

Characteristic of Classification and Prognostic Scoring Based on M6a RNA Methylation Modification in Lung Adenocarcinoma and Tumor Microenvironment Association

Xiaoming Yu

Division of Pulmonary Medicine, Key Laboratory of Heart and Lung, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang

Mingming Han

Division of Pulmonary Medicine, Key Laboratory of Heart and Lung, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang

Zhucheng Zhao

Division of Pulmonary Medicine, Key Laboratory of Heart and Lung, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang

Yanfan Chen

Division of Pulmonary Medicine, Key Laboratory of Heart and Lung, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang

Liangxing Wang (✉ wangliangxing@wzhospital.cn)

Division of Pulmonary Medicine, Key Laboratory of Heart and Lung, The First Affiliated Hospital of Wenzhou Medical University <https://orcid.org/0000-0002-1312-9554>

Research Article

Keywords: lung adenocarcinoma, m6A modification, prognosis, tumor microenvironment, PD-L1, PD-1, CTLA4

Posted Date: February 23rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1355529/v1>

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

Abstract

Accumulating evidence links m6A modification with lung adenocarcinoma (LUAD) tumorigenesis development and tumor immune microenvironment. However, the deep impact of m6A modification in tumor-immune interactions of LUAD still remains largely unknown. In the present study, the landscape of m6A modification features were drawn by analyzing the multi-omics features of 20 m6A regulators from TCGA database. Many m6A regulators were aberrantly expressed in LUAD, whose mutations and copy number variation (CNV) frequency were difference. Besides, overall survival showed many m6A regulators are associated with LUAD prognosis. However, the three m6A modification clusters (m6ACluster) determined by unsupervised consensus clustering based on m6A regulators was hard to distinguish. And then the common intersection differentially expressed genes (inter-DEGs) with significant prognosis among m6AClusters were obtained by univariate cox regression analysis. Encouragingly, the m6A-related geneClusters based on significant inter-DEGs were distinguished from each other by PCA analysis. The three geneClusters were depicted the high-, middle- and low- immune cell infiltration subtypes by ssGSEA analysis. Specially, we established a scoring tool named genescore, which was negatively correlated with PD-1, PD-L1, and CTLA4 expression and could reflect both the tumor microenvironment characterization and prognosis of LUAD patients. Majority of the DEGs between high and low genescore group were closely related to the tumor immune response. Low genescore group associated with poor prognosis exhibited higher immuneScore and lower tumorPurity applied by the ESTIMATE algorithm. Application of genescore on estimation of drug sensitivity confirmed that low genescore was more sensitive to Cisplatin, Docetaxel, Gemcitabine, Paclitaxel and Vinblastine, while the high genescore group was more sensitive to Erlotinib. Our study indicated that m6A modification related classification closely associated with tumor microenvironment and will provide novel insights into the molecular mechanisms of LUAD and contribute to develop more effective personalized treatment decisions.

1 Introduction

With an estimated 2.2 million new cancer cases and 1.8 million deaths, lung cancer is the second most commonly diagnosed cancer and the leading cause of cancer death in 2020[1]. Lung adenocarcinoma (LUAD), the major pathological type of lung cancer of non-small cell lung cancer (NSCLC) [2], accounting for approximately 40% of all lung cancer cases[3]. In the past few decades, the advancements in treatment of lung cancer have developed surprisingly, especially in the aspects of targeted therapy and immunotherapy. However, the final prognosis of the same LUAD stage displayed significant difference due to tumor heterogeneity. Therefore, an in-depth understanding of molecular mechanisms underlying heterogeneity is indispensable for more promising therapies.

Recent years, wide attention is attracted in RNA m6A modification (m6A), which is a ubiquitous post-transcriptional modification mechanism of RNA in eukaryotic cells[4; 5]. The dynamic reversible m6A RNA methylation process is regulated by three types of regulators, i.e. methyltransferases (writers), demethylases (erasers) and binding proteins (readers)[6]. The writers, including METTL3, METTL14,

METTL16, WTAP, VIRMA, ZC3H13, RBM15 and RBM15B, are responsible for adding m6A to mRNA. The role of readers is recognizing and binding to m6A sites of mRNAs and generating corresponding biological signals. The erasers, which are demethylases that convert m6A into A109, currently refer to FTO and ALKBH5[4].

Increasing evidence showed that m6A regulators are involved in tumorigenesis and the role in pan-cancer were complex and different[7]. Aberrant expression of m6A regulators, including METTL3, YTHDFs, HNRNPA2B1, ALKBH5 and FTO, has been identified in lung cancer[8]. The m6A-related patterns and association with TME in lung cancer and other different cancers had been reported in previous studies[9; 10; 11; 12; 13; 14; 15; 16]. However, there are still largely information about the influence of m6A modification on tumor-immune interactions in LUAD remained to be explored.

In the present study, we first evaluated the multi-omics features of 20 m6A regulators from TCGA dataset. And then we comprehensively evaluated the association between m6A-related geneClusters and TME cell-infiltrating characteristics by integrating the transcriptomic and genomic data of 1509 LUAD samples from TCGA and GEO databases. Importantly, we established a scoring tool called genescore, which is strongly correlated with the tumor microenvironment characterization in LUAD and displays the potential in predicting response in PD-1, PD-L1 and CTLA4.

2 Materials And Methods

Data Acquisition

The RNA-sequencing data from the TCGA-LUAD dataset which contains 497 LUAD-ADC samples and 54 normal tissues, the corresponding clinical data and simple nucleotide variation data processed by Varscan2 software were all downloaded from TCGA website (<https://cancergenome.nih.gov/>). The TCGA RNA-sequencing data (FPKM format) were transformed into transcripts per kilobase million (TPM) format. CNV data, the RNA-sequencing data of pan-cancer and the corresponding clinical data were all downloaded from Xena public data hubs (<https://xena.ucsc.edu>). Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) was thoroughly queried for all eligible LUAD datasets contained survival information, and the ADC subtype was screened from 7 datasets (GSE19188, GSE30219, GSE31210, GSE37745, GSE50081, GSE68465 and GSE72094)for analysis, among which GSE19188, GSE30219, GSE31210, GSE37745 and GSE50081 had the same platform of the dataset GPL570. The probes were converted to gene symbols based on the corresponding annotation files. The batch effects among different GEO datasets were removed by using the ComBat method from the 'SVA' R package[17]. 20 m6A regulators were collected from previously published studies (Supplementary Table 1)[5; 18; 19; 20; 21; 22]. Immunophenoscore (IPS) of 515 TCGA-LUAD patients was download from TCIA database (<https://tcia.at/home>). To ensure reproducibility of our study, we selected TCGA + GSE68465 as training set and Meta-cohort (GSE19188, GSE30219, GSE31210, GSE37745 and GSE50081) as validation set for further analysis. The clinical information of the TCGA + GSE68465 datasets are respectively listed in Supplementary Table 2.

Aberrant Expression, Mutation and CNV of m6A regulators in TCGA-LUAD

To identify differentially expression of m6A regulators in the TCGA-LUAD dataset, the expression profiles of genes were compared between tumor and normal tissues by the “limma” and “reshape2” R packages and box plot was drew by the “ggpubr” R package. The “maftools” package was used to extract mutations in those genes from mutation annotation format (MAF) files and draw waterfall plot. The copy number variation landscape of 20 m6A regulators was plotted by the “ggpubr” R package.

Unsupervised consensus clustering

We performed unsupervised consensus clustering with “ConsensusClusterPlus” R package[23] to identify distinct m6A modification cluster (m6ACluster) based on the expression matrix of m6A regulators. The R package “limma” was used to evaluate DEGs in LUAD samples among different m6AClusters. The Venn diagram and univariate Cox regression analysis were performed to screen out the common inter-DEGs with prognostic significance for identifying new different cluster (geneCluster) by the same way. The consensus matrix, the cumulative distribution function (CDF) and the CDF Delta area curves were in the Supplementary Fig. 1.

Survival analysis and principal component analysis (PCA) analysis

Kaplan-Meier survival analysis was conducted to estimate the prognostic impact of the three clusters on OS by the “survival” R package and Kaplan-Meier survival curves were plotted by the “survminer” R package. PCA was used to verify the clusters by “limma” R package and visualized by “ggplot2” R package.

Single-Sample Gene Set Enrichment Analysis (ssGSEA) and Gene set variation analysis (GSVA)

Single-sample gene set enrichment analysis (ssGSEA) applies gene signatures expressed by immune cell populations to individual cancer samples. ssGSEA in R package “GSVA” was introduced to quantify the infiltration levels of the 23 immune cell types in tumor microenvironment (TME). Enrichment of 29 immune signatures for each LUAD sample was quantified using the ssGSEA score (enrichment level) as previously described[24]. The pathway enrichment scores were quantified by the “GSVA” R package. The well-defined biological signatures derived from the Hallmarker gene set [25] were download from MSigDB database v7.4 (<http://www.gsea-msigdb.org/gsea/downloads.jsp>). Differential analysis was then conducted to determine the significantly enriched components in each cluster with “limma” package in R software. The boxplot and heatmap were drawn with the corresponding R package respectively. Supplementary File 1 shows the detailed information of immune infiltration, immune signatures and pathway signatures used in this study.

Construction of genescore

To quantitatively evaluate of m6A-related geneCluster for individual LUAD patients, we established a scoring tool. The establishment procedures of gene scoring tool were as follows: Differential analysis

and Venn diagram screened out the common differential genes among three geneClusters. And then univariate Cox regression analysis was performed for each gene. The genes with the significant prognosis were extracted for PCA analysis to calculate the genescore using the formula:

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

where i is the expression value of each selected genes.

Application of genescore in clusters and clinical features

Kaplan-Meier survival analysis was applied to evaluate the prognostic impact of the high and low genescore on OS. Differential analysis was performed between genescore and cluster by the "limma" R package and plotted by the "ggpubr" R package. Application of genescore into clinical features and visualized with the "plyr", "ggplot2" and "ggpubr" R packages. Alluvial diagram was analysis and drawn by the "ggalluvial", "ggplot2" and "dplyr" R packages to reveal the changes of m6AClusters, geneClusters, genescores, and respond to fustat.

Identification of Differentially Expressed Genes (DEGs) between genescores and functional analysis

To identify the DEGs between the genescores, the empirical Bayesian approach of the "limma" R package was applied in the standard comparison mode. The significance criteria for determining DEGs was set to adjusted P-value < 0.001. Venn analysis of the common inter-DEGs from the training set and validation set database was displayed in the Supplementary File2. To investigate the pathways enriched in the subgroups, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene Ontology (GO) biological processes by applying a threshold P-value < 0.05 with "clusterProfiler", "org.Hs.eg.db", "enrichplot" and "ggplot2" packages in R software.

Tumor microenvironment estimation

IPS is a superior predictor of response to anti-CTLA-4 and anti-PD-1 regimens, which quantify the determinants of tumor immunogenicity and characterize the intratumoral immune landscapes and cancer antigenomes[26]. The performance of IPS in different Genescore groups is illustrated with a violin diagram by the "ggpubr" R package. The StromalScore, ImmuneScore, ESTIMATEScore and TumorPurity for each patient were obtained by using the "estimate" R package. Tumor tissues with abundant immune cell infiltration represented a higher immune score and lower level of tumor purity. The differential expression of genes at the immune checkpoint sites were analyzed by the "limma" R package. And the representative immune checkpoint genes PD-1, PD-L1 and CTLA4 of LUAD were listed separately. The correlation between TMB and genescore or geneCluster in TCGA-LUAD was analyzed by the "ggpubr" and "reshape2" R packages.

