m^6 A Regulator-Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Early Cervical Cancer

Dan Sun
Third Xiangya Hospital of Central South University
Xingping Zhao
Third Xiangya Hospital of Central South University
Aiqian Zhang
Third Xiangya Hospital of Central South University
Lingxiao Zou
Third Xiangya Hospital of Central South University
Huan Huang
Third Xiangya Hospital of Central South University
Jiangtao Fan
The First Affiliated Hospital of Guangxi Medical University
Dabao Xu (dabaoxu@yahoo.com)
Third Xiangya Hospital of Central South University

Research Article

Keywords: m6A, tumor microenvironment, immune infiltration, early cervical cancer

Posted Date: November 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1037825/v1

Abstract

Background

Cervical cancer (CC) is the malignancy of female and almost cases of cervical cancer were caused by high-risk human papillomavirus (HPV) infection. Understanding the pathogenesis and characteristics of the postinfection microenvironment (PIM) in early-stage cervical cancer is needed. The mechanism of N6-methyladenosine (m^6A) in the regulation of immune microenvironment in cervical cancer is unclear.

Methods

Messenger RNA (mRNA) expression profiles and clinical information of cervical cancer were downloaded from the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) dataset GSE44001. We comprehensively evaluated the m^6A modification patterns in early cervical cancer and systematically correlated these modification patterns with tumor microenvironment (TME) cell-infiltrating characteristics. Analysis of tumor mutational signatures and biological enrichment analysis were also conducted.

Results

LRPPRC had the highest mutation frequency. Writers METTL14 and ZC3H13, as well as reader YTHDF3, were prognostic risk-related genes. DEGs were significantly enriched in the C-type lectin receptor signaling pathway. The m^6A cluster A showed a higher level of infiltration immunocytes, and the activity of most immune cells increased. The low-m^6Sig score group was poor in prognosis compared with the high-m^6Sig score group. Further, we found that the 23 immunocytes, excluding plasmacytoid dendritic cells, negatively correlated with the m^6Sig score.

Conclusions

Dysregulation of m^6A lays a critical foundation for understanding the regulation of early CC immunity. What's more, evaluating the m^6A modification pattern of early CC contributes to enhancing our knowledge of the characteristics of PIM and provides an important insight into the efficacy of HPV treatment.

Introduction

Cervical cancer (CC) is the fourth most common female malignant tumor in the world [1]. High-risk human papillomavirus (HPV) infection was the most important cause of cervical cancer [2, 3]. About 31% of those patients suffer from tumor recurrence after treatment [4]. Treatments for recurrent and advanced cervical cancer are limited, and the 5-year survival rate is less than 20% [2]. Therefore, effective diagnosis

and treatment in early stage of cervical cancer are necessary, and it is important to understand the etiology and features of the immune microenvironment in early-stage cervical cancer.

The tumor microenvironment (TME) contains stromal cells, fibroblasts, and endothelial cells, innate and adaptive immune cells that is an area of complex and evolving processes [5–7]. It is recommended that the mere occurrence of HPV infection is not sufficient for malignant progression and postinfection microenvironment (PIM) is required, because PIM is becoming regarded as a key factor in viral persistence, propagation, and malignant progression [8]. The PIM is initiated and established by a complex interplay among virus-infected cells, immune cells, and host stroma, as well as chemokines, cytokines, extracellular vesicles, and metabolites [9, 10].

N6-methyladenosine (m^6A) is one of the most abundant methylation mechanisms in eukaryotic messenger RNA (mRNA) that contains 3 forms called "writers","readers", and "erasers" [11]. There are reports of m^6A modification abnormalities involved in multiple diseases, including cervical cancer [12–14]. Several recent studies have found that m^6A is closely related to the malignant progression of cervical cancer [15–17]. However, the role of m^6A in the regulation of immune microenvironment of cervical cancer is unknown.

This study made comprehensive evaluation on the relationship between m^6A modification patterns and TME cell-infiltrating characteristics of early cervical cancer by integrating transcriptome and genome data obtained from the public databases. Distinct m^6A modification patterns were identified using the "ConsensusClusterPlus" package in R. The landscape, survival outcomes, and immunocyte infiltration of these distinct patterns was analyzed. Moreover, we constructed a scoring scheme to quantify the m^6A modification patterns of early cervical cancer and compared the immunocyte infiltration and m^6Sig score. Biological enrichment analysis for distinct m^6A modification patterns was also conducted. Our findings suggested that the role of m^6A modification in shaping early cervical cancer TME characterizations was not negligible.

Methods

Data selection and processing

Gene expression profiles, clinical information, and single nucleotide polymorphisms (SNPs) for cervical cancer were downloaded from TCGA (TCGA-CC, https://portal.gov/). Copy number variation (CNV) data for TCGA-CC were obtained from the University of California Santa Cruz (UCSC) Xena database (https://xena.edu/). Expression matrices of GSE44001 datasets and corresponding clinical data were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), as detected by Affymetrix Human Genome U133 Plus 2.0 Array platform [18]. TCGA RNA sequence data (FPKM format) were transformed into transcripts per kilobase million (TPM) format by the "limma" package of R. The data used in this study consisted of 3 healthy samples and 306 tumor samples. Patients without survival

information were excluded. A total of 23 m^{6}A regulators included 8 writers (METTL3, METTL14, METTL16, WTAP, VIRMA, ZC3H13, RBM15, and RBM15B), 2 erasers (FTO and ALKBH5), and 13 readers (YTHDC1, YTHDC2, YTHDF1, YTHDF2, YTHDF3, HNRNPC, FMR1, LRPPRC, HNRNPA2B1, IGFBP1, IGFBP2, IGFBP3, and RBMX).

Identification of DEGs in m^{6}A patterns

The gene expression of 23 m^{6}A regulators in TCGA-CC was extracted by R software (version x644.1.0). Analysis of differentially expressed genes (DEGs) was applied to assess the differential expression levels of m^{6}A regulators between tumor and normal tissues using the "limma" package in R [19]. A box plot of DEGs was drawn using the R package "ggpubr".

Estimation of TMB and CNV in TCGA-CC

Tumor mutation load (TMB) was evaluated by the total number of somatic mutations / total covered bases [20]. The SNP dataset was based on VarScan2 variant aggregation and masking data in TCGA. The mutational landscape of 23 m^{6}A modification genes and top 10 significantly mutated genes in low- and high-m^{6}A score groups were depicted by the waterfall function of the R package "'maftools" [21]. We constructed a model based on TMB status and classified cervical cancer patients into high- and low-TMB groups based on the mean TMB. We then analyzed whether there was a difference in TMB between the high- and low-m6A score subtypes. Correlation between TMB and m^{6}A score in the 3 m^{6}A clusters was also analyzed. The R package "RCircos"' was employed to plot the CNV landscape of the 23 m^{6}A regulators in human chromosomes.

Consensus clustering of m^{6}A regulators

To identify m^{6}A regulator-mediated subtypes, unsupervised consensus clustering was performed to cluster tumor samples into 3 subgroups based on the expression matrix of m6A regulators using the "ConsensusClusterPlus" package in R [22]. The following parameters were used for clustering: number of repetitions =1000 bootstraps, pItem =0.8 (resampling 80% of any sample), pFeature =1 (resampling 100% of any protein), and clustering algorithm = k-means method. Graphical output results included heatmaps of the consensus matrices that can determine the approximate number of clusters. The number of clusters was confirmed by the following conditons: relatively high consistency in clusters, relatively low variation coefficient, and no appreciable increase in the area under the cumulative distribution function curve.

Identification of DEGs among distinct m^{6}A modification phenotypes

The consensus clustering algorithm described the above-classified patients into 3 distinct m^{6}A modification patterns, and we next determined m^{6}A modification-related DEGs among distinct m^{6}A phenotypes. The R package "limma" evaluated DEGs in cervical cancer samples among different

modification clusters (Table 1). A Venn diagram showed the overlapping m^6A regulator-mediated genes. The importance of filtering DEGs was set as an adjusted P-value less than 0.05.

