

Received: 2020.11.03

Accepted: 2021.02.17

Available online: 2021.04.19

Published: 2021.08.05

# N6-Methyladenosine Regulators Are Involved in the Progression of and Have Clinical Impact on Breast Cancer

Authors' Contribution:  
Study Design A  
Data Collection B  
Statistical Analysis C  
Data Interpretation D  
Manuscript Preparation E  
Literature Search F  
Funds Collection G

**BCEFG 1,2 Yanni Song**  
**C 3 Chaojing Zheng**  
**D 3 Yangbao Tao**  
**E 3 Rui Huang**  
**AE 4 Qian Zhang**

1 Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang, PR China  
2 Department of Plastic and Cosmetic Surgery, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, PR China  
3 Department of Colorectal Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, PR China  
4 Department of Colorectal Surgery, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, Zhejiang, PR China

**Corresponding Author:** Qian Zhang, e-mail: Zhangqian1171@126.com

**Source of support:** The study was funded by the Haiyan Research Fund of Harbin Medical University Cancer Hospital (JJZD2020-04), Heilongjiang Natural Science Foundation (LH2020H120), and Ministry of Education Chunhui Project Cooperative Research Project (HLJ2020-10)

**Background:** N6-methyladenosine (m<sup>6</sup>A) modification has been widely studied in various cancers, and m<sup>6</sup>A regulators, such as METTL3, METTL14, WTAP, and YTHDF1, play crucial roles in breast cancer. However, a comprehensive study of m<sup>6</sup>A regulators in breast cancer is still lacking.


**Material/Methods:** Expression data of m<sup>6</sup>A regulators and clinicopathological information were acquired from The Cancer Genome Atlas (TCGA) program. Protein interaction was collected from the STRING database. Data on tumor purity and correlation among m<sup>6</sup>A regulators were obtained from the TIMER database. LASSO, consensus clustering, and gene set enrichment analysis (GSEA) were used to evaluate the role of m<sup>6</sup>A regulators. Moreover, the prognostic value of m<sup>6</sup>A-related genomic targets in breast cancer was analyzed by Kaplan-Meier analysis and Cox regression models.

**Results:** We found most m<sup>6</sup>A regulators were associated with key clinicopathological parameters, such as tumor staging, Nottingham prognostic index (NPI), and cellularity. Also, consensus clustering analysis-based grouping could effectively predict patients' overall survival. Correlation analysis also showed that these regulators interacted with each other. Patients were further split into a high-risk group and low-risk group based on Cox and LASSO analysis. High-risk patients had a significantly worse overall survival than did low-risk patients. Moreover, AKT1 and MYC were enriched in patients in the high-risk group, according to GSEA analysis. The patients in the high-risk group also displayed resistance to chemoradiotherapy or hormone therapy.

**Conclusions:** The m<sup>6</sup>A regulators are critical participants in the development and progression of breast cancer and are likely to be used to predict prognosis and develop treatment strategies.

**Keywords:** Breast Neoplasms • Chemoradiotherapy • Methylation • Prognosis

**Full-text PDF:** <https://www.medscimonit.com/abstract/index/idArt/929615>

 2979

 2

 4

 42



## Background

Breast cancer is one of the most common and lethal malignancies in women worldwide. It is estimated that about 1 806 590 new cases and 606 520 deaths will occur in 2020 [1], and therefore immediate action is needed. Although enormous progress has been achieved in the past decades, with the development of surgical skills, diagnosis, and treatment modalities, many questions remain to be answered, such as specific risk factors of breast cancer, limits of immunotherapy, and identification of patients in need of treatment [2].

Epigenetics, mainly including histone modification, chromosome remodeling, acetylation, and methylation, could influence genes' transcription and translation without changing sequences of DNA [3]. Epigenetics is also involved in the carcinogenesis, development, and prognosis of breast cancer [4]. As a notable methylation form, N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) was first discovered and studied in the 1970s and since then has been viewed as the main methylation modification in eukaryotic cells [5]. The m<sup>6</sup>A regulators are a group of molecules that regulate m<sup>6</sup>A by adding, erasing, or recognizing the methyl group at the N<sup>6</sup> position of adenosine. The m<sup>6</sup>A process depends on these regulators: methyltransferases ("writers"), demethylases ("erasers"), and binding proteins ("readers"). "Writers" install m<sup>6</sup>A units to RNA and are mainly composed of Wilms tumor 1- associated protein (WTAP) [6], methyltransferase-like 3 (METTL3) [7], and METTL14 [7]. "Erasers" can remove the m<sup>6</sup>A unit from molecules, and consist of  $\alpha$ -ketoglutarate-dependent dioxygenase alkB homolog 5 (ALKBH5) [8] and fat-mass and obesity-associated protein (FTO) [9]. The modification of m<sup>6</sup>A is recognized by the "readers", which contain heterogeneous nuclear ribonucleoproteins (HNRNPs), YTH521-B homology (YTH) family members, and insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs) [10]. In addition to the above, METTL16, RBM15, KIAA1429, RBM15B, EIF3, ZC3H13, and CBLL1 were included in our analysis. In all, 20 m<sup>6</sup>A regulators were selected in our study, based on previous research and reviews [11,12].

A growing number of studies are focusing on the role of m<sup>6</sup>A in breast cancer. Data have shown that m<sup>6</sup>A modification regulates breast cancer progression [13] and is correlated with prognosis [14]. The knockdown of METTL3 could decrease the m<sup>6</sup>A methylation level, and therefore inhibit breast cancer proliferation and tumor growth [13]. It has been demonstrated that upregulated YTHDF1, YTHDF3, and KIAA1429 can predict poor overall survival (OS). Overexpression of YTHDF3 is an independent factor for OS in patients with breast cancer [14]. To fully understand the role of m<sup>6</sup>A in breast cancer, we used data from The Cancer Genome Atlas (TCGA) to comprehensively analyze RNA expression patterns and their potential relevance to clinical characteristics in breast cancer. Our goal was to provide new evidence about the potential

mechanisms of m<sup>6</sup>A regulators and their involvement in the development and progression of breast cancer and to clarify how this evidence could be used as a tool to predict prognosis of breast cancer patients.

## Material and Methods

### Data Collection and Analysis

The R package `getFirehoseData` was used to acquire breast cancer data from the TCGA dataset (Breast Cancer, METABRIC, Nature 2012 & Nature Commun 2016). A total of 1904 cases with gene expression and clinical data were included in our analysis. Clinical information and RNA expression of breast cancer patients were downloaded and processed with R version 3.6.1.

Protein interaction data were downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database [15] (<https://string-db.org/>) and then analyzed with Cytoscape (<https://cytoscape.org/>).

The Tumor Immune Estimation Resource (TIMER) [16] (<https://cistrome.shinyapps.io/timer/>) was adopted to obtain the correlation among different m<sup>6</sup>A regulators after adjusting by tumor purity.

### Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) (<http://software.broadinstitute.org/gsea/index.jsp>) was performed to determine whether a series of a previously defined set of genes were enriched in different biological groups [17]. Terms with  $P < 0.05$  and  $FDR < 0.25$  were identified. The hallmark gene set "h.all.v6.0.symbols.gmt" was used in this study.

### Least Absolute Shrinkage and Selection Operator

Least absolute shrinkage and selection operator (LASSO) is a regression analysis method that performs both variable selection and regularization to enhance the prediction accuracy and interpretability of the statistical model it produces. In our study, LASSO was utilized to evaluate the weight of each m<sup>6</sup>A regulator in the prognosis of patients [18]. The risk score of each patient was calculated with the following equation: risk score  $\sum_{i=1}^n Coef(i) * \chi(i)$ , where  $coef(i)$  was the value of each gene from LASSO analysis and  $\chi(i)$  was the expression value of each gene.

