HHS Public Access

High RAD51 gene expression is associated with aggressive biology and with poor survival in breast cancer

Rongrong Wu[1,2], **Ankit Patel**[1], **Yoshihisa Tokumaru**[3], **Mariko Asaoka**[2], **Masanori Oshi**[1,4], **Li Yan**[5], **Takashi Ishikawa**[2], **Kazuaki Takabe**[1,2,4,6,7,8,*]

[1]Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263, USA

[2]Department of Breast Surgery and Oncology, Tokyo Medical University, Tokyo 160-8402, Japan

[3]Department of Surgical Oncology, Graduate School of Medicine, Gifu University, Gifu 501-1193, Japan

[4]Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa, 236-0004, Japan

[5]Department of Biostatistics & Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263, USA

[6]Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, State University of New York, Buffalo, NY 14263, USA

[7]Department of Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan

[8]Department of Breast Surgery, Fukushima Medical University, Fukushima, Japan

Abstract

*Corresponding Author: Kazuaki Takabe, MD, PhD, FACS, Roswell Park Comprehensive Cancer Center, Elm & Carlton Streets, Buffalo, New York 14263, USA, phone: +1-716-845-5128; kazuaki.takabe@roswellpark.org.

Conflict of interest
No conflicts of interest to disclose.

Ethical approval
Ethical review and approval were waived for this study, due to that all cohort information was de-identified and had passed ethical review at the initial publication.

Consent to publication
Patient consent was waived due to that all data was obtained from de-identified publicly available cohorts, and patient consent was obtained at the time of initial publication.

Purpose: Although the DNA repair mechanism is important in preventing carcinogenesis, its activation in established cancer cells may support their proliferation and aggravate cancer progression. RAD51 cooperates with BRCA2 and is essential in the homologous recombination of DNA repair. To this end, we hypothesized that RAD51 gene expression is associated with cancer cell proliferation and poor prognosis of breast cancer (BC) patients.

Methods: A total of 8515 primary BC patients with transcriptome and clinical data from 17 independent cohorts were analyzed. The median value was used to divide each cohort into high and low RAD51 expression groups.

Results: High RAD51 expression enriched the DNA repair gene set and was correlated with DNA repair-related genes. Nottingham histological grade, Ki67 expression and cell proliferation-related gene sets (E2F Targets, G2M Checkpoint and Myc Targets) were all significantly associated with the high RAD51 BC group. RAD51 expression was positively correlated with Homologous Recombination Deficiency, as well as both mutational burden and neoantigens that accompanied a higher infiltration of immune cells. Primary BC with lymph node metastases was associated with high expression of RAD51 in two cohorts. There was no strong correlation between RAD51 expression and drug sensitivity in cell lines, and RAD51 expression was lower after the neoadjuvant chemotherapy compared to before the treatment. High RAD51 BC was associated with poor prognosis consistently in three independent cohorts.

Conclusion: RAD51 gene expression is associated with aggressive cancer biology, cancer cell proliferation, and poor survival in breast cancer.

INTRODUCTION

Homologous recombination repair is an important DNA repair mechanism to address DNA double-strand breaks caused by various external and internal stressors [1]. *BRCA1* and *BRCA2* genes are identified to be essential for homologous recombination repair [2], as mutations in germline *BRCA1* and/or *BRCA2* genes induce genomic instability due to homologous recombination deficiency (HRD), leading to an increased risk of breast and/or ovarian carcinogenesis [2]. HRD is not only an important cause of hereditary breast cancer but also contributes to "BRCA-ness", which are the traits of *BRCA1* genetic disorder found in some sporadic breast cancers [3–5]. HRD is a critical therapeutic target in breast cancer because nearly 70% of the most aggressive triple negative breast cancer (TNBC) subtype contain characteristic "BRCA-ness" features [6]. In addition, poly ADP-ribose polymerase (PARP) has also been identified to be essential in DNA repair. PARP inhibitors are used to induce DNA double-strand breaks and destroy cancer cells with HRD. The effectiveness of PARP inhibitors against breast cancer with germline BRCA mutation have been confirmed in multiple clinical trials [7–9].

RAD51 is an ATPase that forms helical nucleoprotein filaments on single or double-stranded DNA [10] and plays a critical role in the early stages of DNA double-strand break recognition during homologous recombination repair. BRCA2 activates the homologous

recombination cascade in a RAD51-dependent manner, particularly during mitosis [11]. BRCA2 recognizes nuclear filaments in single-stranded DNA loaded with RAD51 during DNA damage and invades the homologous DNA duplex to pair up and initiate homologous recombination repair [12, 13]. Although *RAD51* expression is tightly regulated in normal cells to avoid aberrant DNA recombination [14], its expression is strongly upregulated in several types of cancer including breast [15–18]. High levels of RAD51 over-activate homologous recombination, resulting in uncontrolled double-stranded DNA break repairs and cancer cell persistence [19]. Therefore, high expression of RAD51 confers resistance to radiation and drugs whose typical function is to induce double-stranded breaks in cancer cells [20–22]. Based on these mechanisms, studies have reported the involvement of RAD51 in cancer resistance to PARP inhibitors [23, 24]. Studies have even suggested RAD51 to be a candidate as a biomarker of drug sensitivity and as a therapeutic target to combat drug resistance.

Our group has been pursuing translational research that addresses the clinical relevance of gene expression using in-silico analysis of large patient cohorts of transcriptomes associated with clinical parameters [25–32]. Previously, we have reported that increased expression of the *BRCA2* gene is associated with enhanced cancer cell proliferation and immunogenicity in breast cancer [33]. In cancer cells, high expression of *BRCA2* correlated with HRD and was also associated with aggressive breast cancer. Noting that RAD51 acts together with BRCA1 and/or BRCA2 as a key player in homologous recombination repair, we hypothesized that *RAD51* mRNA expression is associated with increased cancer cell proliferation, and thus with poor prognosis. In addition, we hypothesized that *RAD51* might be highly expressed in the treatment non-responder group due to its involvement in drug resistance. To date, studies of *RAD51* have been limited to experiments with cell lines, animals, and retrospective studies with small cohorts. In contrast, we analyzed the relationship between *RAD51* gene expression and breast cancer using three large primary breast cancer cohorts containing several thousand patients. In addition, we analyzed *RAD51* expression by treatment response using multiple neoadjuvant chemotherapy (NAC) treated breast cancer cohorts to explore its potential as a predictor and biomarker of treatment response in breast cancer.

METHODS

Patient data acquisition

All cohorts were downloaded in September 2021. A total of 8515 patients were included in the analysis. Clinicopathological factors and mRNA sequencing data for 1077 breast cancer patients of The Cancer Genome Atlas (TCGA) were downloaded from cBiopotal [34–36]. Batch-normalized RNA sequencing data from Illumina HiSeq_RNASeqV2 to HUGO symbols were used with log2 conversion. We also downloaded the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) cohort [37] of 1904 breast cancer from cBiopotal and microarray RNA expression data annotated from illumina Human v3 to HUGO symbols was used. GSE96058 is a validation cohort of 3069 breast cancer patients from the Sweden Cancerome Analysis Network- Breast study that assessed the genomic profile of early breast cancer [38]. Clinicopathological factors for GSE96058 were

downloaded using the R package GEOquary and RNA sequence data annotated with the HUGO symbol were downloaded directly from NCBI Gene Expression Omnibus database [39]. To analyze neoadjuvant chemotherapy response, the following primary breast cancer cohorts were analyzed: GSE21974 [40], GSE28844 [41], GSE114403 [42], GSE87455 [43, 44], GSE25066 [45, 46], GSE50948 [47], GSE20271 [48, 49], GSE20194 [50, 51], GSE180962 [52], GSE22358 [53], GSE22226 [54], GSE163882 [55], GSE34138 [56, 57], and GSE16446 [58]. These cohorts were downloaded from the GEO database via the R package GEOquary as well. For the GSE180962 cohort, only the control group was used in the analysis. The expression of *RAD51* was calculated from the mean value of probes assigned to *RAD51* from the platform corresponding to each expression data series. Details of the treatment information, the number of patients included in the study, the access number of the platform used for annotation for each cohort are summarized in Supplementary Table 1 and 2.

BRCA mutation data

Genetic mutation data were available for TCGA and METABRIC. TCGA had 30 patients with BRCA1 mutations, 29 patients with BRCA2 mutations, and a total of 57 patients with either mutation. METABRIC had 55, 60, and 114 patients, respectively.

Breast cancer cell line RAD51 expression and drug sensitivity data

Breast cancer cell line RNA sequence data and drug susceptibility data were obtained from the Depmap portal, as we reported previously [59, 60]. This included 64 breast cancer cell lines, and immunohistochemistry staining data were downloaded as well. Expression 21Q3 Public data was used for *RAD51* expression and AUC data from PRISM primary or secondary screening. GDSC1 and GDSC2 data was used to determine drug sensitivity.

Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) [61] was performed on the gene expression data by dividing the analysis dataset into two groups based on the median expression of *RAD51*. This approach examines how strongly pathways defined by particular genes are expressed between two sets. GSEA 4.1.0 (free software from Broad Institute) was used for the analysis and Hallmark was selected as the gene set from the major collection of the Molecular Signatures Database [62]. Following the recommendations of the Broad Institute, FDR q-values below 25% were used as cut-off values for significance, and the Normalized Enrichment Score (NES) was used to assess the strength of the correlation with the gene set.

Immune cell fractionation, HRD, and mutation score analysis

TCGA HRD score, intratumoral heterogeneity score, mutation burden score, and immune activity score were calculated and reported by Thorsson et al. in 2018 [63]. Fractionation of intratumoral immune cells and stromal cells was calculated using the xCell web tool [64], an algorithm for enumerating immune cell subsets from the transcriptome, as previously reported [65–67]. xCell estimates immune cell fraction for each cohort by comparing 489 gene signatures corresponding to 64 cell types, including adaptive and innate immune cells, hematopoietic progenitor cells, epithelial cells, and extracellular matrix cells, with the input

of bulk gene expression dataset. CYT score was used as a measure of immune activity, as previously reported [67, 68].

