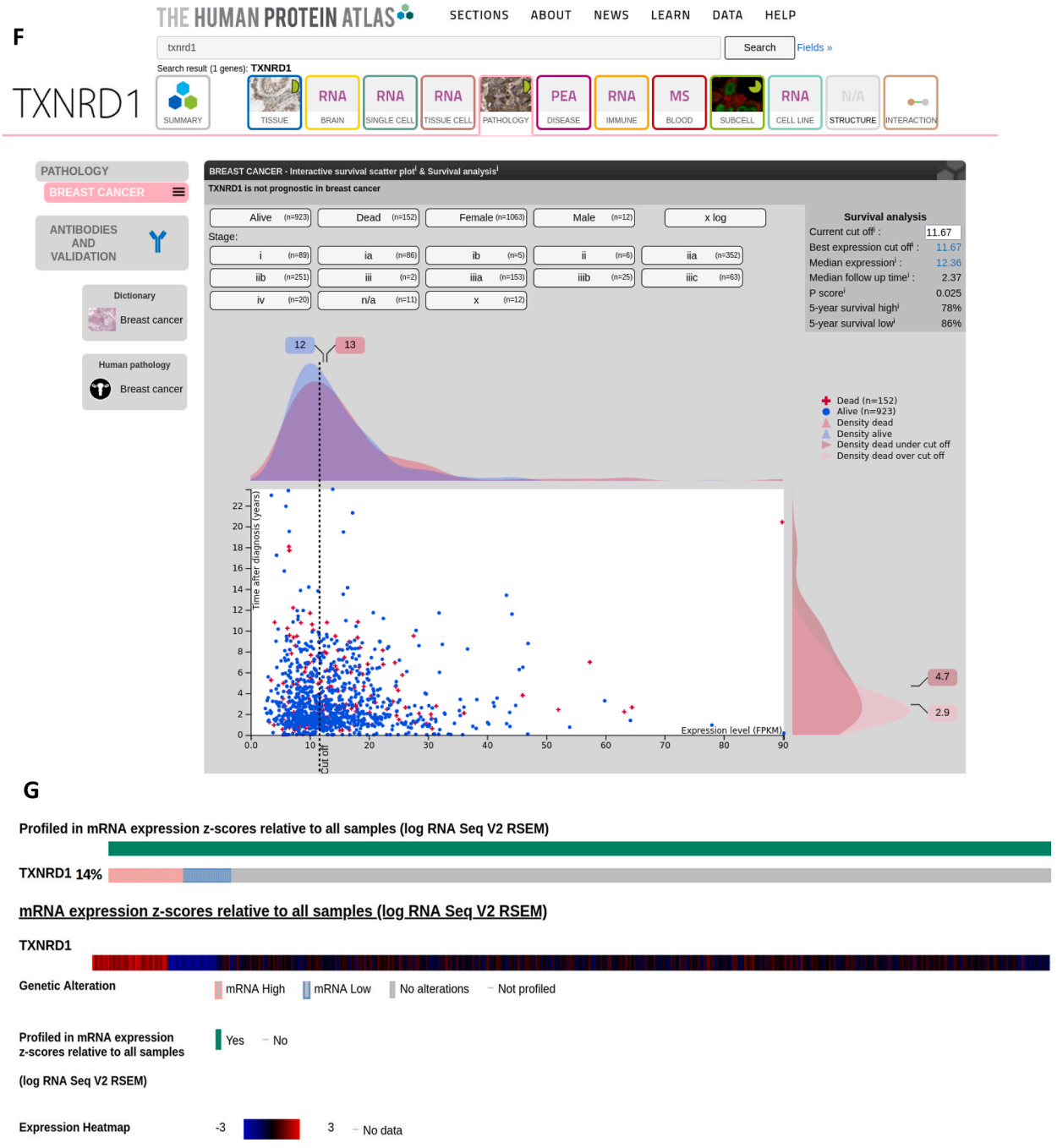


Fig. 1. (*continued*).

Plasmid (#sc-400994-KO-2) using lipofectamine 3000 reagent following manufacturer's protocol. Briefly, plasmid DNA was mixed with transfection reagent and incubated at RT for 10 min. Original media was replaced with DMEM without serum and antibiotic, and DNA:transfection reagent mix was added to the cells. DMEM with 2X serum was added after 16 h, and cells were cultured for 24 h. Transfection was confirmed visually under the microscope with green fluorescence, and knockout was validated by estimating the gene expression, and protein levels of Txnrd1 were quantified by ELISA using a human TrxR1 ELISA kit (Cayman chemicals). Although Txnrd1 mRNA levels were undetectable in transfected MDA-MB-231 cells, a small amount of protein was detected by ELISA. We could achieve 70% efficiency in reducing Txnrd1 protein levels, and hence, we have referred to these cells as Txnrd1 depleted cells.

Table 2
Mean Txnrd1 expression based on hormone receptor and HER2 status.

Dataset	Mean ± SEM Txnrd1 gene expression								
	ER			PR			HER2		
	Negative	Positive	Sig.	Negative	Positive	Sig.	Negative	Positive	Sig.
TCGA	**11.00 ± 0.044** (260)	10.65 ±0.023 (892)	<0.001	**10.96 ± 0.036** (376)	10.62 ±0.026 (773)	<0.001	10.81 ±0.028 (652)	**11.24 ± 0.059** (114)	<0.001
METABRIC	**0.249 ± 0.058** (416)	−0.124 ±0.029 (1160)	<0.001	**0.097 ± 0.049** (709)	−0.154 ±0.033 (805)	<0.001	−0.069 ±0.027 (1325)	**0.19 ± 0.082** (189)	0.001
GSE25066	**10.51 ± 0.032** (205)	10.264 ± 0.037 (297)	<0.001	**10.502 ± 0.032** (258)	10.218 ±0.04 (243)	<0.001	10.365 ±0.027 (485)	10.428 ±0.22 (6)	0.919
Chin 2006	**0.159 ± 0.09** (43)	−0.119 ±0.086 (75)	0.039	**0.101 ± 0.073** (51)	−0.128 ±0.098 (66)	0.051	**NA**		
Hess 2006	**0.111 ± 0.077 (51)**	−0.069 ±0.085 (82)	0.151	**0.098 ± 0.075** (75)	−0.164 ±0.1 (55)	0.043	−0.097 ±0.07 (99)	**0.263 ± 0.11** (33)	0.016
Miller 2005	**0.187 ± 0.053** (34)	−0.065 ±0.022 (213)	<0.001	**0.133 ± 0.043** (61)	−0.085 ±0.022 (190)	<0.001	**NA**		
Yau 2010	**0.261 ± 0.044** (235)	−0.207 ±0.04 (447)	<0.001	**NA**					
vantVeer	**0.027 ± 0.018** (46)	−0.045 ±0.019 (71)	0.014						
Node Negative	**0.209 ± 0.054** (64)	−0.099 ±0.037 (134)	<0.001						
Lymph Node Negative Wang	**0.233 ± 0.047** (77)	−0.055 ±0.035 (209)	<0.001						
GSE3494	**8.182 ± 0.053** (34)	7.93 ±0.022 (213)	<0.001	**8.127 ± 0.043** (61)	7.91 ±0.022 (190)	<0.001	**NA**		
GSE80999	**0.268 ± 0.05** (65)	−0.072 ±0.02 (284)	<0.001	**NA**					
GSE18229	**−1.194 ± 0.054 (92)**	−1.387 ±0.056 (104)	0.051	**−1.231 ± 0.056 (116)**	−1.411 ± 0.05 (79)	0.025	−1.33 ± 0.046 (155)	−1.202 ± 0.07 (40)	0.193
GSE41998	**9.242 ± 0.033 (171)**	9.021 ±0.057 (108)	<0.001	**9.245 ± 0.033 (179)**	8.999 ±0.059 (99)	0.001	9.156 ±0.032 (251)	9.159 ±0.079 (28)	0.975
GSE21653	**8.577 ± 0.0.054** (116)	8.242 ±0.063 (150)	<0.001	**8.584 ± .055** (130)	8.201 ±.064 (136)	<0.001	**NA**		
GSE2990	**9.239 ± 0.087(40)**	8.991 ±0.0408 (85)	0.004	**NA**					
GSE31448	**8.634 ± 0.046 (161)**	8.271 ± 0.053 (188)	<0.001	**8.619 ± 0.046** (177)	8.253 ± 0.055 (172)	<0.001	**NA**		
GSE30682	**8.067 ± 0.052 (75)**	7.882 ±0.027 (268)	0.002	**8.074 ± 0.039** (138)	7.820 ±0.029 (205)	<0.001	7.914 ±0.026 (290)	7.966 ±0.060 (53)	0.45
GSE10886	0.486 ± 0.077 (44)	0.607 ± 0.054 (70)	0.192	0.46 ± 0.081 (43)	0.65 ± 0.075 (41)	0.091	0.548 ± 0.063 (73)	0.58 ± 0.075 (11)	0.851
GSE22219	**9.196 ± 0.051** (82)	8.993 ± 0.038 (134)	0.002	**NA**					
GSE22226	**−1.744 ± 0.113 (70)**	−2.061 ±0.123 (51)	0.065	−1.827 ±0.110 (72)	−1.925 ±0.154 (38)	0.838	**NA**		
GSE25055	**10.533 ± 0.039 (131)**	10.195 ±0.044 (174)	<0.001	**10.517 ± 0.038** (162)	10.137 ±0.046 (142)	<0.001	10.341 ±0.032 (291)	10.51 ±0.258 (4)	0.773
GSE25065	**10.469 ± 0.057 (74)**	10.361 ± 0.064 (123)	<0.01	10.476 ± 0.056 (96)	10.331 ± 0.071 (101)	0.28	10.4 ± 0.046 (194)	10.263 ± 0.55 (2)	0.88
GSE119295	**0.215 ± 0.012 (91)**	0.12 ± 0.003 (672)	<0.001	**0.172 ± 0.007** (210)	0.118 ±0.003 (555)	<0.001	0.128 ±0.003 (711)	**0.197 ± 0.019 (54)**	<0.001
GSE103744	**8.128 ± 0.088 (50)**	7.773 ± 0.044 (122)	<0.001	**NA**			**NA**		
GSE31863	**−1.397 ± 0.107 (29)**	−2.026 ±0.057 (114)	<0.001	**−1.642 ± 0.093** (47)	−2.043 ±0.068 (88)	0.003			
GSE7390	**10.547 ± 0.054 (64)**	10.238 ± 0.037 (134)	<0.001	**NA**					

Mean ± SEM of Txnrd1 expression values is given in front of each dataset for respective ER, PR and HER2 groups. Values in bracket after Mean ± SEM indicates number of patients in that group. Values are given in bold if $p < 0.05$.
NA: Data not available.

Table 3
Mean survival time of breast cancer patients between Txnrd1 expression subgroups.

DSS			
	$Txnrd1^{High}$	$Txnrd1^{Low}$	Sig
METABRIC	13.491 ± 0.735 years (215)	13.951 ± 0.968 years (86)	0.498
Miller 2005	**9.596 ± 0.621 years (62)**	10.954 ± 0.377 years (57)	**0.002**
GSE3494	10.665 ± 0.55 years (58)	10.996 ± 0.472 years (62)	0.761
Node Negative T2	10.493 ± 1.1 years (79)	21.377 ± 1.537 years (42)	0.08
TCGA-BRCA	16.581 ± 1.073 years (303)	17.628 ± 1.335 years (299)	0.238
Chin 2006	8.968 ± 0.713 years (30)	12.533 ± 0.753 years (30)	0.165
Node negative	**9.836 ± 1.003 years (50)**	13.921 ± 0.864 years (50)	**0.02**
RFS			
METABRIC	**12.705 ± 0.622 years (285)**	17.563 ± 0.784 years (283)	**<0.001**
GSE25066	**5.588 ± 0.268 years (127)**	6.729 ±.173 years (127)	**<0.001**
GSE32603	**4.865 ± 0.315 years (62)**	5.838 ± 0.262 years (62)	**0.015**
GSE22219	**6.032 ± 0.485 years (54)**	9.186 ± 0.264 years (54)	**<0.001**
GSE25055	**5.559 ± 0.352 years (78)**	6.859 ± 0.186 years (78)	**0.001**
Wang 2005	**8.151 ± 0.727 years (72)**	10.085 ± 0.481 years (72)	**0.004**
GSE30682	**2.623 ± 0.458 years (13)**	4.359 ± 0.621 years (17)	**0.023**
GSE22226	4.65 ± 0.459 years (33)	4.657 ± 0.408 years (33)	0.527
GSE9893	**3.774 ± 0.22 years (16)**	6.552 ± 0.764 years (13)	**0.04**
Caldas	10.904 ± 0.867 years (26)	10.422 ± 0.84 years (27)	0.532
GSE18229	6.496 ± 0.618 years (41)	5.646 ± 0.289 years (40)	0.172
GSE7390	**9.83 ± 1.002 years (50)**	13.911 ± 0.863 years (50)	**0.02**
DMFS			
GSE158309	**18.261 ± 1.163 years (116)**	22.225 ± 0.718 years (117)	**<0.001**
Vijver 2002	**10.061 ± 0.899 years (74)**	12.823 ± 0.768 years (75)	**0.001**
Yau 2010	**11.307 ± 0.599 years (170)**	16.891 ± 1.233 years (171)	**<0.001**
vantVeer 2002	**6.602 ± 1.065 years (30)**	9.813 ± 0.88 years (30)	**0.019**
GSE11121	14.469 ± 1.236 years (50)	13.61 ± 0.799 years (50)	0.177
Node negative	**12.707 ± 0.989 years (50)**	16.719 ± 0.775 years (50)	**0.033**
GSE9893	**2.932 ± 0.26 years (16)**	5.182 ± 0.696 years (13)	**0.002**
OS			
METABRIC	**12.498 ± 0.66 years (285)**	15.694 ± 0.568 years (283)	**<0.001**
TCGA	12.588 ± 1.131 years (308)	13.741 ± 1.167 years (308)	0.19
GSE124647	**2.53 ± 0.512 years (35)**	4.2 ± 0.552 years (35)	**0.009**
Caldas	9.230 ± 0.908 years (26)	11.058 ± 0.76 years (26)	0.185
GSE22226	5.182 ± 0.365 years (33)	4.932 ± 0.361 years (33)	0.816
GSE9893	8.964 ± 0.887 years (39)	8.049 ± 0.498 years (39)	0.245
Vijver 2002	**10.83 ± 0.824 years (74)**	14.537 ± 0.579 years (75)	**<0.001**
GSE10886	6.198 ± 0.188 years (25)	7.21 ± 0.938 years (26)	0.085
GSE39004	4.965 ± 0.838 years (16)	7.614 ± 1.281 years (16)	0.475
GSE7390	**13.288 ± 0.877 years (50)**	20.254 ± 0.809 years (50)	**0.008**

Mean ± SEM of survival time expressed in years is given in front of each dataset for $Txnrd1^{High}$ and $Txnrd1^{Low}$ subgroups. Values in bracket after Mean ± SEM indicates number of patients in that subgroup. Significance values are given in bold if $p < 0.05$. Values in $Txnrd1^{High}$ sub-group are also given in bold where $p < 0.05$.