### Consensus Clustering Analysis

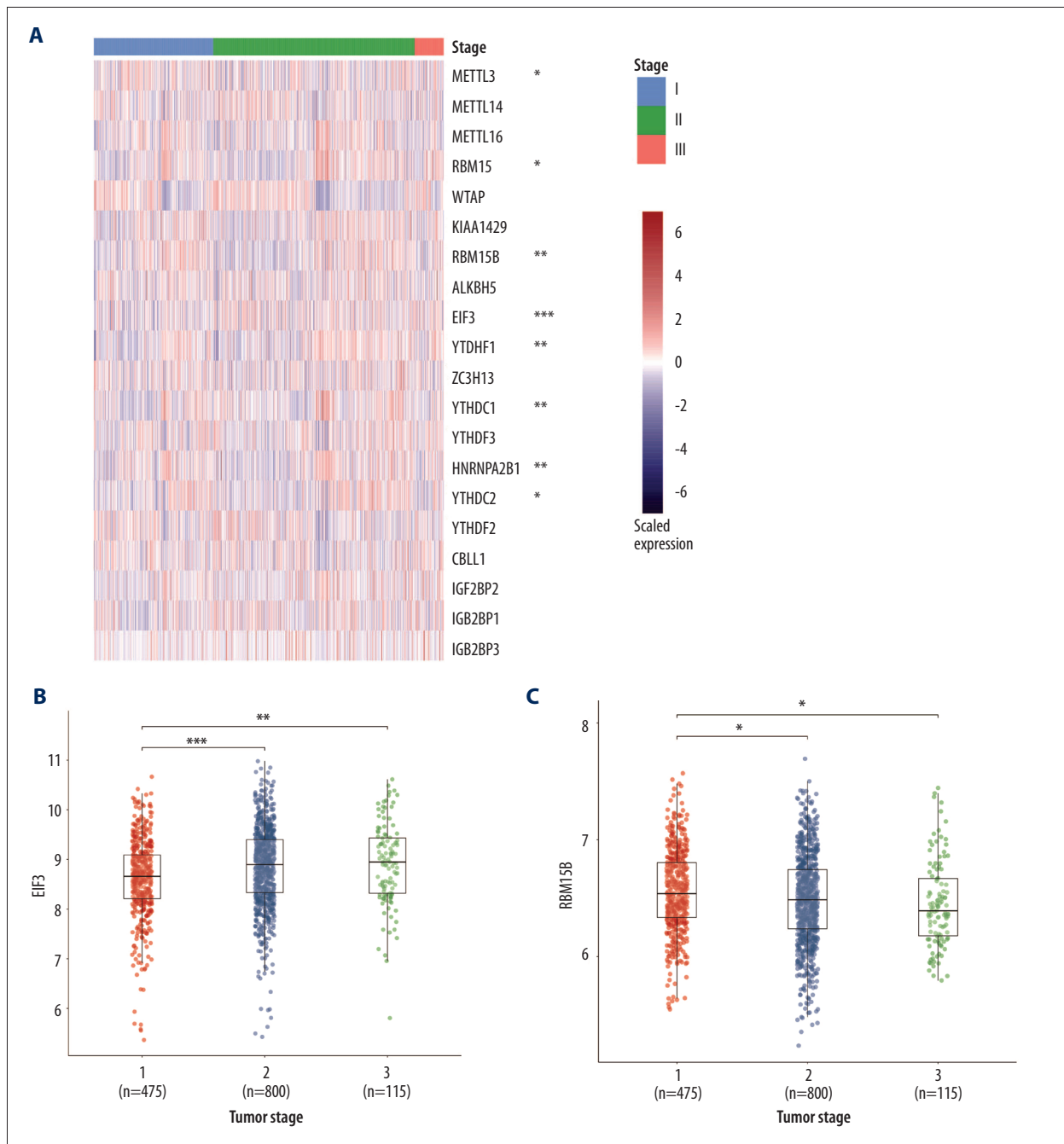
The `ConsensusClusterPlus` package in R was used to perform consensus cluster analysis in our study. The algorithm started

with subsetting a part of the characteristics and items in the data file, and each subset was divided into k groups according to k-means, the clustering algorithm used. This procedure was performed for a customized number of repetitions to yield a way to represent the consensus among several runs of the clustering and evaluate the solidity of the discovered clustering. Pairwise consensus values, the ratio of clustering runs in which 2 items are grouped together, were computed and kept in a consensus matrix for each k [19]. The clusters of patients with breast cancer were obtained according to the expression

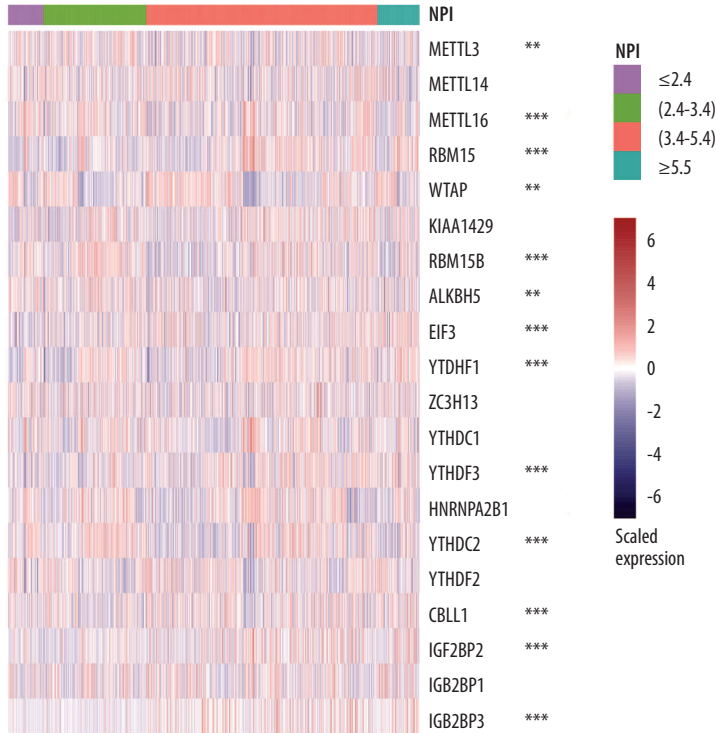
of m<sup>6</sup>A regulators. The prognosis of the different clusters was then compared.

### Ethics Statement

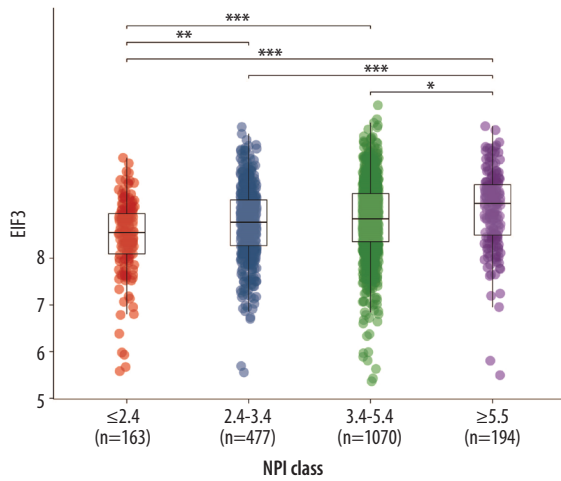
The study was approved by the institutional review board of our institution.