Estimation of drug sensitivity

Drug sensitivity was estimated by the “pRRophetic” R package and determined by the concentration required for 50% of cellular growth inhibition (IC50).

Statistical analysis

The initial processing of the expression data downloaded from TCGA and GEO was conducted by the Perl (version 5.28.2). All statistical analyses and figures were generated by R software (version 4.1.0). A P-value < 0.05 was considered to have statistical significance.

3 Results

Features of m6A regulators in TCGA-LUAD

Mutations of m6A regulators were rare in LUAD, ranging from 0 to 3% (Fig. 1A). We explored the differential expression levels of m6A regulators between tumors and adjacent tissues, and it revealed that 15 out of the 20 m6A regulators were aberrantly expressed (Fig. 1B). Copy number frequency chart showed that YTHDF1/3, VIRMA, FMR1, METTL3, HNRNPA2B1, LRPPRC and FTO displayed widespread copy number gains. Inversely, YTHDF2, WTAP, YTHDC2, ALKBH5, ZC3H13, RBM15, METTL14, RBM15B and METTL16 displayed prevalent copy number losses (Fig. 1C). The 20 m6A regulators in TCGA-LUAD were positively correlated with each other in network (Fig. 1D). CNV alteration of m6A regulators were widely distributed on 23 chromosomes (Fig. 1E).

The prognosis of m6A regulators in LUAD patients

Survival analyses revealed that most m6A regulators had impact on OS in the TCGA + GSE68465 cohort. The impact of most m6A regulators on OS in Meta-cohort from GEO was consistent with the results of TCGA + GSE68465 cohort. The common m6A regulators in different cohorts associated with good survival in LUAD patients were METTL3, FMR1, YTHDC2 and YTHDF2. While the common m6A regulators associated with poor survival were RBM15, WTAP and HNRNPC (Fig. 2).

Identification of m6A-related phenotypes in LUAD

To explore the heterogeneity of different m6A regulators acted in LUAD, we performed unsupervised clustering in TCGA + GSE68465 cohort based on the expression of 20 m6A regulators together with 3 IGFBP genes mentioned in the discussion of our study. And then we identified three distinct m6A-related modification clusters (m6AClusters) with the best clustering stability based on the similarity between the expression level of m6A regulators and the proportion of fuzzy clustering measures (Supplementary Fig. 1A). Survival analysis showed that the prognostic indicator OS varied among m6AClusters (Fig. 3A). However, the PCA plot showed the distribution directions of the three m6AClusters were mixed and there was no clear boundary, suggesting that the m6A regulators couldn't well divide LUAD patients into three groups (Fig. 3B). To explore the difference among three m6AClusters, difference analysis was performed to screen for the DEGs between each other and Venn diagram was performed to circle out the inter-DEGs (Supplementary File 3) for the following analysis. We obtained the significant prognostic inter-DEGs

(Supplementary File 3) by univariate Cox regression analysis. Unsupervised consensus clustering was subsequently implemented in R with the expression of 6 significant prognostic inter-DEGs, and $k = 3$ was the optimal result, with clustering stability increasing from $k = 2$ to 9 (Supplementary Fig. 1B). Strikingly, according to the PCA result of three independent samples, LUAD in TCGA + GSE68465 was successfully divided into three clusters (geneClusters, Fig. 3B). And geneCluster C showed the best survival prognosis among the three phenotypes (Fig. 3A). To validate the result of geneCluster above, the same way was performed to process the dataset in Meta-cohort. Comfortingly, the results in the Meta-cohort were consistent with the training set.

Association of Tumor Microenvironment (TME) Immune Cell Infiltration, GSVA and expression of immune checkpoint genes (PD-1, PD-L1 and CTLA4) with geneCluster

To explore the association of TME immune cell infiltration in geneClusters, we evaluated the proportion difference of 23 immune cell types in the three geneClusters based on the ssGSEA. The three geneClusters exhibited significantly different characteristics of immune cell infiltration (Fig. 4A). The immune cell infiltrations were the most abundant in geneCluster A of the training set and in geneCluster B of validation set. And the degree of most immune cell infiltration in both geneCluster C was significantly lowest. To better understand the immune functions and immune response-related pathways, 29 immune signatures were enriched for each LUAD sample and differential analysis was performed among the three geneClusters. Interestingly, both cohorts showed similar clustering results, with three clusters being relatively clearly separated (Fig. 4B). High immunoactivity was exhibited in geneCluster A of the training set or in geneCluster B of validation set. Low immunoactivity was exhibited both in geneCluster C. Middle immunoactivity was exhibited in rest geneCluster. To depict the biological behavior differences among geneClusters, functional annotations were performed using GSVA algorithm. The results showed that the geneClusters displayed significant differences in the pathways enriched. 32 signature pathways were chosen from the Venn of 46 signature pathways in training set and 88 signature pathways in validation set for the next analysis (Fig. 4C). And then we constructed a heatmap with GSVA to visualize and compare the 32 signature pathways enriched among distinct geneClusters (Fig. 4D). GSVA indicated that geneCluster A in training set was significantly enriched in immune activation-related processes, including antigen processing and presentation, toll like receptor signaling pathway, T cell receptor signaling pathway, B cell receptor signaling pathway, graft versus host disease, allograft rejection, intestinal immune network for IGA production, primary immunodeficiency and so on. Similarly, the geneCluster B in validation set was significantly enriched in immune activation-related processes. However, geneCluster C with comfortable prognosis was significantly enriched in metabolism, including alpha linolenic acid metabolism, ascorbate and aldarate metabolism, valine leucine and isoleucine degradation and propanoate metabolism.

We then explored the expression of hotspot immune checkpoint genes PD-1, PD-L1 and CTLA4 in geneClusters according to the differential expression of check-point signature in TME immune above. Meanwhile, the expression of the hotspot genes in m6AClusters were presented. The expression of PD-L1 was merely depicted in TCGA dataset instead of TCGA + GSE68465 dataset on account of there was no

PD-L1 expressed in GSE68465 dataset. The classification of geneCluster was relatively better to distinguish three check-point genes than m6ACluster in LUAD. Investigation of the expression of PD-1 in geneCluster showed the expression in geneCluster C was lower than geneCluster A. However, pairwise comparisons of other geneClusters are not available (TCGA + GSE68465). As for the validation set, the expression of PD-1 in geneCluster C was the lowest. Other pairwise comparisons were also not available. The expression of PD-L1 in geneCluster A (TCGA) or geneCluster B (Meta-cohort) was the highest among three geneClusters. And the expression of PD-L1 in geneCluster C (TCGA) was lowest. However, there were no significant difference expression of PD-L1 between geneCluster A and C in Meta-cohort. Comfortingly, the expression of CTLA4 was significantly distinguished from each other among geneClusters. The expression of CTLA4 was highest in geneCluster A (TCGA + GSE68465) or in geneCluster B (Meta-cohort), and it was lowest expressed in geneCluster C (Fig. 5).

Construction of geneCluster-related genescore and exploration of its clinical and genetic relevance

The analyses above have revealed the impact of m6A-related phenotypes on tumor microenvironment in LUAD. To systematically assess such impacts better, we established a scoring system named genescore base on the geneCluster. The DEGs of geneCluster were screened out by the differential analysis. And inter-DEGs were obtained by Venn analysis (Supplementary File 4). To assess the prognosis of intersect DEGs, univariate Cox regression was performed sequently. Finally, the inter-DEGs with significant prognosis (Supplementary File 4) were extracted for PCA analysis to calculate the genescore. Survival analysis demonstrated that the high genescore group showed an increased survival than the low genescore group in TCGA + GSE68465, and the result in validation set was consistent with the training set (Fig. 6A, Supplementary File 5). Further analyses revealed that m6ACluster B exhibited the highest genescore, followed by m6ACluster C and m6ACluster A (Fig. 6B, upper panel). As for geneCluster in training set, geneCluster C exhibited the highest genescore, followed by geneCluster B and geneCluster A. And the result in validation set was similar with training set (Fig. 6B, lower panel). The clinical (age, gender, T, N, smoking, fustat and recurrence) and genetic features of high and low genescore patients were analyzed. It was noted that clinical data were obtained from TCGA + GSE68465 cohort and genetic data were obtained from TCGA cohort. The genescore in group T1-2 was higher than that in group T3-4. And group N0 had higher genescore than group N1-3. In addition, group Alive had higher genescore than group Dead. This suggested the genescore decreased with the clinical stage progression and indicated the increased risk in prognose. However, genescore did not vary significantly with age, gender, smoking or recurrence in LUAD patients. Specifically, KRAS mutated group had a higher level of genescore than WT patients and TP53 mutated group had a lower level of genescore than WT patients. However, no significant differences were observed in terms of EGFR and ALK mutations (Fig. 6C). And the attribute changes of individual patients can be clearly presented by an alluvial diagram (Fig. 6D).

GSVA analysis between genescore groups and functional analysis of DEGs

The GSVA results showed that the genescores displayed significant differences in the pathways enriched. In high and low genescore groups, 76 signature pathways were enriched in training set and 100 signature

pathways were enriched in validation set (Supplementary File 6). 46 intersect signature pathways (Fig. 7A) were screened out by Venn diagram and visualized by heatmap. Low genescore group was markedly enriched in immune-related pathways, displayed stronger activities of mismatch repair, cell cycle, DNA replication and P53 signaling pathway. High genescore group showed higher enrichment in metabolic activities (Fig. 7B). To better understand the difference between high and low genescore groups, DEGs were identified by differential analysis and 2017 inter-DEGs were screened from the two cohorts for functional analysis (Fig. 7A, Supplementary File 7). The results of GO analysis showed that the intersect DEGs were significantly enriched in immune functions, cell-cell adhesion functions, cytokine activity and chemokine activity (Fig. 7C, Supplementary File 7). The results of KEGG analysis showed that the intersect DEGs were significantly enriched in immune-related pathways and infection-related pathways (Fig. 7D, Supplementary File 7).

The association of genescore with TME

The proportion difference of 23 immune cell types was explored by ssGSEA analysis in high and low genescore group. In training set, most immune cell infiltration in low genescore group was significantly higher than that in high genescore group. The same result was verified with the verification set (Fig. 8A). To validate the practicability of genescore, Stromal Score, Immune Score, ESTIMATE Score and Tumor Purity were calculated according to the expression level of LUAD using the ESTIMATE algorithm. The violin plot showed that Stromal Score, Immune Score and ESTIMATE Score in low genescore group were higher than that in high genescore group, while Tumor Purity was lower than that in high genescore group (Fig. 8B). Differential analysis of previous 29 immune signatures was performed between high and low genescore group to systematically present the immune functions and immune response-related pathways. The heatmap showed immune functions and immune response-related pathways were more active in low genescore group (Fig. 8C). In addition, we investigated the expression of immune check-point genes in high and low genescore groups based on the difference of check-point genes visualized in genescore heatmap. The boxplot showed the expressions of many check-point genes were higher in low genescore group than that in high genescore group (Fig. 8D). And then we displayed the three LUAD-related focus immune check-point genes (PD-1, PD-L1 and CTLA4) alone (Fig. 9A). It was worth noting that the expression of PD-L1 was lack in GSE68465 dataset, we investigated it in TCGA dataset. In recent years, ICI treatment represented by PD-1/CTLA4 inhibitor has made a significant breakthrough in antitumor therapy. IPS as a newly identified predictor, was widely used and strongly recommended to evaluate the immune response. Our analysis revealed that ips-CTLA4-negative-PD1-positive and ips-CTLA4-positive-PD1-positive were significantly elevated in the low genescore group than that in high genescore group. While there were no significant difference association with ips-CTLA4-negative-PD1-negative or ips-CTLA4-positive-PD1-negative in the two genescore groups (Fig. 9B). The role of tumor burden mutation (TMB) in tumor immune microenvironment has been concerned and controversial. Next, we compared the TMB in different genescore groups. As shown in Fig. 9C, there were no significant difference of TMB between genescore groups. Moreover, there was also no correlation with TMB in different geneClusters (Fig. 9D). The prognostic impact on TMB was then evaluated by the survival analysis. Kaplan-Meier survival curves showed there was no significant difference of the prognostic

impact on TMB (Fig. 9E). To better understand the prognostic impact of genescore on TMB, the crosstalk between genescore and TMB on patients' survival was investigated. High genescore and high TMB expression group had better survival than high genescore and low TMB group, and low genescore and high TMB expression group was associated with better survival relative to those with low genescore and low TMB expression group (Fig. 9F).