Statistical analysis

Data downloading, organization, analysis, and visualization were done using R 4.0.1. The following packages were used in this study: Survival 3.2–11, survAUC 1.0–5, S4Vectors 0.30. 0, MatrixGenerics 1.4.3, Biobase 2.52.0, grayzoneSurv 1.0, RcmdrPlugin.EZR 1.54, RcmdrMisc 2.7–1. ggplot2 3.3.5, backports 1.2.1, tidyverse 1.3.1, GEOquery 2.60.0, SummarizedExperiment 1.22.0. Median values were used for all cut-offs for comparisons between high and low *RAD51* groups. All p-values were calculated by a two-sided test and the cut-off for statistical significance was set at 0.05.

RESULTS

RAD51 gene expression was associated with DNA repair activity in breast cancer

RAD51 is known to play an essential role in the DNA repair mechanism. Therefore, we first investigated whether *RAD51* gene expression was associated with the DNA repair pathway and with the expression of its member genes. Comparison of *RAD51* expression between normal breast and tumor tissues in the TCGA cohort showed that *RAD51* was highly expressed in breast cancer (p<0.001, Figure1a). Intratumoral heterogeneity and Homologues Recombination Deficiency (HRD) scores were positively correlated with *RAD51* expression in TCGA (r=0.32 and 0.53, respectively. Figure1b). Further, *RAD51* high breast cancer significantly enriched the DNA repair gene set consistently in TCGA, METABRIC, and GSE96058 cohorts (All p<0.001 and FDR<0.01, Figure1c). It was also associated with high expression of DNA repair genes, such as *BRCA1*, *BRCA2*, *E2F1*, *E2F4*, *E2F7*, and *CDK12*, consistently across all three cohorts (TCGA, METABRIC, and GSE96058. All p<0.05. Figure1c). Since we were unable to find any literature on the correlation between RAD51 and BRCA1 or BRCA2 expressions in in-vitro or in-vivo studies, we analyzed the relationship in breast cancer cell lines. We found that RAD51 gene expression strongly correlated with BRCA1 or BRCA2 expression in analyzing the cancer cell line encyclopedia (Supplementary Figure 1). To this end, we found that *RAD51* expression is associated with DNA repair activity in the breast cancer tumor microenvironment (TME).

RAD51 gene expression was strongly associated with cancer cell proliferation

As cancer with HRD is known to be highly malignant, we investigated the relationship between *RAD51* expression and cancer cell proliferation. Utilizing the score value provided by Thorsson et. al. [63], we found a very strong correlation between *RAD51* expression and the Proliferation score in the TCGA cohort (r=0.879, p<0.001, Figure2a). RAD51 expression strongly correlated with Nottingham histological grades and pathological quantification of cancer cell proliferation consistently across all three cohorts- TCGA, METABRIC, and GSE96058 (all p<0.001, Figure2b). RAD51 expression also correlated with the Tubular score, Nuclear score, and Mitotic score in TCGA (Supplementary Figure2). In agreement, *RAD51* expression was highly correlated with the cell proliferation marker gene, *MKI67*, across all three cohorts (all r>0.4, Figure2b). Strikingly, all five of the cell proliferation-related gene sets in the Hallmark collection (E2F Targets, G2M Checkpoint,

Myc Targets v1 and v2, and Mitotic Spindle) and MTORC1 Signaling were enriched in high-*RAD51* breast cancer group consistently across all cohorts with a strong significance of FDR<0.01 (Figure2c). These results suggested that high *RAD51* breast cancer is associated with high cancer cell proliferation.

RAD51 was associated with a high mutation rate

As *RAD51* mainly co-acts with BRCA2 and partly with BRCA1, it was of interest to investigate whether *RAD51* expression was associated with overall mutation rates and BRCA gene mutations. Silent or Non-silent mutation rates were significantly increased in the high *RAD51* expression breast cancer group in TCGA (both p<0.001, Figure3a). In addition, we compared wild type group to the mutation bearing group (in *BRCA1*, *BRCA2*, or both). *RAD51* expression was significantly higher in patients with mutations in *BRCA1*, *BRCA2*, or in both across the METABRIC cohort (all p<0.01). However, this was not validated in TCGA cohort (Figure3b). To this end, *RAD51* expression correlated with cancer mutation level, but not consistently, with BRCA mutations.

RAD51 high breast cancer was immunogenic and elicited cancer immunity in the tumor microenvironment

We have previously reported that cancers with high mutation rates elicit immunogenicity and cancer immunity. Having identified increased levels of mutations in high *RAD51* breast cancers, it was of interest to investigate the association of *RAD51* expression with cancer immunity. As expected, single-nucleotide variant (SNV) neoantigens and Indel neoantigens were both significantly higher in breast cancer with high *RAD51* expression. Several factors related to cancer immunity (interferon (IFN)-gamma response, tumor infiltrating lymphocytes (TIL) regional fraction, Wound Healing, B-Cell Receptor (BCR) Richness, BCR Shannon, and Fraction altered) were all significantly higher in the high *RAD51* group (All p<0.001, Figure 4a). Further, we investigated the amount of immune cells in TME, and several immune cell types (CD4 naive T-cells, CD4+ memory T-cells, T helper type1 cells, T helper type2 cells, Plasma cells, M1 macrophage, and activated dendric cells) were significantly infiltrated in the high *RAD51* breast cancer group. Cytolytic Activity score (CYT), which reflects overall immune cell killing, was also significantly increased consistently across all the three cohorts (All p<0.001, Figure4b). Thus, we could conclude that high *RAD51* expressing breast cancer is highly immunogenic and has activated cancer immunity.

RAD51 gene expression was associated with triple-negative breast cancer and with lymph node metastasis

To further elucidate the characteristics of high *RAD51* breast cancer, we analyzed its association with clinicopathological factors. Consistently among the three cohorts, *RAD51* was highest expressed in triple negative breast cancer (TNBC) compared to all other immunohistochemical subtypes of breast cancer (all p<0.001, Figure5). In contrast, the estrogen receptor (ER)-positive/epidermal growth factor receptor 2 (HER2)-negative subtype had the lowest expression of *RAD51*. *RAD51* expression was higher in advanced stages in TCGA, but this was not validated in the METABRIC cohort. *RAD51* expression was significantly increased in the primary tumors of patients with more metastatic lymph nodes

in both the METABRIC and GSE96058 cohorts (both p<0.02), which was not validated in TCGA. On the other hand, the primary breast cancer *RAD51* expression did not change with the presence of distant metastases. These results suggest that *RAD51* is highly expressed in aggressive TNBC and in primary breast cancer with lymph node metastasis.

RAD51 expression is high in tumor that achieved pathological complete response after NAC

Breast cancer containing a *BRCA1* and/or *BRCA2* mutation with HRD is known to be sensitive to platinum-based cytotoxic chemotherapy and PARP inhibitors. In addition, *RAD51* expression was reported to be associated with resistance to PARP inhibitors. To this end, the relationship between *RAD51* expression and treatment drug sensitivity was of interest to investigate. We analyzed the sensitivity to cytotoxic chemotherapies and multiple PARP inhibitors in comparison to *RAD51* gene expression across breast cancer cell lines from the Depmap portal. In TNBC cell lines, *RAD51* expression was positively correlated with sensitivity to docetaxel and epirubicin, but not with cisplatin (both p<0.05 and r>0.5, Figure6a). However, none of the sensitivity to PARP inhibitors correlated with *RAD51* expression (Figure6a). On the other hand, *RAD51* expression significantly correlated with sensitivity to niraparib in ER-positive/HER2-negative cell lines (p<0.05 and r=0.9, Figure6a).

As *RAD51* was reported to have a role in drug resistance, it was of interest to investigate its association with pathological complete response (pCR) after neoadjuvant chemotherapy (NAC). Interestingly, *RAD51* expression consistently decreased after NAC in all 4 cohorts (all p<0.01, Figure6b). A comparison in paired samples before and after NAC showed the same results (Supplementary Figure2). *RAD51* expression between groups that did achieve pCR versus those that did not was investigated by immunohistochemical subtype (Figure6c). Although we expected that *RAD51* expression to be higher in the residual disease (RD) group and particularly in TNBC, that scenario was only found in a single cohort (GSE20271 p=0.042, Figure6c). The opposite relationship was found in other cohorts (GSE25066 p=0.001, Figure6c), and most of the cohorts did not show any significant difference in RAD51 expression as related to pCR in TNBC. In contrast, *RAD51* expression was higher in pCR group in the ER+HER2- subtype across two cohorts (GSE50948 and GSE20271, both p<0.05, Figure6c). These results suggest that *RAD51* expression of a bulk tumor does not predict response to NAC.

RAD51 high breast cancer has worse survival consistently across all three cohorts

Given that breast cancers with high expression of *RAD51* are more aggressive, it was of interest to investigate whether these characteristics translated into survival disparities. To this end, we compared the survival between high and low *RAD51* expression groups. Surprisingly, overall survival (OS) was significantly worse in the high-*RAD51* breast cancer group consistently across all three cohorts, and the same was observed in disease-specific survival (DSS) in TCGA and METABRIC. Disease-free survival (DFS) was only significant in METABRIC alone (Figure7). These differences may be because the number of patients and follow-up period are approximately half of that found in the METABRIC compared to TCGA. In short, the expression of *RAD51* was associated with a worse prognosis.

DISCUSSION

In this study, we investigated the characteristics of breast cancers with high *RAD51* expression through functional analysis of clinical, immunohistochemical, and transcriptomic data using multiple large breast cancer patient cohorts. In line with previous reports, we found that *RAD51* was highly expressed in cancer compared to normal tissues, and strongly correlated with HRD and intratumor heterogeneity. We also showed that the DNA repair gene set, as well as multiple genes related to homologous recombination repair, were significantly associated with high *RAD51* expression. Further, breast cancers with high *RAD51* expression were significantly correlated with histological grade and all five Hallmark cell proliferation-related gene sets, indicating that *RAD51* high tumors are highly proliferative. *RAD51* was also positively correlated with mutation rates. RAD51 expression in BRCA-mutant tumors was significantly higher than in BRCA-wild-type tumors in METABRIC, but this was not validated in TCGA cohort. This was most likely because of a lack of power due to a low number of mutant cases in TCGA (about half of that of METBARIC). Martin et. al. also reported a significant correlation between RAD51 expression and BRCA1 mutation, reporting that tumors with BRCA1 mutation had 2.5-fold higher expression of RAD51 compared with wild type in the gene expression microarray of 117 primary breast tumors [69]. Cancer cell immunogenicity and cancer immune activity were all significantly enhanced in high-*RAD51* tumors across all three cohorts, and the infiltration of each immune cell was also observed in all cohorts. Primary tumors from patients with lymph node metastases were associated with high expression of *RAD51* in both TCGA and METABRIC cohorts. There was no strong correlation between *RAD51* expression and treatment drug sensitivity other than Niraparib in the ER-positive/HER2-negative subtype. Contrary to our expectation, *RAD51* expression was lower after NAC compared to prior to treatment consistently across three independent cohorts. *RAD51* expression was higher in primary tumors that did not achieve pCR after NAC compared to tumors that did achieve pCR in only one among ten independent TNBC NAC cohorts analyzed, whereas this was not validated in any other subtypes in the other cohorts. Finally, overall survival was significantly worse in high *RAD51* breast cancer across all three large cohorts. DSS was also worse in TCGA and METABRIC, and DFS was also worse in METABRIC.

