A Risk Signature Established From Three Coagulation-fibrinolysis Genes Predicts Prognosis in Digestive System Pancancer

Xingyun Wang
Mudanjiang Medical University
Jinli Ji
Mudanjiang Medical University
Ying Jiang
Mudanjiang Medical University
Yiyang Zhao
Southern Medical University
Zheyao Song
Mudanjiang Medical University
Zikang He
Mudanjiang Medical University
Ping Shen
Mudanjiang Medical University
Huan Wang
Mudanjiang Medical University
Liangyu Luo
Hongqi Hospital affiliated to Mudanjiang Medical College
Huilin Guan
Mudanjiang Medical University
Rongjun Cui (cuirongjun@mdjmu.edu.cn)
Mudanjiang Medical University

Research Article

Keywords: VTE, KEGG, YTHDC1, YTHDC2, immune microenvironment components

Posted Date: January 10th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1157925/v1

Abstract

Venous thromboembolism (VTE) is one of the major complications of digestive system cancer, and coagulation-fibrinolysis genes play an important role in VTE. We used univariate Cox analysis, least absolute shrinkage and selection operator (LASSO), and multivariate Cox analysis to construct 3-PCFGs (prognostic coagulation-fibrinolysis genes) model based on six prognostic coagulation-fibrinolysis genes. Gene set enrichment analysis (GSEA) was used to analyze the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the high- and low-risk groups. In addition, we classified digestive system pancancer patients into three clusters A, B, and C based on 3-PCFGs by K means. High-risk group and cluster C were associated with poor prognosis in digestive system pancancer. The m^6A-related genes ALKBH5, FTO, RBM15, YTHDC1, and YTHDC2 ($P<0.001$) were highly expressed in the high-risk group and cluster C. The risk score was positively correlated with cancer-associated fibroblasts and endothelial cells. Cluster C had the highest immune score and stromal score. The poor prognosis in the high-risk group and cluster C may be affected by m^6A epigenetic modification and immune microenvironment components in the digestive system pancancer.

Introduction

The digestive system pancancer mainly includes esophageal cancer, gastric cancer, colorectal cancer, liver cancer, and pancreatic cancer. According to the analysis of global cancer statistical reports in 2020, the number of new cases of colorectal cancer, gastric cancer and liver cancer ranked 3, 5 and 6, and the number of deaths ranked 2, 4 and 3, respectively[1]. It is enough to see that digestive system cancer still endangers human life and health.

Venous thromboembolism (VTE) is one of the causes of patient death. It is a common complication of digestive system cancer patients, including colorectal cancer[2, 3] and pancreatic adenocarcinoma[4]patients. The pathogenesis of VTE is bleeding, thrombosis, and vascular occlusion, which occur in both the tumor capillary network and the peripheral vasculature[5]. There was a report that patients with cancer and VTE, cancer alone, or venous thromboembolism alone have a 30 times, 7 times, and 3 times higher risk of death than the control group respectively[6]. The causes of thrombosis in different cancers and the molecules involved are similar to the traditional coagulation mechanism in the peripheral circulatory system. Therefore, the same diagnostic methods, clinical analysis, and preventive and therapeutic countermeasures can be applied to different cancer thromboses[5, 7]. However, there is a lack of markers in the digestive system pancancer. With the continuous maturation of sequencing technology, the Cancer Genome Atlas (TCGA) program database provides a rare convenience for the study of pancancer[8, 9].

Epigenetic modifications have played a well-documented role in the dysregulation of genes controlling tumor angiogenesis or coagulant phenotypes. N^6-methyladenosine (m^6A) is a methylation modification occurring at the sixth nitrogen (N) atom of adenine (A)and is the most abundant RNA modification in eukaryotes[10]. M^6A methylation modification is a dynamically reversible modification process

that is mainly regulated by m^6A methyltransferase complexes (writers) such as METTL3, m^6A demethylases (erasers) such as ALKBH5, and m^6A binding proteins (readers) such as YTHDC1[11]. Recently, it has been shown that m^6A plays an important role in the tumor immune microenvironment (TIME), providing new ideas for clinical immunotherapy[12, 13]. However, the m^6A level and TIME in digestive system pancancer patients with a high incidence of venous thrombosis remain ambiguous.

In this study, we sought to construct prognostic markers based on six coagulation-fibrinolysis genes (F3, PLAT, PLAU, PLAUR, SERPINE1 and SERPINB2) for personalized treatment of patients with digestive system pancancer. Our results strongly suggest that the correlation among prognostic models, m^6A regulators and TIME components provides a new perspective for immunotherapy.

Materials And Methods

Data source and arrangement

Gene expression RNAseq data as well as clinical data (including phenotype, survival data and somatic mutation) of digestive system pancancer were downloaded by UCSC at https://xenabrowser.net/datapages/. Digestive system pancancer types and samples were as follows: colorectal adenocarcinoma (COAD, T=471), esophageal adenocarcinoma (ESCA, T=162), liver hepatocellular carcinoma (LIHC, T=374), pancreatic adenocarcinoma (PAAD, T=178), rectal adenocarcinoma (READ, T=167) and stomach adenocarcinoma (STAD, T=375). Among them, there were a total of 1662 samples with survival data. Copy number variation (CNV) was download at https://gdc.cancer.gov/about-data/publications/panimmune. Tumor mutational burden (TMB) was obtained based on somatic mutation (VarScan2 Variant Aggregation and Masking GDC hub) by Perl.

Construction of a prognostic coagulation-fibrinolysis gene (PCFG) risk-score model

The PCFGs were selected by univariate Cox regression analysis. To prevent model overfitting, 3-PCFGs were screened by least absolute shrinkage and selection operator (LASSO)[14]. Next, a scoring model was constructed based on prognostically relevant CFGs by univariate Cox regression analysis and multivariate Cox regression. The scoring model is shown as follows (β “regression coefficient”)[15]:

risk score

Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (tSNE) were used for effective dimensionality reduction, model identification, and grouping visualization of high-dimensional data of 3-PCFG expression profiles.

Independence of the 3-PCFGs risk-score model

Multivariate Cox regression and univariate Cox regression analyses were conducted to test whether the prognostic pattern was an independent variable considering other clinical characteristics (age, TNM stage,T stage, N stage, M stage, and gender) with the patients of digestive system pancancer[16].

Establishing and proving a predictive nomogram

The predictive ability of the nomogram and other predictors (age, gender, TNM stage, T stage, N stage, M stage, and risk score) for the 1-, 2-, and 3-year OS was set up by the "rms" package (https://CRAN.R-project.org/package=rms). Calibration curves were applied to illustrate the uniformity between the practical outcome and model prediction outcome.

Gene Set Enrichment Analysis

To explore the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the 3-PCFG scoring model, a gene set enrichment analysis (GSEA Version 4.1.0)[17] was used to pursue the enrichment terms in the digestive system pancancer with the filtration criteria of NOM p-val < 0.05.

Immune-related analysis

The immune microenvironment scores of digestive system pancancer were obtained by https://bioinformatics.mdanderson.org/estimate/. The quantified immune cell matrix is given at TIMER 2.0 web tool (http://timer.cistrome.org)[18]. Immune cell abundance data (calculated using the CIBERSORT algorithm[19]) were downloaded by https://gdc.cancer.gov/about-data/publications/panimmune.

K means cluster analysis

To find new classification criteria, we performed cluster analysis based on 3-PCFG expression in digestive system pancancer by the "ConsensusClusterPlus" R package[20].

Statistical Analysis

The statistical analyses were performed by R software (version 4.1.1) (https://www.R-project.org/) with R studio (version 1.4.1717) (http://www.rstudio.com/).

