



Article

# G2M Cell Cycle Pathway Score as a Prognostic Biomarker of Metastasis in Estrogen Receptor (ER)-Positive Breast Cancer

Masanori Oshi <sup>1,2</sup>, Hideo Takahashi <sup>1</sup>, Yoshihisa Tokumaru <sup>1,3</sup>, Li Yan <sup>4</sup>,  
Omar M. Rashid <sup>5,6</sup>, Ryusei Matsuyama <sup>2</sup>, Itaru Endo <sup>2</sup> and Kazuaki Takabe <sup>1,2,7,8,9,10,\*</sup>

<sup>1</sup> Department of Surgical Oncology, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263, USA; masa1101oshi@gmail.com (M.O.); hideo.takahashi@roswellpark.org (H.T.);

yoshihisa.tokumaru@roswellpark.org (Y.T.)

<sup>2</sup> Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; ryusei@yokohama-cu.ac.jp (R.M.); endoit@yokohama-cu.ac.jp (I.E.)

<sup>3</sup> Department of Surgical Oncology, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan

<sup>4</sup> Department of Biostatistics & Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY 14263, USA; li.yan@roswellpark.org

<sup>5</sup> Department of Surgery, Holy Cross Hospital, Michael and Dianne Bienes Comprehensive Cancer Center, Fort Lauderdale, FL 33308, USA; omarmrashidmdjd@gmail.com

<sup>6</sup> Department of Surgery, Massachusetts General Hospital, Boston, MA 02114, USA

<sup>7</sup> Department of Gastrointestinal Tract Surgery, Fukushima Medical University School of Medicine, Fukushima 960-1295, Japan

<sup>8</sup> Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, State University of New York, Buffalo, NY 14263, USA

<sup>9</sup> Department of Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata 951-8510, Japan

<sup>10</sup> Department of Breast Surgery and Oncology, Tokyo Medical University, Tokyo 160-8402, Japan

\* Correspondence: kazuaki.takabe@roswellpark.org; Tel.: +1-716-8455540; Fax: +1-716-8451668

Received: 31 March 2020; Accepted: 20 April 2020; Published: 22 April 2020



**Abstract:** The vast majority of breast cancer death is a result of metastasis. Thus, accurate identification of patients who are likely to have metastasis is expected to improve survival. The G2M checkpoint plays a critical role in cell cycle. We hypothesized that breast cancer tumors with high activity of G2M pathway genes are more aggressive and likely to metastasize. To test this, we used the single-sample gene set variation analysis method to calculate the score for the Hallmark G2M checkpoint pathway using gene expression data of a total of 4626 samples from 12 human breast cancer cohorts. As expected, a high G2M pathway score correlated with enriched tumor expression of other cell proliferation-related gene sets. The score was significantly associated with clinical aggressive features of tumors and patient survival in estrogen receptor (ER)-positive/human epidermal growth factor receptor 2 (HER2)-negative breast cancer. Interestingly, a high G2M score of metastasis tumors was also significantly associated with worse survival. In primary as well as metastasis tumors with high scores, the infiltration of both pro- and anti-cancerous immune cells increased. Tumor G2M score was also associated with treatment response to systemic chemotherapy in ER-positive/HER2-negative cancer, and was predictive of response to cyclin-dependent kinase inhibition therapy.

**Keywords:** breast cancer; biomarker; cell cycle; gene set; G2M; metastasis; pathway analysis; tumor gene expression

## 1. Introduction

Currently, there is no cure that can eradicate metastatic breast cancer. In breast cancer, metastasis eventually occurs in 40% of patients, and it is the major cause of the 40,000 breast cancer-related deaths that occur annually in the United States [1]. Thus, precise identification of patients who have a specific treatable target, and who develop metastasis and are at high risk of succumbing to the disease, will have a practical impact on improvement of survival.

Cell cycle progression plays a crucial role in cell proliferation, the alteration of which has been acknowledged as one of the hallmarks of cancer [2]. The G2-M (G2M) checkpoint in cell cycle control mediates G2-M cell cycle transition through cyclin B-cdc2 (CDK1) complexes [3,4], and G2M regulation mediated by CDK1 is a critical factor in tumorigenic events [5]. In recent years, the CDK inhibitors palbociclib, ribociclib, and abemaciclib have emerged as a standard of care in combination with endocrine therapies for patients with estrogen receptor (ER)-positive/human epidermal growth factor receptor 2 (HER2)-negative metastatic breast cancer [6]. ER-positive breast cancer is the most common breast cancer subtype, accounting for approximately 70% of all metastatic breast cancers. Therefore, quantification of the activity of genes involved in the G2M pathway is expected to identify patients who are likely to develop metastasis.

Microarray and sequencing technologies have evolved rapidly in the past few years and revolutionized the collection of DNA and RNA data, making them essential for biomedical research. For example, the computational models involving features like gene expression [7] have been established to predict drug sensitivity for cancer treatment. Yet, it has been reported that findings based on gene-level expression can have limited reproducibility across independent studies. It is also challenging to interpret the biological meaning of changes in the expression of a single gene [8]. Biological phenomena such as drug responses are modulated by the concerted behavior of multiple genes [9], and gene expression scoring capturing multiple genes involved in the same pathway can provide a more accurate value than single-gene scoring [10]. A pathway or gene-sets based approach can take into account such coordination of genes, reduce model complexity, and increase the explanatory power of prediction models [11–13]. The single-sample gene set variation analysis (GSVA) computational method can be used to score gene expression at the level of a pathway of interest instead of a single gene, and has been utilized to measure pathway activities in samples [10,14].

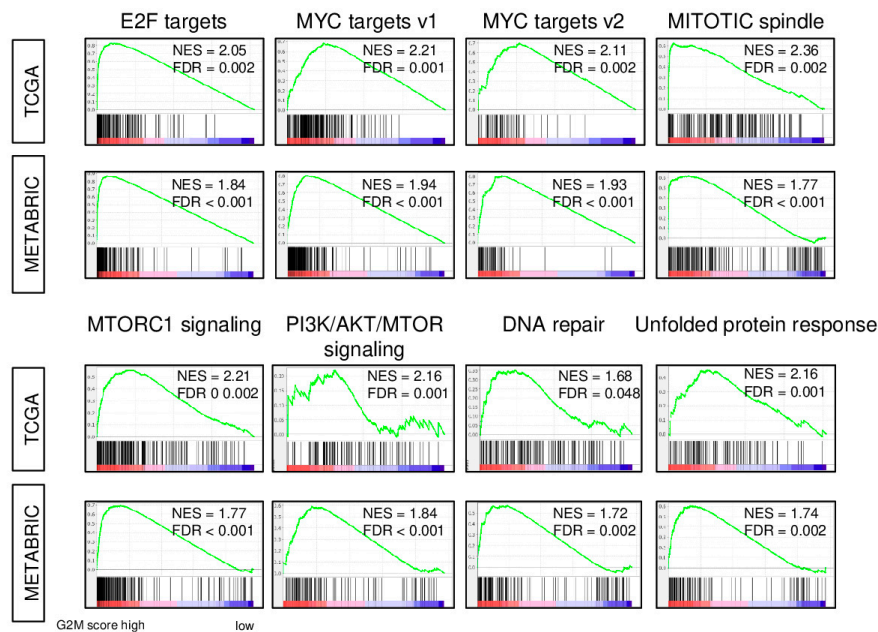
In the current study, we hypothesized that the upregulation of the G2M pathway is associated with aggressive cancer biology, a higher likelihood of metastasis, and a worse survival in patients with ER-positive/HER2-negative breast cancer. We used the GSVA method to measure G2M pathway activity of breast cancer tumors, and examined the association of the G2M score with various pathological and clinical features of breast cancer.

## 2. Results

### 2.1. Expression of Cell Proliferation-Related Genes is increased in Breast Cancer with a High G2M Pathway Score

We calculated the G2M scores for 1065 primary tumors of The Cancer Genome Atlas (TCGA) breast invasive carcinoma (BRCA) cohort as the GSVA score for the Hallmark G2-M checkpoint gene set [15]. Tumors were dichotomized into high and low score groups using the median score value. As expected, gene set enrichment analysis showed that tumors with a high G2M score also had enrichment of the expression of genes of multiple gene sets related to cell cycle and cell proliferation. Specifically, in the examination of the 50-member Hallmark gene set collection, significant enrichment was seen for the E2F target, MYC targets v1 and v2, mitotic spindle, MTORC1 signaling, PI3K AKT MTOR, and DNA repair gene sets (false discovery rate (FDR) < 0.05; Figure 1). Because the same gene can be a member of multiple sets, we examined the overlap of genes of these gene sets. The percentage of the genes overlapping with the G2M pathway set was highest for the E2F target (36.5%), but less than 20% for the other gene sets (Table S1). The association of the G2M pathway score with enriched expression of cell

