

Bioinformatics Analysis the Expression Characteristics of RNA m6A Methylation Regulator in Breast Cancer Progression

Ping Zhao

The Third Affiliated Hospital of Kunming Medical University: Yunnan Cancer Hospital <https://orcid.org/0000-0003-4805-3542>

Xinwei Huang

Key Laboratory of The Second Affiliated Hospital of Kunming Medical College, Kunming, 650101, China

Anhao Wu

Department of Mammary Surgery I, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Xin Yang

Department of Blood Transfusion, The First People's Hospital of Yunnan Province, The Affiliated Hospital of Kunming University of Science and Technology, Kunming, 650032, China

Yang Fu

Department of Radiology, The First People's Hospital of Kunming, Kunming 650100, China

Yuhang Quan

Department of Anesthesiology, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Zhen Li

Department of Mammary Surgery III, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Qi Tang

Department of Mammary Surgery II, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Ji Zhang

Department of Mammary Surgery III, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Maohua Wang (✉ wangmaohua324@126.com)

Department of Mammary Surgery I, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming 650118, China

Research

Keywords: m6A methylation-related genes, breast cancer, prognosis, clinical pathological features, infiltrated immune cell, stemness index

Posted Date: January 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-154340/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: RNA m6A methylation regulator has been shown to play an important regulatory role in tumorigenesis and progression. However, the role and characteristics of RNA m6A methylation in breast cancer are not yet clear.

Objective: To analyze the effect of the expression pattern of M6A methylation factor on the progression of breast cancer by bioinformatics.

Results: This study uses bioinformatics to analyze three data sets of TCGA-BRCA, GSE96058 and GSE25066. According to the expression level of m6A-related genes, breast cancer samples are divided into 4 subtypes: quiescent, m6A-methylation, protein-binding, and mixed. The results of R-survival analysis showed that the survival of breast cancer samples of the four subtypes was significantly different. The analysis of breast cancer gene expression features showed that there are significant differences in the number of exon skip among the four subtypes. For tumor driver genes, the degree of TP53 mutation and copy number loss are the most in the protein-binding subtype. Among DNA damage repair genes, the copy number of RAD54B in the protein binding subtype is significantly increased, while other DNA damage repair-related genes have fewer mutations and copy number deletions are common. The effect of m6A methylation on the proportion of infiltrated immune cells indicates that Macrophages M0 and Mast cells resting are significantly different in the four m6A subgroups, and are related to patient prognosis. The number of sensitive samples of chemotherapeutics (taxane anthracyclines) is the lowest in the quiescent subtype. In addition, the highest tumor stemness index in breast cancer samples is the protein-bound type, and the lowest is the m6A methylated type.

Conclusion: The results of our analysis suggest that the reduced expression level of m6A methylation writer gene and the high expression of m6A reader protein in the four methylation subtypes of breast cancer are key to the progression of breast cancer.

Introduction

Worldwide, breast cancer is the most common malignancy cancer in women. Globally, there are more than 1.7 million new cases occur each year, with an increasing rate of 3 to 4 percent per year[1, 2]. Gene mutation and abnormal expression are important reasons for the occurrence and progression of breast cancer[3]. Studies have shown that each stage of tumor occurrence and progression is accompanied by changes in genetic information[4]. Compared with the lower frequency of gene mutation, the frequency of epigenetic modification changes is greatly increased[5]. Epigenetic modifications lead to changes in gene expression that are common in tumor cells. Among them, the abnormal modification of RNA plays a key role in tumor progression. RNA modifications include 5-methylcytosine (m5C), pseudouridine, 2'-O-methylation, N1-methyladenosine (m1A), N7-methyladenosine and N6-methyladenosine (m6A), etc[3, 6]. Among these modifications, RNA m6A methylation is considered to be the most significant and abundant modification form in eukaryotic cells, with its abundance accounting for 0.1–0.4% of total adenosine residues[7]. Since fat-mass and fat-associated protein was reported in 2011, the study on m6A methylation in tumors has attracted extensive attention[8].

M6A modification is a reversible process dynamically regulated by methyltransferases (“Writers”), binding proteins (“Readers”) and demethylases (“Erasers”)[9, 10]. M6A methylation-related proteins are usually abnormally expressed in a variety of human cancer tissues, which in turn leads to abnormal expression levels of cancer-driving or tumor suppressor genes[11]. For example, YTHDF2 promotes the stem cell phenotype and tumor metastasis of hepatocellular carcinoma by recognizing the m6A methylated 5'-UTR of OCT4 mRNA[12]. M6A reader YTHDF1 promotes ovarian cancer progression by increasing EIF3C translation[13]. METTL3 protein and mRNA levels are highly expressed in pancreatic cancer tissues and cancer cells, high METTL3 expression is associated with high pathological stage and high N stage[14]. Abnormal expression of m6A-related proteins has a similar effect in breast cancer. M6A demethylase FTO promoted breast tumor progression by inhibiting BNIP3 expression[15]. In the hypoxic microenvironment, ALKBH5 was highly expressed in breast cancer, leading to the demethylation of m6A on NANOG mRNA and increasing the stability of NANOG mRNA, thus

promoting the enrichment of breast tumor stem cells in [16]. In addition, it has been reported that the abnormal expression of M6A-related genes is related to the prognosis and survival of breast cancer patients, for example, the up-regulated expression of YTHDF1, YTHDF3 and KIAA1429 is related to the poor prognosis of breast cancer[17]. The increase or decrease of m6A modification showed different results in different tumors. At present, the relationship between the abnormal expression of M6A-related genes and clinical outcomes in patients is not clear, and the effect of M6A-related factors on the molecular characteristics of breast cancer has not been systematically reported. It has been previously reported that the mechanism of m6A modification in promoting cancer depends on the expression of m6A methylation-related proteins. Therefore, breast cancer is divided into several subtypes according to the expression pattern of m6A gene. In this study, the relationship between different m6A subtypes and survival, prognosis, mutation, copy number variation, immune cell infiltration, drug sensitivity and tumor stemness index in breast cancer samples was analyzed.

1 Data And Methods

1.1 Data collection

The data of this study are from TCGA and GEO, respectively TCGA-BRCA, GSE96058 and GSE25066. The transcriptome data of the TCGA-BRCA dataset consisted of 1,217 samples, including 1,104 sequenced tumor tissue samples. The GSE96058 data set contains 3409 sequenced samples, including 136 biological replicate samples. The GSE25066 data set contains 508 samples. ComBat (SVA package, V3.36.0) and RUV Seq (V1.22.0) are used for batch effect removal[18]. Two different standardized algorithms, RUVr and RUVg, are used to remove batch effects (Figure S1).

