Prognostic Value of A Risk Score Model Constructed By m6A Methylation Regulators In Breast Cancer

Xiaowei Qiu
Beijing University of Chinese Medicine Third Affiliated Hospital
Qiaoli Zhang
Beijing University of Chinese Medicine Third Affiliated Hospital
Jingnan Xu
Beijing University of Chinese Medicine Third Affiliated Hospital
Xin Jiang
Beijing University of Chinese Medicine Third Affiliated Hospital
Xuewei Qi
Beijing University of Chinese Medicine Third Affiliated Hospital
Xin Chang
Beijing University of Chinese Medicine Third Affiliated Hospital
Yong Lei
Beijing Hospital of traditional Chinese Medicine
Jinchang Huang
Beijing University of Chinese Medicine Third Affiliated Hospital
Yehong Tian (tyehong@163.com)
Shaanxi University of Chinese Medicine https://orcid.org/0000-0001-8821-4123

Research

Keywords: Breast cancer, m6A, TCGA, Prognosis.

Posted Date: September 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-832380/v1

Abstract

Background: N6-methyladenosine (m6A) methylation modification can affect the tumorigenesis, progression, and metastasis of breast cancer (BC). Up to now, a prognostic model based on m6A methylation regulators for BC is still lacking. This study aimed to construct an accurate prediction prognosis model by m6A methylation regulators for BC patients.

Methods: After processing of The Cancer Genome Atlas (TCGA) datasets, the differential expression and correlation analysis of m6A RNA methylation regulators were applied. Next, tumor samples were clustered into different groups and clinicopathologic features in different clusters were explored. By univariate Cox and Least Absolute Shrinkage and Selection Operator (LASSO) analysis, m6A regulators with prognostic value were identified to develop a prediction model. Furthermore, we constructed and validated a predictive nomogram to predict the prognosis of BC patients.

Results: 19 m6A related genes were extracted and 908 BC patients enrolled from TCGA dataset. After univariate Cox and LASSO analysis, 3 m6A RNA methylation regulators (YTHDF3, ZC3H13 and HNRNPC) were selected to establish the prognosis model based on median risk score (RS) in training and validation cohort. With the increasing of RS, the expression levels of YTHDF3 and ZC3H13 were individually elevated, while the HNRNPC expressed decreasingly. By survival analysis and Receiver Operating Characteristic (ROC) curve, we found that the overall survival (OS) of high-risk group was significantly shorter than that of the low-risk group based on Kaplan-Meier (KM) analysis in each cohort. Univariate and multivariate analysis identified the RS, age, and pathological stage are independent prognostic factors. A nomogram was constructed to predict 1- and 3-year OS and the calibration plots validate the performance. The C-index of nomogram reached 0.757 (95% CI:0.7-0.814) in training cohort and 0.749 (95% CI:0.648-0.85) in validation cohort, respectively.

Conclusions: We successfully constructed a predictive prognosis model by m6A RNA methylation regulators. These results indicated that the m6A RNA methylation regulators are potential therapeutic targets of BC patients.

Background

Breast cancer (BC) has transcended lung cancer to become the most commonly diagnosed cancer and the fifth leading cause of cancer-related death worldwide [1]. Although regular mammographic screening can substantially reduce BC mortality [2], the incidence of BC has been steadily increasing [3]. At present, cancer immunotherapy, especially immune checkpoint inhibitors including programmed death 1 (PD-1), programmed cell death 1 ligand 1 (PD-L1) and cytotoxic T-lymphocyte-associated antigen (CTLA-4) [4], has been first approved to use for treating triple-negative BC in 2019. However, single monoclonal antibodies against PD-1 and PD-L1 have shown little efficacy in patients with metastatic BC [5] and indicated a poor clinical prognosis. Hence, accurate prediction of prognosis has great value in guiding individualized treatment of BC.

The changes of epigenetic has immensely facilitate the development of human cancer. Reversible chemical modifications of RNA are identified as a novel approach of epigenetic regulation [6]. m6A, which was first reported by groundbreaking research of geneticists in the 1970s [7, 8], is the most abundant internal chemical modification of message RNA (mRNA) and long non-coding RNA (lncRNA) in eukaryotes. It plays critical roles in a variety of biological functions, which included the regulation of RNA stability, splicing and translation [9, 10]. The function of m6A RNA methylation modification is determined by regulators included methyltransferase, de-methyltransferase, and m6A recognized RNA binding protein, which named “writers”, “erasers”, and “readers”, respectively [11, 12]. The m6A writers are primarily consist of METTL3, METTL14, WTAP and other subtypes, which can create m6A marks [13, 14]. ALKBH5 and FTO proteins are considered as m6A erasers, which can remove methyltransferase from RNA and maintain dynamic balance of m6A modification [15]. The mainly discovered m6A readers, can decode m6A modification to dominate the fate of modified transcripts, include the YTH domain protein (YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3) and the nuclear heterogeneous protein HNRNPA2B1 and HNRNPC. A newly recognized protein family of m6A readers, including IGF2BP1, IGF2BP2, and IGF2BP3, can promote the stability and translation function of most target RNAs [16].

BC is thought to be driven by epigenetic regulation of m6A related genes. Aberrant expression of m6A methylation regulators may expedite the progression of BC patients, including complicated biological processes such as proliferation, growth, invasion, and metastasis. Cai et al. found that HBXIP can interact with METTL3 to form a positive feedback loop of HBXIP/LET-7G/METTL3/HBXIP, which accelerated the proliferation of BC cells [17]. Through bioinformatics data mining, Gong et al. found METTL14 and ZC3H13 were negatively correlated with the OS of BC patients, which may play a vital role in the tumor inhibition [18]. An animal study shown that repression of HNRNPC in MCF7 and T47D cells can suppress tumor cell proliferation and growth [19]. Previous study indicates that m6A RNA methylation modification is an important epigenetic regulation of BC and may become a new target for the therapy of BC patients [20]. Therefore, making use of m6A methylation regulators as prognostic markers to establish a predictive model have important clinical values for the study of patients with BC.

Our current study aimed to construct a prognostic model for BC patients by m6A related genes. We extracted the expressions of m6A RNA methylation from TCGA dataset and analyzed the molecular features of m6A regulators. By univariate Cox analysis and LASSO analysis, a series of m6A regulators with prognostic value were identified to develop a prediction model to investigated the relationship between the prognostic model with OS in combination with clinicopathological characteristics. Furthermore, we constructed and validated a predictive nomogram which shown a favorable performance to the prognosis of patients with BC.

Methods

Data acquisition and the selection of m6A methylation regulator gene

We downloaded transcriptome RNA-sequencing data of human BC samples and related clinical data from TCGA data portal (https://porta l.gdc.cance r.gov/). By using Perl language (http://www.perl.org/), RNA-seq data of BC patients were converted into a matrix file and the expression of 19 m6A RNA methylation regulators were extracted from the TCGA cohort according to published literature [38]. These m6A regulators included 7 writers (KIAA1429, METTL3, METTL14, METTL16, RBM15, WTAP, ZC3H13), 2 erasers (ALKBH5, FTO) and 10 readers (HNRNPA2B1, HNRNPC, IGF2BP1, IGF2BP2, IGF2BP3, YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3). The clinical data of inclusion and exclusion criteria were as follows. Inclusion criteria: 1. Data of BC samples including males and females; 2. The gene expression of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) was normalized in the transcriptome. Exclusion criteria: 1. Survival time is less than or equal to 30 days; 2. Clinical information is not clear; 3. Annotation information is not matched with the clinical sample information. 4. Duplicate samples. Because the data were obtained from public databases of TCGA, no ethical approval was required.

The differential expression and correlation analysis of m6A RNA methylation regulators

To compare the differential expressions of m6A RNA methylation-related genes between the normal and tumor samples of TCGA BC cohort, heatmap and violin plot were applied through “pheatmap” and “vioplot” R package. Then, the Pearson correlation analysis was used to assess the relationship of m6A RNA methylation regulators expression between normal group and tumor group by “corrplot” R package.

