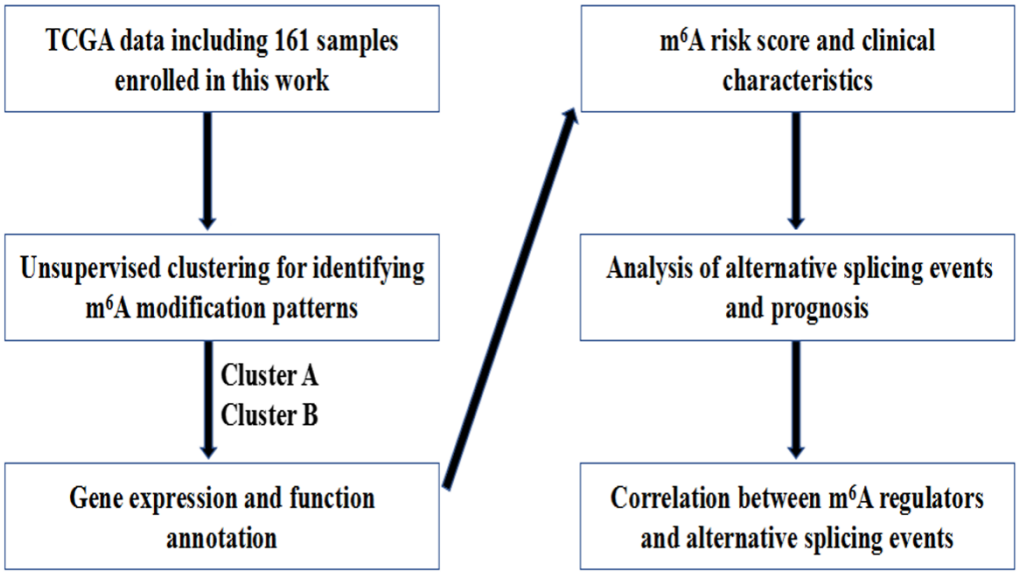
Supplementary Information for

Profiling the m^{6}A regulated RNA expression patterns and alternative splicing features in esophageal carcinoma

Supplementary Information

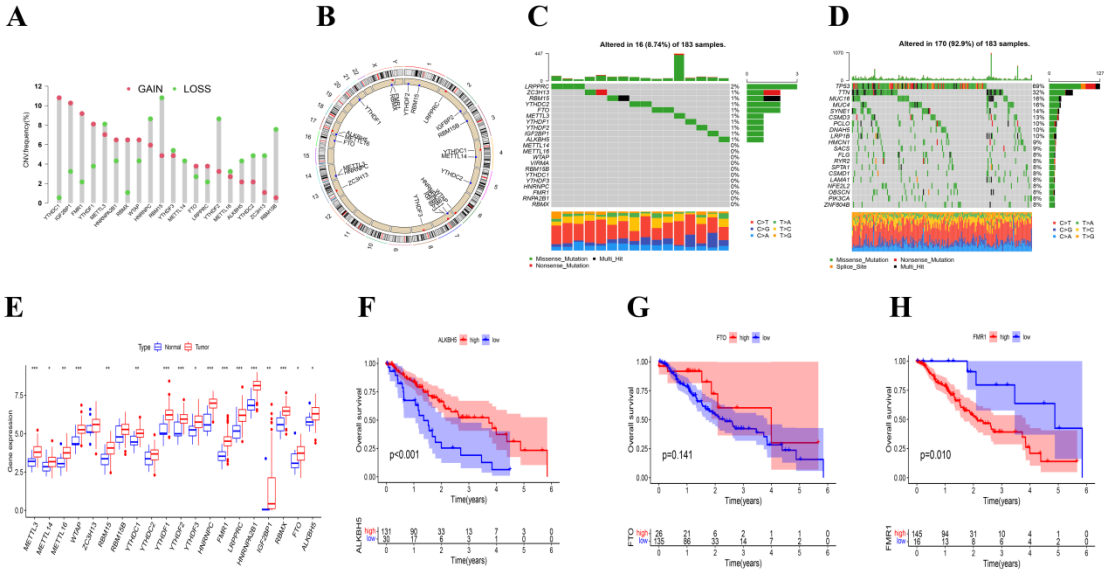
Supplementary Figures:

S1.



The workflow of this study.

S2.



Landscape of genetic and expression variation and prognosis of m^{6}A regulators in EC.

(**A**).The location of CNV alteration of m^{6}A regulators on 22 chromosomes using TCGA-ESCA cohort.

(**B**). The CNV variation frequency of m^{6}A regulators in TCGA-ESCA cohort. The height of the column represented the alteration frequency. The deletion frequency,

blue dot; the amplification frequency, red dot.

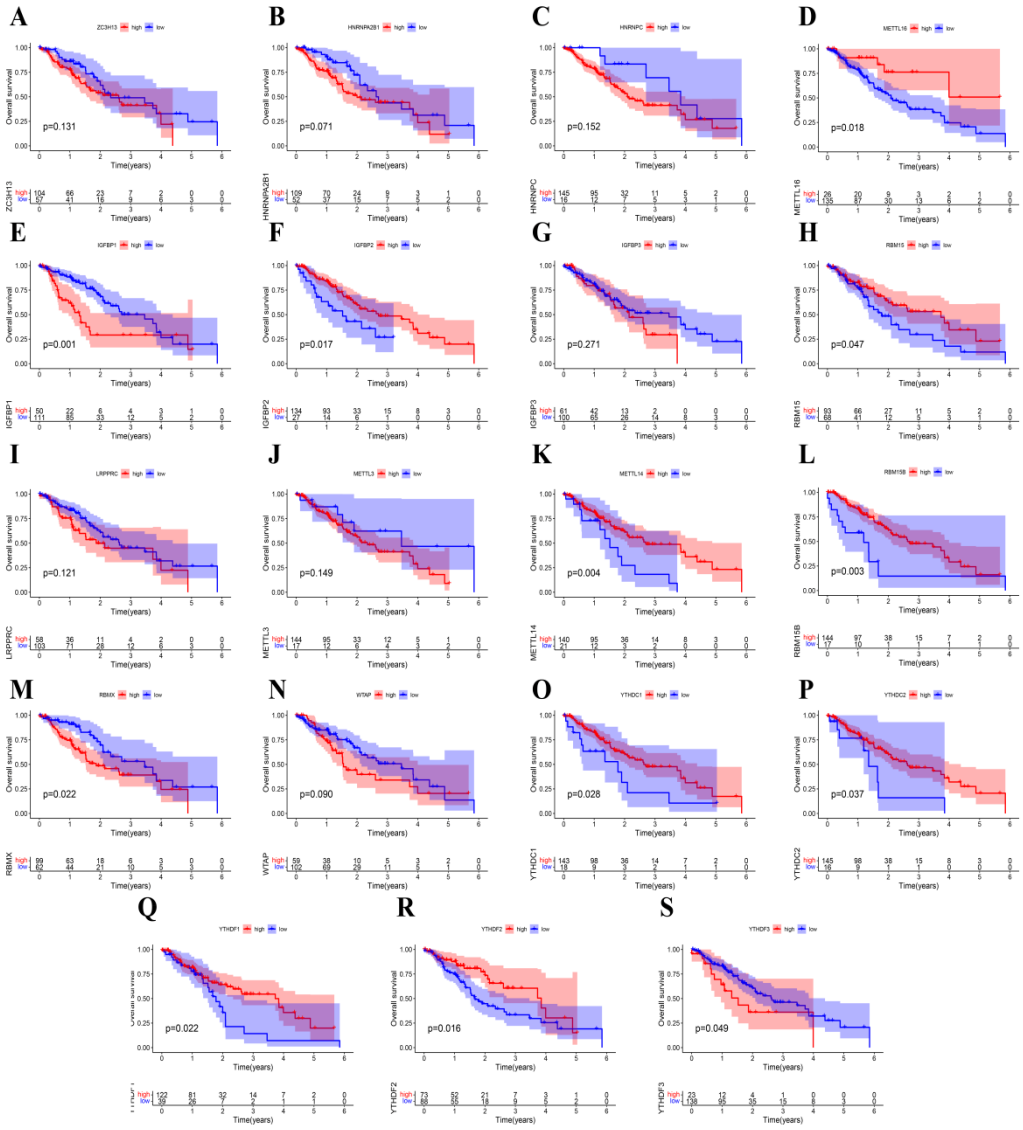
(**C**).The mutation frequency of 22 m^6A regulators in 161 patients with EC from TCGA-ESCA cohort. Each column represented individual patients. The upper barplot showed TMB, the number on the right indicated the mutation frequency in each regulator. The right barplot showed the proportion of each variant type. The stacked barplot below showed fraction of conversions in each sample.

(**D**).The top 20 of mutation frequency genes in 184 patients with EC from TCGA-ESCA cohort. Each column represented individual patients. The upper barplot showed TMB, the number on the right indicated the mutation frequency in each regulator. The right barplot showed the proportion of each variant type. The stacked barplot below showed fraction of conversions in each sample. The asterisks represented the statistical p value (*p< 0.05; **p< 0.01; ***p < 0.001).