Ethics declarations

UCSC, which patients involved in the database have obtained ethical approval, belongs to public databases.All methods were carried out in accordance with relevant guidelines and regulation.

Results

Risk model constuction and validation based on CFGs

Our workflow is shown in Fig.1a. All six CFGs were associated with the prognosis of digestive system pancancer (P□0.05) (Table 1). Four CFGs were identified by a LASSO regression analysis (Fig. 1b-c). Furthermore, 3-PCFGs including PLAT, SERPINB2 and SERPINB1, were identified by univariate and multivariate Cox regression analysis (Additional file 1:Fig. S1a and Fig. 1d). The model was calculated as follows: risk score=0.046×the expression of PLAU+0.069×the expression of SERPINB2+0.185×the expression of SERPINB1.

Table 1 Univariate Cox regression analysis for six CFGs.

Gene	HR	HR.95L	HR.95H	P value
F3	1.098	1.047	1.152	1.13E-04
PLAU	1.111	1.053	1.171	1.01E-04
PLAUR	1.089	1.030	1.152	2.90E-03
PLAT	1.160	1.098	1.226	1.51E-07
SERPINB2	1.166	1.106	1.229	9.95E-09
SERPINE1	1.240	1.183	1.299	1.45E-19

In addition, we divided patients with digestive system pancancer into two groups of high and low risk according to the optimal cutoff value 0.938. (Fig. 1e). In order to observe the data characteristics of the high- and low- risk groups, we performed PCA and tSNE, which showed that the high and low scores could distinguish the two groups of patients (Fig. 1f and Additional file 1: Fig. S1b). Time-dependent ROC curves showed that the scores had a better predictive performance for overall survival at 1 year (AUC=0.619), 2 years (AUC=0.646), and 3 years (AUC=0.655) in patients with digestive system pancancer (Fig. 1g). The survival analysis and risk score status indicated that patients in the high-risk group had a poor prognosis (Fig. 1h). The mortality rate increased with increasing risk score (Fig. 1i).

Clinical relevance and mechanism based on risk scores

We performed univariate and multivariate Cox regression analysis to explore whether the risk score was independent of other clinical characteristics, and the results showed that the risk score was an independent prognostic factor for patients with digestive system pancancer (Fig. 2a and Additional file 1: Fig.S1c). The ROC curve showed that the AUC (0.619)of the risk score was higher than that of other clinical features, which confirmed its better predictive performance (Fig. 2b). To facilitate clinical application, we established a nomogram with predictability, and the calibration curve was able to almost coincide with the ideal state (Fig. 2c-d). The violin plot demonstrated that stage □ and stage □ (P=4.1e-11), T1-T2 and T3-T4 (P=0.0019), N0 and N1 (P=6.7e-15) were the statistically most significant (Fig. 2e-g). However, the relationship among gender, M stage and risk score was not significant (Additional file 1: Fig. S1d-f). In addition, we performed GSEA analysis in order to determine the KEGG pathways enriched in high and low scores. The significant differential

pathways (NOM P value < 0.05) are shown in Tables 2 and 3. The high-risk group was mainly enriched in dilated cardiomyopathy, extracellular matrix (ECM) receptor interaction, focal adhesion, glycosaminoglycan biosynthesis chondroitinsulfate, mitogen-activated protein kinase (MAPK) signaling pathway, and pathway in cancer (Fig. 2h) while the low-risk group was mainly enriched in arcitrate cycle, tricarboxylic acid cycle (TCA) cycle, ginine and proline metabolism, glyoxylate and dicarboxylate metabolism (Fig. 2i).

Table 2 Top10 KEGG pathways for high risk group.

NAME	ES	NES	NOM p-val
KEGG_OXIDATIVE_PHOSPHORYLATION	0.692	2.219	0.006
KEGG_PARKINSONS_DISEASE	0.667	2.162	0.006
KEGG_PYRUVATE_METABOLISM	0.733	2.098	0.000
KEGG_ARGININE_AND_PROLINE_METABOLISM	0.726	2.115	0.000
KEGG_CITRATE_CYCLE_TCA_CYCLE	0.883	2.233	0.000
KEGG_PEROXISOME	0.813	2.065	0.000
KEGG_GLYOXYLATE_AND _DICARBOXYLATE_METABOLISM	0.847	2.056	0.000
KEGG_PROPANOATE_METABOLISM	0.846	2.020	0.000
KEGG_HUNTINGTONS_DISEASE	0.533	1.968	0.014
KEGG_ALZHEIMERS_DISEASE	0.515	1.960	0.014

Table 3 Top10 KEGG pathways for low risk group.

NAME	ES	NES	NOM p-val
KEGG_ECM_RECEPTOR_INTERACTION	-0.685	-1.977	0.004
KEGG_FOCAL_ADHESION	-0.604	-1.929	0.006
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	-0.550	-1.886	0.004
KEGG_MELANOMA	-0.560	-1.859	0.010
KEGG_GLYCOSAMINOGLYCAN_BIOSYNTHESIS _CHONDROITIN_SULFATE	-0.706	-1.804	0.010
KEGG_RENAL_CELL_CARCINOMA	-0.576	-1.795	0.020
KEGG_MAPK_SIGNALING_PATHWAY	-0.488	-1.759	0.016
KEGG_PATHWAYS_IN_CANCER	-0.509	-1.741	0.035
KEGG_DILATED_CARDIOMYOPATHY	-0.538	-1.707	0.026
KEGG_MTOR_SIGNALING_PATHWAY	-0.564	-1.695	0.030

CNV/MSI/TMB and m^6A regulator differences between high- and low- risk groups

Compared with the low-risk group, the CNV, MSI and TMB were significantly decreased in the high-risk group (Fig. 3a-c). The m^6A-related regulators ALKBH5, FTO, RBM15 , YTHDC1, YTHDC2 and 3-PCFGs ($P<0.001$) were highly expressed in the high-risk group (Fig. 3d). METTL3 and YTHDF21 ($P<0.001$) expression was low in the high-risk group (Fig. 3d). Fig. 3e showed the correlation between m6A-related genes and 3-PCFGs. We selected genes with a correlation coefficient $R>0.4$ to draw a correlation network diagram (Fig. 3f). Key genes included eleven m^6A regulators: FTO, METTL14, YTHDC1, YTHDC2, YTHDF1, YTHDF2, RBM15, ZC3H13, HNRNPC, METTL3, WTAP and one 3-PCFG gene, PLAT. Next, we plotted the correlation heatmap between key genes and clinical characteristics and found that the expression of key genes was consistent with the distribution trend of T, N and risk score (Fig. 3g).

Relationship between immune response and high-low scores in TIME

The expression of the general immune checkpoints PD-1 and PD-L1 between high- and low- risk groups is shown in Fig. 4a-b. Next, we expanded the scope and found 51 genes[21] involved in the immune response, including 15 receptors, 3 cell adhesion molecules, 4 antigen presentation-related genes, 3 costimulatory molecules, 6 coinhibitory molecules, and 20 ligands.A further clinical correlation heatmap was plotted and showed that these genes were highly expressed in the high-risk groups (Fig. 4c). Since immune cells played a key role in the TIME, we further explored the correlation between the 3-PCFG signature and immune cells. The bubble plot (Fig. 4d) demonstrated that the riskscore was positively correlated with cancer-ssociated fibroblasts, endothelial cells, monocytes,and

microenvironment scores, but negatively correlated with common lymphoid progenitor, T cell $CD8^+$ naive, T cell $CD4^+$ memory, etc.