Consensus clustering analysis of m6A RNA methylation regulators

In order to investigate the expression characteristics of m6A methylation regulators in BC tumor group, we deleted the normal samples and clustered the tumor samples into different groups using the “ConsensusClusterPlus” R package. After that, the relationship between expression of m6A related genes and clinicopathologic features in different clusters was visualized by “pheatmap” R package.

Construction of prognostic network and univariate Cox prognostic analysis

To explore the prognosis correlation with m6A RNA methylation modification in TCGA BC cohort, a prognostic network was performed based on the 4 R packages included “igraph”, “psych”, “reshape2” and “RColorBrewer”. Next, univariate Cox regression analysis was used to identify the prognosis of m6A RNA methylation regulators by “survival” R package and genes with $p < 0.1$ were considered as prognostic factors.

Biological function analysis of m6A RNA methylation regulators

To evaluate the potential biological functions of m6A RNA methylation regulators in BC, we performed Gene Ontology (GO) enrichment analysis based on the “clusterprofiler” R package. A p-value < 0.05 was set as the significant screening criterion, and “Goplot” was used to visualize the results.

Construction and validation of the m6A-related RS prognostic model

Based on univariate Cox regression analysis screened prognostic factors, LASSO analysis was utilized to construct a m6A related RS prognostic model. We carried out a stepwise Cox proportional hazards regression model to make our model more optimized and practical by “survival” and “glmnet” R packages. At last, the expression of m6A RNA regulators with prognostic value were estimated by multivariate Cox regression coefficients. RS formula calculated as follow:

$$RS= \sum_{i-1}^{N} Expi * Coei.$$

Through “caret” R package, all BC patients with complete survival information were randomly divided into the training cohort and validation cohort at a ratio of 7:3. According to the median RS, which was considered as a cutoff value, the BC patients were classified into the high- or low-risk groups in each cohort. By “time ROC” R package, we plotted the ROC curves to assess the sensitivity and specificity of prognosis. An area under the curve (AUC) > 0.60 was served as acceptable. KM plot was performed to visualize the OS probability of BC patients between the high-risk and low-risk group.

Identification of Independent Prognostic Factors

We conducted univariate and multivariate Cox analysis to identify the independent risk factors. The complete survival information of BC patients in entire TCGA cohort included relevant clinical data were displayed to define the independent prognostic factors by “survival” R package. $p < 0.05$ was considered as statistically significant.

Construction and validation of a Predictive Nomogram

We performed a predictive nomogram by “rms” R package consisting of independent prognostic factors and related clinical parameters as variables, which is based on multivariate Cox regression analysis in entire TCGA cohort. By drawing a vertical line between the total point coordinate axis and each prognostic coordinate axis, we calculated OS rates at 1 and 3 years of BC patients. We applied the concordance index (C-index) by “survcomp” R package and calibration plot by “rms” R package to validate the performance of the nomogram in each cohort.

Statistical Analysis

The statistical analysis of all data was performed by R version 4.0.2 and corresponding packages. Wilcox test was employed to compare m6A RNA expression visualized by heatmap and violin plot. Correlation analysis was performed by Pearson correlation test. The chi-squared test was utilized to assess the significance of the correlation of m6A RNA expression with clinical parameters in different BC clusters. OS analysis was used KM plotter in which survival differences were evaluated by a two-sided log-rank test. Cox proportional hazard regression model was performed to determine whether RS can be used as an independent risk factor. The significance level was selected as $p < 0.05$. In addition, univariate Cox

regression was used to analyze the prognostic value of m6A RNA regulators and the significance level selected as $p < 0.1$. Statistical significance was showed in figures as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Results

m6A RNA methylation regulators expression and clinical data in TCGA BC cohort

FPKM normalized gene expressions used as the RNA-seq results of 1109 BC samples and 113 normal samples were obtained from TCGA database. To explore the expression pattern of m6A RNA methylation regulators in BC, the sequencing data of 19 m6A related genes, shown in table 1, were extracted from TCGA BC cohort. Besides, there were 908 BC patients with complete survival information enrolled from TCGA dataset. The baseline characteristics of these BC patients was shown in table 2.

Differential expression and correlation of m6A RNA methylation regulators

In total, 19 m6A related genes were extracted from TCGA BC cohort to explore the expression pattern of m6A RNA methylation regulators. There were 11 upregulated genes (ALKBH5, HNRNPA2B1, YTHDC2, YTHDF1, HNRNPC, IGF2BP1, KIAA1429, RBM15, YTHDF2, YTHDF3 and IGF2BP3) and 8 downregulated genes (FTO, IGF2BP2, ZC3H13, METTL3, METTL16, METTL14, WTAP and YTHDC1) in tumor samples compared with normal samples of BC patients shown in Figure 1A and 1B. Moreover, the correlation analysis indicated that most parts of the different m6A RNA methylation regulators exhibited weakly to moderately positive correlation and some of them represent the weakly negative correlation. Among them, METTL14 had the strongest positive correlation with YTHDC1 and had the most negative correlation with IGF2BP2 combined with YTHDC2 (Figure 1C).

Consensus clustering of m6A RNA methylation regulators identified two clusters of BC patients with clinicopathological features

We clustered the tumor samples into different groups to investigate the expression similarity of m6A RNA methylation regulators based on the clinicopathological features of BC. According to consensus cluster analysis, k = 2 seemed to be a most optimal choice with clustering stability increasing from k = 2 to 9 in the TCGA BC cohort as shown in Figure 2A, B, C and Supplementary Figure 1. The tumor samples were divided into 2 different groups which were named cluster 1 and cluster 2. To explore the association between the two clusters and clinical outcomes, the heatmap was performed for further analysis. The relationship of clinicopathological characteristics between the two clusters demonstrated significant differences for the age, pathological stage, and T stage ($p < 0.05$). These results indicated that the two clusters were closely related to clinical outcomes and malignancy of BC patients.

Prognostic Network of m6A RNA methylation regulators

A prognostic network was constructed to depict the comprehensive landscape of 19 m6A regulators interaction, regulator connection and their prognostic significance for BC patients (Figure 3A). The results

indicated that the expression of m6A regulators was significantly correlated not only in the same functional category, but also among writers, ereasers and readers. IGF2BP2 was negatively correlated with YTHDF1, YTHDF2, METTL16, METTL14, and FTO with HNRNPA2B1 and HNRNPC, respectively. Other m6A regulators were positively correlated, among which METTL14 had the strongest positive correlation with YTHDF1 and YTHDF2. Furthermore, we conducted a univariate Cox regression in patients with BC. The results viewed that 5 m6A regulators (YTHDF3, ZC3H13, METTL16, HNRNPC and KIAA1429) can be selected as prognostic factors with $p < 0.1$ (Figure 3B).

GO enrichment analysis of m6A RNA methylation regulators

GO enrichment analysis indicated that m6A RNA methylation regulators were enriched in multiple BP (biological process), MF (molecular function), and CC (cellular component) terms. The most enriched top terms of BP, MF and CC are regulation of mRNA metabolic process, catalytic activity (acting on RNA) and nuclear speck, respectively (supplementary table 1). As the top 10 enriched GO terms shown in Figure 4A, 6 terms (included regulation of mRNA metabolic process, regulation of mRNA stability, regulation of RNA stability, regulation of mRNA catabolic process, mRNA catabolic process, and RNA catabolic process) were upregulated and 4 terms (included negative regulation of translation, negative regulation of translation, negative regulation of cellular amide metabolic process, and RNA destabilization) were downregulated. The description of the top 10 enriched GO terms was shown in Figure 4B.