Table 1
23 m^6A regulators associated with OS of patients with early cervical cancer

ID	HR	HR.95L	HR.95H	P value	KM
METTL3	1.049565	0.772363	1.426254	0.757194	0.102246
METTL14	1.635832	1.121101	2.386891	0.010682	0.000686
METTL16	0.962178	0.771383	1.200165	0.732417	0.008791
WTAP	1.086629	0.715734	1.649721	0.696540	0.138942
VIRMA	1.174593	0.844573	1.633571	0.338968	0.125269
ZC3H13	1.567539	1.073304	2.289360	0.020017	0.000060
RBM15	1.182565	0.644490	2.169870	0.588188	0.132704
RBM15B	1.109463	0.866482	1.420580	0.410150	0.034250
YTHDC1	1.032603	0.664827	1.603831	0.886439	0.191452
YTHDC2	0.971446	0.714847	1.320153	0.853134	0.055853
YTHDF1	1.065500	0.739204	1.535828	0.733782	0.045496
YTHDF2	1.292234	0.945673	1.765801	0.107546	0.011599
YTHDF3	1.477872	1.007076	2.168759	0.045935	0.000051
HNRNPC	1.182880	0.838407	1.668885	0.338893	0.000628
FMR1	1.022565	0.715608	1.461191	0.902479	0.024925
LRPPRC	1.193722	0.939965	1.515985	0.146441	0.075837
HNRNPA2B1	0.946153	0.670267	1.335596	0.752990	0.025420
IGFBP1	1.149944	0.833625	1.586291	0.394634	0.036067
IGFBP2	0.886909	0.784531	1.002648	0.055147	0.000969
IGFBP3	1.051132	0.924181	1.195521	0.447648	0.159252
RBMX	1.202304	0.692712	2.086776	0.512528	0.077702
FTO	1.244355	0.953931	1.623198	0.106927	0.000100
ALKBH5	0.902680	0.580803	1.402939	0.649044	0.110021

Estimation of TME cell infiltration

In order to investigate the immune infiltration levels in 23 immune cell subtypes in multiple m^{6}A clauses, the enrichment scores were calculated by single-sample gene set enrichment analysis (ssGSEA) and normalized to a unified distribution from 0 to 1 [23]. The correlation of m^{6}Sig scores with immunocyte fractions was determined by R package "corrplot".

Generation of m^{6}A gene signature

After identifying DEGs in different m^{6}A clusters, we performed prognostic analysis for each DEG in the signature using the R "survival" package. Genes with prognostic significance (P <0.05) were extracted for further analysis. Next, principal component analysis (PCA) was carried out to construct the m^{6}A-related gene signature. Both principal components 1 and 2 were selected to act as signature scores. We defined the m^{6}A score using the following method [24, 25]:

$$m^6A\ score=\sum(PC1i+PC2i)$$

where i is the expression of m^{6}A phenotype-related genes. The cases were divided into high- and low-m^{6}Sig score groups based on a cut-off of 0.3120783.

GSVA analysis for distinct m^{6}A modification patterns

The pathway activation scores between two m^{6}A clusters were compared by the R package "limma" [23]. The Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were carried out using R package "clusterprofiler" and "enrichplot" by applying a threshold of P <0.05. A bar chart and bubble diagram were used to visualize GO enrichment's biological processes, cellular components, and molecular functions.

Statistical analysis

Statistical analyses in this study were generated by R (version 4.1.0). The R package "survival" and "survminer" were used to estimate survival curves. A univariate Cox regression model was performed to calculate the hazard ratio (HR) for each m^{6}A regulator. For all analyses, a P value < 0.05 was regarded as statistical significance.

Results

The landscape of m^{6}A regulators in healthy and cervical cancer samples

In this study, we investigated the role of 23 m^{6}A RNA methylation regulatory genes in cervical cancer. As shown in Figure 1B, the expression level of 23 m^{6}A regulators varied greatly between two groups. The expression levels of METTL16 and FTO were significantly decreased, while expression levels of RBM15, YTHDF2, and HNRNPA2B1 were significantly increased in tumor tissues (P <0.05). Further, we determined

the prevalence of somatic mutations of 23 m^6A regulators in cervical cancer. A total of 37 of the 289 (12.8%) samples had genetic alterations of m^6A regulators, including 6 types of mutations which are depicted in different colors (Figure 1A). LRPPRC showed the highest mutation frequency, followed by ZC3H13, YTHDC2, and RBM15, while VIRMA, RNPA2B1, and GFBP1/2/3 did not show any mutations in cervical cancer samples (Fig. 1A). Next, the 4 highest mutated genes containing 2 readers (LRPPRC and YTHDC2) and 2 writers (ZC3H13 and RBM15) were divided into a wild-type group and mutation group. The gene expression of other m6A regulators was calculated for these 2 groups. We found that tumors with mutations of YTHDC2 showed a low expression of writer gene RBM15 (Fig. 1E), and reader genes IGFBP2, RBMX, and YTHDF1 were significantly downregulated in ZC3H13-mutant as well as RBM15-mutant tumors compared to wild-type tumors (Fig. 1F-H). Other m^6A regulators had no expression differences in the wild-type and mutation groups ($P > 0.05$).

Further analysis of 23 m^6A regulators revealed that CNV mutations were common. FMR1, RBMX, HNRNPC, METTL3, YTHDC1, VIRMA, LRPPRC, and HNRNPA2B1 showed widespread CNV amplification. In contrast, YTHDF1/2/3, ALKBH5, METTL14/16, IGFBP1/2/3, WTAP, HNRNPA2B1, YTHDC2, RBM15/15B, ZC3H13, and FTO had frequent CNV deletions (Fig. 1D). Interestingly, YTHDF1 and YTHDF3 had CNV deletions only. The location of CNV alterations of 23 m^6A regulators on chromosomes is shown in Figure 1C. These results indicate significant differences between normal and cervical cancer samples in the genomic and transcriptomic landscape of m^6A regulators, indicating that expression alterations and genetic variation in m^6A regulators played a important role in regulating the malignant progression of cervical cancer.

Unsupervised clustering of 23 m^6A regulators in the early cervical cancer cohort

OS data and clinical information from 1 GEO dataset (GSE44001) and the TCGA-CC cohort were included in our study. To further explore the characteristics of the m^6A modification phenotypes, we selected m^6A regulators to identify subgroups of early cervical cancer samples. We utilized consensus clustering analysis to stratify samples with qualitatively different m^6A modification patterns based on the expression of 23 m^6A regulators. The coefficient of variation among clusters was calculated according to the number of categories, and K = 3 for early cervical cancer was identified as the optimal choice (Fig. 2B-C). We named these 3 clusters as m^6A clusters A, B, and C, respectively, and included 208 cases in pattern cluster A, 195 cases in cluster B, and 201 cases in cluster C (Fig. 2A). However, Kaplan-Meier survival analysis showed no significant difference in prognosis among the 3 clusters ($P = 0.067$, Fig. 2D).

m^6A methylation modification regulatory network and enrichment analysis

The comprehensive landscape of the interactions of the 23 m^6A regulators, the regulator connections, and their prognostic significance in early CC patients was shown in an m^6A regulator network (Fig. 3B). In

this network, positive relationships between m^{6}A regulators are indicated with pink lines, and the sole negative correlation between HNRNPA2B1 and METTL16 is shown with a blue line. Writers METTL14 and ZC3H13 and the reader YTHDF3 were prognostic risk-related genes, while no relationship with prognosis was found among the rest of the regulators. The results revealed that cross-talk among the regulators of writers and readers may play a critical role in the formation of different m^{6}A modification patterns.

We further investigated the potential m^{6}A-related transcriptional expression change across 3 m^{6}A modification patterns of early cervical cancer to estimate the underlying genetic alterations and expression perturbations within these phenotypes. A total of 115 DEGs representing the significant distinguishing index of the 3 m^{6}A modification patterns were regarded as an m^{6}A-related signature and depicted in the Venn diagram (Fig. 3A). We then compared the GO analyses and KEGG pathways between each of them to investigate biological responses in the 3 m^{6}A modification patterns. The activation state of biological pathways was evaluated by GSVA enrichment analysis.