1.2 m6A gene grouping and m6A subtype classification of breast cancer

The M6A gene set contains 8 writer genes (METTL3, METL14, RBM15, RBM15B, WTAP, KIAA1429, CBLL1, ZC3H13), 2 eraser genes (ALKBH5, FTO) and 11 reader genes (YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, IGF2BP1, HRNPA2B1, HNRNPC, FMR1, LRPPRC, ELAVL1). According to biological function, m6A-related genes are divided into two groups: Set1 (writer + reader) and Set2 (eraser). According to the Z-score median of Set1 and Set2 gene expression of the sample, the samples is divided into 4 types. They are static Quiescent ($set1 \leq 0, set2 \leq 0$), m6A methylation ($set1 > 0, set2 \leq 0$), Protein binding ($set1 \leq 0, set2 > 0$) and Mixed ($set1 > 0, set2 > 0$).

1.3 Analysis of the prognosis and clinicopathological indicators of the four m6A subtypes of breast cancer

The survival package of R is used to analyze the prognosis of the four m6A subtypes of breast cancer, and circoS is used to analyze the correlation between the four m6A subtypes of breast cancer and different clinicopathological indicators.

1.4 Analysis of variable splicing, mutation and copy number variation of four breast cancer m6A subtypes

The alternative splicing data of the TCGA dataset is downloaded from TCGA Spliceseq 7 types of alternative splicing are provided by the database: Exon Skip (ES), Alternate Promoter (AP), Mutually Exclusive Exons (ME), Alternate Terminator (AT), Retained Intron (RD), Alternate Donor site (AD) and Alternate Acceptor site (AA). The download parameters are set to: Percentage of Samples with PSI Value = 0, Minimum PSI Range(delta across samples) = 0, Minimum PSI Standard Deviation = 0. Percentage of Samples with PSI Value = 0.75 is used to filter PSI data[19]. The MsigDB database (V7.1) is used to download the gene annotation data of Genes annotated by the GO term GO: 0006281. The R package Maftools was used to extract the top 10 genes with the most mutations in the TCGA-BRCA dataset and GO:0006281, as well as the mutation type and copy number mutation type information of these 20 genes. Heat maps are used to show the distribution of mutations and copy number variations of the top 10 tumor driver genes and DN damage repair genes in the four m6A subtypes.

1.5 Analysis of infiltrated immune cell and immune efficacy of four breast cancer m6A subtypes

Cibersort (V1.01) is used to analyze the proportion of immune cells in the four subtypes of tumor samples[20]. The coxph function in the survival (V3.2-3) package of R is used in the univariate cox analysis of the proportion of 22 immune cells to obtain the immune cell infiltration types that are significantly related to the prognosis of the four m6A subtypes of breast cancer. Meanwhile, the proportion of 22 immune cells in the four m6A subtypes of breast cancer are analyzed by ANOVA. The infiltration types of immune cells in the four m6A subtypes of breast cancer were obtained. Multivariate cox analysis is performed using the covariance of the immune infiltration ratio of different immune cells calculated by Cibersort, and the sum of the value multiplied by the corresponding immune infiltration ratio is used as the sample risk score. TIDE is used to analyze the differences in immune efficacy of the four m6A subtypes in the TCGA-BRCA and GES96058 data sets. The analysis is performed using the TIDE ([HTTPS: /github.com/liulab-dfci/TIDEpy](https://github.com/liulab-dfci/TIDEpy)) default parameter[21].

1.6 Analysis of the sensitivity difference of four m6A subtypes with anthracyclines

According to the method described in 1.2, Breast cancer samples of GSE25066 are divided into 4 m6A types. Twelve m6A-related genes are used for sample typing: Set1 (FTO, WTAP, METTL3, RBM15B, ZC3H13), Set2 (HNRNPA2B1, HNRNPC, ELAVL1, YTHDC1, LRPPRC, FMR1, YTHDC2). In addition, based on the clinical data from the GSE25066 dataset, the sensitive and non-sensitive samples of the anthracyclines were extracted. The differences in the number of drug-sensitive and non-sensitive samples among the four m6A subtypes of breast cancer were compared.

1.7 Analysis of the stemness index of the four m6A subtypes of breast cancer

According to the mRNasi value of TCGA-BRCA samples, the t test was used to analyze the difference in mRNasi dryness index among the 4 m6A subtypes. Cell stemness index mRNasi was downloaded from PMID 29625051[22].

1.8 Statistical analysis

R (V4.0) is used for statistical tests. ANOVA or T-test are used to analyze the differences among different m6A subtypes. P-value < 0.05 is considered statistically significant.

2 Results

2.1 Breast cancer typing results based on the expression level of m6A-related genes

ComBat (SVA package, V3.36.0) and RUVSeq (V1.22.0) are used to remove batch effects between TCGA and GSE data sets. Then, the Z-score of TCGA and GSE data sets is extracted. The expression levels of M6A-related genes were first standardized by Z-score. Then, breast cancer samples were divided into four subtypes, according to the Z-score median of m6A-related genes expression level in the samples in Set1 and Set2 (Fig. 1A). The sample sizes of the four m6A methylation subtypes show similar distributions (proportions) in each data set, which indicates that the typing results of breast cancer samples based on the expression levels of m6A-related genes showed good consistency in different data sets (Table 1). In addition, the expression levels of 21 m6A methylation-related genes in four m6A subtype breast cancer samples are shown in Fig. 1B.

Table 1
Quantity statistics of m6A subtypes in different data sets

	GSE	TCGA	Grand Total	precents GSE	precents of TCGA	TOTAL precents
m6A_methylation	568	158	726	0.17354109	0.14311594	0.16586703
Mixed	882	385	1267	0.26947754	0.34873188	0.28946767
Protein binding	612	152	764	0.18698442	0.13768116	0.17454878
Quiescent	1211	409	1620	0.36999694	0.37047101	0.37011652
Grand Total	3237	1104	4377			

2.2 Analysis of prognosis and clinical pathological features of 4 m6A breast cancer subtypes

The results of survival analysis showed that there were significant prognostic differences among the 4 subtypes (Fig. 2A). Among them, the survival rate of mixed subtype samples is higher than the other three subtypes. In the triple-negative breast cancer TNBC samples, the survival of the four subtypes of breast cancer samples has more significant differences (Fig. 2B). In addition, the analysis results between m6A subtypes and clinical pathological features show that there is a significant quantitative difference between the m6A methylated subtypes and protein binding subtypes of ER+. However, there were no significant differences among the four m6A subtypes in HER+/-, TNM staging, and age (Fig. 2C).

