Am J Cancer Res 2021;11(9):4421-4437
www.ajcr.us /ISSN:2156-6976/ajcr0135977

Original Article

MELK expression in breast cancer is associated with infiltration of immune cell and pathological compete response (pCR) after neoadjuvant chemotherapy

Masanori Oshi[1,2], Shipra Gandhi[3], Michelle R Huyser[1], Yoshihisa Tokumaru[1,4], Li Yan[5], Akimitsu Yamada[2], Ryusei Matsuyama[2], Itaru Endo[2], Kazuaki Takabe[1,2,6,7,8,9]

[1]Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14263, USA; [2]Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; [3]Department of Medical Oncology, Roswell Park Comprehensive Cancer Center, Elm & Carlton Streets, Buffalo, NY 14263, USA; [4]Department of Surgical Oncology, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan; [5]Department of Biostatistics & Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, New York 14263, USA; [6]Division of Digestive and General Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata 951-8520, Japan; [7]Department of Breast Surgery, Fukushima Medical University School of Medicine, Fukushima 960-1295, Japan; [8]Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, State University of New York, Buffalo, New York 14263, USA; [9]Department of Breast Surgery and Oncology, Tokyo Medical University, Tokyo 160-8402, Japan

Received June 3, 2021; Accepted August 20, 2021; Epub September 15, 2021; Published September 30, 2021

Abstract: In experimental settings, maternal embryonic leucine zipper kinase (*MELK*), an apical member of the snf1/AMPK serine-threonine kinases family, plays a role in tumor growth. We investigated the clinical relevance of *MELK* expression by performing silico analyses of 7,135 breast cancer patients using multiple independent large cohorts. In triple negative breast cancer (TNBC) found that elevated *MELK* expression significantly correlates with Nottingham histologic grade and tumor growth according to American Joint Committee Cancer (AJCC) stage. High *MELK* tumor enriched cell proliferation-related gene sets as well as DNA repair, unfolded protein response, and MTORC signaling gene sets. In two independent cohorts a high mutation rate and worse survival was significantly associated with high *MELK* tumor. In immune-related gene sets including, allograft rejection, interferon (IFN)-α response, and IFN-γ response, high *MELK* tumor significantly enriched. Pro-cancer regulatory T cells, T helper type 2 cells and anti-cancer immune cells including $CD4^+$ memory T cells, T helper type1 cells, $CD8^+$ T cells, M1 macrophages, gamma-delta T cells, and dendritic cells with high levels of cytolytic activity (CYT) were highly infiltrated. *MELK* expression did not correlate with the responses to any of the drugs tested in cell lines. However, pathologic complete response was significantly associated with high *MELK* following NAC in both TNBC and ER-positive plus HER2-negative breast cancer. In conclusion, cell proliferation, immune response, and NAC breast cancer response was associated with *MELK* expression.

Keywords: Breast cancer, drug response, ER+/HER2-, MELK, neoadjuvant chemotherapy, pathological complete response, predictive biomarker, tumor immune microenvironment

Introduction

Neoadjuvant chemotherapy (NAC), systemic chemotherapy which is given prior to definitive surgery, was originally intended to decrease the size of locally advanced or unresectable breast cancer to facilitate breast conserving surgery and reduce surgical morbidity [1]. In addition to the above, NAC is now viewed as an opportunity to assess the drug responses of the tumors in situ. Pathological complete response (pCR) after NAC in breast cancer, specifically aggressive triple-negative breast cancer (TNBC), has now been commonly used as a surrogate marker for predicting better long-term outcomes [1]. Knowledge of achievement of pCR or residual disease is vital for decision making in the adjuvant setting after surgery. When needed, neoadjuvant treatment escalation for TNBC, with use of drugs like carboplatin [2, 3] or pembroli-

zumab [4]/atezolizumab [5], can be associated with significant long-term toxicities including immune-related adverse events. Therefore, it is critical to appropriately select patients to recommend NAC (especially when treatment escalation is needed) where there is a low probability to achieve pCR to maximize the chances for a good response to improve outcomes and avoid additional chemotherapy in the adjuvant setting. Similarly, utilizing genomic assays in hormone-receptor positive breast cancer, such as, oncotypeDx or mammaprint can help predict the benefit of adjuvant chemotherapy. However, if a biomarker were available to identify the benefit of chemotherapy in the neoadjuvant setting, patients could potentially avoid non-beneficial chemotherapy prior to surgery. Breast cancer is the most common cancer affecting women worldwide, with a majority (70%) of them being Estrogen receptor (ER)-positive/human epidermal growth factor receptor 2 (HER2)-negative. Although this is the least aggressive breast cancer subtype, it has a poor response to NAC compared to TNBC where pCR has been observed in around 30%-40% tumors. Therefore, a novel biomarker that predicts NAC response and allows for precise patient selection will be helpful for both ER-positive and TNBC patients.

Maternal embryonic leucine zipper kinase *(MELK)*, an apical member of the snf1/AMPK serine-threonine kinases is highly expressed in several malignancies, including lung cancer, colorectal cancer, and breast cancer [6]. It plays a vital role in cell cycle and proliferation in cell culture settings [7-9]. In terms of its clinical relevance, *MELK* was reported as one of the genes that correlated with worse survival [10] but also showed the highest odds ratio for pCR after NAC in basal-like subtype among the 12 genes tested [11]. However, limitation of these previous studies is that they used high throughput approach that demonstrated the association, but no insight on the mechanism. Additionally, the role of *MELK* expression in a tumor microenvironment (TME), which plays a key role in regulating cancer progression and response to several drugs, is largely unknown since reproduction of human TME in experimental settings are difficult to say the least.

In silico approach is one of the powerful tools for studying clinical relevance of a gene expression in cancer, because it analyzes actual patient's cancer. Due to recent improvement and popularity of sequencing technology, algorithms have been developed that can capture the cancer-related signaling, and the status of immune cells and cancer-related stromal cells in TME using transcriptome data [12, 13]. For instance, we recently discovered that glucocorticoid receptor gene high expression was associated with higher infiltration of immune cells and better survival [14]. Similarly, high expression of Nuclear factor erythroid 2-related factor 2 (*NRF2*) was shown to be associated with high level of tumor infiltrating lymphocytes in ER-positive/HER2-negative breast cancer [15]. We, therefore, hypothesized that *MELK* expression association with clinical outcomes, cell proliferation and drug response in breast cancer using transcriptome data from thousands of patients.

Materials and methods

Breast cancer cohorts and associated data

The transcriptome and clinical data of breast cancer in the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) study (n = 1904) [16] was obtained from cBioPortal [17], and that of GSE96058 cohort (n = 3,273) was obtained from the Swedish Breast Cancer Analysis Network (SCAN-B) [18]. Other study, Shi et al. (GSE20194; n = 248, regimens; cyclophosphamide, doxorubicin, paclitaxel, and 5-fluorouracil) [19] and Symmans et al. (GSE25066; n = 508, regimen; anthracycline and taxane) [20], were also obtained from The Gene Expression Omnibus (GEO) repository. A data studied by Hess et al. (n = 133, regimens; cyclophosphamide, doxorubicin, fluorouracil, and paclitaxel) [21] was accessed from University of California Santa Cruz (UCSC) Xena. Mutation-related score; intratumor heterogeneity, homologous recombination deficiency, silent and non-silent mutation rate, fraction altered, single nucleotide variant (SNV) and indel neoantigens, were accessed from a study by Thorsson et al. [22] in The Cancer Genome Atlas (TCGA) Pan-Cancer study (TCGA-BRCA; n = 1069) [23], which selected female breast cancer patients, as we previously reported [24-27]. The $\log_2$-transform of gene expression data were used in all analyses.

Gene set expression analyses

To investigate biological function, low and high MELK expression groups in breast cancer were

compared using the gene set enrichment analysis (GSEA) [28] with hallmark gene sets of the Molecular Signatures Database (MSigDB) [29].

Immune cell fraction estimation

The fraction of several cell types, including immune and stromal cells, in the tumor microenvironment (TME) was calculated using the xCell algorithm [12]. This score was calculated using R software, as we previously reported [30-35].

Statistical analysis

R software (version 4.0.1) was used to conduct statistical analyses. Group comparisons were conducted using Fisher's exact test, Mann-Whitney U test, or Kruskal-Wallis test, accordingly. Survival analysis was conducted using Kaplan-Meier curves with log-rank test. We determined statistical significance using a *P*-value < 0.05.

Results

MELK expression significantly correlated with tumor growth and survival of breast cancer patients

Since *MELK* gene expression was linked with cell proliferation in cell culture and animal model systems, we investigated whether that is translatable to human breast cancer patients. In the METABRIC cohort, we found that higher *MELK* expression significantly correlated with higher stages of American Joint Committee on Cancer (AJCC) (**Figure 1A**; $P < 0.001$). Also in both METABRIC and and GSE96058 cohorts, Nottingham histologic grade correlated with higher *MELK* expression (**Figure 1A**; both $P < 0.001$). Triple-negative breast cancer (TNBC), which is the most aggressive subtype of breast cancer, was noted consistently in both cohorts to have the highest *MELK* expression compared to the other subtypes (**Figure 1A**; all $P < 0.001$). Additionally, *MELK* expression strongly correlated with *MKI67* expression, which is a well-established cell proliferation marker used in clinics, consistently in both cohorts (Spearman rank correlation (r) = 0.704 and 0.888, respectively, both $P < 0.001$).

