Comprehensive Analysis of Genomic Mutation Signature and Tumor Mutation Burden for Prognosis of Intrahepatic Cholangiocarcinoma

Rui Zhang
Xi'an Jiaotong University Medical College First Affiliated Hospital https://orcid.org/0000-0002-1727-5524

Qi Li
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Jialu Fu
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Zhechuan Jin
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Jingbo Su
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Jian Zhang
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Chen Chen
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Zhimin Geng
Xi'an Jiaotong University Medical College First Affiliated Hospital Department of Hepatobiliary Surgery

Dong Zhang (zhangdong811021@126.com)
The First Affiliated Hospital of Xi'an Jiaotong University https://orcid.org/0000-0001-8254-7568

Research article

Keywords: Genomic mutation signature, Tumor mutation burden, Intrahepatic cholangiocarcinoma, Prognostic biomarker, Nomogram

Posted Date: October 14th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-89482/v1

Version of Record: A version of this preprint was published on February 3rd, 2021. See the published version at https://doi.org/10.1186/s12885-021-07788-7.

Abstract

Background: The intrahepatic cholangiocarcinoma (iCCA) is highly lethal malignancy of biliary tract cancer. Analysis of somatic mutational profiling could help to reveal new prognostic markers and actionable targets for treatment. We aim to explore the impact of genomic mutation signature and tumor mutation burden (TMB) on prognosis of iCCA patients.

Methods: The whole-exome sequencing and corresponding clinical data were collected from ICGC portal and cBioPortal database to detect the mutation prognostic genes and TMB values. To identify the hub prognostic mutant signature, Cox regression and Lasso feature selection were conducted. We built a mutation related signature (MRS) through multivariate Cox regression. The predictive performance of MRS and TMB were assessed using Kaplan-Meier (KM) analysis and receiver operating characteristic (ROC). We performed a functional enrichment pathway analysis with gene set enrichment analysis (GSEA) for mutated genes were conducted. Moreover, on the base of MRS, TMB and TNM stage, a nomogram was constructed to visualize the prognosis of iCCA patients.

Results: The mutation landscape illustrated the distributions of mutation frequencies and types on iCCA, and revealed a list of most frequent mutation genes (such as Tp53, KRAS, ARID1A, IDH1). We obtain a 6-gene signature using the Lasso and Cox method. The AUC of MRS in 1, 3, 5-year OS prediction were 0.759, 0.732, 0.728, respectively. Moreover, Kaplan-Meier analysis showed a significant difference on the prognosis of iCCA with high and low MRS score ($P < 0.001$). Interestingly, GSEA was utilized to show several signaling pathways including MAPK signaling pathway, PI3K-AKT signaling pathway and proteoglycans in cancer. On the other hand, survival analysis indicated that TMB was significantly associated with prognosis. And GSEA indicated that samples with high MRS or TMB upregulated signaling pathways involved in tumor signaling and immune system. At last, we constructed a predictive nomogram (included MRS, TMB and TNM stage) with satisfactory performance in survival prediction.

Conclusions: The mutation genes signature and TMB were associated with prognosis in patients with iCCA. Our study provides a valuable prognostic predictor for further uncovering molecular pathogenesis in iCCA.

Background

Cholangiocarcinoma (CCA) is a highly lethal and aggressive malignancy originating from biliary epithelium. Based on the anatomical site of origin, CCA can be classified into three subtypes including perihilar (pCCA), distal subtypes (dCCA), and intrahepatic (iCCA)[1]. The iCCA represents 10–20% of primary liver cancers and approximately 20% of biliary tract cancers[2]. The postoperative 5-year overall survival (OS) rate of iCCA patients remains dismal (only 30–40%) [3, 4]. Meanwhile, a global trend of increasing incidence rate and mortality by iCCA has been observed in the last decade, which is in contrast with the decreasing trends of pCCA and dCCA [5]. Surgical resection is currently the mainstay of curative-intent treatment for patients at an early stage, however, the vast majority of patients has missed the

opportunity for radical surgery [6, 7]. Nevertheless, iCCA patients, even after resection, are always associated with high incidence of recurrence rate [8–10]. Given that dismal survival in patients, it is imperative to detect prognostic biomarkers for treatment decision-making and improving patient prognosis. Therefore, it is badly needed to explore a distinct prognostic predictive model for patients with iCCA.