2.3 Analysis of variable splicing, mutation and copy number variation of four m6A breast cancer subtypes

According to the statistical results of TCGA SpliceSeq, we found that there are differences in the number of ES between the four m6A subtypes, and the other six alternative splicing subtypes have no significant changes among the four m6A subtypes (Fig. 3A). Furthermore, the effect of the number of ES type variable shear on the prognosis of the four m6A subtypes was analyzed, and the results showed that the m6A methylated subgroup was significantly different from the protein-bound subgroup in the prognosis, no matter in all samples or in triple-negative breast cancer samples (Fig. 3B and C). Next, the mutations and copy number variations of tumor driver factors and DNA damage repair genes of the four m6A subtype samples were analyzed. The 10 oncogenes with the most mutations in the TCGA-BRCA dataset were selected. The 10 genes are TP53, PIK3CA, TTN, CDH1, GATA3, MUC16, KMT2C, MAP3K1, RYR2, HMCN1. The results showed that TP53 was the gene with the most mutations and copy number loss in protein-binding type breast cancer samples (Fig. 3D). Moreover, the analysis of GO: 0006281 gene set showed that the 10 DNA damage repair genes with the most mutations were TP53, ATRX, BRCA2, ATM, BRCA1, POLQ, SETX, BLM, RAD54B, and ATR. Further analysis of the mutation types and copy number variation types of these 10 genes revealed that except for TP53, DNA damage repair genes were rarely mutated in breast cancer samples. However in terms of copy number variation, the copy number of RAD54B in the protein-bound subtype was increased, while the other 9 genes all showed different degrees of copy number deletion in the 4 m6A subtypes (Fig. 3E).

2.4 Analysis of immune cell infiltration in four m6A subtypes of breast cancer

Cibersort was used to analyze the types and proportions of infiltrated immune cells for TCGA and GSE96058 samples. The results showed that a total of 22 types of immune cell infiltration, of which 11 types showed significantly different infiltration rates among the 4 subgroups (Fig. 4A). Considering the influence of immune cell infiltration on the prognosis of the sample, univariate cox is performed to analyze the relationship between the proportion of immune cell infiltration and patient survival. The results indicated that, in all the samples, the different proportion of NK cells resting and macrophages M0 led to significant differences in patient survival time, with HR < 1 (Fig. 4B). In the triple-negative breast cancer samples, patient survival was significantly affected by the ratio of dendritic cell resting and mast cell resting, and HR < 1 (Fig. 4C). Next, the

types of immune cell infiltration with significant prognostic value in univariate cox analysis were demonstrated by multivariate prognostic analysis (risk value = proportion of immune cell infiltration *HR). The results showed that NK cells resting and macrophages M0 were related to the survival of all samples in univariate prognostic analysis. The risk prediction results of high and low risk samples obtained by using the median risk value as the grouping index does not match the actual risk (risk value = β -NK cells resting \times NK cells resting + β -Macrophages M0 \times Macrophages M0) (Fig. 4D). However, the multi-factor risk prediction of dendritic cells resting and mast cells resting showed good risk prediction effects for triple-negative breast cancer samples (Fig. 4E). Based on the above analysis, mast cells resting and macrophages M0 was taken as the follow-up research object, because of the significant differences between the two types of immune cell infiltration in the m6A subgroup, which also showed significant influence on the prognosis in univariate cox analysis. Further survival analysis revealed that the survival of all breast cancer samples was related to changes in the proportion of Macrophages M0 (Fig. 4F). While, mast cells resting with significant prognostic value in triple-negative breast cancer (Fig. 4G).

2.5 Differences in immune efficacy and sensitivity to anthracycline among the 4 m6A subtypes

TIDE is used as an indicator of differences in immune efficacy. The results showed that there was no significant difference in TIDE between the resting group and the m6A methylation group, as well as the mixed type and the m6A methylation group, and there were significant differences in immune efficacy between the other groups (Fig. 5A). Based on the grouping results of GSE25066 samples (Table 2), the number of anthracyclines resistance and sensitivity samples of the 4 m6A subtype groups were compared. The results showed that the number of drug-sensitive samples in the resting group was the least among the 4 subtypes. This result indicates that the expression of m6A gene contributes to the response of breast cancer to anthracyclines (Fig. 5B).

Table 2
Distribution of four m6A subtypes in GSE25066 samples

	Count of type	precent
m6A_methylation	108	0.21259843
Mixed	150	0.29527559
Protein_binding	118	0.23228346
Quiescent	132	0.25984252
Grand Total	508	

2.6 Differences in stemness index (mRNAsi) among the 4 m6A subtype groups

The higher the stemness index, the stronger the progression of the tumor. Elevated stemness promotes the progression of malignant behaviors of tumor cells, such as rapid and durable self-renewal, poor differentiation, and high malignancy. Our analysis results show that there is no significant difference in the stemness index between mixed and protein-bonding breast cancer samples. There are significant differences in tumor stemness index between breast cancer samples of other m6A subtypes. Specifically, the stemness index of the four subtypes of tumor samples from low to high are: protein-bonding, mixed, quiescent, M6A methylated (Fig. 6). This result suggests that the expression of m6A writer and eraser alone will not lead to the increase of tumor stemness, while the expression of m6A reader is related to the increase of tumor cell stemness. In addition, the reason why the dryness index of resting breast cancer samples is higher than that of m6A methylated samples may be due to non-m6A factors.

3 Discussion

Based on the expression level of m6A in breast cancer, this study took multi-platform and diverse data as the object, and the breast cancer samples were divided into 4 m6A subtypes according to the expression level of m6A. The results of this analysis indicate that m6A has important application value in the prognosis, detection of tumor progression, evaluation of chemotherapy effect and risk prediction of breast cancer patients, and may be a new reference standard for the evaluation of breast cancer progression and prognosis. The formation process of m6A methylation is catalyzed by methyltransferases composed of RBM15, ZC3H13, METTL3, METTL14, WTAP and KIAA1429, while the removal process is mediated by demethylases such as FTO and ALKBH5[23]. M6A methylation level is related to the expression of methyltransferase and demethylase in cells. In addition, a group of specific RNA binding proteins consisting of YTHDF1/2/3, YTHDC1/2, HNRNPA2B1, LRPRC, FMR1, etc. are necessary for m6A modification to exert effects[24]. The binding protein of m6A is a "double-edged sword", which can not only promote the translation of mRNA, but also reduce the stability and accelerate degradation of mRNA. More and more evidences indicate that the dysregulation of m6A regulatory factor is a key factor in the development of tumor.