When top one-third *MELK* gene expression level within a cohort was determined as high *MELK* group, we found in both cohorts that high *MELK* breast cancer was strongly enriched for all five cell proliferation-related gene sets including Mitotic spindle by GSEA, MYC targets v1 and v2, G2M checkpoint, and E2F targets within the MSigDB Hallmark (**Figure 1C**; all normalized enrichment score (NES) > 1.70, all false discovery rate (FDR) < 0.01 in both cohorts).

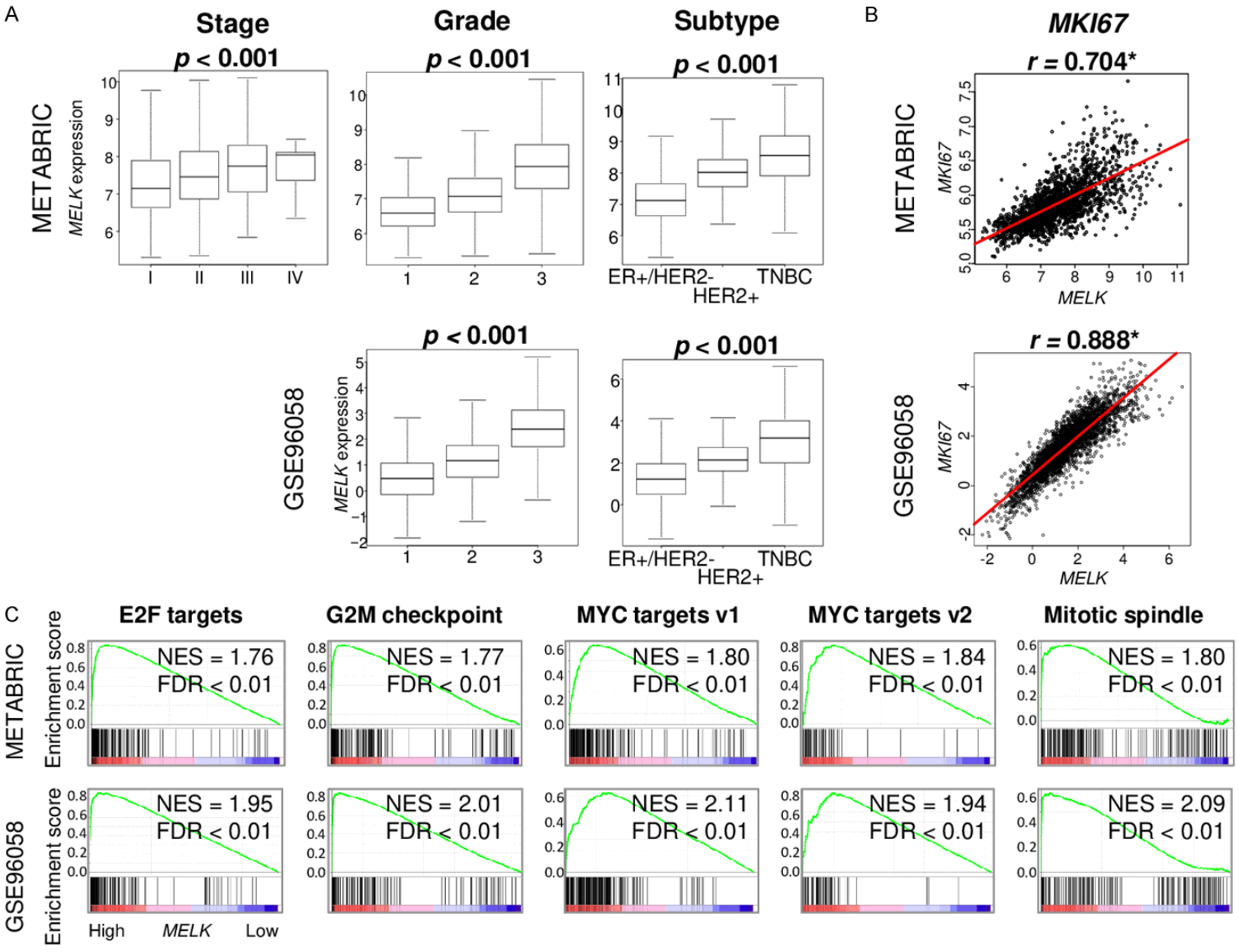
Next, the role of MELK expression in survival was investigated. We found a significant association with worse disease-free and disease-specific survival in high *MELK* breast cancer, within the METABRIC cohort, (**Figure 1D**; both $P < 0.001$) and a similar trend was observed in the GSE96058 cohort with overall survival (OS) in (**Figure 1D**; $P < 0.001$). Therefore, we observed that *MELK* expression was significantly associated with tumor growth assessed clinically (stage), pathologically (grade), and molecular biologically (*MKI67* expression and GSEA). Overall in breast cancer, these findings translated into poor clinical outcomes with high *MELK*.

There was no correlation between MELK expression and drug response in breast cancer cell lines

Given the fact that there was enhanced cell proliferation associated with high *MELK* expression in breast cancer, as evident by histological grade, *MKI67* expression and gene set enrichment, we expected that *MELK* expression would correlate with response to chemotherapy because chemotherapy is effective on high proliferating cells. The correlation of *MELK* expression of breast cancer cell lines with drug sensitivity area under the curve (AUC) for commonly used chemotherapies in breast cancer, namely, carboplatin, cisplatin, cyclophosphamide, doxorubicin, paclitaxel, and 5-FU, from DepMap portal was assessed. We found no correlation between *MELK* expression and AUC levels in ER-positive/HER2-negative breast cancer nor in TNBC (**Figure 2**).

Intratumor heterogeneity, homologous recombination deficiency (HRD), mutation load, and cancer aggressiveness-related gene sets were significantly associated with high MELK breast cancer

Given that high *MELK* breast cancer was significantly associated with worse patient survival, it was of interest to investigate the underlying


A
Stage
p < 0.001
Grade
p < 0.001
Subtype
p < 0.001
METABRIC
MELK expression
I
II
III
IV
1
2
3
ER+/HER2-
HER2+
TNBC
GSE96058
B
MKI67
r = 0.704*
r = 0.888*
MKI67
MELK
C
E2F targets
G2M checkpoint
MYC targets v1
MYC targets v2
Mitotic spindle
NES = 1.76
FDR < 0.01
NES = 1.77
FDR < 0.01
NES = 1.80
FDR < 0.01
NES = 1.84
FDR < 0.01
NES = 1.80
FDR < 0.01
NES = 1.95
FDR < 0.01
NES = 2.01
FDR < 0.01
NES = 2.11
FDR < 0.01
NES = 1.94
FDR < 0.01
NES = 2.09
FDR < 0.01
Enrichment score
High
MELK
Low

D

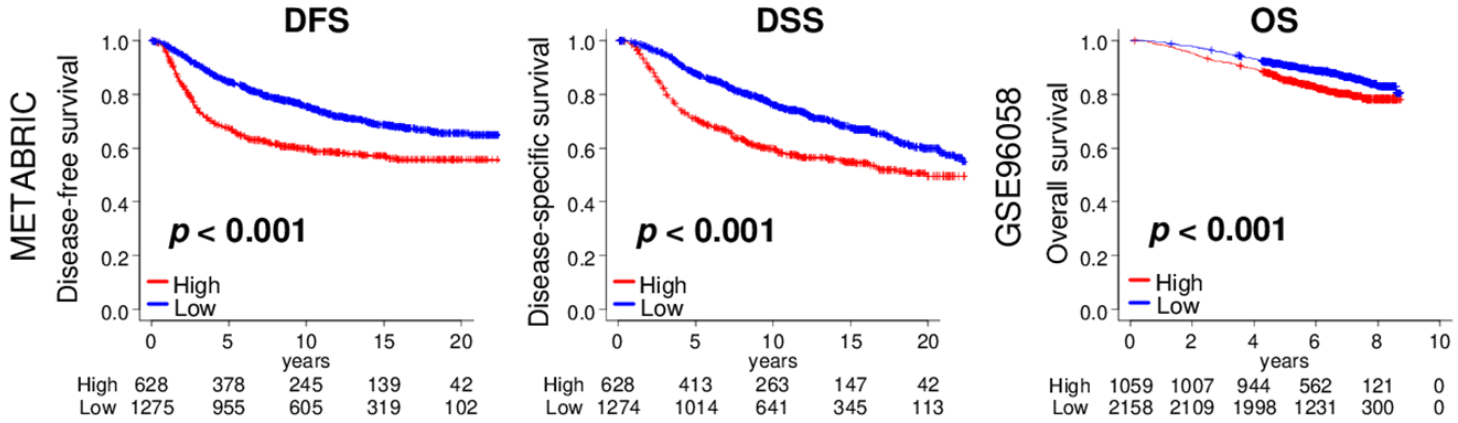


Figure 1. Association of the *MELK* expression with clinicopathological and biological aggressiveness in breast cancer. A. Using METABRIC and GSE96058 cohorts, boxplots of the *MELK* expression showing medians and interquartile ranges by breast cancer subtypes using the American Joint Committee on Cancer (AJCC) staging and Nottingham histological grade (GSE96058 cohort don't have the AJCC stage data). Analysis was performed using the Kruskal-Wallis test. B. Correlation plots between *MELK* and *MKI67* expression by spearman rank correlation. *$P < 0.01$. C. Enrichment plots of cell proliferation-related hallmark gene sets comparing high and low *MELK* expression groups, including false discovery rate (FDR) and normalized enrichment score (NES). D. Kaplan-Meier plots in METABRIC showing disease-free survival (DFS) and disease-specific survival (DSS), also in the GSE96058 cohort overall survival (OS) between high and low *MELK* expression groups for with log-rank test *p*-value.