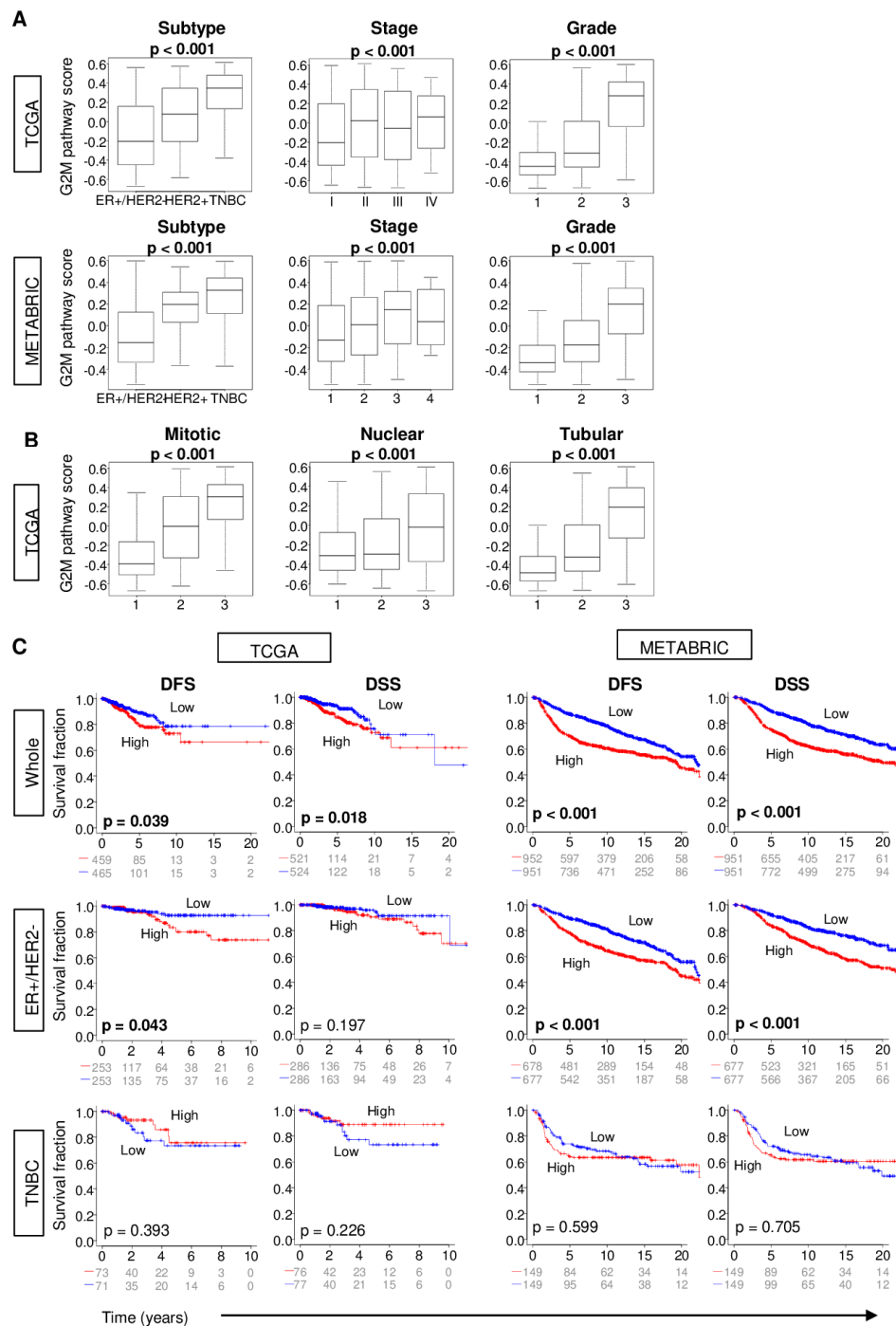
cycle and cell proliferation gene sets that was observed for the TCGA cohort was completely mirrored in a totally independent cohort, the Molecular Taxonomy of breast Cancer International Consortium (METABRIC; Figure 1). These findings suggest that the G2M pathway score reflects cell proliferation in breast cancer.



**Figure 1.** Hallmark gene sets with a significant enrichment in high G2M pathway score tumors. Gene set enrichment plots along with normalized enrichment score (NES) and false discovery rate (FDR) are shown for the eight gene sets for which enrichment was seen in tumors with high compared with low G2M pathway activity score in both The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of breast Cancer International Consortium (METABRIC) cohorts. The G2M score was calculated from tumor gene expression as the single-sample gene set variation analysis score for the Hallmark G2M gene set, and the within-cohort median value was used to identify tumors with high and low scores. NES and FDR were determined with the classical gene set enrichment analysis (GSEA) method.

## 2.2. Increased G2M Pathway Activity in Breast Cancer Tumors is Associated with Worse Clinico-Pathologic Features

Given that the G2M pathway score reflects cell proliferation, we hypothesized that a high G2M score is indicative of aggressive cancer biology, and thus associated with worse clinical outcomes. Indeed, in the TCGA cohort, the tumor G2M score was higher in triple negative breast cancer (TNBC) and HER2-positive subtypes ( $p < 0.001$ ), which are both known to be more aggressive than ER-positive/HER2-negative subtype (Figure 2A;  $p < 0.001$ ). Similarly, the score was significantly higher in tumors with later American Joint Committee on Cancer (AJCC) cancer stage ( $p < 0.001$ ) and greater Nottingham pathological grade as well as the individual Nottingham scores for mitotic count, nuclear pleomorphism, and tubular formation ( $p < 0.001$ ). Strikingly consistent results were seen with the METABRIC cohort (Figure 2A).

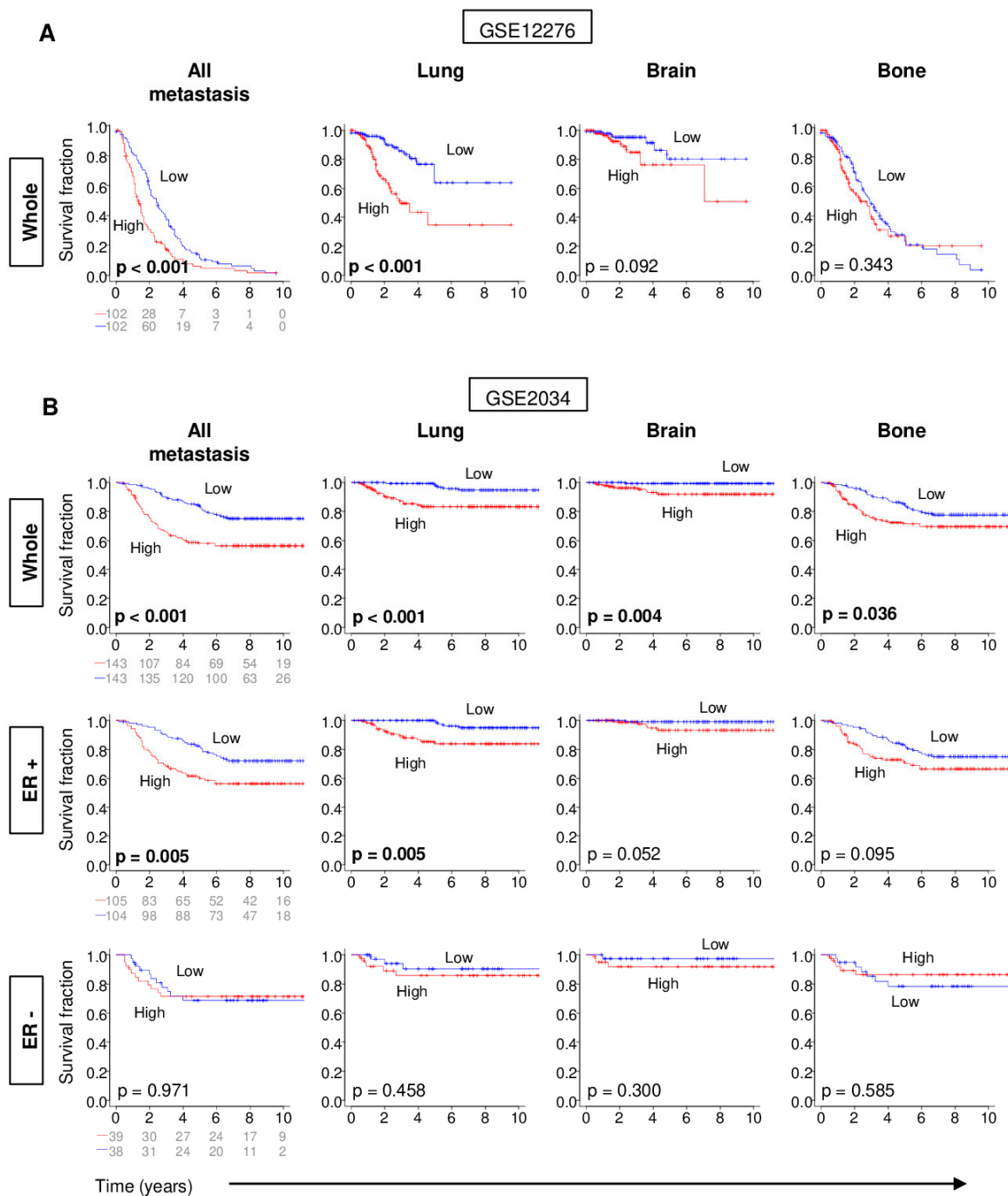


**Figure 2.** G2M pathway activity score and tumor characteristics and patient survival. (A). Tukey boxplots depict G2M scores among TCGA and METABRIC tumors of different grades, American Joint Committee on Cancer clinical stages, and subtypes. (B). Tukey boxplots depict G2M scores among TCGA tumors of different mitotic, nuclear, and tubular Nottingham histological scores. *p*-values for A and B are calculated using one-way analysis of variance (ANOVA). (C). Kaplan–Meier survival plots comparing patients with tumors with high and low G2M scores along with logrank test *p*-values are shown for disease-free (DFS) and disease-specific (DSS) survival for the entire cohort (Whole), or its sub-groups of estrogen receptor (ER)-positive/human epidermal growth factor receptor (HER2)-negative and triple negative (TNBC) patients. The G2M score was calculated from tumor gene expression as the single-sample gene set variation analysis score for the Hallmark G2M gene set, and within-cohort median value was used to identify tumors with high and low scores.