Construction and validation of the m6A related survival prediction model

After performed univariate Cox regression analysis, LASSO analysis was used to further screen the prognosis-related m6A regulators. Totally, 3 m6A RNA methylation regulators (YTHDF3, ZC3H13 and HNRNPC) were selected to establish the prognosis RS model (Figure 5A and 5B). The heatmap of risk distribution among the 908 BC patients with distinct clinicopathologic features in entire TCGA cohort was shown in Figure 5C. Then, 908 BC patients were randomly separated to the training cohort (n=636) and validation cohort (n=272). According to the calculated median RS, BC patients were divided into the high-risk group and the low-risk group in training cohort and validation cohort (Figure 6A1 and 6A2). The mortality rate was constantly increasing with the enhanced RS (Figure 6B1 and 6B2). And with the increase of RS, the expression levels of YTHDF3 and ZC3H13 were individually elevated, while the HNRNPC expressed decreasingly (Figure 6C1 and 6C2). Subsequently, we performed OS and ROC to evaluate and validate the prognostic value of the RS model. We found that the OS of the high-risk group was significantly shorter than that of the low-risk group based on KM analysis ($p < 0.001$, Figure 6D1 and 6D2). The AUC values of time-dependent ROC curves were 0.61 and 0.624 in the training cohort and validation cohort, respectively (Figure 6E1 and 6E2).

Independent Prognostic Factors Identification and the Prognostic Nomogram Construction and validation

The results of univariate and multivariate Cox regression analysis showed in table 3 indicated that age, pathological stage, and RS were significantly correlated with OS and prognostic models constructed with the m6A related genes may be better predictors of OS in BC. As a result, we constructed a nomogram to

predict OS based on the RS combined age and pathological stage in entire TCGA BC cohort (Figure 7A). The calibration plots showed that the performance of the nomogram was the best in predicting the 1- and 3-year OS in the training cohort (Figure 7B and 7C) and validation cohort (Figure 7D and 7E). The C-index of our nomogram reached 0.757 (95% CI:0.7-0.814) in training cohort and 0.749 (95% CI:0.648-0.85) in validation cohort, respectively. These results demonstrated that our nomogram can be applied to predict the OS rate of BC patients based on their own conditions to improve the prediction efficiency and accuracy.

Discussion

RNA methylation modifications account for more than 60% of all RNA modifications, and m6A is the most well studied post-transcriptional modification on mRNAs and lncRNAs, which play a crucial biological function in mammals [21]. Accumulating studies have shown that abnormal regulation of m6A regulators may closely associated with the occurrence of human cancers. Upregulation of METTL14 can inhibit the metastasis of hepatocellular carcinoma by regulating the m6A modification of primary microRNA 126 and affecting the generation and processing of microRNA-126 in an m6A-dependent manner [22]. Aberrant expression of METTL3 can promote the growth, survival, and invasion of lung adenocarcinoma cells [23]. Mutations or copy number variations in m6A regulators are closely related to TP53 mutations, which predict poor prognosis of acute myelocytic leukemia [24]. In glioblastoma stem-like cells, ALKBH5 maintained the tumorigenicity of glioma cells through lncRNA FOXM1 mediated m6A modification on pre-mRNA [25]. In addition, knockout of METTL3 or METTL14 can alter the enrichment of m6A and the expression of ADAM19, which vastly promotes the growth, self-renewal, and tumorigenesis of glioma cells [26].

A mount number of studies have confirmed that the dysregulation of m6A regulators is responsible for tumorigenesis, molecular typing, and progression in BC patients. METTL14 and KIAA1429 can greatly promote the migration and invasion of BC [27, 28]. METTL3 can induce proliferation, inhibit apoptosis, and accelerate tumor growth by targeting Bcl-2 [29]. Hypoxic stimulation can promote the expression of ALKBH5 dependent on Hypoxia-inducible factors. Overexpression of ALKBH5 can reduce the m6A modification of NANOG mRNA, thus stabilizing the mRNA and improving the expression level of NANOG, and ultimately increasing the proportion of BC stem cells [30]. Notably, m6A RNA modification was associated with clinicopathological features and prognosis of BC patients. Recent research reported by Wu et al. found that patients with overexpression of METTL3, METTL14, WTAP, and FTO had better metastase-relapsed-free survival [31]. What's more, upregulation of FTO was significantly associated with shorter OS in advanced BC patients [32]. In early-stage of BC, overexpression of IGF2BP2 is closely related with shorter survival [33].

In the present study, we comprehensively compared the difference expression of 19 m6A regulators between BC tumor samples and adjacent normal samples. The results shown that there were 11 upregulated genes and 8 downregulated genes between 19 m6A regulators. Consistent with previous research, HNRNPC and YTHDF3 were upregulated and ZC3H13 was downregulated in our current study

[18, 19, 34]. The results of Pearson correlation analysis suggested that 19 m6A regulators are intrinsically correlated. To investigate the biological features of these m6A regulators in tumor samples, the genes were divided into two clusters based on the expression of m6A. The results found that these two clusters were significantly correlated with the clinicopathology of age, pathological stage, and T stage in BC patients. In order to assess the potential molecular mechanism and biological functions of m6A RNA methylation regulators, GO enrichment analysis was performed based on the differential expressions of m6A related genes. The results indicated that top enriched GO term of m6A regulators was regulation of mRNA metabolic process. Biological function of YTH family proteins is almost involved in every step of mRNA metabolism. YTHDF3 mainly involved in mRNA translation and degradation [35, 36].

To assess the prognosis value of m6A regulators, we performed prognostic network analysis of 19 m6A related genes and determined the prognostic function of m6A regulators as well as the interrelationship between themselves. The results of prognostic network analysis confirmed the prognostic value of 19 m6A RNA methylation regulators. Ultimately, YTHDF3, ZC3H13 and HNRNPC were selected to establish the RS prognostic model through univariate Cox analysis and Lasso Cox analysis. Wu et al. [19] suggested that upregulated HNRNPC can promote the proliferation of BC cells by preventing the export of Alu sequences to the cytosol. ZC3H13 was the downregulated m6A methylation regulator and it can serve as tumor suppressor genes predicted poor prognosis in BC [18]. While, YTHDF3 overexpression was associated with metastasis and predicted poor prognosis in BC. Chang et al. [37] found that YTHDF3 can induce the translation of m6A-enriched gene transcripts to promote BC brain metastasis. Therefore, YTHDF3, ZC3H13 and HNRNPC have prognostic value and can be used as effective predictors to predict the clinical prognosis of BC patients.

The accuracy and stability of the model were analyzed by OS analysis, ROC curves and risk curves, which proved that the RS prognostic model we constructed could accurately distinguish the high-risk and low-risk groups. The stratification analysis of different clinical features has shown that the RS model could effectively predict the prognosis of BC patients with age, and pathological stage. Next, we conducted univariate and multivariate Cox regression analysis to determine whether the RS and corresponding clinicopathologic features were independent prognostic factors. The results shown that the RS combined with age and pathological stage could be considered as independent prognostic factors for BC patients. Additionally, a nomogram was constructed to predict the 1-year and 3-year survival rate of BC patients, and calibration plot and C-index were applied to validate the performance of the nomogram. These results determined that our RS prognostic model is relatively reliable and can be used to identify the risk and prognosis of BC patients, which is conducive to the early intervention treatment and can better guide clinical treatment.

The significance of our current study is to validate previous studies, which suggested that m6A related regulators with vital clinical implication in predicting the prognosis of patients with BC. Nevertheless, there are still existing some limitations to our research. To date, the biological role of some m6A regulators in tumors remains controversial. The genes screened in this study are only based on published

studies, and more m6A regulators need to be investigated in depth. Moreover, the effectiveness of this RS prognostic model remains to be further verified by subsequent *in vitro* and *in vivo* experiments.

Conclusions

In summary, we obtained 3 m6A regulators with prognostic value for the construction and validation of a RS prognostic model in BC patients. These m6A RNA regulators may become newly biomarkers and potential therapeutic targets to inhibit the progress of patients with BC.

Abbreviations

AUC, area under the curve; BC, breast cancer; BP, biological process; CC, cellular component; FPKM, fragments per kilobase million; GO, Gene Ontology; LASSO, Least Absolute Shrinkage and Selection Operator; lncRNA, long non-coding RNA; m6A, N6-methyladenosine; mRNA, message RNA; MF, molecular function; OS, overall survival; PD-1, programmed death 1; PD-L1, programmed cell death 1 ligand 1; ROC, Receiver Operating Characteristic; RS, risk scores; TCGA, The Cancer Genome Atlas.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

We declare that the data supporting the findings of this study are available in the TCGA database (https ://portal.gdc.cancer.gov/).