(**E**).The expression of 22 m^6A regulators between normal tissues and EC. Tumor, red; Normal, blue. The upper and lower ends of the boxes represented inter quartile range of values. The lines in the boxes represented median value, and black dots showed outliers.

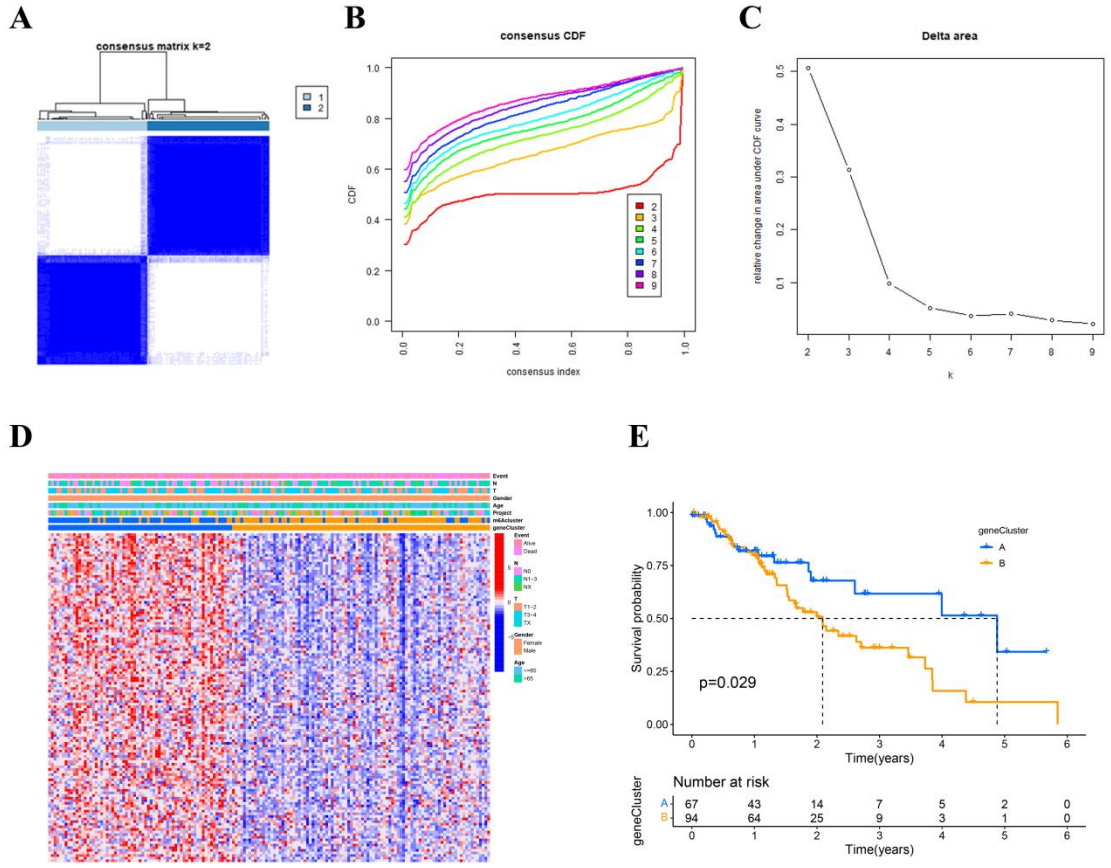
(**F-H**). Prognostic signatures based on ALKBH5, FTO and FMR1 in ES for OS. The figure contains three parts: [1] survival differences estimated by Kaplan-Meier survival curve; [2] number of patients in different groups; [3] number censored at different times.

S3.



(**A-S**).**Prognostic signatures based on expression of 20 m^{6}A regulators gene in EC for OS.**

S4.



Patterns of m^{6}A methylation modification and gene expression of each pattern.

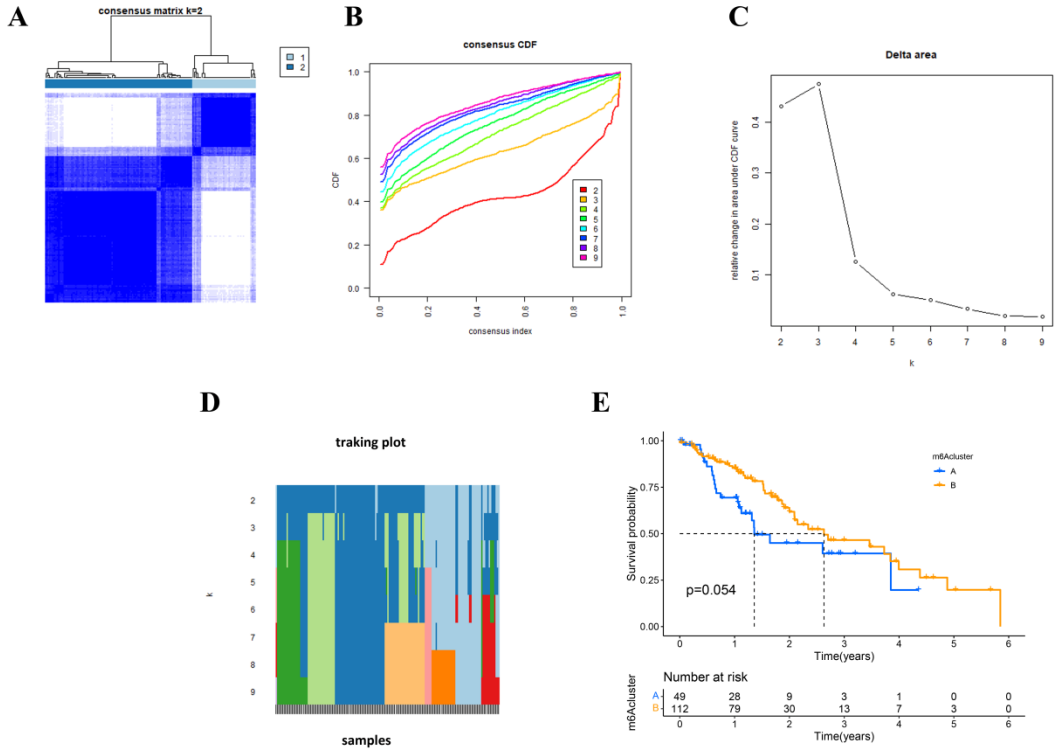
(**A**). Consensus matrices of the TCGA cohort for k = 2.

(**B and C**). The corresponding relative change in area under the cumulative distribution function (CDF) curves when cluster number changing from k to k+1(**B**). The range of k changed from 2 to 9, and the optimal k = 2(**C**).

(**D**). Principal component analysis for the m^{6}A phenotype-related genes of two m^{6}A modification patterns, showing a remarkable difference on transcriptome between different modification patterns.

(**E**). Prognostic signatures based on gene cluster in ES for OS. The figure contains three parts: [1] survival differences estimated by Kaplan-Meier survival curve; [2] number of patients in different groups; [3] number censored at different times.