2.12. Clonogenic assay

Wild type or Txnrd1 depleted MDA-MB-231 cells (1000 cells per well) grown overnight in a 6-well plate (or 35 mm dish) were treated with vehicle or Doxorubicin (10 nM) or exposed to radiation (4 Gy) and cultured for 10–14 days till colonies appear. Colonies were fixed using formaldehyde, stained with crystal violet, and imaged under Syngene:GBox Gel documentation system. Colonies were counted using ImageJ.

2.13. Wound healing assay

Using a sterile microtip, a scratch was made in a plate containing wild type or Txnrd1 depleted MDA-MB-231 cells grown to confluency. Cells were washed with 1X PBS and overlaid with 2% FBS containing media. Images were acquired (0 h), and then the cells were incubated for 24 h to allow gap closure. after 24 h scratches were re-imaged under microscope and gap closure was estimated using ImageJ.

Table 4
Hazard ratio (HR) for Txnrd1 gene expression calculated by accounting for the clinico-pathological covariates available in the respective datasets along with 95% CI, and significance values is provided for disease-specific survival (DSS), overall survival (OS), recurrence-free survival (RFS) and distant metastasis-free survival (DMFS) are provided. Values for HR and significance are indicated in bold if $p<0.05$.

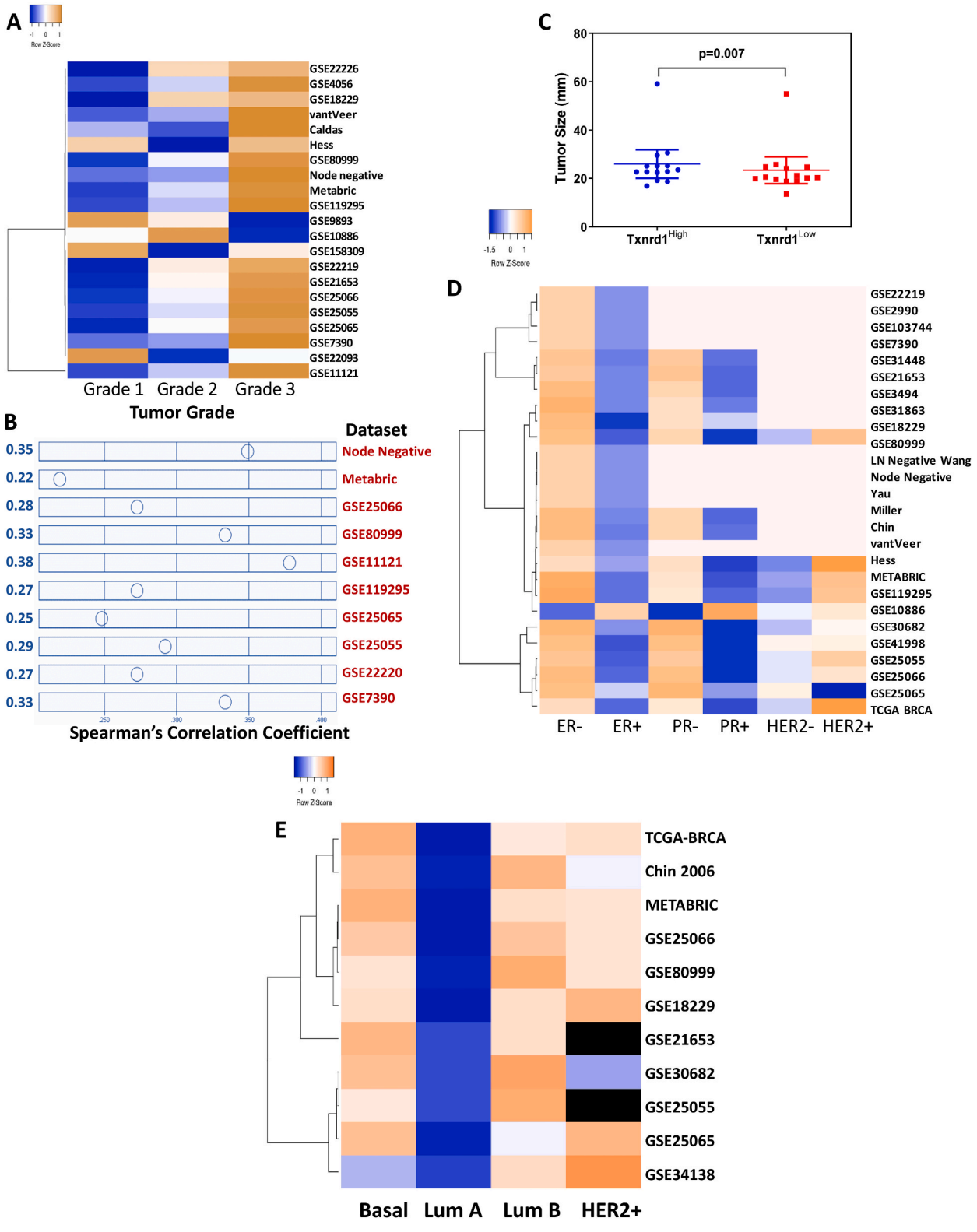
Dataset	Hazard Ratio (HR)	CI (95%)	Significance	Covariates
DSS				
GSE3494	**2.321**	1.066–5.052	**0.034**	ER, PR, age, size, LN
Miller	2.165	0.971–4.825	0.059	Age, ER, PR, LN, p53, size
Chin 2006	1.006	0.481–2.102	0.988	ER, PR, Size, Stage, Subtype
Node Negative	**1.623**	0.999–2.638	**0.05**	Age, Size, ER, stage
OS				
METABRIC	**1.106**	1.036–1.181	**0.002**	ER, PR, HER2, Histology, Menopause
TCGA	1.171	0.325–4.22	0.81	Age, weight, menopause, histology, node, T stage
GSE124647	1.303	0.943–1.8	0.11	PR, therapy, Stage
Caldas	3.674	0.644–20.948	0.143	Age, stage, ER, grade, size
GSE22226	0.836	0.478–1.461	0.529	Age, size, histology, ER, PR, HER2, PAM50, T stage
GSE9893	**1.417**	1.139–1.762	**0.002**	Age, size, grade, histology, T, N
GSE10886	0.003	0.000–2.751	0.096	Age, PR, HER2, node, size
GSE39004	2.242	0.52–9.664	0.28	Age, ER, TNBC, stage, grade, node
GSE7390	1.623	0.88–2.995	0.12	Age, size, histology, grade, ER
RFS				
GSE22219	**1.791**	1.104–2.904	**0.018**	Age, size, node, ER, grade
GSE25055	**1.752**	1.088–2.824	**0.021**	ER, PR, HER2, ESR, ERBB2
GSE25066	**1.788**	1.357–2.356	**<0.001**	Age, grade, pCR, RCB, ERBB2
GSE7390	**1.685**	1.045–2.717	**0.032**	Age, size, grade, ER, histology
GSE22226	0.993	0.622–1.586	0.978	Age, histology, size, T stage, ER, PR, HER2, subtype
GSE32603	1.217	0.813–1.823	0.34	HR, HER2
Wang 2005	**1.819**	1.23–2.691	**0.003**	ER, LN
GSE30682	**1.556**	1.03–2.352	**0.036**	Age, T, N
Caldas 2007	0.764	0.183–3.191	0.764	Age, ER, Grade, Stage, size
GSE18229	1.537	0.788–2.995	0.207	Age, ER, PR, HER2, grade, subtype
DMFS				
GSE158309	**1.798**	1.295–2.497	**<0.001**	Age, node, grade, size, subtype
Yau 2010	**1.182**	1.001–1.395	**0.048**	ER, ERBB, PAM50
Node Negative	**1.623**	0.999–2.638	**0.05**	Age, size, ER, grade, histology
GSE9893	**1.496**	1.197–1.871	**<0.001**	Histology, grade, size, T, N

Table 5
Pooled Hazard Ratio and ln (RR) after meta-analysis. Values for pooled hazard ratio (HR) and log-transformed relative risk ln(RR) , 95% CI along with associated z-value and p-value for disease-specific survival (DSS), overall survival (OS), recurrence-free survival (RFS) and distant metastasis-free survival (DMFS) are provided. Values of pooled HR, ln(RR) and significance are indicated in bold if $p<0.05$.

HR				
	Pooled HR	95%CI	z-value	p-value
DSS	**1.45**	0.92–1.98	5.33	**<0.001**
OS	**1.23**	0.97–1.48	9.36	**<0.001**
RFS	**1.47**	1.25–1.69	13.15	**<0.001**
DMFS	**1.42**	1.14–1.70	16	**<0.001**
ln(RR)				
	Pooled ln(RR)	**95%CI**	**z-value**	**p-value**
DSS	0.26	−0.06–0.58	1.6	0.109
OS	0.28	−0.05–0.60	1.68	0.092
RFS	**0.52**	0.23–0.80	3.53	**<0.001**
DMFS	**0.62**	0.36–0.88	4.63	**<0.001**

2.14. Invasion assay

To test the invasiveness of the cells, a transwell matrigel invasion assay was performed (Corning, Cat# 354,480). MDA-MB-231 cells were seeded at a density of 2×10^4 cells per well in the upper chamber of a 24-well plate with 8 μm pores. The upper chamber contained DMEM without FBS, while the lower chamber of the transwells was filled with DMEM supplemented with 30% FBS. In treated groups, both the upper and lower chamber media contained 1 μM AF. Whereas control groups were treated with DMSO. The upper chamber was rinsed with 1X PBS after 19 h, and cells remaining at the top of the polycarbonate membrane were removed by scrapping with the cotton swab. The cells that had invaded through pores to the lower surface of the membrane were fixed with ethanol and stained with 0.5% crystal violet solution. Images were acquired at 10× magnification.



(caption on next page)

Fig. 2. Higher grade and large size tumors exhibit over-expression of Txnrd1
The expression heatmap of Txnrd1 across tumor grades from different datasets with row z-scores and clustering is shown in (A). Spearman's rank correlation coefficient between Txnrd1 expression and tumor grade is shown in (B). The mean tumor size in Txnrd1 High and Low expression subgroups is shown in (C). Expression heatmap of Txnrd1 across ER, PR & HER2 expression status along with row z-scores and clustering is shown in (D), and molecular subtypes is shown in (E).

2.15. Statistical analysis

Estimation of differences in gene expression among different groups, calculation of correlation coefficient, survival curve comparison, hazard ratio and relative risk estimation followed by meta-analysis was performed as stated above. Statistical analysis was performed using IBM SPSS ver 21 software or GraphPad Prism software version 9 (La Jolla, USA). The significance between the mean of two groups was compared by *t*-test and for more than two groups One-way ANOVA was used. Log-rank (Mantel-Cox) test was used for the comparison of survival curves. The difference between groups is significant if $p < 0.05$. Data are presented as mean ± S.E.M.

3. Results

3.1. Txnrd1 is not over-expressed at mRNA or protein level in breast primary tumor

Overall workflow adopted for meta-analysis is outlined in Fig. 1A. Txnrd1 is under transcriptional control of the Nrf2-ARE pathway. Pathway analysis using cBioportal revealed alterations in the oxidative stress response pathway in 17.2% of TCGA-Firehose patients (7% Keap1, 5.6% Cul3, and 4.6% Nfe2l2) (Fig. 1B) and 13.8% of Metabric patients (5.4% Keap1, 4.2% Cul3, and 4.2% Nfe2l2) (Fig. 1C). Limma_voom analysis in TCGA-BRCA shows a $\log_2$ fold change of 0.143, average expression 6.477, and adjusted p-value of 0.059 for Txnrd1 in breast cancer vs. normal tissue. The B-statistic suggests the odds of differential expression as 0.00154, indicating a 0.15% probability that Txnrd1 is differentially expressed in primary breast tumors. Txnrd1 gene expression does not vary significantly between normal tissue and breast primary or metastatic tumors in TCGA data (Fig. 1D). Protein level z-scores (mass spectrometry) from CPTAC data revealed no significant difference in Txnrd1 protein expression between normal tissue and breast cancer primary tumors (Fig. 1E). Further, Txnrd1 is not considered to have prognostic significance for breast cancer according to the human protein Atlas (Fig. 1F).

Oncoprint analysis from TCGA-Firehose legacy and Metabric revealed that 14% of breast cancer patients harbor high Txnrd1 mRNA levels (exp>1) (Fig. 1G). The expression heatmap shows mRNA expression z-scores (−3 to 3) relative to all samples among these 14% of patients, indicating wide variation in Txnrd1 expression among breast cancer patients. While Txnrd1 is not differentially expressed between normal and breast tumor tissues, its expression shows considerable variation among breast cancer patients.

3.2. Txnrd1 expression exhibited positive correlation with tumor grade and size

Across 21 datasets, Txnrd1 gene expression varied significantly ($p < 0.05$) according to tumor grade in 15 datasets, with 13 of them having p-values ≤0.01 (Supplementary Table 1). The pattern observed in 15 datasets indicated that Txnrd1 expression followed the pattern: grade 3 > grade 2 > grade 1 (Fig. 2A). Positive correlation between Txnrd1 gene expression and tumor grade was found in 10 datasets at $p < 0.01$, as indicated by Spearman's correlation coefficient (Fig. 2B & Supplementary Table 2), with correlation coefficients consistently >0.2.