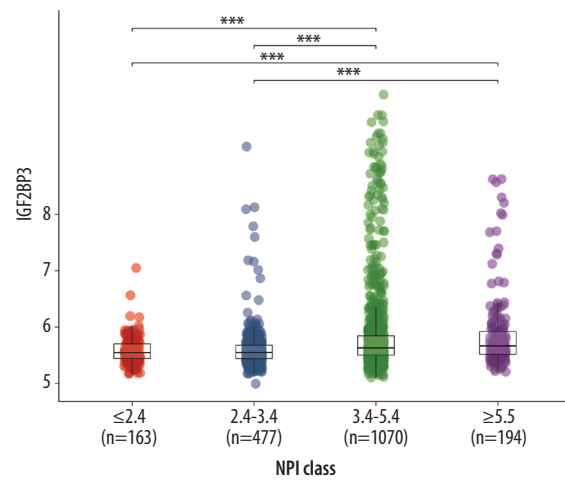
**D**

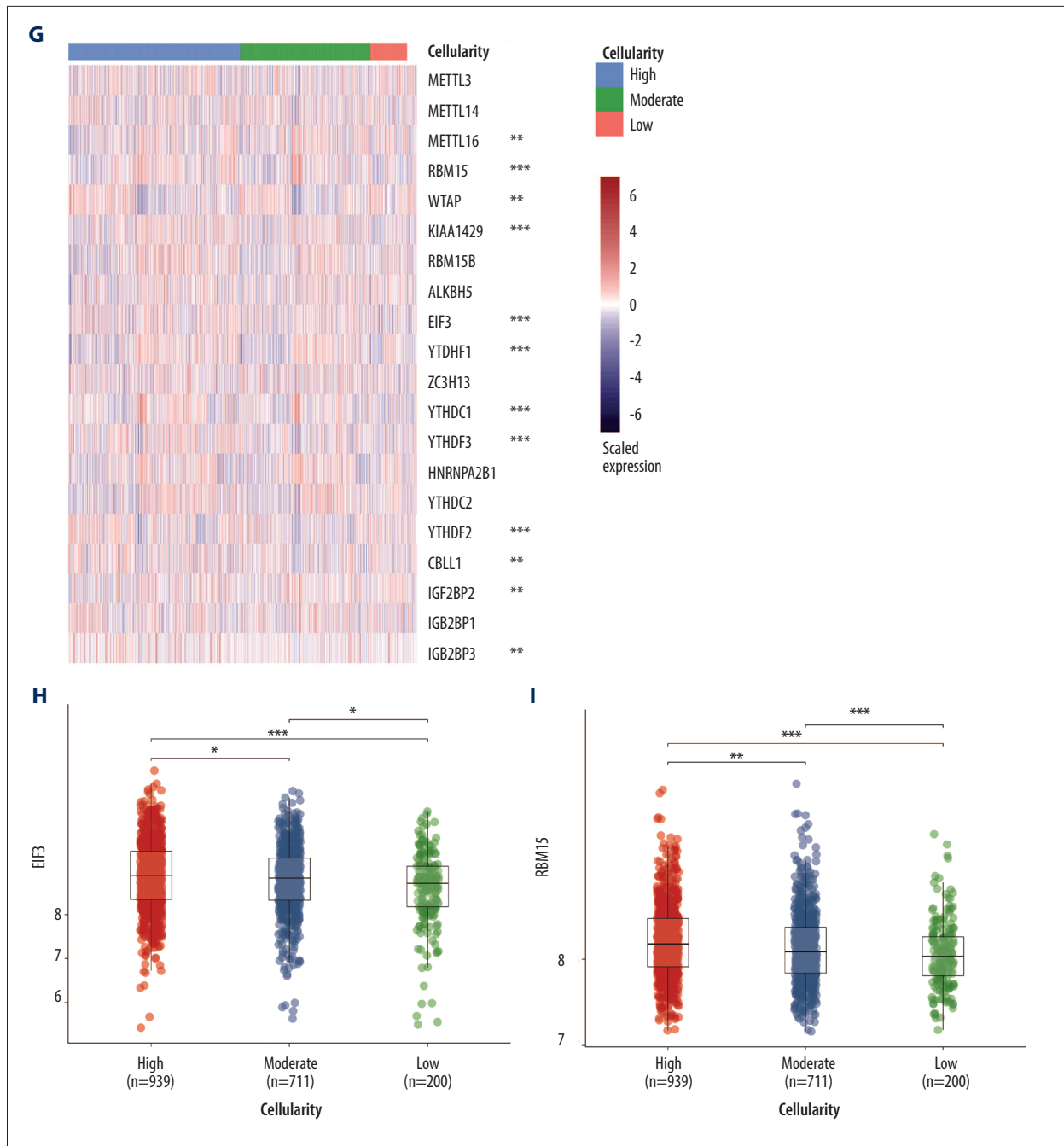


**E**



**F**





**Figure 1.** RNA expression and clinicopathological characteristics. (A) Heatmap of m<sup>6</sup>A regulators expression in patients with different tumor stages and significance. (B) EIF3 RNA expression in patients with stages I, II, and III. (C) RBM15B RNA expression in patients with stages I, II, and III. (D) Heatmap of m<sup>6</sup>A regulators expression in patients of different Nottingham prognostic index (NPI). (E) EIF3 RNA expression in patients with different NPI. (F) IGF2BP3 RNA expression patients with different NPI. (G) Heatmap of m<sup>6</sup>A regulators expression in patients of different cellularity. (H) EIF3 RNA expression in patients of different cellularity. (I) RBM15 RNA expression patients of different cellularity. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .

## Statistics

Chi-square tests or one-way ANOVA were applied to compare clinical and pathological data. Univariate and multivariate Cox regression analyses were conducted to evaluate the prognostic value of multiple parameters and characteristics. The Kaplan-Meier method was used to compare the OS of patients in different groups. All statistical analyses were conducted using R v3.6.1 (<https://www.r-project.org/>) or SPSS v19.0 (SPSS Inc., Chicago, IL, USA).

## Results

### RNA Expression and Clinicopathological Characteristics

The m<sup>6</sup>A regulators used in our study were collected from published studies, and 20 molecules were included. The Nottingham prognostic index (NPI) was used to predict the outcomes of breast cancers following surgery [20]. First, we evaluated the potential relationships between the RNA expression of m<sup>6</sup>A regulators and NPI, cancer cellularity, and tumor staging. As shown in **Figure 1A**, about half of the 20 m<sup>6</sup>A regulators were statistically significantly related with tumor staging. In particular, EIF3 expression increased as tumor staging increased from stage I to stage III (**Figure 1B**). By contrast, the gradually decreased expression of RBM15B was observed as tumor staging increased (**Figure 1C**). Also, the link of NPI to m<sup>6</sup>A regulators was determined. Interestingly, most of the genes were associated with the NPI value (**Figure 1D**). EIF3 and IGF2BP3 expression were further invalidated in different NPI groups. Results showed a growing trend of EIF3 expression as NPI value became greater (**Figure 1E**), and IGF2BP3 showed the same expression pattern as EIF3 (**Figure 1F**). In addition, the role of m<sup>6</sup>A regulators in cellularity was assessed. Those results indicated that more than half of these genes were involved in cellularity distribution at a statistically significant level (**Figure 1G**). EIF3 (**Figure 1H**) and RBM15 (**Figure 1I**) showed similar expression patterns, in which both decreased as the cellularity proportion shifted from high to low. We also displayed the expression of m<sup>6</sup>A regulators across different tumor stages, cellularity, and NPI scores in **Supplementary Table 1**.

### Consensus Clustering of m<sup>6</sup>A Regulators

We also performed consensus clustering to yield an appropriate k value that could perfectly separate patients with breast cancer. k=4 was selected as an adequate choice, with the stability of clustering rising from k = 2 to 10, according to the expression likeness of the m<sup>6</sup>A regulators (**Figure 2A, 2B**). From the k value of 4, patients were divided into 4 clusters based on the consensus clustering results. Survival analysis was then conducted to compare whether the 4 groups of patients

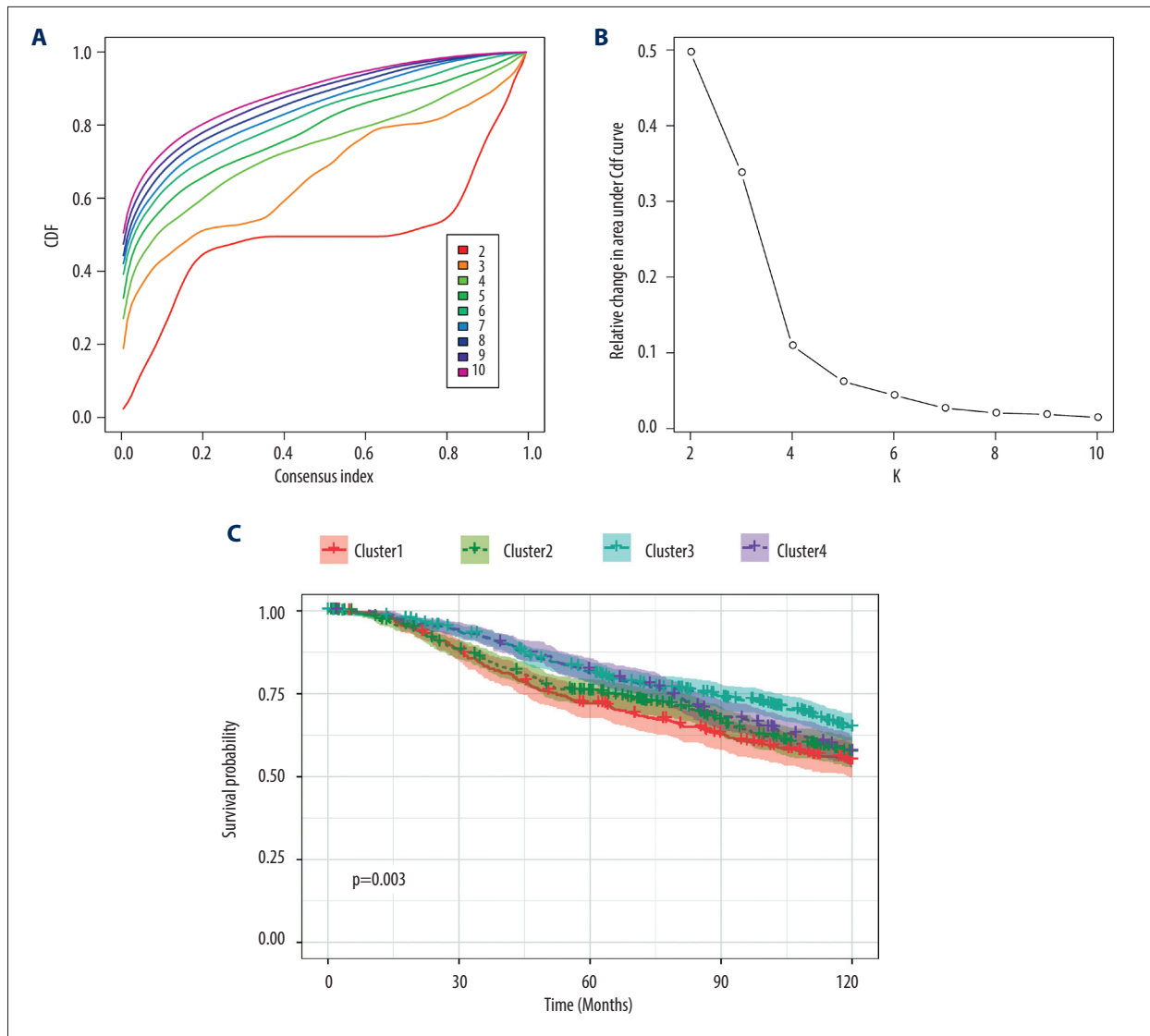
showed distinct 10-year OS. As shown in **Figure 2C**, patients in different clusters had significantly different outcomes, indicating that the clustering was powerful enough to predict the prognosis of patients with breast cancer.