TCGA patients whose tumors had high G2M scores had significantly shorter disease-free survival (DFS) as well as disease-specific survival (DSS) compared with patients with G2M pathway low score groups in TCGA cohort (Figure 2C; logrank test,  $p = 0.039$  and  $0.018$ , respectively). The association of score with survival was more significant in the METABRIC cohort, which had double the cohort size (Figure 2C; both  $p < 0.001$ ). As the G2M score was higher in aggressive cancer subtypes, as demonstrated in Figure 2A, this survival difference between the groups could reflect the inclusion of aggressive subtypes, such as TNBC, in the high G2M score group. Therefore, we performed survival analyses by cancer subtype. Interestingly, a high G2M score was associated with a shorter survival only in the ER-positive/HER2-negative (Figure 2C; DFS:  $p = 0.043$ , DSS:  $p = 0.197$ ), but not in the TNBC subtype (DFS:  $p = 0.393$ , DSS:  $p = 0.226$ ). This result was validated in the METABRIC sub-groups of cases with a stronger significance. For both DFS and DSS among ER-positive/HER2-negative cases, the logrank test  $p$  was  $< 0.001$ , whereas for TNBC, the  $p$ -values were  $0.599$  and  $0.705$ , respectively. It is known that low ER expressing tumors demonstrate worse clinical outcomes compared with high ER expressing tumors [16]. To this end, it was of interest whether the association of the G2M score with clinical outcome is confounded by ER expression levels. There was no correlation between the G2M score and ER expression (ESR1 and ESR2 gene expressions for ER-alpha and ER-beta, respectively), which suggests that low ER expression is not confounding the G2M score as it relates to clinical outcome (Figure S1). These findings suggest that a G2M pathway high score reflects aggressive tumor biology, particularly in ER-positive/HER2-negative breast cancer.

### 2.3. Distant Metastasis is More Likely to Occur in Tumors with High G2M Pathway Activity

With shorter DFS associated with a high G2M pathway score, we hypothesized that tumors with a high G2M score develop recurrence or distant metastasis earlier than those with a low score. In order to test this hypothesis, we identified two independent cohorts with previously published primary tumor gene expression and site-specific metastasis-free survival data (Gene Expression Omnibus studies GSE12276 [17] and GSE2034 [18]). We focused on the lung, brain, and bone as specific metastasis sites as distant metastasis is most common at these sites in breast cancer [19]. Primary tumors with a high G2M score had a significantly shorter overall as well as lung-specific (both logrank test  $p < 0.001$ ), but not for brain- or bone-specific metastasis-free survival in the GSE12276 cohort (Figure 3A). On the other hand, a high G2M pathway score was associated with significantly shorter metastasis-free survival for metastasis to any site ( $p < 0.001$ ), or to the lung ( $p < 0.001$ ), brain ( $p = 0.004$ ), and bone ( $p = 0.036$ ) in the GSE2034 cohort, which has a larger size than GSE12276 (Figure 3B). When examined by ER positivity, a high G2M pathway score was associated with shorter metastasis-free survival only in ER-positive cancer, although the ER-negative sample size was small in GSE2034 cohort, and the ER status of the primary tumors was unavailable for the other cohort (Figure 3B). Thus, the tumor G2M pathway score appears to be useful in predicting distant metastasis among patients with ER-positive breast cancer.

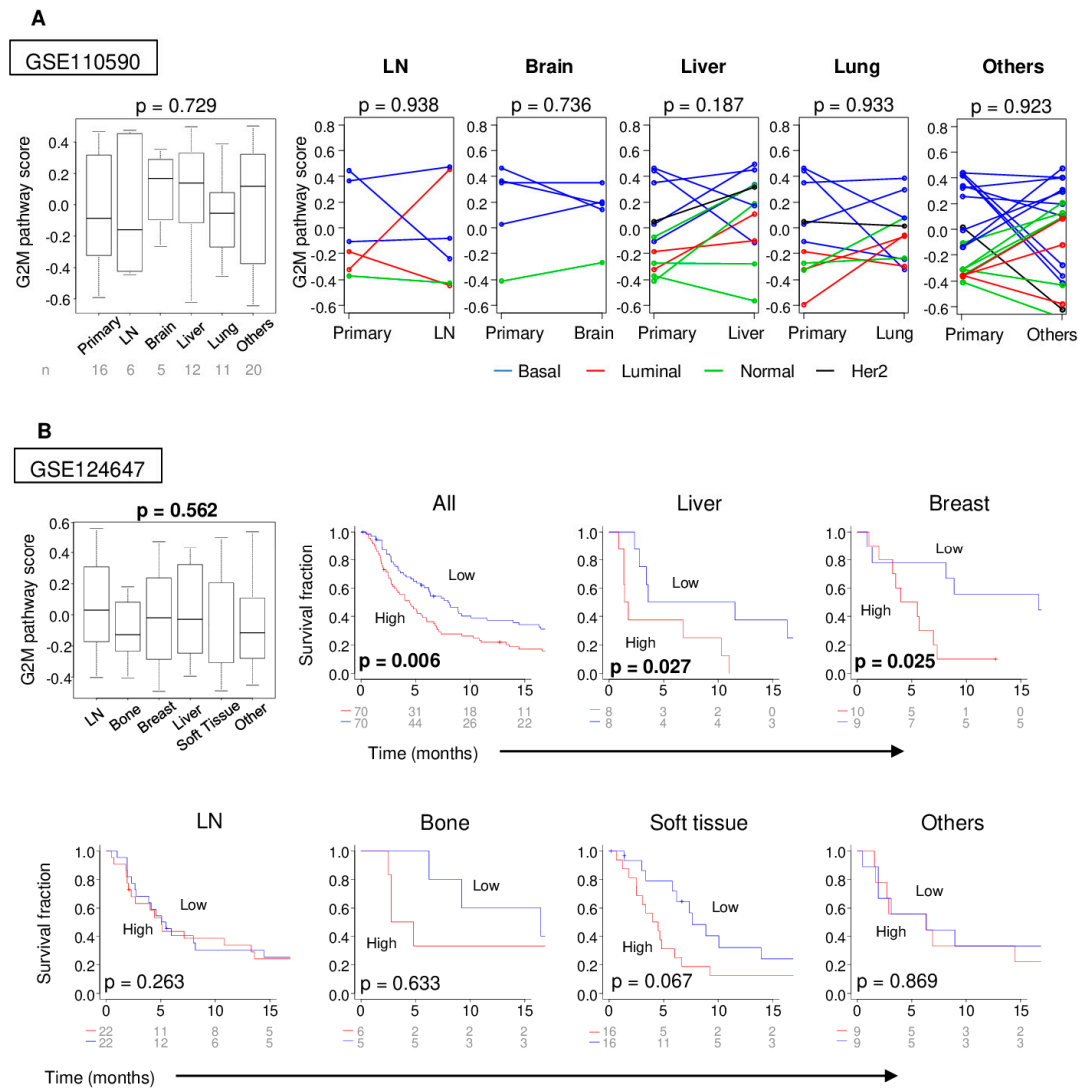


**Figure 3.** (A) The Kaplan–Meier survival plots depicting metastasis-free survival for metastasis to lung, brain, or bone based on G2M pathway score of the primary tumors in the GSE12276 and (B) in the GSE2034 cohort. In the GSE2034, metastasis-free survival was demonstrated in the whole cohort, ER+, and TNBC subtypes. The median value of the G2M pathway score was used to divide patients into low and high groups. ER, estrogen receptor; TNBC, triple negative breast cancer.

#### 2.4. Metastatic Tumors with a High G2M Pathway Score were Associated with Significantly Worse Survival

As tumor metastases are generally more aggressive than their primary tumors, we hypothesized that the G2M pathway activity score would accordingly be higher in metastases. To confirm this, we compared the scores of primary and matched metastasis tumors of the GSE110590 cohort [20]. There were no significant differences in the G2M pathway score among metastases at different sites

(Figure 4A) and, contrary to our hypothesis, there was no significant difference in the G2M pathway score between the primary tumors and their matched metastasis for any site (Figure 4A). In this cohort, the basal cancer subtype tended to have higher scores than other subtypes not only for primary tumors, which is consistent with Figure 2, but also in metastatic tumors.