In this study, differences in clinicopathological indicators, gene mutations, copy number variations, types and proportions of immune cell infiltration, immune efficacy, chemotherapeutic drug sensitivity, and tumor stemness are analyzed among the four m6A subtypes. Specifically, among the four breast cancer subtypes, the survival rate of the protein-binding subtype samples is significantly lower than the other samples. It has been shown that the decreased expression of m6A methylase (METTL3, METTL14 and WTAP) and the increased expression of m6A methylase (FTO and ALKBH5) in BC are closely related to the progression and poor survival rate of breast cancer[25]. The up-regulated expressions of m6A readers YTHDF1 and YTHDF3 are associated with metastasis and poor prognosis in breast cancer[26]. High expression of m6A binding protein has also been reported in other tumors to be associated with poor prognosis and reduced survival of patients except breast cancer. For instance, highly expressed LRPRC is an effective marker for shorter BCP-free survival and OS in patients with metastatic PCa after androgen deprivation therapy[27]. These results indicated that the high expression of M6A-related proteins, especially mA binding proteins, is an important factor for poor prognosis of tumors. In terms of molecular characteristics, TP53 mutations and copy number loss were most common in protein-bound breast cancer samples. It is known that abnormal expression of M6A-related proteins regulates the expression of p53 by changing m6A modification levels. It is reported that the up-regulated expression of YTHDF1 and HNRNPA2B1 in melanoma resulted in the inhibition of p53 expression[28]. However, the regulatory effect of TP53 on m6A-related protein expression has not been reported. We hypothesized that the mutation and copy number loss of TP53, a key tumor suppressor gene, reduced the inhibition of tumor suppressor genes, which may be the key to promoting the up-regulation of m6A binding protein in breast cancer. It is confirmed that the expression of m6A binding protein IGF2BP1 mRNA is increased by 100 times in the study of fibrolamellar hepatocellular carcinoma, while the p53 cancer-suppressing pathway is significantly inactivated in the FL-HCC cells[29]. In addition, CNV deletions of DNA damage repair genes increase genomic instability, which is also an important factor for abnormal expression of M6A-related genes.

Considering the influence of the proportion and type of infiltrating immune cells on the prognosis of the sample, m6A gene expression and the proportion and subtypes of infiltrating immune cells were analyzed. The results revealed that the proportion of NK cells and macrophages M0 was significantly different in the 4 m6A subtypes, and related to the patient survival time. However, in triple-negative breast cancer samples, the ratio of dendritic cells resting and mast cells resting are the main factor for the significant difference in survival time of the four m6A subtype samples. In fact, abnormal expression of m6A-related factors determines the type and proportion of infiltrating immune cells. For example, the infiltration of CD8 + T cells in pancreatic cancer was reduced by the increase and loss of ALKBH5 at arm-level[30]. High expression of WTAP reduces tumor-associated T lymphocyte infiltration and immune response, leading to poor prognosis of gastric cancer[31]. In addition, this study revealed that the immune efficacy of m6A protein-binding breast cancer was significantly lower than that of other subtypes. This approach to predicting tumor therapy based on m6A-related factor expression has also been demonstrated in other tumor species. For example, glioma patients were divided into high-risk

and low-risk groups according to the expression of m6A regulatory factor, and a good predictive effect was obtained in terms of prognosis and treatment effect [32]. In the comparison of chemotherapy resistance, the most chemotherapy-sensitive samples were protein binding subtype breast cancer, while quiescent subtype had the least chemotherapy-sensitive samples. Similar studies have yielded different results in different tumors. Similarly, high expression of METTL3, YTHDF3, and YTHDF1 induced NSCLC drug resistance and metastasis by promoting m6A methylation of YAP mRNA and translating[33]. Finally, the analysis of tumor stemness showed that the m6A protein-binding breast cancer samples had the strongest tumor stemness among the four subtypes, which was consistent with the previous analysis.

In conclusion, the effects of different m6A-related factor expression levels on the prognosis, mutation, immunity and chemotherapy of breast cancer samples were systematically and comprehensively analyzed. This result has positive significance for understanding the role of m6A molecular events in breast cancer.

4 Conclusion

At present, the effect of the overall expression pattern of m6A methylation factor on breast cancer has not been reported. In this study, the classification of breast cancer based on the expression characteristics of m6A methylation regulator can be used to evaluate the disease progression, prognosis, chemotherapy effect, immune cell invasion and stemness index of patients. Our results indicate that Quiescent, M6A methylation, Protein binding and Mixed expression patterns predict different progression and outcome of breast cancer. These findings confirm that the different expression patterns of M6A methylation related factors in breast cancer contribute to provide potential molecular targets for clinical healthcare workers, and are expected to become a new molecular diagnostic and therapy method.

5 Declarations

5.1 Ethics approval and consent to participate

Clinical and animal experiments are not involved in this study. Therefore, ethical and participatory consent elements are not included.

5.2 Consent for publication

This study did not involve the collection of patient samples and data.

5.3 Availability of data and materials

This research involves data from 3 data sets: TCGA-BRCA, GSE96058 and GSE25066. TCGA - BRCA gene expression data, copy number variation and gene mutation data downloaded from the ucsc xena:

[https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Breast%20Cancer%20\(BRCA\)&removeHub=https%3A%2F%2Fxcna.treehouse.gi.ucsc.edu%3A443](https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Breast%20Cancer%20(BRCA)&removeHub=https%3A%2F%2Fxcna.treehouse.gi.ucsc.edu%3A443)

GSE96058 and GSE25066 are downloaded from GEO:

<https://www.ncbi.nlm.nih.gov/geo/>

Clinical data of TCGA-BRCA and GSE96058 were collected:

~/20200729_BRCA_m6A/runtime_ver1.1/data/00_raw_data/GSE96058/GSE96058_clinical.txt

~/20200729_BRCA_m6A/runtime_ver1.1/data/00_raw_data/TCGA-BRCA/TCGA-BRCA.survival.tsv

~/20200729_BRCA_m6A/runtime_ver1.1/data/00_raw_data/TCGA-BRCA/TCGA-BRCA.GDC_phenotype.tsv

Alternative splicing data for the TCGA dataset is downloaded from the TCGA Spliceseq database:

<https://bioinformatics.mdanderson.org/TCGASpliceSeq/PSIdownload.jsp>

In addition, the results analysed during the current study are available from the corresponding author on reasonable request.

5.4 Competing interests

The authors declare that they have no competing interests.

5.5 Funding

The project was funded by Education Department Scientific Research Fund of Yunnan Province [2020J0204 to P.Zhao] Yunnan Fundamental Research Projects (202001AY070001-238 to P. Zhao), Applied Basic Research Programs of Yunnan Province (2018FB129 to XW. Huang), National Natural Science Foundation of China (31860256 to XW. Huang).

5.6 Authors' contributions

Ping Zhao, Xinwei Huang, Anhao Wu contributed to the conception and designed the study. Yuhang Quan, Zhen Li contributed significantly to data analysis and manuscript preparation. Qi Tang performed the data analyses and wrote the initial draft of the manuscript. Xin Yang, Ji Zhang completed statistical analysis and the visualization of the results. Yang Fu, Maohua Wang discussed the results and revised the manuscript.

All authors read and approved the final manuscript.

5.7 Acknowledgements

The authors gratefully acknowledge the help and support of Third Affiliated Hospital of Kunming Medical University in this study. Grateful acknowledgement is made to all colleagues and friends for their contributions to this paper in various ways.