Clustering based on PLAT/SERPINB2/SERPINB1 expression in digestive system pancancer

We performed K-means unsupervised clustering of patients with digestive system tumors based on 3-PCFG expression by thre "ConsensusClusterPlus" R package with selected optimal K value=3 (Fig. 5a-b). Survival analysis of digestive system pancancer showed that patients in cluster C had the worst prognosis (Fig. 5c). As shown in Fig. 5d, we surprisingly found that cluster C was mainly distributed in the high-risk group.

CNV/MSI/TMB and m^6A regulator differences among clusters

Among the clusters, CNV was found to be most significantly reduced in cluster C than in other clusters, and MSI and TMB were slightly reduced (Fig. 6a-c), which was similar to the high-risk group. Subsequently, we similarly analyzed the differences in the expression of m^6A regulators among different clusters. The m^6A regulators ALKBH5, FTO, RBM15, YTHDC1, YTHDC2 and 3-PCFGs (P□0.001) were highly expressed in cluster C, similar to the high-risk group (Fig. 6d). We next plotted the correlation heatmap between clusters and clinical characteristics and found that patients in cluster C were mostly distributed in the group with high expression of key genes (Fig. 6e).

Correlation between clustering and components of TIME

When analyzing of PD-1 and PD-L1 in each cluster, we found that PD-1 and PD-L1 were highly expressed in cluster C (Fig. 7a-b). The expression of 51 immune-related genes was increased accompanied by clinical progression (Fig. 7c). Cluster B, C had similar results as the high-risk group. Next, we analyzed the differences of TIME components between clusters. Differences in immune component scoring between clusters are shown in Fig. 7d. One notable exception was that cluster C had the highest immune score and stromal score. The clinical correlation heatmap of TIME components showed that the higher the clinical grade was, the higher the immune score and stromal score. Clusters B and cluster C had the higher immune scores and stromal scores (Fig. 7e).

Discussion

The treatment modalities for malignant tumors of the digestive system include surgery, chemotherapy, and radiotherapy, but some patients still die of complications such as VTE. The process of coagulation and fibrosis is closely related to the occurrence of VTE[5, 22]. TF, encoded by the F3 gene, is the initiator of the coagulation cascade and has receptor activity for factor □, so it can rapidly stop bleeding upon organ damage[23, 24]. PLAT and PLAU encode serine proteases including tissue plasminogen activator (tPA) and urokinase plasminogen activator (uPA) to inhibit the activity of the fibrinolytic system. Their activity is inhibited by the SERPINE1 and SERPINB2 genes encoding proenzyme activator inhibitor-1 (PAI-1) and proenzyme activator inhibitor-2 (PAI-2), respectively[25, 26]. The expression of both uPAR(encoded

by the gene PLAUR) and uPA is significantly upregulated during cancer progression and uPAR plays an active role in binding to uPA on the cell surface [27].

Hence, we developed and verified a prognostic scoring model for digestive system pancancer based on the study of Zuzana Saidak et al.[25]. Univariate Cox and multivariate Cox analysis were used to determine whether genes were associated with prognosis during the modeling process. LASSO solved the collinearity problem of gene expression data. PCA analysis indicated that patients could be well divided into two groups according to high and low scores. In the relationship betweent the risk score and clinical outcomes, we found that in early patients, the risk scores were greater than those in advanced patients, which might lead to VTE because the fact that early patients are more likely to have abnormal coagulation. In addition, we established a nomogram to provide a basis for determining the clinical stage in which the patient is located. GESA analysis of high and low scores showed that patients with low scores were mainly enriched in metabolic pathways (including ginine and proline metabolism, glyoxylate and dicarboxylate metabolism), suggesting that metabolic disorders may occur. Proline metabolism plays a key role in cancer reprogramming[28]. Simultaneous inhibition of proline metabolism and lipogenesis can inhibit tumor growth[29]. The main enrichment of patients with high scores lies in pathways related to cancer development, such as the MAPK signaling pathway[30]. This further suggests that our metabolic disorders are closely related to the development of cancer.

Nevertheless, the prognosis was not satisfactory in the high-risk group, and the reason for this we speculate may be related to genetic alterations of genes. CNV refers to a single nucleotide polymorphism resulting from a rearrangement of the genome. TMB is defined as the total number of mutations per million bases in tumor tissue. A higher of TMB indicates that tumor cells are more likely to be found by immune cells because the characteristics of tumor cells are more prominent. MSI refers to the phenomenon of the appearance of new microsatellite alleles in tumors due to changes in microsatellite length caused by the insertion or deletion of repeat units compared with normal tissues. Recently, in a study on the immune-related signature of gastric cancer, TMB was found to be higher in the low-risk group[31].It has been demonstrated that CNV, TMB, and MSI are associated with poor prognosis of cancer[32, 33]. In breast cancer, Xin Jin et al performed a comprehensive analysis of CNV, MSI, and TMB data from TCGA, which resulted in two genetic variation associated. They concluded that cluster 2 was marked by a worse prognosis and lower TMB[34]. In our study, similarly, CNV, TMB, and MSI were reduced in the high-risk group.

CNV may lead to disturbed expression of m^6A regulator proteins across 33 tumors[35]. Epigenetic modification played a nonnegligible role in the dysregulation of genes controlling tumor angiogenesis or the coagulant phenotype. This was verified in a study by Feng Xu et al.[16]. In 2011, JIA et al.[36] first found that FTO induced RNA m^6A demethylation in vivo, demonstrating the dynamic reversibility of m^6A modification. It has been demonstrated that rs9939609, located in the first intron of FTO, is associated with the development of recurrent VTE in male patients[37]. Reducing obesity can reduce the occurrence of VTE[38]. A study has suggested that ALKBH5 promotes ESCC cell proliferation by regulating cell cycle progression, ALKBH5 was highly expressed, and was associated with lower 5-year overall survival, and

ALKBH5 could be considered as a factor for poor prognosis in esophageal squamous cell carcinoma (ESCC) patients and might be a new therapeutic target for ESCC patients[39]. Our findings suggested that METTL3 and YTHDF2 expression was lower in the high-risk group than in the low-risk group,. The aforementioned findings suggest that alterations in the expression of m^6A regulators are associated with digestive system pancancer patients with a high-risk of VTE.

There is convincing evidence [13] that m^6A occurring on mRNA plays an indispensable role in the function of immune cells and innate immune response in the tumor microenvironment. First, we analyzed the correlation of 51 genes[21] encoding immunomodulatory molecules with risk scores and found that these genes were highly expressed in most of the high-risk patients. The application of PD-1 or PD-L1 monoclonal antibodies and anti-PD-1/PD-L1 antibody combination therapy have greatly advanced the process of immunotherapy[40].The work of Sun M[12] showed that the expression of PD-L1 and PD-1 was positively correlated with the m6A score. While PD-1 and PD-L1 are highly expressed in the high-risk group, our epigenetic results indicate that most m^6A regulators are highly expressed in the high-risk group, and combining these two results, we conclude that the expression of PD-1 and PD-L1 is positively correlated with PCFG scores. To provide ideas for clinical personalized immunotherapy, we further analyzed the correlation of immune cell infiltration in the TIME with risk scores and suggested that patients with high scores might have plentiful immune cells incuding fibroblasts, endothelial cells and monocytes infiltration, which were similar to previous studies[41]. Consequently, the high-risk group may be insensitive to immunotherapy.

To further profile the predictive power of the three participating model prognostic genes, we performed clustering based on the expression of 3-PCFGs and explored similar content to the risk model. It was found that cluster C had similar results to patients in the high-risk group on m^6A and TIME components. Our prognostic model results and clustering results are sufficient to demonstrate that our prognostic model has good predictive performance even without external dataset validation. Of course, external data validation is highly beneficial.