### Interaction Among m<sup>6</sup>A Regulators

Because the regulators were all involved in m<sup>6</sup>A modification, they were expected to be closely connected with each other. We first searched the STRING database to evaluate the genes' protein interaction with each other. As shown in **Figure 3A**, the identified genes in STRING were closely correlated, indicating that these m<sup>6</sup>A regulators were functionally interrelated at the protein level. Then, the correlation between the regulators' RNA in the TCGA database was further evaluated. The results indicated that most of the RNAs were related with each other, although with a small coefficient (**Figure 3B**), which was not consistent with protein interactions or with the results from a previously published study [11]. This might be due to the tumor heterogeneity across different types of cancers or tumor purity. Because the TCGA dataset was not adjusted by tumor purity, we analyzed data from the TIMER database, in which gene expression is adjusted by tumor purity, to thoroughly measure the link among the regulators. With the TIMER database, we found some stronger associations among these regulators than with raw data from the TCGA database. Four pairs of regulators that were closely related to each other are shown in **Figure 3C**.

### Clinical Relevance of m<sup>6</sup>A Regulators

Based on the above findings, we believed that these regulators contributed to patients' outcomes. Cox regression analysis was then used to estimate the role of these genes in the OS of breast cancer patients. The results showed that most of the modification regulators were related with poor OS and 5 of them could predict OS ( $P < 0.05$ ) (**Figure 4A**). LASSO analysis was then used to additionally assess the effect of the 5 genes on prognosis, and the risk coefficient was acquired for each of the 5 regulators (**Figure 4B**). The results were consistent with the Cox regression analysis, showing that the LASSO risk coefficient was not zero. Subsequently, patients were divided into 2 groups according to the median risk score, which was calculated with the following equation: risk score  $\sum_{i=1}^n Coef(i) * \chi(i)$ . The survival analysis indicated that the risk score could effectively separate patients' survival ( $P < 0.0001$ ) (**Figure 4C**). Multivariate analysis also demonstrated that the risk score was an independent prognosis prediction factor (hazard ratio 1.23, 95% confidence interval [CI] 1.08-1.39;  $P = 0.002$ ) (**Table 1**). These results showed the strong power of m<sup>6</sup>A regulators in predicting prognosis of breast cancer. In addition, other independent prognosis prediction factors in our study were patient age, radiotherapy, HER2 status, positive lymph node number,



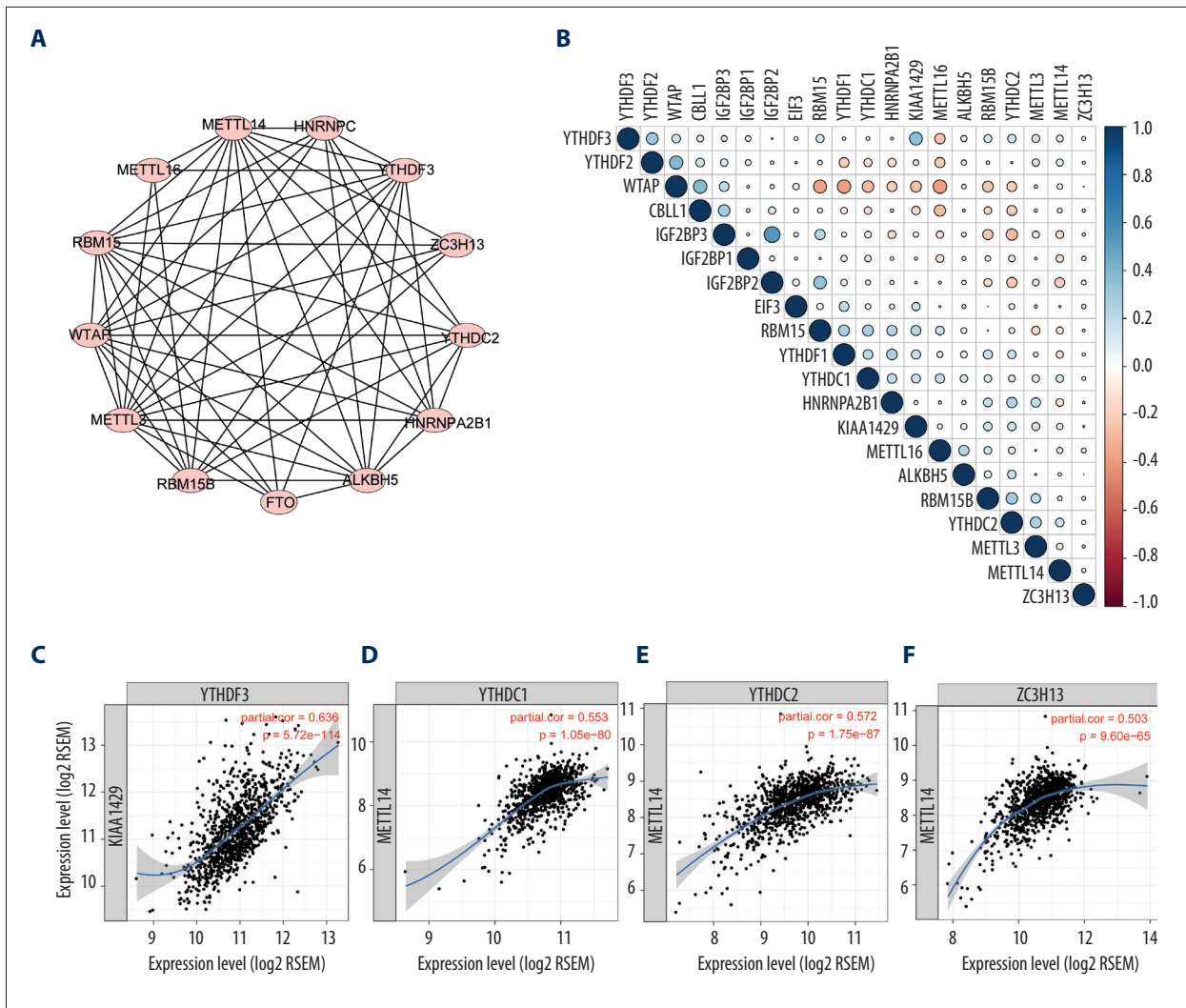
**Figure 2.** Consensus clustering of m<sup>6</sup>A regulators. (A) Consensus clustering cumulative distribution function (CDF) for k=2 to 10. (B) Relative change in area under CDF curve for k=2 to 10. (C) Overall survival analysis for breast cancer patients in 4 clusters.

and tumor size (Table 1). GSEA analysis was then conducted to explore the potential underlying mechanisms, pathways, and cancer hallmarks. As shown in Figure 4D, 4 common cancer-related pathways, AKT1 signaling, cytokinesis, RNA transcription, and APC-MYC pathway, were enriched in patients with a high-risk score. Next, we explored whether the risk score influenced the effect of radiotherapy, chemotherapy, and hormone therapy on breast cancer. First, we identified patients who had received chemotherapy, radiotherapy, or hormone therapy. Then, survival analysis was performed in these patients. The results suggested that patients with a high risk score had relatively worse survival compared with those with a low risk score (Figure 4E), indicating that high-risk cases were not as sensitive to chemoradiotherapy or hormone therapies as were

low-risk cases. Therefore, we recommend that patients with a high risk should receive closer follow-up to aid in early diagnosis and treatment of recurrence.