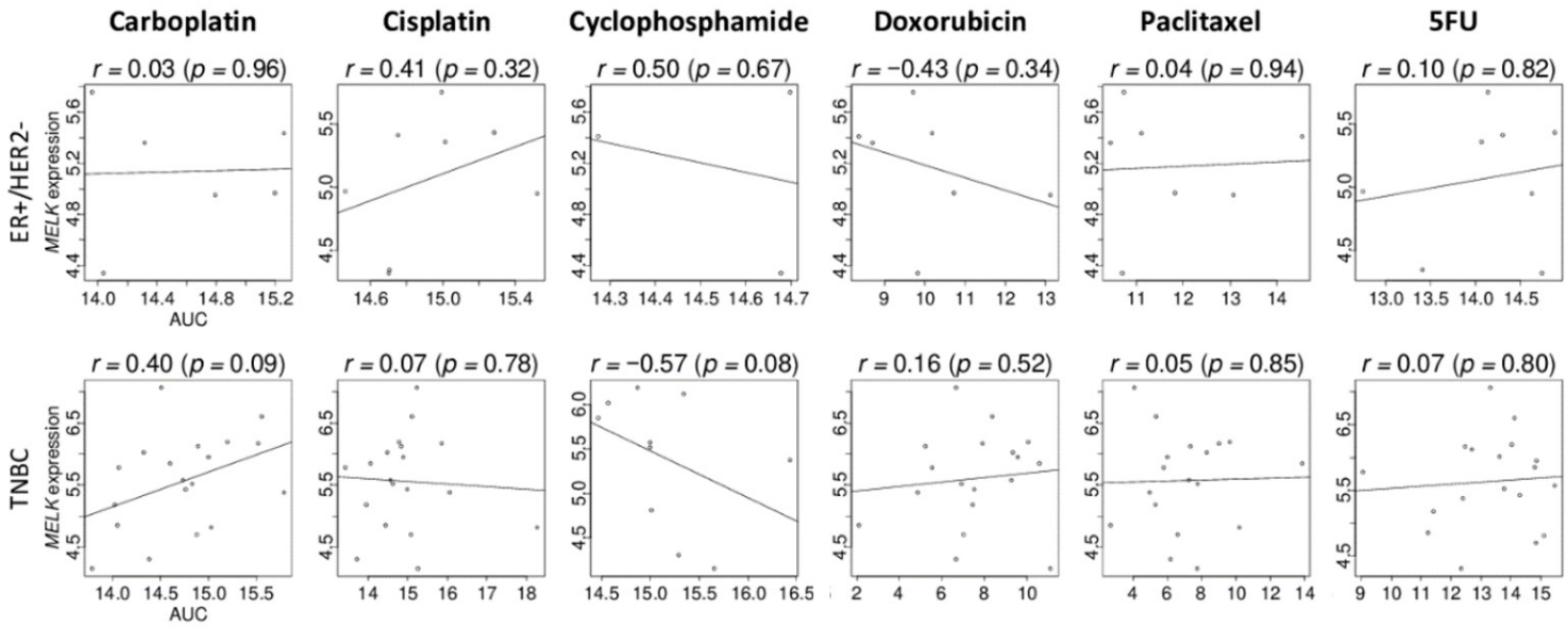


Figure 2. Correlation of the *MELK* expression with drug sensitivity of different chemotherapies in multiple cell lines. Correlation plots of *MELK* expression with drug sensitivity using area under the curve (AUC) of carboplatin, cisplatin, cyclophosphamide, doxorubicin, paclitaxel, and 5-FU in ER-positive/HER2-negative breast cancer and TNBC cell lines from the DepMap portal. The analysis was performed using Spearman rank correlation.

mechanism involved. Given that breast cancer with high mutation rate was associated with aggressive cancer [36], we expected *MELK* expression to correlate with mutation levels. High *MELK* breast cancer, in the TCGA cohort, was noted to be significantly associated with high level of intratumor heterogeneity, HRD, silent and non-silent mutation rate, fraction altered, single-nucleotide variant (SNV) and indel neoantigens score (**Figure 3A**; all $p < 0.03$). Using METABRIC and GSE96058 cohorts, we found that high *MELK* breast cancer was significantly enriched in cancer cell survival pathways including MTORC1 signaling, DNA repair, and unfolded protein response gene sets (**Figure 3B**, all NES > 1.50). Therefore, we observe that high mutation rate and enhanced cancer cell survival pathways were significantly associated with *MELK* expression.

High MELK breast cancer was associated with enhanced immune response and significantly enriched for immune-related gene sets

High mutation rate in breast cancer we previously reported as associated with increased cell proliferation and anti-cancer immunity [36]. Additionally, we also reported enhanced immune response association with high histological grade breast cancer [35]. Given that high *MELK* breast cancer was associated with high mutation rate as well as high histological grade, we expected that it would also be associated with enhanced immune response. As expected, using calculation provided by Thorsson et al on TCGA we found a significant association with high levels of immune-related scores, including tumor lymphocyte infiltration, infiltrating lymphocyte (TIL) regional fraction, T cell receptor (TCR), leukocyte fraction, B cell receptor (BCR) richness, and interferon (IFN)-γ response (**Figure 4A**). Further, we found that high *MELK* breast cancer significantly enriched for allograft rejection, IFN-γ response and IFN-α response, which are all immune response-related gene sets, consistently in both cohorts (**Figure 4B**, all NES > 1.30). These findings suggest a significant association with immune cell infiltration as well as immune response in breast cancer *MELK* expression.

There was higher anti-cancer immune cell infiltration in breast cancer with high MELK expression

Given that *MELK* expression significantly enriched for immune-related gene sets, we expected TME to have an association with immune cell infiltration. However, we found that there was no difference in *MELK* expression between tumor and immune cells in a single cell-sequencing cohort (GSE75688), which suggests that *MELK* is not a mere immune cell marker (**Figure 5A**). Next, we investigated which fraction of cells were infiltrated in the TME of high *MELK* expression breast cancer using xCell algorithm. We found that except for fibroblasts, in both METABRIC and GSE96058 cohorts, high *MELK* breast cancer was significantly associated with less stromal cells (Figure S1). Further, in both the METABRIC and GSE96058 cohorts, we observed high infiltration of pro-cancer immune cells, such as T helper type2 (Th2) cells and regulatory T cells (Tregs). High infiltration also was seen in anti-cancer immune cells, including $CD4^+$ memory T cells, T helper type1 (Th1) cells, $CD8^+$ T cells, M1 macrophages, gamma-delta T (γδT) cells, and dendritic cells (DC) in high *MELK* breast cancer (**Figure 5B** and **5C**). In both cohorts, low fraction of neutrophils was associated with high *MELK* expression (**Figure 5B**). Furthermore, in both cohorts consistently there was high level of cytolytic activity score (CYT) observed in high *MELK* breast cancer (**Figure 5D**). Interestingly, high *MELK* in immune cells was significantly associated with high expression of immune response-related genes, including *IFNG*, *GZMA*, and *PRF1*, compared to low *MELK* group (Figure S3). On the other hand, no such difference was observed between low and high groups in tumor cells. These findings again suggest that high *MELK* was associated with immune cell infiltration and enhanced immune killing, but not the stromal cells.

After neoadjuvant chemotherapy (NAC), MELK expression predicts pathological complete response (pCR) in both ER-positive/HER2-negative and TNBC

We expected since *MELK* expression was significantly associated with both cell proliferation and immune cell infiltration in breast cancer, it may also associate with NAC treatment given the higher proliferative nature. We examined three independent NAC cohorts with different chemotherapy combinations; GSE25066 (n = 508, underwent anthracycline and taxane), GSE20194 (n = 248, underwent cyclophosphamide, doxorubicin, paclitaxel, and 5-fluorouracil), and HESS (n = 133, underwent cyclophos-

