

# Comprehensive Characterization of m6A Regulators-Mediated Methylation Modification Patterns and Tumor Microenvironment Signature in Breast Cancer

**Jian Zhang**

Department of Radiation Oncology, Affiliated Cancer Hospital of Guangzhou Medical University

**Huali Jiang**

Tungwah Hospital of Sun Yat-Sen University

**Xiaoshan Hu**

Internal Medicine of Oncology, Affiliated Cancer Hospital of Guangzhou Medical University

**Tao Xie**

Department of Radiation Oncology, Affiliated Cancer Hospital of Guangzhou Medical University

**Kunpeng Du**

Guangzhou Medical University Affiliated Cancer Hospital

**Xi Zhong**

Affiliated Cancer Hospital of Guangzhou Medical University

**Baiyao Wang**

Affiliated Cancer Hospital of Guangzhou Medical University

**Bohong Cen**

Affiliated Hospital of Guangzhou Medical University

**Chengcong Chen**

Affiliated Cancer Hospital & Institute of Guangzhou Medical University

**Yawei Yuan** (✉ [yuanyawei@gzhmu.edu.cn](mailto:yuanyawei@gzhmu.edu.cn))

Department of radiation oncology, Affiliated Cancer Hospital and institute of Guangzhou Medical University <https://orcid.org/0000-0002-8761-1140>

---

**Research article**

**Keywords:** breast cancer, m6A, tumour microenvironment, immunotherapy, tumor mutation burden

**Posted Date:** March 30th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-332570/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

Recent studies have identified the epigenetic regulation in the tumor immune microenvironment (TME). Nonetheless, RNA N6-methyladenosine (m6A) modification mediated TME cell infiltration in breast cancer is still unknown.

## Methods

Herein, we systematically explored m6A modification TME infiltration characteristics of 2249 breast cancer patients and comprehensively correlated the m6A modification patterns with tumor immune landscape. The m6AScore was established using principal component analysis. Multivariate Cox regression analysis was performed to evaluate the prognostic value of the m6AScore.

## Results

Three m6A modification clusters were identified. The TME cell infiltration characteristics under m6A clusters were consistent with the tumor immune phenotypes (-excluded, -inflamed and -desert). The assessment of m6A modification clusters was able to estimate tumor stages, stromal activity (angiogenesis, CD8+ T effector and EMT), clinical prognosis and genomic mutation. The genetic alternation mediated by m6AScore was also associated with enhanced response to targeted therapy and immunotherapy.

## Conclusion

This study demonstrated that m6A regulator-mediated methylation modification plays an important role in TME modulation. Thus, comprehensive evaluation of m6A modification patterns may enhance our understanding of TME modulation and presents a new approach toward targeted therapy and immunotherapeutic strategies for breast patients.

## Introduction

Breast cancer is the most commonly diagnosed cancer worldwide, and its incidence rates have increased rapidly in recent years [1]. Metastatic disease accounts for more than 90% of breast cancer-related deaths, and its molecular and clinical features are characterized by a high degree of heterogeneity, affecting disease progression and treatment response [2, 3]. Therefore, further elucidation of the genetic and epigenetic aberrations underlying the intertumoral heterogeneity of breast cancer is still necessary.

The dynamic epigenetic modifications on mRNAs are an important post-transcriptional regulation of gene expression [4, 5]. N6-methyladenosine (m6A) methylation has been identified as one of the most common modifications that mainly occurs in eukaryotic mRNAs [6]. m6A modification is mainly mediated by m6A methyltransferases (“writers”), binding proteins (“erasers”), and demethylases (“readers”), which is linked

to RNA fate control by regulating pre-mRNA alternative splicing, RNA stability, and mRNA translation [7]. Accumulating evidence identified that m6A modification was associated with a broad impact on DNA damage response, circadian clock control, cell differentiation, and tumorigenesis [8, 9, 10, 11]. However, the correlation between m6A regulators mediated methylation modification patterns and tumor progression and immunomodulatory abnormality in breast cancer remains unknown.

The tumor microenvironment (TME) comprises a complex milieu of non-malignant cells, including blood endothelial cells and lymphatic endothelial cells, mesenchymal cells, and immune cells, along with the extracellular matrix and inflammatory mediators they secrete. Tumor-infiltrating immune cells substantially influence therapeutic responses and clinical outcomes [12, 13, 14]. Chang et al. reported that the m6A reader YTHDF3 promoted breast cancer brain metastasis by inducing the translation of m6A-enriched gene transcripts and controlling the interaction of cancers and the brain microenvironment [15]. The m6A RNA methyltransferases METTL3/14 regulate immune responses to anti-PD-1 therapy by decreasing cytotoxic tumor-infiltrating CD8 + T cells and increasing the secretion of IFN- $\gamma$ , CXCL9, and CXCL10 in the TME [16]. The m6A demethylase ALKBH5 regulates anti-PD-1 therapy response by modulating lactate and suppressive immune cell accumulation in the TME [17]. However, the comprehensive role of TME cell infiltration mediated by multiple m6A regulators requires systematic clarification.

In this study, we integrated the molecular and clinical information of 2249 breast cancer samples to comprehensively analyze the m6A modification patterns, and correlated the patterns with the TME cell infiltration characteristics. Three m6A modification clusters were identified, and we found that the TME characteristics under these three clusters were highly consistent with the immune phenotypes, indicating that m6A modification plays an important role in modulating TME characterization in breast cancer. Moreover, we determined that the m6AScore is an effective prognostic biomarker, and it appears useful for predicting the immunotherapy outcome in breast cancer.

## Materials And Methods

### Dataset acquisition and preprocessing

The gene expression data and clinical annotation of breast cancer were searched in Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) database. Patients without survival information were removed from further evaluation. In total, we used six eligible breast cancer cohorts (GSE20713, GSE42568, GSE20685, GSE25066, GSE124647, TCGA-BRCA [TCGA-Breast Invasive Carcinoma]) for further analysis. The raw CEL data of these six datasets were uniformly processed using the Robust Multichip Average algorithm for background correction and normalization [18]. Next, they were integrated into a meta-cohort using the ComBat function from the R sva package for batch removal [19].

We collected 21 m6A regulators: eight writers (CBLL1, KIAA1429, METTL3, METTL14, RBM15, RBM15B, WTAP, ZC3H13), two erasers (ALKBH5, FTO), and 11 readers (ELAVL1, FMR1, YTHDC1, HNRNPA2B1, HNRNPC, IGF2BP1, LRPPRC, YTHDC2, YTHDF1, YTHDF2, YTHDF3), and 23 tumor-infiltrating immune

cells from previously published studies [20, 21, 22]. Somatic mutation data, transcriptomics data, copy number variation (CNV), and phenotypic data were collected using UCSC Xena database (<https://xenabrowser.net>) and the GDC data portal (<https://portal.gdc.cancer.gov/>). Data were analyzed with R (version 3.6.1) and R Bioconductor packages.

### **Unsupervised consensus clustering**

To identify m6A regulator-mediated m6A modification sub-clusters, tumor samples were clustered into sub-clusters with unsupervised consensus clustering based on the expression matrix of 21 m6A regulators using the R Consensus ClusterPlus package. To guarantee the stability of classification, the parameters were used for clustering as following: number of repetitions = 1000 bootstraps, pltem = 0.8, pFeature = 1.0, clustering algorithm = k-means method. The clustering that exhibited the most significant survival difference was considered.

### **Gene set variation analysis and gene set enrichment analysis**

To investigate the biological processes difference between m6A modification sub-clusters, gene set variation analysis (GSVA) was performed to estimate the pathway enrichment scores using the R GSVA package [23]. The biological processes related to m6A regulators were further analyzed with gene set enrichment analysis (GSEA) using the R clusterProfiler package. The gene ontology (GO) gene sets of c2.cp.kegg.v7.1.symbols were downloaded from the MSigDB database. Adjusted P-values of <0.05 were considered significant.