Biological functions of the m^{6}A regulator were classified as biological processes, cellular components, or molecular functions (Fig. 3C-D). GO analysis showed that DEGs were mainly enriched in the biological processes associated with the regulation of protein serine / threonine kinase activity, cell leading edge, and SH3 domain binding. Furthermore, the analysis using "clusterProfiler" showed that these genes were significantly enriched in the C-type lectin receptor signaling pathway, an immune pathway associated with antigen presentation (Fig. 3E-F). These findings suggested that m^{6}A modifications had a role in immune regulation in the TME.

m^{6}A phenotype-related DEGs in early cervical cancer

Based on the 115 most representative m^{6}A phenotype-related signature genes, unsupervised consensus clustering analysis was performed to obtain 4 stable transcriptomic phenotypes. Then the patients were divided into 4 distinct m^{6}A gene signature subgroups with different clinicopathologic features and were defined as m^{6}A gene clusters A, B, C, and D (Fig. 4B-C). The heatmap indicated that downregulated m^{6}A regulators were mainly concentrated in m^{6}A gene cluster A while most m^{6}A regulators in cluster D were upregulated. There was no significant difference in the expression of m^{6}A regulators in clusters B or C (Fig. 4A). Further survival analysis showed a significant prognostic difference among the 4 m^{6}A gene signatures in early cervical cancer patients ($P < 0.001$). The signature of m^{6}A gene cluster D was related to the best prognosis, while the m^{6}A gene cluster A was associated with the worst survival outcomes (Fig. 4D).

Expression levels of the 23 m^{6}A regulators were also compared among the 4 gene signature subgroups (Fig. 4E). We observed significant differences in the expression of 17 m^{6}A regulators among the 4 m^{6}A gene signature subgroups (Fig. 4E). Thirteen prognostic-related m^{6}A regulators were extracted using the Kaplan-Meier method ($P < 0.05$, Fig. 5A-M). Low expression of FMR1, HNRNPA2B1, IGFBP2, and METTL16

was related to a worse prognosis. High expression of FTO, HNRNPC, IGFBP1, METTL14, RBM15B, ZC3H13, and YTHDF1/2/3 also showed worse OS outcomes.

Construction of the m^{6}Sig score

Although our study identified the role of m^{6}A modification in prognosis, the findings were based on a patient population and did not accurately predict the patterns of m^{6}A methylation modification in individual tumors. Therefore, we established m^{6}Sig, a scoring system to assess such impacts systematically. Based on an optimal cut-off value of 0.3120783, early cervical cancer patients were separated into a high- or low-m^{6}Sig score group. Considering the complexity of m^{6}A modification quantification, we depicted the workflow of m^{6}Sig score construction with an alluvial diagram (Fig. 6C). Most patients in m^{6}A gene cluster D were in the high-m^{6}Sig score group, whereas most patients in m^{6}A gene cluster A were in the low-m^{6}Sig score group. Survival analysis indicated a significant prognostic difference between the 2 groups (P =0.002, Fig. 6D). The low-m^{6}Sig score group was associated with a worse prognosis, while the high-m^{6}Sig score group had better survival outcomes. We further evaluated the relationship between known biological signatures and the m^{6}Sig score. Significant differences were found among the m^{6}A gene cluster groups in the m^{6}Sig score (P <0.001, Fig. 6B). Notably, m^{6}A gene cluster D showed the highest m^{6}Sig score, followed by m^{6}A gene clusters B, C, and A. There was also a significant difference in m^{6}Sig score between m^{6}A clusters A and B and between A and C (P <0.001), but there was no differences between m^{6}A clusters B and C (P =0.93, Fig. 6A). Figure 6A shows that m^{6}A cluster A had the lowest m^{6}Sig score.

Immune microenvironment characteristics in distinct m^{6}A modification patterns

The immune infiltration of 23 immunocytes was evaluated to detect differences in immune microenvironment characteristics among the distinct m^{6}A modification patterns. We found that 15 types of immune cells presented differential expression profiles among the 3 unsupervised clusters (Fig. 6E). The results showed that compared with m^{6}A clusters B and C, cluster A had a higher level of infiltrated activated CD4 T cells, activated CD8 T cells, activated dendritic cells, gamma delta T cells, immature B cells, MDSC, macrophages, monocytes, natural killer T cells, neutrophils, plasmacytoid dendritic cells, regulatory T cells, T follicular helper cells, and type I T helper cells (P <0.05, Fig. 6E). These results revealed that the activity of most immunocytes increased in cluster A. Further, we analyzed the correlation between the m^{6}Sig score and 23 immunocytes. We found that except for plasmacytoid dendritic cells, the immunocytes all had a negative relationship with the m^{6}Sig score (Fig. 6F). These results strongly suggested that the unsupervised clustering of 23 m^{6}A regulators was significantly correlated with TME.

Significantly mutated genes and tumor mutational signatures

Then, we analyzed the difference in the distribution of somatic mutations between the low- and high-m^6Sig score groups in early cervical cancer using the "maftools" package in R. The 20 genes with the highest mutation rates and their mutation classification of early cervical cancer were summarized in a waterfall map. As shown in figures 7A and 7B, the top 5 most frequently mutated genes were TTN (29%), PIK3CA (24%), KMT2C (23%), MUC4 (20%), and MUC16 (17%) in the low-m^6Sig score group, and PIK3CA (31%), TTN (27%), MUC16 (14%), MUC4 (13%), and FBXW7 (13%) in the high-m^6Sig score group. Interestingly, mutations of the TTN, PIK3CA, MUC4, and MUC16 genes were more common in both subgroups. The missense mutation was the highest mutation in the 7 variant classifications. The box plot in Figure 7D shows that no difference between the low- and high-m^6Sig score groups was found. m^6Sig score and TMB exhibited a significant negative correlation (Fig. 7E). However, Kaplan-Meier analysis indicated that there was no significant difference in OS between the high- and low-TMB groups (P =0.165, Fig. 7C), suggesting that TMB may not be a risk-independent prognostic factor in early cervical cancer. The results provided a new perspective to explore the mechanisms of m^6A methylation modification in tumor somatic mutations and the shaping of the TME landscape.

Discussion

m^6A is the most common internal modification of long noncoding RNAs (lncRNAs), transcripts and mammalian mRNAs [26–27]. The dysregulation of m^6A is closely related to the carcinogenesis and pathogenesis [28–32]. What's more, m^6A mRNA is important in the development and function of immune cells and other stromal cells within the TME [33]. However, it is unclear how m^6A mRNA regulates the TME in cervical cancer pathogenesis. Here, the interaction between m^6A and TME characteristics in early cervical cancer was analyzed.

We found that DEGs of distinct m^6A clusters were significantly enriched in the C-type lectin receptor signaling pathway. C-type lectins have been implicated in a diverse range of physiological functions, including antigen presentation in preventing invading pathogen viruses [34, 35]. The C-type lectin receptor (CLR) family contains a large group of pattern recognition receptors (PRRS) present on dendritic cells and Langerhans cells that form an immune response [36]. HPVs have employed various strategies to interfere with multiple signaling pathways in virus-infected cervical keratinocytes and evade surveillance from immune sentinels [37]. Subsequently, reeducated keratinocytes communicate with the local components via their regulated molecular factors and direct cell-cell interaction. These components work jointly to promote the initiation and establishment of PIM [38]. However, the function of the C-type lectin receptor signaling pathway in PIM is unclear.

The results showed that m^6A cluster A had a higher level of several infiltrated immune cells than clusters B and C, and its low-m^6Sig score indicated worse survival outcomes. However, some research has found PIM to be an immunosuppressive environment because dampened immune responses in the HPV-infected keratinocytes may affect the capacity to alert local immune cells or lower immune cell trafficking, thereby preventing infected cells from being recognized and eliminated [39, 40]. We speculated

that the immune cells were not functioning normally. Previous studies have illustrated that tumors with an immune-excluded phenotype contained abundant immune cells, but these immune cells were retained in the stroma surrounding tumor cell nests rather than penetrating their parenchyma [41]. The mutual relationship between activated C-type lectin receptor signaling pathway and immunosuppressive PIM needs to be studied.

Leucine rich pentatricopeptide repeat (PPR)-containing (LRPPRC) gene was firstly identified in 1994 [42]. LRPPRC belongs to the PPR motif-containing proteins family. PPR proteins bind to RNA and regulate transcription, RNA processing, splicing, stability, editing, and translation [43]. Some studies have demonstrated that LRPPRC expression increases in various cancer tissues and cell lines [44–46], and LRPPRC level is positively related to tumor grade and metastasis [47]. It is reported that LRPPRC plays a crucial role in antiviral ability and provides a potential therapeutic target for treating viral diseases [48]. Interestingly, the most frequent mutation in early cervical cancer was LRPPRC, but there was no difference in the normal level and the expression level of tumor samples. It is unclear whether the mutation led to changes in protein structure and function. Further research is needed to investigate how the function of LRPPRC changes.