Conclusion

In conclusion, we found a prognostic scoring model of 3-PCFGs and identified novel tumor subtypes for patients with digestive system pancancer. More importantly, we explained that poor prognosis in high-scoring patients and cluster C patients may be affected by m^6A epigenetic modification and TIME components.

Declarations

Availability of data and materials

The datasets used and/or analyzed during the current study are available

from UCSC: https://xenabrowser.net/datapages/; GDC repository https://gdc.cancer.gov/about-data/publications/panimmune; TIMER 2.0 http://timer.cistrome.org; the KEGG pathway database: www.kegg. jp/ kegg/ kegg1. html

Competing interests

None of the authors has any conflicts of interest to disclose.

Funding

2019 Basic Scientific Research Projects of Provincial Higher Education Institutions in Heilongjiang Province (2019-KYYWFMY-0001).

Authors' contributions

XYW, JL J, Y J and RJ C conceived and designed the experiments, performed the experiments, prepared figures and/or tables, and authored or reviewed drafts of the paper.YY Z, ZY S, ZK H, P S, H W, LY L, HL G performed the experiments, analyzed the data, prepared figures and/or tables, and authored or reviewed drafts of the paper. All authors read and approved the final manuscript.

Acknowledgments

We acknowledge UCSC databases for providing their platforms and contributors for uploading their meaningful datasets.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F: Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021, 71:209-249. https://doi.org/10.3322/caac.21660.
2. Achebe I, Mbachi C, Palacios P, Wang Y, Asotibe J, Ofori-Kuragu A, Gandhi S: Predictors of venous thromboembolism in hospitalized patients with inflammatory bowel disease and colon cancer: A retrospective cohort study. Thrombosis research 2021, 199:14-18. https://doi.org/10.1016/j.thromres.2020.12.017.
3. Lundbech M, Krag AE, Iversen LH, Hvas A-M: Postoperative bleeding and venous thromboembolism in colorectal cancer patients undergoing cytoreductive surgery with hyperthermic intraperitoneal chemotherapy: a systematic review and meta-analysis. International journal of colorectal disease 2021. https://doi.org/10.1007/s00384-021-04021-6.
4. Frere C, Bournet B, Gourgou S, Fraisse J, Canivet C, Connors JM, Buscail L, Farge D: Incidence of Venous Thromboembolism in Patients With Newly Diagnosed Pancreatic Cancer and Factors Associated With Outcomes. Gastroenterology 2020, 158. https://doi.org/10.1053/j.gastro.2019.12.009.

5. Tawil N, Bassawon R, Rak J: Oncogenes and Clotting Factors: The Emerging Role of Tumor Cell Genome and Epigenome in Cancer-Associated Thrombosis. Semin Thromb Hemost 2019, 45:373-384. https://doi.org/10.1055/s-0039-1687891.
6. Timp JF, Braekkan SK, Versteeg HH, Cannegieter SC: Epidemiology of cancer-associated venous thrombosis. Blood 2013, 122:1712-1723. https://doi.org/10.1182/blood-2013-04-460121.
7. Akl EA, Ramly EP, Kahale LA, Yosuico VE, Barba M, Sperati F, Cook D, Schunemann H: Anticoagulation for people with cancer and central venous catheters. Cochrane Database Syst Rev 2014:CD006468. https://doi.org/10.1002/14651858.CD006468.pub5.
8. Hutter C, Zenklusen JC: The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. Cell 2018, 173:283-285. https://doi.org/10.1016/j.cell.2018.03.042.
9. Blum A, Wang P, Zenklusen JC: SnapShot: TCGA-Analyzed Tumors. Cell 2018, 173:530. https://doi.org/10.1016/j.cell.2018.03.059.
10. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, de Crecy-Lagard V, Ross R, Limbach PA, Kotter A, et al: MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 2018, 46:D303-D307. https://doi.org/10.1093/nar/gkx1030.
11. Liu C, Yang S, Zhang Y, Wang C, Du D, Wang X, Liu T, Liang G: Emerging Roles of N6-Methyladenosine Demethylases and Its Interaction with Environmental Toxicants in Digestive System Cancers. Cancer Manag Res 2021, 13:7101-7114. https://doi.org/10.2147/CMAR.S328188.
12. Sun M, Xie M, Zhang T, Wang Y, Huang W, Xia L: mA Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Pancreatic Cancer. Frontiers in immunology 2021, 12:739768. https://doi.org/10.3389/fimmu.2021.739768.
13. Zhan L, Zhang J, Zhu S, Liu X, Zhang J, Wang W, Fan Y, Sun S, Wei B, Cao Y: N-Methyladenosine RNA Modification: An Emerging Immunotherapeutic Approach to Turning Up Cold Tumors. Frontiers in cell and developmental biology 2021, 9:736298. https://doi.org/10.3389/fcell.2021.736298.
14. Sauerbrei W, Royston P, Binder H: Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Stat Med 2007, 26:5512-5528. https://doi.org/10.1002/sim.3148.
15. Russo AE, Strong VE: Gastric Cancer Etiology and Management in Asia and the West. Annu Rev Med 2019, 70:353-367. https://doi.org/10.1146/annurev-med-081117-043436.
16. Xu F, Huang X, Li Y, Chen Y, Lin L: m(6)A-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with LUAD. Mol Ther Nucleic Acids 2021, 24:780-791. https://doi.org/10.1016/j.omtn.2021.04.003.
17. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005, 102:15545-15550. https://doi.org/10.1073/pnas.0506580102.
18. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS: TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res 2017, 77:e108-e110.

https://doi.org/10.1158/0008-5472.CAN-17-0307.

19. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015, 12:453-457. https://doi.org/10.1038/nmeth.3337.
20. Wilkerson MD, Hayes DN: ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics (Oxford, England) 2010, 26:1572-1573. https://doi.org/10.1093/bioinformatics/btq170.
21. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, et al: The Immune Landscape of Cancer. Immunity 2018, 48:812-830 e814. https://doi.org/10.1016/j.immuni.2018.03.023.
22. Falanga A, Schieppati F, Russo L: Pathophysiology 1. Mechanisms of Thrombosis in Cancer Patients. Cancer Treat Res 2019, 179:11-36. https://doi.org/10.1007/978-3-030-20315-3_2.
23. Grover SP, Mackman N: Tissue Factor: An Essential Mediator of Hemostasis and Trigger of Thrombosis. Arterioscler Thromb Vasc Biol 2018, 38:709-725. https://doi.org/10.1161/ATVBAHA.117.309846.
24. Witkowski M, Landmesser U, Rauch U: Tissue factor as a link between inflammation and coagulation. Trends Cardiovasc Med 2016, 26:297-303. https://doi.org/10.1016/j.tcm.2015.12.001.
25. Saidak Z, Soudet S, Lottin M, Salle V, Sevestre MA, Clatot F, Galmiche A: A pancancer analysis of the human tumor coagulome and its link to the tumor immune microenvironment. Cancer Immunol Immunother 2021, 70:923-933. https://doi.org/10.1007/s00262-020-02739-w.
26. Yao Y, Xu Q: Progress in the study of cancer-associated venous thromboembolism. Vascular 2021, 29:408-414. https://doi.org/10.1177/1708538120957443.
27. Lund IK, Illemann M, Thurison T, Christensen IJ, Høyer-Hansen G: uPAR as anti-cancer target: evaluation of biomarker potential, histological localization, and antibody-based therapy. Curr Drug Targets 2011, 12:1744-1760. https://doi.org/10.2174/138945011797635902.
28. Phang JM: Proline Metabolism in Cell Regulation and Cancer Biology: Recent Advances and Hypotheses. Antioxid Redox Signal 2019, 30:635-649. https://doi.org/10.1089/ars.2017.7350.
29. Liu M, Wang Y, Yang C, Ruan Y, Bai C, Chu Q, Cui Y, Chen C, Ying G, Li B: Inhibiting both proline biosynthesis and lipogenesis synergistically suppresses tumor growth. J Exp Med 2020, 217. https://doi.org/10.1084/jem.20191226.
30. Chang WT, Bow YD, Fu PJ, Li CY, Wu CY, Chang YH, Teng YN, Li RN, Lu MC, Liu YC, Chiu CC: A Marine Terpenoid, Heteronemin, Induces Both the Apoptosis and Ferroptosis of Hepatocellular Carcinoma Cells and Involves the ROS and MAPK Pathways. Oxid Med Cell Longev 2021, 2021:7689045. https://doi.org/10.1155/2021/7689045.
31. Dai S, Liu T, Liu XQ, Li XY, Xu K, Ren T, Luo F: Identification of an Immune-Related Signature Predicting Survival Risk and Immune Microenvironment in Gastric Cancer. Front Cell Dev Biol 2021, 9:687473. https://doi.org/10.3389/fcell.2021.687473.