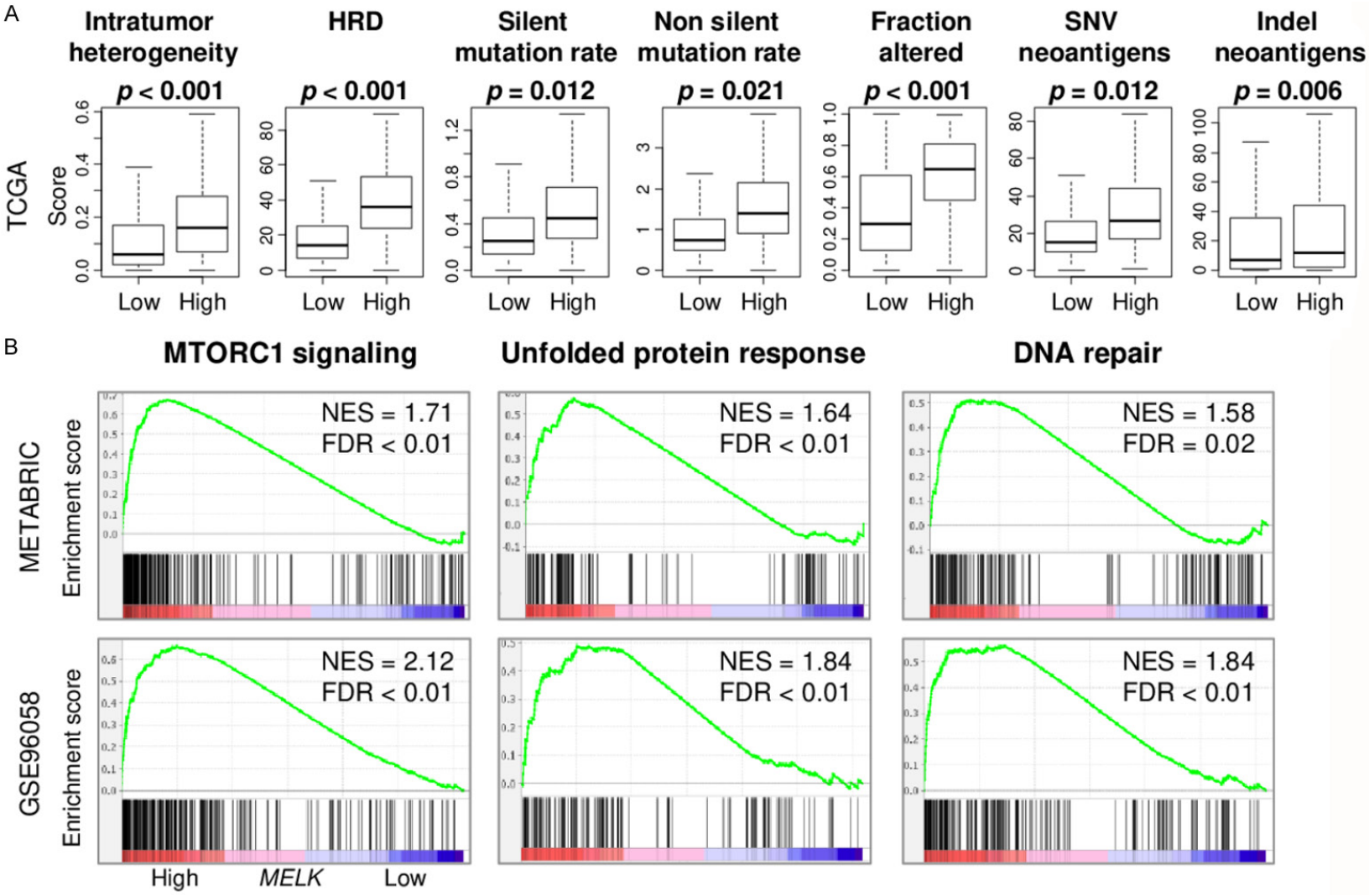


Figure 3. Association of the *MELK* expression with cancer aggressiveness-related factors in breast cancer. A. Boxplots of the intratumor heterogeneity, HRD, mutation rate, fraction altered, neoantigens showing medians and interquartile ranges for high and low *MELK* groups in the TCGA cohort. Analysis was performed using the Mann-Whitney U test. B. Enrichment plots of several malignant-related hallmark gene sets, including MTORC1 signaling, unfolded protein response, and DNA repair, which were significantly enriched in the high *MELK* groups consistently in both METABRIC and GSE96058 cohorts. FDR, false discovery rate; HRD, homologous recombination deficiency; NES, normalized enrichment score; SNV, single nucleotide.

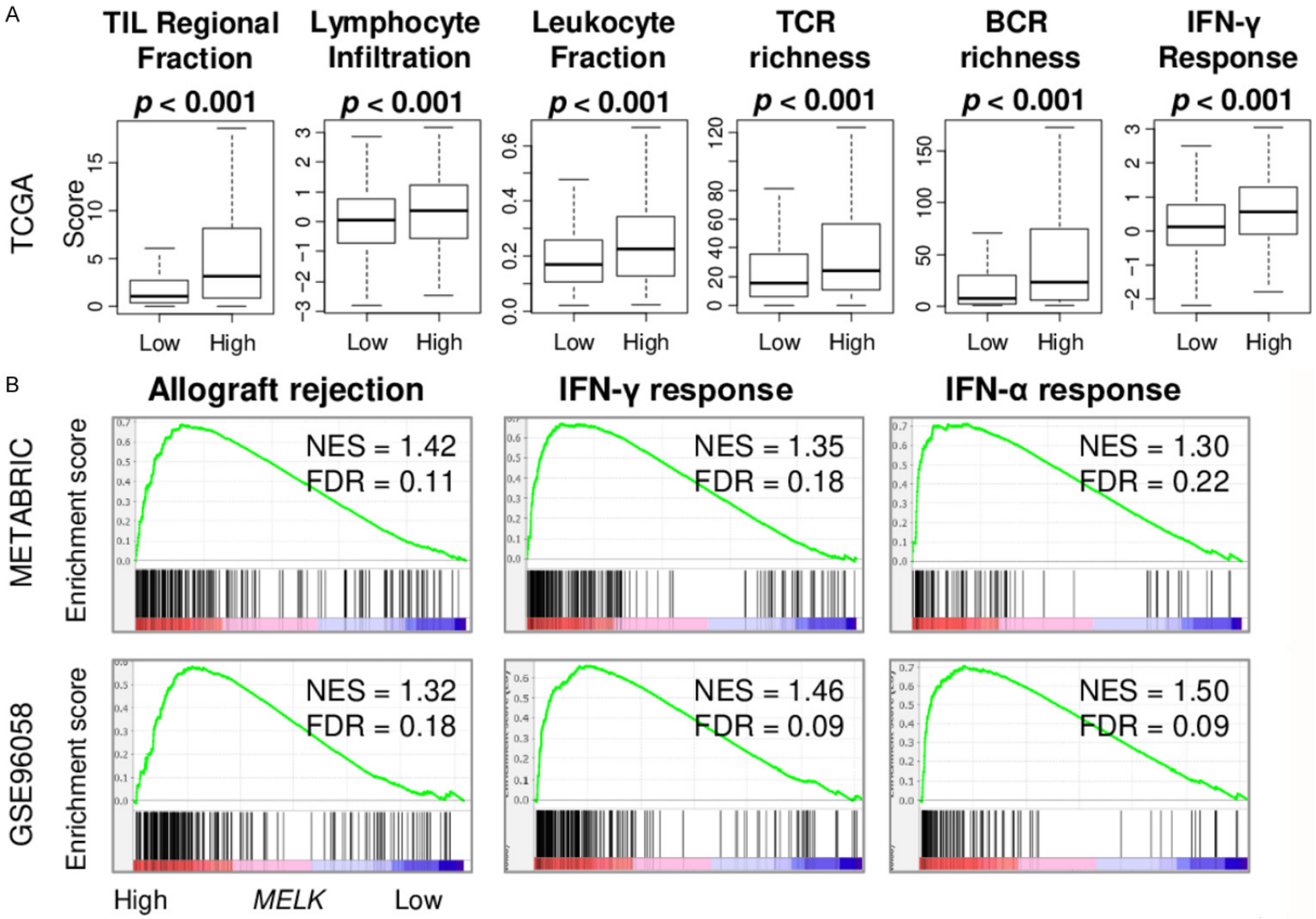


Figure 4. *MELK* expression and immune response in breast cancer. A. Box plots of the immune-related scores, including tumor infiltrating lymphocyte (TIL) regional fraction, leukocyte fraction, T cell receptor (TCR), lymphocyte infiltration, B cell receptor (BCR) richness, and IFN-γ response, showing medians and interquartile ranges between high and low *MELK* expression in the TCGA cohort in breast cancer. Analysis was performed using Mann-Whitney U test. B. Enrichment plots of immune-related hallmark gene sets, including allograft rejection, α response and interferon (IFN)-γ, were significantly enriched in the high *MELK* groups consistently in both METABRIC and GSE96058 cohorts.