### **TME estimation**

The relative abundance of the various immune cells in the TME was quantified using the single-sample GSEA (ssGSEA) algorithm [24]. Stromal cell populations, including the angiogenesis information, CD8+ T effector signature, epithelial-mesenchymal transition (EMT), and pan-fibroblast TGF- $\beta$  response (Pan-F-TBRS) signatures, in the TME were quantified as per the immune infiltration estimation (microenvironment cell populations' counter) method described by Becht et al [25].

### **Identification of differentially expressed genes**

To identify m6A-related differentially expressed genes (DEGs), three distinct m6A modification clusters were classified based on the expression of the 21 m6A regulators. DEGs between the modification clusters were determined using the empirical Bayesian approach of the R limma package [26]. The significance criteria for determining DEGs were set as adjusted P-values of <0.05.

### **Construction of the m6A gene signatures**

To quantify the m6A modification clusters of individual tumors, we constructed a set of scoring systems to evaluate the m6A modification clusters of individual patients with breast cancer: the m6A gene signature, also termed the m6Ascore. The m6A gene signatures were established as follows:

The DEGs identified from the different m6A clusters were first normalized, and the overlap genes were extracted. The patients were classified into groups with the unsupervised clustering method for deeper analysis of the overlap DEGs. The number of gene clusters and their stability were defined using the consensus clustering algorithm. The prognostic analysis for each gene in the signature was evaluated using the univariate Cox regression model. The genes with significant prognosis were extracted for further analysis.

As per the method described by Zhang et al. and Teskey et al. [27, 28], the m6A-relevant gene signatures were constructed using principal component analysis (PCA). After obtaining the prognostic value of each gene signature score, the m6Ascore of each patient was obtained with the gene expression grade index method:

$$\text{glyScore} = \sum (\text{PC1}_i + \text{PC2}_i),$$

where  $i$  is the expression of m6A phenotype-related genes.

### **Estimation of immunotherapy sensitivity**

Treatment sensitivity was estimated using the R pRRophetic package. Treatment sensitivity was determined by the concentration required for inhibiting 50% of cellular growth (IC50) [29, 30]. The immunotherapy response was predicted based on the TIDE (tumor immune dysfunction and exclusion) database (<http://tide.dfci.harvard.edu/>).

### **Statistical analysis**

The DEGs were analyzed using the R limma package. Correlation analysis between the TME-infiltrating immune cells and the expression of m6A regulators was conducted using the Spearman method. Composition differences of >2 groups were calculated using one-way analysis of variance (ANOVA) and Kruskal-Wallis tests. The cut-off point of each dataset subgroup was determined using the R survminer package to construct the correlation between survival and m6Ascore. Further, patients were divided into high and low m6Ascore groups based on the maximally selected log-rank statistics. The survival curves for the prognostic analysis were generated via the Kaplan–Meier method, and the significance of differences was identified using log-rank tests. The hazard ratios (HR) for TME-infiltrating immune cells and patient survival were calculated using a univariate Cox regression model. Two-sided  $P < 0.05$  was regarded as statistically significant. All data processing was done in R 3.6.1 software.

## **Results**

### **Genetic variation of m6A regulators in breast cancer**

In this study, we identified and analyzed 21 m6A regulators. **Fig. 1A** shows the dynamic m6A methylation processes mediated by the m6A regulators and their potential biological roles in RNA post-transcriptional regulation. We firstly explored the status of CNVs and somatic mutations of the 21 regulators in breast

cancer. Among 986 breast cancer samples, 84 had m6A regulator mutations, with a frequency of 8.52%. Four writers (KIAA1429, RBM15, WTAP, ZC3H13) and eight readers (ELAVL1, FMR1, YTHDC1, HNRNPA2B1, LRPPRC, YTHDC2, YTHDF1, YTHDF3) exhibited similar mutation frequency and accounted for 1%. Four methylases (CBLL1, METTL3, METTL14, RBM15B), two demethylases (ALKBH5, FTO), and three binding proteins (HNRNPC, IGF2BP1, YTHDF2) did not show any mutations (**Fig. 1B**).

To assess the CNV alteration frequency of the 21 m6A regulators, a prevalent CNV alteration was analyzed, and showed that RBM15B, YTHDF2, WTAP, and ZC3H13 had the highest CNV amplification, with frequencies of 13.4%, 13.4%, 8.8%, and 8.3%, respectively. KIAA1429, IGF2BP1, YTHDF1, YTHDF3, and HNRNPC had a widespread CNV deletion at frequencies of 16.1%, 12.5%, 12.2%, 8.6%, and 8.0%, respectively (**Fig. 1C**). Analysis of the distribution of CNV alterations of the 21 regulators on 23 chromosomes showed that the distribution was irregular and diffused (**Fig. 1D**). Based on the expression of the 21 regulators, we were able to completely distinguish the breast cancer samples from the normal samples (**Fig. 1E**). To ascertain whether the above genetic variations influenced the expression of the m6A regulators, we investigated the regulators' mRNA expression levels between the normal and cancer samples, and found that the CNV alterations could be the prominent factors resulting in the abnormal regulator expression, i.e., YTHDF2, WTAP, ZC3H13, IGF2BP1, YTHDF1, and HNRNPC (**Fig. 1F**). The above analyses reveal the high genetic and expression heterogeneity of m6A regulators in breast cancer. The abnormal expression of m6A regulators may play a crucial role in breast cancer tumorigenesis and progression.

### Identification of m6A-related phenotypes in breast cancer

Six independent BRCA datasets with available overall survival (OS) data and clinical information (GSE20713, GSE42568, GSE20685, GSE25066, GSE124647, TCGA-BRCA) were integrated into one meta-cohort (**Additional file 1**). We first evaluated the m6A regulator interactions, m6A regulator connections, and prognostic values in 2049 breast cancer samples. **Fig. 2A** shows that HNRNPC, METTL3, and YTHDF2 were favorable factors of OS, while LRPPRC, YTHDC1, and YTHDF3 were risk factors of OS. The strongest positive correlation was observed for METTL14 and YTHDC1, and METTL3 and HNRNPA2B1 ( $P < 0.001$ ). On the other hand, several erasers and writers showed negative correlations: KIAA1429 and METTL3 expression correlated negatively with YTHDC1 and METTL14. These results indicate that cross-talk among the 21 m6A regulators may play vital roles in the formation of m6A modification clusters and in the characterization of TME cell infiltration.

To classify patients with different m6A modification clusters based on writer, reader, and eraser expression (**Additional file 2**), three distinct modification clusters were identified with unsupervised clustering: cluster A ( $n = 576$ ), cluster B ( $n = 839$ ) and cluster C ( $n = 834$ ) (**Fig. 2B**). Prognostic analysis of the three clusters indicated a particularly prominent survival advantage in cluster C ( $P = 0.0082$ ) (**Fig. 2C**).

### TME cell infiltration characteristics in three m6A modification clusters

We further conducted GSVA enrichment analysis to clarify the biological behaviors among the three m6A modification clusters. **Fig. 2D** shows that the stromal and carcinogenic activation pathways, i.e., the adherens junction, GAP junction, epithelial cell signaling, cell cycle, interaction, TGF- $\beta$  signaling, cell adhesion, and p53 signaling pathways, were markedly enriched in cluster A vs. cluster B. Meanwhile, the focal adhesion, adherens junction, and cell cycle pathways were enriched in cluster A vs. cluster C. These results indicate that the three clusters are involved in the stromal and carcinogenic activation pathways.