## Discussion

Prior work has documented the role of m<sup>6</sup>A regulators in breast carcinogenesis, development, and drug resistance; for example, researchers found that FTO can promote the progression of breast tumors [21] and HNRNPA2/B1 is altered in endocrine-resistant breast cancer [22]. However, these studies have either included a single m<sup>6</sup>A regulator or were not focused on clinical relevance. In the present study, we studied the m<sup>6</sup>A



**Figure 3.** Interaction among m<sup>6</sup>A regulators. **(A)** Interactions across the m<sup>6</sup>A regulators from STRING database. **(B)** Correlation plot of m<sup>6</sup>A regulators RNA from the TCGA database. **(C)** Representative figures of correlations of m<sup>6</sup>A regulators adjusted by tumor purity from the TIMER database.

regulators as a whole in breast cancer to test their relevance with clinicopathological features.

We identified that the RNA expression of m<sup>6</sup>A regulators is closely related to prognosis and other clinical features such as the NPI and cellularity of breast cancer. Here, we first divided breast cancer patients into 4 clusters by consensus clustering and found that the clusters significantly affected the prognosis of patients with breast cancer. Also, patients were further split into 2 subgroups according to the risk score value. Risk was correlated with prognosis and influenced cancer crucial pathways and biological processes of breast cancer.

Twenty m<sup>6</sup>A regulators have been included in this study. These regulators were searched from previously published data and have been validated to be involved in m<sup>6</sup>A modification in

various cancer types besides breast cancer. According to the theory of tumor heterogeneity [23], which states that different tumor cells and tissues of distinct morphological and phenotypic profiles exist in all cancers, the function of some m<sup>6</sup>A regulators in breast cancer might be different from that in other cancers, where the role of the m<sup>6</sup>A regulators have been validated. This could lead to some inconsistencies between our results and those of other studies. For example, WTAP was reported to promote genes in various cancers, including hepatocellular cancer [24], bladder cancer [25], and ovarian cancer [26]. However, in the present study, WTAP was found to be a protective factor of OS. The inconsistency could be the result of the limited sample size and gap between clinical and experimental differences in the present study. CBL1 is mainly studied in lung cancer and has been recognized to promote cancer progression in lung cancer [27]. It was also related to a

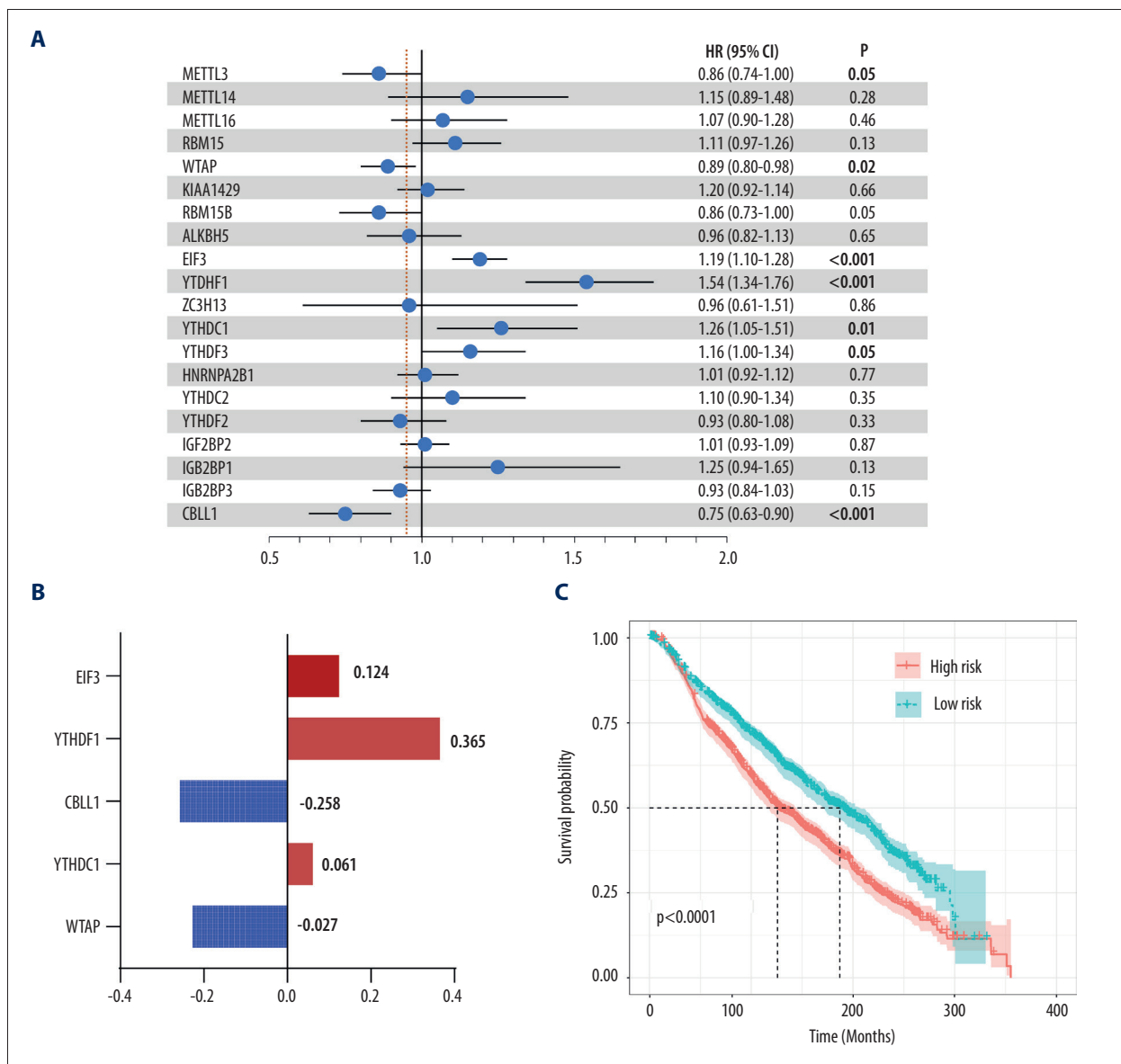


favorable prognosis in our study, which warrants more studies to explore the potential effect of CBLL1 in breast cancer and the tumor heterogeneity between lung cancer and breast cancer. Our findings also indicated that dysregulation of certain m<sup>6</sup>A regulators resulted in alterations of RNAs in cancers and validated that the same regulators might exert divergent functions in different tumors.

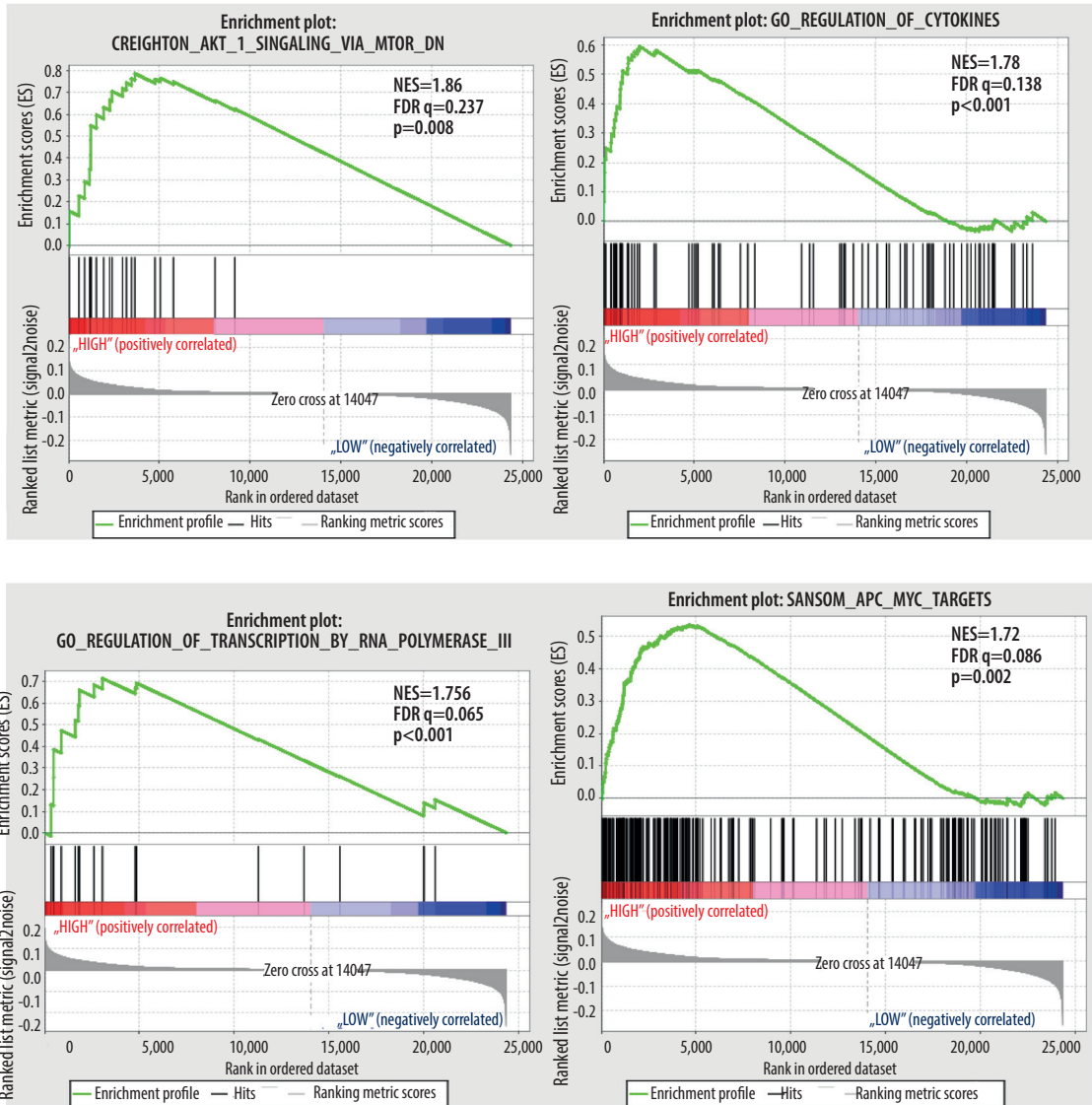
In our study, the relationship between m<sup>6</sup>A regulators and clinicopathological characteristics was also analyzed. As a “reader”, EIF3 could be directly recruited by m<sup>6</sup>A methylation, which is followed by a cap-independent translation in the cellular stress response [28]. EIF3 is closely linked to the tumor stage, NPI, and cellularity of breast cancer. Also, EIF3 was associated with poor survival of breast cancer in the present study, which