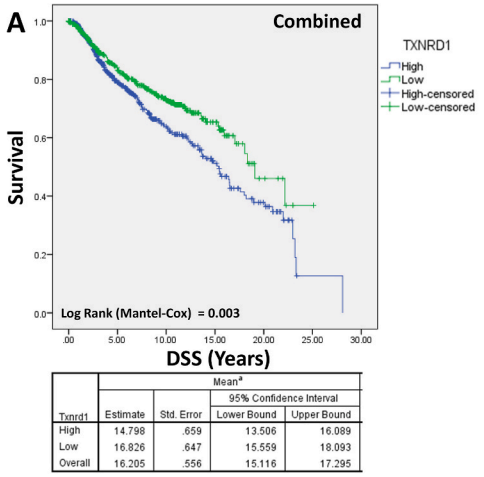
In 13 datasets, tumors with high Txnrd1 expression had a larger size than those with low expression, with significance observed in 5 datasets (Fig. 2C & Supplementary Table 3). Spearman's correlation coefficient analysis in 5 datasets showed a significant positive correlation ($p < 0.05$) between Txnrd1 gene expression and tumor size (Supplementary Table 2). The average tumor size was significantly larger in $Txnrd1^{High}$ expression patients (25.99 ± 2.63 mm) compared to $Txnrd1^{Low}$ expression cohort (23.40 ± 2.479 mm) with a mean size difference of 2.59 ± 0.8 mm (95% CI, 0.849–4.331) at $p < 0.007$ indicating larger size tumors in $Txnrd1^{High}$ expression patients.

In majority of datasets, no significant variation in Txnrd1 expression was observed across different pathologic T stages or pathologic N stages (Supplementary Tables 4 and 5). There was no significant difference in Txnrd1 gene expression across histological subtypes or menopause status (data not shown).

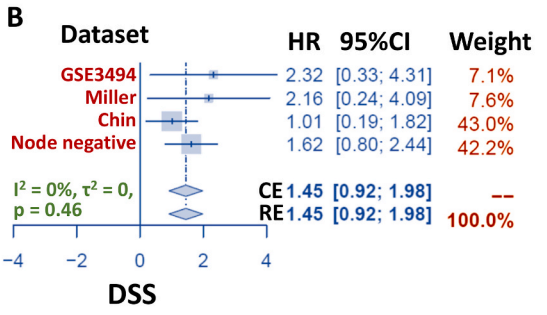
3.3. Txnrd1 is over-expressed in hormone receptor-negative and HER2-positive breast cancer

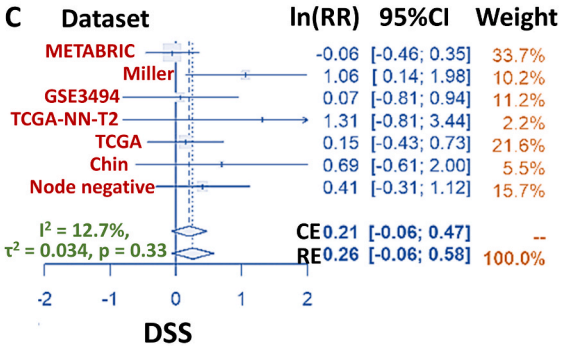
In a comprehensive analysis across 27 datasets, Txnrd1 gene expression was consistently higher ($p < 0.05$) in ER-negative breast tumors compared to ER-positive patients (Fig. 2D & Table 2). Similarly, across 18 datasets, Txnrd1 expression was significantly elevated ($p < 0.05$) in PR-negative breast cancer patients compared to PR-positive counterparts. In 11 datasets, Txnrd1 gene expression was significantly higher in HER2-positive breast tumors compared to HER2-negative patients ($p < 0.05$). Overall, Txnrd1 expression followed the pattern: ER-negative > ER-positive, PR-negative > PR-positive, and HER2-positive > HER2-negative (Fig. 2D & Table 2). Overall, Txnrd1 showed a differential expression pattern with respect to breast cancer subtypes. Interestingly, it is overexpressed in hormone-receptor-negative and HER2-positive breast cancers.

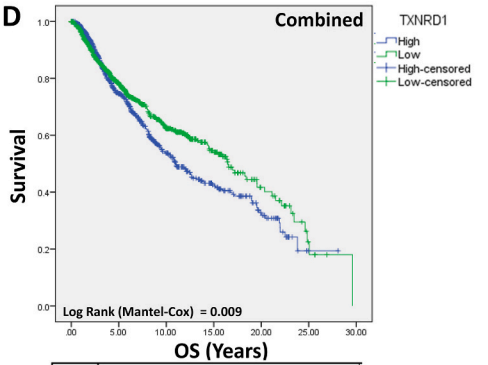
Analyzing Txnrd1 gene expression across PAM50 molecular subtypes in 11 datasets revealed that mean expression varied



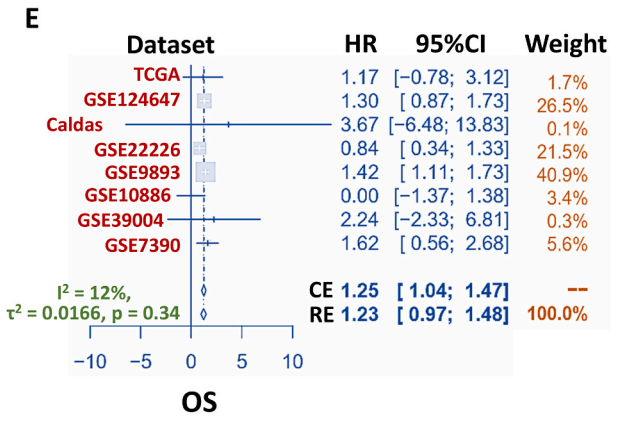
Txnrd1	Mean[a] Estimate	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
High	14.798	.659	13.506	16.089
Low	16.826	.647	15.559	18.093
Overall	16.205	.556	15.116	17.295

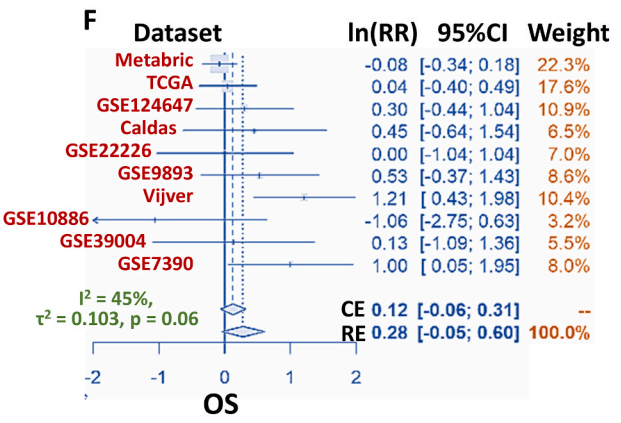






Txnrd1	Mean[a] Estimate	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
High	13.819	.532	12.777	14.861
Low	15.856	.607	14.665	17.047
Overall	14.955	.429	14.114	15.796





(caption on next page)

Fig. 3. Txnrd1 over-expression is associated with an increased hazard of shorter disease-specific survival and overall survival
Kaplan-Meier survival curve for breast cancer-specific survival generated by pooling the quantile normalized data from all the datasets along with the Log-Rank (Mantel-Cox) test p-value is shown in (A). The cumulative mean survival time between high (n = 722) and low (n = 714) expression subgroups is shown below the survival curve. Multivariate Cox-regression analysis considering various clinicopathological covariates yielded different values for hazard ratio in independent datasets. These results were combined in a meta-analysis performed employing random effects model, and the results of the pooled hazard ratio, 95%CI, weight assignment and heterogeneity assessment are shown in (B). By enumerating the number of events (death due to disease) in high vs low expression cohort in view of the total cohort size, meta-analysis for log-transformed relative risk was performed. Forest plot of overall log-transformed relative risk, 95%CI, weight assignment and heterogeneity are shown in (C). Kaplan-Meier survival curve for overall survival generated by pooling the quantile normalized data from all the datasets along with Log-Rank (Mantel-Cox) test p-value for $Txnrd1^{High}$ (n = 888) and $Txnrd1^{Low}$ (n = 886) is shown in (D). The pooled hazard ratio for overall survival, along with the statistical significance and heterogeneity parameters, is shown in (E). Similarly, a forest plot with results of overall log-transformed relative risk for overall survival is shown in (F). Associated statistical significance parameters and values for (B, C, E & F) are shown in Table 5. CI, confidence interval; CE, common effects model; RE, random effects model, HR, hazard ratio; ln(RR), log-transformed relative risk.

significantly ($p < 0.05$) in all datasets (Supplementary Table 6). Consistent with hormone receptor and HER2 status, luminal A (ER + PR + HER2-) breast cancer subtype exhibited the least Txnrd1 expression across datasets. In 3 out of 11 datasets, HER2+ subtypes, and in another 3 datasets, basal (ER-PR-HER2-) subtype exhibited the highest Txnrd1 gene expression. Interestingly, in the remaining 5 datasets, luminal B (ER + PR ± HER2+) subtype showed the highest Txnrd1 gene expression. Overall, 8 HER2+ subtypes and 9 HR-subtypes exhibited the highest Txnrd1 gene expression, consistent with the overarching pattern observed (Fig. 2E & Supplementary Table 6).

3.4. Txnrd1 over-expression indicated a poor prognosis for breast cancer-specific survival

In 2 out of 7 datasets, $Txnrd1^{High}$ expression subgroup of patients exhibited significantly shorter ($p \leq 0.01$) breast cancer-specific survival as compared to $Txnrd1^{Low}$ breast cancer patients (S1A-G & Table 3). A combined survival analysis performed after pooling data from all seven datasets, demonstrated that $Txnrd1^{High}$ patients (n = 722) exhibited a mean breast cancer-specific survival of 14.79 ± 0.66 years, while $Txnrd1^{Low}$ subgroup (n = 714) had a longer survival of 16.83 ± 0.65 years (Log-Rank test, p = 0.003) (Fig. 3A). Thus, $Txnrd1^{High}$ patients are more likely to succumb to disease related death almost 2 years earlier than $Txnrd1^{Low}$ patients. Multivariate Cox-regression analysis, considering covariates such as ER, PR, p53, age, tumor size, lymph node status, molecular subtype, menopause status, and histology, revealed hazard ratios (HR) of 2.32 in GSE 3494 (p = 0.03), 2.16 in Miller 2005 (p = 0.06), 1.006 in Chin 2006 (p = 0.988), and 1.623 in Node negative (p = 0.05), indicating a significant association between Txnrd1 over-expression and decreased breast cancer-specific survival (Table 4). Meta-analysis of hazard ratio employing random effects model revealed a pooled hazard ratio of 1.45 (95% CI, 0.92; 1.98), z = 5.33, at $p < 0.001$ (Fig. 3B). This indicated a statistically significant hazard associated with Txnrd1 over-expressing breast cancer patients for shorter disease-specific survival. Further, a meta-analysis of log-transformed relative risks, considering the number of events (death due to disease) in $Txnrd1^{High}$ and $Txnrd1^{Low}$ expression cohorts, showed a pooled log-transformed relative risk of 0.26 (95% CI, −0.06; 0.58) at p = 0.109 with a z-value of 1.6 (Fig. 3C). However, the non-significant p-value and the proximity of the result to the null value indicated a non-significant risk for death due to disease events in Txnrd1 over-expressing patients. A pairwise comparison of the cumulative mean disease-specific survival between $Txnrd1^{High}$ and $Txnrd1^{Low}$ cohort revealed no statistical significance in this cross-study comparison (S1H). This implies that although $Txnrd1^{High}$ patients are statistically more likely to experience early mortality there are other confounding factors that are also important determinants of disease-specific survival which is explained in detail in the discussion section.

Significant differences in overall survival among breast cancer patients based on Txnrd1 gene expression were observed in 4 out of 10 datasets (S2A-J & Table 3). Survival distribution differences were studied by the combined survival analysis conducted after pooling data from all ten datasets. Breast cancer patients with Txnrd1 over-expression exhibited significantly shorter overall survival (13.819 ± 0.532 years) compared to those with low Txnrd1 expression (15.856 ± 0.607 years) (Log-Rank test, p = 0.009) (Fig. 3D) indicating $Txnrd1^{High}$ group of patients are more likely to succumb to death almost 2 years ealier than $Txnrd1^{Low}$ patients.

Multivariate Cox-regression analysis, accounting for covariates such as ER, PR, HER2, histology, menopause status, age, tumor size, and grade, showed varying hazard ratios (HR) across different datasets. Notably, Metabric (HR = 1.106, p = 0.002) and GSE9893 (HR = 1.417, p = 0.002) demonstrated significant associations, while other datasets exhibited mixed results (Table 4). Meta-analysis of HR across all datasets revealed a pooled HR of 1.23 (95% CI, 0.97; 1.48), z = 9.36 at $p < 0.001$ (Fig. 3E), indicating a statistically significant hazard associated with Txnrd1 over-expression for shorter overall survival. Associated I^2, Tau^2 statistic along with Chochran's Q statistic, and heterogeneity factor are summarized in Supplementary Table 7. These metrics indicated very low heterogeneity and between study variance attributing the results of meta-analysis to the true effect size.