S5



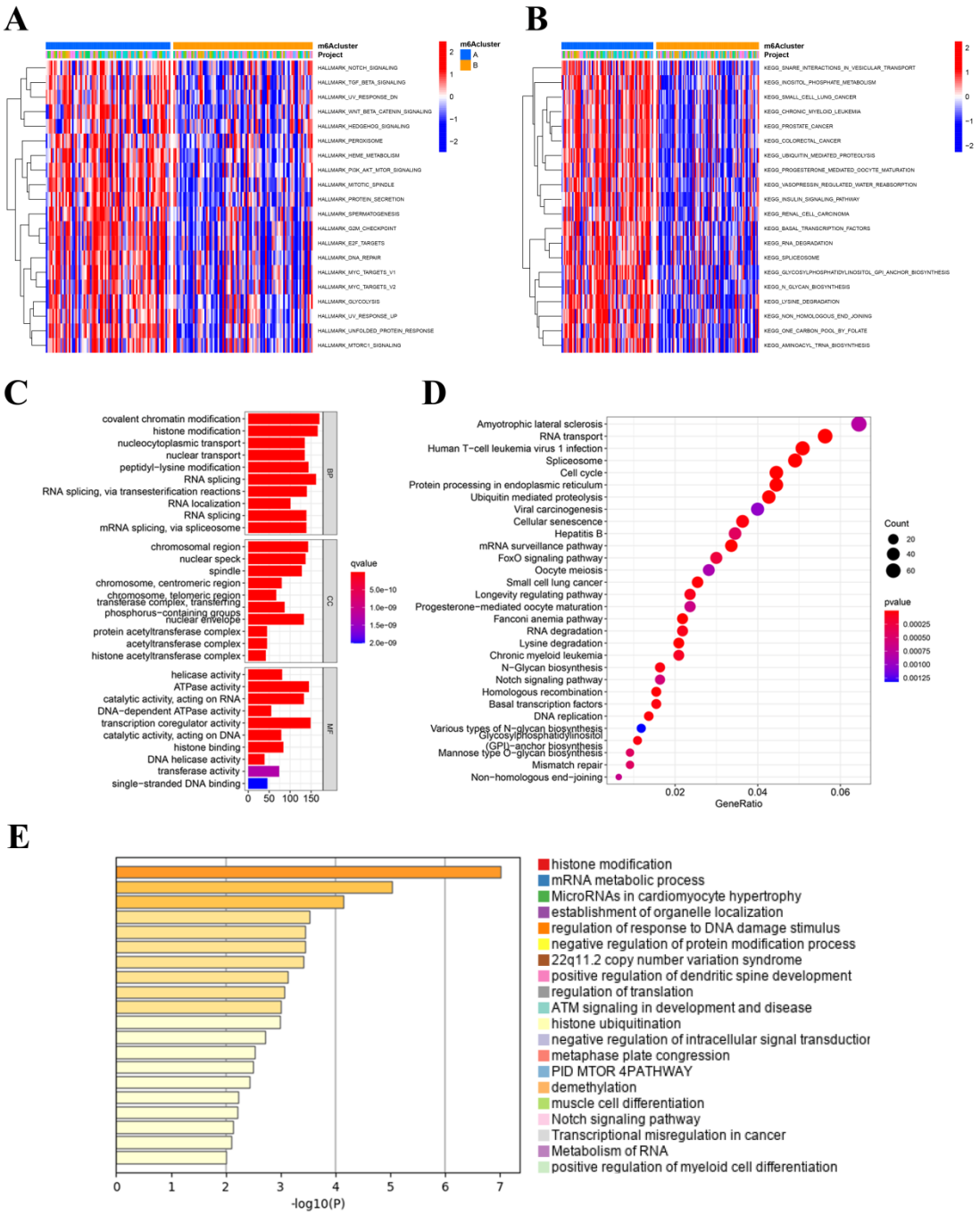
Patterns of m^{6}A methylation modification and gene expression of each pattern.

(**A**). Consensus matrices of the m^{6}A risk gene in TCGA cohort for k = 2.

(**B-D**). The corresponding relative change in area under the cumulative distribution function (CDF) curves when cluster number changing from k to k+1. The range of k changed from 2 to 9, and the optimal k = 2.

(**E**). Prognostic signatures based on genecluster in ES for OS. The figure contains three parts: [1]survival differences estimated by Kaplan-Meier survival curve; [2] number of patients in different groups; [3] number censored at different times.

S6.



Post transcriptional characteristics of RNA in two different m^{6}A modification patterns.

(**A and B**). GSVA enrichment analysis showing the activation states of biological pathways in distinct m^{6}A modification patterns. The heatmap was used to visualize these biological processes, the red represented activated pathways and blue represented inhibited pathways. The ESCA cohorts were used as sample annotations.

(**A**) pathway in HALLMARK;

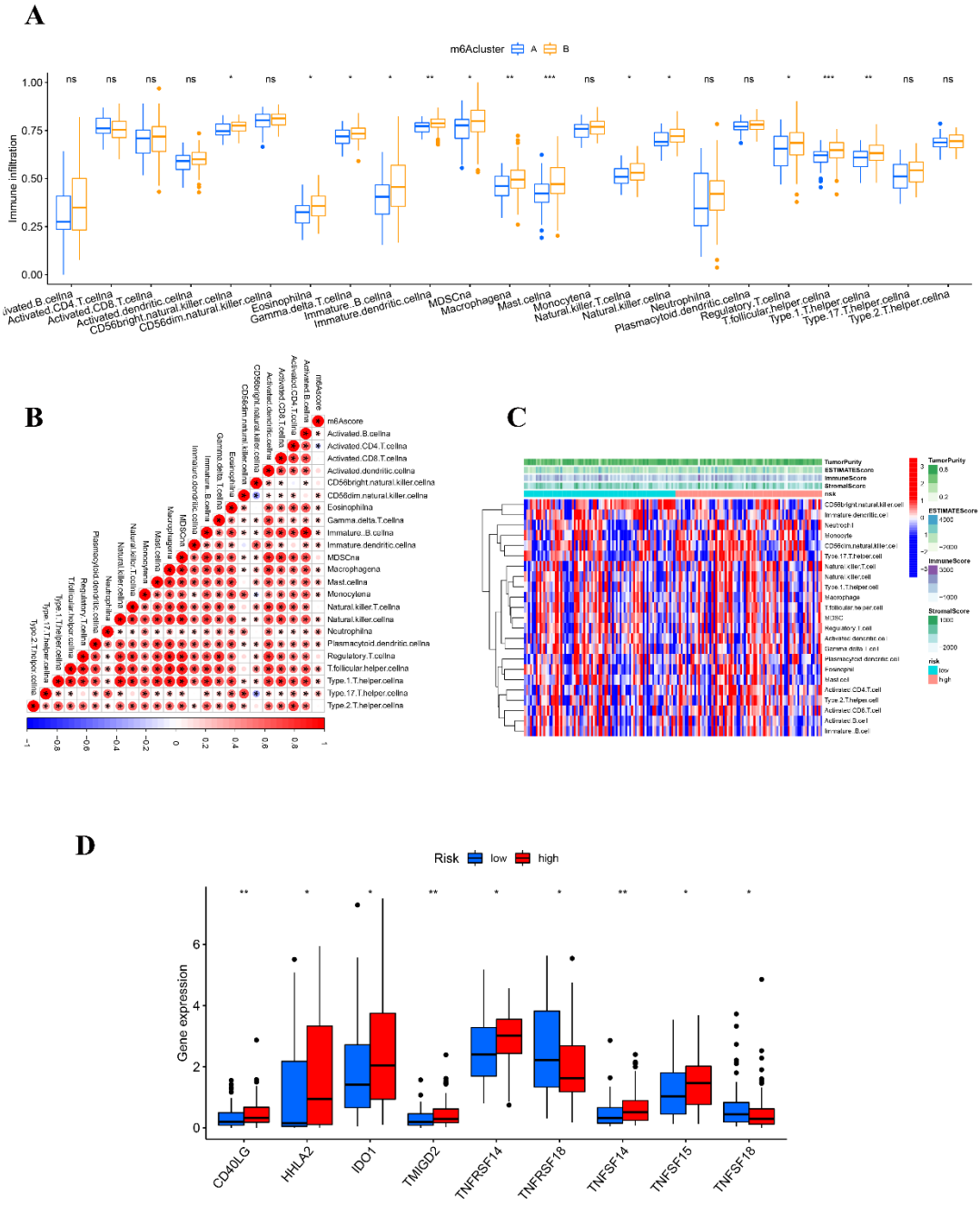
(**B**) pathway in KEGG.

(**C**). Functional annotation for m^{6}A-related genes using GO enrichment analysis. The color depth of the barplots represented the number of genes enriched.

(**D**). Functional annotation for m^{6}A-related genes using KEGG enrichment analysis. The color depth of the dot plots represented the number of genes enriched.

(**E**). Functional annotation for m^{6}A-related genes using GO enrichment analysis and KEGG enrichment analysis by Metascape.

S7.



(**A**).The abundance of each TME infiltrating cell in two m^{6}A modification patterns.