Verification of genescore in different datasets from GEO and exploration the role of genescore in pan-cancer cohort

Since the GEO datasets from meta-cohort shared the same microarray sequencing platform GPL570, we chose the inter-DEGs from geneCluster of meta-cohort for the following analysis. Univariate Cox regression analysis was performed to figure out the inter-DEGs with significant prognosis in each cohort. And then the signature inter-DEGs were extracted for PCA analysis to calculate the genescore. LUAD-TCGA dataset along with two independent LUAD-GEO datasets were performed to verify the prognosis in different genescore groups by survival analysis. We next applied genescore into the pan-cancer cohort to look for the possibility of predicting prognosis. Of the 33 cancer types from TCGA, the results of genescore associated with the prognosis in 13 cancers were consistent with those in LUAD, including ACC, BRCA, CHOL, ESCA, GBM, HNSC, KIRC, LAML, LGG, MESO, READ, SKCM and UCEC (Fig. 10).

Prediction of chemotherapy and targeted therapy sensitivity between high and low genescore group patients

Besides immunotherapy, the treatment widely used in LUAD are chemotherapy and targeted therapy. It is important to identify the subgroup patients that could be more sensitive to certain drugs. Prediction of several commonly used drugs sensitivity was conducted in high and low genescore group patients. As shown in Fig. 11A and B, the high genescore group was more sensitive to Erlotinib, while the low genescore group was more sensitive to Cisplatin, Docetaxel, Gemcitabine, Paclitaxel and Vinblastine. The result was verified by the application of genescore into three independent cohort GSE72094, TCGA and GSE68465 (Fig. 11C, D and E). The results could provide guidance for clinicians to determinate personalized treatment strategies for patients with LUAD.

4 Discussion

Accumulating evidence indicates that m6A modification, the most abundant and predominant eukaryotic RNA modification, take a critically important part in various biological processes[4; 27]. These processes are realized by regulating the transcriptome, such as RNA splicing, nuclear export, stability, translation, and decay. Importantly, m6A modification is highlighted to be involved in cancer biology. In the research field of lung cancer, m6A modification of RNA as a biomarker can thus serve the dual purpose of timely screening and therapy[28]. Although the role of m6A-related patterns and its association with TME in lung cancer has been revealed[14; 29], there is much more to be explored.

A previous study of colon cancer revealed the mutations of m6A regulators occurred in 1 to 8% of patients[13]. Consistent to the low occurrence, we found that the mutations of m6A regulators in LUAD were quite low, ranging from 0 to 3%. Investigation of the CNV features of the 20 m6A regulators demonstrated that more than half were copy-number-gain-dominant. Moreover, most m6A regulators were differentially expressed in tumor tissues, and the interaction associations among which were close and complex. In addition, our study found that 12 m6A regulators had impact on OS in two different cohort, respectively, among which 7 m6A regulators had the same impact on OS. The above evidence suggested that m6A regulators may play an essential role in LUAD.

Constant attention has been focused on the role of Insulin-like growth factor binding proteins (IGFBPs) in lung cancer during recent years. It is widely known that IGFBPs play the functions of modulating the availability of unbound insulin-like growth factor (IGF) for interaction with IGF receptor (IGFR). And they modulate important biological processes, including cell proliferation, survival, migration, senescence, autophagy, and angiogenesis. Their actions have been implicated in growth, metabolism, cancer, stem cell maintenance and differentiation, and immune regulation[30]. In this paper, three study hotspot IGFBP genes (IGFBP1, IGFBP2 and IGFBP3)[31; 32; 33] with significant prognosis in lung cancer (supplementary) were chosen to participate in the construction of m6A-related classification.

Previous studies analyzing the features of m6A regulators in LUAD focused on the construction of m6A-related classification and the clinical prognostic model simply by m6A regulators[9; 29; 34; 35; 36; 37]. In our analysis, we integrated different independent studies with more than 1500 samples to investigate the shared key features of LUAD based on m6A regulators. Different from previous studies, m6A regulators, along with above IGFBPs were fitted into unsupervised consensus clustering to identify the heterogeneity of LUAD. However, individual samples classified by the m6ACluster were not well to distinguish from each other. To better understand the heterogeneity in LUAD impacted by m6A regulators, we screened the DEGs with significant prognosis from the three distinct m6AClusters to construct the new classification, i.e. geneCluster. Satisfyingly, individual samples classified by m6A-related geneCluster were distinguished from each other. Mounting studies in favor of the result of our research that m6A related patterns had critical important role in tumor immune microenvironment[11; 13; 14]. We next explore the interested immune checkpoint genes in the geneClusters and m6AClusters. CTLA4 classified by geneCluster was able to distinguish from each other well, while it was far from satisfaction for differentiating PD-1 or PD-L1 by any clusters. Finally, we constructed the scoring tool called genescore based on the DEGs with significant prognosis screened from geneClusters, which was able to differentiate PD-1, PD-L1 and CTLA4. Jinyang Li et al. reported that tumor cell-intrinsic factors shape the tumor immune microenvironment and influence the outcome of immunotherapy. T cell high tumors are associated with an increased abundance of PD-1 + CD8 + T cells and predict better response to immunotherapy[38]. In our study, patients in low genescore group bearing T cell high tumors exhibited improved survival, characterized by increased abundance of CD8 + T cells and PD-1; While patients in high genescore group bearing T cell low tumors exhibited poor survival, characterized by decreased abundance of CD8 + T cells and PD-1. Moreover, the results of ESTIMATE indicated patients in low genescore group preferred to have high degree of immune infiltration and low tumor purity. Conversely, patients in high genescore group

tended to have low degree of immune infiltration and high tumor purity. More importantly, patients in low genescore group scored higher on ips-CTLA4-negative-PD1-positive and ips-CTLA4-positive-PD1-positive than that in high genescore group. Thus, it was the tumor cell intrinsic DEGs which generated the genescore that shaped the immune TME and influenced the response to immunotherapy. Besides, it is worth mentioning that most of the DEGs were enriched in immune-related pathways. The viewpoint that immunotherapy-refractory tumors may utilize molecular mechanisms for immune suppression distinct from those observed in immunotherapy-sensitive malignancies to break the immunosuppression[38] inspired us to explore the molecular mechanisms between the different genescore groups. And then GSVA analysis was performed to explore the functional pathways in genescore groups. Low genescore group characterized by T cell high tumors was markedly displayed stronger activities of immunity, mismatch repair, cell cycle, DNA replication and P53 signaling pathway. While high genescore group characterized by T cell low tumors showed higher enrichment in metabolic activities, such as amino acid metabolism and lipid metabolism. The immune evasion phenomenon in patients with high genescore agreed with the view that the abnormal metabolism of lipids or amino acids in tumor cells will affect the metabolic level of immune cells and the expression of metabolism-related genes, leading to immune cell dysfunction or the increased expression of molecules that promote tumor proliferation and weaken anti-tumor immune effects, achieving immune evasion and weakening anti-tumor immunity[39]. We further applied the genescore model to the prediction of chemotherapy and targeted therapy response and found that there were significant differences in the IC50 of drugs between the high- and low-genescore groups. The high genescore group was more sensitive to Erlotinib, while the low genescore group was more sensitive to chemotherapeutics drugs. All heterogeneity features generated by genescore emphasized the importance of m6A modification in shaping the tumor immune microenvironment.

Limitations

Several inevitable limitations in our study should be concerned. Firstly, some data was missing in the integrated data from TCGA database and GEO database. Secondly, we should be aware of the possibility of studying biases due to utilize the infiltration of tumor immune cells obtained based on algorithms. Thirdly, larger prospective clinical trials are warranted to further verify the predictive ability of the genescore model in the future. In addition, due to the lack of clinical LUAD cohorts with ICI treatment, we did not evaluate the correlation between the genescore model and the response to immunotherapy.

5 Conclusion

In this study, multi-levels of unsupervised consensus clustering base on m6A regulators and IGFBPs were performed. More importantly, we constructed genescore to better distinguish expression of genes at the immune checkpoint sites and to predict the response to immunotherapy, chemotherapy and targeted therapy. Essentially, m6A modification are involved in shaping the heterogeneity and complexity of tumor microenvironment.

Declarations

Conflict of Interest

The authors declare that they have no competing interests.

Author Contributions

Conception and design: XM.Y. and LX.W.; methodology: XM.Y. and ZC.Z.; data analysis: XM.Y. and YF.C.; manuscript writing: XM.Y. and MM.H. All authors read and approved the final manuscript.

Funding

None.

Acknowledgments

All authors would like to genuinely thank the specimen donors and research groups for TCGA and GEO, which provided data for this collection.

References

- [1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 71 (2021) 209-249.
- [2] J.J. Lee, S. Park, H. Park, S. Kim, J. Lee, J. Lee, J. Youk, K. Yi, Y. An, I.K. Park, C.H. Kang, D.H. Chung, T.M. Kim, Y.K. Jeon, D. Hong, P.J. Park, Y.S. Ju, and Y.T. Kim, Tracing Oncogene Rearrangements in the Mutational History of Lung Adenocarcinoma. *Cell* 177 (2019) 1842-1857 e21.
- [3] C. Zappa, and S.A. Mousa, Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res* 5 (2016) 288-300.
- [4] S. Zaccara, R.J. Ries, and S.R. Jaffrey, Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* 20 (2019) 608-624.
- [5] X.Y. Chen, J. Zhang, and J.S. Zhu, The role of m(6)A RNA methylation in human cancer. *Mol Cancer* 18 (2019) 103.
- [6] Y. Yang, P.J. Hsu, Y.S. Chen, and Y.G. Yang, Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell Res* 28 (2018) 616-624.
- [7] G. Yang, Z. Sun, and N. Zhang, Reshaping the role of m6A modification in cancer transcriptome: a review. *Cancer Cell Int* 20 (2020) 353.
- [8] R.I.N. Khan, and W.A. Malla, m(6)A modification of RNA and its role in cancer, with a special focus on lung cancer. *Genomics* 113 (2021) 2860-2869.