6 References

1. Goldberg MS, Villeneuve PJ, Crouse D, To T, Weichenthal SA, Wall C, Miller AB: **Associations between incident breast cancer and ambient concentrations of nitrogen dioxide from a national land use regression model in the Canadian National Breast Screening Study.** *Environ Int* 2019, **133**:105182.
2. Thorat MA, Balasubramanian R: **Breast cancer prevention in high-risk women.** *Best Pract Res Clin Obstet Gynaecol* 2020, **65**:18-31.
3. Zhao BS, Roundtree IA, He C: **Post-transcriptional gene regulation by mRNA modifications.** *Nat Rev Mol Cell Biol* 2017, **18**:31-42.
4. Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, Nik-Zainal S, Steenbruggen TG, Tjan-Heijnen VCG, Labots M, et al: **The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies.** *Nat Genet* 2019, **51**:1450-1458.
5. Dumitrescu RG: **Interplay Between Genetic and Epigenetic Changes in Breast Cancer Subtypes.** *Methods Mol Biol* 2018, **1856**:19-34.
6. Roundtree IA, Evans ME, Pan T, He C: **Dynamic RNA Modifications in Gene Expression Regulation.** *Cell* 2017, **169**:1187-1200.
7. Hong K: **Emerging function of N6-methyladenosine in cancer.** *Oncol Lett* 2018, **16**:5519-5524.

8. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang YG, He C: **N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO.** *Nat Chem Biol* 2011, **7**:885-887.
9. Sun T, Wu R, Ming L: **The role of m6A RNA methylation in cancer.** *Biomed Pharmacother* 2019, **112**:108613.
10. Chen XY, Zhang J, Zhu JS: **The role of m(6)A RNA methylation in human cancer.** *Mol Cancer* 2019, **18**:103.
11. Zhu GQ, Yu L, Zhou YJ, Du JX, Dong SS, Wu YM, Shi YH, Zhou J, Fan J, Dai Z: **Genetic Alterations and Transcriptional Expression of m(6)A RNA Methylation Regulators Drive a Malignant Phenotype and Have Clinical Prognostic Impact in Hepatocellular Carcinoma.** *Front Oncol* 2020, **10**:900.
12. Zhang C, Huang S, Zhuang H, Ruan S, Zhou Z, Huang K, Ji F, Ma Z, Hou B, He X: **YTHDF2 promotes the liver cancer stem cell phenotype and cancer metastasis by regulating OCT4 expression via m6A RNA methylation.** *Oncogene* 2020, **39**:4507-4518.
13. Liu T, Wei Q, Jin J, Luo Q, Liu Y, Yang Y, Cheng C, Li L, Pi J, Si Y, et al: **The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation.** *Nucleic Acids Res* 2020, **48**:3816-3831.
14. Xia T, Wu X, Cao M, Zhang P, Shi G, Zhang J, Lu Z, Wu P, Cai B, Miao Y, Jiang K: **The RNA m6A methyltransferase METTL3 promotes pancreatic cancer cell proliferation and invasion.** *Pathol Res Pract* 2019, **215**:152666.
15. Niu Y, Lin Z, Wan A, Chen H, Liang H, Sun L, Wang Y, Li X, Xiong XF, Wei B, et al: **RNA N6-methyladenosine demethylase FTO promotes breast tumor progression through inhibiting BNIP3.** *Mol Cancer* 2019, **18**:46.
16. Zhang C, Samanta D, Lu H, Bullen JW, Zhang H, Chen I, He X, Semenza GL: **Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m6A-demethylation of NANOG mRNA.** *Proc Natl Acad Sci U S A* 2016, **113**:E2047-2056.
17. Liu L, Liu X, Dong Z, Li J, Yu Y, Chen X, Ren F, Cui G, Sun R: **N6-methyladenosine-related Genomic Targets are Altered in Breast Cancer Tissue and Associated with Poor Survival.** *J Cancer* 2019, **10**:5447-5459.
18. Zhang Y, Parmigiani G, Johnson WE: **ComBat-seq: batch effect adjustment for RNA-seq count data.** *NAR Genom Bioinform* 2020, **2**:lqaa078.
19. Ryan M, Wong WC, Brown R, Akbani R, Su X, Broom B, Melott J, Weinstein J: **TCGASpliceSeq a compendium of alternative mRNA splicing in cancer.** *Nucleic Acids Res* 2016, **44**:D1018-1022.
20. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA: **Profiling Tumor Infiltrating Immune Cells with CIBERSORT.** *Methods Mol Biol* 2018, **1711**:243-259.
21. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, et al: **Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response.** *Nat Med* 2018, **24**:1550-1558.
22. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, et al: **Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation.** *Cell* 2018, **173**:338-354.e315.
23. 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**:616-624.
24. Patil DP, Pickering BF, Jaffrey SR: **Reading m(6)A in the Transcriptome: m(6)A-Binding Proteins.** *Trends Cell Biol* 2018, **28**:113-127.
25. Wu L, Wu D, Ning J, Liu W, Zhang D: **Changes of N6-methyladenosine modulators promote breast cancer progression.** *BMC Cancer* 2019, **19**:326.
26. 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.** *Am J Cancer Res* 2020, **10**:2546-2554.
27. Zhang HY, Ma YD, Zhang Y, Cui J, Wang ZM: **Elevated levels of autophagy-related marker ULK1 and mitochondrion-associated autophagy inhibitor LRPPRC are associated with biochemical progression and overall survival after androgen deprivation therapy in patients with metastatic prostate cancer.** *J Clin Pathol* 2017, **70**:383-389.

28. Li T, Gu M, Deng A, Qian C: **Increased expression of YTHDF1 and HNRNPA2B1 as potent biomarkers for melanoma: a systematic analysis.** *Cancer Cell Int* 2020, **20**:239.
29. Sorenson EC, Khanin R, Bamboat ZM, Cavnar MJ, Kim TS, Sadot E, Zeng S, Greer JB, Seifert AM, Cohen NA, et al: **Genome and transcriptome profiling of fibrolamellar hepatocellular carcinoma demonstrates p53 and IGF2BP1 dysregulation.** *PLoS One* 2017, **12**:e0176562.
30. Tang R, Zhang Y, Liang C, Xu J, Meng Q, Hua J, Liu J, Zhang B, Yu X, Shi S: **The role of m6A-related genes in the prognosis and immune microenvironment of pancreatic adenocarcinoma.** *PeerJ* 2020, **8**:e9602.
31. Li H, Su Q, Li B, Lan L, Wang C, Li W, Wang G, Chen W, He Y, Zhang C: **High expression of WTAP leads to poor prognosis of gastric cancer by influencing tumour-associated T lymphocyte infiltration.** *J Cell Mol Med* 2020, **24**:4452-4465.
32. Xu S, Tang L, Dai G, Luo C, Liu Z: **Expression of m6A Regulators Correlated With Immune Microenvironment Predicts Therapeutic Efficacy and Prognosis in Gliomas.** *Front Cell Dev Biol* 2020, **8**:594112.
33. Jin D, Guo J, Wu Y, Du J, Yang L, Wang X, Di W, Hu B, An J, Kong L, et al: **Correction to: m6A mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis.** *J Hematol Oncol* 2020, **13**:106.