Cancer is often accompanied with accumulation of various genetic mutations, and the accumulation of somatic mutation will contribute to the tumorigenesis and progression of malignancy. Gene mutation is consistently the critical factor determining function of gene on biological behavior of malignant tumors [11, 12]. Multiple studies suggest that particular gene mutations may be prognostic indicators to predicting survival and response of adjunctive therapy [13–16]. There are several high frequent genetic alterations have been identified in iCCA including TP53, KRAS, ARID1A, IDH1/2, BAP1 and PBRM1[17, 18]. However, it is poorly understood that the prognostic implications of these somatic alterations on iCCA. Therefore, an enhanced understanding of the genetic mutations is imperative to explore prognostic genetic biomarkers for selecting high risk CCA patients harboring pertinent genetic mutations, and tailor appropriate treatment plans in clinical practice.

Tumor mutation burden (TMB) was defined as the number of somatic (such as missense, deletion or insertion) per megabase of genome examined[19]. Recently, TMB has arisen as a robust biomarker that can predict response to immune checkpoint inhibitor (ICI) therapy [20, 21]. Furthermore, many studies have established the prognostic role of TMB on the efficiency of immunotherapy in many tumors, including non-small cell lung carcinoma (NSCLC) [22, 23], melanoma [24], esophagogastric cancer [25] and colorectal cancer [26]. However, although the role of TMB on the efficiency of ICIs had been well-studied, limited studies investigated the prognostic impact of TMB on patients suffered from malignancies. No published study clarifies the clinical and survival impact of TMB on iCCA.

Whole exome sequencing (WES) is regarded as the gold standard method to assess the value of TMB. In recent years, bioinformatic sources of WES was available from public databases such as International Cancer Genome Consortium (ICGC, https://dcc.icgc.org/) and cBioPortal (https://www.cbioportal.org/), which made it possible for large-scale cancer genomic integration and comprehensive bioinformatics analysis. Based on these public databases, many researches had been performed to find the affecting factors of cancer immunotherapy.

We sought to investigate the mutational landscape on iCCA, and explore potential impact of mutation related signature on survival using WES data from ICGC and cBioPortal database, and tried to establish a reliable nomogram model based on mutant gene signature, TMB and other clinical characteristics to predict overall survival (OS) of patients with iCCA. By this way, we could explore new potential prognostic biomarkers and provide potential therapeutic targets for iCCA.

Methods

Mutation Data Collection

Somatic mutations data and the corresponding clinicopathological characteristics information for iCCA patients were acquired from cBioPortal (http://www.cbioportal.org) and ICGC portal (http://dcc.icgc.org/releases/current/Projects). We only chose the WES dataset of iCCA patients. The repositories used were BTCA-JP (Japan, Nat Genet 2015) [27], BTCA-SG (Singapore, Cancer Discov 2017) [28], TCGA-CHOL (TCGA, PanCancer Atlas) [12] and Intrahepatic Cholangiocarcinoma (Shanghai, Nat Commun 2014) [29]. Only those with complete clinicopathological information were included. Clinical characteristics includes age, gender, TNM stage, survival status, and survival time. Then, the Perl scripts was used to extract the somatic mutation information in iCCA. We utilized “GenVisR” and “karyoploteR” package in R software to achieve the visualization.

Construction of Protein-Protein Interaction (PPI) Network

Top 200 most frequency mutated genes were selected to investigate the potential interactive relationships. The genes were inputted into STRING database (https://stringdb.org/) to generate the PPI network, with a confidence score >0.7 were considered as the cut-off criterion. We also modified the PPI network by Cytoscape. Then the mutation gene nodes with edge of >5 were extracted as the most important targets.