- [9] F. Li, H. Wang, H. Huang, L. Zhang, D. Wang, and Y. Wan, m6A RNA Methylation Regulators Participate in the Malignant Progression and Have Clinical Prognostic Value in Lung Adenocarcinoma. *Front Genet* 11 (2020) 994.
- [10] L. Yang, S. Wu, C. Ma, S. Song, F. Jin, Y. Niu, and W.M. Tong, RNA m(6)A Methylation Regulators Subclassify Luminal Subtype in Breast Cancer. *Front Oncol* 10 (2020) 611191.
- [11] J. Li, W. Wang, Y. Zhou, L. Liu, G. Zhang, K. Guan, X. Cui, X. Liu, M. Huang, G. Cui, and R. Sun, m6A Regulator-Associated Modification Patterns and Immune Infiltration of the Tumor Microenvironment in Hepatocarcinoma. *Front Cell Dev Biol* 9 (2021) 687756.
- [12] Y. Du, Y. Ma, Q. Zhu, T. Liu, Y. Jiao, P. Yuan, and X. Wang, An m6A-Related Prognostic Biomarker Associated With the Hepatocellular Carcinoma Immune Microenvironment. *Front Pharmacol* 12 (2021) 707930.
- [13] W. Chong, L. Shang, J. Liu, Z. Fang, F. Du, H. Wu, Y. Liu, Z. Wang, Y. Chen, S. Jia, L. Chen, L. Li, and H. Chen, m(6)A regulator-based methylation modification patterns characterized by distinct tumor microenvironment immune profiles in colon cancer. *Theranostics* 11 (2021) 2201-2217.
- [14] H. Zhou, M. Zheng, M. Shi, J. Wang, Z. Huang, H. Zhang, Y. Zhou, and J. Shi, Characteristic of molecular subtypes in lung adenocarcinoma based on m6A RNA methylation modification and immune microenvironment. *BMC Cancer* 21 (2021) 938.
- [15] Y. Jin, Z. Wang, D. He, Y. Zhu, X. Hu, L. Gong, M. Xiao, X. Chen, Y. Cheng, and K. Cao, Analysis of m6A-Related Signatures in the Tumor Immune Microenvironment and Identification of Clinical Prognostic Regulators in Adrenocortical Carcinoma. *Front Immunol* 12 (2021) 637933.
- [16] J. Zhong, Z. Liu, C. Cai, X. Duan, T. Deng, and G. Zeng, m(6)A modification patterns and tumor immune landscape in clear cell renal carcinoma. *J Immunother Cancer* 9 (2021).
- [17] W. Dai, Y. Li, S. Mo, Y. Feng, L. Zhang, Y. Xu, Q. Li, and G. Cai, A robust gene signature for the prediction of early relapse in stage III colon cancer. *Mol Oncol* 12 (2018) 463-475.
- [18] S. Liu, Q. Li, K. Chen, Q. Zhang, G. Li, L. Zhuo, B. Zhai, X. Sui, X. Hu, and T. Xie, The emerging molecular mechanism of m(6)A modulators in tumorigenesis and cancer progression. *Biomed Pharmacother* 127 (2020) 110098.
- [19] A.E. Arguello, A.N. DeLiberto, and R.E. Kleiner, RNA Chemical Proteomics Reveals the N(6)-Methyladenosine (m(6)A)-Regulated Protein-RNA Interactome. *J Am Chem Soc* 139 (2017) 17249-17252.
- [20] Y. Li, J. Xiao, J. Bai, Y. Tian, Y. Qu, X. Chen, Q. Wang, X. Li, Y. Zhang, and J. Xu, Molecular characterization and clinical relevance of m(6)A regulators across 33 cancer types. *Mol Cancer* 18 (2019) 137.

- [21] Z. Zhou, J. Lv, H. Yu, J. Han, X. Yang, D. Feng, Q. Wu, B. Yuan, Q. Lu, and H. Yang, Mechanism of RNA modification N6-methyladenosine in human cancer. *Mol Cancer* 19 (2020) 104.
- [22] Y. Niu, X. Zhao, Y.S. Wu, M.M. Li, X.J. Wang, and Y.G. Yang, N6-methyl-adenosine (m6A) in RNA: an old modification with a novel epigenetic function. *Genomics Proteomics Bioinformatics* 11 (2013) 8-17.
- [23] M.D. Wilkerson, and D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (2010) 1572-3.
- [24] Y. He, Z. Jiang, C. Chen, and X. Wang, Classification of triple-negative breast cancers based on Immunogenomic profiling. *J Exp Clin Cancer Res* 37 (2018) 327.
- [25] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102 (2005) 15545-50.
- [26] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, H. Hackl, and Z. Trajanoski, Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* 18 (2017) 248-262.
- [27] J. Chen, C. Wang, W. Fei, X. Fang, and X. Hu, Epitranscriptomic m6A modification in the stem cell field and its effects on cell death and survival. *Am J Cancer Res* 9 (2019) 752-764.
- [28] L.P. Vu, B.F. Pickering, Y. Cheng, S. Zaccara, D. Nguyen, G. Minuesa, T. Chou, A. Chow, Y. Saleture, M. MacKay, J. Schulman, C. Famulare, M. Patel, V.M. Klimek, F.E. Garrett-Bakelman, A. Melnick, M. Carroll, C.E. Mason, S.R. Jaffrey, and M.G. Kharas, The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat Med* 23 (2017) 1369-1376.
- [29] Y. Li, J. Gu, F. Xu, Q. Zhu, Y. Chen, D. Ge, and C. Lu, Molecular characterization, biological function, tumor microenvironment association and clinical significance of m6A regulators in lung adenocarcinoma. *Brief Bioinform* 22 (2021).
- [30] L.A. Bach, What Happened to the IGF Binding Proteins? *Endocrinology* 159 (2018) 570-578.
- [31] F. Zheng, Q. Tang, X.H. Zheng, J. Wu, H. Huang, H. Zhang, and S.S. Hann, Inactivation of Stat3 and crosstalk of miRNA155-5p and FOXO3a contribute to the induction of IGFBP1 expression by beta-elemene in human lung cancer. *Exp Mol Med* 50 (2018) 1-14.
- [32] J. Dong, Y. Zeng, P. Zhang, C. Li, Y. Chen, Y. Li, and K. Wang, Serum IGFBP2 Level Is a New Candidate Biomarker of Severe Malnutrition in Advanced Lung Cancer. *Nutr Cancer* 72 (2020) 858-863.

- [33] F. Tas, E. Bilgin, D. Tastekin, K. Erturk, and D. Duranyildiz, Serum IGF-1 and IGFBP-3 levels as clinical markers for patients with lung cancer. *Biomed Rep* 4 (2016) 609-614.
- [34] X. Li, Z. Feng, R. Wang, J. Hu, X. He, and Z. Shen, Expression Status and Prognostic Value of m(6)A RNA Methylation Regulators in Lung Adenocarcinoma. *Life (Basel)* 11 (2021).
- [35] X. Wu, H. Sheng, L. Wang, P. Xia, Y. Wang, L. Yu, W. Lv, and J. Hu, A five-m6A regulatory gene signature is a prognostic biomarker in lung adenocarcinoma patients. *Aging (Albany NY)* 13 (2021) 10034-10057.
- [36] H. Wang, X. Zhao, and Z. Lu, m(6)A RNA Methylation Regulators Act as Potential Prognostic Biomarkers in Lung Adenocarcinoma. *Front Genet* 12 (2021) 622233.
- [37] Z. Zhuang, L. Chen, Y. Mao, Q. Zheng, H. Li, Y. Huang, Z. Hu, and Y. Jin, Diagnostic, progressive and prognostic performance of m(6)A methylation RNA regulators in lung adenocarcinoma. *Int J Biol Sci* 16 (2020) 1785-1797.
- [38] J. Li, K.T. Byrne, F. Yan, T. Yamazoe, Z. Chen, T. Baslan, L.P. Richman, J.H. Lin, Y.H. Sun, A.J. Rech, D. Balli, C.A. Hay, Y. Sela, A.J. Merrell, S.M. Liudahl, N. Gordon, R.J. Norgard, S. Yuan, S. Yu, T. Chao, S. Ye, T.S.K. Eisinger-Mathason, R.B. Faryabi, J.W. Tobias, S.W. Lowe, L.M. Coussens, E.J. Wherry, R.H. Vonderheide, and B.Z. Stanger, Tumor Cell-Intrinsic Factors Underlie Heterogeneity of Immune Cell Infiltration and Response to Immunotherapy. *Immunity* 49 (2018) 178-193 e7.
- [39] F. Wu, Y. Cheng, L. Wu, W. Zhang, W. Zheng, Q. Wang, H. Cao, X. Pan, and W. Tang, Emerging Landscapes of Tumor Immunity and Metabolism. *Front Oncol* 10 (2020) 575037.

Figures

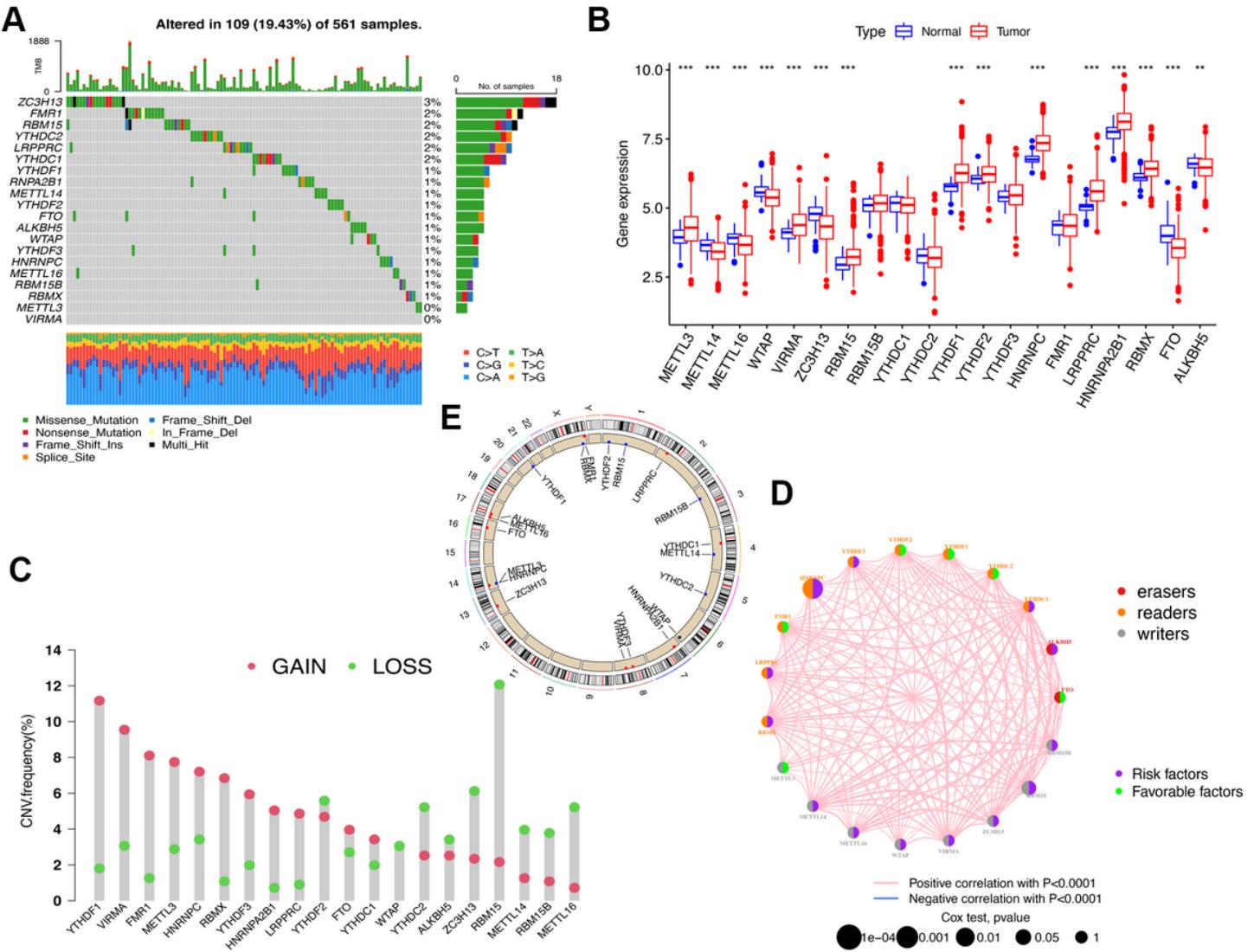


Figure 1

Features of m6A regulators based on the TCGA LUAD cohort. (A) Mutations of m6A regulators. (B) Comparison of the gene expression level of m6A regulators in normal and tumor samples in TCGA LUAD cohort ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$). (C) The CNV variation frequency of 20 m6A regulators. (D) Interaction associations among 20 m6A regulators are visualized in the network and the relationship between m6A regulators and prognosis. The size of each circle represented survival impact of each m6A regulator. Green (purple) dots represented favorable (risk) factors of overall survival. The lines connecting m6A regulators presented their interactions, and thickness of the lines represented the correlation strength among regulators. Positive (negative) correlation was indicated in red (blue) ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$). (E) The location of CNV alteration of m6A regulators on 23 chromosomes.