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by Key Research Project of Beijing University of Chinese Medicine (2020-JYB-ZDGG-143-1), Beijing Municipal Natural Science Foundation (7202122) and Young Teachers Research Project of Beijing University of Chinese Medicine (303-02-01-05-122).

Author contributions

Jinchang Huang, Yong Lei, and Yehong Tian conceived and designed the study. Xiaowei Qiu and Qiaoli Zhang wrote the paper. Jingnan Xu, Xin Jiang performed data analysis. Xuewei Qi and Xin Chang downloaded and organized data, and drew the figures. All authors read and approved the manuscript.

Acknowledgements

Not applicable.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Tabar L, Dean PB, Chen TH, Yen AM, Chen SL, Fann JC, et al. The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening. Cancer. 2019;125(4):515–23.
3. Britt KL, Cuzick J, Phillips KA. Key steps for effective breast cancer prevention. Nat Rev Cancer. 2020;20(8):417–36.
4. Wei SC, Duffy CR, Allison JP. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. Cancer discovery. 2018;8(9):1069–86.
5. Esteva FJ, Hubbard-Lucey VM, Tang J, Pusztai L. Immunotherapy and targeted therapy combinations in metastatic breast cancer. The Lancet Oncology. 2019;20(3):e175-e86.
6. He C. Grand challenge commentary: RNA epigenetics? Nat Chem Biol. 2010;6(12):863–5.
7. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci U S A. 1974;71(10):3971–5.
8. Schafer KP. RNA synthesis and processing reactions in a subcellular system from mouse L cells. Hoppe Seylers Z Physiol Chem. 1982;363(1):33–43.
9. Liu Z, Zhang J. Human C-to-U Coding RNA Editing Is Largely Nonadaptive. Mol Biol Evol. 2018;35(4):963–9.
10. Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. Nat Rev Mol Cell Biol. 2014;15(5):313–26.
11. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m^6A RNA methylation. (1471-0064 (Electronic)).
12. Liu ZX, Li LM, Sun HL, Liu SM. Link Between m6A Modification and Cancers. Front Bioeng Biotechnol. 2018;6:89.
13. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA Modifications in Gene Expression Regulation. Cell. 2017;169(7):1187–200.

14. Warda AS, Kretschmer J, Hackert P, Lenz C, Urlaub H, Hobartner C, et al. Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. 2017;18(11):2004–14.
15. He L, Li H, Wu A, Peng Y, Shu G, Yin G. Functions of N6-methyladenosine and its role in cancer. Mol Cancer. 2019;18(1):176.
16. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol. 2018;20(3):285–95.
17. Cai X, Wang X, Cao C, Gao Y, Zhang S, Yang Z, et al. HBXIP-elevated methyltransferase METTL3 promotes the progression of breast cancer via inhibiting tumor suppressor let-7g. Cancer Lett. 2018;415:11–9.
18. Gong PJ, Shao YC, Yang Y, Song WJ, He X, Zeng YF, et al. Analysis of N6-Methyladenosine Methyltransferase Reveals METTL14 and ZC3H13 as Tumor Suppressor Genes in Breast Cancer. Front Oncol. 2020;10:578963.
19. Wu Y, Zhao W, Liu Y, Tan X, Li X, Zou Q, et al. Function of HNRNPC in breast cancer cells by controlling the dsRNA-induced interferon response. EMBO J. 2018;37:23.
20. Fang R, Ye L, Shi H. Understanding the roles of N(6)-methyladenosine writers, readers and erasers in breast cancer. Neoplasia. 2021;23(6):551–60.
21. Zheng F, Du F, Zhao J, Wang X, Si Y, Jin P, et al. The emerging role of RNA N6-methyladenosine methylation in breast cancer. Biomark Res. 2021;9(1):39.
22. Ma JZ, Yang F, Zhou CC, Liu F, Yuan JH, Wang F, et al. METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N(6) -methyladenosine-dependent primary MicroRNA processing. Hepatology. 2017;65(2):529–43.
23. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol Cell. 2016;62(3):335–45.
24. Kwok CT, Marshall AD, Rasko JE, Wong JJ. Genetic alterations of m(6)A regulators predict poorer survival in acute myeloid leukemia. J Hematol Oncol. 2017;10(1):39.
25. Zhang S, Zhao BS, Zhou A, Lin K, Zheng S, Lu Z, et al. m(6)A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. Cancer Cell. 2017;31(4):591–606. e6.
26. Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, et al. m(6)A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. Cell Rep. 2017;18(11):2622–34.
27. Qian J-Y, Gao J, Sun X, Cao M-D, Shi L, Xia T-S, et al. KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. Oncogene. 2019;38(33):6123–41.
28. Yi D, Wang R, Shi X, Xu L, Yilihamu Y, Sang J. METTL14 promotes the migration and invasion of breast cancer cells by modulating N6methyladenosine and hsamiR146a5p expression. Oncol Rep. 2020;43(5):1375–86.

29. Wang H, Xu B, Shi J. N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. (1879-0038 (Electronic)).
30. Zhang C, Samanta D, Lu H, Bullen JW, Zhang H, Chen I, et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m(6)A-demethylation of NANOG mRNA. Proc Natl Acad Sci U S A. 2016;113(14):E2047-56.
31. Wu L, Wu D, Ning J, Liu W, Zhang D. Changes of N6-methyladenosine modulators promote breast cancer progression. BMC Cancer. 2019;19(1):326.
32. Niu Y, Lin Z, Wan A, Chen H, Liang H, Sun L, et al. RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3. Mol Cancer. 2019;18(1):46.
33. Barghash A, Helms V, Kessler SM. Overexpression of IGF2 mRNA-Binding Protein 2 (IMP2/p62) as a Feature of Basal-like Breast Cancer Correlates with Short Survival. Scand J Immunol. 2015;82(2):142–3.
34. Anita R, Paramasivam A, Priyadharsini JV, Chitra S. The m6A readers YTHDF1 and YTHDF3 aberrations associated with metastasis and predict poor prognosis in breast cancer patients. American journal of cancer research. 2020;10(8):2546–54.
35. Li A, Chen YS, Ping XL, Yang X, Xiao W, Yang Y, et al. Cytoplasmic m(6)A reader YTHDF3 promotes mRNA translation. Cell Res. 2017;27(3):444–7.
36. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res. 2017;27(3):315–28.
37. Chang G, Shi L, Ye Y, Shi H, Zeng L, Tiwary S, et al. YTHDF3 Induces the Translation of m(6)A-Enriched Gene Transcripts to Promote Breast Cancer Brain Metastasis. Cancer Cell. 2020;38(6):857–71. e7.
38. Wang P, Wang X, Zheng L, Zhuang C. Gene Signatures and Prognostic Values of m6A Regulators in Hepatocellular Carcinoma. Front Genet. 2020;11:540186.

Tables

Table 1 The m6A regulator gene expressions in BC patient samples from TCGA.

Gene function	Gene id	Normal mean	Tumor mean	Log FC	p-value
Writer	KIAA1429	9.115622	12.964357	0.508138	1.33E-07
	METTL3	5.49215	4.57853	-0.262487	6.34E-11
	METTL14	5.81065	4.670613	-0.315088	1.62E-12
	METTL16	6.593373	6.483965	-0.02414	0.216999
	RBM15	2.669866	3.196371	0.259668	0.000336
	WTAP	16.804125	14.756599	-0.187455	8.54E-09
	ZC3H13	12.397454	9.575718	-0.372591	6.02E-17
Ereaser	ALKBH5	32.44575	33.49177	0.045777	0.463316
	FTO	11.03887	5.784148	-0.93242	1.55E-38
Reader	HNRNPA2B1	72.12021	107.0777	0.570183	3.96E-41
	HNRNPC	48.00371	66.38669	0.467748	9.60E-32
	IGF2BP1	0.008226	0.104377	3.665529	1.91E-21
	IGF2BP2	2.221743	1.617101	-0.458283	3.95E-34
	IGF2BP3	0.022774	0.143698	2.657563	0.012057
	YTHDC1	15.43507	13.762006	-0.165522	5.77E-05
	YTHDC2	4.24979	4.536147	0.094076	0.535395
	YTHDF1	17.657637	26.843499	0.60428	3.45E-36
	YTHDF2	23.898463	26.197081	0.132488	0.009291
	YTHDF3	15.936446	17.421101	0.128506	0.658005

Table 2 Baseline characteristics of BC patients in TCGA cohort (n=908).