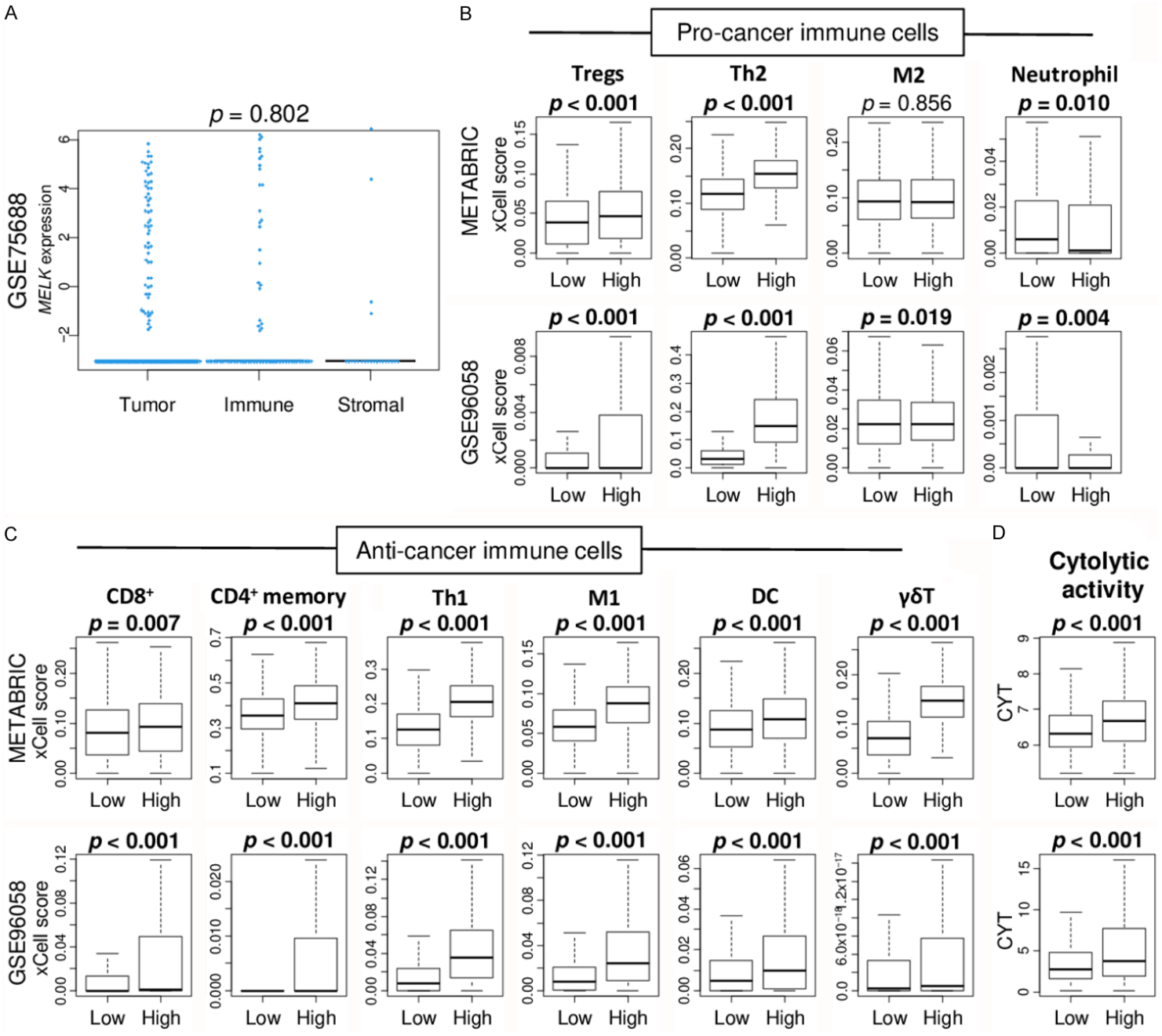

A
p = 0.802
GSE75688
MELK expression
Tumor
Immune
Stromal
B
Pro-cancer immune cells
Tregs
p < 0.001
Th2
p < 0.001
M2
p = 0.856
Neutrophil
p = 0.010
METABRIC
xCell score
p < 0.001
p < 0.001
p = 0.019
p = 0.004
GSE96058
xCell score
Low
High
C
Anti-cancer immune cells
CD8+
p = 0.007
CD4+ memory
p < 0.001
Th1
p < 0.001
M1
p < 0.001
DC
p < 0.001
γδT
p < 0.001
METABRIC
xCell score
p < 0.001
p < 0.001
p < 0.001
p < 0.001
p < 0.001
p < 0.001
GSE96058
xCell score
Low
High
D
Cytolytic activity
p < 0.001
CYT
p < 0.001
CYT
Low
High

Figure 5. *MELK* expression and association with immune cells in breast cancer. (A) Dot plots of the *MELK* expression by single cells of tumor, immune and stromal cells in the GSE75688 single-sequence cohort using Kruskal-Wallis test. (B) Boxplots showing medians and interquartile ranges of the infiltrating fraction of pro-cancer immune cells (T helper type2 [Th2] cells, regulatory T cells [Tregs], M2 macrophages [M2], and Neutrophil cells) and (C) anti-cancer immune cells ($CD4^+$ memory T cells, T helper type1 [Th1] cells, $CD8^+$ T cells, M1 macrophages [M1], gamma delta T cells [γδT], and dendritic cells [DC]) categorized by high and low *MELK* groups in the METABRIC and GSE96058 cohorts. (D) Boxplots, in both METABRIC and GSE96058 cohorts, of the cytolytic activity score (CYT) by high and low *MELK* groups. Mann-Whitney U test was used to calculate the *p*-values.

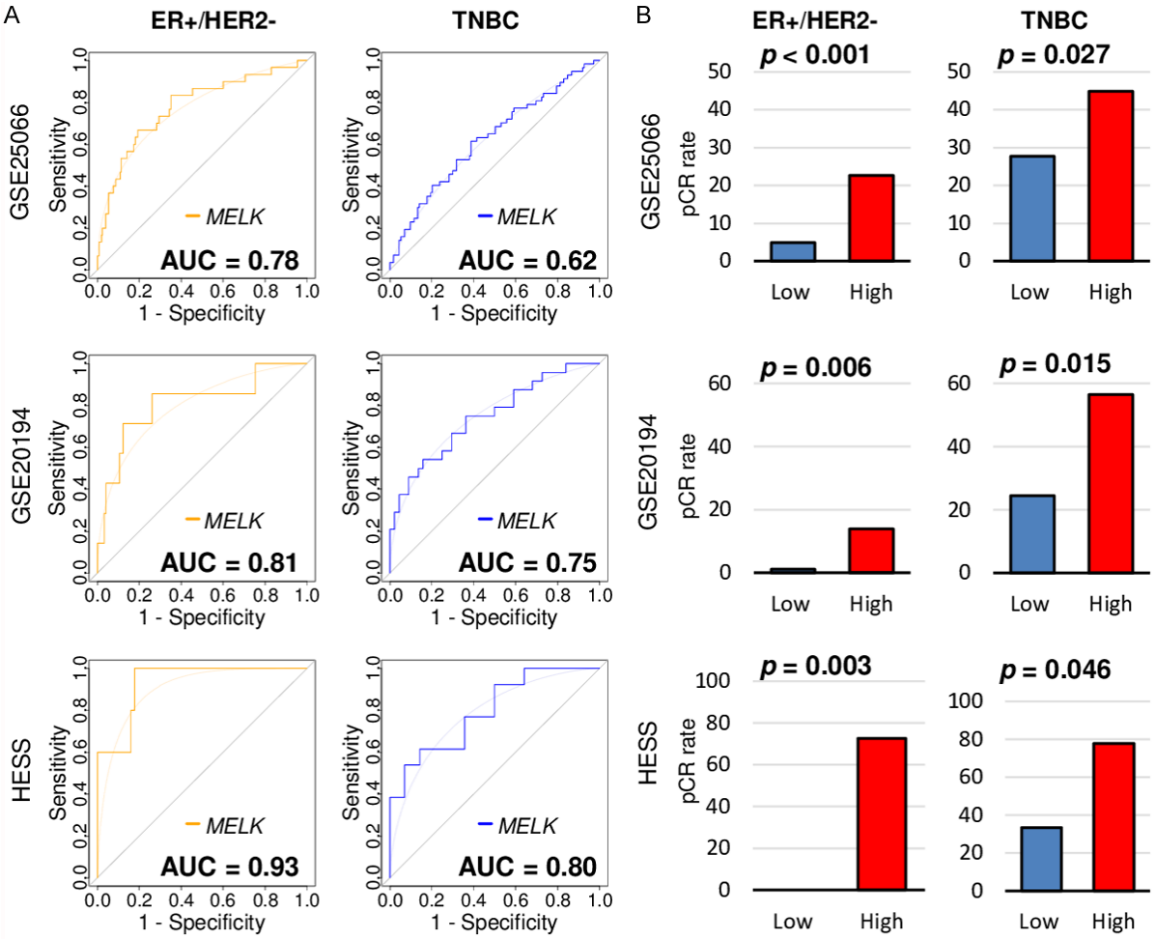


Figure 6. Association of *MELK* expression with drug response for neoadjuvant chemotherapy (NAC) in the GSE25066, GSE20194, and HESS cohorts. A. Receiver operating characteristic (ROC) curve of *MELK* expression between pathological complete response (pCR) and non pCR groups with area under the curve (AUC) in ER-positive/HER2-negative and TNBC subgroups. B. pCR bar plot rates after NAC by high and low *MELK* groups in three cohorts, *p*-values calculated using the Fisher's exact test.

phamide, doxorubicin, fluorouracil, and paclitaxel). In three NAC cohorts, we found that area under the curve (AUC) of *MELK* was more than 0.70 in ER-positive plus HER2-negative breast cancer in GSE25066 (AUC = 0.78, *n* = 278), GSE20194 (AUC = 0.81, *n* = 129), and HESS (AUC = 0.93, *n* = 67) (**Figure 6A**). AUC of *MELK* was also high in TNBC in three cohorts; GSE25066 (AUC = 0.62, *n* = 170), GSE20194 (AUC = 0.75, *n* = 68), and HESS (AUC = 0.80, *n* = 27) (**Figure 6A**). Finally, we investigated the NAC predictive use of *MELK* as a biomarker. In all the above three cohorts, high *MELK* achieved a higher pCR rate with both ER-positive plus HER2-negative breast cancer as well as TNBC breast subtypes (**Figure 6B**; ER-positive/HER2-negative; $P < 0.001$, $P = 0.006$, and $P = 0.003$, TNBC; $P = 0.027$, 0.015, and 0.046, respectively). These results suggest, following NAC, that *MELK* expression can be utilized as a biomark-

er for attaining pCR in both ER-positive plus HER2-negative as well as TNBC subtypes.