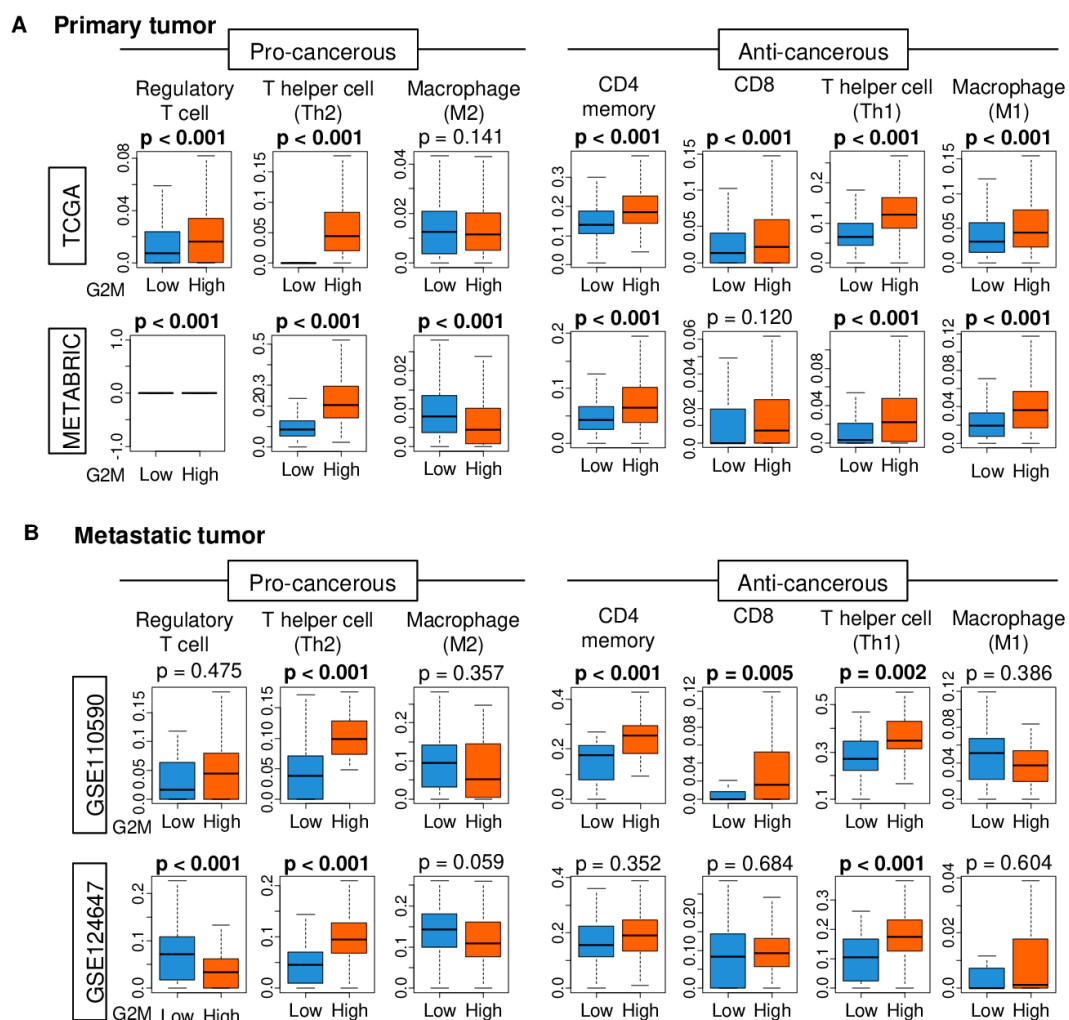
**Figure 4.** (A) Boxplots of the G2M pathway score of primary and each metastatic tumor, and matched pair comparison of G2M pathway scores between the primary and each metastatic tumor, including lymph node, brain, liver, lung and others in GSE110590 cohort. The PAM50 subtypes of cancer are indicated. (B) Boxplots of G2M pathway score of primary and each metastatic tumor, and Kaplan–Meier plots with logrank test  $p$ -values; progression-free survival (PFS) between the high- and low-G2M pathway score groups within different metastatic sites in the GSE124647 cohort. The median value of the G2M pathway score was used to divide patients into low and high groups.

Next, we hypothesized that metastatic tumors with a high G2M pathway score were also associated with a poor prognosis, like the primary tumors with a high score. In order to test this hypothesis, we used the GSE124647 cohort, which includes the gene expression of the various metastatic breast cancers associated with their progression-free survival (PFS). As with Figure 4A, there were no significant differences in the G2M pathway score among the sites of metastatic tumors. Metastatic lesions with a high G2M score were associated with a significantly worse PFS in the whole cohort of metastases as well as sub-groups of local or liver metastases, but not metastases in lymph node, bone,

soft tissue, or other metastasis sites. These findings indicate that the G2M pathway activity in tumor metastases also reflects tumor aggressiveness, as it does for primary tumors.

2.5. Immune Cell Infiltration is Higher in Tumors with High G2M Pathway Activity

As the tumor immune microenvironment is deeply involved in cancer progression, it was of interest whether primary and metastatic tumors with a high G2M pathway score are associated with unfavorable immune cell infiltration. Using the xCell algorithm on tumor gene expression data of TCGA and METABRIC primary tumor cohorts and GSE110590 and GSE124647 tumor metastasis cohorts, we found that the primary as well as metastasis tumors with a high G2M activity had significantly higher infiltration by both pro- and anti-cancerous immune cells. For primary tumors, in particular, pro-cancerous regulatory T and Th2 helper T cells, and anti-cancerous CD4+ memory T, CD8+ T, Th1 helper T, and M1 macrophage cells were increased in the high-score tumor group of both TCGA and METABRIC cohorts (Figure 5 and Figure S2). For metastasis tumors, Th1 and Th2 T cells were increased in the high-score group of metastases in both cohorts (Figure 5). This was also observed in examinations of site-specific metastases of the two cohorts (Figure S3).



**Figure 5.** Boxplots of immune cells, CD8 T cell, CD4 memory T cell, T helper cell type 1 and 2, M1 and M2 macrophage, and regulatory T cells depicting the high and low G2M scores (A) in primary tumors and (B) in metastatic tumors.



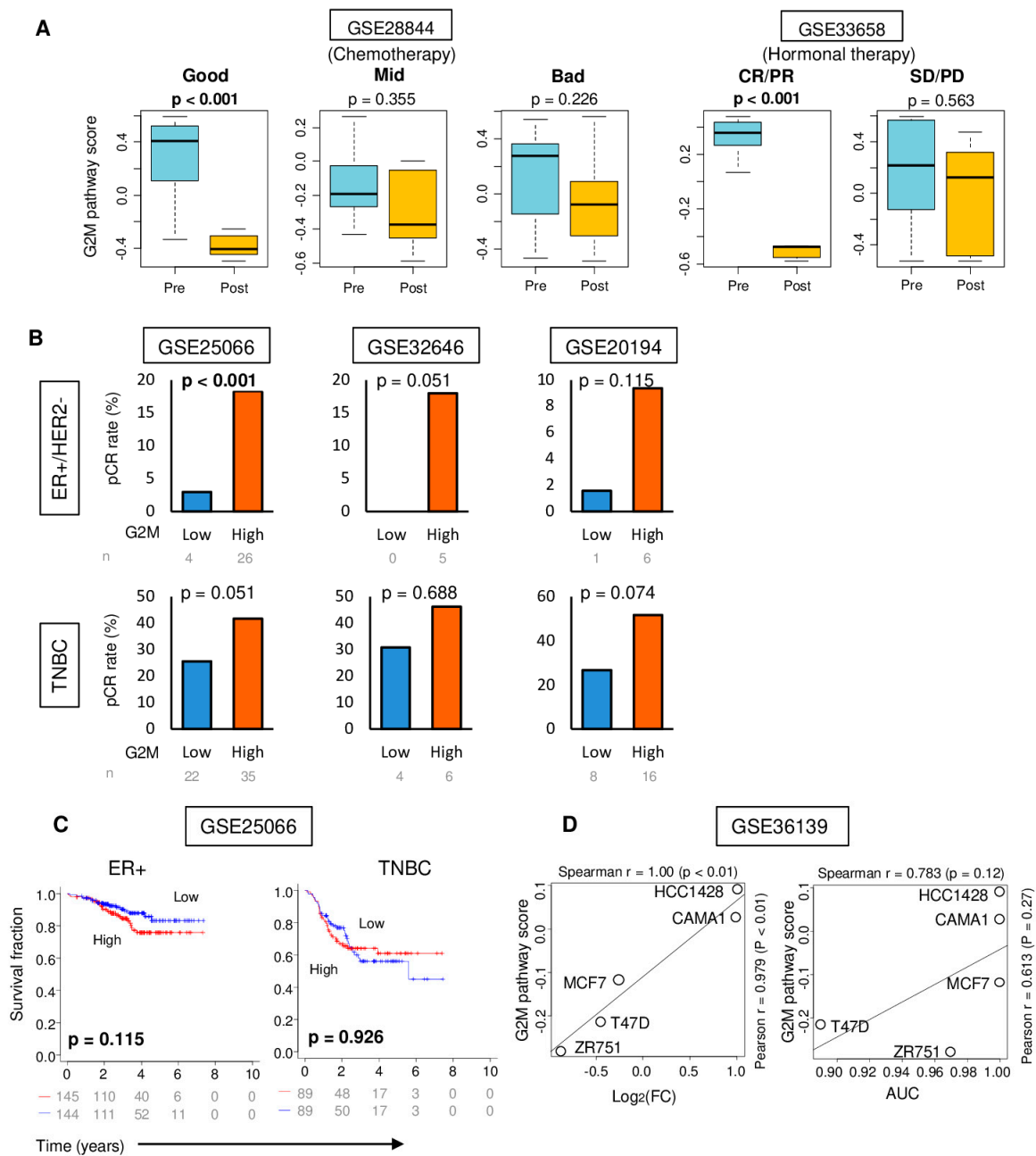
### 2.6. High G2M Pathway Score was Associated with Significantly Better Response to Chemotherapy, but Not with Improved Survival