Clinical Characteristics	All	Training test	Validation test	*p*-value
No.	908	636	272	-
Age, median (IQR)	58 (48 to 66)	58 (48 to 66)	58 (49 to 66)	0.83
Gender, No. (%)				
Female	897 (98.79%)	628 (98.74 %)	269 (98.90%)	0.85
Male	11 (1.21%)	8 (1.26 %)	3 (1.10%)	
T Stage, No. (%)				
T1	234 (25.77%)	164 (25.79%)	70 (25.74%)	0.50
T2	539 (59.36%)	377 (59.28%)	162 (59.56%)	
T3	103 (11.34%)	69 (10.85%)	34 (12.50%)	
T4	32 (3.52%)	26 (4.09 %)	6 (2,21%)	
N Stage, No. (%)				
N0	449 (49.45%)	306 (48.11%)	143 (52.57%)	0.66
N1	301 (33.15%)	215 (33.81%)	86 (31.62%)	
N2	103 (11.34%)	75 (11.79%)	28 (10.29%)	
N3	55 (6.06%)	40 (6.29 %)	15 (5.51%)	
Stage, No. (%)				
Stage1	159 (17.51%)	109(17.14%)	50 (18.38%)	0.40
Stage2	531 (58.48%)	368 (57.86%)	163 (59.93%)	
Stage3	201 (22.14%)	149 (23.43%)	52 (19.12%)	
Stage4	17 (1.87%)	10 (1.57 %)	7 (2.57%)	
Futime, median (IQR)	882 (429.5 to 1796.5)	880.5 (429.5 to 1796)	885.5 (426 to 1829.5)	0.78
Fustat, No. (%)				
Live	784 (86.34%)	553 (86.95%)	231 (84.93%)	0.42
Death	124 (13.66%)	83 (13.05%)	41 (15.07%)	

Note: IQR, interquartile range; Futime, survival time; Fustat, survival state. *p*-value: *p* value represents the difference of clinical characteristics between high and low risk groups.

Table 3 Univariate and multivariate analyses for BC patients' OS (n=908).

Variables	Univariate Cox regression analysis				Multivariate Cox regression analysis			
	HR	Low 95%CI	High 95%CI	*p*-value	HR	Low 95%CI	High 95%CI	*p*-value
Age	1.035	1.02	1.05	<0.001	1.04	1.02	1.05	<0.001
Gender	0.86	0.12	6.15	0.88	0.55	0.076	4.00	0.56
Stage	2.15	1.70	2.72	<0.001	1.71	1.13	2.59	0.01
T	1.53	1.24	1.90	<0.001	1.02	0.77	1.35	0.89
N	1.70	1.42	2.05	<0.001	1.25	0.94	1.67	0.12
Risk Score	1.84	1.45	2.34	<0.001	1.64	1.24	2.17	<0.001

Figures

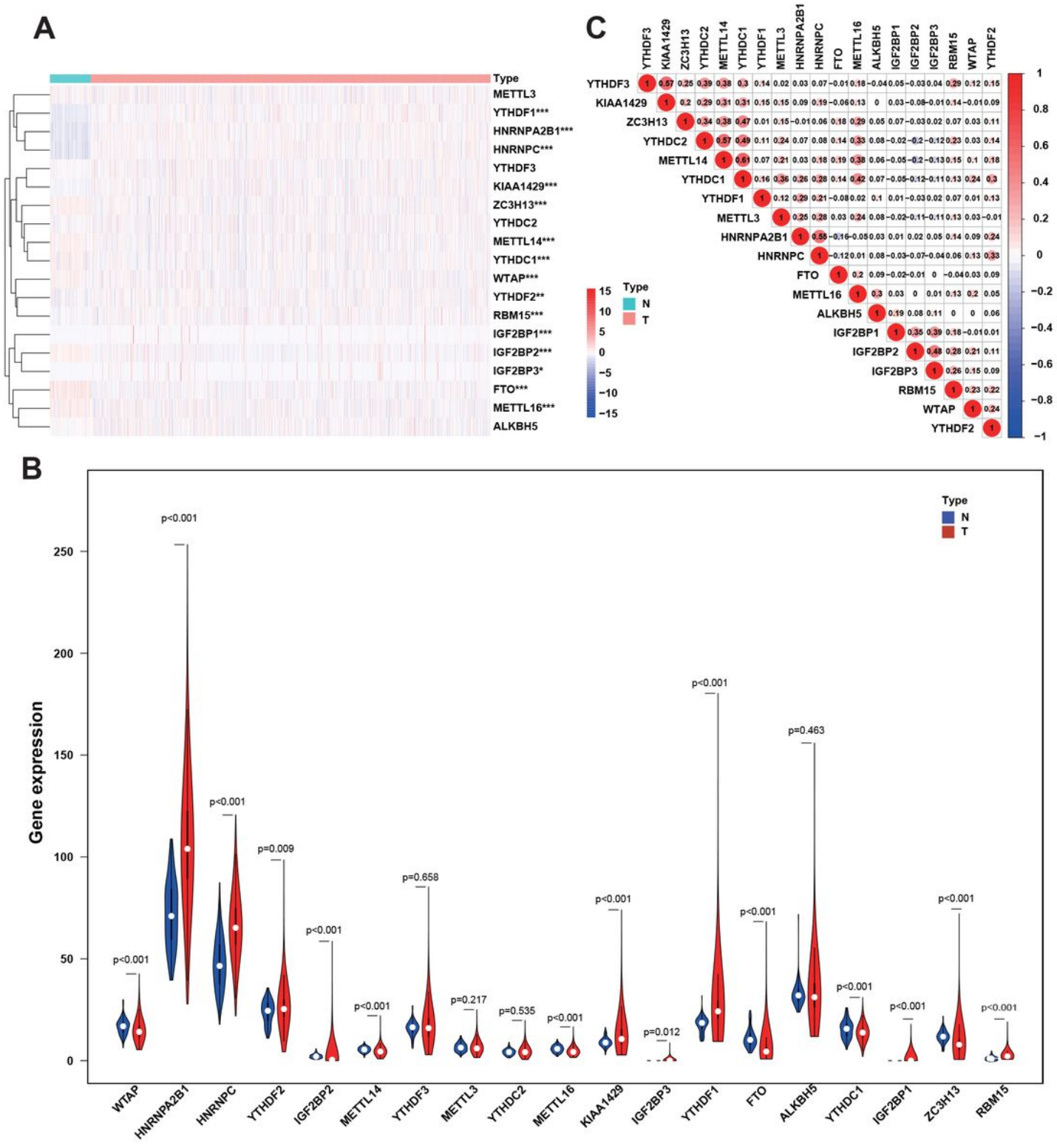


Figure 1

The differential expression and correlation analysis of 19 selected m6A RNA methylation regulators in TCGA BC cohort (N: normal samples, T: tumor samples). A. Heatmap showed the expression levels of m6A RNA methylation regulators in normal and tumor group. B. Violin plot visualizing the differential expression of m6A RNA methylation regulators in tumor and normal samples. C. The Pearson correlation

analysis of the 19 selected m6A RNA methylation regulators. Red dots: positive correlation, blue dots: negative correlation, blank: no correlation.