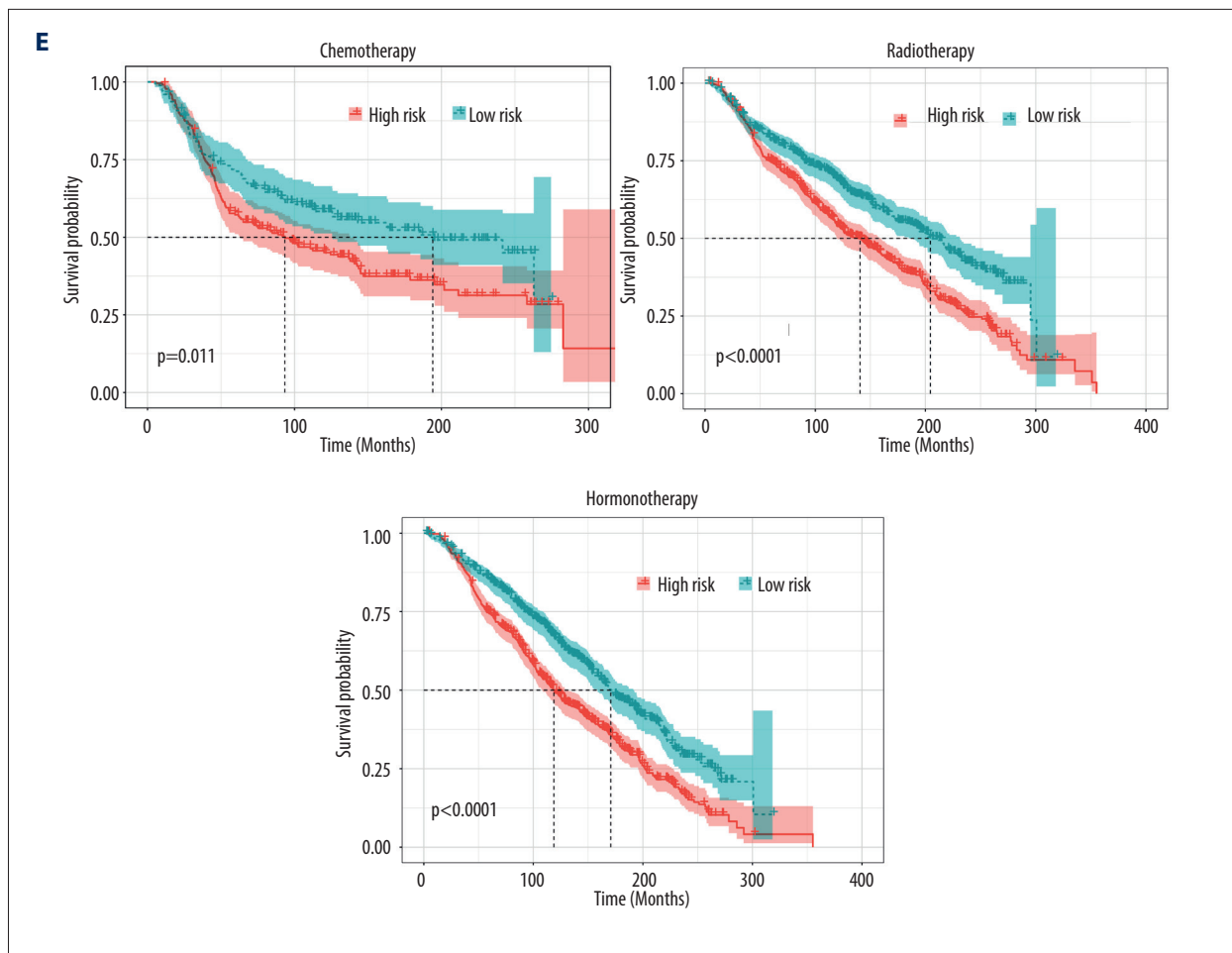
is consistent with previous research showing that EIF3 contributes to poor survival of patients with HER2 (+) breast cancer [29] and triple-negative breast cancer [30]. These pieces of evidence indicate that EIF3 has a pivotal role of in breast cancer. In addition, most of the regulators were found to be associated with breast cancer staging, NPI, and cellularity in our study, which is consistent with previous studies that found m<sup>6</sup>A is widely involved in cancer progression and prognosis [31].

It has been demonstrated that FTO can regulate cell migration and invasion in breast cancer via the miR-181b-3p/ARL5B signaling pathway, and high FTO expression is associated with advanced breast cancer staging [32]. A recent study also showed that tumor status and stage were relevant to the expression level of m<sup>6</sup>A RNA methylation regulators [33]. The m<sup>6</sup>A in peripheral



D





**Figure 4.** Clinical relevance of m<sup>6</sup>A regulators. (A) Forest plot of univariate Cox regression results of m<sup>6</sup>A regulators. (B) Coefficients calculated by multivariate Cox regression using Least absolute shrinkage and selection operator are shown. (C) Kaplan-Meier overall survival curves of patients that were divided into 2 groups based on the median risk score. (D) Gene set enrichment analysis results of patients in the high-risk group versus low-risk group. (E) Kaplan-Meier overall survival curves of patients that received chemoradiotherapy or hormone therapy.

blood RNA is closely associated with breast cancer stage [34]. The regulators promote the invasiveness of breast cancer, and therefore are correlated with an advanced tumor stage. NPI is calculated based on 3 pathological criteria: tumor size, the number of involved lymph nodes, and tumor grade [20]. The m<sup>6</sup>A regulators are involved in lymph node metastasis [32] and are therefore associated with NPI. However, low expression of some m<sup>6</sup>A regulators, such as METTL14 and ZC3H13, is associated with tumor progression, including NPI [35]. This could be explained by the potential dual function of m<sup>6</sup>A regulators in breast cancer. We also searched data on the relationship between m<sup>6</sup>A and cellularity in breast cancer but found little evidence about that relationship. This indicates that our findings are novel and might provide some clues for future studies.