It is well-known that highly proliferative cancers are more likely to respond to cytotoxic chemotherapy [21]. As we found that a high G2M pathway score reflects cell proliferation, we speculated that the G2M score would decrease when tumors respond to chemotherapy. To examine this, we evaluated data from the GSE28844 cohort [22] in which tumor gene expression data were obtained before and after chemotherapy with doxorubicin. Indeed, the G2M pathway score of primary tumors was significantly reduced by chemotherapy in patients who achieved a good response (t-test  $p < 0.001$ ), while the score did not change in those who had an intermediate ( $p = 0.355$ ) or poor ( $p = 0.226$ ) response (Figure 6A). Furthermore, we found that the G2M score significantly decreased ( $p < 0.001$ ) following successful endocrine therapy with anastrozole and fulvestrant in combination with gefitinib in the GSE33658 cohort [23] (Figure 6A).

As reduction in the G2M score was associated with response to chemotherapy, we further hypothesized that a high score would be associated with pathological complete response (pCR) to chemotherapy. In the GSE25066 cohort [24], we found that the high G2M pathway score prior to the treatment was associated with a significantly higher rate in achieving pCR compared with low score in ER-positive patients (Figure 6B;  $p < 0.001$ ). Although there were no statistically significant differences, there were trends towards a higher pCR rate among the high G2M pathway score groups in ER-positive patients in two other cohorts with very small sample sizes (Figure 6B). Although similar trends were observed in the high G2M pathway score groups among the TNBC patients, it was less prominent and none of them was statistically significant.

As pCR after neoadjuvant chemotherapy is often considered a surrogate marker for better survival, we further hypothesized that a high G2M pathway score was associated with a higher pCR rate and it would also translate into better survival in patients using neoadjuvant chemotherapy. To evaluate this question, we examined the GSE25066 cohort [24]. Disappointingly, the G2M pathway score demonstrated no association with DFS in neither ER-positive or TNBC subtype of cancer (Figure 6C). These findings implicate that while the G2M pathway score may be a potential predictive biomarker for response to neoadjuvant chemotherapy or endocrine therapy, it does not prognosticate improvement in survival owing to therapy.

Given that the G2M checkpoint is an important mechanism of cell cycle control, we assessed if the G2M score is associated with response to CDK inhibitors that directly affect the cell cycle. Because we were unable to identify a CDK inhibitor-treated breast cancer cohort with publicly available gene expression and treatment response data, we examined gene expression and drug response data in the Cancer Cell Line Encyclopedia [25] for five ER-positive/HER2-negative human breast cancer cell lines. The G2M score demonstrated extremely strong correlations with fold-change (FC) and area under the curve (AUC) values for CDK inhibitors (Figure 6D; Spearman  $r = 1$  and  $0.783$  with  $p < 0.01$  and  $0.12$ , respectively). This result suggests that high G2M pathway activity in breast cancer may be predictive of good response to therapy with CDK inhibitors.



**Figure 6.** (A) Boxplots depicting the scores of pre- and post- treatment using the GSE28844 and GSE33658 cohorts. GSE28844 used neoadjuvant chemotherapy and GSE33658 used endocrine therapy. In GSE28844, response was categorized as Miller–Payne criteria. *RECIST* criteria were used for GSE33658. (B) Bar plots depicting the pCR rates between high- and low G2M pathway score groups among patients with ER-positive/HER2-negative tumors and TNBCs in the GSE25066 ( $n = 508$ ), GSE32646 ( $n = 115$ ), and GSE20194 ( $n = 278$ ) cohorts. (C) Kaplan–Meier plots with logrank test  $p$ -values; DFS between the high- and low-G2M pathway score groups. ER-positive/HER2-negative and TNBC patients in the GSE25066 cohort. (D) Correlation plot between the G2M pathway score and fold-change (FC) and area under the curve (AUC) using the GSE36139 cohort. AUC, area under the curve; CR, complete response; DFS, disease-free survival; ER, estrogen-receptor; FC, fold-change; pCR, pathological complete response; PD, progressive disease; PR, partial response; *RECIST*, The Response Evaluation Criteria in Solid Tumors; SD, stable disease; TNBC, triple-negative breast cancer.

### 3. Discussion

We studied a total of 4626 breast cancer tumors for the association between their cell cycle G2M checkpoint activity and with cancer aggressiveness, metastasis, and treatment response. To the best of our knowledge, this is the first study to utilize single-sample gene set expression scoring for elucidating the clinical relevance of the G2M checkpoint.

The G2M activity score was defined as the GSVA score of the “HALLMARK\_G2M\_checkpoint” gene set [15] using its median as the cut-off. The G2M score significantly enriched expression for cell proliferation-related gene sets such as E2F targets, MYC targets v1 and v2, and the mitotic spindle in GSEA analyses. High G2M score tumor had more aggressive clinical characteristics, such as hormone receptor negative status, higher AJCC cancer staging, and higher Nottingham pathological grade. For ER-positive/HER2-negative breast cancer, high G2M score patients were significantly associated with worse survival. A high G2M score was associated with shorter metastasis-free survival, especially to the lung in ER-positive breast cancer, but not TNBC. A high score was also associated with greater infiltration by both pro- and anti-cancerous immune cells in primary as well as metastatic tumors. The G2M score decreased with a good response to chemotherapy or hormonal therapy. A high G2M score tumor was predictive of a good response to neoadjuvant chemotherapy, but this was not associated with survival benefit. In breast cancer cell-lines, we found that the G2M pathway score had strong correlations with CDK inhibitor cytotoxicity.

As CDK is considered a key molecule for several cell cycle transitions, targeting this pathway has been extensively studied in multiple cancer types during the last decade [26–29]. CDK1, the only CDK that can initiate the onset of M phase mitosis [30], is essential for the growth of cancer cells as well as normal cells. G2-M regulation mediated by CDK1 has widely been studied for cancer therapeutics in breast cancer [5,31]. Because there are so many molecules involved in the cell cycle, analysis of a single gene expression will not depict the entire picture. In order to overcome this challenge, we used the GSVA to derive a G2M pathway activity score that reflected the activity of 200 genes that are involved in the G2-M checkpoint. As demonstrated in Figure 1, the score represented underlying cell proliferation ability and was associated with other cell cycle gene sets.

The ER-positive subtype accounts for approximately 70% of advanced breast cancer, and is thus responsible for the majority of the deaths from the disease. Identification of the patients who have a higher risks for distant metastasis, especially in the ER-positive subtype, is essential to improve patient survival. While many efforts have been made to predict the response to systemic treatment or the development of metastasis, such as Oncotype DX and MammaPrint [32,33], their utility is still limited. For instance, there is no doubt that Oncotype DX is clinically useful to stratify ER-positive patients to whether they benefit from adjuvant chemotherapy, where now we know that 70% of them do not [33]. However, Oncotype DX only analyzes the expressions of 21 genes, and among them, only 5 genes (Ki67, STK15, Survivin, Cyclin B1, and MYBL2) are proliferation related. This is clearly less accurate in grasping a specific pathway, whereas the G2M score analyzes 200 genes specifically related to the pathway, and thus demonstrates an extremely strong correlation with CDK inhibition. The ability of the G2M pathway score to predict aggressiveness and metastasis shown in multiple cohorts is expected to be useful in patient selection for therapy. This is particularly important in ER-positive breast cancer, where it is well-known that they do not respond to chemotherapy like ER-negative tumors. Thus, a predictive biomarker like the G2M pathway score that identifies aggressive tumors that metastasize followed by a poor prognosis, and that correlate with response to neoadjuvant chemotherapy, is expected to improve efficacy, reduce the toxicity of chemotherapy, and improve patient quality of life. In the present study, the G2M pathway score correlated with the clinical response to both systemic chemotherapy and endocrine therapy.