Figures

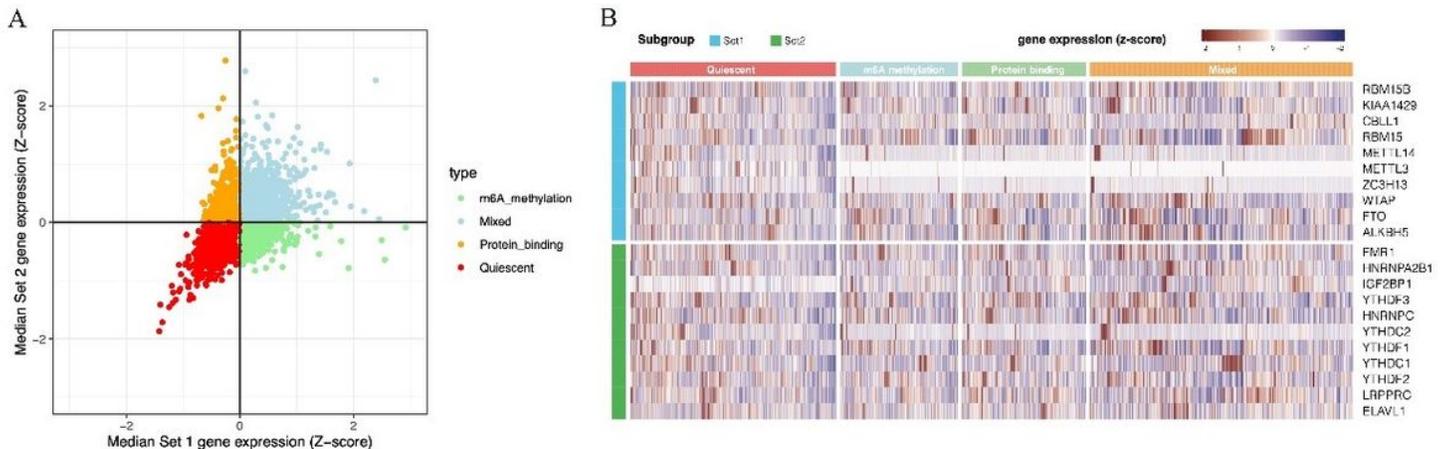


Figure 1

The expression levels of m6A-related genes in the four m6A subtype samples of breast cancer. A: Based on the Z-score of m6A gene expression level, the expression differences of different m6A gene sets among the four breast cancer subtypes are shown. B: Based on the Z-score of the m6A gene expression level, the expression differences of different m6A gene sets among the four subtypes were demonstrated.

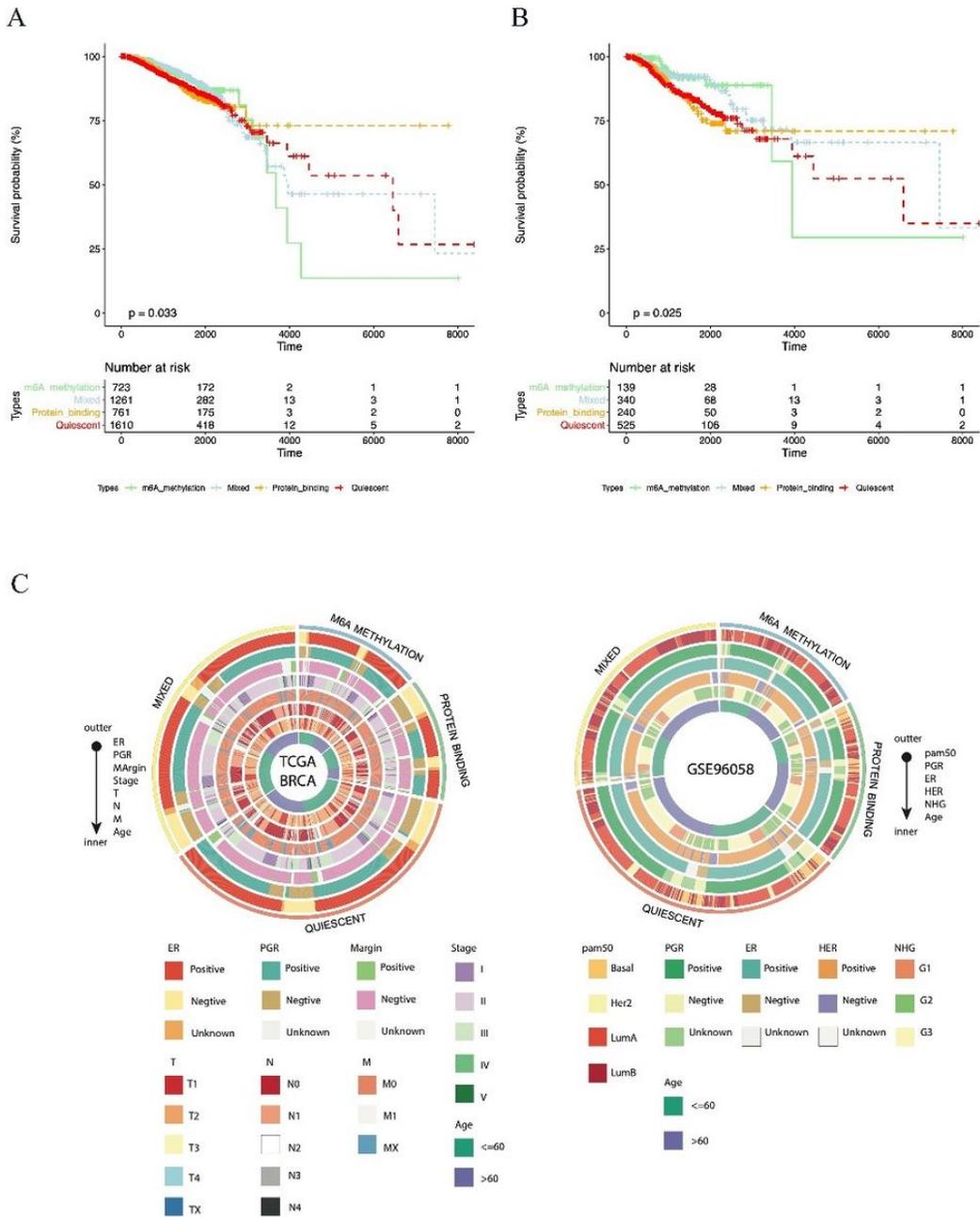


Figure 2

Correlation analysis of prognosis and clinicopathological indicators of four m6A breast cancer subtypes A: Survival analysis of all samples. B: Survival analysis of triple negative breast cancer (TNBC) samples of. C: Correlation of four m6A breast cancer subtypes in TCGA-BRCA and GSE 96058 samples with clinicopathological indicators.