32. Li J, Tong Y, Wang Z, Liu Y, Dai X, Zhu Y: A Comprehensive Multiomics Analysis Identified Ubiquilin 4 as a Promising Prognostic Biomarker of Immune-Related Therapy in pancancer. Journal of oncology 2021, 2021:7404927. https://doi.org/10.1155/2021/7404927.
33. Zhang N, Li P, Wu X, Xia S, Zhao X, Chen L: Analysis of Threshold Changes of Tumor Mutation Burden of Gastric Cancer and Its Relationship with Patients' Prognosis. Journal of oncology 2021, 2021:9026610. https://doi.org/10.1155/2021/9026610.
34. Jin X, Yan J, Chen C, Chen Y, Huang WK: Integrated Analysis of Copy Number Variation, Microsatellite Instability, and Tumor Mutation Burden Identifies an 11-Gene Signature Predicting Survival in Breast Cancer. Front Cell Dev Biol 2021, 9:721505. https://doi.org/10.3389/fcell.2021.721505.
35. Li Y, Xiao J, Bai J, Tian Y, Qu Y, Chen X, Wang Q, Li X, Zhang Y, Xu J: Molecular characterization and clinical relevance of m(6)A regulators across 33 cancer types. Mol Cancer 2019, 18:137. https://doi.org/10.1186/s12943-019-1066-3.
36. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, Yi C, Lindahl T, Pan T, Yang YG, He C: N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 2011, 7:885-887. https://doi.org/10.1038/nchembio.687.
37. Ahmad A, Memon A, Sundquist J, Svensson P, Zöller B, Sundquist K: Fat mass and obesity-associated gene rs9939609 polymorphism is a potential biomarker of recurrent venous thromboembolism in male but not in female patients. Gene 2018, 647:136-142. https://doi.org/10.1016/j.gene.2018.01.013.
38. Lindström S, Germain M, Crous-Bou M, Smith E, Morange P, van Hylckama Vlieg A, de Haan H, Chasman D, Ridker P, Brody J, et al: Assessing the causal relationship between obesity and venous thromboembolism through a Mendelian Randomization study. Human genetics 2017, 136:897-902. https://doi.org/10.1007/s00439-017-1811-x.
39. Nagaki Y, Motoyama S, Yamaguchi T, Hoshizaki M, Sato Y, Sato T, Koizumi Y, Wakita A, Kawakita Y, Imai K, et al: m A demethylase ALKBH5 promotes proliferation of esophageal squamous cell carcinoma associated with poor prognosis. Genes to cells : devoted to molecular & cellular mechanisms 2020, 25:547-561. https://doi.org/10.1111/gtc.12792.
40. Jiang Y, Chen M, Nie H, Yuan Y: PD-1 and PD-L1 in cancer immunotherapy: clinical implications and future considerations. Hum Vaccin Immunother 2019, 15:1111-1122. https://doi.org/10.1080/21645515.2019.1571892.
41. Wang Y, Guo S, Chen Z, Bai B, Wang S, Gao Y: Re-Clustering and Profiling of Digestive System Tumors According to Microenvironment Components. Front Oncol 2020, 10:607742. https://doi.org/10.3389/fonc.2020.607742.