Chemotherapy, radiation, and hormone therapy have been shown to greatly improve the prognosis of breast cancer. In

addition to predicting OS, the risk score also performs well when predicting the response of patients with breast cancer after receiving the above treatment modalities. Therefore, patients with a high risk score should receive a higher dose of adjuvant therapies and close surveillance after treatment. Preclinical studies have indicated that the activation of PI3K/AKT/mTOR contributes to an acquired resistance to hormone therapy [36,37]. Data from other randomized trials also demonstrated that mTOR inhibition could improve hormone therapy resistance [38,39]. It has been reported that MYC is frequently increased in chemotherapy-resistant breast cancer patients and that MYC is able to maintain breast cancer stem cells in chemotherapy-resistant patients [40]. High-risk breast cancer patients are harboring the enrichment of the above pathways, which probably leads to radiation, chemotherapy, and hormone therapy resistance. These molecules and hallmarks have been targeted to develop new drugs and overcome resistance to

**Table 1.** Cox regression analysis of risk scores and clinicopathological factors for patients' overall survival.

Characteristics	Univariate			Multivariate		
	HR (95%CI)	P	HR (95%CI)	P		
Age (≥62 vs <62 years)	2.01 (1.77-2.27)	<0.001	2.02 (1.73-2.37)	<0.001		
Risk score (high vs low)	1.48 (1.32-1.67)	<0.001	1.23 (1.08-1.39)	0.002		
Hormone therapy (yes vs no)	1.28 (1.13-1.44)	<0.001	0.93 (0.79-1.09)	0.372		
Chemotherapy (yes vs no)	1.23 (1.06-1.43)	0.007	1.14 (0.91-1.42)	0.261		
Radiotherapy (yes vs no)	0.84 (0.74-0.94)	0.003	0.78 (0.68-0.89)	<0.001		
Histologic grade						
1	1		1			
2	1.28 (1-1.64)	0.046	1.07 (0.69-1.67)	0.764		
3	1.63 (1.28-2.07)	<0.001	0.96 (0.56-1.64)	0.878		
ER Status (positive vs negative)	0.85 (0.74-0.98)	0.022	0.91 (0.74-1.11)	0.347		
PR Status (positive vs negative)	0.79 (0.7-0.89)	<0.001	0.9 (0.78-1.04)	0.169		
HER2 Status (positive vs negative)	1.45 (1.22-1.73)	<0.001	1.36 (1.13-1.65)	0.001		
Menopausal State (pre vs post)	0.59 (0.5-0.7)	<0.001	0.87 (0.71-1.07)	0.192		
Tumor Laterality (right vs left)	0.93 (0.82-1.05)	0.252				
NPI						
≤2.4	1		1			
(2.4-3.4]	0.97 (0.75-1.24)	0.787	0.87 (0.52-1.46)	0.609		
(3.4-5.4]	1.39 (1.1-1.74)	0.005	1.14 (0.6-2.16)	0.688		
≥5.4	3.07 (2.35-4.01)	<0.001	1.51 (0.65-3.51)	0.337		
No. of positive LN						
0	1		1			
1-3	1.36 (1.19-1.56)	<0.001	1.21 (0.98-1.5)	0.076		
4-9	2.14 (1.78-2.58)	<0.001	1.51 (1.03-2.2)	0.034		
≥10	4.27 (3.38-5.38)	<0.001	2.48 (1.62-3.79)	<0.001		
Tumor Size (>23 vs ≤23)	1.86 (1.65-2.1)	<0.001	1.53 (1.35-1.75)	<0.001		

CI – confidence interval; ER – estrogen receptor; HER2 – human epidermal growth factor receptor 2; HR – hazard ratio; LN – lymph node; NPI – Nottingham prognostic index; PR – progesterone receptor.

adjuvant therapies. Further studies targeting m<sup>6</sup>A methylation regulators are warranted.

Solid tumor tissues consist of cancerous and noncancerous components, which contain epithelial, stromal, and endothelial cells [41]. Tumor purity, characterized as the proportion of cancerous cells in a solid tumor sample, is a key feature in cancer transcriptomics and metabolomics data analysis. Noncancerous components are viewed as prevailing pollutants in commixture and constitute a massive part of tumor masses. Studies have found that m<sup>6</sup>A modulators are closely related to each other, and our results from the STRING and TIMER databases also invalidated the findings. However, the correlation result of RNA data from the TCGA database without tumor purity adjusted was not consistent with the results of STRING or TIMER, an inconsistency

that could have resulted from tumor purity. Typically, harvested surgical samples could be of a purity of lower than 70%, giving rise to systematic biases in genomic analysis [42]. Tumor purity is an important confounder and therefore should be considered to increase data accuracy and reliability.

### Conclusions

In conclusion, our data demonstrated that m<sup>6</sup>A regulators were related to clinicopathological characteristics and showed prognostic significance in breast cancer. Future studies to evaluate the role of m<sup>6</sup>A regulators in other types of cancer are needed. The development of targeted therapeutics based on m<sup>6</sup>A regulators is also promising.

## Supplementary Data

**Supplementary Table 1.** Expression of N6-methyladenosine regulators in different cellularity, tumor stage, or Nottingham prognostic index groups.

m <sup>6</sup> A regulators	Cellularity			NPI Class				Tumor stage		
	High	Moderate	Low	≤2.4	2.4-3.4	3.4-5.4	≥5.5	1	2	3
METTL3 (mean (SD))	8.30 (0.40)	8.27 (0.39)	8.34 (0.39)	8.32 (0.36)	8.34 (0.38)	8.28 (0.41)	8.25 (0.40)	8.34 (0.39)	8.27 (0.40)	8.28 (0.40)
METTL14 (mean (SD))	6.08 (0.24)	6.06 (0.23)	6.03 (0.23)	6.07 (0.23)	6.07 (0.21)	6.06 (0.24)	6.02 (0.26)	6.05 (0.22)	6.06 (0.24)	6.05 (0.26)
METTL16 (mean (SD))	6.63 (0.33)	6.68 (0.33)	6.66 (0.32)	6.72 (0.32)	6.68 (0.31)	6.63 (0.34)	6.58 (0.31)	6.64 (0.31)	6.64 (0.33)	6.56 (0.29)
RBM15 (mean (SD))	8.21 (0.42)	8.13 (0.43)	8.05 (0.39)	8.03 (0.43)	8.07 (0.38)	8.21 (0.43)	8.24 (0.44)	8.11 (0.39)	8.15 (0.44)	8.20 (0.44)
WTAP (mean (SD))	7.01 (0.59)	6.98 (0.56)	7.13 (0.52)	7.04 (0.59)	6.94 (0.54)	7.04 (0.58)	7.06 (0.60)	7.07 (0.59)	7.02 (0.59)	7.12 (0.61)
KIAA1429 (mean (SD))	8.07 (0.55)	8.03 (0.55)	7.88 (0.44)	7.95 (0.56)	8.02 (0.56)	8.05 (0.53)	8.01 (0.52)	8.00 (0.55)	8.01 (0.52)	8.01 (0.50)
RBM15B (mean (SD))	6.55 (0.40)	6.50 (0.36)	6.49 (0.35)	6.58 (0.33)	6.62 (0.36)	6.49 (0.38)	6.44 (0.39)	6.56 (0.36)	6.49 (0.39)	6.45 (0.37)
ALKBH5 (mean (SD))	11.23 (0.38)	11.22 (0.38)	11.27 (0.33)	11.30 (0.35)	11.26 (0.36)	11.21 (0.36)	11.21 (0.47)	11.24 (0.36)	11.22 (0.38)	11.17 (0.41)
EIF3 (mean (SD))	8.88 (0.79)	8.79 (0.77)	8.59 (0.84)	8.47 (0.79)	8.75 (0.78)	8.84 (0.79)	9.04 (0.82)	8.62 (0.80)	8.84 (0.81)	8.91 (0.83)
YTHDF1 (mean (SD))	9.35 (0.45)	9.29 (0.41)	9.25 (0.39)	9.20 (0.39)	9.28 (0.42)	9.32 (0.44)	9.41 (0.40)	9.23 (0.42)	9.31 (0.44)	9.34 (0.41)
ZC3H13 (mean (SD))	5.23 (0.13)	5.24 (0.13)	5.23 (0.12)	5.24 (0.13)	5.24 (0.13)	5.23 (0.13)	5.24 (0.13)	5.24 (0.13)	5.23 (0.14)	5.22 (0.13)
YTHDC1 (mean (SD))	6.08 (0.32)	6.04 (0.29)	5.96 (0.26)	6.06 (0.34)	6.05 (0.30)	6.04 (0.30)	6.05 (0.30)	6.01 (0.30)	6.07 (0.30)	6.01 (0.27)
YTHDF3 (mean (SD))	7.29 (0.40)	7.19 (0.37)	7.15 (0.34)	7.12 (0.40)	7.17 (0.35)	7.27 (0.39)	7.28 (0.42)	7.22 (0.38)	7.24 (0.39)	7.30 (0.36)
HNRNPA2B1 (mean (SD))	8.45 (0.62)	8.47 (0.62)	8.46 (0.55)	8.52 (0.52)	8.50 (0.58)	8.45 (0.64)	8.43 (0.58)	8.50 (0.57)	8.47 (0.60)	8.45 (0.58)
YTHDC2 (mean (SD))	6.23 (0.31)	6.22 (0.29)	6.22 (0.24)	6.27 (0.28)	6.30 (0.26)	6.20 (0.30)	6.14 (0.29)	6.21 (0.29)	6.22 (0.30)	6.14 (0.31)
YTHDF2 (mean (SD))	8.77 (0.40)	8.70 (0.38)	8.75 (0.36)	8.75 (0.42)	8.72 (0.37)	8.75 (0.39)	8.76 (0.38)	8.76 (0.36)	8.75 (0.40)	8.74 (0.34)
CBLL1 (mean (SD))	7.56 (0.36)	7.50 (0.33)	7.53 (0.29)	7.49 (0.29)	7.48 (0.32)	7.55 (0.35)	7.59 (0.35)	7.55 (0.32)	7.52 (0.35)	7.63 (0.34)
IGF2BP2 (mean (SD))	6.89 (0.84)	6.84 (0.70)	7.02 (0.72)	6.72 (0.53)	6.73 (0.53)	6.97 (0.86)	6.93 (0.84)	6.81 (0.68)	6.89 (0.78)	6.94 (0.86)
IGF2BP1 (mean (SD))	5.60 (0.21)	5.59 (0.20)	5.58 (0.20)	5.57 (0.20)	5.59 (0.21)	5.59 (0.20)	5.61 (0.22)	5.57 (0.19)	5.61 (0.21)	5.62 (0.22)
IGF2BP3 (mean (SD))	5.81 (0.74)	5.71 (0.52)	5.80 (0.64)	5.58 (0.24)	5.60 (0.36)	5.86 (0.76)	5.86 (0.64)	5.73 (0.51)	5.78 (0.68)	5.91 (0.78)