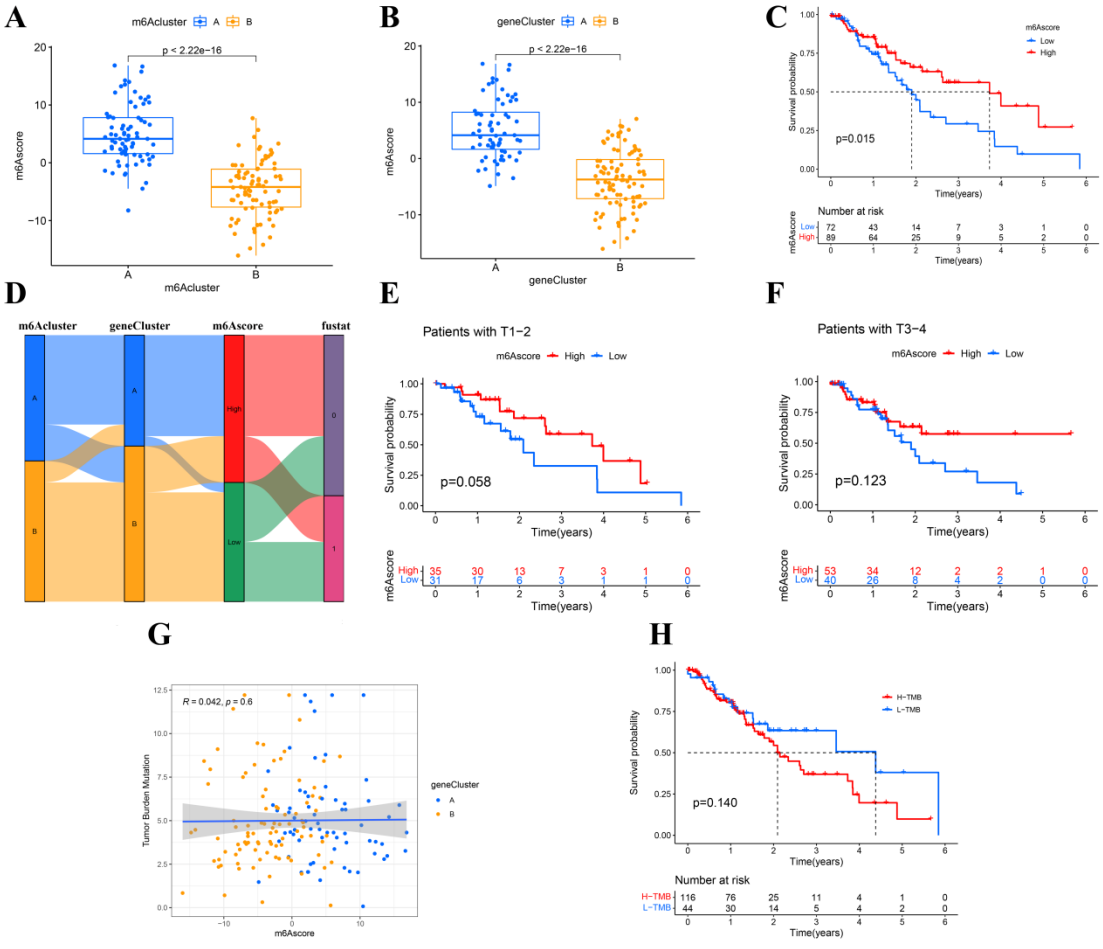
The upper and lower ends of the boxes represented interquartile range of values. The lines in the boxes represented median value, and black dots showed outliers. The asterisks represented the statistical p value (*p< 0.05; **p < 0.01; ***p< 0.001).

(**B**). Correlations between m^6Ascore and the known gene signatures in TCGA cohort using Spearman analysis. Negative correlation was marked with blue and positive correlation with orange.

(**C**).The heatmap shows the degree of immune cell infiltration in the high and low groups of riskcore.

(**D**).Thebargraph shows the correlation of immune checkpoint molecular expression in the high and low groups of riskcore.

S8.



Construction of m^6A signatures and the role in TCGA molecular subtypes and tumor ESCA mutation.

(**A and B**).Differences in m^6Ascore among two gene clusters or two m^6A clusters in TCGA cohort. The Kruskal-Wall is test was used to compare the statistical difference between three gene clusters (p< 0.001).

(**C**).Alluvial diagram showing the changes of m^6Aclusters, TCGA molecular subtypes, gene cluster and m^6Ascore.

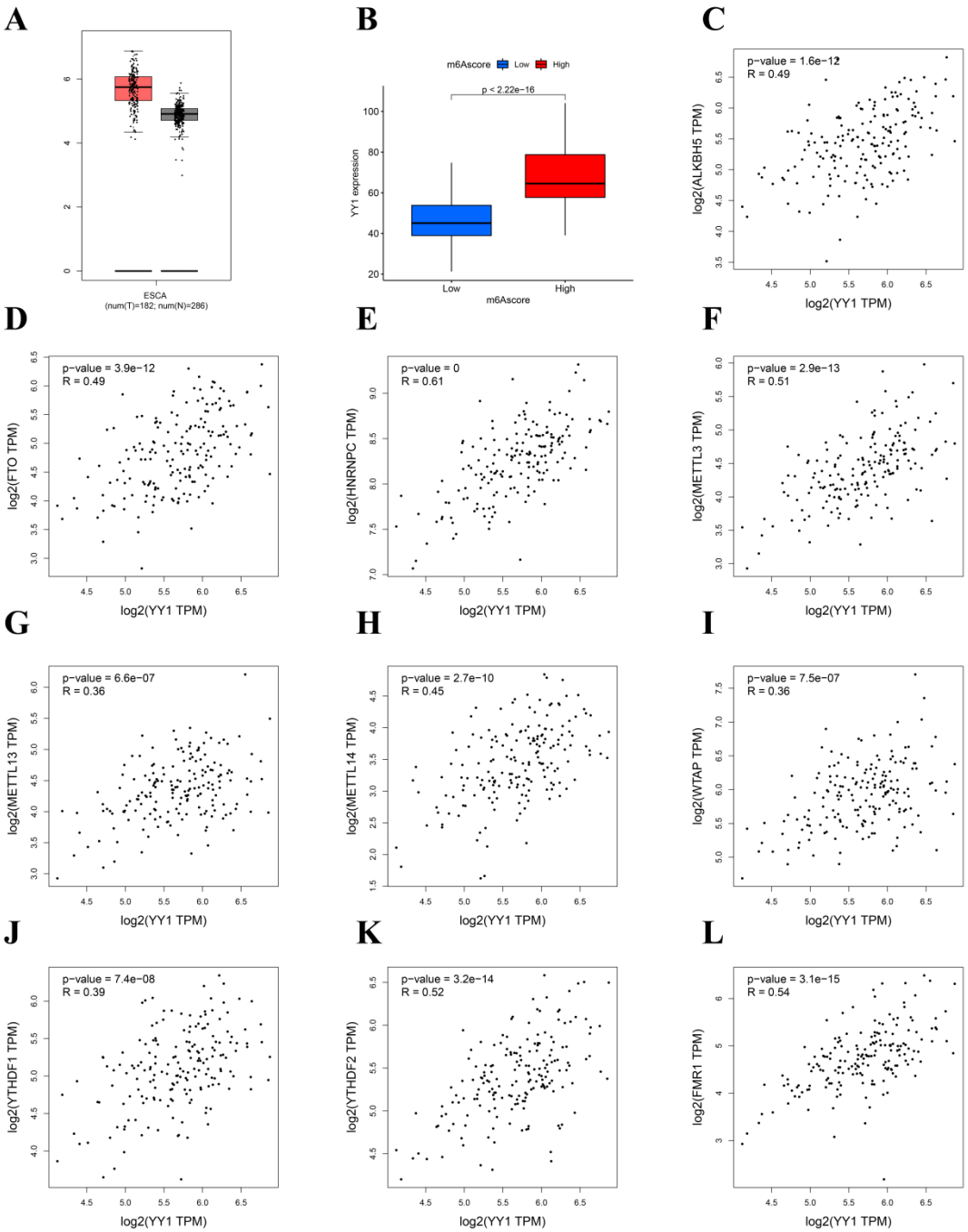
(**D**).Correlation between m^6Ascore and TMB among two gene clusters.

(**E**). Prognostic signatures based on m^6Ascore in ES for OS.

(**F and G**).Prognostic signatures based on Tumor stage in ES for OS.

(**H**).Prognostic signatures based on TMB and m^6Ascore in ES for OS.

S9.



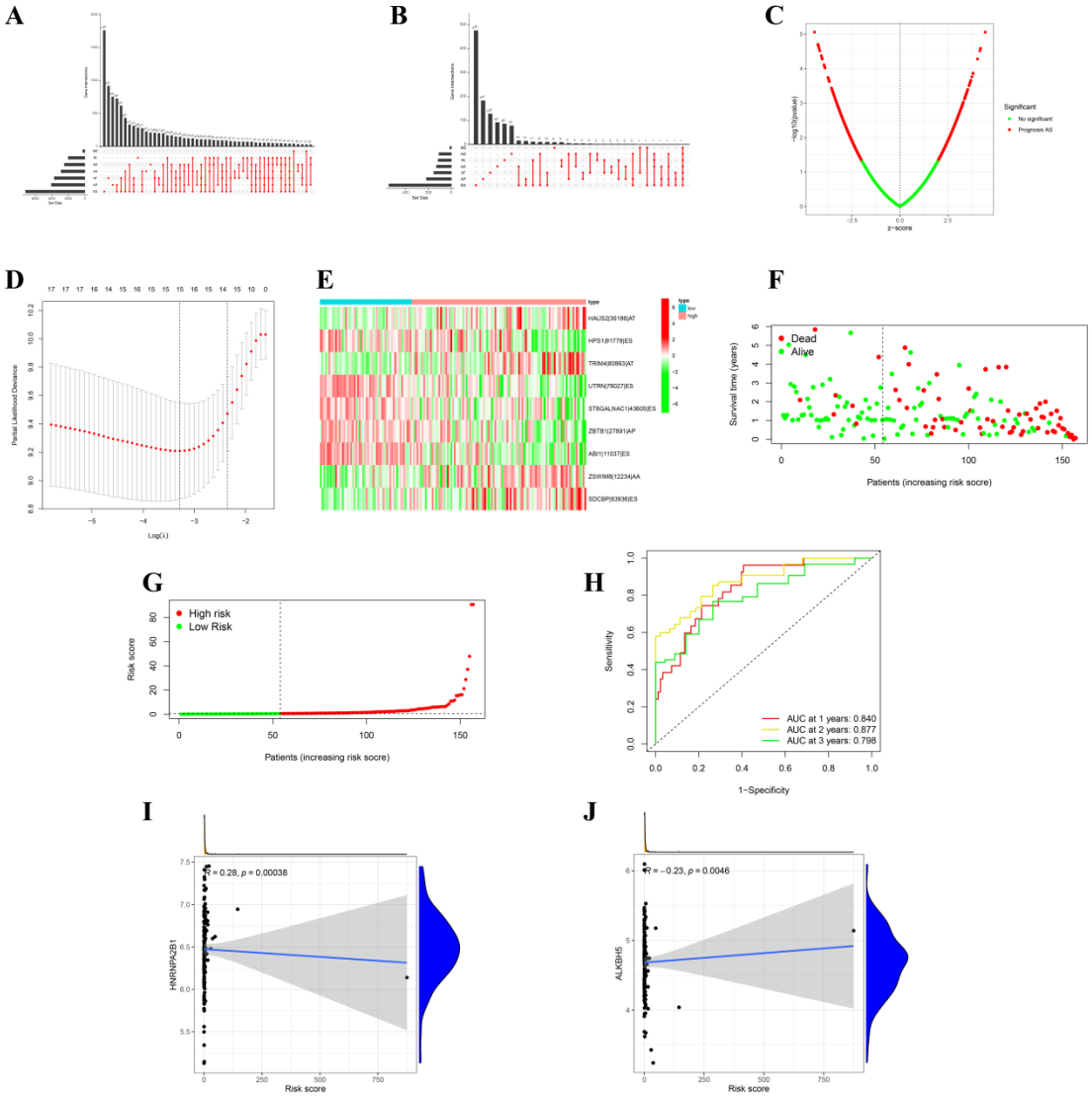
Generation of m^{6}A gene signatures and functional annotation in m^{6}A-related phenotypes.