Additionally, a meta-analysis of log-transformed relative risks, considering the number of events (deaths) in $Txnrd1^{High}$ and $Txnrd1^{Low}$ expression cohorts, yielded a pooled log-transformed relative risk of 0.28 (95% CI, −0.05; 0.60) at p = 0.09 with a z-value of 1.68 (Fig. 3F). The non-significant p-value and the proximity of the result to the null value indicated a non-significant risk for death in patients with Txnrd1 over-expression. Heterogeneity assessment revealed moderate between study variance and cross-study heterogeneity. Sensitivity analysis performed by excluding the studies with relatively higher weight assignments revealed moderate changes in the overall heterogeneity. Egger's test result illustrated no strong evidence of publication bias (Supplementary Table 13), indicating robustness in the meta-analysis findings. All these findings suggest the statistical significance and the stability of the observed effects across various studies. Overall this analysis implies that although there is a strong likelihood of early mortality in $Txnrd1^{High}$ breast

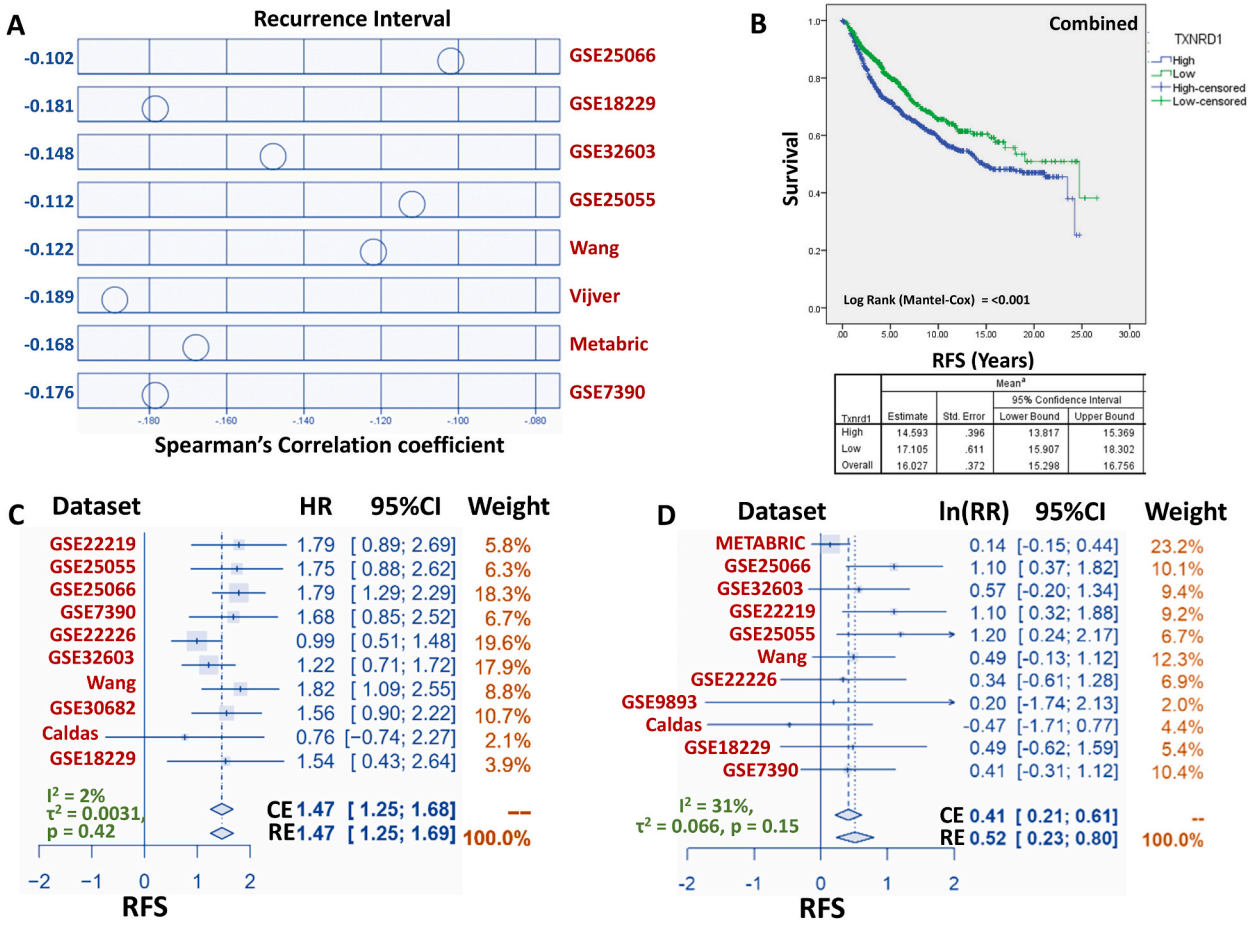


Fig. 4. Txnrd1 overexpression is associated with an increased risk of early recurrence
Spearman's correlation coefficient for the correlation between Txnrd1 gene expression and recurrence interval in various datasets is shown in (A). Kaplan-Meier survival curve for recurrence-free survival generated by pooling the data for all the datasets after quantile normalization along with Log-Rank (Mantel-Cox) test p-value is shown in (B). The cumulative mean survival for $Txnrd1^{High}$ (n = 903) and $Txnrd1^{Low}$ (n = 807) cohorts is shown below the survival curve. The hazard ratio for RFS obtained by multivariate analysis for datasets included in this study was examined collectively by performing meta-analysis. A forest plot showing the pooled hazard ratio, 95% CI, weight assignment and heterogeneity is shown in (C). Meta-analysis of relative risk for recurrence for individual datasets is shown in (D). Associated statistical significance parameters and values for (C & D) are shown in Table 5. CI, confidence interval; CE, common effects model; RE, random effects model.

cancer patients, the risk is not associated with only Txnrd1 expression status, which is conceivable considering the role of other clinico-pathological factors in determining overall survival.

Txnrd1 over-expressing breast cancer patients are at high risk of early recurrence and metastasis.

In 10 out of 11 datasets, Txnrd1 expression is higher in breast cancer patients with recurrence than in patients with no recurrence event (Supplementary Table 8A). In 7 out of 11 datasets, Txnrd1 expression is significantly higher (p, <0.05) in breast cancer patients with recurrence of the disease as compared to those with no recurrence event (Supplementary Table 8A). Conversely, in one dataset (GSE10886) $Txnrd1^{Low}$ expression group has significantly higher recurrence (Supplementary Table 8A). Interestingly, in 8 of these 11 datasets, Txnrd1 expression exhibited a significant negative correlation with recurrence interval at p, 0.05 level (2-tailed) (Fig. 4A & Supplementary Table 2).

In 9 out of 12 datasets, $Txnrd1^{High}$ breast cancer patients exhibited significantly shorter recurrence-free survival ($p < 0.05$) compared to $Txnrd1^{Low}$ sub-group (S3A-I & Table 3). Combined survival analysis, after pooling data from all 12 datasets, showed that $Txnrd1^{High}$ patients experience recurrence after 14.593 ± 0.39 years, while $Txnrd1^{Low}$ breast cancer patients have a longer recurrence-free period of 17.1 ± 0.61 years (Fig. 4B). Multivariate Cox-regression analysis, considering various covariates, revealed HR > 1 at $p < 0.05$ in 6 out of 10 datasets. Meta-analysis of HR, using random-effects model, resulted in a pooled hazard ratio of 1.47 (95% CI, 1.25; 1.69), $z = 13.15$ at $p < 0.001$) (Fig. 4C). Meta-analysis of log-transformed relative risks (lnRR) revealed a pooled log-transformed relative risk of 0.52 (95% CI, 0.23; 0.8), $z = 3.53$ at $p < 0.001$ (Fig. 4D). Test of heterogeneity indicated low to moderate heterogeneity in the studies included in the meta-analysis (Supplementary Table 7). These analyses demonstrated a statistically significant hazard associated with Txnrd1 over-expression, indicating an increased risk of early recurrence in breast cancer patients.

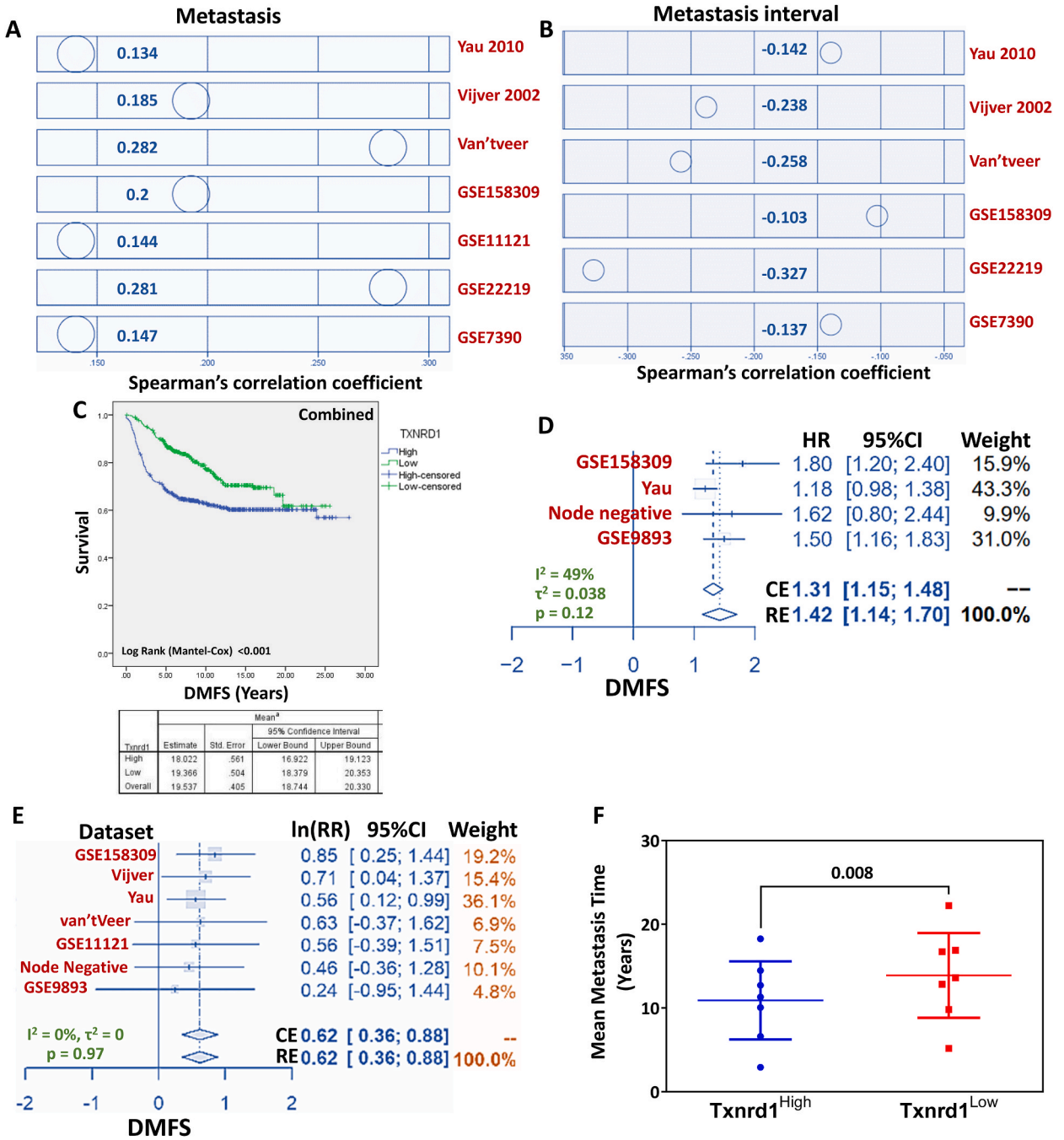
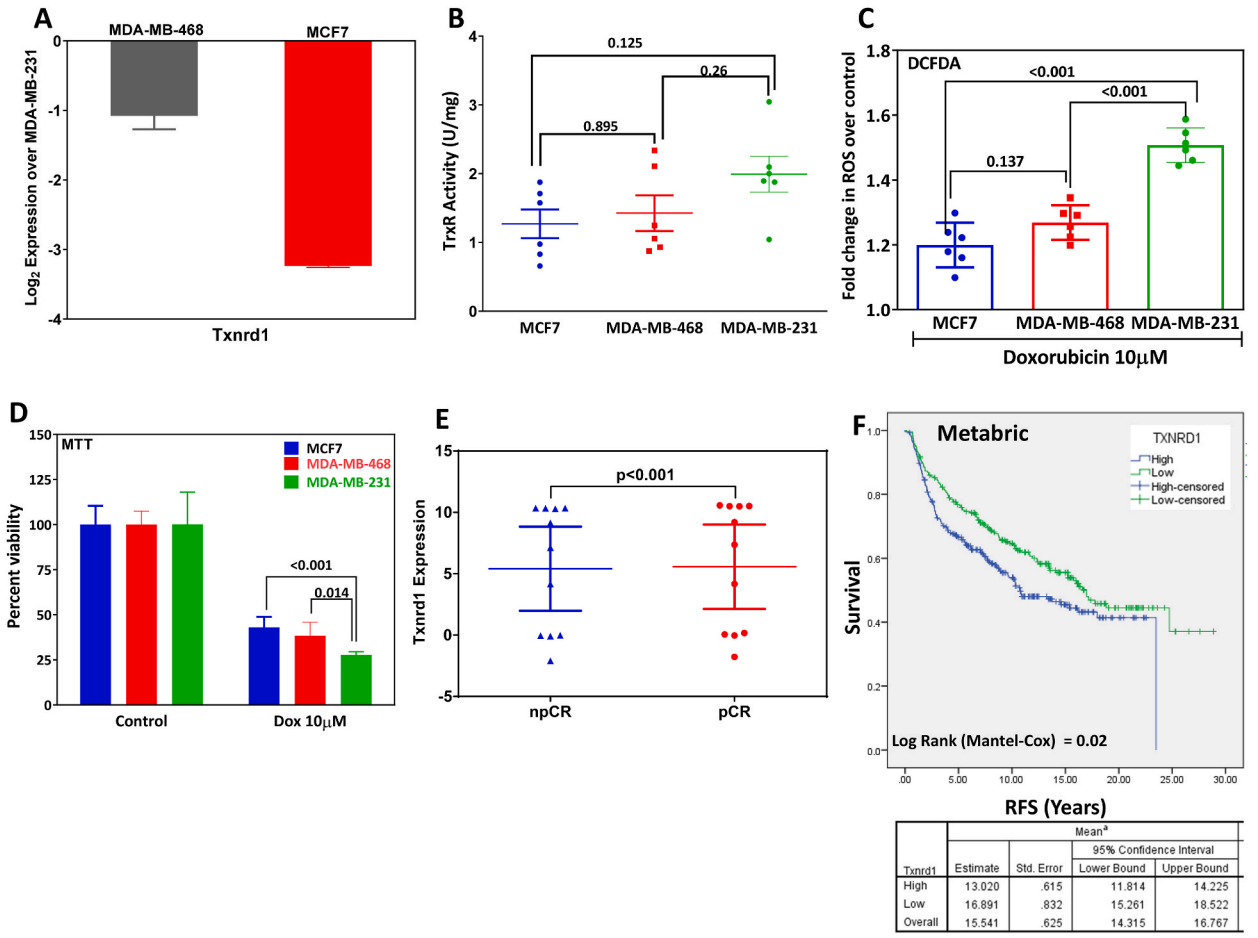


Fig. 5. Txnrd1 overexpressing breast cancer patients are at elevated risk for early metastasis
Spearman's correlation coefficient for correlation between Txnrd1 gene expression and metastasis and metastasis interval for different datasets is shown in (A) & (B) respectively. Kaplan-Meier survival curve for metastasis-free survival by combining the data from all the datasets after quantile normalization along with the Log-Rank (Mantel-Cox) test p-value is shown in (C). The cumulative mean survival in $Txnrd1^{High}$ (n = 527) and $Txnrd1^{Low}$ (n = 517) cohort is shown below the survival curve. Meta-analysis for hazard ratio and relative risk is shown in (D) & (E), respectively. Associated statistical significance parameters and values for (D & E) are shown in Table 5. Survival distribution differences between $Txnrd1^{High}$ and $Txnrd1^{Low}$ groups in terms of mean time to metastasis in individual datasets is shown in (F). Statistical significance is calculated by paired *t*-test. CI, confidence interval; CE, common effects model; RE, random effects model.