To further know the effect of m6A regulators on TME cell infiltration, 23 TME-infiltrating cells among the three clusters were analyzed. Except monocytes, 22 immune infiltration cells were remarkably enriched (**Fig. 3A**). Univariate Cox regression analysis indicated that activated CD8 T cells ( $P = 0.019$ ), CD56<sup>dim</sup> natural killer cells ( $P = 0.0084$ ), eosinophils ( $P = 0.00025$ ), mast cells ( $P = 0.027$ ), and natural killer cells ( $P = 0.027$ ) were associated with the prognosis of the three clusters (**Fig. 3B**). The activation of stromal cells in the TME, such as cancer-associated fibroblasts, as well as the overexpression of immune checkpoint genes such as PD1/PDL1, can significantly dampen the anti-tumor immunity mediated by cytotoxic T cells [31, 32]. Subsequent investigation confirmed that the m6A modification clusters were characterized by higher infiltration of stromal-related activities such as angiogenesis, EMT, CD8+ T effectors, and pan-fibroblast TGF- $\beta$  (**Fig. 3C**). To determine hormone receptor expression, we analyzed the ER, PR, and HER2 statuses. m6A modification cluster C had a higher ER+ and PR+ rate than clusters B and C, and a higher percentage of triple-negative breast cancer (TNBC) ( $P < 0.001$ ) (**Additional file 3**). Unsupervised clustering also revealed that the 21 regulators were significantly different among the three clusters (**Fig. 3D**). These results show that the DEGs in the m6A modification clusters may play a critical role in maintaining the TME.

### Identification of m6A gene signatures

To investigate the potential biological behavior of each m6A modification cluster, 178 m6A phenotype-related DEGs were overlapped using the R limma package (**Additional file 4**). GO enrichment analysis of the DEGs showed that the biological processes were remarkably associated with m6A modification and immunity, which further showed that m6A modification plays a non-negligible role in TME immune regulation (**Fig. 3E and F**). To explore the potential mechanism, the 178 DEGs underwent unsupervised clustering analyses, which identified three distinct m6A modification genomic clusters: gene cluster A ( $n = 939$ ), gene cluster B ( $n = 606$ ), and gene cluster C ( $n = 704$ ) (**Fig. 4A**). Prognostic analysis showed that the patients with breast cancer in gene cluster C had better prognoses, while patients in gene cluster B had poorer prognoses (**Fig. 4B**). We observed prominent differences in m6A regulator expression in the three gene clusters (\*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ ), which was similar to the results of the m6A methylation modification clusters (**Fig. 4C**).

To accurately predict the pattern of m6A methylation modification in individual patients, a set of scoring systems (m6AScore) was used to quantify the m6A modification pattern of individual patients with breast cancer based on the phenotype-related genes. The attribute changes of the patients were illustrated using an alluvial diagram (**Fig. 4D**). To better illustrate the characteristics of the m6A signature, we also tested

the correlation between the known signatures and the m6Ascore (**Fig. 4E**). **Fig. 4F** shows that gene cluster B had the highest median m6Ascore, while gene cluster C had the lowest median score (\*\*\*\*P < 0.0001). These results indicate that a low m6Ascore could be closely linked to the immune activation-related signatures, whereas a high m6Ascore could be linked to the stromal activation-related signatures. More significantly, compared to the other m6A modification clusters and to m6A modification cluster C, which had the lowest median score, m6A cluster B had a significantly increased m6Ascore (\*\*\*\*P < 0.0001) (**Fig. 4G**). Analysis of stroma-related pathway activity indicated that high scores were significantly associated with enhanced stromal pathway activation (Figure 4H). These results reveal that a low m6Ascore is significantly associated with immune activation and that a high m6Ascore is associated with stromal activation. The m6Ascore was able to better discriminate the m6A modification patterns of breast cancer, and identify TME cell infiltration characterization.

### Clinical characteristics of m6Ascore phenotypes

To determine the clinical characteristics of the m6Ascore, we first analyzed the distribution of the m6Ascore and OS status of patients with breast cancer (**Fig. 5A**). The m6Ascore was calculated for all patients, and the m6Ascore score breakpoint was used to classify patients into high m6Ascore (n = 1124) and low m6Ascore groups (n = 1125). A high m6A score was associated with poor survival, and a low m6A score was associated with favorable survival (**Fig. 5B**). We evaluated the difference in m6Ascore between histological subtypes, and found that the ER+, PR+, HER2-, and non-TNBC subtypes were characterized by better prognosis and correlated significantly with lower m6Ascore, whereas the ER-, PR-, HER2+, and TNBC subtypes were correlated with a higher m6Ascore (**Fig. 5C**). To explore the clinical prognosis value, Kaplan–Meier curve analyses suggested that patients in the low m6Ascore group had significantly better OS probability than those in the high m6Ascore group in TCGA-BRCA (P = 0.0047), GSE42568 (P = 0.00064), GSE124647 (P = 0.00015), GSE25066 (P < 0.0001), and GSE20685 (P = 0.0026) (**Fig. 5D–H**), even though there were no survival differences between the two m6Ascore groups in GSE20713 (P = 0.21) (**Fig. 5I**).

### Genomic mutations of the m6Ascore phenotypes

To determine the genetic and clinicopathological characteristics of patients in the m6Ascore groups, we first compared the CNVs in the m6Ascore groups. **Fig. 6A and B** showed the distribution of the segment score across all chromosomes; there was a significant difference for chromosome 8, 11, and 17. We then compared the proportion of the genome significantly altered on the three chromosomes, and found that the high m6Ascore group showed increased levels of CNVs than the low m6Ascore group. Identification of the somatic mutation landscapes in the high and low m6Ascore groups (**Fig. 6C and D**) showed that the TP53-mutated group had more patients with high m6Ascore than patients with low m6Ascore. The patients with high m6Ascore had a significantly higher level of tumor mutation burden (TMB) than the patients with low m6Ascore, with the rate of the tenth most significant mutated gene being 8% versus 5%. As TMB is closely related to clinical targeted therapy and immunotherapy, it is important to identify the patient subgroup that could be more sensitive to certain drugs.

## Targeted therapy and immunotherapy sensitivity between patients with high and low m6Ascore

We predicted the treatment response to several commonly used targeted therapy drugs in the high- and low-m6Ascore patients with breast cancer. **Fig. 6E and F** showed that the high m6Ascore group in the breast cancer cohorts was more sensitive to lapatinib and sorafenib (\*\*\*\*P < 0.0001). We investigated whether the m6A modification signatures could predict the patient response to immune checkpoint blockade therapy based on the TIDE prediction. The high m6Ascore group had lower TIDE expression, which indicated that the group is much more sensitive to immunotherapy (\*\*\*\*P < 0.0001) (**Fig. 6G**). Further research implied that quantifying the m6A modification patterns is a potential and robust biomarker of prognosis and clinical response assessment of immunotherapy (**Fig. 6H**). This result could yield more clues for determining personalized treatment strategies for patients with breast cancer.

## Discussion

The dysregulation of epigenetic modifications is a common feature of most human cancers, including breast cancer [33, 34, 35]. Apart from DNA methylation and protein phosphorylation or histone modifications, RNA m6A modification has been identified as another important layer of epigenetic modification in transcription and translation [36], which are involved in regulating RNA stability, RNA export, protein translation, and stem cell renewal [4, 37]. The alteration of epigenetic m6A modification is often associated with tumorigenesis and cancer progression. However, recent studies have mostly focused on single TME cell types or single m6A modification regulators; the overall TME infiltration characterization mediated by the integrated roles of multiple m6A regulators is not comprehensively recognized.

In the present study, we comprehensively identified three distinct m6A methylation modification clusters in breast cancer based on 21 m6A modification regulators. We found that the three clusters had differential TME cell infiltration characteristics. Cluster A was identified as innate immunity and stromal activation, corresponding to the immune-excluded phenotype, which is significantly associated with stromal activation status, including the angiogenesis, EMT, and TGF- $\beta$  pathways; cluster B was identified as the suppression of immunity, corresponding to the immune desert phenotype, which is significantly associated with immune tolerance and ignorance, and the lack of activated and priming T cells [38]; cluster C was identified as adaptive immunity activation, corresponding to the immune-inflamed phenotype, which is associated with a large amount of immune cell infiltration in the TME [39, 40]. Tumors lacking the infiltration of immune cells, particularly CD8 + T cells, are usually defined as cold tumors, and such tumors are often associated with fewer benefits from immune checkpoint inhibitors [41]. Prognostic analysis showed that, compared to cluster A and C, cluster B had the immune desert phenotype and poorer prognosis. This could provide novel checkpoints for m6A modification pattern mediation of the TME.