We found that *RAD51* expression was highly associated with cancer cell proliferation across multiple cohorts, which agrees with Maack et. al. who reported that RAD51 was more highly expressed in invasive breast cancer with higher grades [15]. Although multigene assay risk scores were not available in the cohorts examined in this study, our results were consistent with previous studies in that RAD51 expression was proportional to clinical proliferation indices such as Nottingham histological grade across all cohorts and Tubular score, Nuclear score and Mitosis score in TCGA. *RAD51* was most highly expressed in TNBC, which is known to be the most aggressive subtype of breast cancer. Although not consistent in all cohorts, our study suggested high *RAD51* expression occurred at more advanced stages of breast cancer, such as the presence of multiple lymph node metastases, which is consistent with a previous report that *RAD51* protein was associated with cancer progression and metastasis of sporadic breast cancer [70]. High-*RAD51* breast cancer had

a higher mutational burden and increased number of neoantigens, and thus, were more immunogenic. Although there was an increased immune cell infiltration in high-*RAD51* breast cancer, none of the immune-related gene sets enriched to *RAD51* high tumor, suggesting anti-cancer immunity was not truly activated. As a result of its strong reflection of cancer aggressors, *RAD51* high expression was significantly associated with a poor prognosis across all the large cohorts analyzed in this study.

RAD51 was highlighted as a potential marker for predicting treatment response of breast cancer. BRCA-deficient ovarian and breast cancers with HRD showed sensitivity to PARP inhibitors and DNA-damaging drugs such as platinum, because these drugs arrest a large number of replication forks and lead to synthetic lethality [71]. Since these processes can be circumvented by RAD51, which plays a central role in the repair and restart of replication forks [72, 73], the high expression of RAD51 is thought to lead to resistance to these drugs [74]. RAD51 histological expression as identified by fluorescent immunostaining was found to reflect homologous recombination repair function and was claimed as a predictive marker of pCR after NAC in TNBC [75]. Loss of RAD51 fusion in TNBC correlated with HRD as well as with pCR after platinum-based neoadjuvant chemotherapy [76]. However, *RAD51* gene expression in our study showed discrepant results to the previously reported RAD51 assay, which was a functional HRD marker scored by simultaneous expression of both RAD51 and geminin, a cell proliferation marker [75]. Low-RAD51 tumors determined by RAD51 assay were most frequently of the TNBC subtype, which was inverse to the relation found by our RAD51 gene expression study. Furthermore, high *RAD51* expression was positively correlated with HRD, indicating that there may be a dissociation between these functional HRD markers and the gene expression of *RAD51*. Comparison of drug sensitivity with *RAD51* expression suggested that *RAD51* expression may be positively correlated with chemotherapy sensitivity in TNBC cell lines. Interestingly, no resistance to PARP inhibitors was observed. The original RAD51 assay study also showed that RAD51 was barely expressed in the baseline biopsy samples but was upregulated in samples taken immediately after radiation-induced DNA damage [75]. However, *RAD51* gene expression was downregulated after NAC in our study comparing pre and post NAC samples. *RAD51* was not under-expressed in the group that achieved pCR after NAC, and conversely, was highly expressed in the pCR group in some cohorts. It is unclear whether this difference is due to differences between *RAD51* gene expression in the RAD51 assay and in bulk tumors, but the function of RAD51 as a marker of drug sensitivity is questionable.

The limitations of this study are as follow. First, there is a patient selection bias in the large cohort included in this analysis, because the patient information was collected more than 10 years ago. Patients receiving newly authorized treatments, such as PARP inhibitors, are not included. Second, the in-vitro cohort was all small, with fewer than 30 cell lines, so a larger number of studies of PARP inhibitors across cell lines may give different results. In addition, we did not perform in-vivo or in-vitro experiments, so the mechanisms by which RAD51 induces cell proliferation and drug resistance will require more detailed testing. Given our result that RAD51 expression was associated with immunogenicity, it was of interest to investigate its potential as a biomarker for immunotherapy. However, we were unable to pursue this since we did not have access to breast cancer patient cohorts with information on response to immunotherapy. In addition, as all our studies

have been conducted in retrospective cohorts, prospective studies will need to be designed to investigate the usefulness of RAD51 as a biomarker.

CONLUSION

RAD51 expression is strongly associated with aggressive cancer biology, cancer cell proliferation, and poor survival in breast cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

FUNDING

DATA AVAILABILITY

Publicly available datasets were analyzed in this study. TCGA data can be found here: [https://www.cbioportal.org. /Breast Invasive Carcinoma (TCGA, PanCancer Atlas)]. METABRIC data can be found here: [https://www.cbioportal.org. / Breast Cancer (METABRIC, Nature 2012 & Nat Commun 2016)]. Data sets from each of the GEO databases can be downloaded from the following sites and access numbers: [https://www.ncbi.nlm.nih.gov/geo. / GSE96058/ GSE21974/ GSE28844/ GSE114403/ GSE87455/ GSE25066/ GSE50948/ GSE20271/ GSE20194/ GSE180962/ GSE22358/ GSE22226/ GSE163882/ GSE34138/ GSE16446]. Breast cancer cell line data can be found here: [https://depmap.org/portal/].

REFERENCES

1. Hoeijmakers JH (2001) Genome maintenance mechanisms for preventing cancer. Nature 411:366–374. doi: 10.1038/35077232 [PubMed: 11357144]

2. Venkitaraman AR (2002) Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 108:171–182. doi: 10.1016/s0092-8674(02)00615-3 [PubMed: 11832208]

3. Teraoka S, Muguruma M, Takano N, Miyahara K, Kawate T, Kaise H, Yamada K, Miyazawa K, Ishikawa T (2020) Association of BRCA Mutations and BRCAness Status With Anticancer Drug Sensitivities in Triple-Negative Breast Cancer Cell Lines. J Surg Res 250:200–208. doi: 10.1016/j.jss.2019.12.040 [PubMed: 32092597]

4. Teraoka S, Sato E, Narui K, Yamada A, Fujita T, Yamada K, Oba M, Ishikawa T (2020) Neoadjuvant Chemotherapy With Anthracycline-Based Regimen for BRCAness Tumors in Triple-Negative Breast Cancer. J Surg Res 250:143–147. doi: 10.1016/j.jss.2019.12.047 [PubMed: 32044511]

5. Turner N, Tutt A, Ashworth A (2004) Hallmarks of 'BRCAness' in sporadic cancers. Nat Rev Cancer 4:814–819. doi: 10.1038/nrc1457 [PubMed: 15510162]

6. Lips EH, Mulder L, Oonk A, van der Kolk LE, Hogervorst FB, Imholz AL, Wesseling J, Rodenhuis S, Nederlof PM (2013) Triple-negative breast cancer: BRCAness and concordance of clinical features with BRCA1-mutation carriers. Br J Cancer 108:2172–2177. doi: 10.1038/bjc.2013.144 [PubMed: 23558900]

7. Eikesdal HP, Yndestad S, Elzawahry A, Llop-Guevara A, Gilje B, Blix ES, Espelid H, Lundgren S, Geisler J, Vagstad G, Venizelos A, Minsaas L, Leirvaag B, Gudlaugsson EG, Vintermyr OK, Aase HS, Aas T, Balmana J, Serra V, Janssen EAM, Knappskog S, Lonning PE (2021) Olaparib monotherapy as primary treatment in unselected triple negative breast cancer. Ann Oncol 32:240–249. doi: 10.1016/j.annonc.2020.11.009 [PubMed: 33242536]

8. Litton JK, Rugo HS, Ettl J, Hurvitz SA, Goncalves A, Lee KH, Fehrenbacher L, Yerushalmi R, Mina LA, Martin M, Roche H, Im YH, Quek RGW, Markova D, Tudor IC, Hannah AL, Eiermann W, Blum JL (2018) Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. N Engl J Med 379:753–763. doi: 10.1056/NEJMoa1802905 [PubMed: 30110579]

9. Robson M, Im SA, Senkus E, Xu B, Domchek SM, Masuda N, Delaloge S, Li W, Tung N, Armstrong A, Wu W, Goessl C, Runswick S, Conte P (2017) Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. N Engl J Med 377:523–533. doi: 10.1056/NEJMoa1706450 [PubMed: 28578601]

10. Mimitou EP, Symington LS (2009) Nucleases and helicases take center stage in homologous recombination. Trends Biochem Sci 34:264–272. doi: 10.1016/j.tibs.2009.01.010 [PubMed: 19375328]

11. Kojic M, Kostrub CF, Buchman AR, Holloman WK (2002) BRCA2 homolog required for proficiency in DNA repair, recombination, and genome stability in Ustilago maydis. Mol Cell 10:683–691. doi: 10.1016/s1097-2765(02)00632-9 [PubMed: 12408834]

12. San Filippo J, Sung P, Klein H (2008) Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 77:229–257. doi: 10.1146/annurev.biochem.77.061306.125255 [PubMed: 18275380]

13. Moynahan ME, Pierce AJ, Jasin M (2001) BRCA2 is required for homology-directed repair of chromosomal breaks. Mol Cell 7:263–272. doi: 10.1016/s1097-2765(01)00174-5 [PubMed: 11239455]

14. Richardson C, Stark JM, Ommundsen M, Jasin M (2004) Rad51 overexpression promotes alternative double-strand break repair pathways and genome instability. Oncogene 23:546–553. doi: 10.1038/sj.onc.1207098 [PubMed: 14724582]