Discussion

We found, in this study, that *MELK* expression was significantly correlated with tumor growth assessed by AJCC stage, Nottingham histological grade, *MKI67* expression, TNBC subtype and also with cell proliferation-related gene sets in two large cohorts. Furthermore, we observed worse patient survival, high mutation rate, and enhanced cancer cell survival pathways, including MTORC1 signaling, DNA repair and unfolded protein response in high *MELK* expression breast cancer. Additionally, breast cancer with high *MELK* expression was significantly enriched in immune-related gene sets, including allograft rejection, IFN-α response and IFN-γ response. Furthermore, infiltration of anti-cancer immune cells ($CD4^+$ memory T cells, T helper type1 cells, $CD8^+$ T cells, M1 macrophages, gamma-delta T cells, and dendritic cells), pro-cancer (T helper type 2 cells and regulatory T cells), and high immune cell killing activity (CYT) was associated with high *MELK* expression. Although *MELK* expression did not correlate with sensitivity of any drug tested in cell lines, high *MELK* was significantly associated with high pCR rate after NAC not only in TNBC, the aggressive breast cancer known to be associated with around 30-40% pCR, but also in ER-positive plus HER2-negative breast cancer, a subtype of breast cancer where pCR rates are very low.

As discussed, *MELK* is known to regulate cellular metabolism [37]. In fact, *MELK* is one of the proliferation markers already included in clinically used prognostic panels, such as MammaPrint [38, 39], and PAM50 [40]. *MELK* plays a role in cell proliferation, splicing, cell cycle, cell invasion, and apoptosis [41]. Our study confirmed that *MELK* expression was associated with cancer cell proliferation using human data by analyzing *MELK's* role clinically (stage), pathologically (grade), and molecular biologically (*MKI67* expression and GSEA). This explains why *MELK* expression was also associated with worse breast cancer survival outcomes due to its aggressive biology. Our previous work also demonstrated that cancer cell proliferation marker is a strong poor prognostic biomarker in breast cancer, which agrees with these findings. For example, G2M target pathway [42] and E2F targets [43] were significantly associated with worse patient outcomes. Furthermore, interestingly, our results suggest that *MELK* expression was significantly associated with not only cell proliferation but also other hallmarks of cancer, such as aggressive signaling (PI3K and MTORC1 signaling), DNA repair, glycolysis, and immune response (allograft rejection, interferon response), shown in Figure S2. Taken together, *MELK* expression is associated with worse outcomes in breast cancer explained by its role in enhanced cancer cell proliferation and also other hallmark signaling pathways.

Pathological complete response (pCR) after NAC is a well-established surrogate marker for improved survival in breast cancer [44]; however, predicting response to NAC continues to remain a challenge in ER-positive breast cancer. An additional area of dilemma is determining the appropriate scenario to escalate or de-escalate NAC for TNBC. Our group has previously reported on several biomarkers that can be used to predict response to NAC. We reported after NAC in ER-positive plus HER2-negative breast cancer, that cell cycle pathways, E2F targets and G2M checkpoints, were associated with pCR [42, 43]. We also reported that in TNC achieving pCR after NAC had a higher probability with low Inositol 1,4,5-trisphosphate 3-kinase C (*ITPKC*) expression [45]. *MELK* was previously pointed out, in basal-like subtype, as one of the genes associated with pCR after NAC, but no mechanistic insight was provided [11]. In this study, we also found that *MELK* expression associated with cell proliferation, immune cell infiltration, which explains the higher NAC response in both TNBC and ER-positive/HER2-negative breast cancer with high *MELK* expression. Compared to using tumor profile pathways that calculate 200 gene expressions requiring analysis of comprehensive gene expression; the measurement of a single gene is far more practical as a predictive biomarker, both from a cost and simplicity stand point, especially when it has the ability to assess predictive chemotherapy benefit for two subtypes of breast cancer using the same approach. Additionally, knowledge of the *MELK* expression upfront to predict patients who are unable to achieve pCR would fulfil two goals. Most importantly, as discussed before, it would help us prioritize patients where escalation to standard treatment of doxorubicin, cyclophos-

phamide and paclitaxel could be planned with confidence to improve on the chances to achieve a good response, that is, at least aim for lower residual cancer burden (RCB-I or RCB-II), since poorer outcomes are associated with RCB-III as greater amount of residual disease remains, especially in patients with hormone receptor positive breast cancer where there is always a debate whether to treat with chemotherapy upfront or not [46]. Secondly, it would encourage the treating physician to treat these patients with neoadjuvant chemotherapy to identify the non-responders where information about residual disease is critical to escalate therapy in the adjuvant setting, example, use of capecitabine for residual disease for TNBC, instead of taking these patients for upfront surgery [47]. Thirdly, for some TNBC patients where MELK expression does not predict for pCR, it could help us select those patients for escalating NAC to include checkpoint inhibitors or carboplatin to maximize chances of achieving pCR, especially since their use in the NAC setting in TNBC is still a much debated issue.

We have previously reported that there is enhanced cell proliferation in breast cancer with high mutation load, but this is counterbalanced with immune response and immune cell infiltration [35, 36]. Many including our group have suggested that immune cell infiltration in TME affects response to NAC [48]. We reported that high fraction of regulatory T cells (Tregs) in TME is associated with low NAC response in TNBC [49]. To our knowledge, this is the first study to report association with *MELK* expression and immune response and immune cell infiltration in breast cancer patient TME, a finding that was not known before. Since *MELK* gene was expressed not only in cancer cells but also in immune cells that we found by single cell sequence analysis, it is possible that infiltrated immune cells may be contributing to the *MELK* expression levels. Since tumors in humans are infiltrated with immune cells as opposed to cell lines, this may be responsible for the different outcomes of drug response in cell lines and patient response to NAC. The above findings lay a firm ground to support the clinical predictive utility of *MELK* expression in patient selection as a biomarker for NAC in breast cancer, consistent with our original hypothesis. Future prospective trials are warranted in order to further test the clinical utility of this biomarker and to refute or validate our hypothesis.

We have used multiple independent large cohorts to test and validate that *MELK* expression is associated with NAC response; however, the study still has limitations. Given that we utilized existing publicly available cohorts, this is by nature a retrospective study where we cannot eliminate the possibility of selection bias. Further, current study was conducted by bioinformatics analyses alone, thus the causal relationships are unknown and in vitro and/or in vivo experiments are needed to elucidate the underlying mechanisms. At the same time, a strength of this study is that we analyzed the patient's cancer including the TME that cannot be fully modeled by any experimental setting. Eventually, a prospective study is necessary in order to validate and confirm the utility of *MELK* expression in breast cancer management, both as a predictive and prognostic biomarker.

We demonstrated, in conclusion, in both ER-positive/HER2-negative and TNBC, high *MELK* expression is significantly associated with cell proliferation, immune cell infiltration, and higher incidence of response to NAC.

Acknowledgements

This work was supported by US National Institutes of Health/National Cancer Institute grant R01CA160688, R01CA250412, R37CA-248018, US Department of Defense BCRP grant W81XWH-19-1-0674, as well as the Edward K. Duch Foundation and Paul & Helen Ellis Charitable Trust to K.T., and US National Cancer Institute cancer center support grant P30-CA016056 to Roswell Park Comprehensive Cancer Center.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Kazuaki Takabe, Breast Surgery, Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Elm & Carlton Streets, Buffalo, New York 14263, USA. Tel: 1-716-845-2918; Fax: 1-716-845-1668; E-mail: kazuaki.takabe@roswellpark.org

References

[1] Asaoka M, Gandhi S, Ishikawa T and Takabe K. Neoadjuvant chemotherapy for breast cancer: past, present, and future. Breast Cancer (Auckl) 2020; 14: 1178223420980377.

[2] Sikov WM, Berry DA, Perou CM, Singh B, Cirrincione CT, Tolaney SM, Kuzma CS, Pluard TJ, Somlo G, Port ER, Golshan M, Bellon JR, Collyar D, Hahn OM, Carey LA, Hudis CA and Winer EP. Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (Alliance). J Clin Oncol 2015; 33: 13-21.

[3] von Minckwitz G, Schneeweiss A, Loibl S, Salat C, Denkert C, Rezai M, Blohmer JU, Jackisch C, Paepke S, Gerber B, Zahm DM, Kümmel S, Eidtmann H, Klare P, Huober J, Costa S, Tesch H, Hanusch C, Hilfrich J, Khandan F, Fasching PA, Sinn BV, Engels K, Mehta K, Nekljudova V and Untch M. Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): a randomised phase 2 trial. Lancet Oncol 2014; 15: 747-756.

[4] Schmid P, Cortes J, Pusztai L, McArthur H, Kümmel S, Bergh J, Denkert C, Park YH, Hui R, Harbeck N, Takahashi M, Foukakis T, Fasching PA, Cardoso F, Untch M, Jia L, Karantza V, Zhao J, Aktan G, Dent R and O'Shaughnessy J. Pembrolizumab for early triple-negative breast cancer. N Engl J Med 2020; 382: 810-821.

[5] Mittendorf EA, Zhang H, Barrios CH, Saji S, Jung KH, Hegg R, Koehler A, Sohn J, Iwata H, Telli ML, Ferrario C, Punie K, Penault-Llorca F, Patel S, Duc AN, Liste-Hermoso M, Maiya V, Molinero L, Chui SY and Harbeck N. Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): a randomised, double-blind, phase 3 trial. Lancet 2020; 396: 1090-1100.