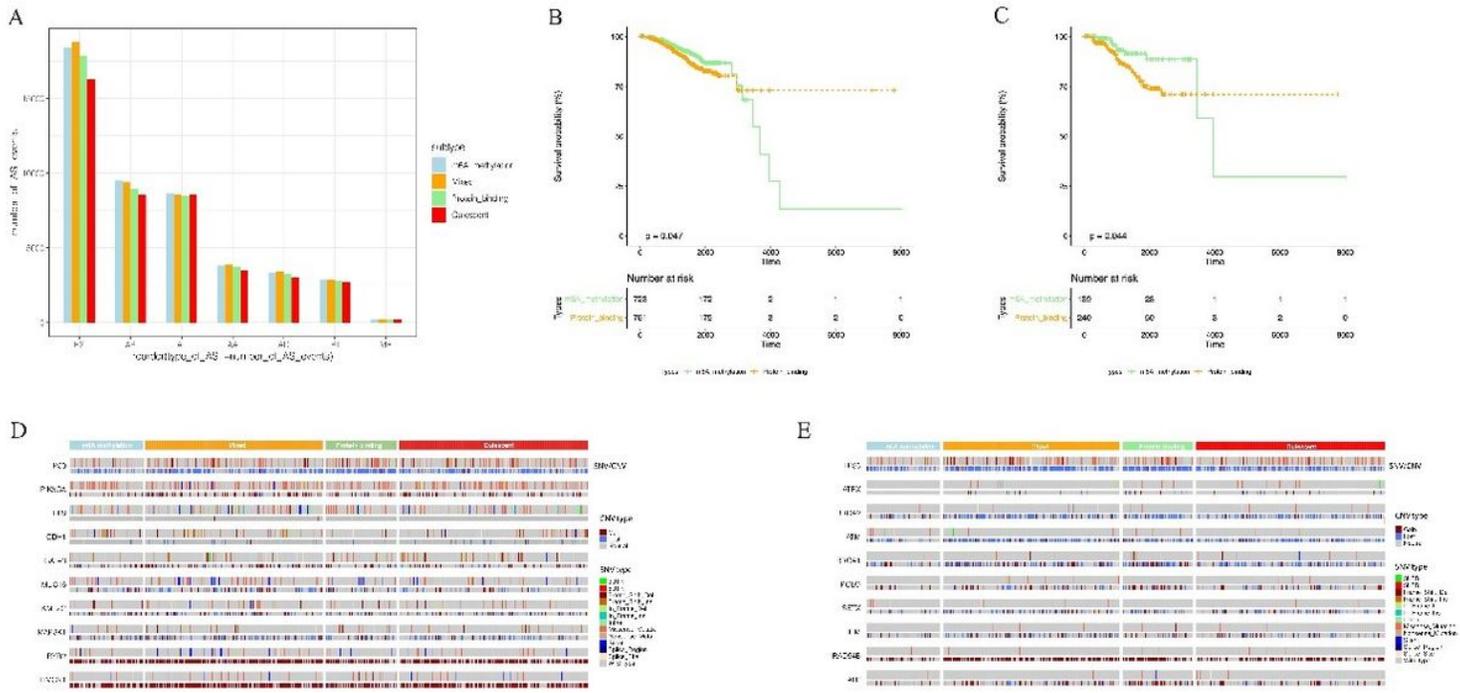


Figure 3

The differences of alternative splicing, mutation and copy number variation in four m6A subtypes of breast cancer A: Statistics of the number of alternative splicing types in 4 m6A breast cancer subtype samples. B: Survival analysis of m6A methylated subtype and protein-bound subtype in all breast cancer samples. C: Survival analysis of m6A methylated subtype and protein-bound subtype in triple negative breast cancer samples. D: Differences in the number and type of tumor driver gene mutations and CNVs among the four m6A subtypes. E: Differences in the number and type of DNA damage repair genes mutations and CNVs among the four m6A subtypes.

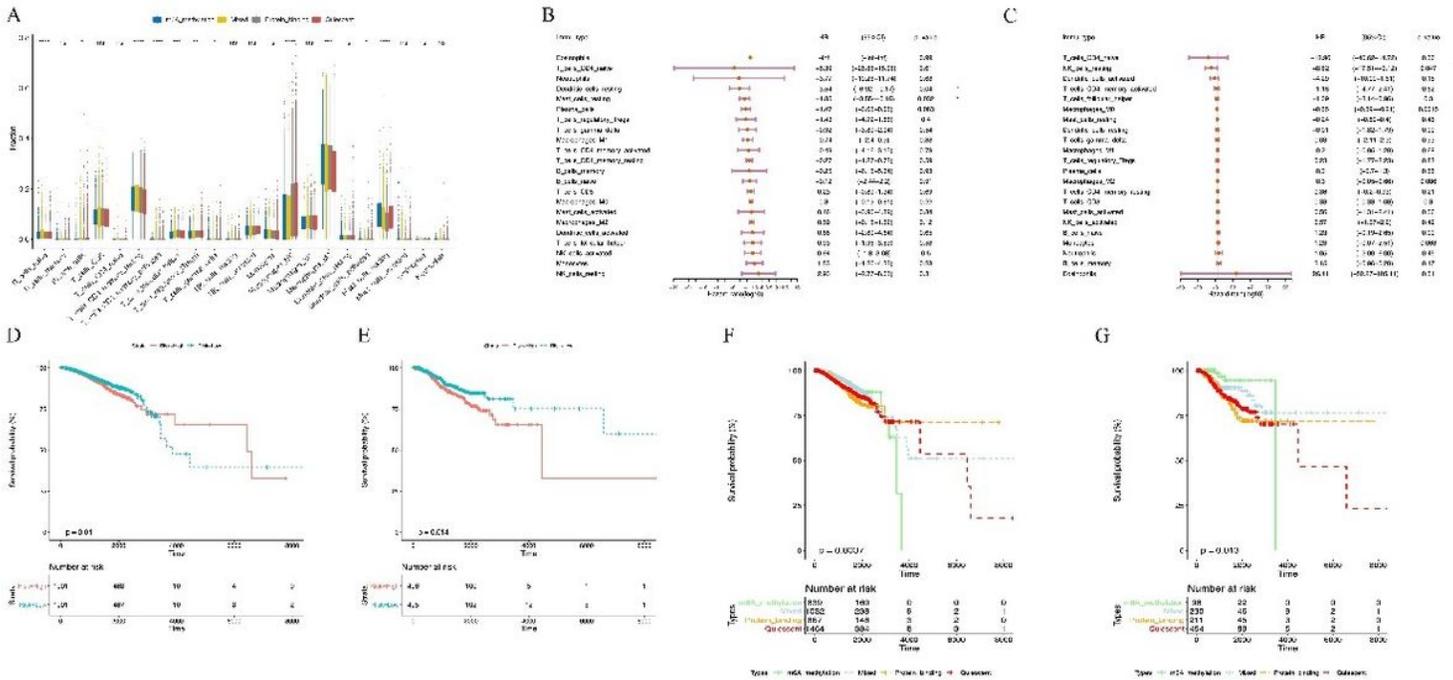


Figure 4

The effect of the types and proportions of infiltrating immune cells on the prognosis of the four m6A subtypes of breast cancer A: Types and proportions of immune cell infiltration in different breast cancer subtypes (ANOVA analysis). B: Univariate COX is used to analyze the effect of immune cell infiltration ratio on the prognosis of breast cancer (all samples). C: Univariate COX is used to analyze the effect of immune cell infiltration ratio on the prognosis of breast cancer (triple negative breast cancer sample). D: Multivariate prognostic analysis of NK cells resting and Macrophages M0 in breast cancer samples. E: Multivariate risk prediction of Dendritic cells resting and Mast cells resting in triple negative breast cancer samples. F: The difference in the proportion of infiltrated macrophages M0 in the prognosis of the four m6A subtypes of breast cancer (all samples). G: The difference in the ratio of mast cell resting infiltration among the 4 m6A breast cancer subtypes (triple negative breast cancer samples)