15. Maacke H, Opitz S, Jost K, Hamdorf W, Henning W, Kruger S, Feller AC, Lopens A, Diedrich K, Schwinger E, Sturzbecher HW (2000) Over-expression of wild-type Rad51 correlates with histological grading of invasive ductal breast cancer. Int J Cancer 88:907–913. doi: 10.1002/1097-0215(20001215)88:6<907::aid-ijc11>3.0.co;2-4 [PubMed: 11093813]

16. Qiao GB, Wu YL, Yang XN, Zhong WZ, Xie D, Guan XY, Fischer D, Kolberg HC, Kruger S, Stuerzbecher HW (2005) High-level expression of Rad51 is an independent prognostic marker of survival in non-small-cell lung cancer patients. Br J Cancer 93:137–143. doi: 10.1038/sj.bjc.6602665 [PubMed: 15956972]

17. Sarwar R, Sheikh AK, Mahjabeen I, Bashir K, Saeed S, Kayani MA (2017) Upregulation of RAD51 expression is associated with progression of thyroid carcinoma. Exp Mol Pathol 102:446–454. doi: 10.1016/j.yexmp.2017.05.001 [PubMed: 28502582]

18. Zhang X, Ma N, Yao W, Li S, Ren Z (2019) RAD51 is a potential marker for prognosis and regulates cell proliferation in pancreatic cancer. Cancer Cell Int 19:356. doi: 10.1186/s12935-019-1077-6 [PubMed: 31889908]

19. Richardson C (2005) RAD51, genomic stability, and tumorigenesis. Cancer Lett 218:127–139. doi: 10.1016/j.canlet.2004.08.009 [PubMed: 15670890]

20. Raderschall E, Stout K, Freier S, Suckow V, Schweiger S, Haaf T (2002) Elevated levels of Rad51 recombination protein in tumor cells. Cancer Res 62:219–225. [PubMed: 11782381]

21. Maacke H, Jost K, Opitz S, Miska S, Yuan Y, Hasselbach L, Luttges J, Kalthoff H, Sturzbecher HW (2000) DNA repair and recombination factor Rad51 is over-expressed in human pancreatic adenocarcinoma. Oncogene 19:2791–2795. doi: 10.1038/sj.onc.1203578 [PubMed: 10851081]

22. Nagathihalli NS, Nagaraju G (2011) RAD51 as a potential biomarker and therapeutic target for pancreatic cancer. Biochim Biophys Acta 1816:209–218. doi: 10.1016/j.bbcan.2011.07.004 [PubMed: 21807066]

23. Cruz C, Castroviejo-Bermejo M, Gutierrez-Enriquez S, Llop-Guevara A, Ibrahim YH, Gris-Oliver A, Bonache S, Morancho B, Bruna A, Rueda OM, Lai Z, Polanska UM, Jones GN, Kristel P,

de Bustos L, Guzman M, Rodriguez O, Grueso J, Montalban G, Caratu G, Mancuso F, Fasani R, Jimenez J, Howat WJ, Dougherty B, Vivancos A, Nuciforo P, Serres-Creixams X, Rubio IT, Oaknin A, Cadogan E, Barrett JC, Caldas C, Baselga J, Saura C, Cortes J, Arribas J, Jonkers J, Diez O, O'Connor MJ, Balmana J, Serra V (2018) RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in germline BRCA-mutated breast cancer. Ann Oncol 29:1203–1210. doi: 10.1093/annonc/mdy099 [PubMed: 29635390]

24. Zhao Q, Guan J, Zhang Z, Lv J, Wang Y, Liu L, Zhou Q, Mao W (2017) Inhibition of Rad51 sensitizes breast cancer cells with wild-type PTEN to olaparib. Biomed Pharmacother 94:165–168. doi: 10.1016/j.biopha.2017.07.090 [PubMed: 28759753]

25. Satyananda V, Oshi M, Tokumaru Y, Maiti A, Hait N, Matsuyama R, Endo I, Takabe K (2021) Sphingosine 1-phosphate (S1P) produced by sphingosine kinase 1 (SphK1) and exported via ABCC1 is related to hepatocellular carcinoma (HCC) progression. American journal of cancer research 11:4394–4407. [PubMed: 34659894]

26. Satyananda V, Oshi M, Endo I, Takabe K (2021) High BRCA2 Gene Expression is Associated with Aggressive and Highly Proliferative Breast Cancer. Annals of surgical oncology 28:7356–7365. doi: 10.1245/s10434-021-10063-5 [PubMed: 33966140]

27. Satyananda V, Oshi M, Endo I, Takabe K (2021) ASO Author Reflections: High BRCA2 Gene Expression is Associated with Aggressive and Highly Proliferative Breast Cancer. Annals of surgical oncology 28:7366–7367. doi: 10.1245/s10434-021-10135-6 [PubMed: 33966141]

28. Oshi M, Tokumaru Y, Mukhopadhyay S, Yan L, Matsuyama R, Endo I, Takabe K (2021) Annexin A1 Expression Is Associated with Epithelial-Mesenchymal Transition (EMT), Cell Proliferation, Prognosis, and Drug Response in Pancreatic Cancer. Cells 10. doi: 10.3390/cells10030653

29. Oshi M, Newman S, Murthy V, Tokumaru Y, Yan L, Matsuyama R, Endo I, Takabe K (2020) ITPKC as a Prognostic and Predictive Biomarker of Neoadjuvant Chemotherapy for Triple Negative Breast Cancer. Cancers 12. doi: 10.3390/cancers12102758

30. Oshi M, Angarita FA, Tokumaru Y, Yan L, Matsuyama R, Endo I, Takabe K (2020) High Expression of NRF2 Is Associated with Increased Tumor-Infiltrating Lymphocytes and Cancer Immunity in ER-Positive/HER2-Negative Breast Cancer. Cancers 12. doi: 10.3390/cancers12123856

31. Takahashi H, Katsuta E, Yan L, Dasgupta S, Takabe K (2019) High expression of Annexin A2 is associated with DNA repair, metabolic alteration, and worse survival in pancreatic ductal adenocarcinoma. Surgery 166:150–156. doi: 10.1016/j.surg.2019.04.011 [PubMed: 31171367]

32. Okano M, Oshi M, Butash AL, Katsuta E, Tachibana K, Saito K, Okayama H, Peng X, Yan L, Kono K, Ohtake T, Takabe K (2019) Triple-Negative Breast Cancer with High Levels of Annexin A1 Expression Is Associated with Mast Cell Infiltration, Inflammation, and Angiogenesis. International journal of molecular sciences 20. doi: 10.3390/ijms20174197

33. Satyananda V, Oshi M, Endo I, Takabe K (2021) High BRCA2 Gene Expression is Associated with Aggressive and Highly Proliferative Breast Cancer. Annals of surgical oncology. doi: 10.1245/s10434-021-10063-5

34. Cancer Genome Atlas N (2012) Comprehensive molecular portraits of human breast tumours. Nature 490:61–70. doi: 10.1038/nature11412 [PubMed: 23000897]

35. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2:401–404. doi: 10.1158/2159-8290.CD-12-0095 [PubMed: 22588877]

36. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 6:pl1. doi: 10.1126/scisignal.2004088 [PubMed: 23550210]

37. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Group M, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C, Aparicio S (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486:346–352. doi: 10.1038/nature10983 [PubMed: 22522925]

38. Saal LH, Vallon-Christersson J, Hakkinen J, Hegardt C, Grabau D, Winter C, Brueffer C, Tang MH, Reutersward C, Schulz R, Karlsson A, Ehinger A, Malina J, Manjer J, Malmberg M, Larsson C, Ryden L, Loman N, Borg A (2015) The Sweden Cancerome Analysis Network - Breast (SCAN-B) Initiative: a large-scale multicenter infrastructure towards implementation of breast cancer genomic analyses in the clinical routine. Genome Med 7:20. doi: 10.1186/s13073-015-0131-9 [PubMed: 25722745]

39. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30:207–210. doi: 10.1093/nar/30.1.207 [PubMed: 11752295]

40. Stickeler E, Pils D, Klar M, Orlowsk-Volk M, Zur Hausen A, Jager M, Watermann D, Gitsch G, Zeillinger R, Tempfer CB (2011) Basal-like molecular subtype and HER4 up-regulation and response to neoadjuvant chemotherapy in breast cancer. Oncol Rep 26:1037–1045. doi: 10.3892/or.2011.1392 [PubMed: 21769435]

41. Vera-Ramirez L, Sanchez-Rovira P, Ramirez-Tortosa CL, Quiles JL, Ramirez-Tortosa M, Lorente JA (2013) Transcriptional shift identifies a set of genes driving breast cancer chemoresistance. PLoS One 8:e53983. doi: 10.1371/journal.pone.0053983 [PubMed: 23326553]

42. Li X, Warren S, Pelekanou V, Wali V, Cesano A, Liu M, Danaher P, Elliott N, Nahleh ZA, Hayes DF, Hortobagyi GN, Barlow WE, Hatzis C, Pusztai L (2019) Immune profiling of pre- and post-treatment breast cancer tissues from the SWOG S0800 neoadjuvant trial. J Immunother Cancer 7:88. doi: 10.1186/s40425-019-0563-7 [PubMed: 30967156]

43. Kimbung S, Markholm I, Bjohle J, Lekberg T, von Wachenfeldt A, Azavedo E, Saracco A, Hellstrom M, Veerla S, Paquet E, Bendahl PO, Ferno M, Bergh J, Loman N, Hatschek T, Hedenfalk I, Group PT (2018) Assessment of early response biomarkers in relation to long-term survival in patients with HER2-negative breast cancer receiving neoadjuvant chemotherapy plus bevacizumab: Results from the Phase II PROMIX trial. Int J Cancer 142:618–628. doi: 10.1002/ijc.31070 [PubMed: 28940389]

44. Palazon A, Tyrakis PA, Macias D, Velica P, Rundqvist H, Fitzpatrick S, Vojnovic N, Phan AT, Loman N, Hedenfalk I, Hatschek T, Lovrot J, Foukakis T, Goldrath AW, Bergh J, Johnson RS (2017) An HIF-1alpha/VEGF-A Axis in Cytotoxic T Cells Regulates Tumor Progression. Cancer Cell 32:669–683 e665. doi: 10.1016/j.ccell.2017.10.003 [PubMed: 29136509]