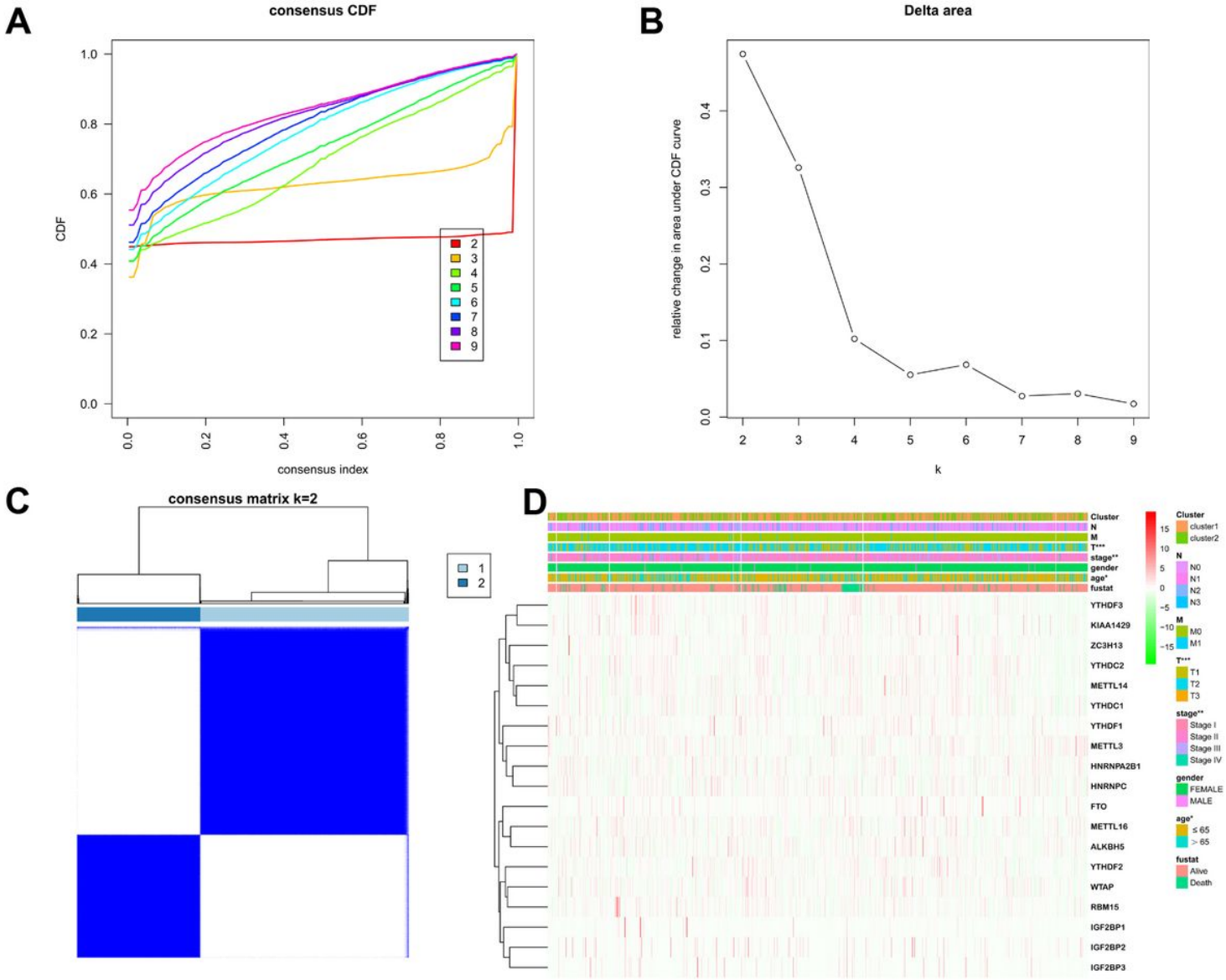
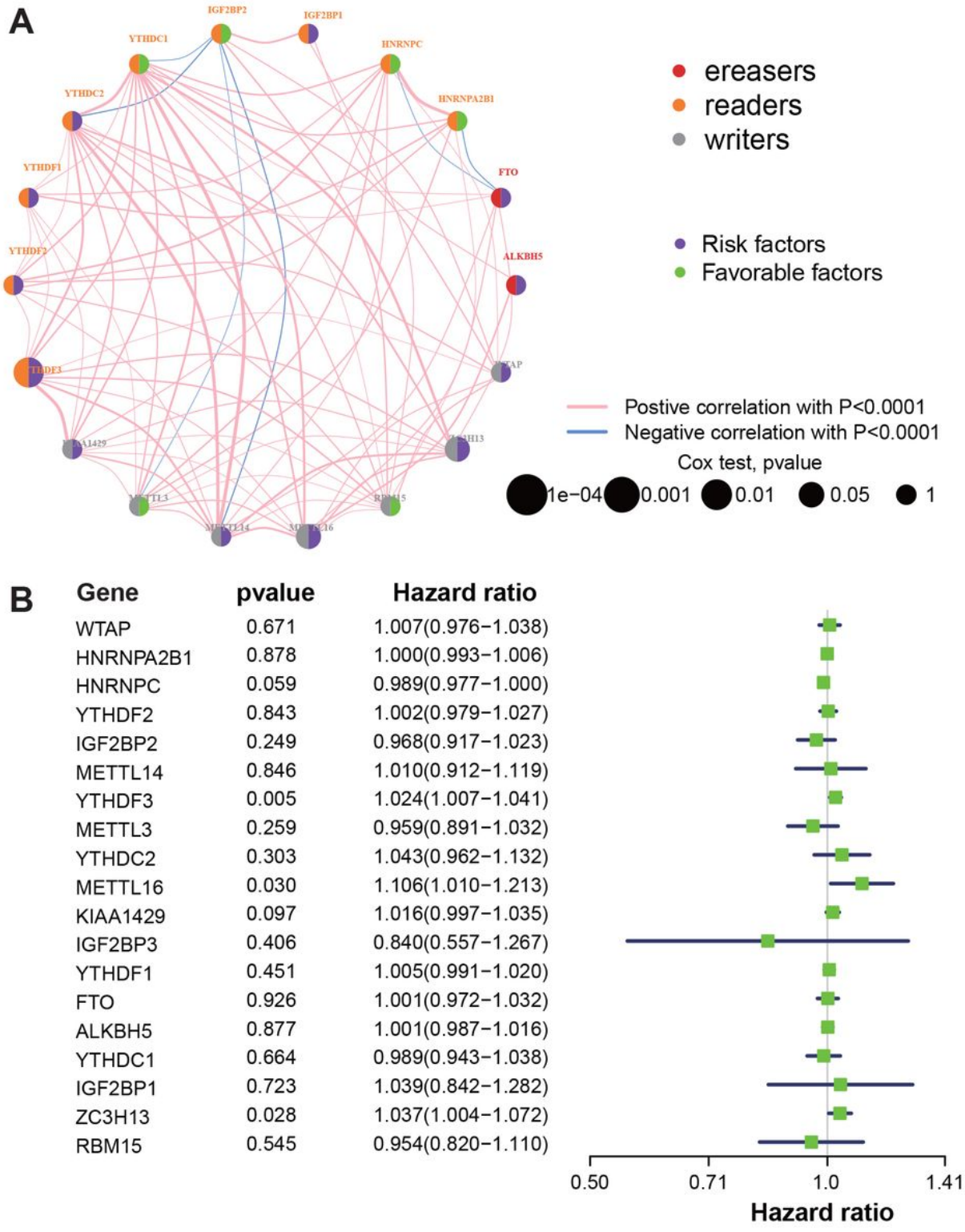


Figure 2

Identification and consensus clustering analysis of 19 selected m6A RNA methylation regulators. A. Consensus clustering cumulative distribution function (CDF) for k = 2-9. B. Relative change in area under CDF curve based on results of consensus clustering for k = 2-9. C. Consensus clustering matrix for k = 2. D. Heatmap and clinicopathologic features of the two clusters identified by the m6A RNA methylation regulators consensus expression.