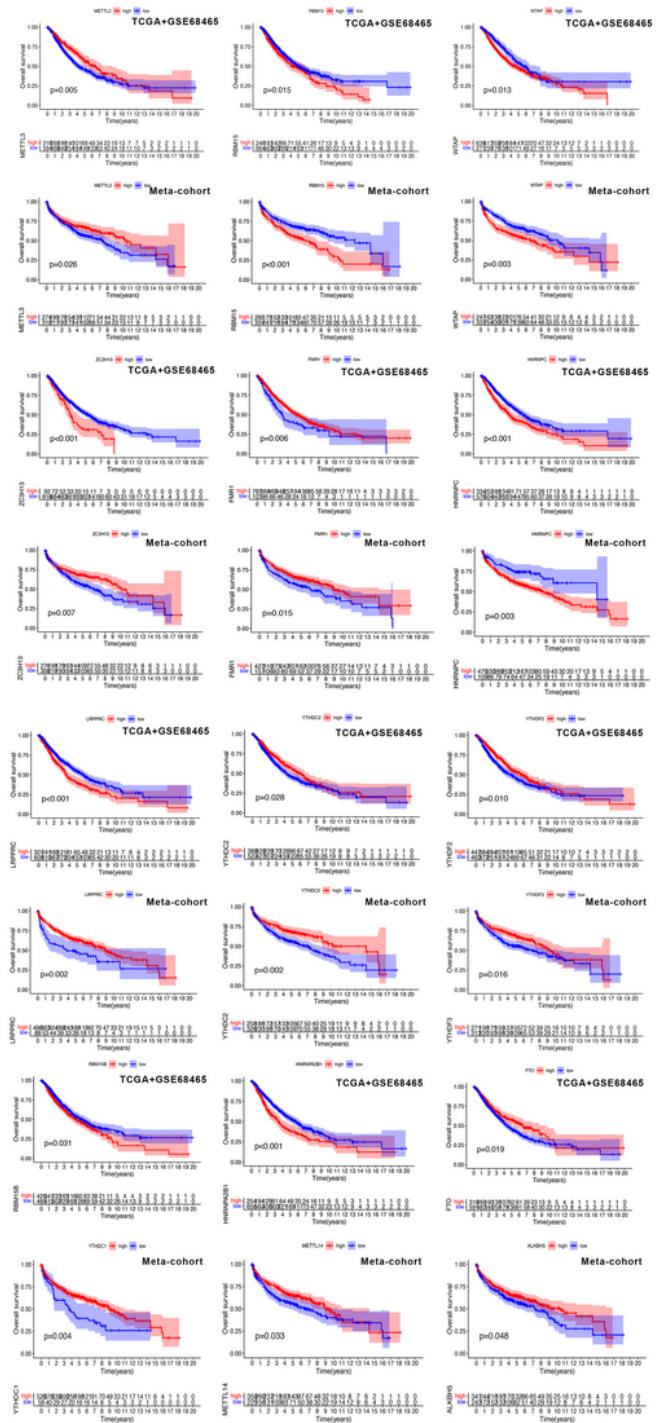


Figure 2

Survival analyses for m6A regulators in LUAD using Kaplan-Meier curves based on the TCGA+GSE68465 cohort and Meta-cohort. The Meta-cohort composition (GSE19188, GSE30219, GSE31210, GSE37745, and GSE50081) were used as sample annotations.

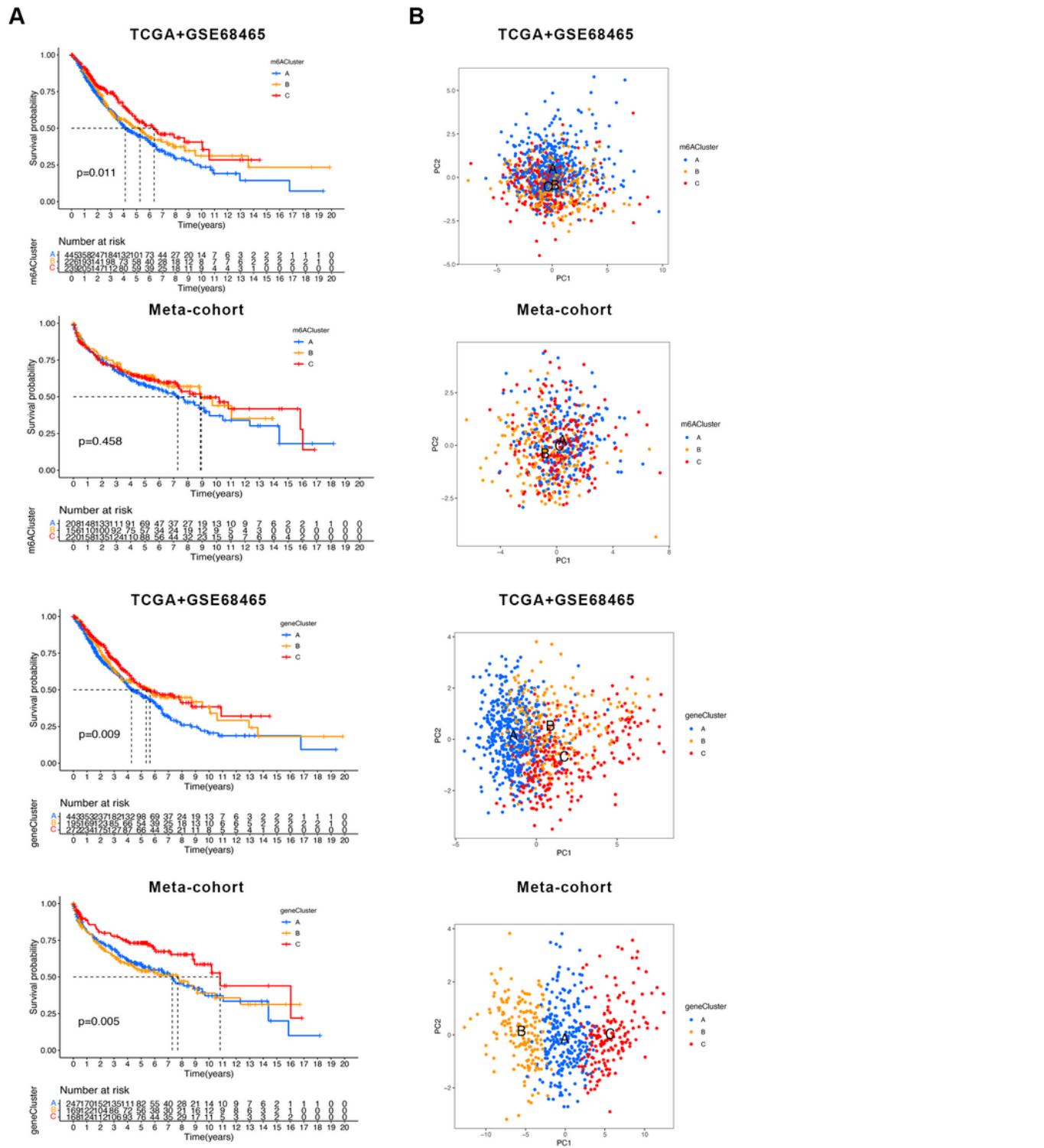


Figure 3

Construction of m6ACluster and geneCluster. (A) Survival analyses for Cluster patterns in LUAD using Kaplan-Meier curves based on the TCGA+GSE68465 cohort and Meta-cohort. (B) Principal component analysis (PCA) of TCGA+GSE68465 cohort and Meta-cohort RNA-sequence profiles for the Clusters.

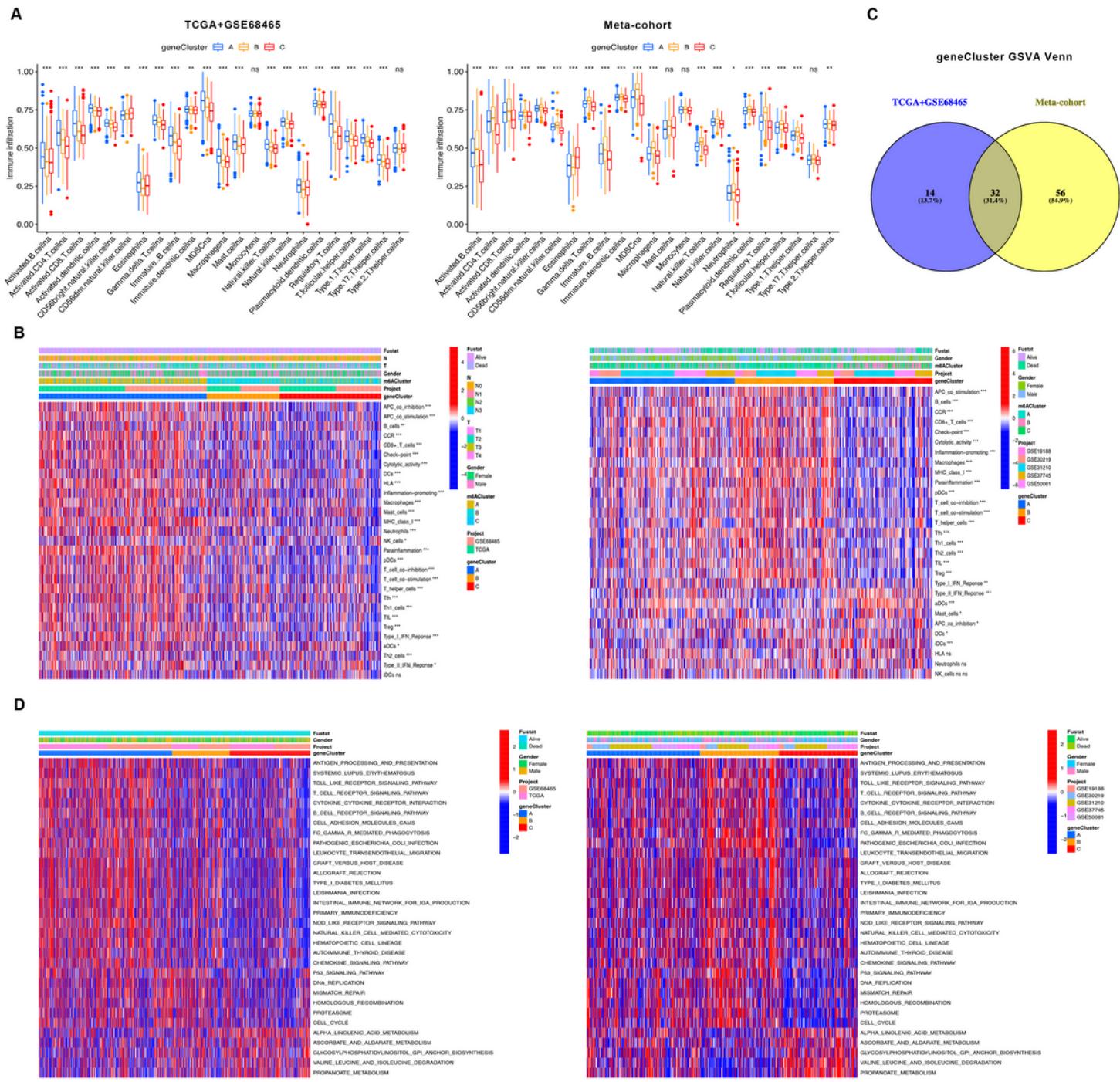


Figure 4

TME immune cells infiltration, immune function and immune response-related pathways and GSVA. (A) Comparison of the abundance of infiltrating immune cells among geneClusters. (B) Immune infiltration, immune function and immune response-related pathways of geneCluster from TCGA+GSE68465 cohort and Meta-cohort were shown in heatmap. (C) 32 intersection GSVA of geneCluster from TCGA+GSE68465 cohort and Meta-cohort were shown in the Venn diagram. (D) Heatmap showing the