Txnrd1	Mean[a] Estimate	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
High	13.020	.615	11.814	14.225
Low	16.891	.832	15.261	18.522
Overall	15.541	.625	14.315	16.767

Fig. 6. Txnrd1 overexpression is associated with better response to chemotherapy and resistance to radiotherapy
qRT-PCR showing Log_2 expression of Txnrd1 in MCF7 and MDA-MB-468 cells relative to MDA-MB-231 is shown in (A). The bar graph shows baseline TrxR activity in MCF7, MDA-MB-468 and MDA-MB-231 cells (B). Fold change in ROS over control in MCF7, MDA-MB-468 & MDA-MB-231 cells after doxorubicin treatment is shown in (C). Cell viability of MCF7, MDA-MB-468 & MDA-MB-231 cells after doxorubicin treatment as assessed by MTT assay is shown in (D). Data represents mean ± S.E.M. from six replicates pooled from two independent experiments for A-D. Mean Txnrd1 expression between breast cancer patients with or without pCR is shown in (E). Kaplan-Meier survival curve for recurrence-free survival of breast cancer patients after radiotherapy between Txnrd1 high or low expression is shown in (F). Significance is calculated by the Log-Rank (Matel-Cox) test, and the cumulative mean survival of $Txnrd1^{High}$ (n = 198) and $Txnrd1^{Low}$ (n = 166) cohorts is shown below the survival curve. Statistical significance for B & C is calculated by One way ANOVA, and Student's t-test was used for D & E.

In 6 out of 7 datasets, Txnrd1 expression is significantly higher ($p < 0.05$) in breast cancer patients with metastasis compared to non-metastatic individuals (Supplementary Table 7B). Positive correlation with metastasis is observed in 7 datasets, and negative correlation with metastasis interval is noted in 6 datasets based on Spearman's correlation coefficient at $p < 0.05$ level (2-tailed) (Fig. 5A and B).

In 6 out of 7 datasets, $Txnrd1^{High}$ breast cancer patients exhibit significantly shorter metastasis-free survival ($p < 0.05$) compared to $Txnrd1^{Low}$ sub-group (S4A-F & Table 3). Combined survival analysis of 7 datasets illustrated that $Txnrd1^{High}$ patients experience a metastasis-free interval of 18.02 ± 0.56 years, while $Txnrd1^{Low}$ cohort records a time to metastasis of 19.36 ± 0.50 years (Fig. 5C). Multivariate Cox-regression analysis, with age, tumor size, node, grade, subtype, ER as covariates, revealed HR = 1.80 in GSE158309 ($p < 0.001$), 1.5 in GSE9893 ($p < 0.001$), 1.62 in Node Negative ($p = 0.05$), and 1.18 in Yau 2010 ($p < 0.048$) (Table 4). Meta-analysis of HR revealed a pooled HR of 1.42 (95% CI, 1.14; 1.70), z = 16, $p < 0.001$ (Fig. 5D) with moderate heterogeneity (Supplementary Table 7). Meta-analysis of log-transformed relative risks, considering the number of events (metastasis) in $Txnrd1^{High}$ and $Txnrd1^{Low}$ sub-groups, identified a pooled log-transformed relative risk of 0.62 (95% CI, 0.36; 0.88), z = 4.63 at $p < 0.001$ (Fig. 5E). Paired comparison of the cumulative mean metastasis time between $Txnrd1^{High}$ and $Txnrd1^{Low}$ groups across 7 datasets revealed statistically significant difference. The bar graph, along with the statistical significance, is shown in (Fig. 5F). These findings collectively demonstrate a consistent and statistically significant association between Txnrd1 over-expression and an increased risk of early metastasis in breast cancer patients.

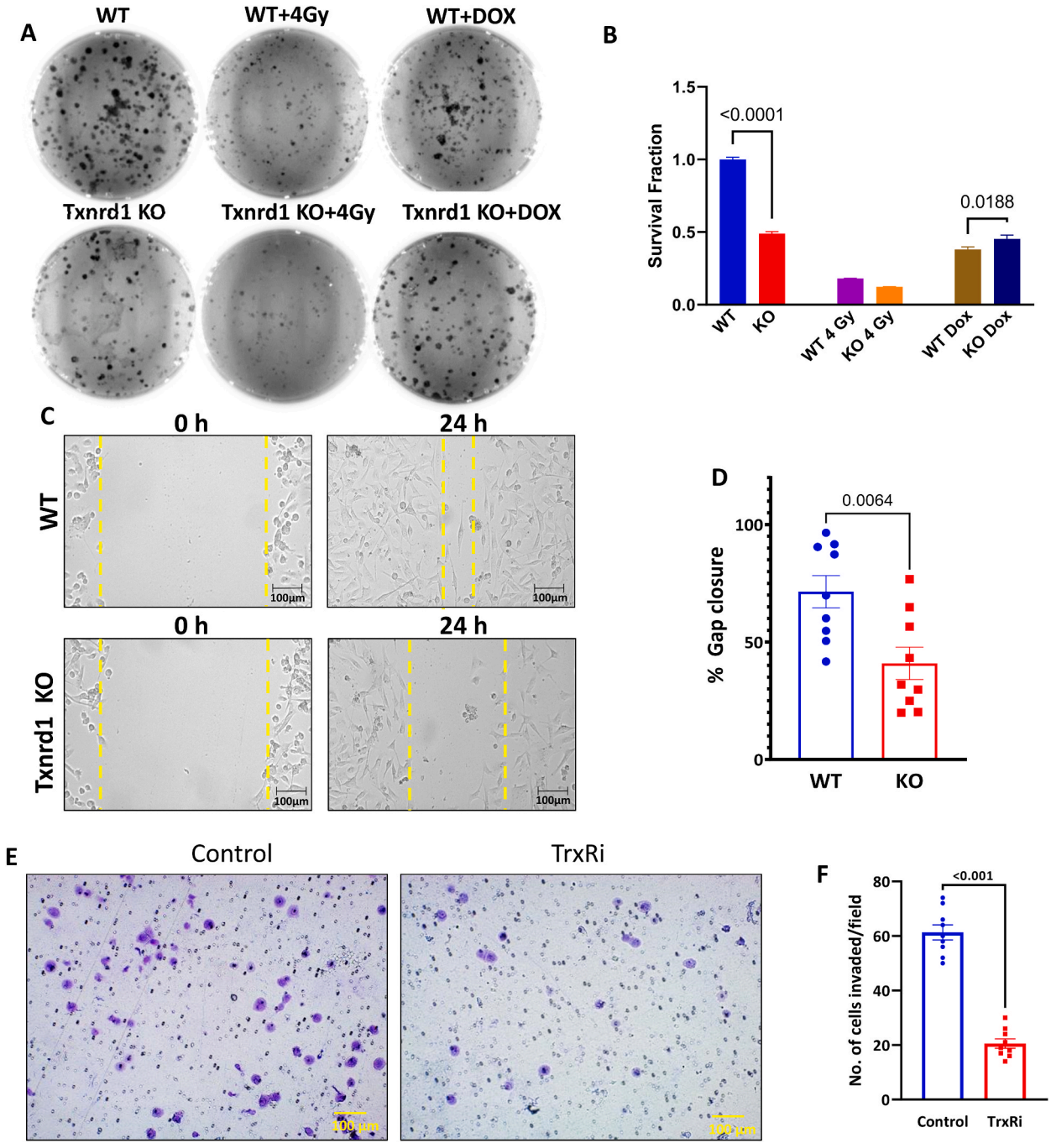


Fig. 7. Multifaceted role of Txnrd1 in regulating survival, treatment sensitivity, motility, and invasion in MDA-MB-231 cells
Representative images of clonogenic assay performed with wild type and Txnrd1 depleted MDA-MB-231 cells treated with or without Doxorubicin or radiation 4 Gy (A). The bar graph shows survival fraction (B). Data represents mean ± S.E.M. from three replicates and two such independent experiments were carried out. Representative images of wound healing scratch assay at 0 h and 24 h for wild type and Txnrd1 depleted cells is shown in (C). The bar graph shows the percent gap closure (D). Data represents mean ± S.E.M. from nine replicates, and two such independent experiments were carried out. Representative images of transwell Matrigel invasion assay performed with MDA-MB-231 cells treated with or without pharmacologic inhibitor of TrxR (TrxRi i.e. AF, 1 μM) are shown in (E). The bar graph shows number of cells invaded per field (F). Data represents mean ± S.E.M from nine replicates, and two such independent experiments were carried out. Statistical significance was calculated using Student's t-test for B, D & F.

3.5. Txnrd1 over-expressing tumors exhibit pCR after neoadjuvant chemotherapy

In 4 out of 10 datasets, patients achieving pCR after taxane-anthracycline neoadjuvant chemotherapy exhibited significantly elevated Txnrd1 gene expression at $p < 0.05$ compared to the non-pCR group (Supplementary Table 9). Across all 10 datasets, patients with pCR consistently showed higher Txnrd1 gene expression compared to those with residual cancer burden (Supplementary Table 9). In 7 out of 10 datasets, patients with Txnrd1 expression lower than the median value were associated with poor response to chemotherapy (npCR). Numerically, in all datasets, patients with pCR consistently exhibited Txnrd1 expression higher than the median Txnrd1 expression value (Supplementary Table 9).

Analysis of Txnrd1 expression in two datasets encompassing over 50 breast cancer cell lines demonstrated that MDA-MB-231 cells (basal) exhibited Txnrd1 gene expression higher than the third quartile (Q3), while MCF7 cells (luminal A) showed Txnrd1 expression lower than the (Q1) first quartile (S5A & B & Supplementary Table 10), whereas MDA-MB-468 cells showed Txnrd1 mRNA expression higher than MCF7 but lower than MDA-MB-231 cells. Consistent with RNAseq data, MCF7 cells exhibited $\log_2$ three-fold lower Txnrd1 mRNA expression, and MDA-MB-468 cells displayed $\log_2$ one-fold lower expression over MDA-MB-231 cells (Fig. 6A). TrxR activity assessment revealed the highest TrxR activity in MDA-MB-231 followed by MDA-MB-468 and MCF7 cells (Fig. 6B). Interestingly, ROS levels followed the similar trend with maximum increase in Txnrd1 over-expressing MDA-MB-231 cells followed by MDA-MB-468 and MCF7 cells after doxorubicin treatment (Fig. 6C). Excess ROS correlated with significantly reduced cell viability of MDA-MB-231 cells after doxorubicin treatment as compared to MDA-MB-468 or MCF7 cells (Fig. 6D). Expression data illustrated those patients achieving pCR exhibited significantly elevated Txnrd1 expression compared to the non-pCR group (Fig. 6E). These findings establish a consistent correlation between Txnrd1 overexpression and the achievement of pCR after neoadjuvant therapy.

3.6. Txnrd1 over-expression indicated a poor prognosis for recurrence after radiotherapy

Although radiotherapy exhibits superior tumor control and better QoL with longer event-free survival, not all patients benefit from radiotherapy due to the activation of radioresistance post-radiotherapy. Analysis of the gene expression profile induced by exposure to radiation in TNBC MDA-MB-231 cells revealed significant upregulation of Txnrd1 in cells exposed to a 9 Gy dose of radiation (Supplementary Table 12) [46]. In Metabric dataset, $Txnrd1^{High}$ breast cancer patients undergoing radiotherapy ($n = 292$) exhibited a mean recurrence interval of 13.02 ± 0.61 (95% CI-11.81; 14.22) years as compared to $Txnrd1^{Low}$ cohort ($n = 244$) which showed mean recurrence interval of 16.89 ± 0.23 (95% CI - 15.26; 18.52) years with log-rank (Mantel-Cox) p-value, 0.019 (Fig. 6F). Thus, Txnrd1 over-expressing breast cancer patients who have undergone radiotherapy exhibited significantly shorter recurrence intervals as compared to those with low expression of Txnrd1 (Fig. 6F). ROS are principal mediators of radiation injury and hence we sought to estimate cellular ROS levels and viability in both the cell lines after exposure to different doses of radiation alone and in combination with doxorubicin. MDA-MB-231 cells harboured elevated ROS levels when exposed to radiation alone or combined with doxorubicin as compared to MCF7 cells (S5C). However, except doxorubicin +4 Gy combination, the decrease in cell viability of MDA-MB-231 and MCF7 cells was comparable (S5D).

3.7. Txnrd1 regulates in vitro cellular traits associated with survival, recurrence, and metastasis in response to chemotherapy & radiotherapy

Our analyses have identified the positive correlation between Txnrd1 gene expression and tumor grade and size, which is associated with proliferation and clonogenic potential in vitro. Corroborating with the findings from gene expression data, our in vitro studies revealed that MDA-MB-231 cells exhibited elevated Txnrd1 gene expression and activity compared to other breast cancer cells. Moreover, MDA-MB-231 cells also displayed higher ROS levels and enhanced sensitivity to Doxorubicin treatment. To further confirm whether high Txnrd1 expression indeed contributes to the effects observed in $Txnrd1^{High}$ cohort and to establish a causal relationship, we performed validation experiments employing depletion of Txnrd1 in Txnrd1 over-expressing MDA-MB-231 cells. Transfection with TrxR1 CRISPR Cas9 knockout plasmid led to a significant reduction in TrxR1 protein levels, confirming the significant depletion in Txnrd1 protein levels (S5E). Txnrd1 depleted MDA-MB-231 cells exhibited significantly reduced colony-forming ability compared to wild type cells (Fig. 7A and B). Exposure to 4 Gy radiation substantially decreased survival fraction in both wild type and Txnrd1 depleted cells. However, Txnrd1 depleted cells showed a marginal decrease in clonogenic potential compared to wild-type cells, suggesting a potential role of Txnrd1 in response to radiation. Consistent with our findings about response to neoadjuvant chemotherapy in $Txnrd1^{High}$ cohort, Txnrd1 depleted cells displayed limited sensitivity to doxorubicin compared to wild type cells. Wild type cells yielded a significantly lower survival fraction than Txnrd1 depleted cells after doxorubicin treatment (Fig. 7A and B).