(**A**).The expression of YY1 in cancer and adjacent species.

(**B**). Differences in YY1 expression between low and high m^{6}Ascore groups ($p < 0.0001$, Wilcoxon test).

(**C-L**). The correlation between YY1 and m^{6}A regulators.

S10.



Correlation between m^{6}A modification and RNA alternative splicing.

(**A**).UpSet plot of the interactions between DEAS events and their parent genes ($p < 0.05$).

(**B**).UpSet plot of the interactions between DEAS events and their parent genes which overlapping phenotype-relate ($p< 0.05$).

(**C**).The DEAS overlapping phenotype-relate identified in ES was visualized in a Volcano plot. The red and green points in the plot represent DEAS with statistical

significance (adj *p* value < 0.05 and |z-score|>2).

(**D**).Cross-validated partial log-likelihood deviance, including upper and lower standard deviations, as a function of log for the DEAS data set. The dotted vertical lines indicate the values with minimal deviance (left) and with the largest value within one standard deviation of the minimal deviance (right).

(**E**).Prognostic signatures based on risk score of DEAS events in ES for OS.

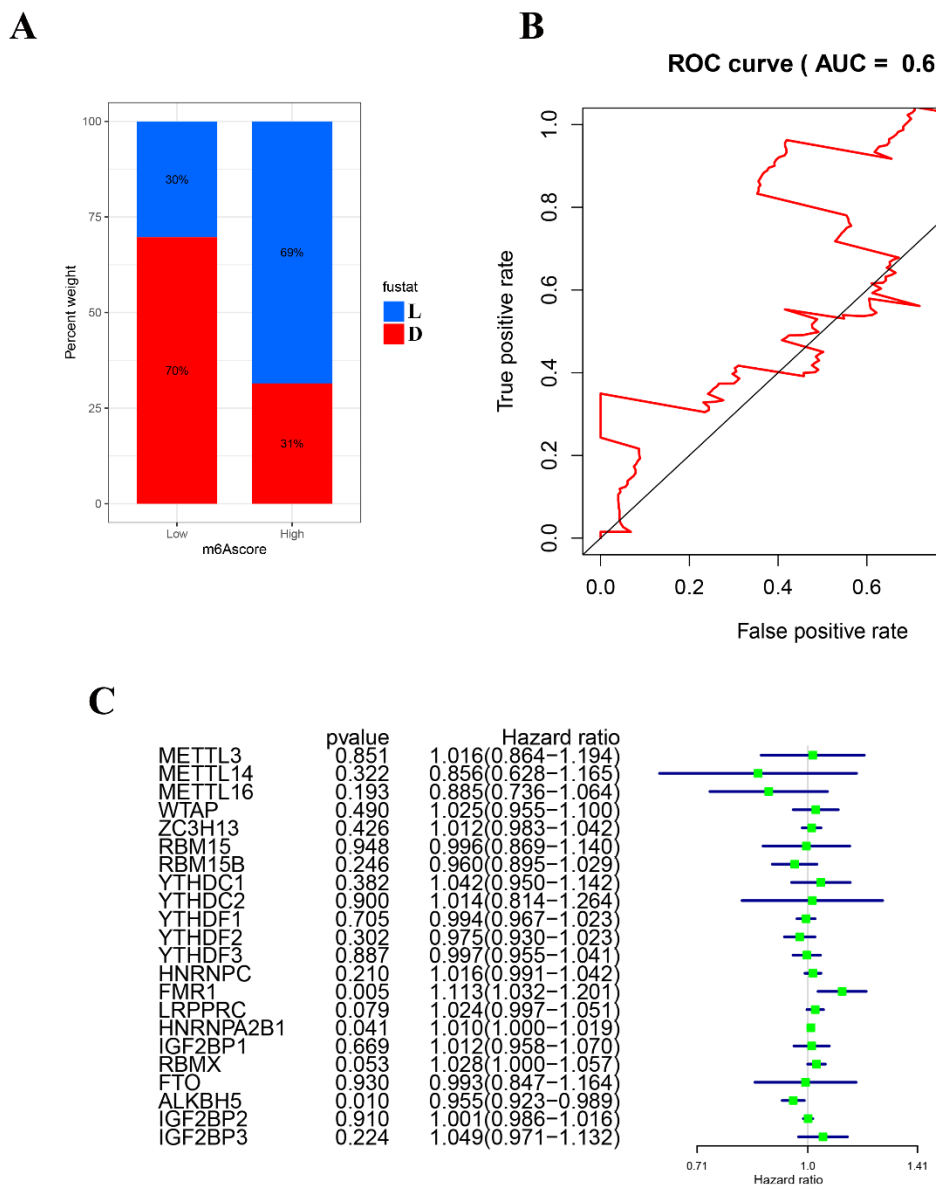
(**F and G**).DFS-related prognostic model. High-risk and low-risk groups were divided based on the median value of risk score. The upper plot illustrated assignment of patients's survival status and survival times, the middle plot showed the risk score curve, and the bottom heatmap represented splicing distribution of the AS in compound prognostic models. Color transition from blue to red indicates the increasing PSI value of corresponding AS event from low to high.

(**H**).The calibration curves of 1-, 2-, and 3-year OS nomogram prediction in the TCGA cohort. The y-axis showed the observed OS, and the red, blue and green line indicated the respective performance of the nomogram with 1-, 2-, and 3-year outcomes in the TCGA cohort.

(**I**).The correlation between HNRNPA2B1 and risk score.

(**J**).The correlation between ALKBH5 and risk score.

S11.