Figures

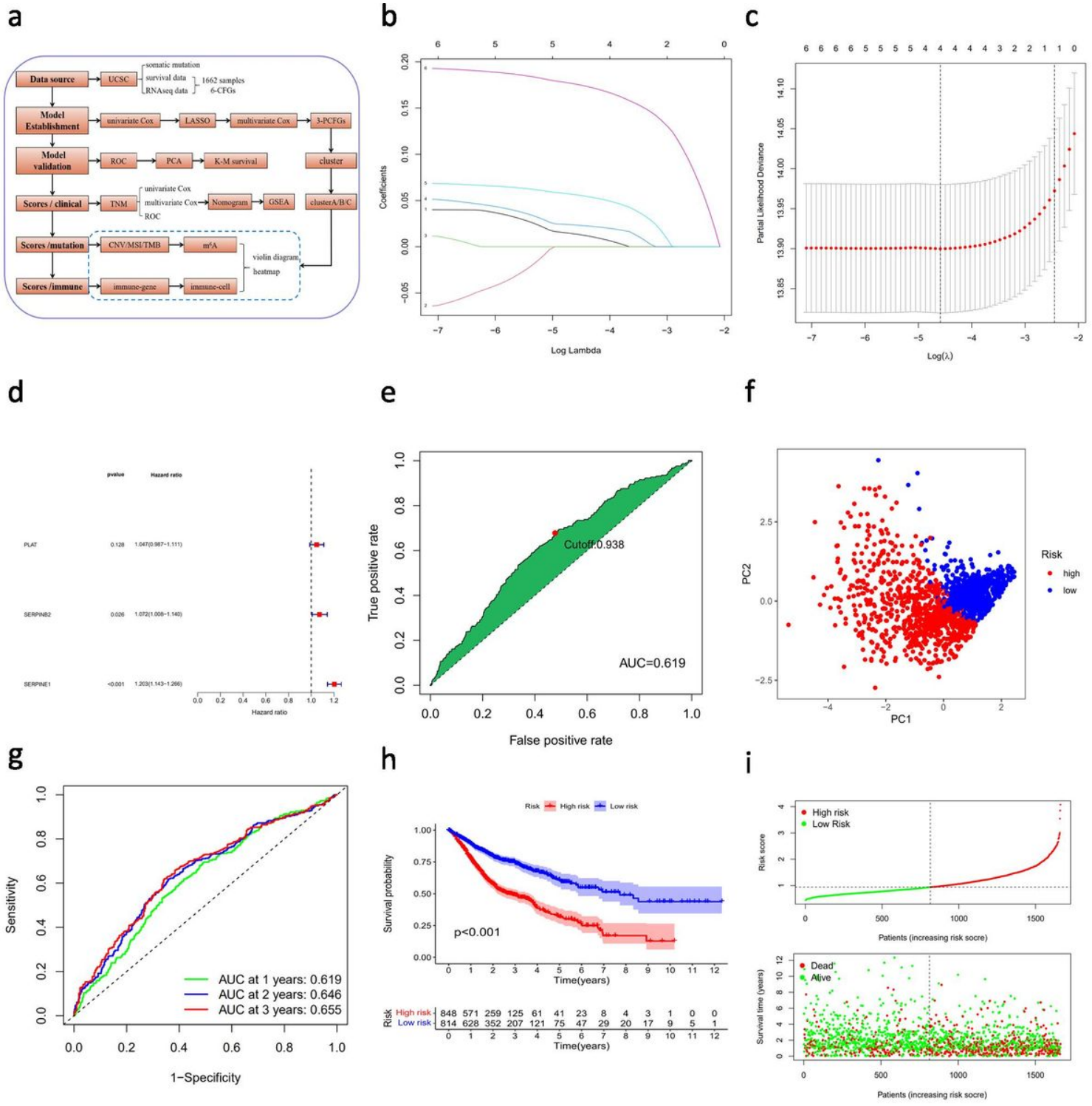

Figure 1

Construction and validation of the 3-PCFG signature for the 1662 digestive system pancancer **a** The workflow of our study. **b** The tuning parameters (log l) of OS-related genes were selected to cross-verify the error curve. According to the minimal criterion and 1-se criterion, perpendicular imaginary lines were drawn at the optimal value. **c** The LASSO coefficient profile of 6 OS-related genes and perpendicular imaginary line were drawn at the value chosen by 10-fold cross-validation. **d** Multivariate Cox regression analysis showed 3 independent prognostic genes. **e** AUC of digestive system pancancer patients and the

best cutoff value obtained by the AIC. **f** Principal component analysis between the high- and low-risk groups. **g** The AUCs for 1-, 2-, and 3-year overall survival. **h** Kaplan-Meier plots of overall survival according to the signature with high and low risk scores. **i** The distribution of patients in the high- and low-risk groups.

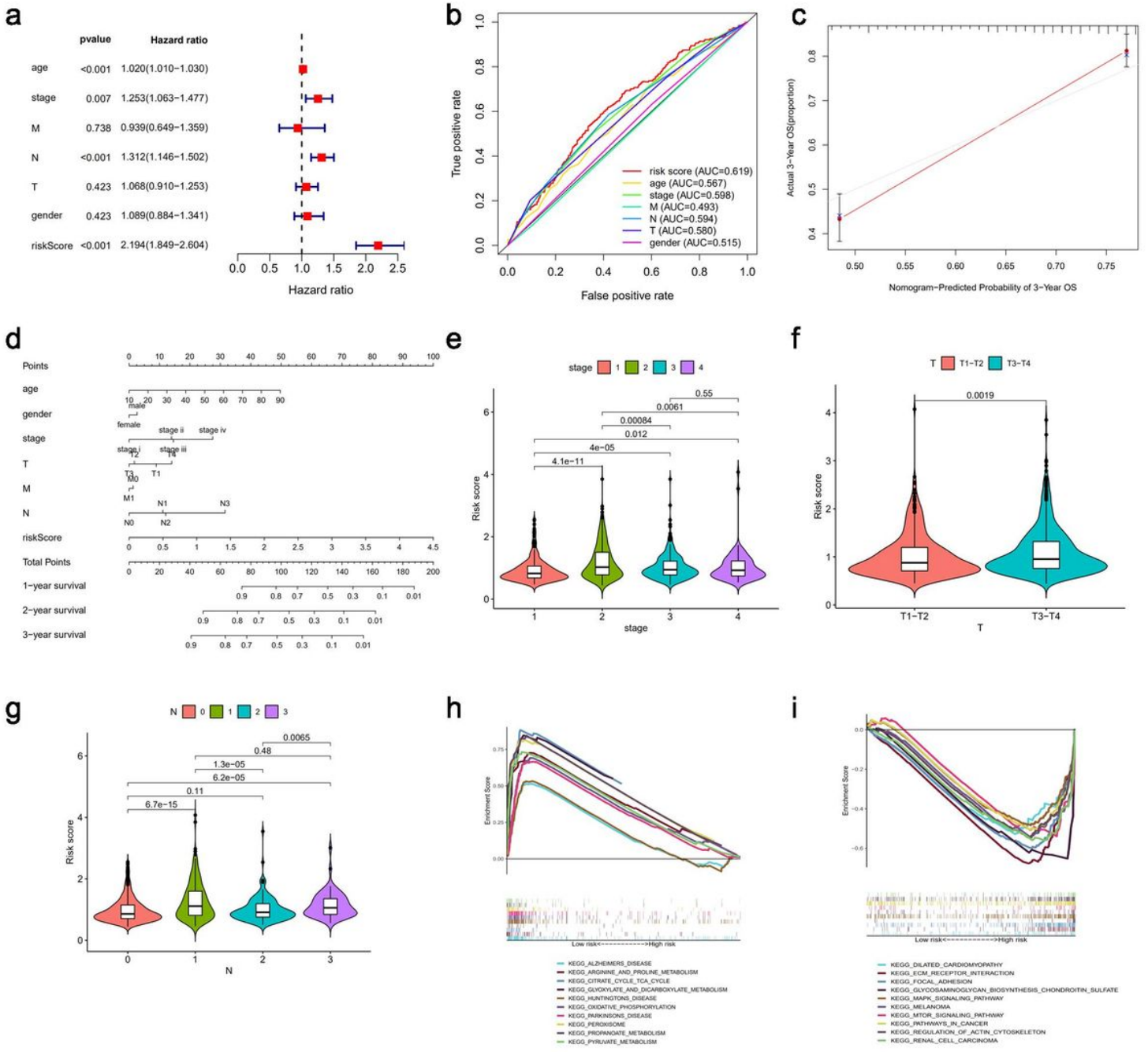

Figure 2

Clinical relevance and KEGG pathways based on risk score. **a** Multivariate Cox regression analysis. **b** Area under the ROC curve (AUC) for risk scores and clinical characteristics. **c-d** Construction and validation of the nomogram. **e** TNM stage **f** T stage **g** N stage **h** Top10 enriched pathways in the low-risk group. **i** Enriched Top10 pathways in high risk group.