inter 32 biological pathway differences among geneClusters based on the TCGA+GSE68465 cohort and Meta-cohort (adjusted P-value <0.05). (*P < 0.05; **P < 0.01; ***P < 0.001; ns, no sense)

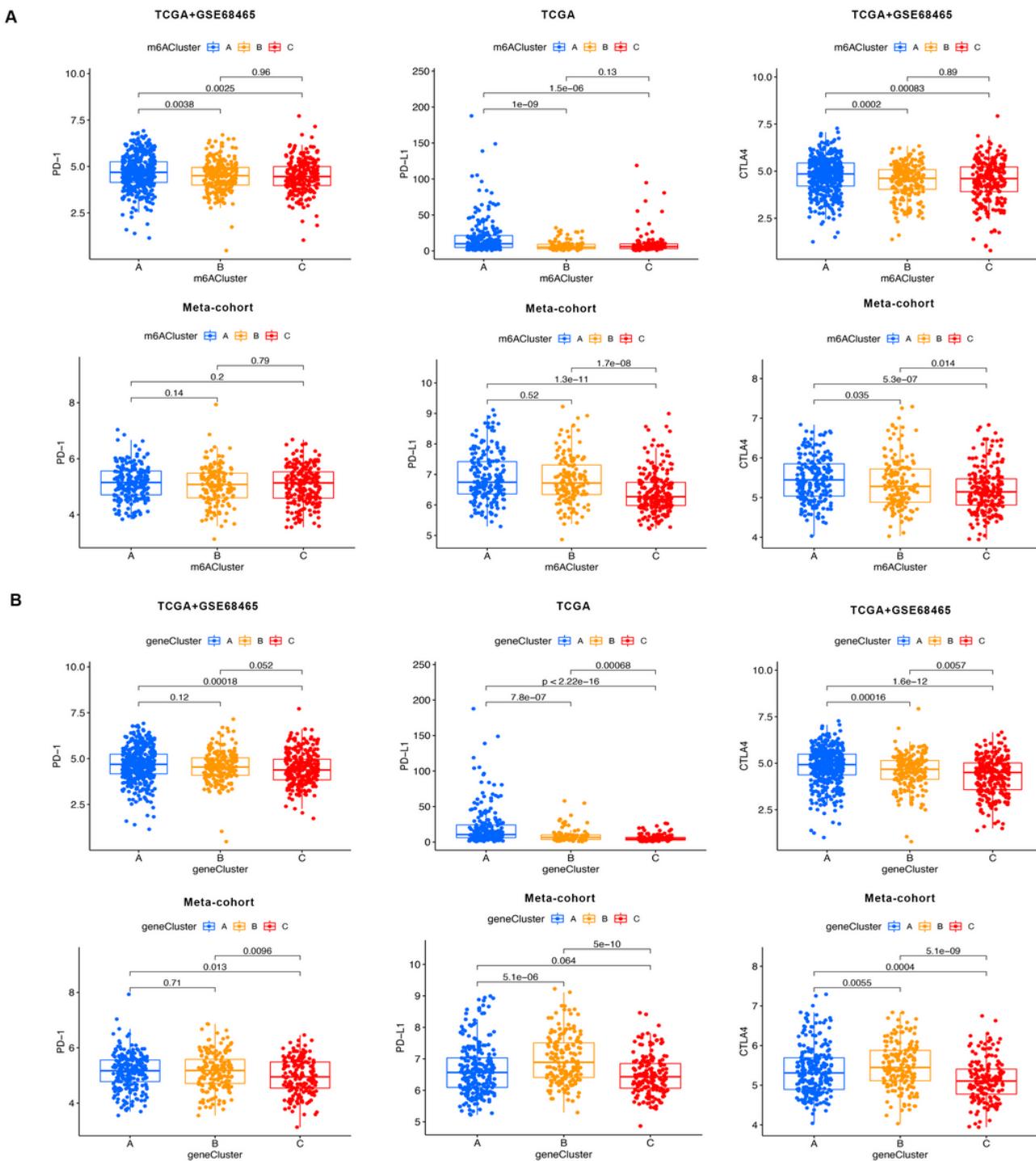


Figure 5

Differences in expression of PD-1, PD-L1 and CTLA4 among Clusters (A) Differences in expression of PD-1, PD-L1 and CTLA4 among three m6AClusters in the TCGA+GSE68465 cohort and Meta-cohort. (B)

Differences in expression of PD-1, PD-L1 and CTLA4 among three geneClusters in the TCGA+GSE68465 cohort and Meta-cohort.

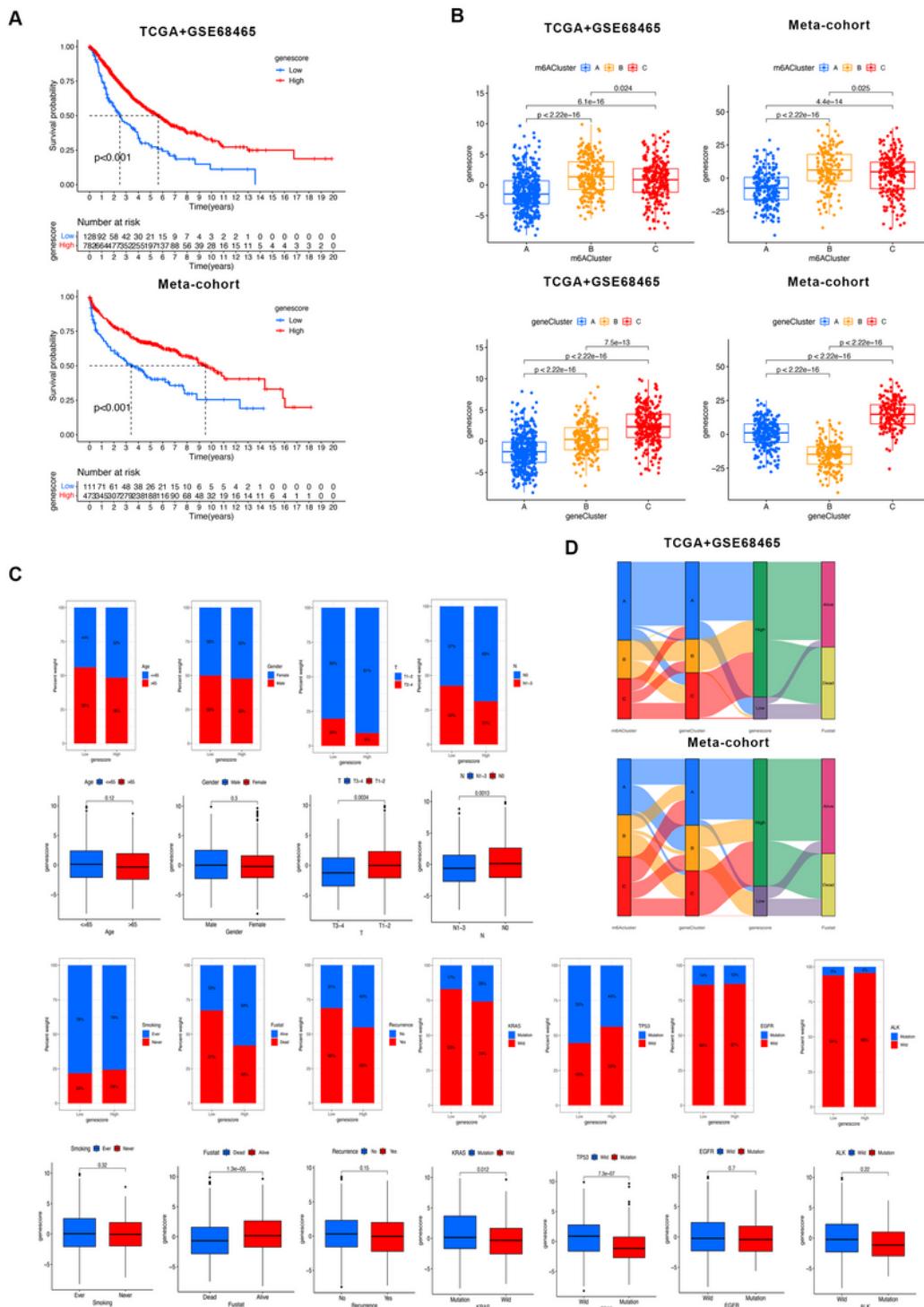


Figure 6

Construction of genescore and comparison of the clinical characteristics. (A) Kaplan-Meier curves for high and low genescore patient groups in the TCGA+GSE68465 cohort and Meta-cohort. (B) Differences

in genescore among three m6AClusters and three geneClusters in the TCGA+GSE68465 cohort and Meta-cohort. (C) The fraction and comparison of patients with clinical characteristics and hotspot mutation genes between high and low genescore groups based on the TCGA+GSE68465 cohort. (D) Alluvial diagram of the changes of m6AClusters, geneClusters, genescores, and respond to fustat.