[6] Gray D, Jubb AM, Hogue D, Dowd P, Kljavin N, Yi S, Bai W, Frantz G, Zhang Z, Koeppen H, de Sauvage FJ and Davis DP. Maternal embryonic leucine zipper kinase/murine protein serine-threonine kinase 38 is a promising therapeutic target for multiple cancers. Cancer Res 2005; 65: 9751-9761.

[7] Badouel C, Chartrain I, Blot J and Tassan JP. Maternal embryonic leucine zipper kinase is stabilized in mitosis by phosphorylation and is partially degraded upon mitotic exit. Exp Cell Res 2010; 316: 2166-2173.

[8] Chen P, Wang J, Wang X, Chen X, Li C and Tan T. Cloning, tissue distribution, expression pattern, and function of porcine maternal embryonic leucine zipper kinase. Ann Transl Med 2020; 8: 239.

[9] McDonald IM and Graves LM. Enigmatic MELK: the controversy surrounding its complex role in cancer. J Biol Chem 2020; 295: 8195-8203.

[10] Yang K, Gao J and Luo M. Identification of key pathways and hub genes in basal-like breast cancer using bioinformatics analysis. Onco Targets Ther 2019; 12: 1319-1331.

[11] Liu R, Lv QL, Yu J, Hu L, Zhang LH, Cheng Y and Zhou HH. Correlating transcriptional networks with pathological complete response following neoadjuvant chemotherapy for breast cancer. Breast Cancer Res Treat 2015; 151: 607-618.

[12] Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017; 18: 220.

[13] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12: 453-457.

[14] Gandhi S, Elkhanany A, Oshi M, Dai T, Opyrchal M, Mohammadpour H, Repasky EA and Takabe K. Contribution of immune cells to glucocorticoid receptor expression in breast cancer. Int J Mol Sci 2020; 21: 4635.

[15] Oshi M, Angarita FA, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. High expression of NRF2 is associated with increased tumor-infiltrating lymphocytes and cancer immunity in ER-positive/HER2-negative breast cancer. Cancers (Basel) 2020; 12: 3856.

[16] Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Gräf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Langerød A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Børresen-Dale AL, Brenton JD, Tavaré S, Caldas C and Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.

[17] 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 and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.

[18] Brueffer C, Gladchuk S, Winter C, Vallon-Christersson J, Hegardt C, Häkkinen J, George AM, Chen Y, Ehinger A, Larsson C, Loman N, Malmberg M, Rydén L, Borg Å and Saal LH. The mutational landscape of the SCAN-B real-world primary breast cancer transcriptome. EMBO Mol Med 2020; 12: e12118.

[19] 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, Cui J, Czika W, Demichelis F, Deng X, Dosymbekov D, Eils R, Feng Y, Fostel J, Fulmer-Smentek S, Fuscoe JC, Gatto L, Ge W, Goldstein DR, Guo L, Halbert DN, Han J, Harris SC, Hatzis C, Herman D, Huang J, Jensen RV, Jiang R, Johnson CD, Jurman G, Kahlert Y, Khuder SA, Kohl M, Li J, Li L, Li M, Li QZ, Li S, Li Z, Liu J, Liu Y, Liu Z, Meng L, Madera M, Martinez-Murillo F, Medina I, Meehan J, Miclaus K, Moffitt RA, Montaner D, Mukherjee P, Mulligan GJ, Neville P, Nikolskaya T, Ning B, Page GP, Parker J, Parry RM, Peng X, Peterson RL, Phan JH, Quanz B, Ren Y, Riccadonna S, Roter AH, Samuelson FW, Schumacher MM, Shambaugh JD, Shi Q, Shippy R, Si S, Smalter A, Sotiriou C, Soukup M, Staedtler F, Steiner G, Stokes TH, Sun Q, Tan PY, Tang R, Tezak Z, Thorn B, Tsyganova M, Turpaz Y, Vega SC, Visintainer R, von Frese J, Wang C, Wang E, Wang J, Wang W, Westermann F, Willey JC, Woods M, Wu S, Xiao N, Xu J, Xu L, Yang L, Zeng X, Zhang J, Zhang L, Zhang M, Zhao C, Puri RK, Scherf U, Tong W and Wolfinger RD. The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. Nat Biotechnol 2010; 28: 827-838.

[20] 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, Chacón 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 and Symmans WF. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA 2011; 305: 1873-1881.

[21] Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, Booser D, Theriault RL, Buzdar AU, Dempsey PJ, Rouzier R, Sneige N, Ross JS, Vidaurre T, Gómez HL, Hortobagyi GN and Pusztai L. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol 2006; 24: 4236-4244.

[22] 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 Network, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG and Shmulevich I. The immune landscape of cancer. Immunity 2019; 51: 411-412.

[23] Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, Omberg L, Wolf DM, Shriver CD, Thorsson V and Hu H. An integrated TCGA Pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell 2018; 173: 400-416, e411.

[24] Oshi M, Tokumaru Y, Angarita FA, Yan L, Matsuyama R, Endo I and Takabe K. Degree of early estrogen response predict survival after endocrine therapy in primary and metastatic er-positive breast cancer. Cancers (Basel) 2020; 12: 3557.

[25] Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. Inflammation is associated with worse outcome in the whole cohort but with better outcome in triple-negative subtype of breast cancer patients. J Immunol Res 2020; 2020: 5618786.

[26] Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Endo I, Nagahashi M and Takabe K. Intra-tumoral angiogenesis is associated with inflammation, immune reaction and metastatic recurrence in breast cancer. Int J Mol Sci 2020; 21: 6708.

[27] Oshi M, Katsuta E, Yan L, Ebos JML, Rashid OM, Matsuyama R, Endo I and Takabe K. A novel 4-gene score to predict survival, distant

metastasis and response to neoadjuvant therapy in breast cancer. Cancers (Basel) 2020; 12: 1148.

[28] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.

[29] Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP and Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015; 1: 417-425.

[30] Oshi M, Tokumaru Y, Asaoka M, Yan L, Satyananda V, Matsuyama R, Matsuhashi N, Futamura M, Ishikawa T, Yoshida K, Endo I and Takabe K. M1 Macrophage and M1/M2 ratio defined by transcriptomic signatures resemble only part of their conventional clinical characteristics in breast cancer. Sci Rep 2020; 10: 16554.

[31] Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Kalinski P, Endo I and Takabe K. 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 (Basel) 2020; 12: 3342.

[32] Oshi M, Asaoka M, Tokumaru Y, Yan L, Matsuyama R, Ishikawa T, Endo I and Takabe K. CD8 T cell score as a prognostic biomarker for triple negative breast cancer. Int J Mol Sci 2020; 21: 6968.

[33] Schulze A, Oshi M, Endo I and Takabe K. MYC targets scores are associated with cancer aggressiveness and poor survival in ER-positive primary and metastatic breast cancer. Int J Mol Sci 2020; 21: 8127.

[34] Tokumaru Y, Oshi M, Katsuta E, Yan L, Satyananda V, Matsuhashi N, Futamura M, Akao Y, Yoshida K and Takabe K. KRAS signaling enriched triple negative breast cancer is associated with favorable tumor immune microenvironment and better survival. Am J Cancer Res 2020; 10: 897-907.

[35] Takahashi H, Oshi M, Asaoka M, Yan L, Endo I and Takabe K. Molecular biological features of nottingham histological grade 3 breast cancers. Ann Surg Oncol 2020; 27: 4475-4485.

[36] Takahashi H, Asaoka M, Yan L, Rashid OM, Oshi M, Ishikawa T, Nagahashi M and Takabe K. Biologically aggressive phenotype and anticancer immunity counterbalance in breast cancer with high mutation rate. Sci Rep 2020; 10: 1852.

[37] Sun X, Gao L, Chien HY, Li WC and Zhao J. The regulation and function of the NUAK family. J Mol Endocrinol 2013; 51: R15-22.

[38] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415: 530-536.

[39] Tian S, Roepman P, Van't Veer LJ, Bernards R, de Snoo F and Glas AM. Biological functions of the genes in the mammaprint breast cancer profile reflect the hallmarks of cancer. Biomark Insights 2010; 5: 129-138.

[40] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM and Bernard PS. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009; 27: 1160-1167.

[41] Chartrain I, Blot J, Lerivray H, Guyot N and Tassan JP. A mitochondrial-targeting signal is present in the non-catalytic domain of the MELK protein kinase. Cell Biol Int 2007; 31: 196-201.

[42] Oshi M, Takahashi H, Tokumaru Y, Yan L, Rashid OM, Matsuyama R, Endo I and Takabe K. G2M cell cycle pathway score as a prognostic biomarker of metastasis in estrogen receptor (ER)-positive breast Cancer. Int J Mol Sci 2020; 21: 2921.

[43] Oshi M, Takahashi H, Tokumaru Y, Yan L, Rashid OM, Nagahashi M, Matsuyama R, Endo I and Takabe K. The E2F pathway score as a predictive biomarker of response to neoadjuvant therapy in ER+/HER2- breast cancer. Cells 2020; 9: 1643.

[44] Pennisi A, Kieber-Emmons T, Makhoul I and Hutchins L. Relevance of pathological complete response after neoadjuvant therapy for breast cancer. Breast Cancer (Auckl) 2016; 10: 103-106.