Since enhanced motility and invasion are the prerequisites for metastatic spread of the disease, to correlate with tumor metastasis, we performed a motility and invasion assay. In the scratch-wound healing assay, Txnrd1 depleted cells showed significantly less gap closure as compared to wild type cells (Fig. 7C and D), suggesting a significant role of Txnrd1 in regulating the motile behaviour of the cells. Further, we performed a transwell matrigel invasion assay. Here, we used auranofin (1 μM) for pharmacologic inhibition of TrxR. We observed that TrxR inhibition significantly attenuated the invasion potential of MDA-MB-231 cells compared to vehicle treated control (Fig. 7E and F). Collectively, these in vitro functional assays substantiated our interpretations from public gene expression data analysis.

4. Discussion

Nrf2 is a cytoprotective transcription factor, which, besides being the master regulator of antioxidant defence, also governs the metabolism of xenobiotics, bioenergetics, metabolic pathways, unfolded protein response, proteostasis, mitobiogenesis, autophagy, DNA damage repair and cellular survival [47]. TrxR is one of the Nrf2-dependent genes which reduces disulfide bonds on proteins involved in cell survival, proliferation, apoptosis, DNA synthesis, cell cycle, antioxidant response and DNA binding. Although the Nrf2-ARE oxidative stress response pathway is altered in >13% of breast cancer patients, Txnrd1 is not differentially expressed between normal tissue and breast tumor. Hence, it is not considered to have prognostic value. However, among breast cancer patients there is wide variation in Txnrd1 mRNA expression. Therefore, there is a need to investigate the underlying factors for observed variation, which may unravel novel associations between Txnrd1 gene expression and breast tumor characteristics. Indeed, our analysis revealed that Txnrd1 gene expression varied significantly according to tumor grades, and followed the pattern grade 3 > grade 2 > grade 1. Cross-study comparison revealed that, Txnrd1 gene expression exhibited a significant positive correlation with tumor grade. Additionally, the mean tumor size in $Txnrd1^{High}$ cohort is ~2.6 mm larger than $Txnrd1^{Low}$ cohort, regardless of patient and tumor characteristics. Moreover, Txnrd1 gene expression exhibited a significant positive correlation with tumor size in many datasets. This underscores the association of Txnrd1 over-expression with tumor grade and size, which was not evident by the non-significant difference of Txnrd1 gene expression between normal and tumor tissue. It is pertinent to note that about 64% of patients have local-stage breast cancer, 27% with regional stage, and 6% have distant stage cancer at the time of diagnosis [48]. Aggressive, high-grade tumors are actively dividing and rely on the supply of dNTPs to meet the tumor requirement. Ribonucleotide reductase is the only enzyme that can convert rNTPs into dNTPs, which is dependent on Txnrd1 for disulfide reduction. Hence, it is anticipated that dividing tumor cells have an increased requirement of Txnrd1 to match their proliferation rate. Further, corroborating this, large tumors also exhibited significantly elevated Txnrd1 expression. Concurrently, Txnrd1 depletion in $Txnrd1^{High}$ MDA-MB-231 cells significantly reduced clonogenic potential under basal conditions (Fig. 7A and B). This underscored the role of Txnrd1 in regulating cell survival and proliferation, potentially linking to advanced tumor grade and large size observed in $Txnrd1^{High}$ aggressive tumors, indicating its prognostic value for locally advanced breast cancers.

It is noteworthy that, lymph node involvement, T stage and tumor histology are governed by the multifactorial interplay between tumor and patient body and factors other than Txnrd1 may be influencing the pathophysiological progression of breast cancer. Hence, we did not observe any correlation between menopause status, histological subtypes or pathologic T and N stage with Txnrd1 expression, indicating breast cancer progression independent of Txnrd1 gene expression across these parameters.

PAM50 molecular subtypes of breast cancer are determined by gene expression signature analysis, which is expensive and not a standard clinical practice. However, clinical evaluation of biomarkers viz., ER, PR & HER2 is used to determine four major subtypes as luminal A (ER + PR + HER2-), luminal B (ER + PR ± HER2+), basal (ER-PR-HER2-) and HER2 enriched (ER-PR-HER2+) with five-year relative survival rates as 92%, 89%, 77% and 83% respectively. Luminal A tumors are the most common, slower-growing, less aggressive and exhibit a favourable prognosis as they are responsive to endocrine therapy. Whereas hormone receptor-negative tumors are higher grade, fast-growing, aggressive and exhibit poor prognosis. Interestingly, ER-negative and PR-negative breast tumors exhibited significantly increased Txnrd1 gene expression, respectively, compared to their respective hormone receptor-positive counterpart. While HER2-positive breast cancer patients from 11 independent cohorts exhibited significantly elevated levels of Txnrd1. The overall trend suggested that HR-HER2+ breast cancer patients, indicative of aggressive breast cancers, consistently exhibited Txnrd1 over-expression. While the luminal A subtype displayed the lowest Txnrd1 gene expression, TNBC, HER2+, and luminal B subtypes exhibited over-expression of Txnrd1. Consistent with this pattern, our in vitro studies showed elevated Txnrd1 gene expression in MDA-MB-231 (TNBC) than MCF7 (luminal) (Fig. 6A), and TrxR activity followed a similar trend (Fig. 6B). This pattern underscores the association of Txnrd1 over-expression with breast cancer subtypes featured by aggressive biology and poor prognosis.

Breast cancer-specific five-year survival rates depend on the stage of diagnosis and stand at 77% and 90% in TNBC and other breast cancers, respectively, for all SEER (The Surveillance, Epidemiology, and End Results Program) stages combined (2011–2017). Combined survival analysis of 7 datasets illustrated that $Txnrd1^{High}$ breast cancer patients succumbed to disease-related mortality ~2.03 years earlier than $Txnrd1^{Low}$ sub-group. Meta-analysis of hazard ratio from 4 datasets highlighted statistically significant hazard in $Txnrd1^{High}$ cohort for shorter disease-specific survival. Although survival analysis was conducted for 7 datasets, meta-analysis could include only 4 datasets due to the non-availability of covariates for calculating hazard ratio in other datasets. Given the limited number of available datasets and skewed weight distribution in meta-analysis, the results need to be interpreted with caution. However, very low heterogeneity and between-study variance signify the robustness of these findings. Meta-analysis of log-transformed relative risks assigned no significant risk for disease-related mortality events in $Txnrd1^{High}$ cohort. Results obtained from meta-analysis of hazard ratio and relative risks are counterintuitive making it challenging to draw confirmatory inferences. Thus, it can be concluded that Txnrd1 over-expression poses a significant hazard for early disease-related mortality. However, Txnrd1 over-expression cannot be considered as a risk factor in determining the death due to disease outcome.

Interestingly, $Txnrd1^{High}$ cohort of patients recorded 2.04 years shorter overall survival as compared to $Txnrd1^{Low}$ cohort. Supporting these findings, meta-analysis identified a statistically significant hazard in Txnrd1 over-expressing breast cancer patients for early mortality. However, meta-analysis of relative risks identified negligible mortality risk due to Txnrd1 over-expression. Sensitivity analysis by excluding studies with more weight influenced the precision of the outcome highlighting the potential involvement of heterogeneity and bias. This is further validated by the moderate heterogeneity present in the meta-analysis. Overall, $Txnrd1^{High}$ breast cancer patients experienced significantly lower overall survival but it could not be classified as a risk factor for mortality. Thus, Txnrd1 cannot be considered as a risk factor in predicting disease-specific or overall survival. These results are reasonable and understandable from the clinical perspective given the heterogeneity of patient and tumor characteristics, treatment modality and standard-of-care

determine both disease-specific and overall survival.

Txnrd1 over-expression has a significant positive correlation with disease recurrence and a negative correlation with recurrence interval. Txnrd1^{High} patients experienced ~2.5 years earlier disease recurrence than Txnrd1^{Low} group. Interestingly, findings from meta-analysis coincided with combined survival analysis illustrating a statistically significant hazard of early recurrence in Txnrd1 over-expressing breast cancer patients. The robustness of these findings can be visualized from very low associated heterogeneity and between-study variance. Moreover, meta-analysis of log-transformed relative risks identified a statistically significant positive risk of early recurrence in Txnrd1^{High} breast cancer patients. Despite moderate heterogeneity, sensitivity analysis and Egger's test revealed no impact on the precision or potential publication bias. Collectively, it can be stated that Txnrd1 can act as a prognosticator for disease recurrence in breast cancer patients.

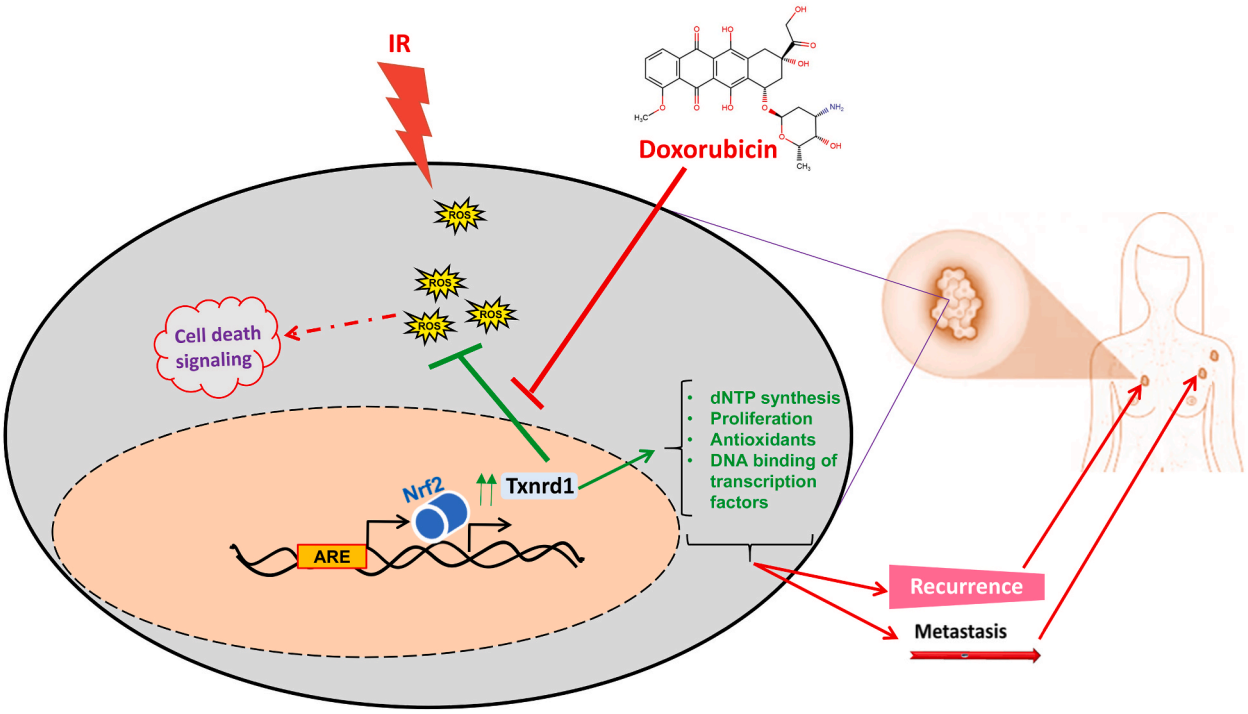
Thioredoxin system is important for the invasion and metastasis of breast cancer cells and exhibits a poor prognosis for distant metastasis in breast cancer patients Txnrd1 overexpression has a significant positive correlation with metastasis event and a negative correlation with metastasis interval. Txnrd1^{High} cohort exhibited ~1.3 years early distant metastasis compared to Txnrd1^{Low} group. Txnrd1^{High} breast cancer patients experienced statistically significant hazard for the shorter metastasis-free period and an elevated risk for early metastasis events. The availability of limited number of datasets resulted in moderate heterogeneity for pooled hazard ratio in meta-analysis. However, there is no heterogeneity in identifying overall relative risk. Validation experiments revealed that Txnrd1 depletion in Txnrd1^{High} MDA-MB-231 cells abrogated the motility compared to wild type cells (Fig. 7C and D). Concomitantly, pharmacologic inhibition of Txnrd1 hampered the invasiveness of MDA-MB-231 cells (Fig. 7E and F), supporting our findings about metastatic behaviour in Txnrd1^{High} cohort. Together the findings from survival and meta-analysis of DSS, OS, RFS & MFS indicated that Txnrd1 over-expression is a risk factor for early recurrence or distant metastasis leading to a significant hazard for shorter recurrence-free or metastasis-free survival. However, elevated expression of Txnrd1, although poses a significant hazard for mortality, it is not a risk factor for the same. In other words, Txnrd1 over-expressing patients are at elevated risk of local or distant disease recurrence, which enhances the probability of death. However, several other contributing factors will ultimately determine the risk of mortality.

There lie inherent challenges in collating the data from diverse sources due to different methodologies employed and reliance on publicly available data. Differences in experimental design, platforms, and data preprocessing methods across studies can introduce technical variability. Differences in patient demographics, treatment histories, and clinical variables may impact gene expression patterns. Additionally, publication bias and selective reporting may lead to overestimation of the actual effect. Here, we have attempted to mitigate these issues by employing multiple strategies to overcome these limitations. Further, we have also performed validation experiments to account for the biological complexity and variability.

Neoadjuvant chemotherapy is used as a preoperative therapy to shrink the tumor mass, facilitating easy and less extensive surgical removal of tumor. It is primarily used in HER2+ and TNBC cancers, and it is shown to completely clear all the clinical evidence of cancer. Summary analysis revealed that it is as effective as adjuvant therapy regarding survival and distant recurrence. RCB scale is used to predict the outcome of neoadjuvant chemotherapy, with RCB 0 indicating pCR and RCB III as residual disease (RD). Taxane-anthracycline neoadjuvant chemotherapy is widely adopted due to its clinical benefit in improved pathological complete response rate and recurrence-free survival. It is essential to develop genomic predictors for response and survival after neoadjuvant chemotherapy due to cases of LABC, which failed to respond to neoadjuvant taxane-anthracycline chemotherapy. Interestingly, breast cancer patients exhibiting pCR harboured elevated Txnrd1 than npCR group in all the datasets. Patients that exhibited npCR harboured lower than median Txnrd1 mRNA levels in the respective cohort. Although the analysis revealed that Txnrd1 over-expression is associated with faster recurrence and poor prognosis for metastasis-free survival, counterintuitively, it is associated with improved pCR for taxane-anthracycline neoadjuvant chemotherapy.