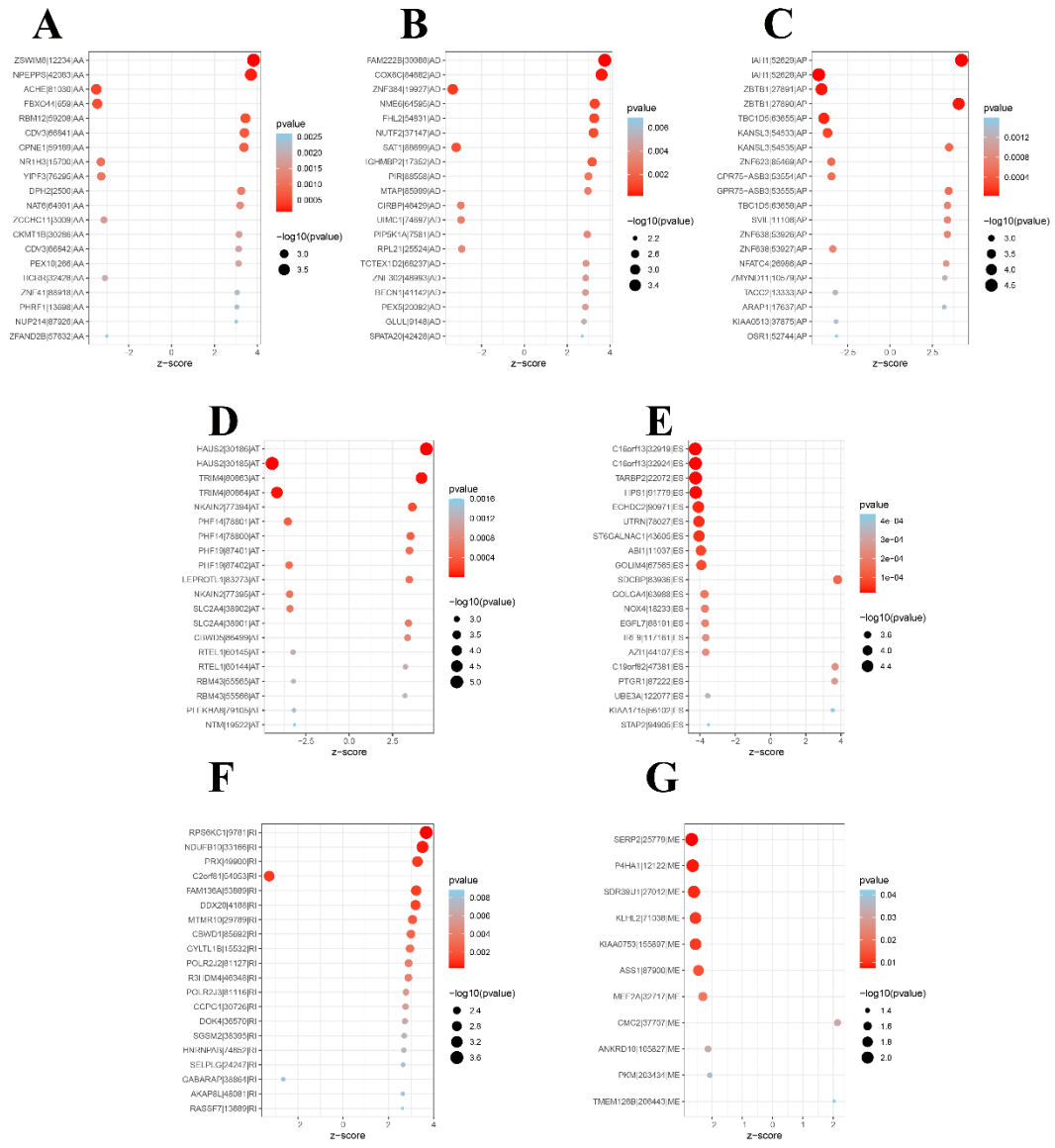


(**A**).The proportion of patients with prognosis in low or high m^6Ascore groups.Dead/Survival: 50%/50% in the low m^6Ascore groups and 31%/69% in the high m^6Ascore groups.

(**B**).Receiver operating characteristic curve for the m^6Ascore of esophageal cancer. Note—AUC, areas under the receiver operating characteristics curve.

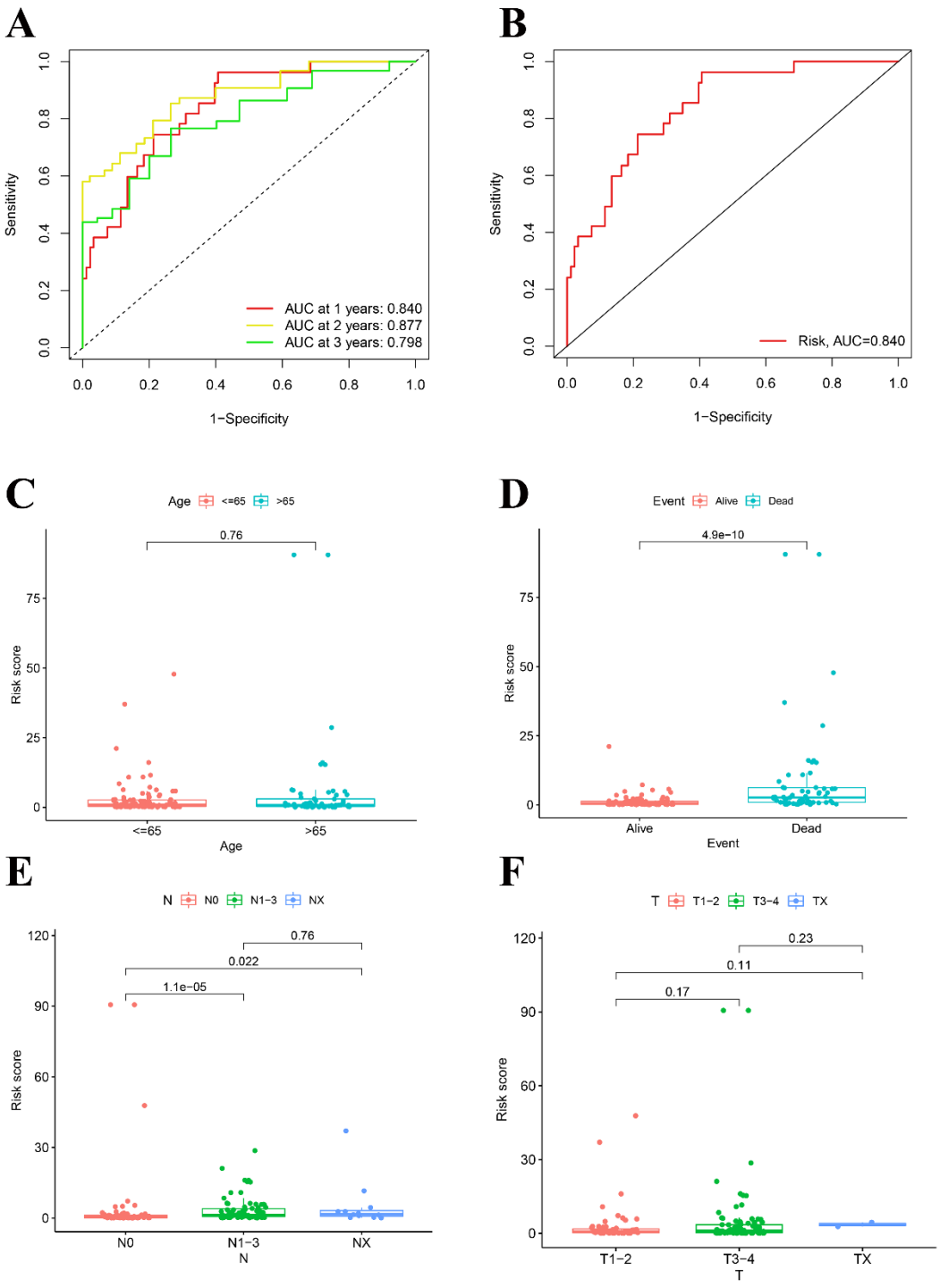
(**C**).DEAS that were simultaneously associated with OS.Univariate and multivariate analysis of DEAS on OS. Unadjusted HRs (boxes) and 95% confidence intervals (horizontal lines) limited to DEAS with $p < 0.05$. Box size is inversely proportional to the width of the confidence interval.

S12.



(**A-G**).The bubble plot shows the most significant target genes involved in seven alternative splicing events in patients with esophageal cancer. The red points in the plot represent Events with statistical significance (adj p value < 0.05 and |z-score|>2).

S13.