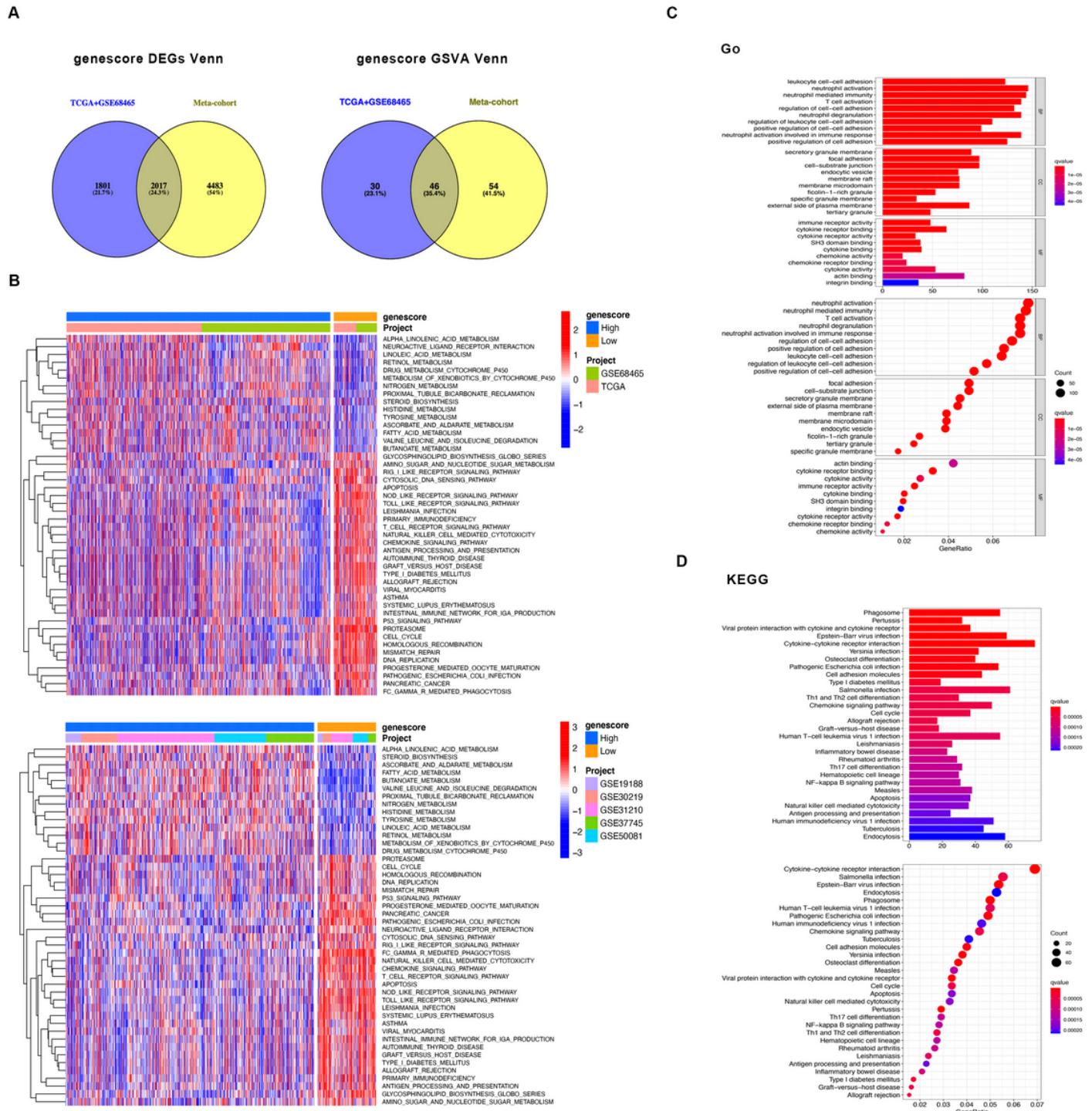


Figure 7

GSVA analysis of genescore and function analysis of DEGs identified from genescore. (A) 2017 intersection DEGs and 46 intersection GSVA of genescore from TCGA+GSE68465 cohort and Meta-cohort were shown in the Venn diagram. (B) Heatmap showing the inter 46 biological pathway differences between high and low genescore based on the TCGA+GSE68465 cohort and Meta-cohort (adjusted P-value <0.05). (C) Gene ontology enrichment analysis of the 2017 intersection DEGs, BP biology process; CC cellular component; MF molecular function. (D) KEGG pathway enrichment analysis of the 2017 intersection DEGs.

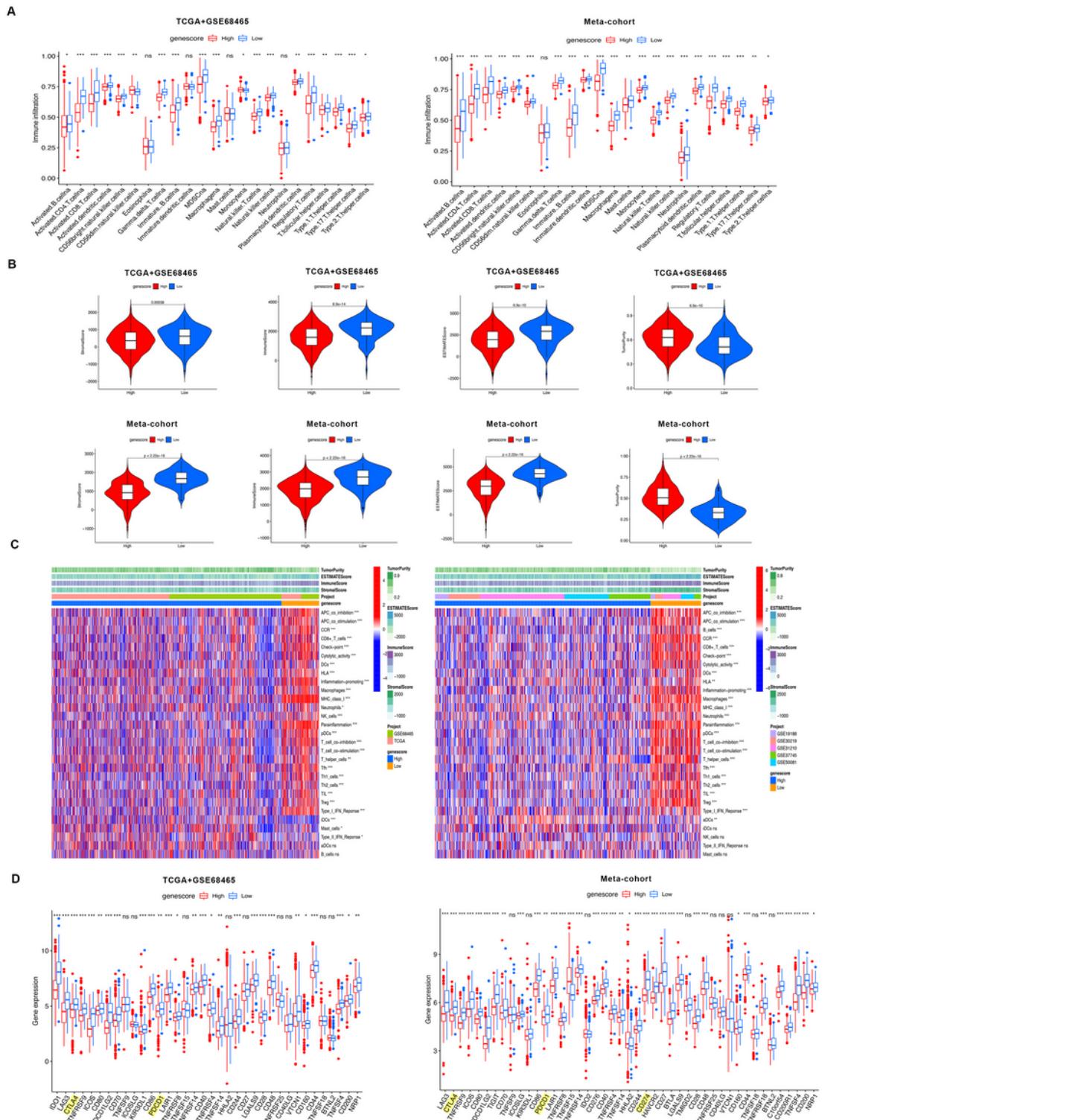
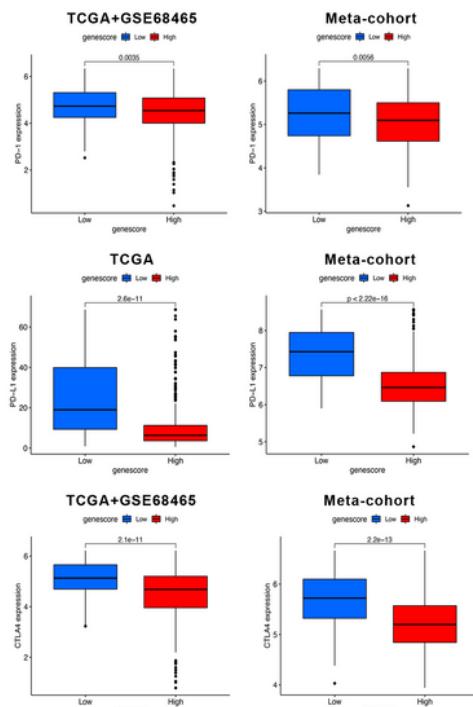


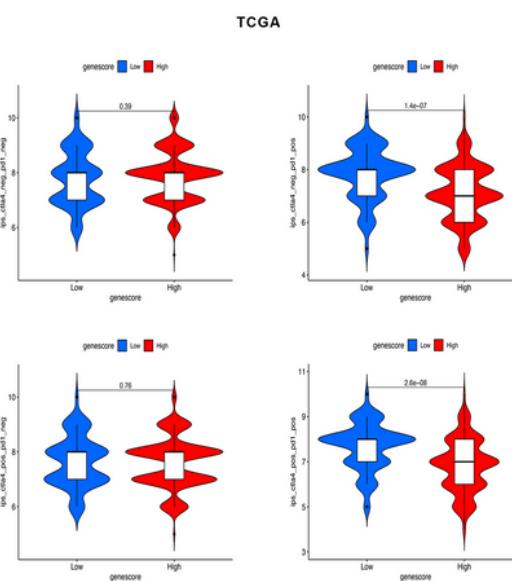
Figure 8

Features of genescore in tumor immune microenvironment and immune checkpoint sites (A) Comparison of the abundance of infiltrating immune cells between high and low genescore groups from TCGA+GSE68465 cohort and Meta-cohort. (B) Discrepancy of stromalscore, immunescore, estimatescore and tumorpurity in high and low genescore groups from TCGA+GSE68465 cohort and Meta-cohort. (C) Immune infiltration, immune function and immune response-related pathways of genescore from TCGA+GSE68465 cohort and Meta-cohort were shown in heatmap. (D) Comparison of the expression of genes at immune checkpoint sites among m6A clusters, and the interested genes were shown in yellow. (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, no sense)

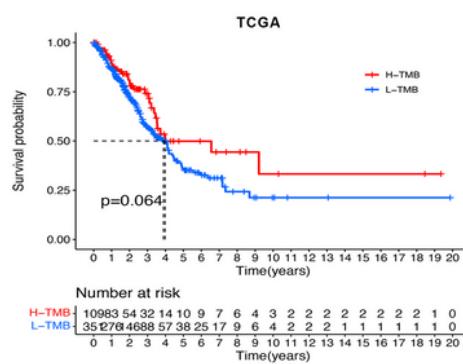
A



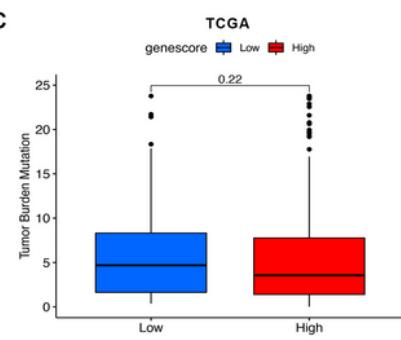
B



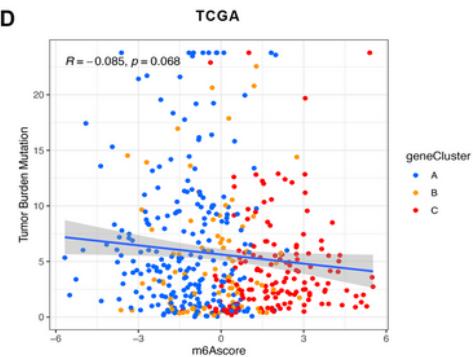
E



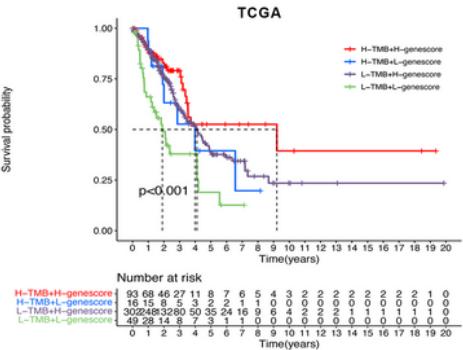
C



D



F

**Figure 9**

Immunotherapy gene expression, IPS and TMB analysis. (A) The genescore was associated with the expression of three immune checkpoints (PD-1, PD-L1 and CTLA4) from TCGA+GSE68465 cohort and Meta-cohort. (B) The association between IPS and the genescore in LUAD patients based on TCGA database. (C) Relative distribution of TMB in high and low genescore groups in the TCGA-LUAD cohort. (D) The correlation between the geneCluster and the TMB in the TCGA-LUAD cohort. (E) Kaplan-Meier

curves for high and low TMB patient groups in the TCGA-LUAD cohort. (F) Survival analysis among four patient groups stratified by the genescore and TMB in the TCGA-LUAD cohort.