Although our study involved 23 recognized regulators of RNA methylation, the inclusion of newly identified regulators into the model is needed to optimize the accuracy of m^6A modification patterns. Further, m^6A modification patterns and m^6Sig scores were identified using retrospective datasets. Due to the limited data of the GEO dataset, advanced cervical cancer data could not be obtained, and so further analysis of other stages of cervical cancer is needed. In addition, the number of stage IA cases was small, and the analysis of m^6A and clinical parameters were insufficient. Therefore, further study needs to be conducted with a larger sample size.

Conclusion

This study showed that the dysregulation of m^6A forms an important foundation for understanding the regulation of early cervical cancer immunity. What's more, evaluating the m^6A modification patterns of early cervical cancer will contribute to enhancing our understanding of the characteristics of PIM and provide crucial insight into the efficacy of anti HPV treatment.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

DB Xu conceived and designed the study with JT Fan. D Sun drafted the manuscript and analyzed the data. XP Zhao and AQ Zhang were responsible for the image and article format. LX Zou and H Huang

reviewed the data. All authors contributed to the article and approved the submitted version.

Funding: This study is supported by the Natural Science Foundation of Hunan Provincial (Grant No. 2021JJ40956), the National Key Research and Development Program of China (no. 2018YFC1004800), the Hunan Provincial Clinical Medical Technology Innovation Guiding Project (no. 2020SK53605 and 2020SK53606) and the Natural Science Foundation of Hunan Provincial (no. 2021JJ40593).

Availability of data and materials

Gene expression profiles, clinical information, and single nucleotide polymorphisms (SNPs) for cervical cancer were downloaded from TCGA (TCGA-CC, https://portal.gov/). Copy number variation (CNV) data for TCGA-CC were obtained from the University of California Santa Cruz (UCSC) Xena database (https://xena.edu/). Expression matrices of GSE44001 datasets and corresponding clinical data were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), as detected by Affymetrix Human Genome U133 Plus 2.0 Array platform. All the data are publicly available.

Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Gynecology, Third Xiangya Hospital of Central South University, Changsha, 410013, China. [2]Department of Gynecology and Obstetrics, The First Affiliated Hospital of Guangxi Medical University, Nanning, 530021, China.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. 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. Cohen PA, Jhingran A, Oaknin A, et al. Lancet. **2019** Jan 12;393(10167):169-182.
3. Canfell K, Kim JJ, Brisson M, et al. Mortality impact of achieving WHO cervical cancer elimination targets: a comparative modelling analysis in 78 low-income and lower-middle-income countries. Lancet. **2020** Feb 22;395(10224):591-603.

4. Crosbie EJ, Einstein MH, Franceschi S, et al. Human papillomavirus and cervical cancer. Lancet. **2013** Sep 7;382(9895):889–99.
5. Kim J, Bae JS. Tumor-Associated Macrophages and Neutrophils in Tumor Microenvironment. Mediators Inflamm. 2016;2016:6058147.
6. Motta JM, Rumjanek VM. Sensitivity of Dendritic Cells to Microenvironment Signals. J Immunol Res. 2016;2016:4753607.
7. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. Cancer Res. **2019** Sep 15;79(18):4557–4566.
8. Smola S. Immunopathogenesis of HPV-Associated Cancers and Prospects for Immunotherapy. Viruses. **2017** Sep 12;9(9):254.
9. Lo Cigno I, Calati F, Albertini S, et al. Subversion of Host Innate Immunity by Human Papillomavirus Oncoproteins. Pathogens. **2020** Apr 17;9(4):292.
10. Cao M, Wang Y, Wang D, et al. Increased High-Risk Human Papillomavirus Viral Load Is Associated With Immunosuppressed Microenvironment and Predicts a Worse Long-Term Survival in Cervical Cancer Patients. Am J Clin Pathol. **2020** Mar 9;153(4):502–512.
11. Boccaletto P, Machnicka MA, Purta E, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. **2018** Jan 4;46(D1):D303-D307.
12. Ben-Haim MS, Moshitch-Moshkovitz S, Rechavi G. FTO: linking m^6A demethylation to adipogenesis. Cell Res. **2015** Jan;25(1):3–4.
13. Cui Q, Shi H, Ye P, et al. m^6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. Cell Rep. **2017** Mar 14;18(11):2622–2634.
14. Wang X, Li Z, Kong B, et al. Reduced m^6A mRNA methylation is correlated with the progression of human cervical cancer. Oncotarget. **2017** Oct 24;8(58):98918–98930.
15. Wang Q, Guo X, Li L, et al. N6-methyladenosine METTL3 promotes cervical cancer tumorigenesis and Warburg effect through YTHDF1/HK2 modification. Cell Death Dis. **2020** Oct 24;11(10):911.
16. Li Z, Peng Y, Li J, et al. N6-methyladenosine regulates glycolysis of cancer cells through PDK4. Nat Commun. **2020** May 22;11(1):2578.
17. Zou D, Dong L, Li C, et al. The m^6A eraser FTO facilitates proliferation and migration of human cervical cancer cells. Cancer Cell Int. **2019** Dec 2;19:321.
18. Lee YY, Kim TJ, Kim JY, et al. Genetic profiling to predict recurrence of early cervical cancer. Gynecol Oncol. 2013 Dec;131(3):650–4.
19. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. **2015** Apr 20;43(7):e47.
20. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. **2017** Jan 18;541(7637):321–330.
21. Mayakonda A, Lin DC, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. **2018** Nov;28(11):1747–1756.

22. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. **2010** Jun 15;26(12):1572–3.
23. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. **2013** Jan 16;14:7.
24. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst. **2006** Feb 15;98(4):262–72.
25. Zeng D, Li M, Zhou R, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. Cancer Immunol Res. 2019 May;7(5):737–750.
26. Ke S, Pandya-Jones A, Saito Y, et al. m^6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. **2017** May 15;31(10):990–1006.
27. Zhao BS, He C. Fate by RNA methylation: m^6A steers stem cell pluripotency. Genome Biol. 2015 Feb 22;16(1):43.
28. He C. Grand challenge commentary: RNA epigenetics? Nat Chem Biol. 2010 Dec;6(12):863–5.
29. Cantara WA, Crain PF, Rozenski J, et la. The RNA Modification Database, RNAMDB: 2011 update. Nucleic Acids Res. 2011 Jan;39(Database issue):D195-201.
30. Machnicka MA, Milanowska K, Osman Oglou O, et al. MODOMICS: a database of RNA modification pathways–2013 update. Nucleic Acids Res. 2013 Jan;41(Database issue):D262-7.
31. Rottman F, Shatkin AJ, Perry RP. Sequences containing methylated nucleotides at the 5' termini of messenger RNAs: possible implications for processing. Cell. 1974 Nov;3(3):197-9.
32. Wei CM, Gershowitz A, Moss B. Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. Cell. **1975** Apr;4(4):379-86.
33. Li M, Zha X, Wang S. The role of N6-methyladenosine mRNA in the tumor microenvironment. Biochim Biophys Acta Rev Cancer. 2021 Apr;1875(2):188522.
34. Zelensky AN, Gready JE. The C-type lectin-like domain superfamily. FEBS J. 2005 Dec;272(24):6179–217.
35. Brown GD, Willment JA, Whitehead L. C-type lectins in immunity and homeostasis. Nat Rev Immunol. 2018 Jun;18(6):374–389.
36. van den Berg LM, Gringhuis SI, Geijtenbeek TB. An evolutionary perspective on C-type lectins in infection and immunity. Ann N Y Acad Sci. 2012 Apr;1253:149–58.
37. Zhou C, Tuong ZK, Frazer IH. Papillomavirus Immune Evasion Strategies Target the Infected Cell and the Local Immune System. Front Oncol. 2019 Aug 2;9:682.
38. Yuan Y, Cai X, Shen F, et al. HPV post-infection microenvironment and cervical cancer. Cancer Lett. **2021** Jan 28;497:243–254.