Unlike in the ER-positive subtype, a high G2M pathway score did not correlate with better response to chemotherapy, nor with worse survival in either primary or metastasis tumors in the TNBC subtype. Given the data that the G2M score was significantly high in TNBC, and the fact that TNBC is biologically aggressive and initially responds to chemotherapy better than ER-positive subtype,

we hypothesize that the G2M checkpoint alone is not strong enough to predict the clinical outcome of TNBC tumors.

The utility of the G2M pathway score as a biomarker to predict pCR after neoadjuvant chemotherapy in ER-positive/HER2-negative breast cancer is significantly hindered by the fact that it did not correlate with DFS of patients who underwent neoadjuvant chemotherapy in this study. This result was disappointing because pCR after neoadjuvant chemotherapy is considered a surrogate for a better prognosis. One may argue that this result is because of the small number of ER-positive breast cancer patients who achieved pCR to chemotherapy. However, on the basis of our results that the G2M score correlated with aggressive cancer biology and worse metastasis and survival, it is more likely that cytotoxic chemotherapy was effective enough to achieve pCR in the neoadjuvant setting, but G2M high score tumors were biologically too aggressive for chemotherapy to control the disease in the adjuvant or metastatic setting. A hope is that the G2M score demonstrated strong positive correlation to cytotoxic effect of CDK inhibitors in the breast cancer cell lines that we examined. CDK, which is one of the most critical molecules for several cell cycle transitions, has recently been considered a key target to treat ER-positive breast cancer. Indeed, the results of recent clinical trials indicate that combinations of cell cycle inhibitors and other drugs may be one of the most promising therapeutic approaches to breast cancer in the future [6,34–38]. To this end, we cannot help but speculate that the G2M pathway score may have a utility to be used for patient selection and as a predictive biomarker for CDK inhibitors among patients with ER-positive breast cancer.

Although our findings are novel, our study has a few limitations. First, even though it examined two very large and well-characterized cohorts (TCGA and METABRIC), it remains a retrospective study. A prospective study will be required in order to establish the G2M pathway score as a predictive biomarker. Although we demonstrated a significant association between the G2M pathway score and the effect of CDK inhibitors in human cell lines, which may provide a clue to understanding the interaction between tumors and the response to CDK inhibitors, we were unable to show a similar predictive value of the score for the response to CDK inhibition in patients owing to limited data availability. A randomized clinical trial will be necessary to examine the biomarker utility of the G2M pathway scoring in predicting the effectiveness of CDK inhibitor treatment.

In conclusion, we have demonstrated that the G2M pathway score may serve as a useful tool for identifying patients who are likely to metastasize and have a poor survival in ER-positive/HER2-negative breast cancer. Our findings also support a clinical trial to evaluate the G2M score as a predictive biomarker for response to CDK inhibition therapy.

## 4. Materials and Methods

### 4.1. Data of The Cancer Genome Atlas Breast Cancer Cohort

Tumor gene expression and corresponding clinical data for the TCGA-BRCA project were obtained from the Pan-Cancer Clinical Data Resource [39] and through cBio Cancer Genomic Portal [40], as we previously reported [41–45]. Data for the 1065 female patients were analyzed in the study. As TCGA data do not include information on pathological grade, we used Nottingham histological grade data that had been previously manually collected from pathology reports for 573 of the 1065 patients using software from Text Information Extraction System Cancer Research Network [46]. The Nottingham histological grade is the sum of individual scores for three parameters, degree of tubular formation, nuclear pleomorphism, and mitosis [47].

### 4.2. Data of METABRIC and Other Breast Cancer Cohorts

Normalized microarray-based tumor gene expression and clinical data for 1903 patients of METABRIC cohort were obtained through the cBio portal, as we previously reported [48–50]. For the following studies, normalized microarray-based tumor gene expression and clinical data were obtained from Gene Expression Omnibus (GEO) repository (<http://www.ncbi.nlm.nih.gov/geo/>): Shi et al.

(accession number GSE20194;  $n = 278$ ) [51], Symmans et al. (GSE25066;  $n = 508$ ) [24], Vera-Ramirez et al. (GSE28844;  $n = 33$ ) [22], Noguchi et al. (GSE32646;  $n = 115$ ) [52], and Massarweh et al. (GSE33658;  $n = 11$ ) [23]. We used metastasis cohorts (Wang et al. GSE2034 [18];  $n = 286$ , Bos et al. GSE12276 [17];  $n = 204$ , Siegel et al. GSE110590 [20];  $n = 16$ , Sinn et al. GSE124647 [53];  $n = 140$ ), which have the data of metastatic tumors, to investigate the G2M pathway scores in metastatic tumors. The GSE110590 cohort contained 16 samples of the primary tumors as well as 62 samples of the metastatic tumors from patients with matching primary tumors. The GSE124647 cohort had 140 sample data about metastasis tumors. We used five sample data of ER-positive/HER2-negative breast cancer cell lines with parvo pharmacological profiles obtained from the GSE36139 cohort [25]. When necessary, probe-level microarray data for genes with multiple probes were summarized using mean.

#### 4.3. Gene Set Expression Analyses

Log<sub>2</sub>-transformed normalized gene expression data were used. The gene set variation analysis (GSVA) method [14] was utilized to obtain a GSVA score from gene expression data for the “HALLMARK\_G2M\_checkpoint” gene set of the Molecular Signatures Database Hallmark gene set collection [15]. GSVA Bioconductor package (version 3.10) was used. Within each cohort, tumor samples were categorized into high and low G2M activity score groups using the median GSVA score as cut-off. For gene set enrichment analysis (GSEA) [54], GSEA software (Java version 4.0) and the Hallmark gene set collection were used, as we described previously [55–60]. As recommended for GSEA, a false discovery rate threshold of 0.25 was used to deem significance.

#### 4.4. Other

Immune features of tumors such as composition of infiltrating immune cells were estimated from whole tumor gene expression data using xCell algorithm [61]. All statistical analyses and data plotting were performed with R software (version 3.6.2) and Microsoft Excel (version 16 for Windows, Redmond, WA, USA). Unless noted otherwise, a  $p$ -value threshold of 0.05 was used to deem significance. One-way ANOVA or Fisher’s exact test were used to determine the significance of differences for groups. The Kaplan–Meier method with logrank test was used for survival analysis. In data visualizations that are presented here, boxplots are of Tukey type, with boxes depicting median and inter-quartile range.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/21/8/2921/s1>.

**Author Contributions:** Conceptualization, I.E., K.T., M.O., R.M.; Methodology, K.T., M.O., and Y.T.; Formal Analysis, L.Y., and M.O.; Writing—Original Draft Preparation, M.O.; Writing—Review & Editing, H.T., K.T., L.Y., and O.M.R.; Supervision, K.T.; Project Administration, K.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by National Institutes of Health, USA grant R01CA160688; Susan G. Komen Foundation, grant CCR17481211 to K.T.; and National Cancer Institute, USA cancer center support grant P30-CA016056, to Roswell Park Comprehensive Cancer Center (RPCCC).

**Acknowledgments:** We thank Santosh Patnaik of RPCCC for feedback.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

#### Abbreviations

AJCC	American Joint Committee on Cancer
DFS	disease-free survival
DSS	disease-specific survival
ER	estrogen receptor
FDR	false discovery rate
GSVA	gene set variation analysis
HER2	human epidermal growth factor receptor 2

METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
NES	normalized enrichment score
OS	overall survival
pCR	pathological complete response
PFS	progression-free survival
TCGA	The Cancer Genome Atlas
TNBC	triple negative breast cancer