Biological pathway analysis of the DEGs among the three distinct m6A methylation modification clusters showed that the m6A-related gene signatures were associated with various activation of immune

pathway. Similar to the clustering results of the m6A modification clusters, three genomic subtypes were identified based on the m6A gene signatures, which were also significantly correlated with stromal and immune activation. To reduce the heterogeneity of the three genomic subtypes, we established a scoring system based on the m6A-related genes to evaluate the m6A gene signatures. Higher m6Ascores exhibited an immune-excluded phenotype, and vice versa. Integrated analyses also revealed that the m6Ascore was an independent biomarker of poor prognosis in breast cancer. The ER + and PR + histological subtypes, i.e., sensitive to checkpoint endocrine and immunotherapy therapy, were significantly related to higher m6Ascore. Therapy sensitivity analysis showed that the lower m6Ascore group was more sensitive to targeted therapy and immunotherapy. Therefore, the findings from the m6Ascore could provide more clues into personalized treatment for breast cancer.

## Conclusions

In conclusion, we comprehensively evaluated the RNA methylation patterns and TME characterization of m6A regulators in breast cancer and highlighted the indispensable role of m6A modification in shaping the heterogeneity and complexity of the TME. Targeting m6A modifications could be a promising strategy against breast cancer.

## Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

### Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

### Funding

This study was supported by grants from the National Natural Science Foundation of China (82003212), the Guangzhou Key Medical Discipline Construction Project Fund (02-412-B205002-1004042), Guangzhou High Level Clinical Key Specialty Construction Project (2019-2021) and Clinical Key Specialty Construction Project of Guangzhou Medical University (YYPT202017).

### Ethics declarations

### Ethics approval and consent to participate

This study has been approved by INSTITUTIONAL REVIEW BOARD of Affiliated Cancer Hospital & Institute of Guangzhou Medical University

## Consent for publication

Not applicable

## Competing interests

The authors declare no competing interests.

## Author contributions

JZ, HLJ, and XSH performed all experiments, prepared the figures, and drafted the manuscript. JZ, TX, XZ, BYW, and BHC participated in the data analysis and results interpretation. JZ, KPD, CCC and YWY designed the study and participated in the data analysis. All authors have read and approved the manuscript.

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424.
2. Nounou MI, Elamrawy F, Ahmed N, Abdelraouf K, Goda S, Syed-Sha-Qhattal H. Breast cancer: Conventional diagnosis and treatment modalities and recent patents and technologies. *Breast Cancer.* 2015;9:17–34.
3. Bertos NR, Park M. Breast cancer - one term, many entities? *J Clin Invest.* 2011;121:3789–96.
4. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m(6)a RNA methylation. *Nat Rev Genet.* 2014;15:293–306.
5. Roundtree IA, He C. RNA epigenetics—chemical messages for posttranscriptional gene regulation. *Curr Opin Chem Biol.* 2016;30:46–51.
6. Meyer KD, Jaffrey SR. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat Rev Mol Cell Biol.* 2014;15:313–26.
7. 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;28:616–24.
8. Xiang Y, Laurent B, Hsu CH, Nachtergaele S, Lu Z, Sheng W, Xu C, Chen H, Ouyang J, Wang S, et al. Rna m(6)a methylation regulates the ultraviolet-induced DNA damage response. *Nature.* 2017;543:573–6.
9. Tong J, Flavell RA, Li HB. Rna m(6)a modification and its function in diseases. *Front Med.* 2018;12:481–9.
10. Li Z, Weng H, Su R, Weng X, Zuo Z, Li L, Huang H, Nachtergaele S, Dong L, Hu C, et al. Fto plays an oncogenic role in acute myeloid leukemia as a n(6)-methyladenosine rna demethylase. *Cancer Cell.* 2017;31:127–41.

11. Zhang S, Zhao BS, Zhou A, Lin K, Zheng S, Lu Z, Chen Y, Sulman E, Xie K, 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;31:591–606.e6.
12. Dong X, Bai X, Ni J, Zhang H, Duan W, Graham P, Li Y. Exosomes and breast cancer drug resistance. *Cell Death Dis*. 2020;11:987.
13. Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Kalinski P, Endo I, Takabe K. (2020). Plasmacytoid dendritic cell (pdc) infiltration correlate with tumor infiltrating lymphocytes, cancer immunity, and better survival in triple negative breast cancer (tnbc) more strongly than conventional dendritic cell (cdc). *Cancers (Basel)*.12, 3342.
14. Coste A, Karagiannis GS, Wang Y, Xue EA, Lin Y, Skobe M, Jones JG, Oktay MH, Condeelis JS, Entenberg D. Hematogenous dissemination of breast cancer cells from lymph nodes is mediated by tumor microenvironment of metastasis doorways. *Front Oncol*. 2020;10:571100.
15. Chang G, Shi L, Ye Y, Shi H, Zeng L, Tiwary S, Huse J, Huo L, Ma L, Ma Y, et al. Ythdf3 induces the translation of m(6)a-enriched gene transcripts to promote breast cancer brain metastasis. *Cancer Cell*. 2020;38:857–71.
16. Wang L, Hui H, Agrawal K, Kang Y, Li N, Tang R, Yuan J, Rana T. M(6) a rna methyltransferases mettl3/14 regulate immune responses to anti-pd-1 therapy. *EMBO J*. 2020;39:e104514.
17. Li N, Kang Y, Wang L, Huff S, Tang R, Hui H, Agrawal K, Gonzalez G, Wang Y, Patel S, et al. (2020). Alkbh5 regulates anti-pd-1 therapy response by modulating lactate and suppressive immune cell accumulation in tumor microenvironment. *Proc. Natl. Acad. Sci. USA*. 117, 20159–20170.
18. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy–analysis of affymetrix genechip data at the probe level. *Bioinformatics*. 2004;20:307–15.
19. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28:882–3.
20. Zhang B, Wu Q, Li B, Wang D, Wang L, Zhou Y. M(6)a regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer*. 2020;19:53.
21. Chen XY, Zhang J, Zhu JS. The role of m(6)a rna methylation in human cancer. *Mol Cancer*. 2019;18:103.
22. Liu S, Li Q, Chen K, Zhang Q, Li G, Zhuo L, Zhai B, Sui X, Hu X, Xie T. The emerging molecular mechanism of m(6)a modulators in tumorigenesis and cancer progression. *Biomed Pharmacother*. 2020;127:110098.
23. Hanzelmann S, Castelo R, Guinney J. Gsva: Gene set variation analysis for microarray and rna-seq data. *BMC Bioinformatics*. 2013;14:7.
24. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn LF, Schinzel AC, Sandy P, Meylan E, Scholl C, et al. Systematic rna interference reveals that oncogenic kras-driven cancers require tbk1. *Nature*. 2009;462:108–12.

25. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016;17:218.
26. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth G. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43:e47.
27. Zhang J, Zhong X, Jiang H, Jiang H, Xie T, Tian Y, Li R, Wang B, Zhang J, Yuan Y. Comprehensive characterization of the tumor microenvironment for assessing immunotherapy outcome in patients with head and neck squamous cell carcinoma. *Aging.* 2020;12:22509–26.
28. Teskey GC, Kavaliers M. Modifications of social conflict-induced analgesic and activity responses in male mice receiving chronic opioid agonist and antagonist treatments. *Pharmacol Biochem Behav.* 1991;38:485–93.
29. Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol.* 2014;15:R47.
30. Geeleher P, Cox N, Huang RS. Prrophetic: An r package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One.* 2014;9:e107468.
31. Lakins MA, Ghorani E, Munir H, Martins CP, Shields JD. Cancer-associated fibroblasts induce antigen-specific deletion of cd8 (+) t cells to protect tumour cells. *Nat Commun.* 2018;9:948.
32. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer.* 2012;12:252–64.
33. Carter b, Zhao K. (2020). The epigenetic basis of cellular heterogeneity. *Nat. Rev. Genet.*
34. Muhammad JS, Guimei M, Jayakumar MN, Shafarin J, Janeeh AS, AbuJabal R, Eladl MA, Ranade AV, Ali A, Hamad M. Estrogen-induced hypomethylation and overexpression of yap1 facilitate breast cancer cell growth and survival. *Neoplasia.* 2011;23:68–79.
35. Tao C, Luo J, Tang J, Zhou D, Feng S, Zhu Q, Putti T, Xiang T, Tao Q, Li L, et al. The tumor suppressor zinc finger protein 471 suppresses breast cancer growth and metastasis through inhibiting akt and wnt/beta-catenin signaling. *Clin Epigenetics.* 2020;12:173.
36. Yang C, Hu Y, Zhou B, Bao Y, Li Z, Gong C, Yang H, Wang S, Xiao Y. The role of m(6)a modification in physiology and disease. *Cell Death Dis.* 2020;11:960.
37. Roundtree IA, Evans ME, Pan T, He C. Dynamic rna modifications in gene expression regulation. *Cell.* 2017;169:1187–200.
38. Kim JM, Chen DS. Immune escape to pd-1/pd-1 blockade: Seven steps to success (or failure). *Ann Oncol.* 2016;27:1492–504.
39. Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol.* 2015;15:669–82.
40. Gajewski TF, Woo SR, Zha Y, Spaapen R, Zheng Y, Corrales L, Spranger S. Cancer immunotherapy strategies based on overcoming barriers within the tumor microenvironment. *Curr Opin Immunol.* 2013;25:268–76.

## Figures

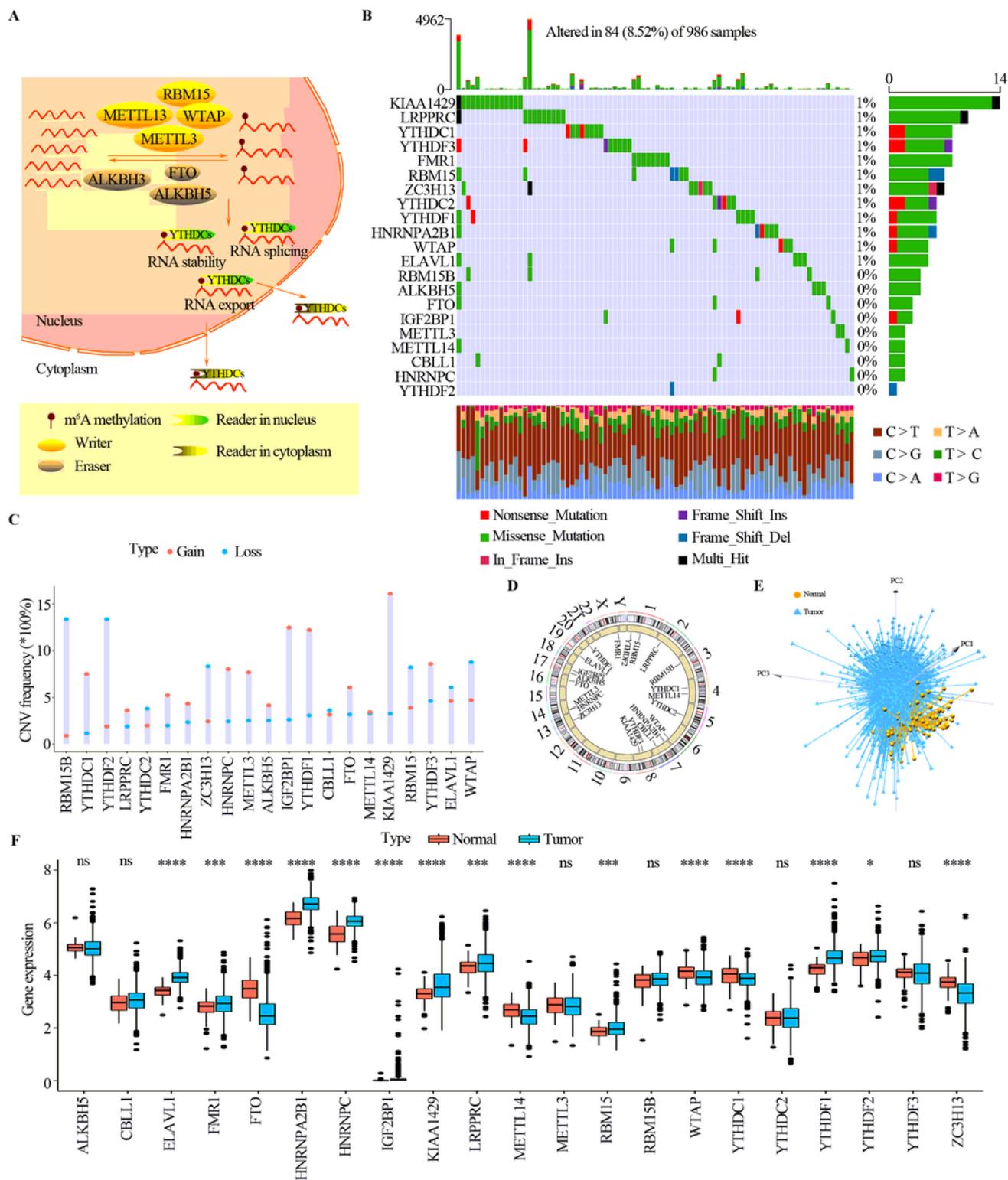
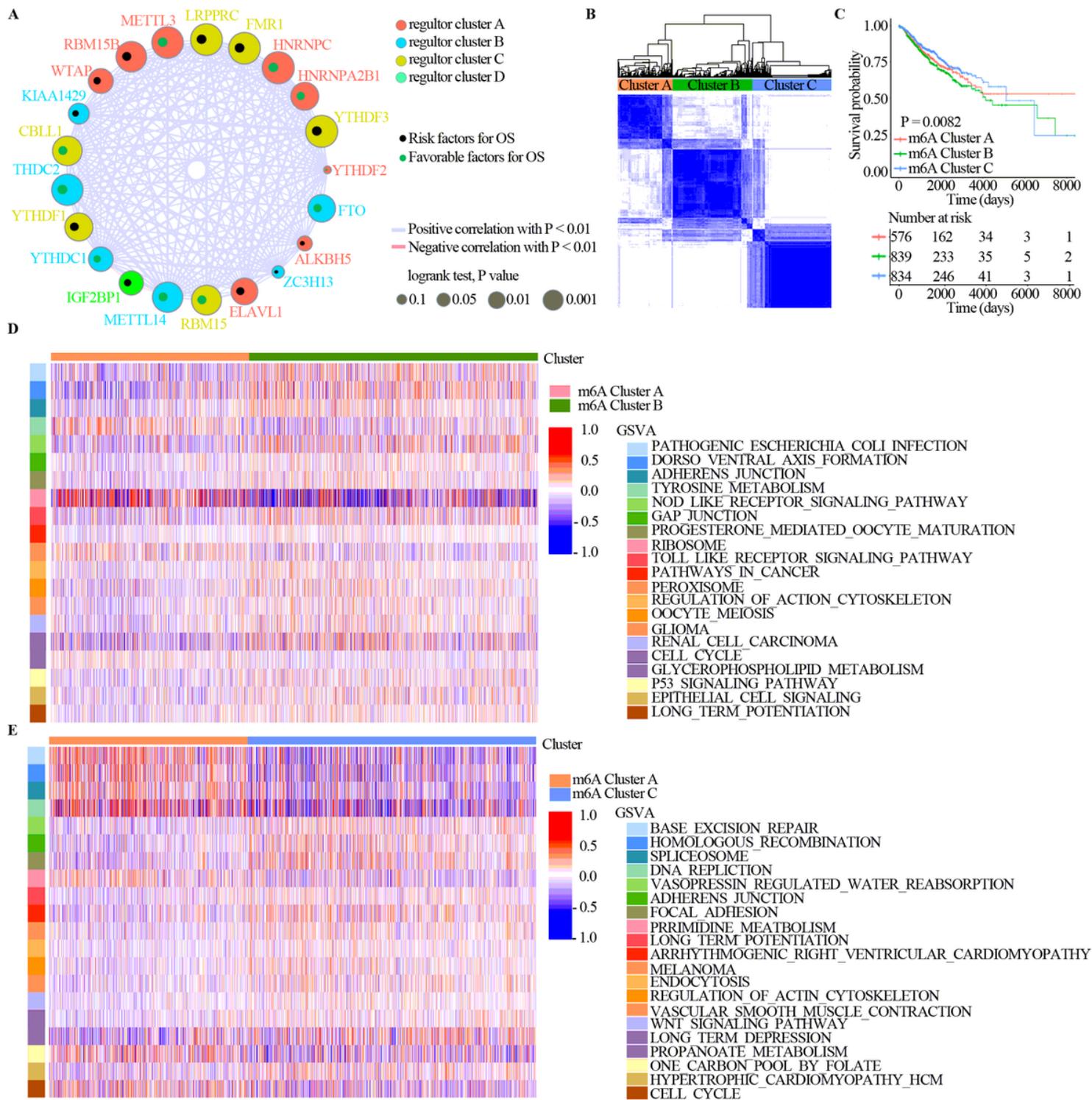