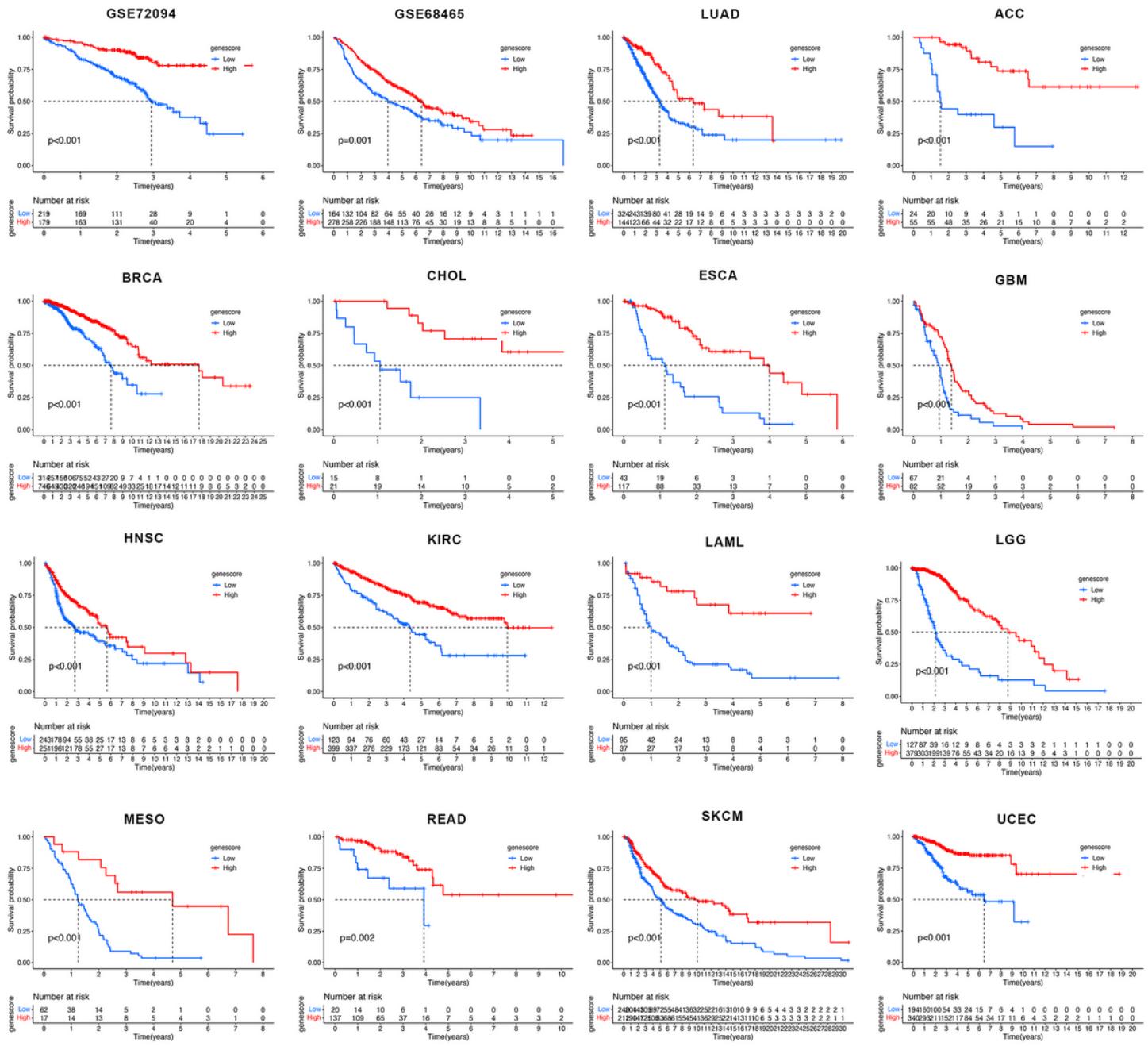
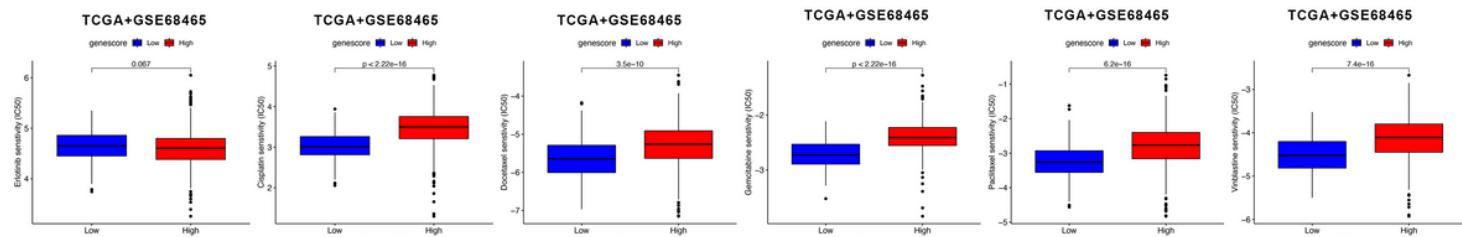
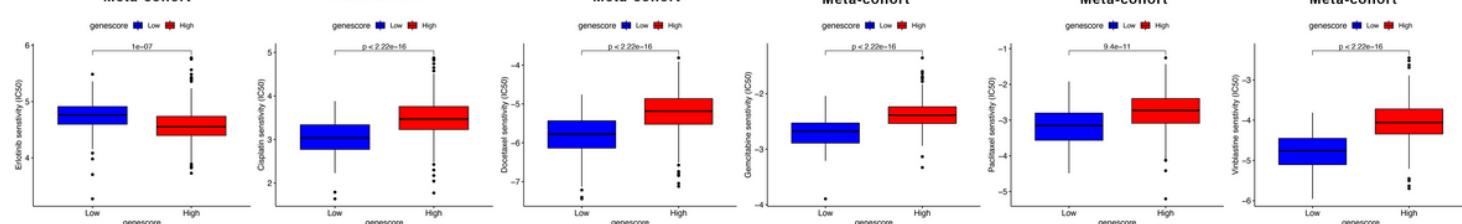
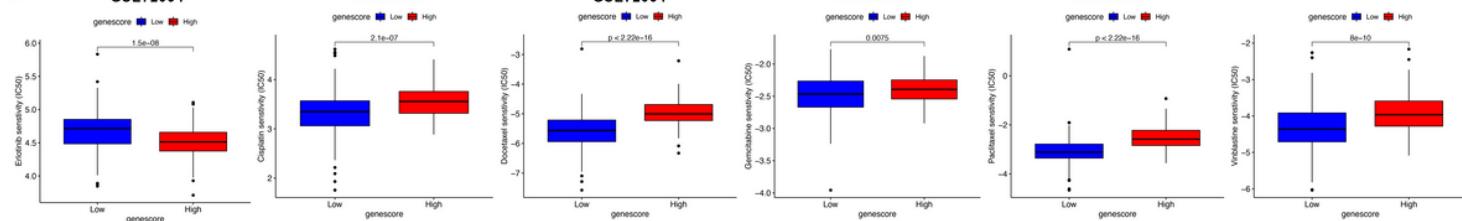
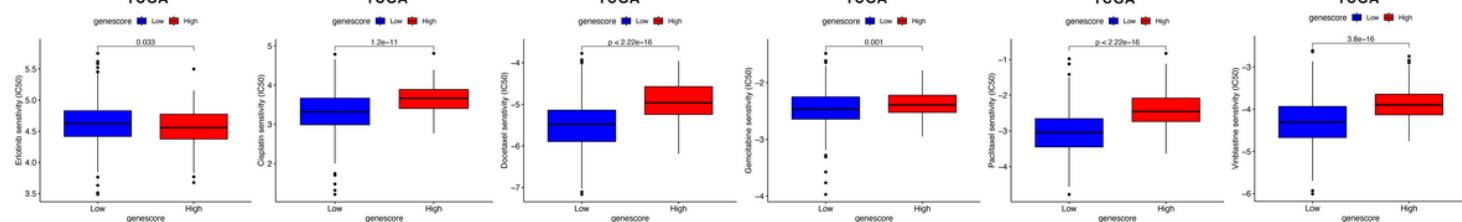
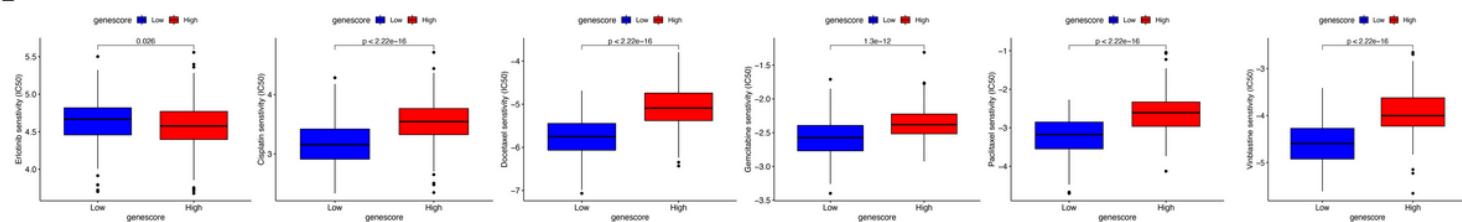


Figure 10

The role of genescore in the GSE72094 cohort, GSE68465 cohort and the pan-cancer cohort.

A**B****C****D****E****Figure 11**

Drug sensitivities comparison between high and low genescore groups. Distribution of the estimated IC₅₀ of Erlotinib, Cisplatin, Docetaxel, Gemcitabine, Paclitaxel and Vinblastine between high and low genescore group in the TCGA+GSE68465 cohort (A) and Meta-cohort (B). Validation of genescore in the GSE72094 cohort (C), TCGA LUAD cohort (D) and GSE68465 cohort(E).

Supplementary Files

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

- [SupplementaryFigure1.tif](#)
- [SupplementaryFile1.rar](#)
- [SupplementaryFile2.rar](#)
- [SupplementaryFile3.rar](#)
- [SupplementaryFile4.rar](#)
- [SupplementaryFile5.rar](#)
- [SupplementaryFile6.rar](#)
- [SupplementaryFile7.rar](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)