39. Lin D, Kouzy R, Abi Jaoude J, et al. Microbiome factors in HPV-driven carcinogenesis and cancers. PLoS Pathog. **2020** Jun 4;16(6):e1008524.
40. Wang X, Huang X, Zhang Y. Involvement of Human Papillomaviruses in Cervical Cancer. Front Microbiol. **2018** Nov 28;9:2896.
41. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. Science. **2015** Apr 3;348(6230):74-80.
42. Ghiso NS, Lennon GG. lrp130 gene assigned to chromosome 2. In Vitro Cell Dev Biol Anim. 1994 Nov;30A(11):744.
43. Manna S. An overview of pentatricopeptide repeat proteins and their applications. Biochimie. 2015 Jun;113:93–9.
44. Jiang X, Zhong W, Huang H, et al. Autophagy defects suggested by low levels of autophagy activator MAP1S and high levels of autophagy inhibitor LRPPRC predict poor prognosis of prostate cancer patients. Mol Carcinog. 2015 Oct;54(10):1194–204.
45. Li X, Lv L, Zheng J, et al. The significance of LRPPRC overexpression in gastric cancer. Med Oncol. 2014 Feb;31(2):818.
46. Nishio T, Kurabe N, Goto-Inoue N, et al. Immunohistochemical expression analysis of leucine-rich PPR-motif-containing protein (LRPPRC), a candidate colorectal cancer biomarker identified by shotgun proteomics using iTRAQ. Clin Chim Acta. 2017 Aug;471:276–282.
47. Zhang HY, Ma YD, Zhang Y, et al. 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** May;70(5):383–389.
48. Refolo G, Ciccosanti F, Di Rienzo M, et al. Negative Regulation of Mitochondrial Antiviral Signaling Protein-Mediated Antiviral Signaling by the Mitochondrial Protein LRPPRC During Hepatitis C Virus Infection. Hepatology. 2019 Jan;69(1):34–50.

Figures

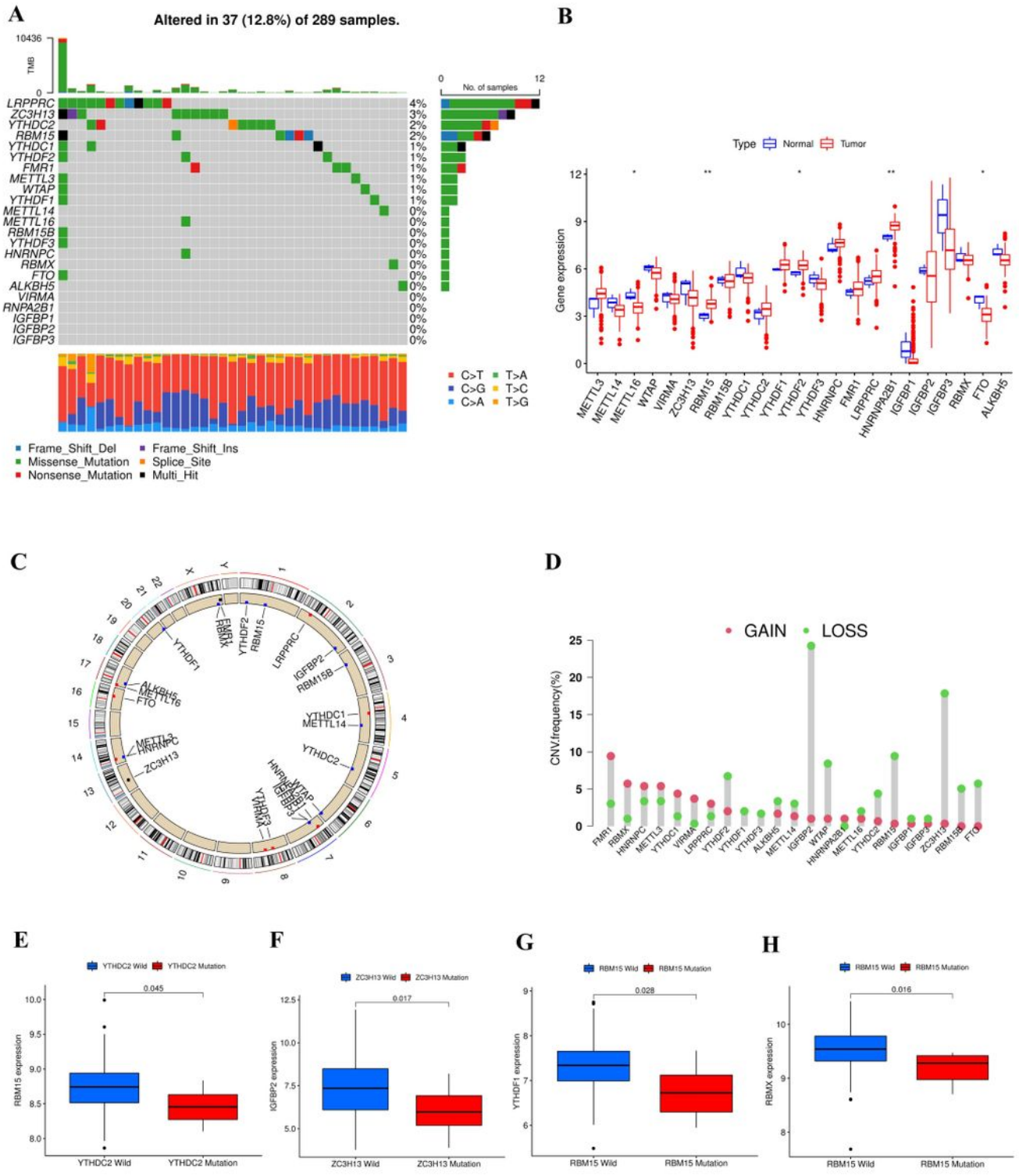

Figure 1

Landscape of genetic and expression variation of m6A regulators in early cervical cancer. (A) The mutation frequency of 23 m6A regulators in 289 patients with early cervical cancer from the TCGA-CC cohort and GSE44001 dataset. Each column represents an individual patient. The upper bar plot shows TMB. The number on the right indicates the mutation frequency in each regulator. The right bar plot shows the proportion of each variant type. The stacked bar plot below shows the percentage of

conversions in each sample. (B) The expression of 23 m6A regulators in two groups. normal, blue. The top and bottom of the boxes represent interquartile range of values. The lines in the boxes represent median value. *P <0.05,**P <0.01, and ***P <0.001. (C) The location of CNV alteration of m6A regulators on 23 chromosomes. (D) CNV mutation of the 23 m6A regulators was prevalent. The column represents the alteration frequency. Deletion frequency, green dot; amplification frequency, red dot. (E-H) Gene expression in wild-type and mutation groups of the 4 highest mutated genes. Wild-type group, blue box; mutation group, red box.