Thioredoxin is known to be responsible for the redox cycling of anthracyclines, thereby generating superoxide and increased DNA damage [49,50]. Breast cancer cells with thioredoxin overexpression exhibit enhanced ROS production and apoptosis upon anthracycline treatment. Thus, thioredoxin is known to sensitize cancer cells to anthracycline toxicity [51]. In line with these findings, it is imperative to understand that Txnrd1 over-expressing breast cancer patients exhibit better pCR after taxane-anthracycline neoadjuvant chemotherapy probably due to enhanced cancer cell killing by excess ROS generation. Corroborating with our meta-analysis, MDA-MB-231 cells exhibited elevated Txnrd1 gene expression, activity and lower viability after doxorubicin treatment (Fig. 6D). Moreover, Txnrd1 depleted MDA-MB-231 cells showed limited sensitivity to doxorubicin as compared to wild type cells (Fig. 7A and B) confirming the causal role of Txnrd1 in determining response to neoadjuvant chemotherapy.

Radiation therapy is used as a post-operative therapy to reduce the risk of recurrence after surgery. Radiation therapy reduces the 10-year risk of recurrence by 50% and the 15-year risk of mortality by 20% [52]. Concurrent chemoradiotherapy is used in high risk, advanced stage cancer patients with or without surgery for better therapeutic benefit [53,54]. Meta-analysis revealed that Txnrd1 over-expressing breast cancer patients treated with radiotherapy exhibited early recurrence of disease by almost 2.8 years as compared to a low-expression cohort. It was observed that Txnrd1^{High} expressing MDA-MB-231 cells harboured elevated ROS levels after exposure to ionizing radiation alone or in combination with doxorubicin in comparison to MCF7 cells (S5C). Additionally, Txnrd1 depleted MDA-MB-231 were more sensitive to radiation induced killing (Fig. 7A and B) establishing the role of Txnrd1 in response to radiotherapy. Further, although doxorubicin alone decreased the viability of MDA-MB-231 cells significantly as compared to MCF7 cells, there was no synergistic decrement in cell viability in doxorubicin + radiation combination treatment in either cell type (S5D). Although further experimental proof is required it may be presumed that elevated oxidative burden in Txnrd1^{High} MDA-MB-231 cells may describe the early recurrence of disease in Txnrd1^{High} expression subgroup of breast cancer patients. Earlier, we have shown that Txnrd1 is over-expressed in radioresistant human lung cancer cells [20], and is an attractive target for radiosensitization of breast cancer [32]. Chlorophyllin, a semisynthetic chlorophyll derivative, has been shown to induce cancer cell killing by targeting



Scheme 1. Txnrd1 over-expression is associated with recurrence, metastasis and therapy response in breast cancer.

mammalian thioredoxin reductase [55]. We have also demonstrated that chlorophyllin could enhance the radiation-induced killing of human breast cancer cells in vitro and in xenograft tumor-bearing SCID mice [56]. Further, pharmacologic inhibition of TrxR using IQ9 enhanced radio-sensitivity of Txnrd1 high expressing MDA-MB-231 but not MCF-7 cells [57]. Another study reports that auranofin, a well-known TrxR inhibitor, exhibited potent radiosensitization of murine mammary carcinoma 4T1 cells [58]. These findings coupled with our meta-analysis warrant further preclinical studies using drug (TrxR inhibitor) -radiotherapy combinations in breast cancer models.

5. Conclusions

Thus, to summarise Txnrd1 over-expression is associated with higher grade, hormone receptor negative, HER2 positive, larger tumors of aggressive subtype and exhibit significant hazard for shorter breast cancer-specific and overall survival. Txnrd1 over-expressing breast cancer patients are at elevated risk of early recurrence and metastasis. Further, Txnrd1^{High}-expression cohort of breast cancer patients exhibit a pathologic complete response to neoadjuvant chemotherapy but are poor responders to radiotherapy (Scheme 1). Hence, in Txnrd1 over-expressing breast cancer patients, radiotherapy can be combined with taxane-anthracycline chemotherapy for better therapeutic outcomes.

Studies in humans

Authors have used data from public gene expression repositories.

Use of generative AI and AI-assisted technologies

Authors have not used any generative AI or AI-assisted technologies during writing of this manuscript.

Availability of data

The data that support the findings of this study are openly available in The Cancer Genome Atlas (http://cancergenome.nih.gov/), cBioPortal (http://www.cbioportal.org/), UCSC Xena (https://xena.ucsc.edu/), UALCAN (http://ualcan.path.uab.edu/analysis-prot.html), GEO ((https://www.ncbi.nlm.nih.gov/geo/) and Human Protein Atlas (https://www.proteinatlas.org/ENSG00000198431-TXNRD1/pathology/breast+cancer). The complete list of the source of data is provided in Table 1. The experimental data that support the findings of this study are available on request from the corresponding author with prior permission of Department.

CRediT authorship contribution statement

Raghavendra S. Patwardhan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Archita Rai:** Methodology, Investigation. **Deepak Sharma:** Writing – review & editing, Visualization, Resources. **Santosh K. Sandur:** Writing – review & editing, Resources, Supervision, Fundimg acquisition. **Sejal Patwardhan:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Investigation, Funding acquisition, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by Basic and Translational Research in Cancer, grant no. (1/3(6)2020/TMC/R&D-II/8823) from Department of Atomic Energy, Government of India, India.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e27011.

References

[1] M.C. Cheang, M. Martin, T.O. Nielsen, A. Prat, D. Voduc, A. Rodriguez-Lescure, et al., Defining breast cancer intrinsic subtypes by quantitative receptor expression, Oncol. 20 (2015) 474–482.
[2] R.A.-O. Siegel, K.A.-O. Miller, H.E. Fuchs, A. Jemal, Cancer Statistics, 2022.
[3] D.A. Berry, K.A. Cronin, S.K. Plevritis, D.G. Fryback, L. Clarke, M. Zelen, et al., Effect of screening and adjuvant therapy on mortality from breast cancer, N. Engl. J. Med. 353 (2005) 1784–1792.
[4] N. Breen, J.F. Gentleman, J.S. Schiller, Update on mammography trends: comparisons of rates in 2000, 2005, and 2008, Cancer 117 (2011) 2209–2218.
[5] P.M. Ravdin, K.A. Cronin, N. Howlader, C.D. Berg, R.T. Chlebowski, E.J. Feuer, et al., The decrease in breast-cancer incidence in 2003 in the United States, N. Engl. J. Med. 356 (2007) 1670–1674.
[6] R.M. Pfeiffer, Y. Webb-Vargas, W. Wheeler, M.H. Gail, Proportion of US trends in breast cancer incidence attributable to long-term changes in risk factor distributions, Cancer Epidemiology and Prevention Biomarkers 27 (2018) 1214–1222.
[7] K.D. Miller, L. Nogueira, A.B. Mariotto, J.H. Rowland, K.R. Yabroff, C.M. Alfano, et al., Cancer treatment and survivorship statistics, CA: a cancer journal for clinicians 2019 69 (2019) 363–385.
[8] R.L. Costa, W.J. Gradishar, Triple-negative breast cancer: current practice and future directions, Journal of oncology practice 13 (2017) 301–303.
[9] A. Witteveen, A.B. Kwast, G.S. Sonke, M.J. Ijzerman, S. Siesling, Survival after locoregional recurrence or second primary breast cancer: impact of the disease-free interval, PLoS One 10 (2015) e0120832.
[10] M. Clarke, A. Coates, Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer', Lancet Oncol. 19 (2018) 27–39.
[11] B. Holleczek, C. Stegmaier, J.C. Radosa, E.-F. Solomayer, H. Brenner, Risk of loco-regional recurrence and distant metastases of patients with invasive breast cancer up to ten years after diagnosis–results from a registry-based study from Germany, BMC Cancer 19 (2019) 1–14.
[12] A. Wallgren, M. Bonetti, R. Gelber, A. Goldhirsch, M. Castiglione-Gertsch, S. Holmberg, et al., Risk factors for locoregional recurrence among breast cancer patients: results from International Breast Cancer Study Group Trials I through VII, J. Clin. Oncol. 21 (2003) 1205–1213.
[13] M.A. Ferreira, E.R. Gamazon, F. Al-Ejeh, K. Aittomäki, I.L. Andrulis, H. Anton-Culver, et al., Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer, Nat. Commun. 10 (2019) 1–18.
[14] J.A. Sparano, R.J. Gray, D.F. Makower, K.I. Pritchard, K.S. Albain, D.F. Hayes, et al., Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer, N. Engl. J. Med. 379 (2018) 111–121.
[15] P. Cortazar, L. Zhang, M. Untch, K. Mehta, J.P. Costantino, N. Wolmark, et al., Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis, Lancet 384 (2014) 164–172.
[16] C. Speers, L.J. Pierce, Postoperative radiotherapy after breast-conserving surgery for early-stage breast cancer: a review, JAMA Oncol. 2 (2016) 1075–1082.
[17] S.A. Castaneda, J. Strasser, Updates in the treatment of breast cancer with radiotherapy, Surgical Oncology Clinics 26 (2017) 371–382.
[18] J. Poleszczuk, K. Luddy, L. Chen, J.K. Lee, L.B. Harrison, B.J. Czerniecki, et al., Neoadjuvant radiotherapy of early-stage breast cancer and long-term disease-free survival, Breast Cancer Res. 19 (2017) 1–7.
[19] P. McGale, C. Correa, D. Cutter, F. Duane, M. Ewertz, R. Gray, et al., Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials, Lancet 383 (2014) 2127–2135.
[20] B. Singh, R.S. Patwardhan, S. Jayakumar, D. Sharma, S.K. Sandur, Oxidative stress associated metabolic adaptations regulate radioresistance in human lung cancer cells, J. Photochem. Photobiol. B Biol. 213 (2020) 112080.
[21] Z. Tian, J. Tang, X. Liao, Q. Yang, Y. Wu, G. Wu, Identification of a 9-gene prognostic signature for breast cancer, Cancer Med. 9 (2020) 9471–9484.
[22] J. Sun, X. Chen, Z. Wang, M. Guo, H. Shi, X. Wang, et al., A potential prognostic long non-coding RNA signature to predict metastasis-free survival of breast cancer patients, Sci. Rep. 5 (2015) 1–11.
[23] M. Chanrion, V. Negre, H. Fontaine, N. Salvetat, F. Bibeau, G. Mac Grogan, et al., A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer, Clin. Cancer Res. 14 (2008) 1744–1752.
[24] R. Patwardhan, D. Sharma, R. Checker, M. Thoh, S. Sandur, Spatio-temporal changes in glutathione and thioredoxin redox couples during ionizing radiation-induced oxidative stress regulate tumor radio-resistance, Free Radic. Res. 49 (2015) 1218–1232.
[25] R.S. Patwardhan, R. Checker, D. Sharma, S.K. Sandur, K.B. Sainis, Involvement of ERK-Nrf-2 signaling in ionizing radiation induced cell death in normal and tumor cells, PLoS One 8 (2013) e65929.
[26] H.A. Joshi, R.S. Patwardhan, D. Sharma, S.K. Sandur, P.V. Devarajan, Pre-clinical evaluation of an innovative oral nano-formulation of baicalein for modulation of radiation responses, Int. J. Pharm. 595 (2021) 120181.