Functional Enrichment Pathway Analysis

The top 300 most frequency mutant genes were selected to perform enrichment pathway analysis. “org.Hs.eg.db” package, “ggplot2”, “clusterProfiler” and “enrichplot” packages were utilized for Gene Ontology (GO) analysis and Kyoto Gene and Genome Encyclopedia (KEGG) pathways with FDR <0.05 as statistically significant.

Hub Prognostic Mutant Genes and Construction of the Prognostic Model

To screen hub mutant genes for prognosis of iCCA, univariate Cox regression analysis was conducted to find prognostic mutant genes. We performed least absolute shrinkage and selection operator (Lasso) regression to conduct dimensionality reduction analysis of the survival associated mutant genes, using the “survival” and “glmnet” packages in R. Lasso sub-selects prognostic mutant genes by providing a penalty proportional to the contraction of the regression coefficient. Then, we performed multivariate Cox regression analysis to establish the mutation related signature (MRS), which was calculated as formula:

$$\mathrm{MRS} = \sum_{i=1}^{n} (\beta_i * \mathrm{Mut}_i)$$

where β_i is the coefficient and Mut_i represents the mutation status of genes (if the status is Mutation, Mut_i=1; if the statue is Wild, Mut_i=0). Subsequently, the 318 iCCA patients were classified into low- and high-groups according to the median MRS. Overall survival (OS) was estimated to compare the differential survival between two groups, with P value <0.05 indicating a significant difference. Furthermore, Receiver operating characteristic (ROC) curves were built to evaluate prediction performance of MRS for the 1-, 3-, and 5-year survival rate. Forest plots were also used to show the hazard ratio (HR) of selected prognostic mutation genes by “survminer” package.

TMB Value of Patients with iCCA Estimation and Prognostic Analysis

TMB was defined as the total number of mutations per coding area. All non-synonymous variants in the coding region were counted, and silent mutations were not considered. The genomic mutations of 318 iCCA patients were specifically extracted. Since 38 Mb is routinely taken as the total length of exon on human, so we calculated the TMB as the total mutation frequency divided by 38 [30]. TMB of each patient was calculated by this method and corresponding survival data was merged. Then, we divided the patients into high- and low- TMB groups according to the cut-off values of TMB, which were determined by the maximum Youden index of ROC curve analysis.

Gene Set Enrichment Analysis (GSEA)

To explore the potential biological pathway among the high- and low- group based on MRS and TMB, we performed GSEA (MSigDB; version 7.1) with “KEGG”, “GO” and “immunologic signatures” gene sets from Molecular Signature Database. The mutation gene list of MRS or TMB status was used as input phenotype data. A total of 1,000 times were performed for gene set permutations and the pathways with P<0.05 were considered as significant.

Statistical Analysis

All statistical tests were completed by SPSS 24.0, R software (version 4.0.2) and GraphPad Prism 8.0. The Student t test was used to compare continuous variables. Meanwhile, the χ^2 test or Fisher's exact test was used to compare categorical data. The effects of AJCC TMN stage, MRS and TMB on survival were assessed with the log-rank test and Kaplan–Meier method. Multivariable Cox regression analysis was

used to determine the independent risk factors. A nomogram model was constructed and the predicted performance of nomogram was estimated by the C-index and calibration plot.

Results

Landscape of Mutation Gene Profiles in iCCA

The demographics and clinicopathological characteristics of 318 iCCA patients who underwent WES are listed in Table 1. There were 192 males and 126 females in the included cohort. The median age at the time of diagnosis was 62 years (range, 26–89 years). The results revealed 15 genes (TP53, TTN, KRAS, MUC2, ARID1A, MUC16, BAP1, OBSCN, CSMD3, EPHA2, IDH1, PCLO, LRP1B, PBRM1, SYNE1) were mutated in more than 20 samples. We visualized the landscape of mutation profiles using “GenVisR” package, which only showed the top 35 most frequency mutation genes across 318 samples (Figure 1A). Moreover, we mapped the mutation genes whose mutation frequency is more than 5 on chromosomes, in which red indicated the high frequency mutation sites (Figure 1B).