45. Itoh M, Iwamoto T, Matsuoka J, Nogami T, Motoki T, Shien T, Taira N, Niikura N, Hayashi N, Ohtani S, Higaki K, Fujiwara T, Doihara H, Symmans WF, Pusztai L (2014) Estrogen receptor (ER) mRNA expression and molecular subtype distribution in ER-negative/progesterone receptor-positive breast cancers. Breast Cancer Res Treat 143:403–409. doi: 10.1007/s10549-013-2763-z [PubMed: 24337596]

46. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, Vidaurre T, Holmes F, Souchon E, Wang H, Martin M, Cotrina J, Gomez H, Hubbard R, Chacon JI, Ferrer-Lozano J, Dyer R, Buxton M, Gong Y, Wu Y, Ibrahim N, Andreopoulou E, Ueno NT, Hunt K, Yang W, Nazario A, DeMichele A, O'Shaughnessy J, Hortobagyi GN, Symmans WF (2011) A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA 305:1873–1881. doi: 10.1001/jama.2011.593 [PubMed: 21558518]

47. Prat A, Bianchini G, Thomas M, Belousov A, Cheang MC, Koehler A, Gomez P, Semiglazov V, Eiermann W, Tjulandin S, Byakhow M, Bermejo B, Zambetti M, Vazquez F, Gianni L, Baselga J (2014) Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. Clin Cancer Res 20:511–521. doi: 10.1158/1078-0432.CCR-13-0239 [PubMed: 24443618]

48. Shen K, Song N, Kim Y, Tian C, Rice SD, Gabrin MJ, Symmans WF, Pusztai L, Lee JK (2012) A systematic evaluation of multi-gene predictors for the pathological response of breast cancer patients to chemotherapy. PLoS One 7:e49529. doi: 10.1371/journal.pone.0049529 [PubMed: 23185353]

49. Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M, Qi Y, Barajas-Figueroa LJ, Souchon E, Coutant C, Doimi FD, Ibrahim NK, Gong Y, Hortobagyi GN, Hess KR, Symmans WF, Pusztai L (2010) Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in

breast cancer. Clin Cancer Res 16:5351–5361. doi: 10.1158/1078-0432.CCR-10-1265 [PubMed: 20829329]

50. Shi L, Campbell G, Jones WD, Campagne F, Wen Z, Walker SJ, Su Z, Chu TM, Goodsaid FM, Pusztai L, Shaughnessy JD Jr., Oberthuer A, Thomas RS, Paules RS, Fielden M, Barlogie B, Chen W, Du P, Fischer M, Furlanello C, Gallas BD, Ge X, Megherbi DB, Symmans WF, Wang MD, Zhang J, Bitter H, Brors B, Bushel PR, Bylesjo M, Chen M, Cheng J, Cheng J, Chou J, Davison TS, Delorenzi M, Deng Y, Devanarayan V, Dix DJ, Dopazo J, Dorff KC, Elloumi F, Fan J, Fan S, Fan X, Fang H, Gonzaludo N, Hess KR, Hong H, Huan J, Irizarry RA, Judson R, Juraeva D, Lababidi S, Lambert CG, Li L, Li Y, Li Z, Lin SM, Liu G, Lobenhofer EK, Luo J, Luo W, McCall MN, Nikolsky Y, Pennello GA, Perkins RG, Philip R, Popovici V, Price ND, Qian F, Scherer A, Shi T, Shi W, Sung J, Thierry-Mieg D, Thierry-Mieg J, Thodima V, Trygg J, Vishnuvajjala L, Wang SJ, Wu J, Wu Y, Xie Q, Yousef WA, Zhang L, Zhang X, Zhong S, Zhou Y, Zhu S, Arasappan D, Bao W, Lucas AB, Berthold F, Brennan RJ, Buness A, Catalano JG, Chang C, Chen R, Cheng Y, et al. (2010) The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol 28:827–838. doi: 10.1038/nbt.1665 [PubMed: 20676074]

51. Popovici V, Chen W, Gallas BG, Hatzis C, Shi W, Samuelson FW, Nikolsky Y, Tsyganova M, Ishkin A, Nikolskaya T, Hess KR, Valero V, Booser D, Delorenzi M, Hortobagyi GN, Shi L, Symmans WF, Pusztai L (2010) Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. Breast Cancer Res 12:R5. doi: 10.1186/bcr2468 [PubMed: 20064235]

52. Yee D, Isaacs C, Wolf DM, Yau C, Haluska P, Giridhar KV, Forero-Torres A, Jo Chien A, Wallace AM, Pusztai L, Albain KS, Ellis ED, Beckwith H, Haley BB, Elias AD, Boughey JC, Kemmer K, Yung RL, Pohlmann PR, Tripathy D, Clark AS, Han HS, Nanda R, Khan QJ, Edmiston KK, Petricoin EF, Stringer-Reasor E, Falkson CI, Majure M, Mukhtar RA, Helsten TL, Moulder SL, Robinson PA, Wulfkuhle JD, Brown-Swigart L, Buxton M, Clennell JL, Paoloni M, Sanil A, Berry S, Asare SM, Wilson A, Hirst GL, Singhrao R, Asare AL, Matthews JB, Hylton NM, DeMichele A, Melisko M, Perlmutter J, Rugo HS, Fraser Symmans W, Van't Veer LJ, Berry DA, Esserman LJ (2021) Ganitumab and metformin plus standard neoadjuvant therapy in stage 2/3 breast cancer. NPJ Breast Cancer 7:131. doi: 10.1038/s41523-021-00337-2 [PubMed: 34611148]

53. Gluck S, Ross JS, Royce M, McKenna EF Jr., Perou CM, Avisar E, Wu L (2012) TP53 genomics predict higher clinical and pathologic tumor response in operable early-stage breast cancer treated with docetaxel-capecitabine +/– trastuzumab. Breast Cancer Res Treat 132:781–791. doi: 10.1007/s10549-011-1412-7 [PubMed: 21373875]

54. Esserman LJ, Berry DA, Cheang MC, Yau C, Perou CM, Carey L, DeMichele A, Gray JW, Conway-Dorsey K, Lenburg ME, Buxton MB, Davis SE, van't Veer LJ, Hudis C, Chin K, Wolf D, Krontiras H, Montgomery L, Tripathy D, Lehman C, Liu MC, Olopade OI, Rugo HS, Carpenter JT, Livasy C, Dressler L, Chhieng D, Singh B, Mies C, Rabban J, Chen YY, Giri D, Au A, Hylton N, Investigators IST (2012) 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 132:1049–1062. doi: 10.1007/s10549-011-1895-2 [PubMed: 22198468]

55. Chen JW, Russell RP, Desai T, Fiel-Gan M, Bhat V, de Fátima Dias Gaui M, Amendola LC, Vasconcelos Z, Brufsky AM, Fournier MV, Tannenbaum SH (2021) RNA expression classifiers from a model of breast epithelial cell organization to predict pathological complete response in triple negative breast cancer. medRxiv:2021.2002.2010.21251517. doi: 10.1101/2021.02.10.21251517

56. Kersten K, Coffelt SB, Hoogstraat M, Verstegen NJM, Vrijland K, Ciampricotti M, Doornebal CW, Hau CS, Wellenstein MD, Salvagno C, Doshi P, Lips EH, Wessels LFA, de Visser KE (2017) Mammary tumor-derived CCL2 enhances pro-metastatic systemic inflammation through upregulation of IL1beta in tumor-associated macrophages. Oncoimmunology 6:e1334744. doi: 10.1080/2162402X.2017.1334744 [PubMed: 28919995]

57. de Ronde JJ, Lips EH, Mulder L, Vincent AD, Wesseling J, Nieuwland M, Kerkhoven R, Vrancken Peeters MJ, Sonke GS, Rodenhuis S, Wessels LF (2013) SERPINA6, BEX1, AGTR1, SLC26A3, and LAPTM4B are markers of resistance to neoadjuvant chemotherapy in HER2-

negative breast cancer. Breast Cancer Res Treat 137:213–223. doi: 10.1007/s10549-012-2340-x [PubMed: 23203637]

58. Juul N, Szallasi Z, Eklund AC, Li Q, Burrell RA, Gerlinger M, Valero V, Andreopoulou E, Esteva FJ, Symmans WF, Desmedt C, Haibe-Kains B, Sotiriou C, Pusztai L, Swanton C (2010) Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials. Lancet Oncol 11:358–365. doi: 10.1016/S1470-2045(10)70018-8 [PubMed: 20189874]

59. Oshi M, Gandhi S, Huyser MR, Tokumaru Y, Yan L, Yamada A, Matsuyama R, Endo I, Takabe K (2021) MELK expression in breast cancer is associated with infiltration of immune cell and pathological compete response (pCR) after neoadjuvant chemotherapy. American journal of cancer research 11:4421–4437. [PubMed: 34659896]

60. Oshi M, Tokumaru Y, Mukhopadhyay S, Yan L, Matsuyama R, Endo I, Takabe K (2021) Annexin A1 Expression Is Associated with Epithelial-Mesenchymal Transition (EMT), Cell Proliferation, Prognosis, and Drug Response in Pancreatic Cancer. Cells 10:653. doi: 10.3390/cells10030653 [PubMed: 33804148]

61. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102:15545–15550. doi: 10.1073/pnas.0506580102 [PubMed: 16199517]

62. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 1:417–425. doi: 10.1016/j.cels.2015.12.004 [PubMed: 26771021]

63. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Mokrab Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS, Cancer Genome Atlas Research N, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG, Shmulevich I (2019) The Immune Landscape of Cancer. Immunity 51:411–412. doi: 10.1016/j.immuni.2019.08.004 [PubMed: 31433971]

64. Aran D, Hu Z, Butte AJ (2017) xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 18:220. doi: 10.1186/s13059-017-1349-1 [PubMed: 29141660]

65. Oshi M, Asaoka M, Tokumaru Y, Yan L, Matsuyama R, Ishikawa T, Endo I, Takabe K (2020) CD8 T Cell Score as a Prognostic Biomarker for Triple Negative Breast Cancer. International journal of molecular sciences 21:6968. doi: 10.3390/ijms21186968

66. Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Kalinski P, Endo I, Takabe K (2020) Plasmacytoid Dendritic Cell (pDC) Infiltration Correlate with Tumor Infiltrating Lymphocytes, Cancer Immunity, and Better Survival in Triple Negative Breast Cancer (TNBC) More Strongly than Conventional Dendritic Cell (cDC). Cancers 12:3342. doi: 10.3390/cancers12113342

67. Takahashi H, Asaoka M, Yan L, Rashid OM, Oshi M, Ishikawa T, Nagahashi M, Takabe K (2020) Biologically Aggressive Phenotype and Anti-cancer Immunity Counterbalance in Breast Cancer with High Mutation Rate. Scientific reports 10:1852. doi: 10.1038/s41598-020-58995-4 [PubMed: 32024876]

68. Le L, Tokumaru Y, Oshi M, Asaoka M, Yan L, Endo I, Ishikawa T, Futamura M, Yoshida K, Takabe K (2021) Th2 cell infiltrations predict neoadjuvant chemotherapy response of estrogen receptor-positive breast cancer. Gland surgery 10:154–165. doi: 10.21037/gs-20-571 [PubMed: 33633972]

69. Martin RW, Orelli BJ, Yamazoe M, Minn AJ, Takeda S, Bishop DK (2007) RAD51 up-regulation bypasses BRCA1 function and is a common feature of BRCA1-deficient breast tumors. Cancer Res 67:9658–9665. doi: 10.1158/0008-5472.CAN-07-0290 [PubMed: 17942895]

70. Wiegmans AP, Al-Ejeh F, Chee N, Yap PY, Gorski JJ, Da Silva L, Bolderson E, Chenevix-Trench G, Anderson R, Simpson PT, Lakhani SR, Khanna KK (2014) Rad51 supports triple negative

breast cancer metastasis. Oncotarget 5:3261–3272. doi: 10.18632/oncotarget.1923 [PubMed: 24811120]

71. Banerjee S, Kaye SB, Ashworth A (2010) Making the best of PARP inhibitors in ovarian cancer. Nat Rev Clin Oncol 7:508–519. doi: 10.1038/nrclinonc.2010.116 [PubMed: 20700108]

72. Bonilla B, Hengel SR, Grundy MK, Bernstein KA (2020) RAD51 Gene Family Structure and Function. Annu Rev Genet 54:25–46. doi: 10.1146/annurev-genet-021920-092410 [PubMed: 32663049]

73. Zeman MK, Cimprich KA (2014) Causes and consequences of replication stress. Nat Cell Biol 16:2–9. doi: 10.1038/ncb2897 [PubMed: 24366029]

74. Orhan E, Velazquez C, Tabet I, Sardet C, Theillet C (2021) Regulation of RAD51 at the Transcriptional and Functional Levels: What Prospects for Cancer Therapy? Cancers 13. doi: 10.3390/cancers13122930

75. Graeser M, McCarthy A, Lord CJ, Savage K, Hills M, Salter J, Orr N, Parton M, Smith IE, Reis-Filho JS, Dowsett M, Ashworth A, Turner NC (2010) A marker of homologous recombination predicts pathologic complete response to neoadjuvant chemotherapy in primary breast cancer. Clin Cancer Res 16:6159–6168. doi: 10.1158/1078-0432.CCR-10-1027 [PubMed: 20802015]