Figure 1

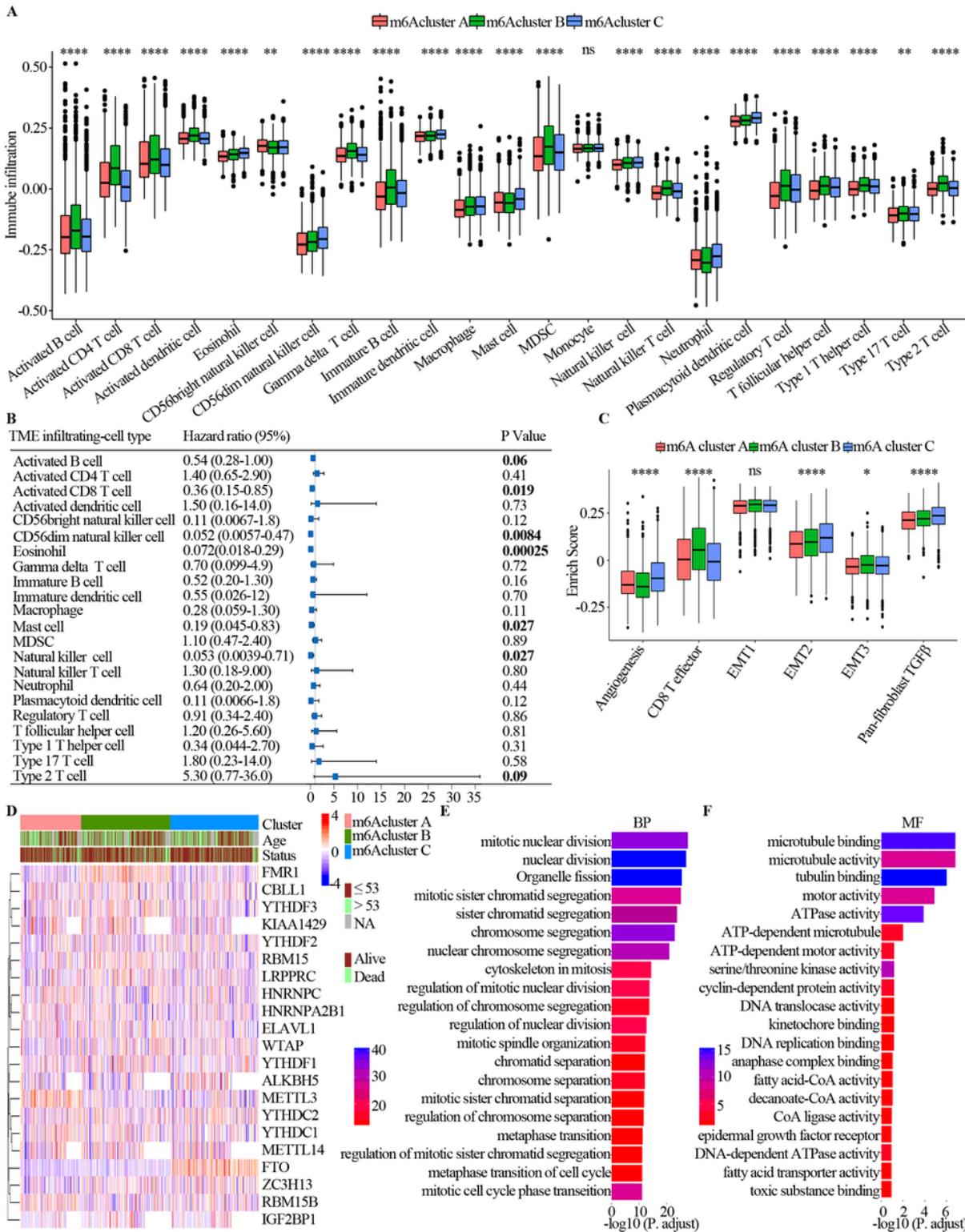
Genetic and expression level of m6A regulators in breast cancer. (A) Summary of the potential biological functions of m6A RNA methylation mediated by their regulators (writers, erasers, readers). (B) The mutation frequency of 21 m6A regulators in 986 patients with breast cancer from TCGA-BRCA cohort. Each column represents an individual patient. The top bar plot shows the TMB; the bottom stacked bar plot shows the fraction of conversions per sample. The numbers on the right indicate the mutation frequency per regulator. The right bar plot shows the proportion of each variant type. (C) The CNV frequency of m6A regulators in breast cancer. The column height represents the amplification frequency (red dot) and deletion frequency (blue dot). (D) The location of CNV alteration of m6A regulators on chromosomes. (E) PCA of the expression profiles of 21 m6A regulators for distinguishing tumor samples from normal samples in TCGA-BRCA cohort. Two subgroups without intersection were identified, indicating that the tumor (blue) and normal (yellow) samples were well distinguished based on the expression profiles of the m6A regulators. (F) The expression of 21 m6A regulators between tumor (red) and normal (blue) tissues. \*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001; ns, not significant.