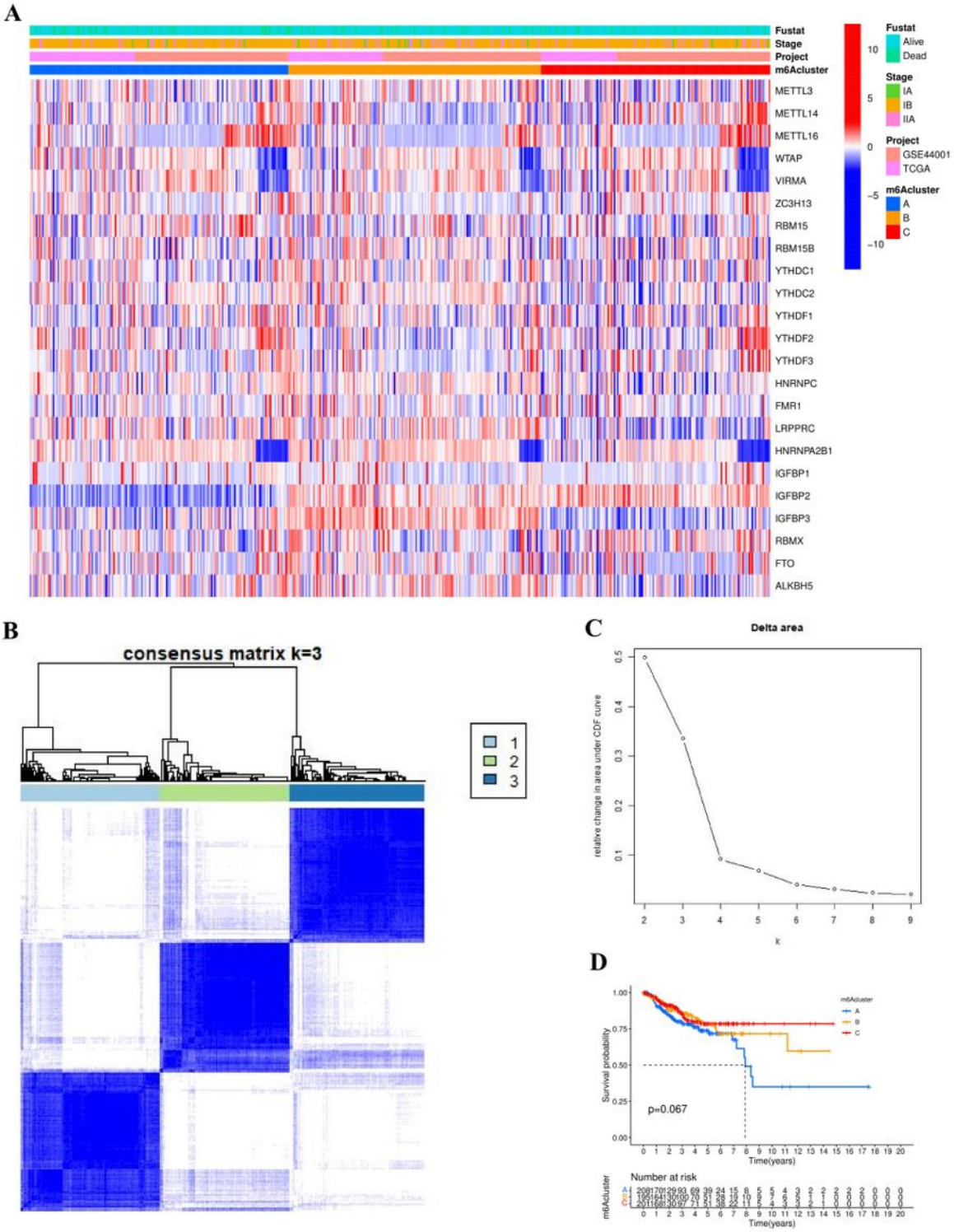

Figure 2

Identifying 3 distinct m6A modification pattern subtypes in early cervical cancer. (A) Unsupervised clustering of 23 m6A regulators in the 3 modification patterns. The heatmap also reveals the relationship of different clinical characteristics within the 3 subtypes. (B) Heatmap of the matrix of cooccurrence proportions for early cervical cancer samples. (C) Relative change in area under consensus clustering cumulative distribution function (CDF) curve for k =2-9. (D) Survival analyses of 3 distinct m6A modification patterns in early cervical cancer samples.

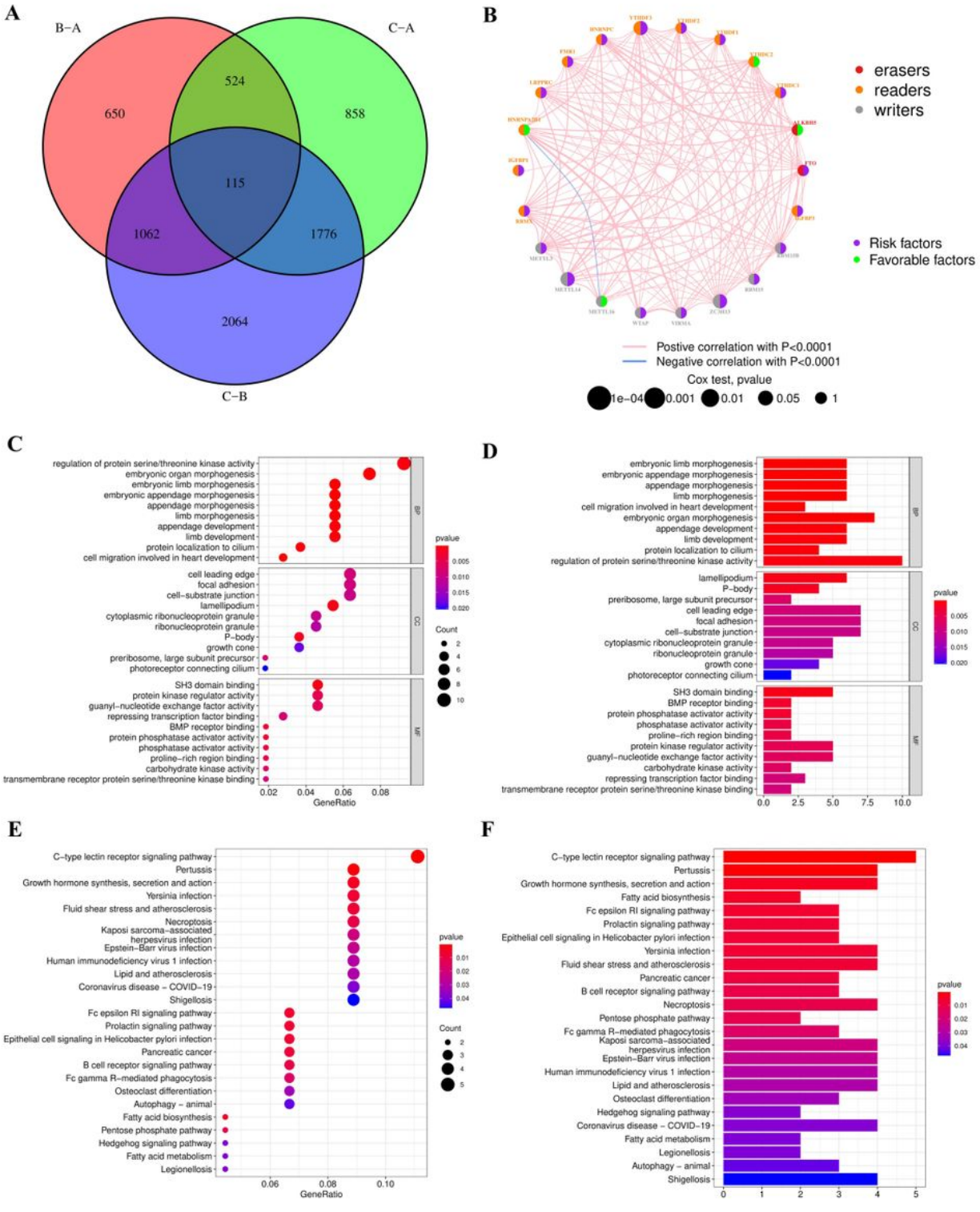

Figure 3

Regulatory network construction and enrichment analysis of m6A regulators. (A) 115 m6A-related DEGs among the 3 m6A clusters are depicted in Venn diagram. (B) The interaction of expression on 23 m6A regulators in early cervical cancer. 3 types of the m6A regulators are shown in different colors in the left half of the circles. erasers, red. The lines connecting m6A regulators represent their interaction with each other. The size of each circle represents the prognosis effect of each regulator, scaled by P value. Protective factors for patient survival are represented by green dots and risk factors by purple dots. (C-F) GO and KEGG analysis exploring the potential function underlying the 115 m6A-related genes in early cervical cancer. In the bubble plot, "Count" represents "number of enriched genes" and "GeneRatio" represents "number of enriched genes/number of total genes." The number of enriched genes was represented by the size of the dots, and the color means P value. The redder the color, the smaller the P value. The color depth of the bar plots represents the number of enriched genes. BP, biological process; CC, cellular component; MF, molecular function.