m<sup>6</sup>A – N6-methyladenosine; NPI – Nottingham prognostic index; SD – standard deviation.

References:

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *Cancer J Clin.* 2020;70(1):7-30
- Woolston C. Breast cancer: 4 big questions. *Nature.* 2015;527(7578):S120
- Zhao Y, Shi Y, Shen H, Xie W. m(6)A-binding proteins: the emerging crucial performers in epigenetics. *J Hematol Oncol.* 2020;13(1):35
- Nandy D, Rajam SM, Dutta D. A three layered histone epigenetics in breast cancer metastasis. *Cell Biosci.* 2020;10:52
- Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci USA.* 1974;71(10):3971-75
- Agarwala SD, Blitzblau HG, Hochwagen A, Fink GR. RNA methylation by the MIS complex regulates a cell fate decision in yeast. *PLoS Genet.* 2012;8(6):e1002732
- Wang X, Feng J, Xue Y, et al. Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. *Nature.* 2016;534(7608):575-78
- Zheng G, Dahl JA, Niu Y, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell.* 2013;49(1):18-29
- Jia G, Fu Y, Zhao X, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat Chem Biol.* 2011;7(12):885-87
- Yang Y, Hsu PJ, Chen YS, Yang YG. Dynamic transcriptomic m(6)A decoration: Writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 2018;28(6):616-24
- Li Y, Xiao J, Bai J, et al. Molecular characterization and clinical relevance of m(6)A regulators across 33 cancer types. *Mol Cancer.* 2019;18(1):137
- Lan Q, Liu PY, Haase J, et al. The critical role of RNA m(6)A methylation in cancer. *Cancer Res.* 2019;79(7):1285-92
- Wang H, Xu B, Shi J. N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. *Gene.* 2020;722:144076
- Liu L, Liu X, Dong Z, et al. N6-methyladenosine-related genomic targets are altered in breast cancer tissue and associated with poor survival. *J Cancer.* 2019;10(22):5447-59
- Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47(D1):D607-13
- Hsiao KY, Lin YC, Gupta SK, et al. Noncoding effects of circular RNA CCDC66 promote colon cancer growth and metastasis. *Cancer Res.* 2017;77(9):2339-50
- Subramanian A, Kuehn H, Gould J, et al. GSEA-P: A desktop application for Gene Set Enrichment Analysis. *Bioinformatics.* 2007;23(23):3251-53
- Chai RC, Wu F, Wang QX, et al. m(6)A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging (Albany NY).* 2019;11(4):1204-25
- Zhang S, Wang Y, Gu Y, et al. Specific breast cancer prognosis-subtype distinctions based on DNA methylation patterns. *Mol Oncol.* 2018;12(7):1047-60
- Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. *Br J Cancer.* 1982;45(3):361-66
- Niu Y, Lin Z, Wan A, et al. RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. *Mol Cancer.* 2019;18(1):46
- Klinge CM, Piell KM, Tooley CS, Rouchka EC. HNRNPA2/B1 is upregulated in endocrine-resistant LCC9 breast cancer cells and alters the miRNA transcriptome when overexpressed in MCF-7 cells. *Sci Rep.* 2019;9(1):9430
- Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta.* 2010;1805(1):105-17
- Chen Y, Peng C, Chen J, et al. WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. *Mol Cancer.* 2019;18(1):127
- Chen L, Wang X. Relationship between the genetic expression of WTAP and bladder cancer and patient prognosis. *Oncol Lett.* 2018;16(6):6966-70
- Yu HL, Ma XD, Tong JF, et al. WTAP is a prognostic marker of high-grade serous ovarian cancer and regulates the progression of ovarian cancer cells. *Oncotargets Ther.* 2019;12:6191-201
- Hui L, Zhang S, Wudu M, et al. CBLL1 is highly expressed in non-small cell lung cancer and promotes cell proliferation and invasion. *Thorac Cancer.* 2019;10(6):1479-88
- Meyer KD, Patil DP, Zhou J, et al. 5' UTR m(6)A promotes cap-independent translation. *Cell.* 2015;163(4):999-1010
- Kar A, Gutierrez-Hartmann A. ESE-1/ELF3 mRNA expression associates with poor survival outcomes in HER2(+) breast cancer patients and is critical for tumorigenesis in HER2(+) breast cancer cells. *Oncotarget.* 2017;8(41):69622-40
- Zhao ZM, Yost SE, Hutchinson KE, et al. CCNE1 amplification is associated with poor prognosis in patients with triple negative breast cancer. *BMC Cancer.* 2019;19(1):96
- Chen XY, Zhang J, Zhu JS. The role of m(6)A RNA methylation in human cancer. *Mol Cancer.* 2019;18(1):103
- Xu Y, Ye S, Zhang N, et al. The FTO/miR-181b-3p/ARL5B signaling pathway regulates cell migration and invasion in breast cancer. *Cancer Commun (Lond).* 2020;40(10):484-500
- Zhang B, Gu Y, Jiang G. Expression and prognostic characteristics of m(6) A RNA methylation regulators in breast cancer. *Front Genet.* 2020;11:604597
- Xiao H, Fan X, Zhang R, Wu G. Upregulated N6-methyladenosine RNA in peripheral blood: potential diagnostic biomarker for breast cancer. *Cancer Res Treat.* 2021;53(2):399-408
- Gong PJ, Shao YC, Yang Y, et al. Analysis of N6-methyladenosine methyltransferase reveals METTL14 and ZC3H13 as tumor suppressor genes in breast cancer. *Front Oncol.* 2020;10:578963
- Miller TW, Hennessy BT, Gonzalez-Angulo AM, et al. Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *J Clin Invest.* 2010;120(7):2406-13
- Sanchez CG, Ma CX, Crowder RJ, et al. Preclinical modeling of combined phosphatidylinositol-3-kinase inhibition with endocrine therapy for estrogen receptor-positive breast cancer. *Breast Cancer Res.* 2011;13(2):R21
- Bachelot T, Bourcier C, Cropet C, et al. Randomized phase II trial of everolimus in combination with tamoxifen in patients with hormone receptor-positive, human epidermal growth factor receptor 2-negative metastatic breast cancer with prior exposure to aromatase inhibitors: a GINECO study. *J Clin Oncol.* 2012;30(22):2718-24
- Baselga J, Campone M, Piccart M, et al. Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer. *N Engl J Med.* 2012;366(6):520-29
- Lee KM, Giltneane JM, Balko JM, et al. MYC and MCL1 cooperatively promote chemotherapy-resistant breast cancer stem cells via regulation of mitochondrial oxidative phosphorylation. *Cell Metab.* 2017;26(4):633-47.e7
- Rhee JK, Jung YC, Kim KR, et al. Impact of tumor purity on immune gene expression and clustering analyses across multiple cancer types. *Cancer Immunol Res.* 2018;6(1):87-97
- Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun.* 2015;6:8971