[45] Oshi M, Newman S, Murthy V, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. ITPKC as a prognostic and predictive biomarker of neoadjuvant chemotherapy for triple negative breast cancer. Cancers (Basel) 2020; 12: 2758.

[46] Hamy AS, Darrigues L, Laas E, De Croze D, Topciu L, Lam GT, Evrevin C, Rozette S, Laot L, Lerebours F, Pierga JY, Osdoit M, Faron M, Feron JG, Laé M and Reyal F. Prognostic value of the Residual Cancer Burden index according to breast cancer subtype: validation on a cohort of BC patients treated by neoadjuvant chemotherapy. PLoS One 2020; 15: e0234191.

[47] Masuda N, Lee SJ, Ohtani S, Im YH, Lee ES, Yokota I, Kuroi K, Im SA, Park BW, Kim SB, Yanagita Y, Ohno S, Takao S, Aogi K, Iwata H,

Jeong J, Kim A, Park KH, Sasano H, Ohashi Y and Toi M. Adjuvant capecitabine for breast cancer after preoperative chemotherapy. N Engl J Med 2017; 376: 2147-2159.

[48] Loi S, Sirtaine N, Piette F, Salgado R, Viale G, Van Eenoo F, Rouas G, Francis P, Crown JP, Hitre E, de Azambuja E, Quinaux E, Di Leo A, Michiels S, Piccart MJ and Sotiriou C. Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. J Clin Oncol 2013; 31: 860-867.

[49] Oshi M, Asaoka M, Tokumaru Y, Angarita FA, Yan L, Matsuyama R, Zsiros E, Ishikawa T, Endo I and Takabe K. Abundance of regulatory T Cell (Treg) as a predictive biomarker for neoadjuvant chemotherapy in triple-negative breast cancer. Cancers (Basel) 2020; 12: 3038.

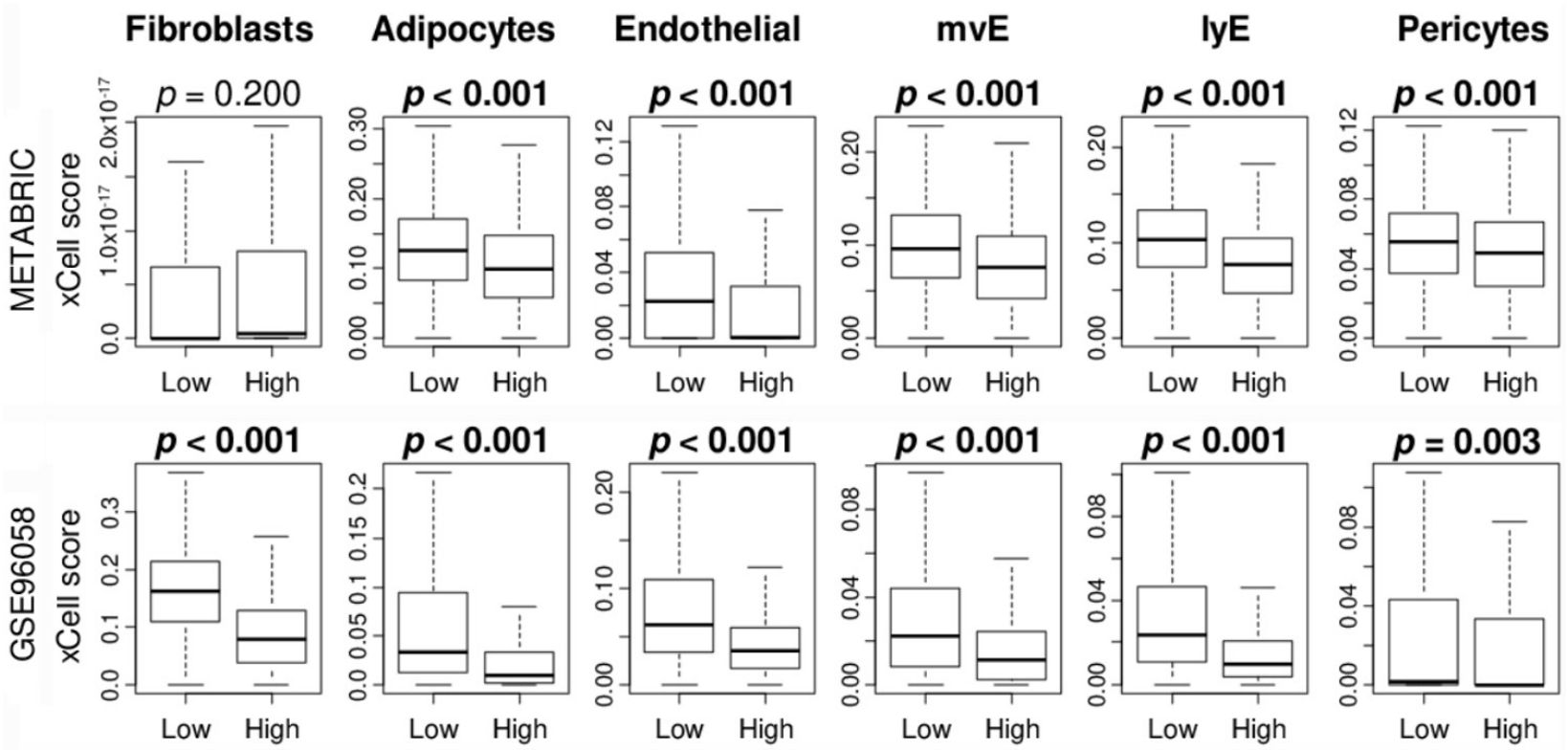


Figure S1. Association of the MELK expression with the fraction of stromal cells in the tumor microenvironment. Boxplots of the fraction of several stromal cells, including fibroblasts, adipocytes, endothelial cells, micro vessel endothelial (mvE) cells, lymphatic endothelial (lyE) cells, and pericytes cells by high and low *MELK* groups in the METABRIC and GSE96058 cohorts. *P* values were calculated by Mann-Whitney U test.

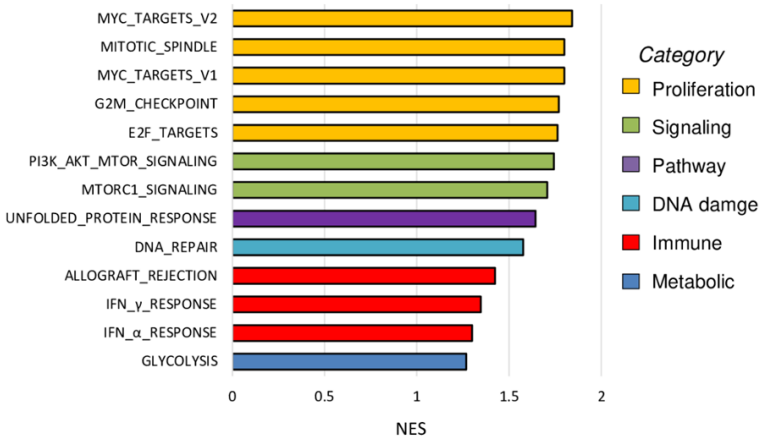


Figure S2. A Gene set enrichment analysis (GSEA) of the Hallmark gene sets with significant enrichment due to high MELK expression in breast cancer. Gene sets were listed by the order of high to low enrichment score for the hallmark gene sets that were statistically significant (false discovery rate (FDR) < 0.25, as recommended by the GSEA software). Category was defined by original paper of hallmark gene set collection.

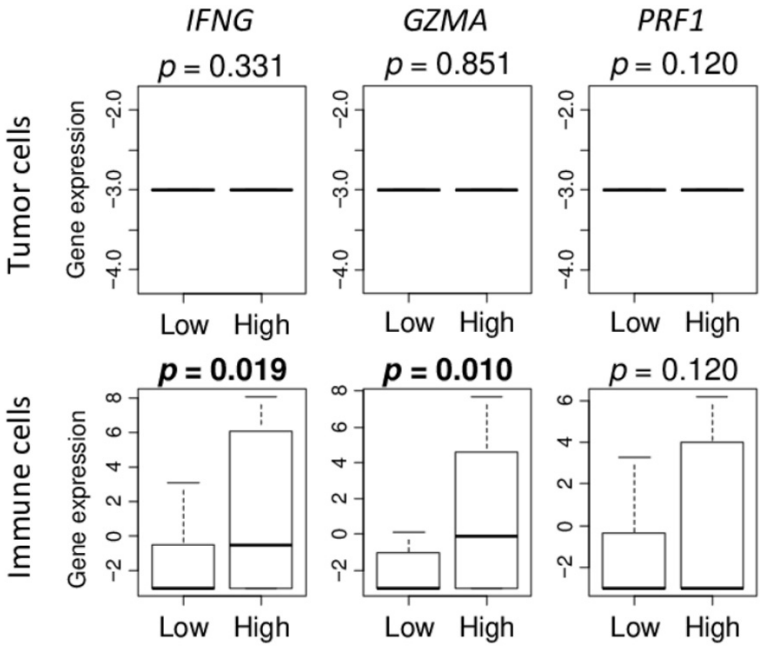


Figure S3. Association between MELK and immune function-related gene expression in tumor cells and immune cells. Boxplots of comparison of immune function-related genes expression, including IFNG, GZMA, and PRF1, by low and high MELK expression groups. The low and high represent groups without and with expression of MELK, respectively. Single cell sequence data was used for the analysis (GSE75688 cohort). *P*-value was calculated by Kruskal-Wallis test.