**Figure 2**

The m6A methylation modification and biological function of the three clusters. (A) The interaction between m6A regulators in breast cancer. The circle size represents the effect of each regulator on the prognosis ( $P < 0.1$ ,  $P < 0.05$ ,  $P < 0.01$ ,  $P < 0.0001$ , log-rank test). Black dots indicate risk factors of OS; green dots indicate protective factors of OS. The lines linking regulators show their interactions, and the line thickness shows the correlation strength between regulators. Blue indicates negative correlation; red indicates positive correlation. (B) Consensus clustering analysis identification of the three m6A

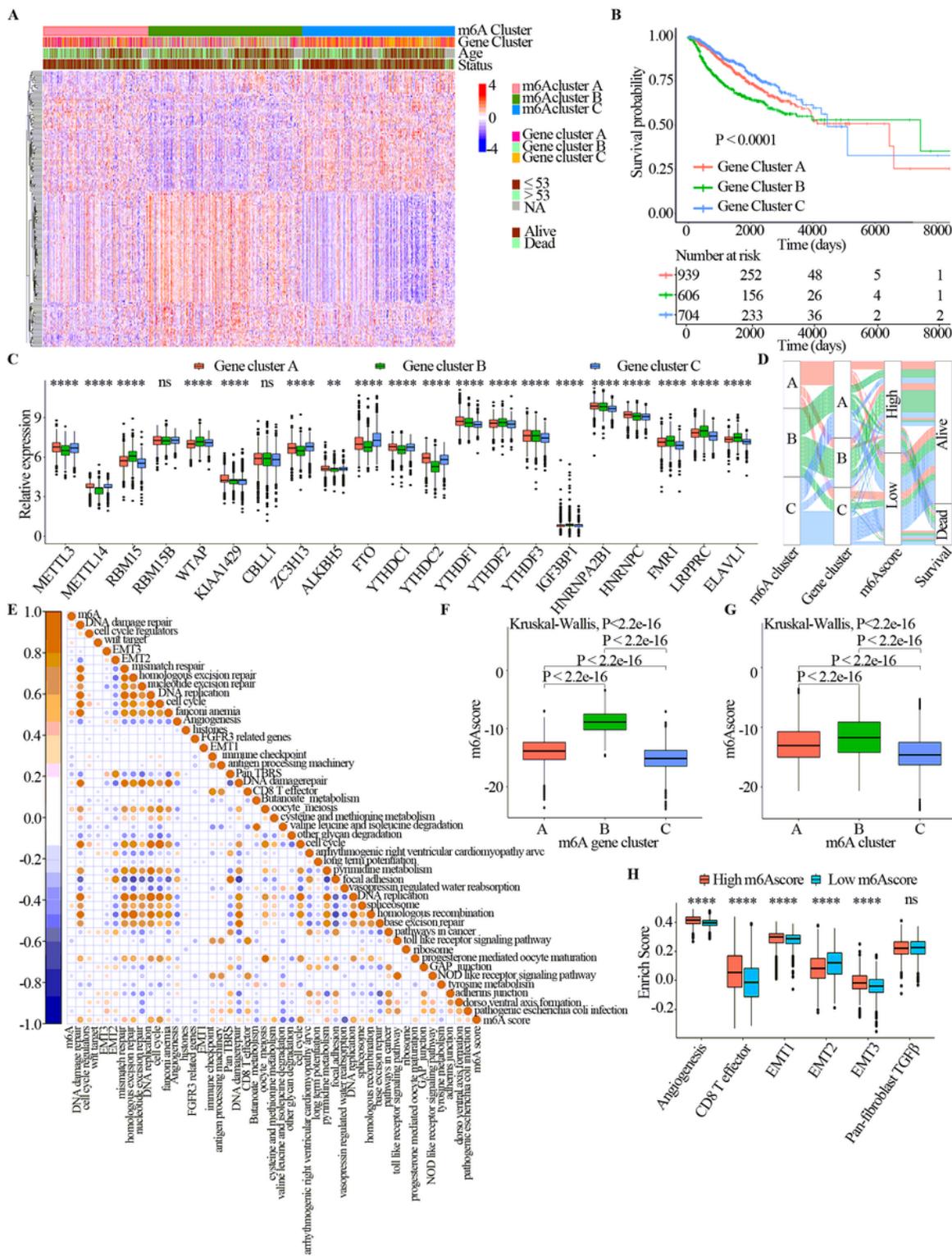
modification clusters. The white (consensus value = 0, samples never clustered together) and blue (consensus value = 1, samples always clustered together) heatmaps illustrate the sample consensus. (C) Survival analyses of the three clusters based on 2249 patients with breast cancer: cluster A, 576 cases; cluster B, 839 cases; cluster C, 834 cases. Kaplan–Meier curves with log-rank P-value of 0.0082 indicate a significant survival difference among the three clusters. Cluster C showed significantly better OS than the other two clusters. (D and E) GSEA enrichment analysis showing the activation states of biological pathways in the three clusters. The biological processes are visualized with the heatmap: red represents activated pathways; blue represents inhibited pathways. The breast cancer cohorts were used as sample annotations: m6A-cluster A vs. m6A-cluster B (D); and m6A-cluster A vs. m6A-cluster C (E).



**Figure 3**

TME cell infiltration characteristics and transcriptome characteristics in the three m6A modification clusters. (A) The abundance of each TME-infiltrating cell in the three clusters. (B) Forest plots of HRs for 22 TME-infiltrating cells associated with OS. (C) Enriched pathways in the three clusters, i.e., the angiogenesis, CD8 T effector, EMT, and TGF- $\beta$  pathways. (D) Unsupervised clustering of 21 m6A regulators in the three clusters. Red represents high regulator expression; blue represents low regulator

expression. (E and F) Functional annotation for m6A-related genes in BP (biological processes) (E) and MF (molecular function) (F). The color depth of the bar plots represents the number of genes enriched. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; ns, not significant.



**Figure 4**

Construction of m6A signatures in breast cancer. (A) Heatmap of overlapping m6A phenotype-related genes in breast cancer for classifying patients into genomic subtypes, termed m6A gene cluster A–C. (B)

Kaplan–Meier curves indicating that the m6A modification genomic phenotypes were markedly related to OS in 2249 patients, of which 939 cases were in gene cluster A, 606 cases in gene cluster B, and 704 cases in gene cluster C ( $P < 0.0001$ , log-rank test). (C) The expression of 21 m6A regulators in the three gene clusters. (D) Alluvial diagram showing the changes of the m6A modification clusters, gene cluster, m6AScore, and OS. (E) Spearman analysis showing correlations between m6AScore and the known gene signatures in breast cancer. Blue indicates negative correlation; orange indicates positive correlation. (F) Differences in m6AScore among the three gene clusters in breast cancer ( $P < 0.001$ , Kruskal-Wallis test). (G) Differences in m6AScore among the three m6A modification clusters in breast cancer ( $P < 0.001$ , Kruskal-Wallis test). (H) Differences in enriched angiogenesis, CD8 T effector, EMT, and TGF- $\beta$  pathways between the high and low m6AScore groups. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ ; ns, not significant.