B

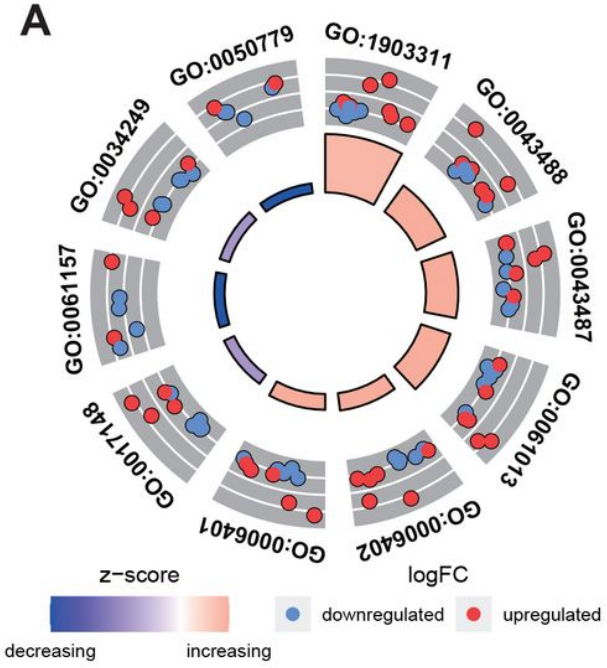
Gene	pvalue	Hazard ratio
WTAP	0.671	1.007(0.976−1.038)
HNRNPA2B1	0.878	1.000(0.993−1.006)
HNRNPC	0.059	0.989(0.977−1.000)
YTHDF2	0.843	1.002(0.979−1.027)
IGF2BP2	0.249	0.968(0.917−1.023)
METTL14	0.846	1.010(0.912−1.119)
YTHDF3	0.005	1.024(1.007−1.041)
METTL3	0.259	0.959(0.891−1.032)
YTHDC2	0.303	1.043(0.962−1.132)
METTL16	0.030	1.106(1.010−1.213)
KIAA1429	0.097	1.016(0.997−1.035)
IGF2BP3	0.406	0.840(0.557−1.267)
YTHDF1	0.451	1.005(0.991−1.020)
FTO	0.926	1.001(0.972−1.032)
ALKBH5	0.877	1.001(0.987−1.016)
YTHDC1	0.664	0.989(0.943−1.038)
IGF2BP1	0.723	1.039(0.842−1.282)
ZC3H13	0.028	1.037(1.004−1.072)
RBM15	0.545	0.954(0.820−1.110)

Figure 3

The prognostic analysis of m6A RNA methylation regulators. A. The landscape and crosslinks between 19 m6A regulators in BC. The circle size represented the effect of each regulator on the prognosis, and the range of values calculated by Cox test was $p < 0.0001$, $p < 0.001$, $p < 0.01$, $p < 0.05$ and $p < 1$, respectively. Purple dots in the circler refer to risk factors of prognosis; Green dots in the circle refer to protective factors of prognosis. The lines linking regulators showed their interactions, and thickness showed the

correlation strength between regulators. Negative correlation was marked with blue and positive correlation with pink. The regulator type of ereasers, readers and writers were marked with red, orange, and gray, respectively. B. Univariate Cox analysis of the 19 selected m6A RNA methylation regulators in TCGA BC cohort.

A



B

ID	Description
GO:1903311	regulation of mRNA metabolic process
GO:0043488	regulation of mRNA stability
GO:0043487	regulation of RNA stability
GO:0061013	regulation of mRNA catabolic process
GO:0006402	mRNA catabolic process
GO:0006401	RNA catabolic process
GO:0017148	negative regulation of translation
GO:0061157	mRNA destabilization
GO:0034249	negative regulation of cellular amide metabolic process
GO:0050779	RNA destabilization

Figure 4

Functional enrichment analysis of prognostic differentially expressed m6A-related regulators in BC patients. A. The outer circle represents the expression (log FC) of m6A-related regulators in each enriched GO term: red dots indicate upregulated m6A-related regulators and blue dots indicate downregulated m6A-related regulators, respectively. The inner circle indicates the significance of GO terms (log10-adjusted P values). B. The description of top 10 enriched GO terms of m6A-related regulators in BC patients.

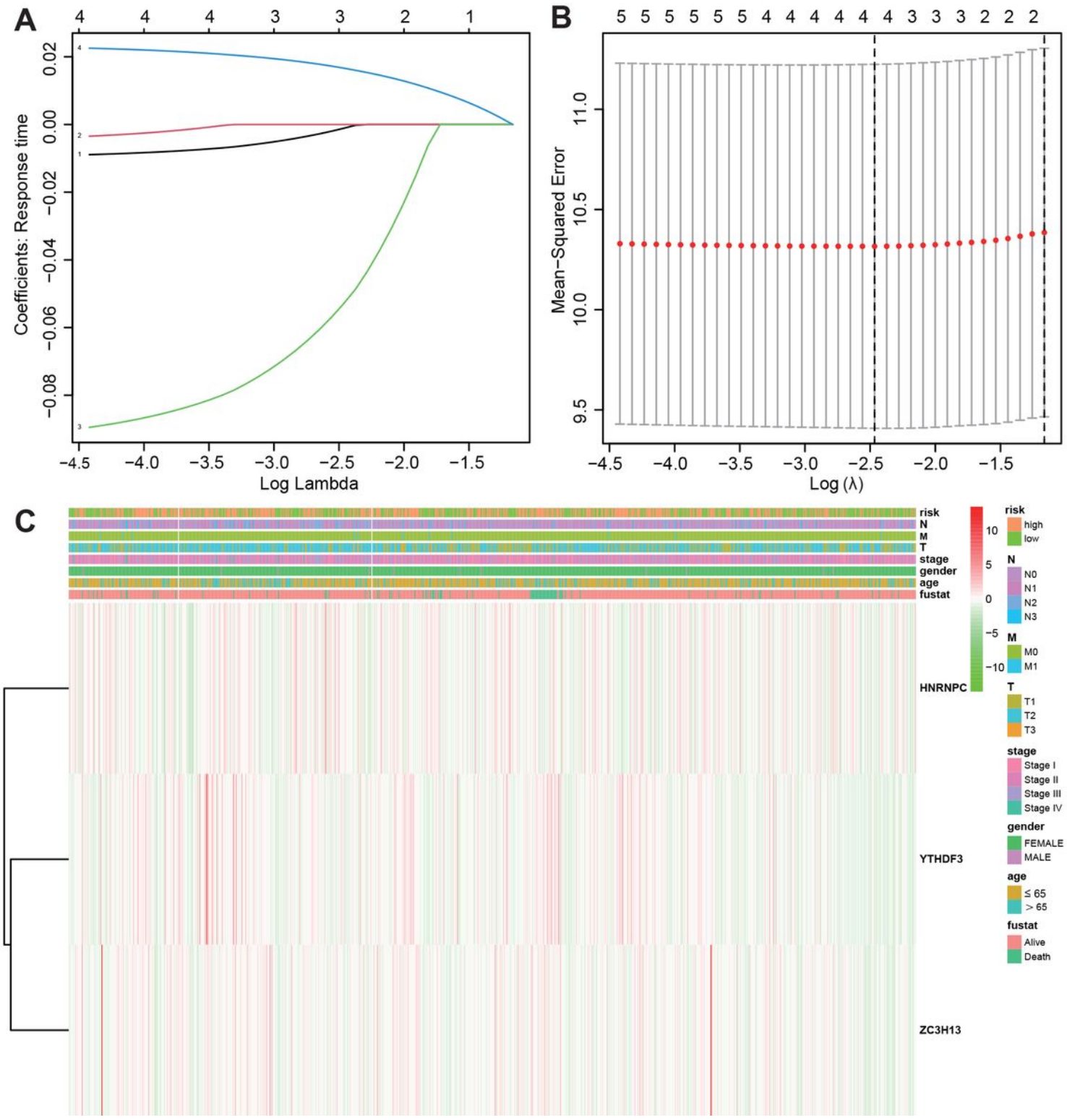


Figure 5

Screening prognostic factors of m6A regulators. A, B. LASSO Cox regression analysis to select m6A RNA methylation regulators with prognostic significance. C. Heatmap showed the risk distribution in BC patients with related clinicopathological characteristics.