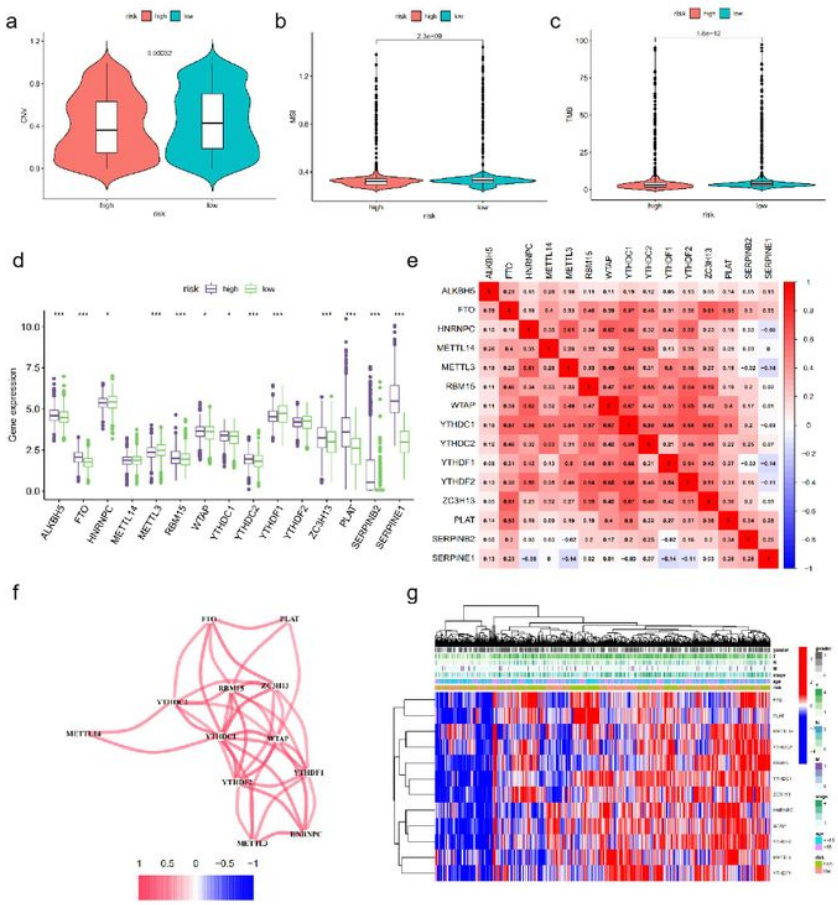

Figure 3

M^{6}A-related genes and mutation differences between high and low risk groups **a-c** Differences in CNV, MSI and TMB **d** Differences in 12 m^{6}A-related genes and 3-PCFGs **e-f** Correlation analysis heatmap of 12 m^{6}A-related genes and 3-PCFGs. **g** The clinically relevant heatmap of 11 m^{6}A-related genes and PLAT.

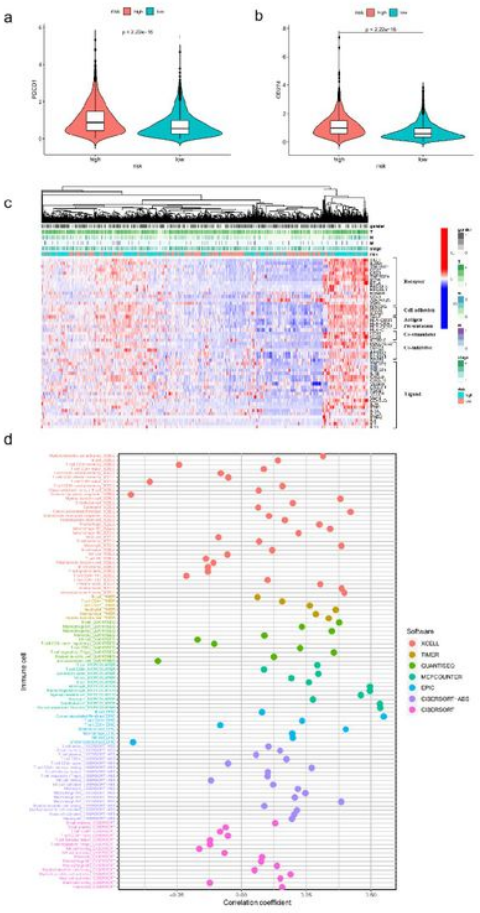

Figure 4

The correlation of immune-related genes and immune cells with risk score **a-b** The differences in PDCD1(PD-1) and CD274(PD-L1) **c** The clinically relevant heatmap of 51 immune-related genes **d** Bubble plot for the correlation of immune cells with risk scores.

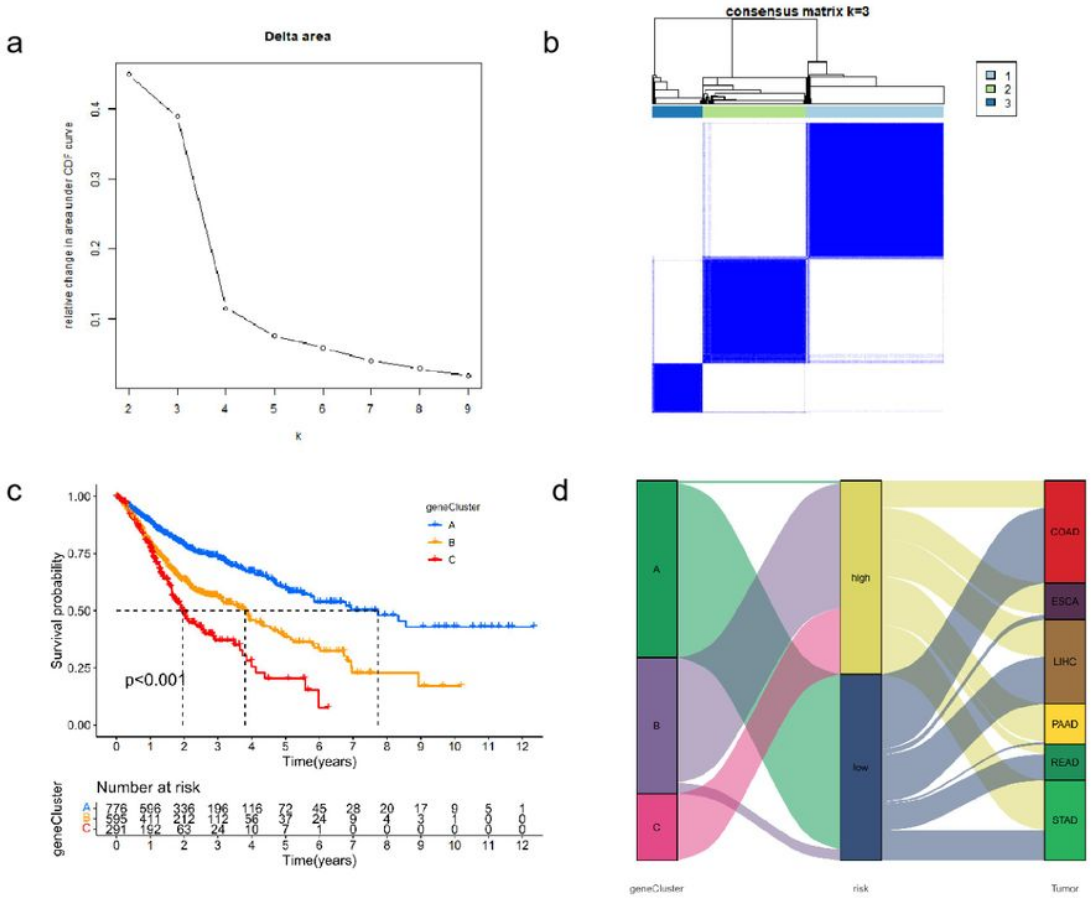

Figure 5

K shows unsupervised clustering of patients with digestive system tumors based on 3-PCFG expression **a** Relative change in area under the CDF curve fork=2–9. **b** Consensus clustering matrix for k=2. **c** Survival analysis of custers **d** Sankey diagram for clusters, high-low group and six types of digestive system cancer.