[27] J.-J. Jia, W.-S. Geng, Z.-Q. Wang, L. Chen, X.-S. Zeng, The role of thioredoxin system in cancer: strategy for cancer therapy, Cancer Chemother. Pharmacol. 84 (2019) 453–470.
[28] M. Dagnell, E.E. Schmidt, E.S. Arnér, The A to Z of modulated cell patterning by mammalian thioredoxin reductases, Free Radic. Biol. Med. 115 (2018) 484–496.
[29] Q. Xu, J. Zhang, Novel strategies for targeting the thioredoxin system for cancer therapy, Expet Opin. Drug Discov. 17 (2022) 437–442.
[30] F. Mohammadi, A. Soltani, A. Ghahremanloo, H. Javid, S.I. Hashemy, The thioredoxin system and cancer therapy: a review, Cancer Chemother. Pharmacol. 84 (2019) 925–935.
[31] R. Gencheva, E.S. Arnér, Thioredoxin reductase inhibition for cancer therapy, Annu. Rev. Pharmacol. Toxicol. 62 (2022) 177–196.
[32] R.S. Patwardhan, D. Sharma, S.K. Sandur, Thioredoxin reductase: an emerging pharmacologic target for radiosensitization of cancer, Translational Oncology 17 (2022) 101341.
[33] D. Chandrashekar, B. Bashel, S. Balasubramanya, C. Creighton, I. Ponce-Rodriguez, B. Chakravarthi, et al., UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses, Neoplasia, New York, NY, 2017.
[34] D.S. Chandrashekar, S.K. Karthikeyan, P.K. Korla, H. Patel, A.R. Shovon, M. Athar, et al., UALCAN: an update to the integrated cancer data analysis platform, Neoplasia 25 (2022) 18–27.
[35] C. Dong, L. Zhang, R. Sun, J. Liu, H. Yin, X. Li, et al., Role of thioredoxin reductase 1 in dysplastic transformation of human breast epithelial cells triggered by chronic oxidative stress, Sci. Rep. 6 (2016) 1–13.
[36] C. Cadenas, D. Franckenstein, M. Schmidt, M. Gehrmann, M. Hermes, B. Geppert, et al., Role of thioredoxin reductase 1 and thioredoxin interacting protein in prognosis of breast cancer, Breast Cancer Res. 12 (2010) 1–15.
[37] M. Bhatia, K.L. McGrath, G. Di Trapani, P. Charoentong, F. Shah, M.M. King, et al., The thioredoxin system in breast cancer cell invasion and migration, Redox Biol. 8 (2016) 68–78.
[38] Brigham, Hospital Ws, 13 HMSCLPPJKR, 25 GdaBCoMCCJDLA, Ilya IfSBRSKRBBBBRETLJTVZWS, Comprehensive molecular portraits of human breast tumours, Nature 490 (2012) 61–70.
[39] N. Holländer, M. Schumacher, On the problem of using 'optimal'cutpoints in the assessment of quantitative prognostic factors, Oncol. Res. Treat. 24 (2001) 194–199.
[40] D.G. Altman, B. Lausen, W. Sauerbrei, M. Schumacher, Dangers of using "optimal" cutpoints in the evaluation of prognostic factors, JNCI (J. Natl. Cancer Inst.): J. Natl. Cancer Inst. 86 (1994) 829–835.
[41] Y.-F. Kuo, Statistical Methods for Determining Single or Multiple Cutpoints of Risk Factors in Survival Data Analysis, The Ohio State University, 1997.
[42] J. Budczies, F. Klauschen, B.V. Sinn, B. Győrffy, W.D. Schmitt, S. Darb-Esfahani, et al., Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization, PLoS One 7 (2012) e51862.
[43] Y.J. Lee, Y.S. Park, H.W. Lee, T.Y. Park, J.K. Lee, E.Y. Heo, Peripheral lymphocyte count as a surrogate marker of immune checkpoint inhibitor therapy outcomes in patients with non-small-cell lung cancer, Sci. Rep. 12 (2022) 1–8.
[44] B.M. Bolstad, R.A. Irizarry, M. Åstrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, Bioinformatics 19 (2003) 185–193.
[45] D. Amaratunga, J. Cabrera, Outlier resistance, standardization, and modeling issues for DNA microarray data, Statistics in Genetics and in the Environmental Sciences (2001) 17–26.
[46] V. Bravata, F.P. Cammarata, L. Minafra, R. Musso, G. Pucci, M. Spada, et al., Gene expression profiles induced by high-dose ionizing radiation in MDA-MB-231 triple-negative breast cancer cell line, Cancer Genomics Proteomics 16 (2019) 257–266.
[47] F. He, X. Ru, T. Wen, NRF2, a transcription factor for stress response and beyond, Int. J. Mol. Sci. 21 (2020) 4777.
[48] A.C. Society, Breast Cancer Facts & Figures 2019–2020, Am Cancer Soc, 2019, pp. 1–44.
[49] D. Ravi, K.C. Das, Redox-cycling of anthracyclines by thioredoxin system: increased superoxide generation and DNA damage, Cancer Chemother. Pharmacol. 54 (2004) 449–458.
[50] D. Ravi, H. Muniyappa, K.C. Das, Endogenous thioredoxin is required for redox cycling of anthracyclines and p53-dependent apoptosis in cancer cells, J. Biol. Chem. 280 (2005) 40084–40096.
[51] K.C. Das, H. Muniyappa, V. Kundumani-Sridharan, J. Subramani, Thioredoxin decreases anthracycline cardiotoxicity, but sensitizes cancer cell apoptosis, Cardiovasc. Toxicol. 21 (2021) 142–151.
[52] S. Darby, P. McGale, C. Correa, C. Taylor, R. Arriagada, M. Clarke, Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials, Lancet 378 (2011) 1707–1716.
[53] V. Mandilaras, N. Bouganim, J. Spayne, R. Dent, A. Arnaout, J. Boileau, et al., Concurrent chemoradiotherapy for locally advanced breast cancer—time for a new paradigm? Curr. Oncol. 22 (2015) 25–32.
[54] K.S. Rallis, T.H.L. Yau, M. Sideris, Chemoradiotherapy in cancer treatment: rationale and clinical applications, Anticancer Res. 41 (2021) 1–7.
[55] S. Sun, Y. Zhang, W. Xu, Y. Zhang, R. Yang, J. Guo, et al., Chlorophyllin inhibits mammalian thioredoxin reductase 1 and triggers cancer cell death, Antioxidants 10 (2021) 1733.
[56] Sharma D., Sandur S.K., Checker R., Patwardhan R.S., Gota V.P., Sundarraj J., et al., US-10183026-B2 (2019). PubChem [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004-. PubChem Patent Summary for US-10183026-B2; [cited 2024 Mar. 4]. Available from: https://pubchem.ncbi.nlm.nih.gov/patent/US-10183026-B2.
[57] N.A. Abdullah, M. Inman, C.J. Moody, S.J. Storr, S.G. Martin, Cytotoxic and radiosensitising effects of a novel thioredoxin reductase inhibitor in breast cancer, Invest. N. Drugs 39 (2021) 1232–1241.
[58] H. Wang, S. Bouzakoura, S. De Mey, H. Jiang, K. Law, I. Dufait, et al., Auranofin radiosensitizes tumor cells through targeting thioredoxin reductase and resulting overproduction of reactive oxygen species, Oncotarget 8 (2017) 35728.
[59] N. Cancer Genome Atlas, Comprehensive molecular portraits of human breast tumours, Nature 490 (2012) 61–70.
[60] C. Curtis, S.P. Shah, S.F. Chin, G. Turashvili, O.M. Rueda, M.J. Dunning, et al., The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups, Nature 486 (2012) 346–352.
[61] M. Schmidt, D. Bohm, C. von Torne, E. Steiner, A. Puhl, H. Pilch, et al., The humoral immune system has a key prognostic impact in node-negative breast cancer, Cancer Res. 68 (2008) 5405–5413.
[62] Y. Wang, J.G. Klijn, Y. Zhang, A.M. Sieuwerts, M.P. Look, F. Yang, et al., Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer, Lancet 365 (2005) 671–679.
[63] M. Chanrion, V. Negre, H. Fontaine, N. Salvetat, F. Bibeau, G. Mac Grogan, et al., A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer, Clin. Cancer Res. 14 (2008) 1744–1752.
[64] J.S. Parker, M. Mullins, M.C. Cheang, S. Leung, D. Voduc, T. Vickery, et al., Supervised risk predictor of breast cancer based on intrinsic subtypes, J. Clin. Oncol. 27 (2009) 1160–1167.
[65] A. Prat, J.S. Parker, O. Karginova, C. Fan, C. Livasy, J.I. Herschkowitz, et al., Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer, Breast Cancer Res. 12 (2010) R68.
[66] R. Sabatier, P. Finetti, J. Adelaide, A. Guille, J.P. Borg, M. Chaffanet, et al., Down-regulation of ECRG4, a candidate tumor suppressor gene, in human breast cancer, PLoS One 6 (2011) e27656.
[67] F.M. Buffa, C. Camps, L. Winchester, C.E. Snell, H.E. Gee, H. Sheldon, et al., microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer, Cancer Res. 71 (2011) 5635–5645.
[68] L.J. Esserman, D.A. Berry, M.C. Cheang, C. Yau, C.M. Perou, L. Carey, et al., Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the I-SPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657), Breast Cancer Res. Treat. 132 (2012) 1049–1062.

[69] N. Servant, M.A. Bollet, H. Halfwerk, K. Bleakley, B. Kreike, L. Jacob, et al., Search for a gene expression signature of breast cancer local recurrence in young women, Clin. Cancer Res. 18 (2012) 1704–1715.
[70] E. Nimeus-Malmstrom, M. Krogh, P. Malmstrom, C. Strand, I. Fredriksson, P. Karlsson, et al., Gene expression profiling in primary breast cancer distinguishes patients developing local recurrence after breast-conservation surgery, with or without postoperative radiotherapy, Breast Cancer Res. 10 (2008) R34.
[71] P. den Hollander, K. Rawls, A. Tsimelzon, J. Shepherd, A. Mazumdar, J. Hill, et al., Phosphatase PTP4A3 promotes triple-negative breast cancer growth and predicts poor patient survival, Cancer Res. 76 (2016) 1942–1953.
[72] A.V. Pladsen, G. Nilsen, O.M. Rueda, M.R. Aure, O. Borgan, K. Liestol, et al., DNA copy number motifs are strong and independent predictors of survival in breast cancer, Commun. Biol. 3 (2020) 153.
[73] M. Sjostrom, J. Staaf, P. Eden, F. Warnberg, J. Bergh, P. Malmstrom, et al., Identification and validation of single-sample breast cancer radiosensitivity gene expression predictors, Breast Cancer Res. 20 (2018) 64.
[74] J. Tutzauer, M. Sjostrom, E. Holmberg, P. Karlsson, F. Killander, L.M.F. Leeb-Lundberg, et al., Breast cancer hypoxia in relation to prognosis and benefit from radiotherapy after breast-conserving surgery in a large, randomised trial with long-term follow-up, Br. J. Cancer 126 (2022) 1145–1156.
[75] D. Yee, C. Isaacs, D.M. Wolf, C. Yau, P. Haluska, K.V. Giridhar, et al., Ganitumab and metformin plus standard neoadjuvant therapy in stage 2/3 breast cancer, NPJ Breast Cancer 7 (2021) 131.
[76] M. Chapellier, E. Bachelard-Cascales, X. Schmidt, F. Clement, I. Treilleux, E. Delay, et al., Disequilibrium of BMP2 levels in the breast stem cell niche launches epithelial transformation by overamplifying BMPR1B cell response, Stem Cell Rep. 4 (2015) 239–254.
[77] K. Shen, N. Song, Y. Kim, C. Tian, S.D. Rice, M.J. Gabrin, et al., A systematic evaluation of multi-gene predictors for the pathological response of breast cancer patients to chemotherapy, PLoS One 7 (2012) e49529.
[78] C.E. Horak, L. Pusztai, G. Xing, O.C. Trifan, C. Saura, L.M. Tseng, et al., Biomarker analysis of neoadjuvant doxorubicin/cyclophosphamide followed by ixabepilone or Paclitaxel in early-stage breast cancer, Clin. Cancer Res. 19 (2013) 1587–1595.
[79] M.J. Magbanua, D.M. Wolf, C. Yau, S.E. Davis, J. Crothers, A. Au, et al., Serial expression analysis of breast tumors during neoadjuvant chemotherapy reveals changes in cell cycle and immune pathways associated with recurrence and response, Breast Cancer Res. 17 (2015) 73.
[80] C. Hatzis, L. Pusztai, V. Valero, D.J. Booser, L. Esserman, A. Lluch, et al., A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer, JAMA 305 (2011) 1873–1881.
[81] K. Kersten, S.B. Coffelt, M. Hoogstraat, N.J.M. Verstegen, K. Vrijland, M. Ciampricotti, et al., Mammary tumor-derived CCL2 enhances pro-metastatic systemic inflammation through upregulation of IL1beta in tumor-associated macrophages, OncoImmunology 6 (2017) e1334744.
[82] J.J. Zoeller, M.F. Press, L.M. Selfors, J. Dering, D.J. Slamon, S.A. Hurvitz, et al., Clinical evaluation of BCL-2/XL levels pre- and post- HER2-targeted therapy, PLoS One 16 (2021) e0251163.
[83] O. Thuerigen, A. Schneeweiss, G. Toedt, P. Warnat, M. Hahn, H. Kramer, et al., Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer, J. Clin. Oncol. 24 (2006) 1839–1845.
[84] A.S. Heimes, F. Hartner, K. Almstedt, S. Krajnak, A. Lebrecht, M.J. Battista, et al., Prognostic significance of interferon-gamma and its signaling pathway in early breast cancer depends on the molecular subtypes, Int. J. Mol. Sci. (2020) 21.
[85] A. Lundberg, L.S. Lindstrom, J.C. Harrell, C. Falato, J.W. Carlson, P.K. Wright, et al., Gene expression signatures and immunohistochemical subtypes add prognostic value to each other in breast cancer cohorts, Clin. Cancer Res. 23 (2017) 7512–7520.
[86] C. Sotiriou, P. Wirapati, S. Loi, A. Harris, S. Fox, J. Smeds, et al., Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis, J. Natl. Cancer Inst. 98 (2006) 262–272.
[87] N. Putluri, S. Maity, R. Kommagani, C.J. Creighton, V. Putluri, F. Chen, et al., Pathway-centric integrative analysis identifies RRM2 as a prognostic marker in breast cancer associated with poor survival and tamoxifen resistance, Neoplasia 16 (2014) 390–402.
[88] A. Naderi, A.E. Teschendorff, N.L. Barbosa-Morais, S.E. Pinder, A.R. Green, D.G. Powe, et al., A gene-expression signature to predict survival in breast cancer across independent data sets, Oncogene 26 (2007) 1507–1516.
[89] K. Chin, S. DeVries, J. Fridlyand, P.T. Spellman, R. Roydasgupta, W.L. Kuo, et al., Genomic and transcriptional aberrations linked to breast cancer pathophysiologies, Cancer Cell 10 (2006) 529–541.
[90] K.R. Hess, K. Anderson, W.F. Symmans, V. Valero, N. Ibrahim, J.A. Mejia, et al., Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer, J. Clin. Oncol. 24 (2006) 4236–4244.
[91] L.D. Miller, J. Smeds, J. George, V.B. Vega, L. Vergara, A. Ploner, et al., An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 13550–13555.
[92] L.J. van 't Veer, H. Dai, M.J. van de Vijver, Y.D. He, A.A. Hart, M. Mao, et al., Gene expression profiling predicts clinical outcome of breast cancer, Nature 415 (2002) 530–536.
[93] M.J. van de Vijver, Y.D. He, L.J. van't Veer, H. Dai, A.A. Hart, D.W. Voskuil, et al., A gene-expression signature as a predictor of survival in breast cancer, N. Engl. J. Med. 347 (2002) 1999–2009.
[94] C. Yau, L. Esserman, D.H. Moore, F. Waldman, J. Sninsky, C.C. Benz, A multigene predictor of metastatic outcome in early stage hormone receptor-negative and triple-negative breast cancer, Breast Cancer Res. 12 (2010) R85.
[95] C. Desmedt, F. Piette, S. Loi, Y. Wang, F. Lallemand, B. Haibe-Kains, et al., Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series, Clin. Cancer Res. 13 (2007) 3207–3214.
[96] B.V. Sinn, C. Fu, R. Lau, J. Litton, T.H. Tsai, R. Murthy, et al., SETER/PR: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer, NPJ Breast Cancer 5 (2019) 16.
[97] C. Desmedt, F. Piette, S. Loi, Y. Wang, F. Lallemand, B. Haibe-Kains, et al., Consortium T: strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series, Clin. Cancer Res. 13 (2007) 3207–3214.