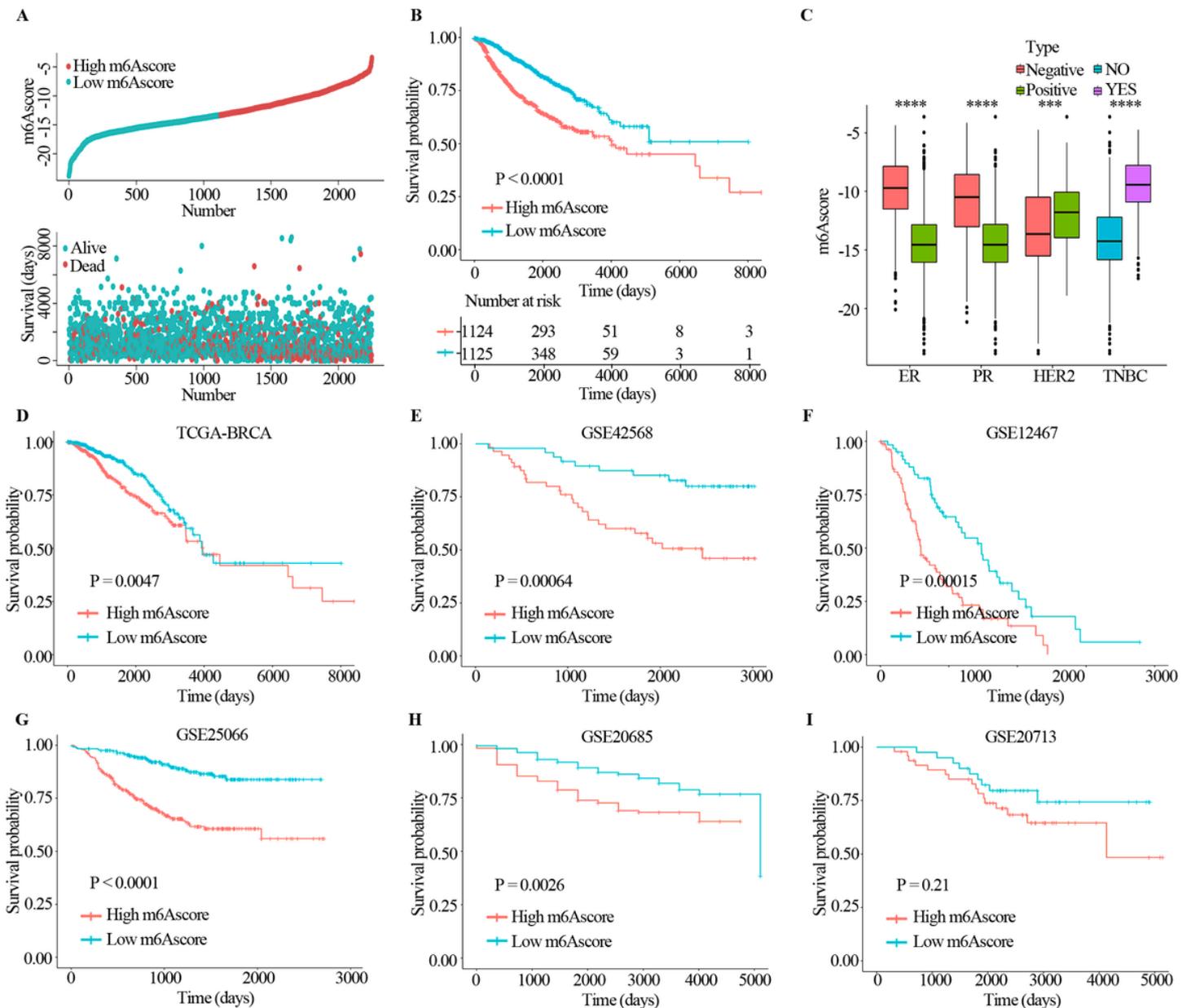


Figure 5

Clinical characteristics and prognosis of m6Ascores. (A) m6Ascore distribution (top) and OS status (bottom) in patients with breast cancer. (B) Kaplan–Meier curves for the high and low m6Ascores. Low m6Ascores were associated with better outcomes than high m6Ascores ( $P < 0.0001$ , log-rank test). (C) Differences in m6Ascore between breast cancer molecular subtypes. (D-I) Survival analyses of patient subgroups stratified by both m6Ascore and treatment in TCGA-BRCA (D), GSE42568 (E), GSE124647 (F), GSE25066 (G), GSE20685 (H), and GSE20713 (I). ( $P < 0.001$ , log-rank test). \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ .

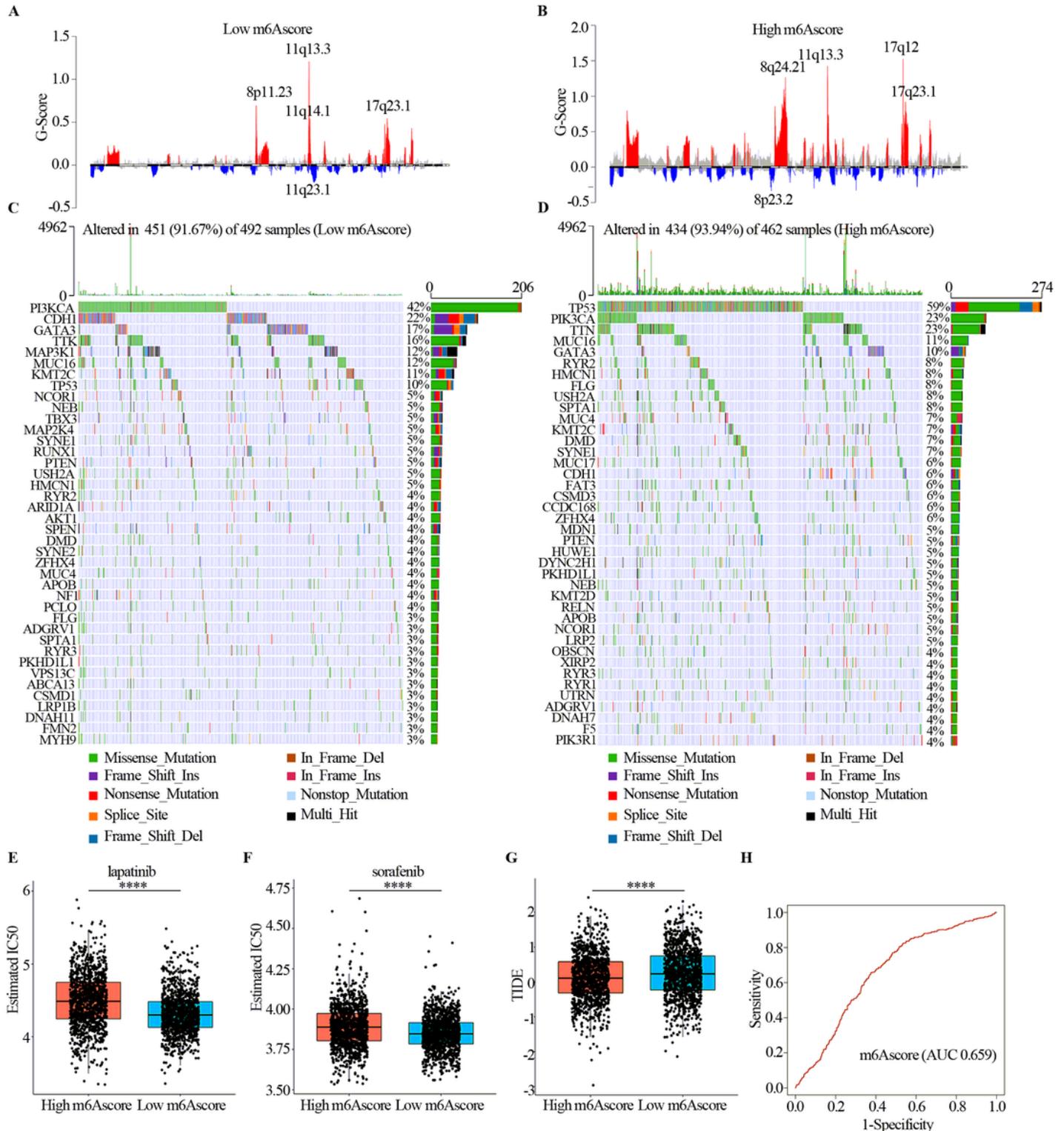


Figure 6

Landscape of tumor somatic mutation and efficacy prediction of targeted therapy in m6Ascore groups. (A and B) CNV at focal regions detected by GISTIC 2.0. Regions of recurrent focal amplifications (red) and focal deletions (blue) in the high (A) and low m6Ascore (B) groups are plotted by the false discovery rate (x-axis) for each chromosome (y-axis). Dashed lines represent the centromere of each chromosome. (C and D) Distribution of highly variant mutated genes correlated with high (C) and low m6Ascore (D) groups. E and F, Differences in m6Ascore among lapatinib (E) and sorafenib (F) response groups ( $P < 0.0001$ , Wilcoxon test). (G) TIDE prediction score differences in m6Ascore groups for anti-PD1 immunotherapy ( $P < 0.0001$ , Wilcoxon test). (H) The predictive value of m6Ascore quantification in breast cancer (area under the curve [AUC] = 0.659). \*\*\*\* $P < 0.0001$ .

## Supplementary Files

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

- [Additionalfile1.xlsx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile3.tif](#)
- [Additionalfile4.xlsx](#)