## References

1. Miller, K.D.; Nogueira, L.; Mariotto, A.B.; Rowland, J.H.; Yabroff, K.R.; Alfano, C.M.; Jemal, A.; Kramer, J.L.; Siegel, R.L. Cancer treatment and survivorship statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 363–385. [[CrossRef](#)]
2. Hanahan, D.; Weinberg, R.A. The hallmarks of cancer. *Cell* **2000**, *100*, 57–70. [[CrossRef](#)]
3. Chen, G.L.; Yang, L.; Rowe, T.C.; Halligan, B.D.; Tewey, K.M.; Liu, L.F. Nonintercalative antitumor drugs interfere with the breakage-reunion reaction of mammalian DNA topoisomerase II. *J. Biol. Chem.* **1984**, *259*, 13560–13566. [[PubMed](#)]
4. Lock, R.B.; Ross, W.E. Inhibition of p34cdc2 kinase activity by etoposide or irradiation as a mechanism of G2 arrest in Chinese hamster ovary cells. *Cancer Res.* **1990**, *50*, 3761–3766. [[PubMed](#)]
5. Asghar, U.; Witkiewicz, A.K.; Turner, N.C.; Knudsen, E.S. The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nat. Rev. Drug Discov.* **2015**, *14*, 130–146. [[CrossRef](#)]
6. Slamon, D.J.; Neven, P.; Chia, S.; Fasching, P.A.; De Laurentiis, M.; Im, S.A.; Petrakova, K.; Bianchi, G.V.; Esteva, F.J.; Martin, M.; et al. Phase III Randomized Study of Ribociclib and Fulvestrant in Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor 2-Negative Advanced Breast Cancer: MONALEESA-3. *J. Clin. Oncol.* **2018**, *36*, 2465–2472. [[CrossRef](#)] [[PubMed](#)]
7. Azuaje, F. Computational models for predicting drug responses in cancer research. *Brief. Bioinform.* **2017**, *18*, 820–829. [[CrossRef](#)]
8. Ein-Dor, L.; Zuk, O.; Domany, E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5923–5928. [[CrossRef](#)]
9. Shi, W.; Jiang, T.; Nuciforo, P.; Hatzis, C.; Holmes, E.; Harbeck, N.; Sotiriou, C.; Pena, L.; Loi, S.; Rosa, D.D.; et al. Pathway level alterations rather than mutations in single genes predict response to HER2-targeted therapies in the neo-ALTTO trial. *Ann. Oncol.* **2017**, *28*, 128–135. [[CrossRef](#)]
10. Wang, X.; Sun, Z.; Zimmermann, M.T.; Bugrim, A.; Kocher, J.P. Predict drug sensitivity of cancer cells with pathway activity inference. *BMC Med. Genom.* **2019**, *12*, 15. [[CrossRef](#)]
11. Lee, E.; Chuang, H.Y.; Kim, J.W.; Ideker, T.; Lee, D. Inferring pathway activity toward precise disease classification. *PLoS Comput. Biol.* **2008**, *4*, e1000217. [[CrossRef](#)] [[PubMed](#)]
12. Su, J.; Yoon, B.J.; Dougherty, E.R. Accurate and reliable cancer classification based on probabilistic inference of pathway activity. *PLoS ONE* **2009**, *4*, e8161. [[CrossRef](#)] [[PubMed](#)]
13. Khatri, P.; Sirota, M.; Butte, A.J. Ten years of pathway analysis: Current approaches and outstanding challenges. *PLoS Comput. Biol.* **2012**, *8*, e1002375. [[CrossRef](#)] [[PubMed](#)]
14. Hanzelmann, S.; Castelo, R.; Guinney, J. GSEA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform.* **2013**, *14*, 7. [[CrossRef](#)]
15. Liberzon, A.; Birger, C.; Thorvaldsdottir, H.; Ghandi, M.; Mesirov, J.P.; Tamayo, P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **2015**, *1*, 417–425. [[CrossRef](#)]
16. Wu, J.R.; Zhao, Y.; Zhou, X.P.; Qin, X. Estrogen receptor 1 and progesterone receptor are distinct biomarkers and prognostic factors in estrogen receptor-positive breast cancer: Evidence from a bioinformatic analysis. *Biomed. Pharmacother.* **2020**, *121*, 109647. [[CrossRef](#)]
17. Bos, P.D.; Zhang, X.H.; Nadal, C.; Shu, W.; Gomis, R.R.; Nguyen, D.X.; Minn, A.J.; van de Vijver, M.J.; Gerald, W.L.; Foekens, J.A.; et al. Genes that mediate breast cancer metastasis to the brain. *Nature* **2009**, *459*, 1005–1009. [[CrossRef](#)]
18. Wang, Y.; Klijn, J.G.; Zhang, Y.; Sieuwerts, A.M.; Look, M.P.; Yang, F.; Talantov, D.; Timmermans, M.; Meijer-van Gelder, M.E.; Yu, J.; et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **2005**, *365*, 671–679. [[CrossRef](#)]

19. Shah, N.; Mohammad, A.S.; Saralkar, P.; Sprowls, S.A.; Vickers, S.D.; John, D.; Tallman, R.M.; Lucke-Wold, B.P.; Jarrell, K.E.; Pinti, M.; et al. Investigational chemotherapy and novel pharmacokinetic mechanisms for the treatment of breast cancer brain metastases. *Pharmacol. Res.* **2018**, *132*, 47–68. [[CrossRef](#)]
20. Siegel, M.B.; He, X.; Hoadley, K.A.; Hoyle, A.; Pearce, J.B.; Garrett, A.L.; Kumar, S.; Moylan, V.J.; Brady, C.M.; Van Swearingen, A.E.; et al. Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Investig.* **2018**, *128*, 1371–1383. [[CrossRef](#)]
21. Witkiewicz, A.K.; Balaji, U.; Knudsen, E.S. Systematically defining single-gene determinants of response to neoadjuvant chemotherapy reveals specific biomarkers. *Clin. Cancer Res.* **2014**, *20*, 4837–4848. [[CrossRef](#)] [[PubMed](#)]
22. Vera-Ramirez, L.; Sanchez-Rovira, P.; Ramirez-Tortosa, C.L.; Quiles, J.L.; Ramirez-Tortosa, M.; Lorente, J.A. Transcriptional shift identifies a set of genes driving breast cancer chemoresistance. *PLoS ONE* **2013**, *8*, e53983. [[CrossRef](#)] [[PubMed](#)]
23. Massarweh, S.; Tham, Y.L.; Huang, J.; Sexton, K.; Weiss, H.; Tsimelzon, A.; Beyer, A.; Rimawi, M.; Cai, W.Y.; Hilsenbeck, S.; et al. A phase II neoadjuvant trial of anastrozole, fulvestrant, and gefitinib in patients with newly diagnosed estrogen receptor positive breast cancer. *Breast Cancer Res. Treat.* **2011**, *129*, 819–827. [[CrossRef](#)] [[PubMed](#)]
24. Itoh, M.; Iwamoto, T.; Matsuoka, J.; Nogami, T.; Motoki, T.; Shien, T.; Taira, N.; Niikura, N.; Hayashi, N.; Ohtani, S.; et al. Estrogen receptor (ER) mRNA expression and molecular subtype distribution in ER-negative/progesterone receptor-positive breast cancers. *Breast Cancer Res. Treat.* **2014**, *143*, 403–409. [[CrossRef](#)]
25. Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A.A.; Kim, S.; Wilson, C.J.; Lehar, J.; Kryukov, G.V.; Sonkin, D.; et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012**, *483*, 603–607. [[CrossRef](#)]
26. Meijer, L. Chemical inhibitors of cyclin-dependent kinases. *Prog. Cell Cycle Res.* **1995**, *1*, 351–363. [[CrossRef](#)]
27. Sedlacek, H.; Czech, J.; Naik, R.; Kaur, G.; Worland, P.; Losiewicz, M.; Parker, B.; Carlson, B.; Smith, A.; Senderowicz, A.; et al. Flavopiridol (L86 8275; NSC 649890), a new kinase inhibitor for tumor therapy. *Int. J. Oncol.* **1996**, *9*, 1143–1168. [[CrossRef](#)]
28. Knudsen, E.S.; Witkiewicz, A.K. The Strange Case of CDK4/6 Inhibitors: Mechanisms, Resistance, and Combination Strategies. *Trends Cancer* **2017**, *3*, 39–55. [[CrossRef](#)]
29. Chung, S.; Vail, P.; Witkiewicz, A.K.; Knudsen, E.S. Coordinately Targeting Cell-Cycle Checkpoint Functions in Integrated Models of Pancreatic Cancer. *Clin. Cancer Res.* **2019**, *25*, 2290–2304. [[CrossRef](#)]
30. Santamaria, D.; Barriere, C.; Cerqueira, A.; Hunt, S.; Tardy, C.; Newton, K.; Caceres, J.F.; Dubus, P.; Malumbres, M.; Barbacid, M. Cdk1 is sufficient to drive the mammalian cell cycle. *Nature* **2007**, *448*, 811–815. [[CrossRef](#)]
31. Johnson, N.; Li, Y.C.; Walton, Z.E.; Cheng, K.A.; Li, D.; Rodig, S.J.; Moreau, L.A.; Unitt, C.; Bronson, R.T.; Thomas, H.D.; et al. Compromised CDK1 activity sensitizes BRCA-proficient cancers to PARP inhibition. *Nat. Med.* **2011**, *17*, 875–882. [[CrossRef](#)] [[PubMed](#)]
32. Audeh, W.; Blumencranz, L.; Kling, H.; Trivedi, H.; Srkalovic, G. Prospective Validation of a Genomic Assay in Breast Cancer: The 70-gene MammaPrint Assay and the MINDACT Trial. *Acta Med. Acad.* **2019**, *48*, 18–34. [[CrossRef](#)] [[PubMed](#)]
33. Sparano, J.A.; Gray, R.J.; Makower, D.F.; Pritchard, K.I.; Albain, K.S.; Hayes, D.F.; Geyer, C.E., Jr.; Dees, E.C.; Goetz, M.P.; Olson, J.A., Jr.; et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *New Engl. J. Med.* **2018**, *379*, 111–121. [[CrossRef](#)] [[PubMed](#)]
34. He, S.; Roberts, P.J.; Sorrentino, J.A.; Bisi, J.E.; Storrer-White, H.; Tiessen, R.G.; Makhuli, K.M.; Wargin, W.A.; Tadema, H.; van Hoogdalem, E.J.; et al. Transient CDK4/6 inhibition protects hematopoietic stem cells from chemotherapy-induced exhaustion. *Sci. Transl. Med.* **2017**, *9*. [[CrossRef](#)] [[PubMed](#)]
35. Johnson, S.M.; Torrice, C.D.; Bell, J.F.; Monahan, K.B.; Jiang, Q.; Wang, Y.; Ramsey, M.R.; Jin, J.; Wong, K.K.; Su, L.; et al. Mitigation of hematologic radiation toxicity in mice through pharmacological quiescence induced by CDK4/6 inhibition. *J. Clin. Investig.* **2010**, *120*, 2528–2536. [[CrossRef](#)]
36. Witkiewicz, A.K.; Chung, S.; Brough, R.; Vail, P.; Franco, J.; Lord, C.J.; Knudsen, E.S. Targeting the Vulnerability of RB Tumor Suppressor Loss in Triple-Negative Breast Cancer. *Cell Rep.* **2018**, *22*, 1185–1199. [[CrossRef](#)]