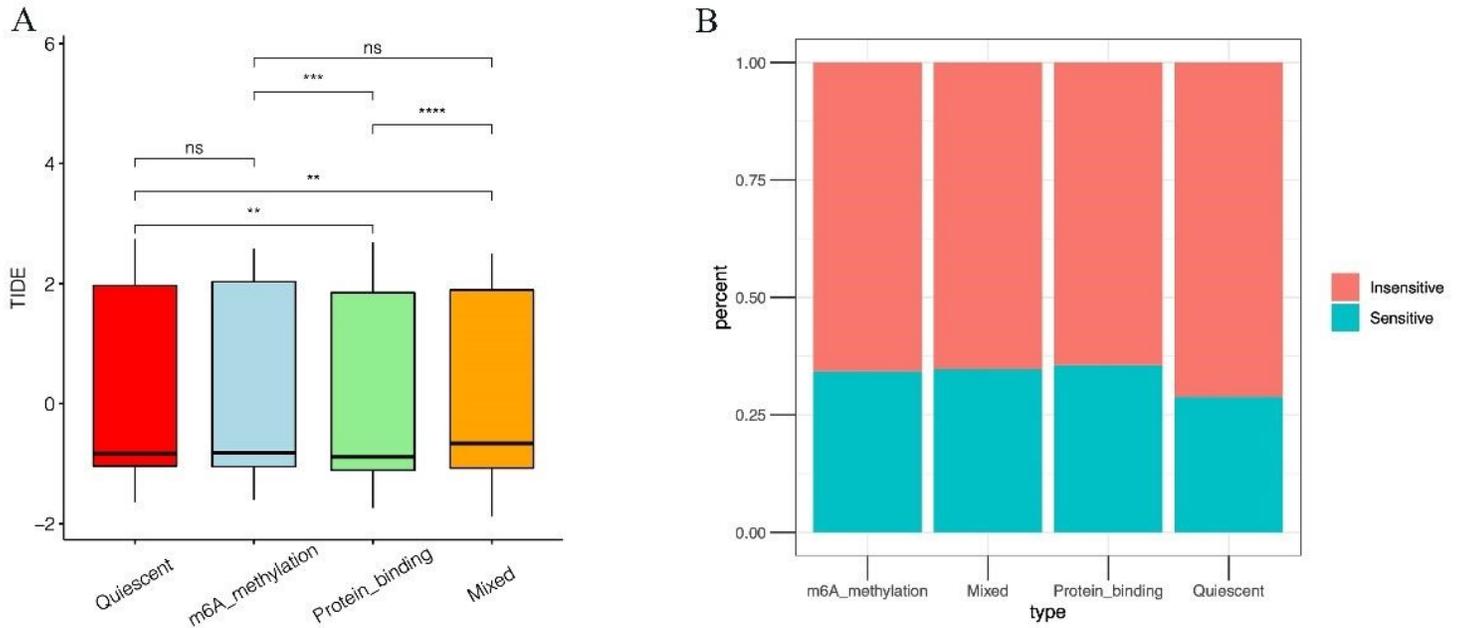


Figure 5

Differences in immune efficacy and sensitivity to anthracycline in 4 breast cancer subtypes of m6A A: Difference in immune response among four m6A subtypes (T-test). B: Differences of resistance and sensitivity to taxane anthracycline of four m6A subtypes.

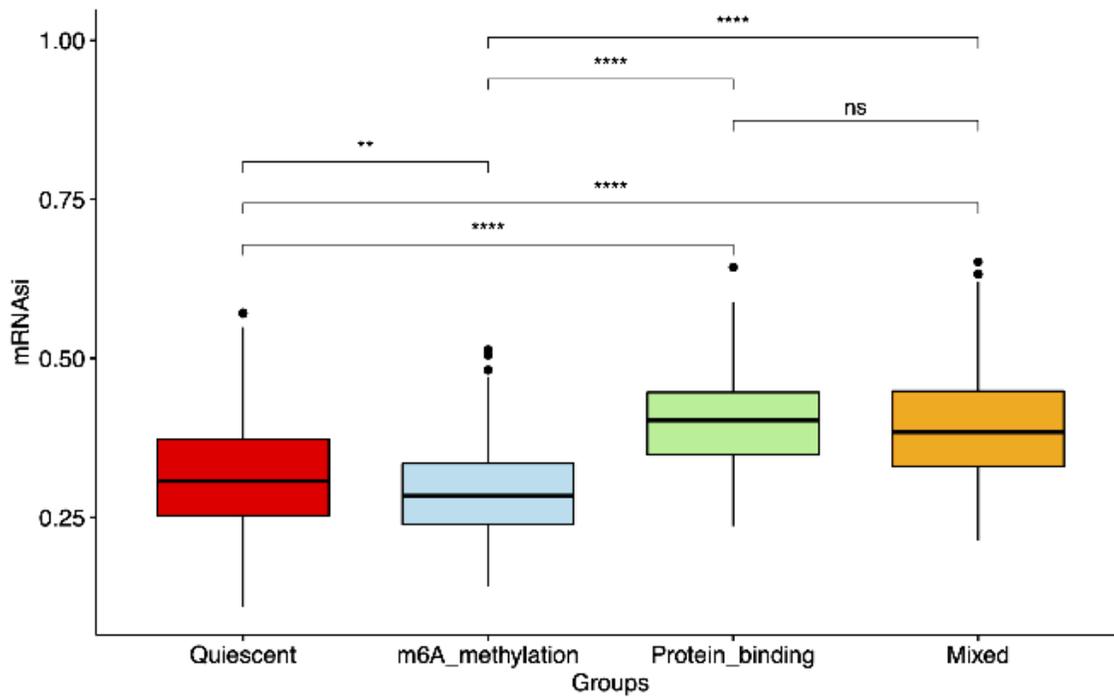


Figure 6

Differences in the tumor dry index between the four m6A subtypes

Supplementary Files

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

- [FigureS1.jpg](#)