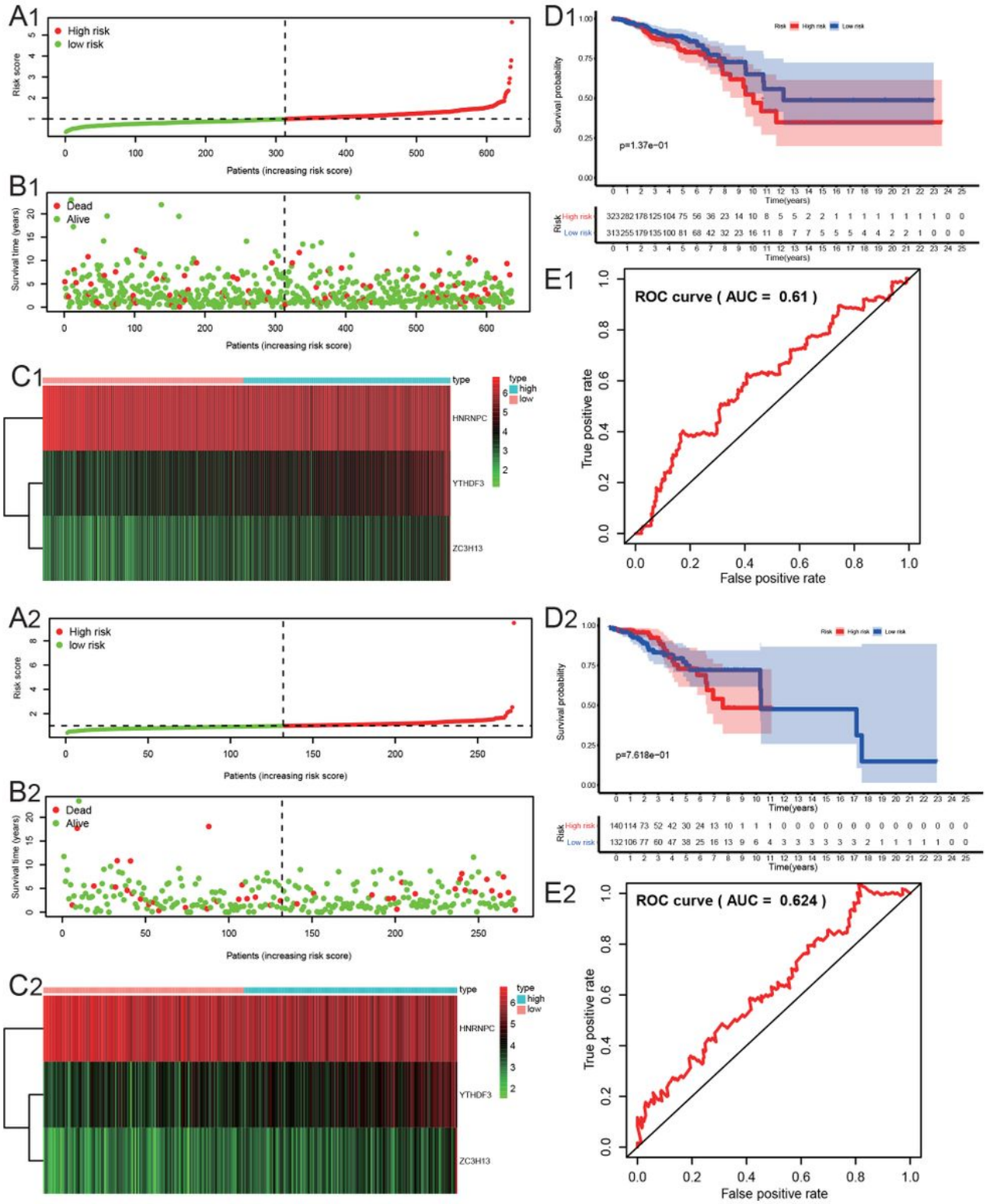


Figure 6

Construction and validation of RS model for survival prediction in training cohort (A1, B1, C1, D1, E1) and validation cohort (A2, B2, C2, D2, E2). A1, A2. Distribution of risk score in BC patients. B1, B2. Survival status and duration of BC patients. C1, C2. Heatmap of the 3 m6A-related genes expression in BC patients. D1, D2. Kaplan-Meier curves showed the OS in the high-risk group and low-risk group. E1, E2. ROC curve evaluated the probability of predicting OS based on the RS model.

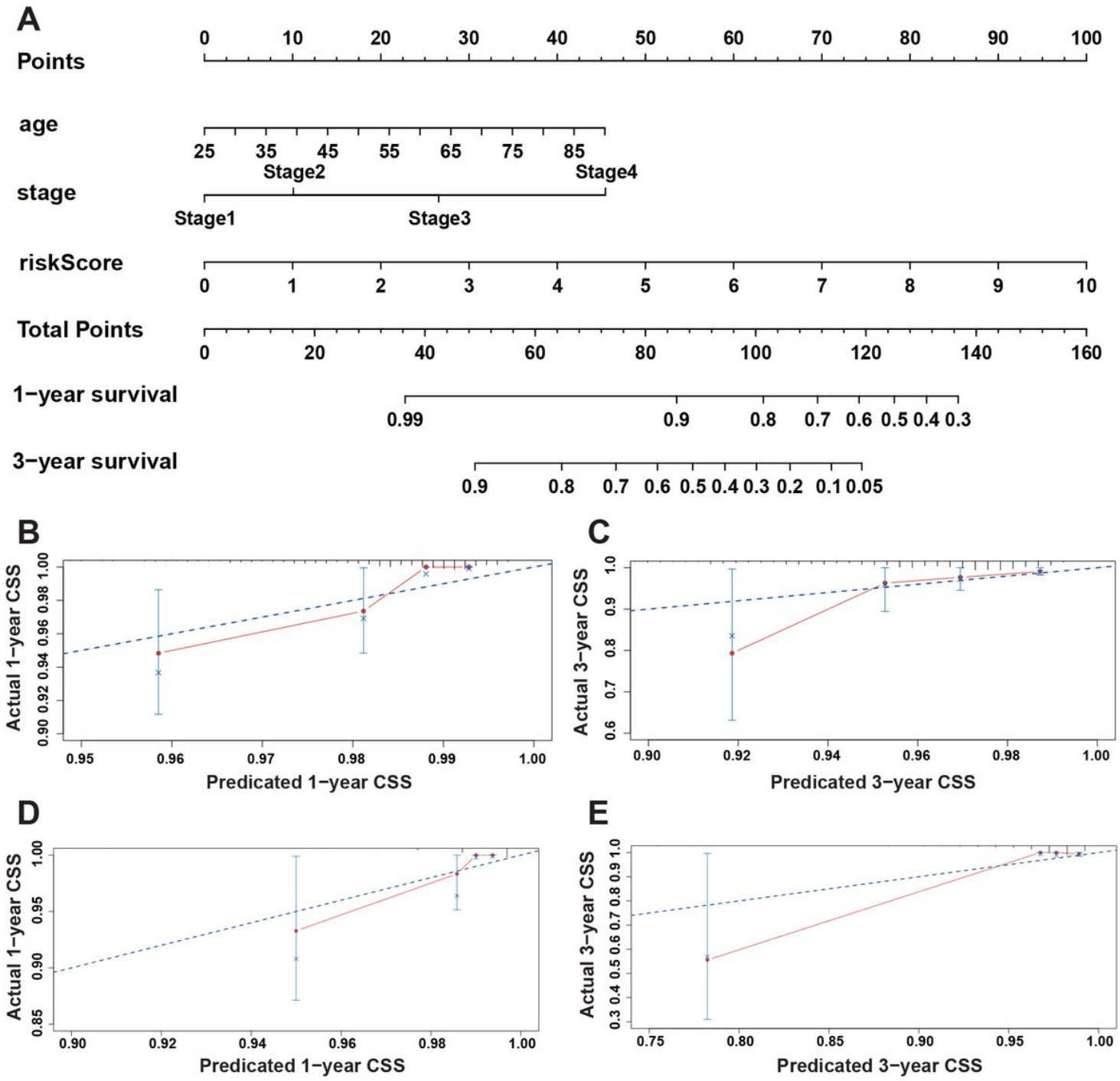


Figure 7

Construction and validation of the prognostic nomogram. A. Construction of the nomogram was based on the RS, patients age and pathological stage in entire TCGA BC cohort. B and C. Calibration plot for the validation of the nomogram in training cohort. D and E. Calibration plot for the validation of the nomogram in validation cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile.docx