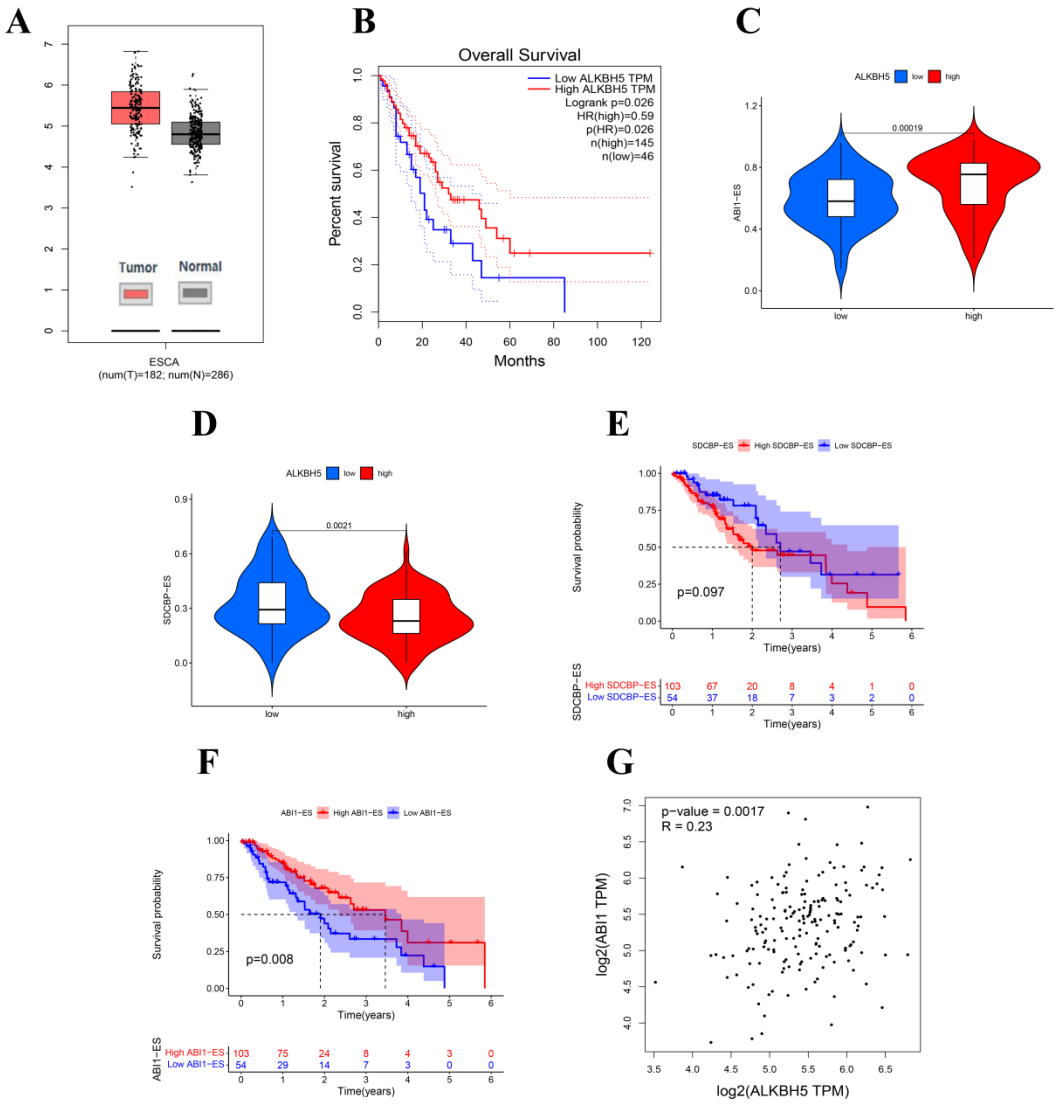


(**A**).The calibration curves of 1-, 2-, and 3-year OS nomogram prediction in the ES cohort. The y-axis showed the observed OS, and the red, yellow and green line indicated the respective performance of the nomogram with 1-, 2-, and 3-year outcomes in the ES cohort.

(**B**).Receiver operating characteristic curve for the alternative splicing risk score of esophageal cancer.

(**C-F**).The correlation between alternative splicing risk score and clinical events in patients with esophageal cancer.

S14.



ALKBH5 regulates ABI1 alternative splicing.

(**A**).The expression of ALKBH5 in cancer and adjacent species.

(**B**). Prognostic signatures based on expression of ALKBH5 in ES for OS.

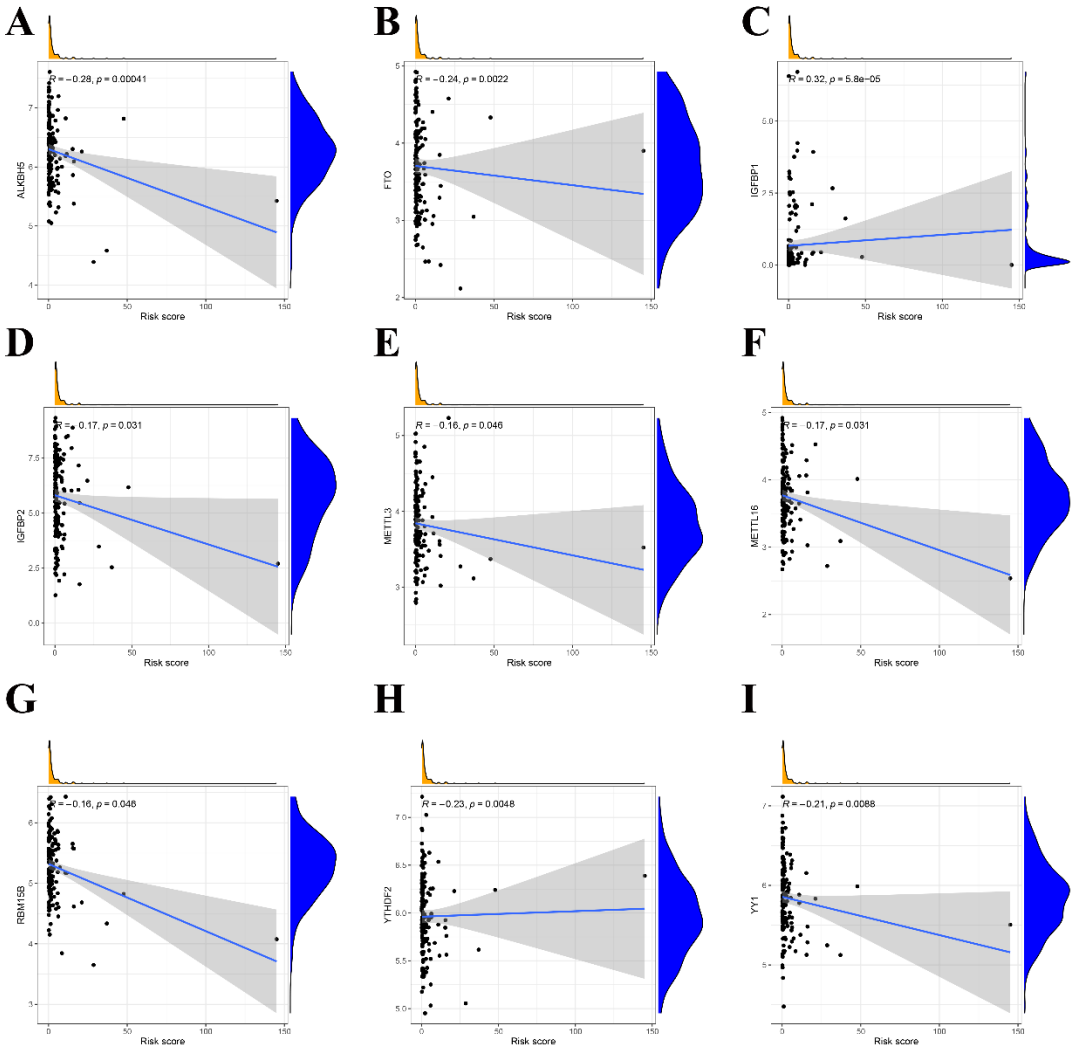
(**C and D**). Differences in ABI-ES and SDCBP-ES among high and low expression of ALKBH5 in TCGA ($p< 0.05$).

(**E**). Prognostic signatures based on ABI-ES and SDCBP-ES events in ES for OS.

(**F**).The correlation between ALKBH5 and ABI1 expression.

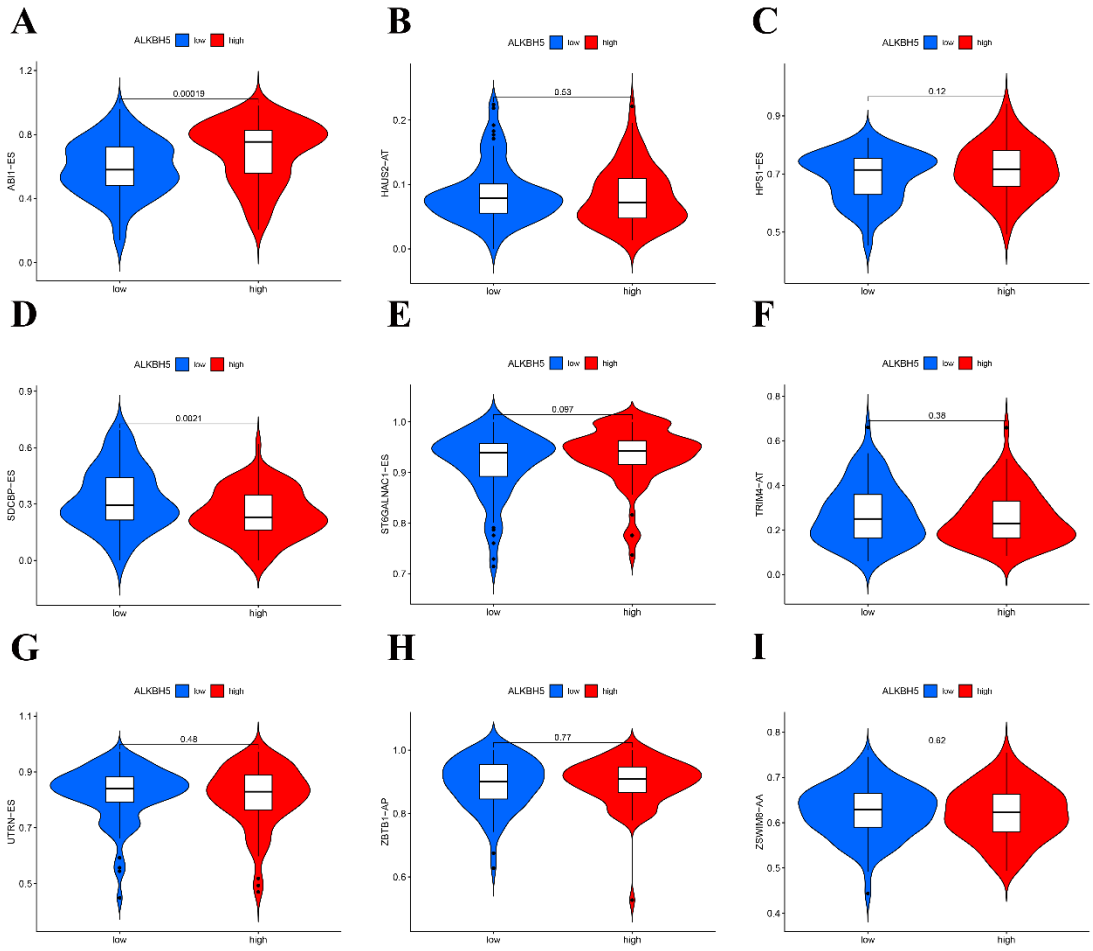
(**G**). Alternative splicing sites and predicted m^6A modification sites.