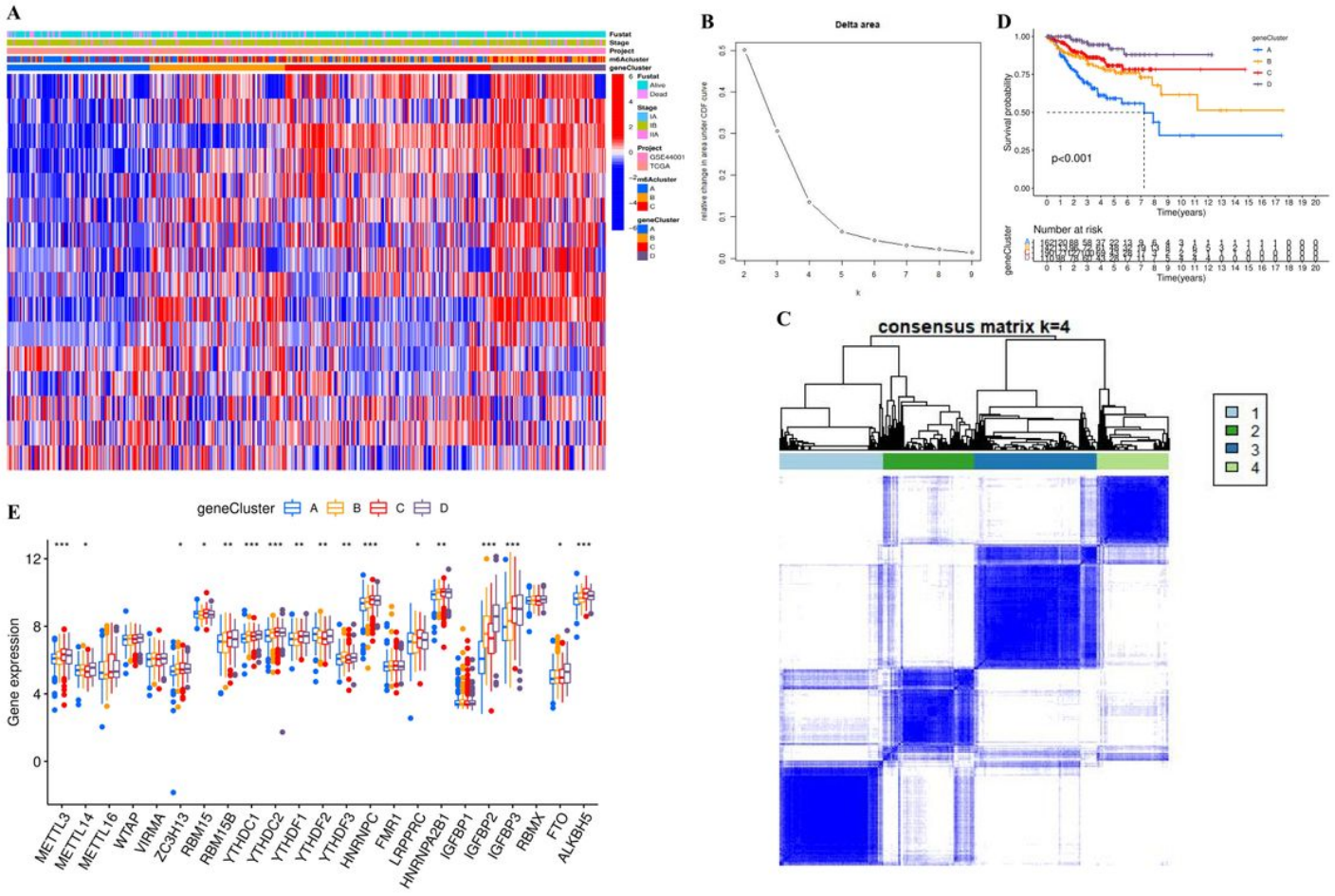

Figure 4

Unsupervised clustering of prognostic-related DEGs. (A) Unsupervised clustering of prognostic-related DEGs in the 4 modification patterns. (B) Relative change in area under consensus clustering CDF curve for k =2-9. (C) Consensus matrices of the 115 m6A prognostic-related genes for early cervical cancer (k

=4). (D) Survival analyses for 4 distinct m6A gene clusters in early cervical cancer samples. (E) The expression of 23 m6A regulators in A-D m6A gene clusters. cluster D, grey. The upper and lower ends of the boxes represent interquartile range of values. The lines in the boxes represent median value.*P <0.05,**P <0.01, and ***P <0.001.

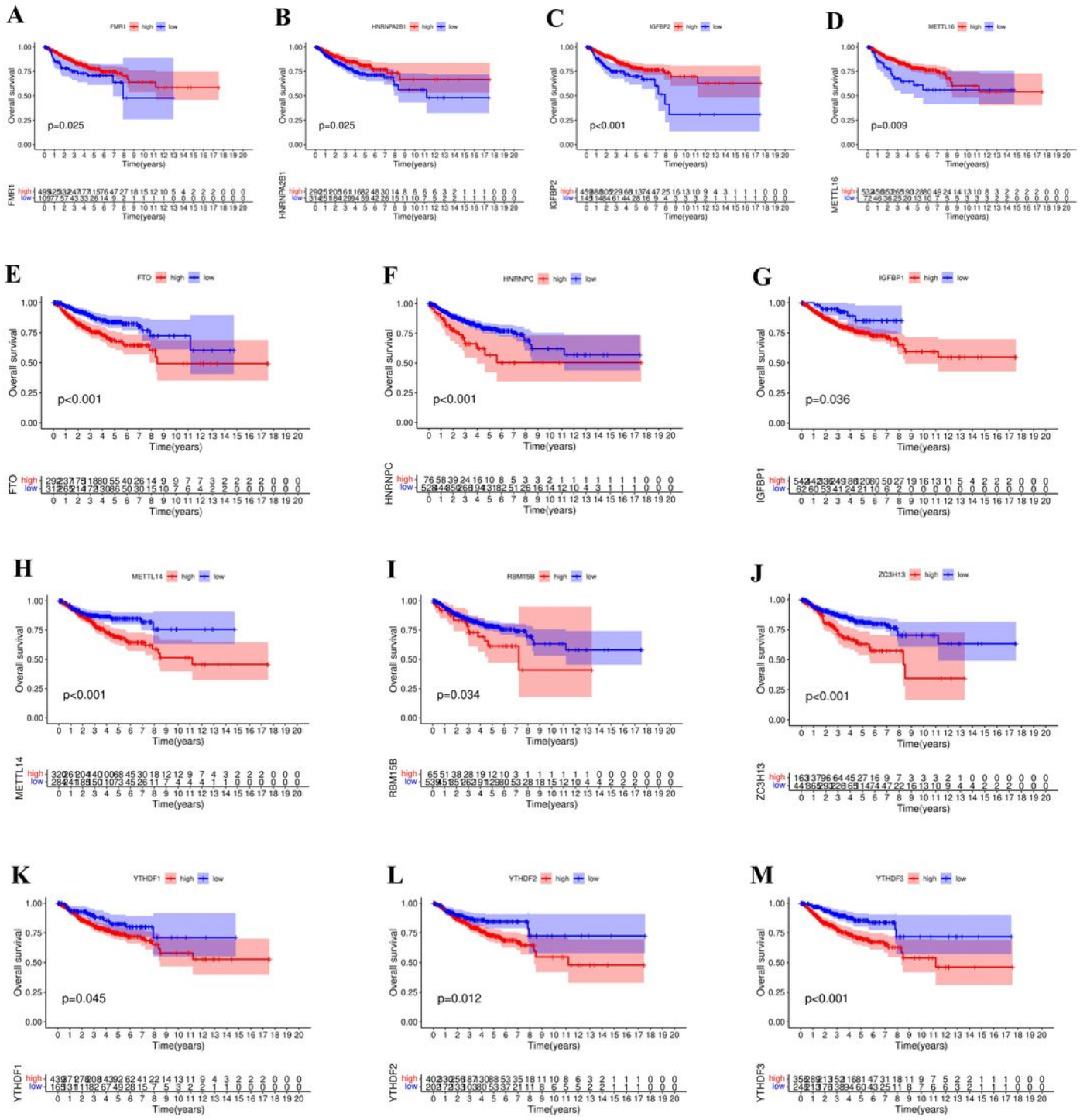

Figure 5

Kaplan-Meier analysis of 13 prognostic-related m6A regulators. Low expression, blue; high expression, red.