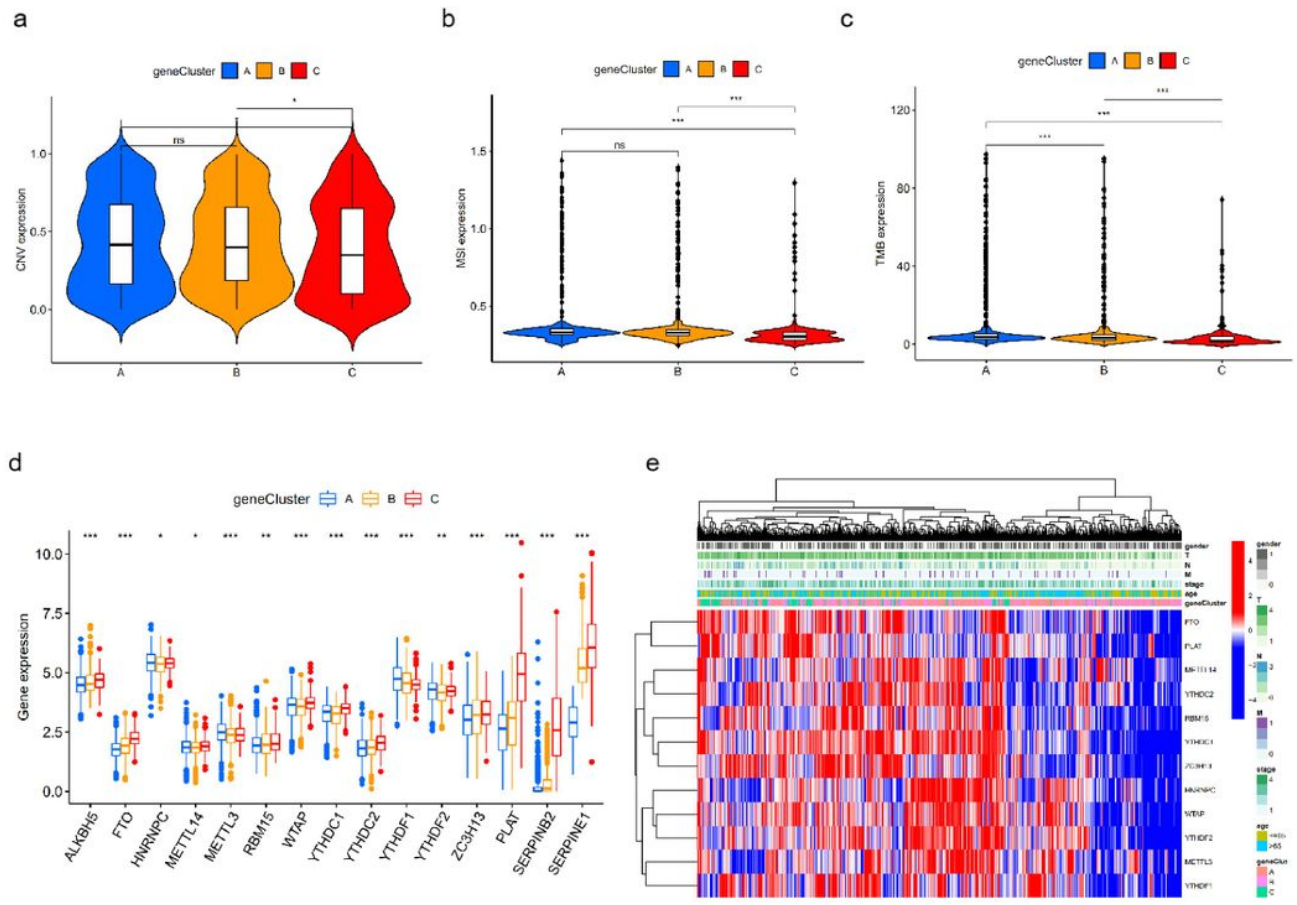

Figure 6

M^{6}A-related genes and mutation differences among clusters **a-c** The differences of CNV, MSI, TMB **d** The differences of 12 m^{6}A-related genes and 3-PCFGs. **e** The clinically relevant heatmap of 11 m^{6}A-related genes and PLAT. (ns indicates no significancce, *indicates P<0.05, **indicates P<0.01,***indicates P<0.001).

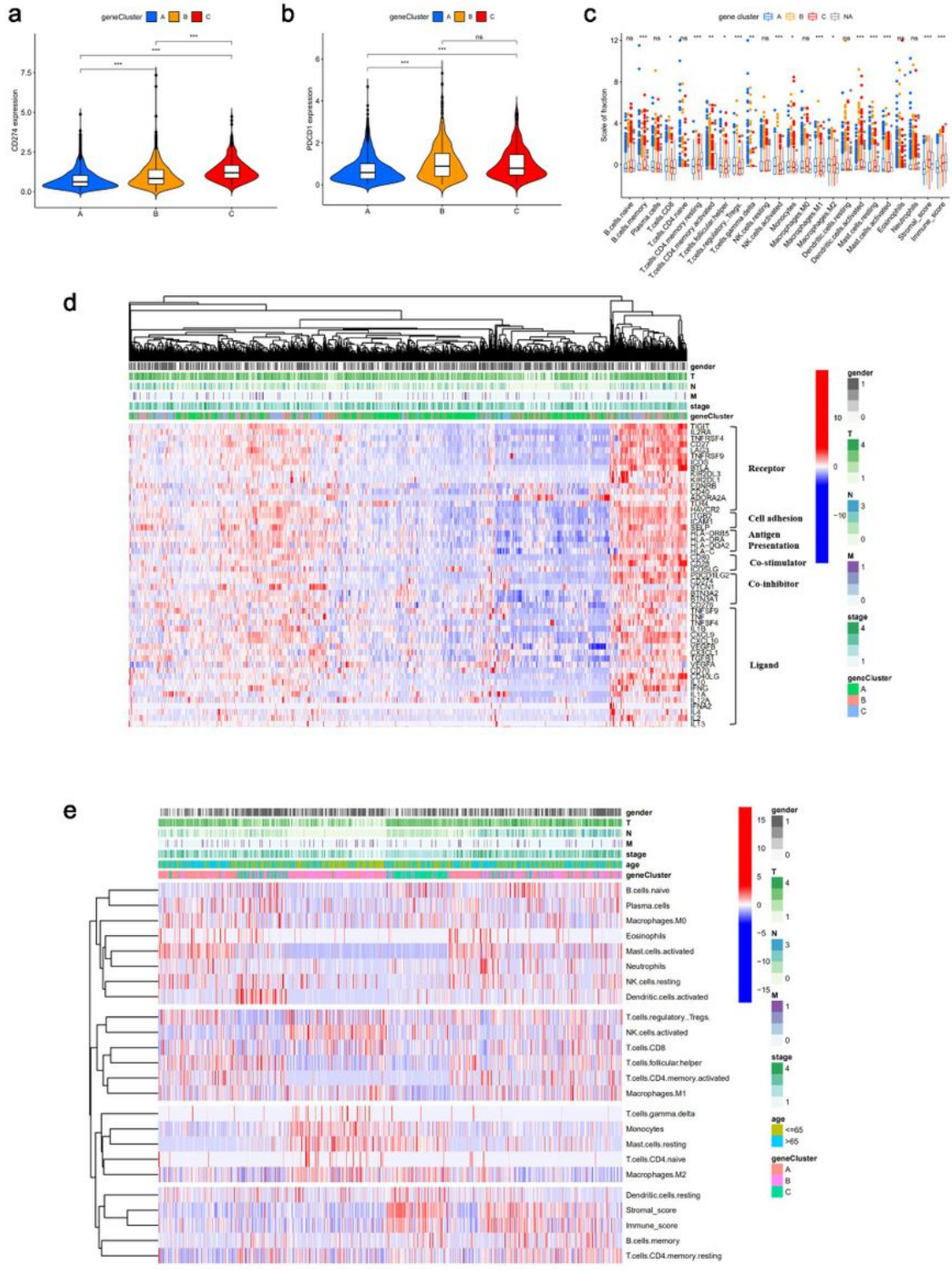

Figure 7

Immune-related analysis among clusters **a-b** The differences of PDCD1(PD-1) and CD274 (PD-L1) **c** The correlation of TIME composition among clusters **d** The clinically relevant heatmap of 51 immune-related genes (ns means no significancce, *means P<0.05, **means P<0.01,***means P<0.001). **e** The clinically relevant heatmap of 24 TIME components (ns indicates no significancce, *indicates P<0.05, **indicates P<0.01,***indicates P<0.001).

Supplementary Files

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

- Additionalfile1.doc