76. Llop-Guevara A, Loibl S, Villacampa G, Vladimirova V, Schneeweiss A, Karn T, Zahm DM, Herencia-Ropero A, Jank P, van Mackelenbergh M, Fasching PA, Marme F, Stickeler E, Schem C, Dienstmann R, Florian S, Nekljudova V, Balmana J, Hahnen E, Denkert C, Serra V (2021) Association of RAD51 with Homologous Recombination Deficiency (HRD) and clinical outcomes in untreated triple-negative breast cancer (TNBC): analysis of the GeparSixto randomized clinical trial. Ann Oncol. doi: 10.1016/j.annonc.2021.09.003
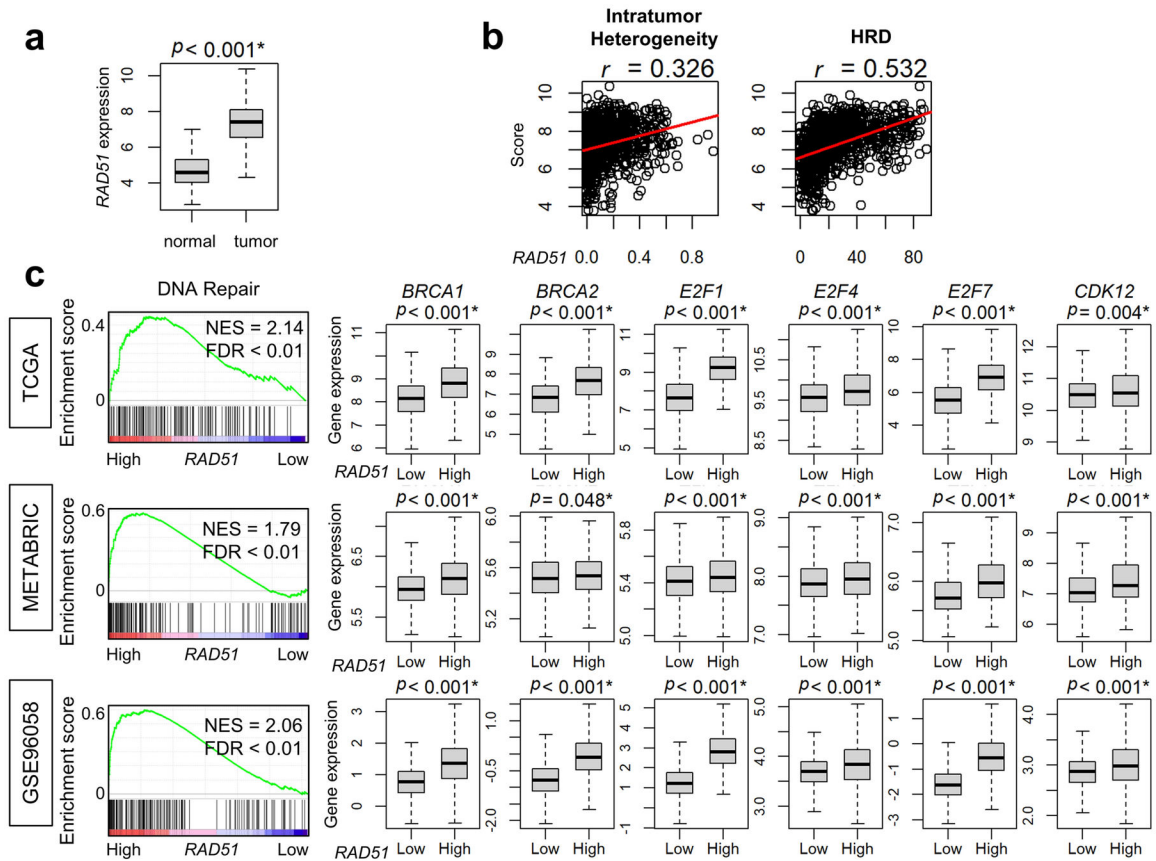
Figure 1. Association between *RAD51* gene expression and DNA repair.
(a) *RAD51* expression between normal breast and tumor tissues in TCGA. n; normal breast, t; tumor tissues. **(b)** The scatter plots between *RAD51* gene expression and Intratumoral heterogeneity (left) and Homologues Recombination Deficiency (HRD) score (right) in TCGA. **(c)** The enrichment plots of DNA repair pathway in gene enrichment analysis (GSEA) comparing high vs low expression of *RAD51* divided by a median cut-off in TCGA, METABRIC, and GSE96058 cohorts. The boxplots show the expression of DNA repair-related genes; *BRCA1*, *BRCA2*, *E2F1*, *E2F4*, *E2F7*, and *CDK12* by high vs low expression of *RAD51*. FDR less than 0.25 is regarded as significant in GSEA. * = p-value of statistical significance. The r-value indicates Spearman's rank correlation coefficient. All two group comparisons are tested by Wilcoxon signed-rank test. The error bars in each boxplot show the 95% confidence interval. The line in the box shows the median, and top and bottom show the 25th and 75th percentiles respectively.
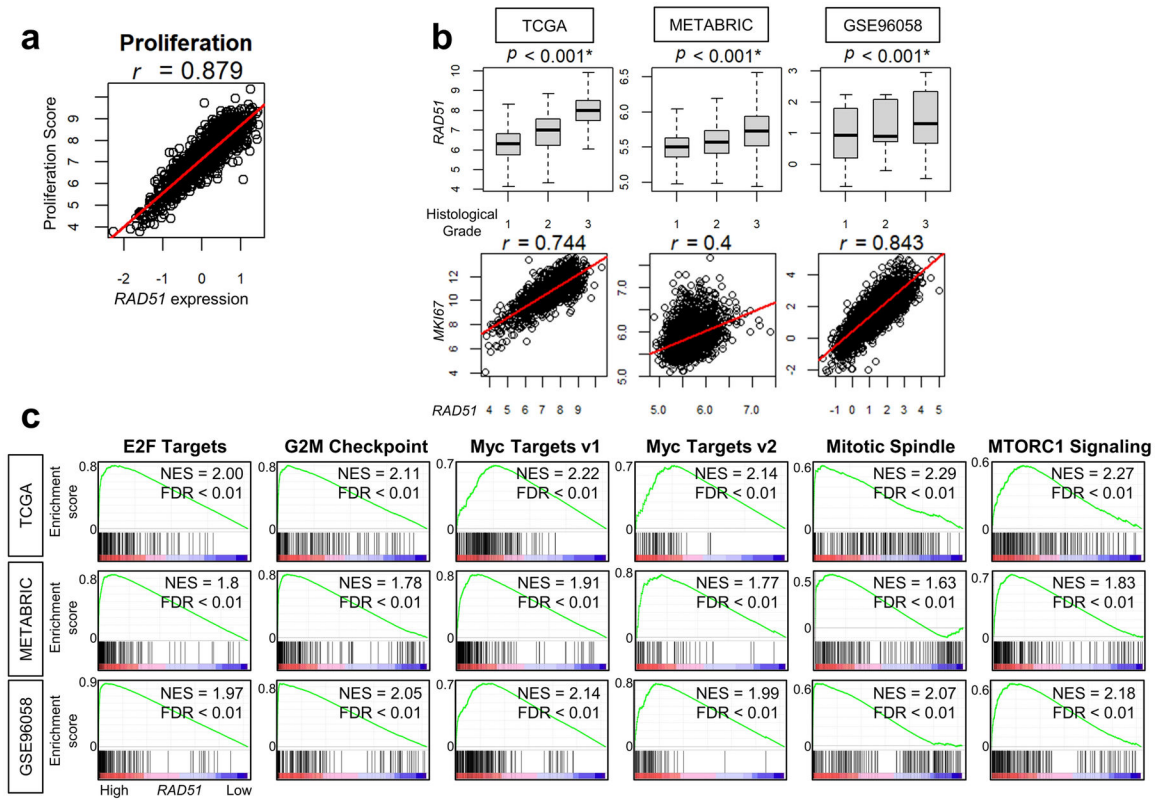
Figure 2. Association between *RAD51* and cancer cell proliferation.
(a) The scatter plot of *RAD51* gene expression and proliferation score in TCGA. **(b)** The boxplots of *RAD51* gene expression by Nottingham histological grade in TCGA, METABRIC, and GSE96058 cohorts. The scatter plots of *MKI67* and *RAD51* gene expressions. **(c)** GSEA of all cell proliferation related gene sets by the high and low expression of *RAD51* with a median cut-off in TCGA, METABRIC, and GSE96058 cohort. FDR less than 0.25 is regarded as significant in GSEA. *= p-value of statistical significance. The r-value indicates Spearman's rank correlation coefficient. All multiple group comparisons are tested by Kruskal–Wallis test. The error bars in each boxplot show the 95% confidence interval. The line in the box shows the median, and top and bottom show the 25th and 75th percentiles respectively.
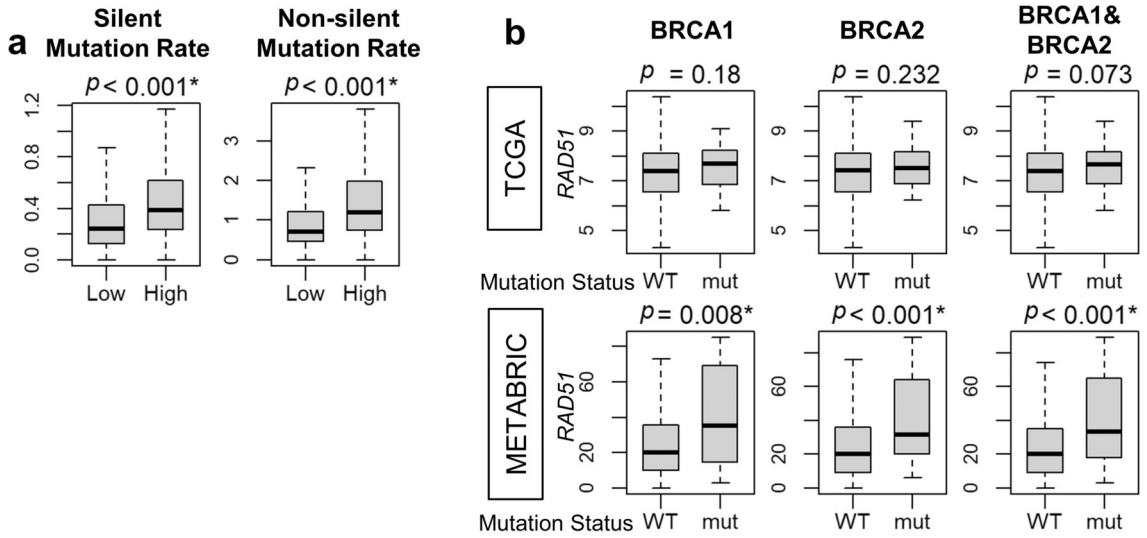
Figure 3. Association of *RAD51* expression with mutation rates and BRCA mutations.
(a) The boxplots of silent and non-silent mutation rate by the high and low *RAD51* expression with a median cut-off in TCGA. **(b)** The boxplots of *RAD51* gene expression in *BRCA1*, *BRCA2*, and both wild-type and mutant breast cancer in TCGA and METABRIC. The number of patients with BRCA1, BRCA2 or either mutation was 30, 29, 57 in TCGA, and 55, 60, 114 in METABRIC. *= p-value of statistical significance. All two group comparisons are tested by Wilcoxon signed-rank test. The error bars in each boxplot show the 95% confidence interval. The line in the box shows the median, and top and bottom show the 25th and 75th percentiles respectively.
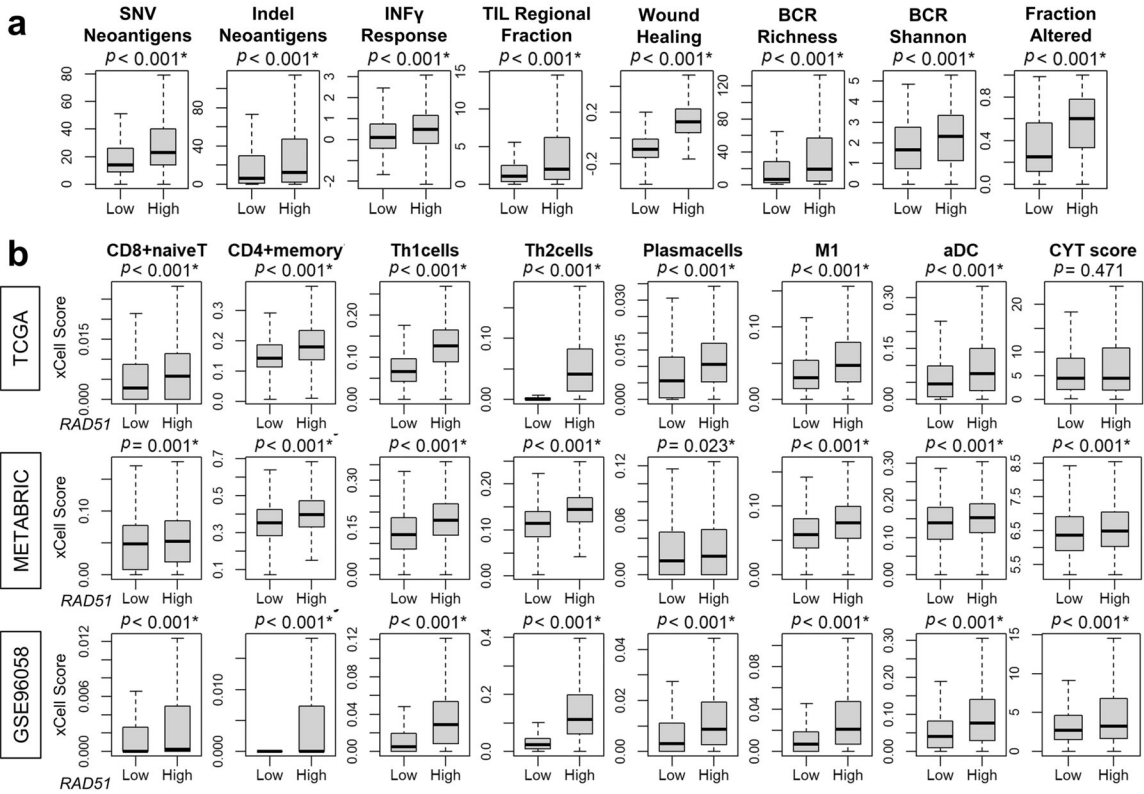
Figure 4. *RAD51* expression and immune activation and immune cell infiltration.
(a) The boxplots show immune activity scores from TCGA. **(b)** Immune cell infiltrations
by the high and low *RAD51* expression with a median cut-off in TCGA, METABRIC,
and GSE96058 cohort. *= p-value of statistical significance. All two group comparisons
are tested by Wilcoxon signed-rank test. The error bars in each boxplot show the 95%
confidence interval. The line in the box shows the median, and top and bottom show the 25th
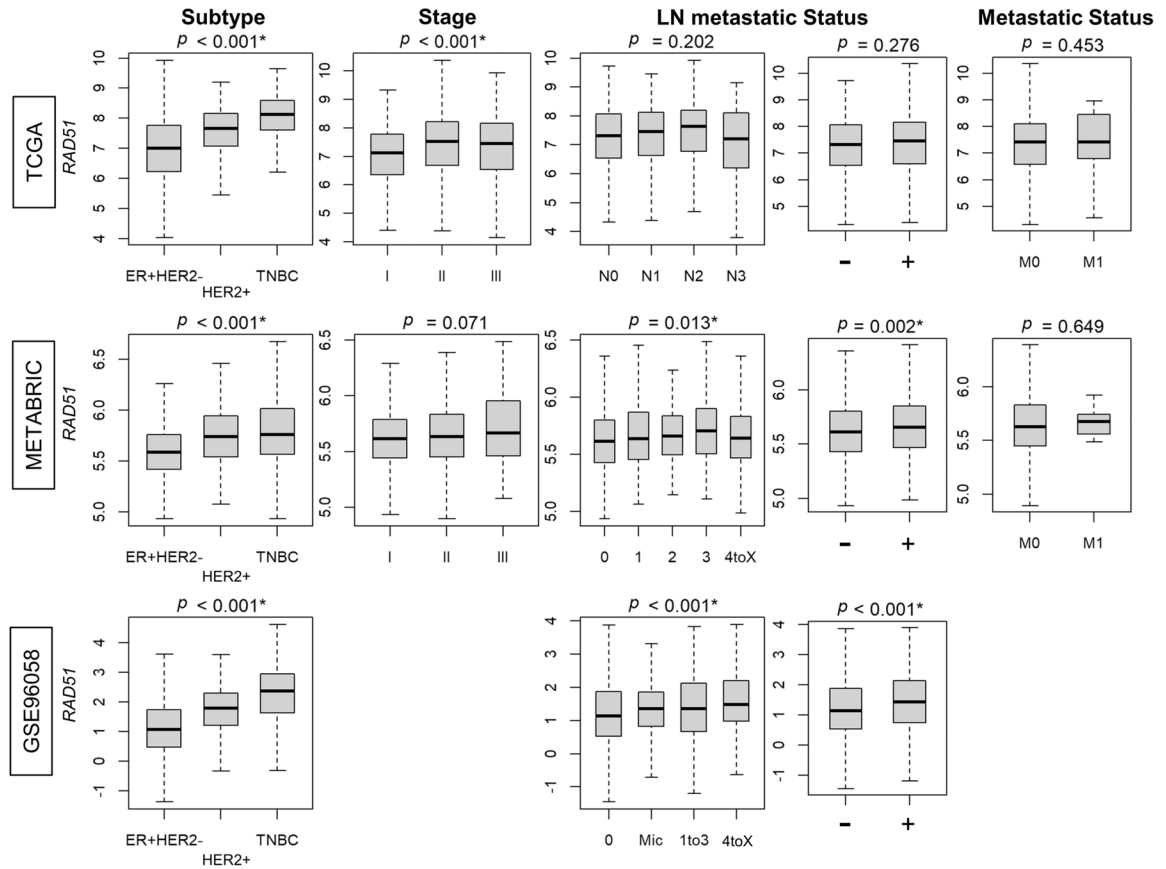and 75th percentiles respectively.