S15.



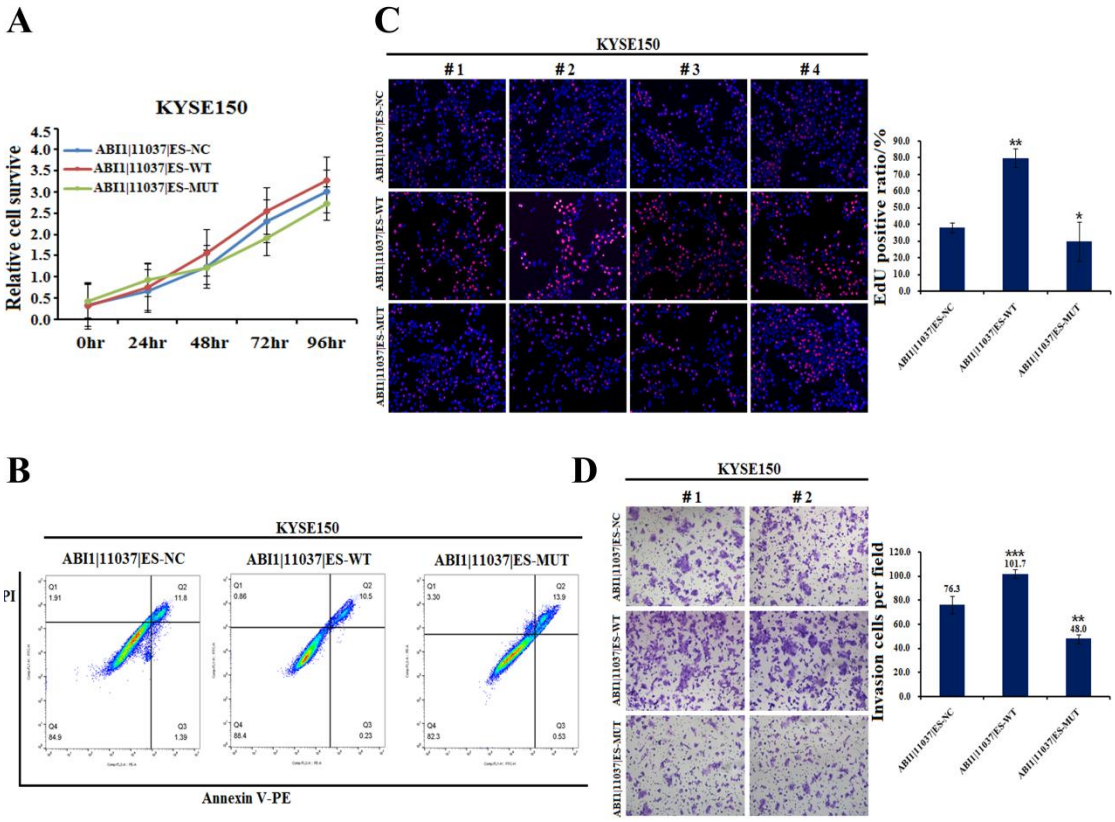
(**A-I**). The correlation between expression of ALKBH5and significant alternative splicing events.

S16.



(**A-I**). The correlation between alternative splicing risk score and the expression of m^6A regulators.

S17.



ABI1|11037|ES event inhibits cell proliferation and migration in ESCC cell lines.

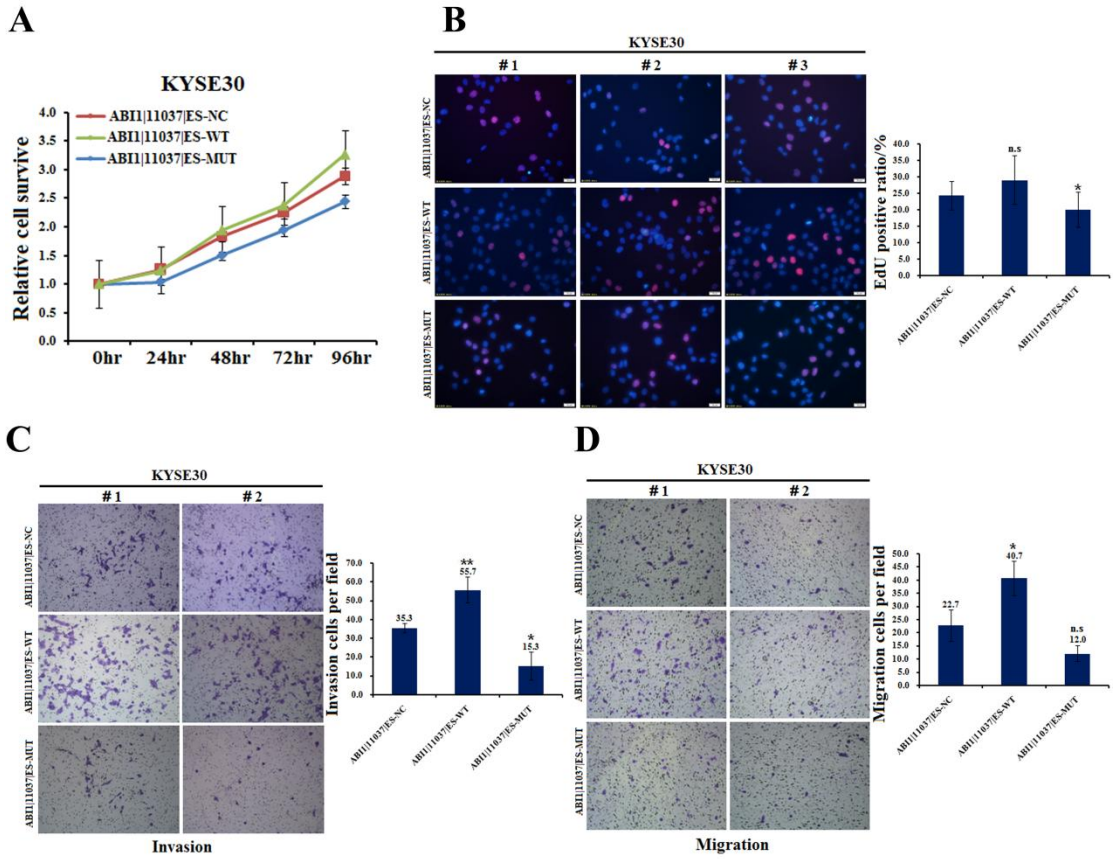
(**A**).CCK-8 assays every 24 hours inKYSE150 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUToverexpressionversus the negative control (ABI1|11037|ES-NC).

(**B**).The cells apoptosis as analyzed by flow cytometry in KYSE150 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT overexpression versus the negative control (ABI1|11037|ES-NC).

(C**).**5-Ethynyl-2′-deoxyuridine (EdU) was analyzed in KYSE150 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT overexpression versus the negative control (ABI1|11037|ES-NC). *, p value < 0.05; **, p value < 0.01.

(**D**).The cells invasion was analyzed in KYSE150 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT overexpression versus the negative control (ABI1|11037|ES-NC). **, p value < 0.01; ***, p value < 0.001.

S18.



ABI1|11037|ES event inhibits cell proliferation and migration in ESCC cell lines.

(**A**).CCK-8 assays every 24 hours inKYSE30 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUToverexperssionversus the negative control (ABI1|11037|ES-NC).

(**B**).5-Ethynyl-2′-deoxyuridine (EdU) was analyzed in KYSE30 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT overexpression versus the negative control (ABI1|11037|ES-NC). *, p value < 0.05; n.s, no statistical significance.

(C**).**The cells invasion was analyzed in KYSE30 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT overexpression versus the negative control (ABI1|11037|ES-NC). *, p value < 0.05; **, p value < 0.01.

(**D**).The cells migration was analyzed in KYSE30 cells treated with ABI1|11037|ES overexpression (ABI1|11037|ES-WT) and ABI1|11037|ES-MUT over expression versus the negative control (ABI1|11037|ES-NC). *, p value < 0.05; n.s, no statistical significance.