37. Kumarasamy, V.; Ruiz, A.; Nambiar, R.; Witkiewicz, A.K.; Knudsen, E.S. Chemotherapy impacts on the cellular response to CDK4/6 inhibition: Distinct mechanisms of interaction and efficacy in models of pancreatic cancer. *Oncogene* **2020**, *39*, 1831–1845. [[CrossRef](#)]
38. Cao, J.; Zhu, Z.; Wang, H.; Nichols, T.C.; Lui, G.Y.L.; Deng, S.; Rejto, P.A.; VanArsdale, T.; Hardwick, J.S.; Weinrich, S.L.; et al. Combining CDK4/6 inhibition with taxanes enhances anti-tumor efficacy by sustained impairment of pRB-E2F pathways in squamous cell lung cancer. *Oncogene* **2019**, *38*, 4125–4141. [[CrossRef](#)]
39. Liu, J.; Lichtenberg, T.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatich, A.J.; Benz, C.C.; Levine, D.A.; Lee, A.V.; et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **2018**, *173*, 400–416.e411. [[CrossRef](#)]
40. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [[CrossRef](#)]
41. Narayanan, S.; Kawaguchi, T.; Yan, L.; Peng, X.; Qi, Q.; Takabe, K. Cytolytic Activity Score to Assess Anticancer Immunity in Colorectal Cancer. *Ann. Surg. Oncol.* **2018**, *25*, 2323–2331. [[CrossRef](#)] [[PubMed](#)]
42. Katsuta, E.; Qi, Q.; Peng, X.; Hochwald, S.N.; Yan, L.; Takabe, K. Pancreatic adenocarcinomas with mature blood vessels have better overall survival. *Sci. Rep.* **2019**, *9*, 1310. [[CrossRef](#)] [[PubMed](#)]
43. McDonald, K.A.; Kawaguchi, T.; Qi, Q.; Peng, X.; Asaoka, M.; Young, J.; Opyrchal, M.; Yan, L.; Patnaik, S.; Otsuji, E.; et al. Tumor Heterogeneity Correlates with Less Immune Response and Worse Survival in Breast Cancer Patients. *Ann. Surg. Oncol.* **2019**, *26*, 2191–2199. [[CrossRef](#)] [[PubMed](#)]
44. Takahashi, H.; Katsuta, E.; Yan, L.; Dasgupta, S.; Takabe, K. High expression of Annexin A2 is associated with DNA repair, metabolic alteration, and worse survival in pancreatic ductal adenocarcinoma. *Surgery* **2019**, *166*, 150–156. [[CrossRef](#)]
45. Katsuta, E.; Maawy, A.A.; Yan, L.; Takabe, K. High expression of bone morphogenetic protein (BMP) 6 and BMP7 are associated with higher immune cell infiltration and better survival in estrogen receptorpositive breast cancer. *Oncol. Rep.* **2019**, *42*, 1413–1421. [[CrossRef](#)]
46. Jacobson, R.S.; Becich, M.J.; Bollag, R.J.; Chavan, G.; Corrigan, J.; Dhir, R.; Feldman, M.D.; Gaudioso, C.; Legowski, E.; Maihle, N.J.; et al. A Federated Network for Translational Cancer Research Using Clinical Data and Biospecimens. *Cancer Res.* **2015**, *75*, 5194–5201. [[CrossRef](#)]
47. Rakha, E.A.; El-Sayed, M.E.; Lee, A.H.; Elston, C.W.; Grainge, M.J.; Hodi, Z.; Blamey, R.W.; Ellis, I.O. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol.* **2008**, *26*, 3153–3158. [[CrossRef](#)]
48. Yamada, A.; Nagahashi, M.; Aoyagi, T.; Huang, W.C.; Lima, S.; Hait, N.C.; Maiti, A.; Kida, K.; Terracina, K.P.; Miyazaki, H.; et al. ABCC1-Exported Sphingosine-1-phosphate, Produced by Sphingosine Kinase 1, Shortens Survival of Mice and Patients with Breast Cancer. *Mol. Cancer Res. Mcr* **2018**, *16*, 1059–1070. [[CrossRef](#)]
49. Okano, M.; Oshi, M.; Butash, A.L.; Asaoka, M.; Katsuta, E.; Peng, X.; Qi, Q.; Yan, L.; Takabe, K. Estrogen Receptor Positive Breast Cancer with High Expression of Androgen Receptor has Less Cytolytic Activity and Worse Response to Neoadjuvant Chemotherapy but Better Survival. *Int. J. Mol. Sci.* **2019**, *20*, 2655. [[CrossRef](#)]
50. Okano, M.; Oshi, M.; Butash, A.L.; Katsuta, E.; Tachibana, K.; Saito, K.; Okayama, H.; Peng, X.; Yan, L.; Kono, K.; et al. Triple-Negative Breast Cancer with High Levels of Annexin A1 Expression Is Associated with Mast Cell Infiltration, Inflammation, and Angiogenesis. *Int. J. Mol. Sci.* **2019**, *20*, 4197. [[CrossRef](#)]
51. Shi, L.; Campbell, G.; Jones, W.D.; Campagne, F.; Wen, Z.; Walker, S.J.; Su, Z.; Chu, T.M.; Goodsaid, F.M.; Pusztai, L.; et al. 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. [[CrossRef](#)] [[PubMed](#)]
52. Miyake, T.; Nakayama, T.; Naoi, Y.; Yamamoto, N.; Otani, Y.; Kim, S.J.; Shimazu, K.; Shimomura, A.; Maruyama, N.; Tamaki, Y.; et al. GSTP1 expression predicts poor pathological complete response to neoadjuvant chemotherapy in ER-negative breast cancer. *Cancer Sci.* **2012**, *103*, 913–920. [[CrossRef](#)] [[PubMed](#)]
53. Sinn, B.V.; Fu, C.; Lau, R.; Litton, J.; Tsai, T.H.; Murthy, R.; Tam, A.; Andreopoulou, E.; Gong, Y.; Murthy, R.; et al. SETER/PR: A robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. *Npj Breast Cancer* **2019**, *5*, 16. [[CrossRef](#)]



54. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 15545–15550. [[CrossRef](#)] [[PubMed](#)]
55. Narayanan, S.; Kawaguchi, T.; Peng, X.; Qi, Q.; Liu, S.; Yan, L.; Takabe, K. Tumor Infiltrating Lymphocytes and Macrophages Improve Survival in Microsatellite Unstable Colorectal Cancer. *Sci. Rep.* **2019**, *9*, 13455. [[CrossRef](#)] [[PubMed](#)]
56. Asaoka, M.; Ishikawa, T.; Takabe, K.; Patnaik, S.K. APOBEC3-Mediated RNA Editing in Breast Cancer is Associated with Heightened Immune Activity and Improved Survival. *Int. J. Mol. Sci.* **2019**, *20*, 5621. [[CrossRef](#)] [[PubMed](#)]
57. Takeshita, T.; Asaoka, M.; Katsuta, E.; Photiadis, S.J.; Narayanan, S.; Yan, L.; Takabe, K. High expression of polo-like kinase 1 is associated with TP53 inactivation, DNA repair deficiency, and worse prognosis in ER positive Her2 negative breast cancer. *Am. J. Transl. Res.* **2019**, *11*, 6507–6521. [[PubMed](#)]
58. Takeshita, T.; Yan, L.; Asaoka, M.; Rashid, O.; Takabe, K. Late recurrence of breast cancer is associated with pro-cancerous immune microenvironment in the primary tumor. *Sci. Rep.* **2019**, *9*, 16942. [[CrossRef](#)]
59. Katsuta, E.; Yan, L.; Takeshita, T.; McDonald, K.A.; Dasgupta, S.; Opyrchal, M.; Takabe, K. High MYC mRNA Expression Is More Clinically Relevant than MYC DNA Amplification in Triple-Negative Breast Cancer. *Int. J. Mol. Sci.* **2019**, *21*, 217. [[CrossRef](#)]
60. Takahashi, H.; Asaoka, M.; Yan, L.; Rashid, O.M.; Oshi, M.; Ishikawa, T.; Nagahashi, M.; Takabe, K. Biologically Aggressive Phenotype and Anti-cancer Immunity Counterbalance in Breast Cancer with High Mutation Rate. *Sci. Rep.* **2020**, *10*, 1852. [[CrossRef](#)]
61. Aran, D.; Hu, Z.; Butte, A.J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* **2017**, *18*, 220. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).