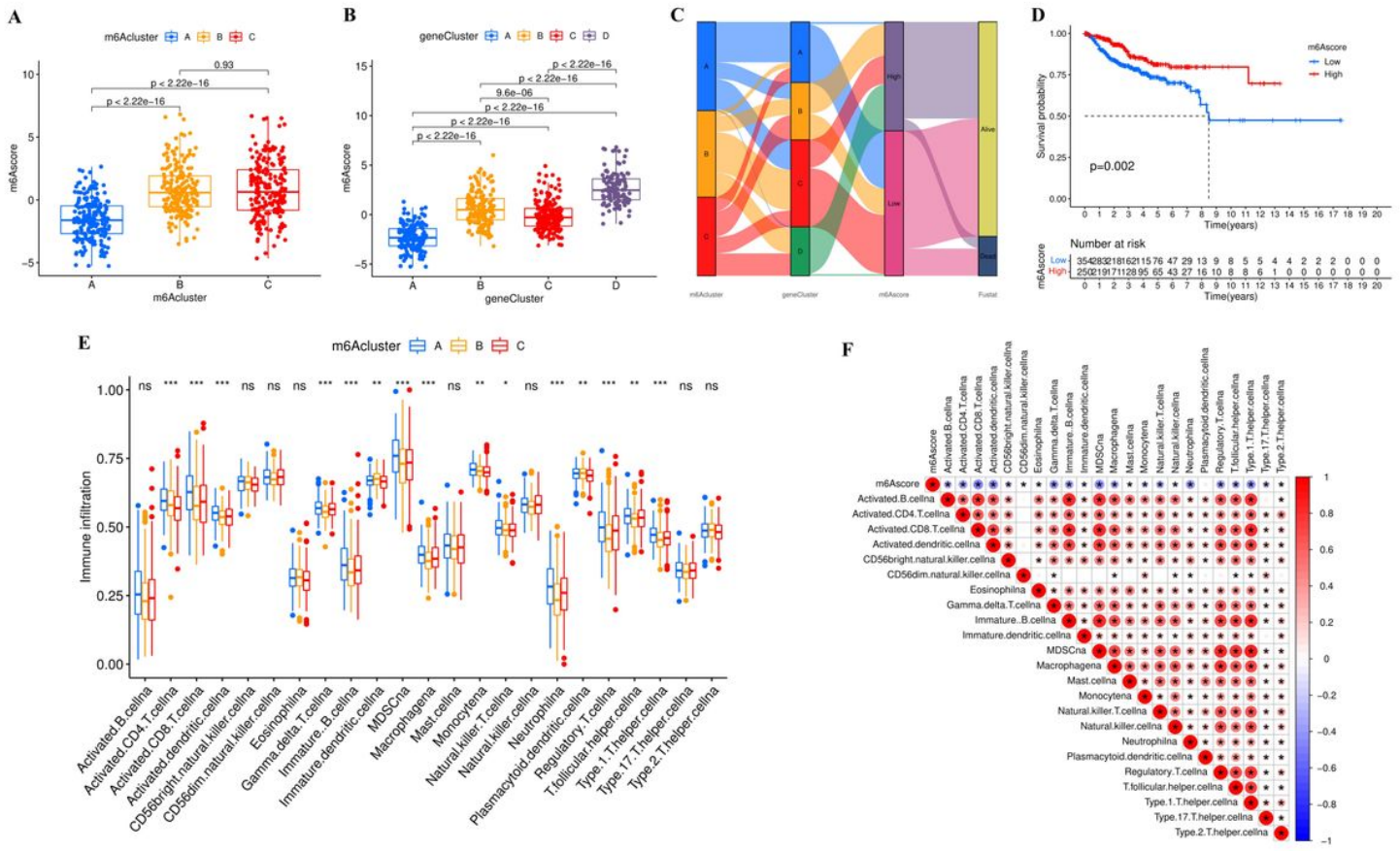

Figure 6

Construction of m6Sig score and exploration of TME characteristics in distinct m6A modification patterns. m6Sig score in different m6A cluster groups (A) and m6A gene cluster groups (B) Cluster A, blue; cluster D, grey. The top and bottom of the boxs represent interquartile range of values. The lines in the boxes represent median value. (C) Alluvial diagram of m6A clusters in groups with m6A gene cluster, m6Sig score, and survival outcomes. (D) Kaplan-Meier curves for patients in high- and low-m6Sig score subgroups. (E) The fraction of tumor-infiltrating immunocytes in 3 m6A clusters. In each group, TME cell expression values are depicted by the scattered dots. The transverse line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. Statistical differences in 3 gene clusters were compared using the R package "limma". *P <0.05,**P <0.01, and ***P <0.001. (F) The correlation of m6Sig score and 23 immunocytes. *Statistically significant; red, positive correlation; blue, negative correlation.

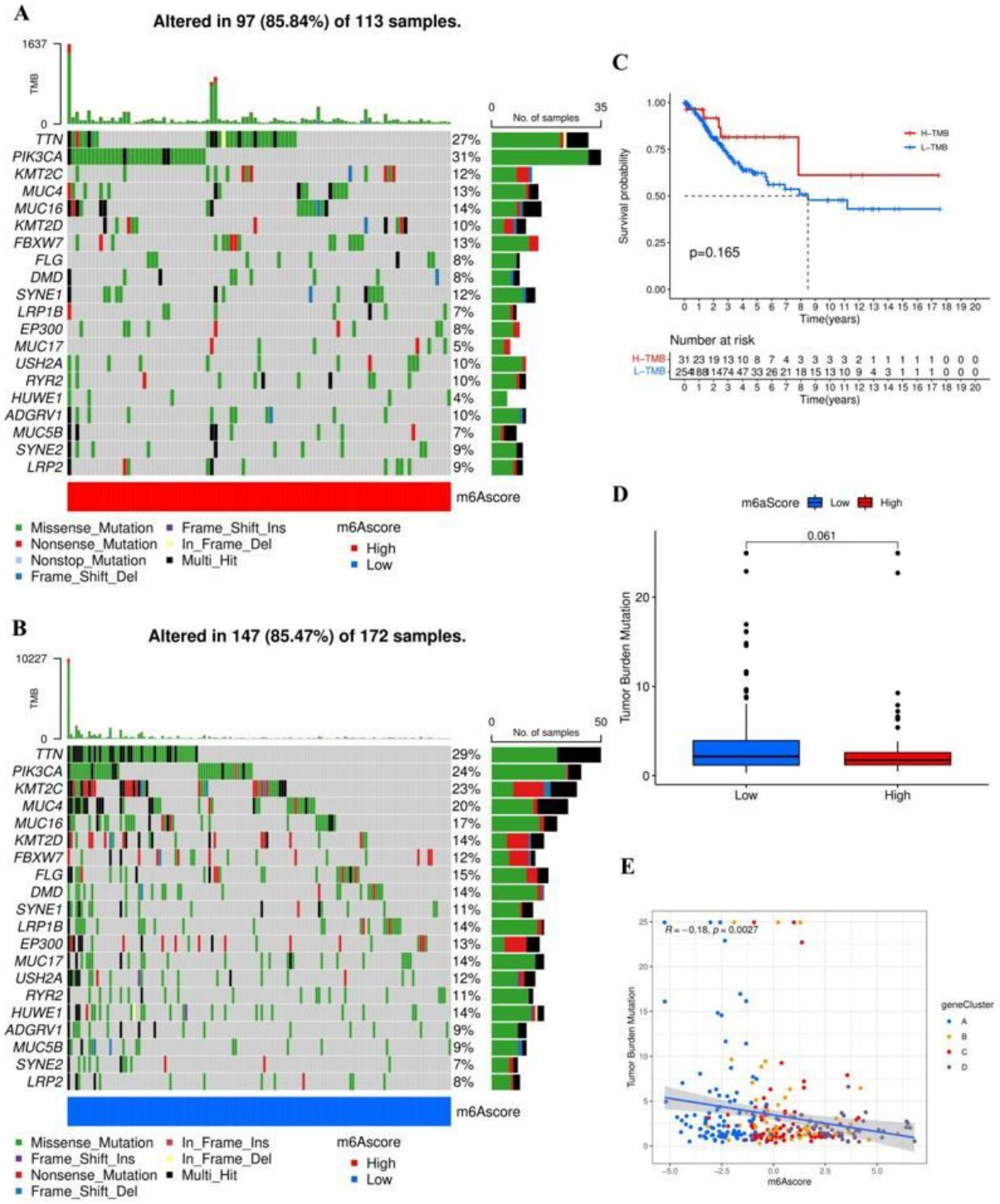

Figure 7

Characteristics of tumor somatic mutations in low- and high-m6Sig score groups. (A, B) Waterfall plot of tumor somatic mutations for the high-m6Sig score (A) and low-m6Sig score groups (B). Each column represents an individual patient. The upper bar plot shows TBM. The right number shows the mutation frequency of each gene. The right bar plot shows the percentage of each variant type. (C) Kaplan-Meier curves for patients in high- and low-TBM subgroups. (D) TBM in high- and low-m6Sig score groups. (E)

The correlation between TBM and m6Sig score in 4 m6A gene clusters. Cluster A, blue; cluster B, yellow; cluster C, red; cluster D, grey.