Figure 5. Relationship between *RAD51* and clinicopathological factors.

RAD51 gene expression by immunohistochemical subtype, stage, lymph node metastasis, and distant metastasis in TCGA, METABRIC, and GSE96058 cohorts. *= p-value of statistical significance. All two group comparisons are tested by Wilcoxon signed-rank test, and multiple groups by Kruskal–Wallis test. The error bars in each boxplot show the 95% confidence interval. The line in the box shows the median, and top and bottom show the 25th and 75th percentiles respectively.
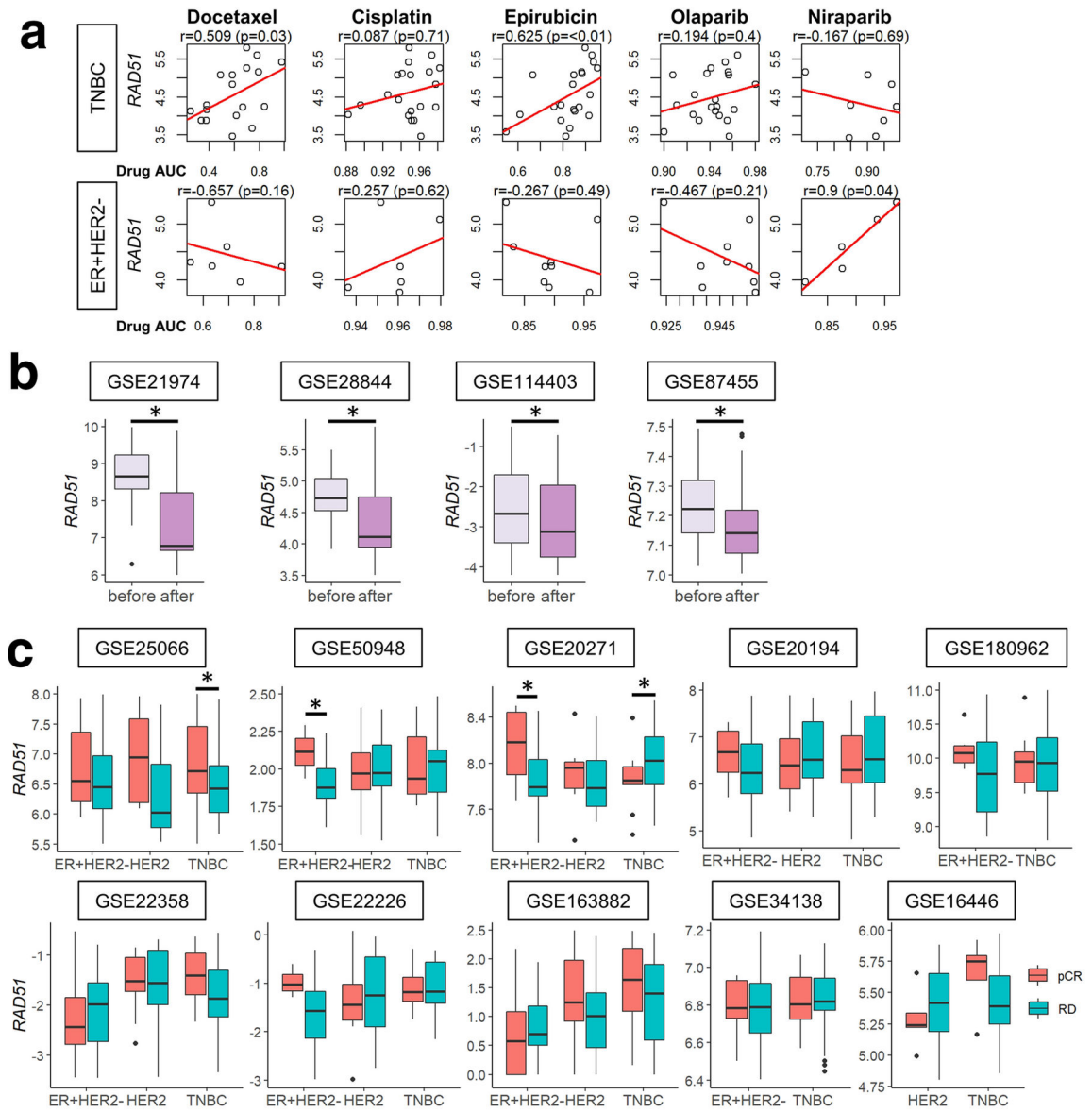
Figure 6. Relationship between *RAD51* and drug response in breast cancer.
(**a**) The scatter plots of correlation between *RAD51* expression and area under the curve (AUC) of each drug. Docetaxel and cisplatin are from the PRISM primary screen; PARP inhibitors are from the GDSC. (b) Boxplots show *RAD51* expression before (light purple boxes) and after (dark purple boxes) NAC. (**c**) All boxplots compare *RAD51* expression by immunohistochemical subtype between the two groups, orange for pCR: pathological complete response and light green for RD: residual tumor. *= p-value of statistical significance. The r-value indicates Spearman's rank correlation coefficient. All two group comparisons are tested by Wilcoxon signed-rank test. The error bars in each boxplot show the 95% confidence interval. The line in the box shows the median, and top and bottom show the 25th and 75th percentiles respectively.
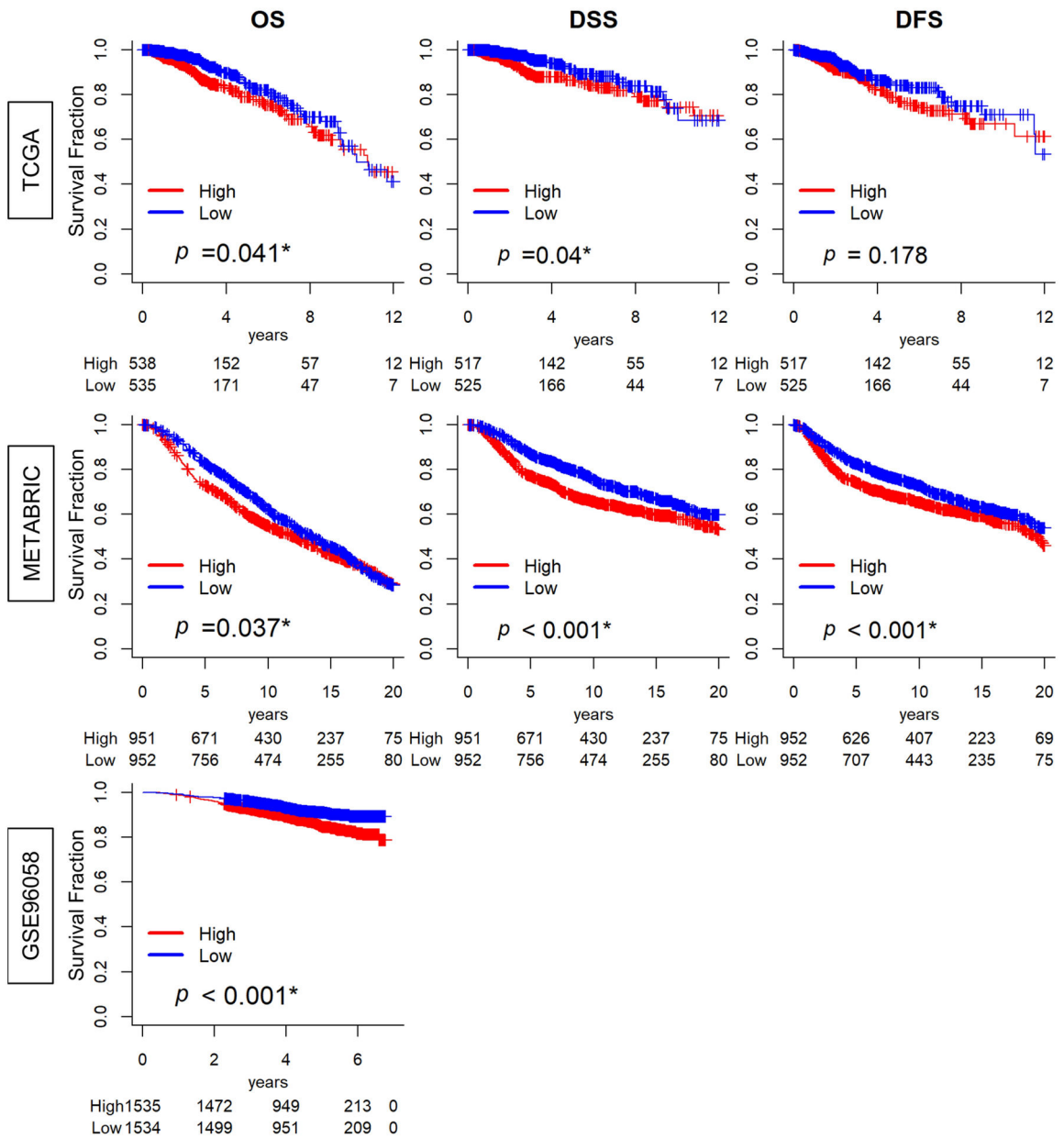
Figure 7. Survival analyses by *RAD51* expression.
Kaplan-Meier survival curves of the DFS, DSS and, OS by *RAD51* high vs. low expressions with a median cut-off in TCGA, METABRIC, and OS of the GSE96058 cohort. High groups are indicated by red lines, low groups by blue lines. *= p-value of statistical significance. Log-rank test was used to test the significance of the survival analysis. The r-value indicates Spearman's rank correlation coefficient.