Table 1
Clinical characteristics of patients from ICGC and cBioPortal database.

Variables	All patients (n = 318)	
	Number (n)	Percent (%)
Age, years		
Median	62	
Range	26-89	
< 65	178	56.0
≥ 65	140	44.0
Gender		
Female	126	39.6
Male	192	60.4
TNM Stage		
Stage I	80	25.2
Stage II	94	29.6
Stage III	37	11.6
Stage IV	107	33.6
Project		
TCGA-CHOL	32	10.0
BTCA-JP	136	42.8
BTCA-SG	48	15.1
SMMU	102	32.1
TMB, mut/Mb		
Median	1.25	
Range	0.03-54.74	
<10	306	96.2
≥10	12	3.8

PPI Network of Mutant Genes

A total of top 200 most frequency mutated genes were selected to construct the protein-to-protein network. We performed the PPI network-based analysis with STRING database to determine mutated genes functionally interacted together involved in tumorigenesis. The networks were also visualized by Cytoscape (Figure 2A). The top hub genes with the highest clustering included TP53, PIK3CA, KRAS, NRAS, PTEN, ANK2, SPTA1, ANK3 and ARID1A (Figure 2B).

Functional Pathway Analysis of Hub Mutant Signature

We conducted functional pathway analysis of the top 300 mutant genes in R software. Figure 2C shows the top 30 enriched GO terms which were associated with regulation of multicellular organismal signaling, ion transmembrane transport, transmembrane transporter complex and ion gated channel activity. In addition, KEGG pathway analysis demonstrated mutant genes enriching in several signaling pathways on malignancy, including PI3K-AKT pathway, MAPK pathway, proteoglycans in cancer or calcium signaling pathway (Figure 2D).

Prognostic Signature of Mutant Genes

To explore the prognostic role of mutation genes on iCCA, we utilized univariate Cox regression to analyze the survival of patients. The patients were categorized into wild-group and mutation- group according to the mutation statue of gene. Thirty significant mutant genes associated with OS were obtained and Kaplan-Meier analysis was used to assess the prognostic value (Supplementary Table 1 and Supplementary Figure 1 and 2). Furthermore, we obtained a 12- prognostic mutant gene signature by using the Lasso Cox method (Figure 3A and 3B). We further utilized multivariate Cox regression analysis to establish a model including six mutant genes as a predictor of survival in iCCA (Figure 3C). In the multivariate Cox analysis, regression coefficients were weighted for the six mutant genes, and the risk prediction model was established. The MRS was calculated as follows: MRS = (0.9772 × CDC27) + (3.3262 × AAK1) + (1.0356 × TP53) + (0.8040 × RBM10) + (0.5645 × KRAS) + (1.4581 ×IPO5) (Table 2). According to MRS value, the patients were divided into high- and low-risk group. Patients with low-risk group showed more significant survival benefit (P <0.001; Figure 3D). Our results revealed that 1, 3, 5-year survival of the MRS to predict prognosis yielded an AUC value of 0.759, 0.732, and 0.728 respectively (Figure 3E), indicating its high prediction efficiency.

Table 2
Clinical characteristics of patients.6- mutation gene risk signature from multivariable Cox regression analysis.

Gene	Coefficient	HR	95% CI	P-value
CDC27	0.9772	2.657	1.463-4.826	0.001333
AAK1	3.3263	27.835	6.606-117.292	5.83e-06
TP53	1.0356	2.817	2.050-3.871	1.70e-10
RBM10	0.8040	2.235	1.001-4.987	0.049659
KRAS	0.5645	1.759	1.227-2.520	0.002095
IPO5	1.4581	4.298	1.743-10.600	0.001547
Abbreviations: CI, confidence interval; HR, hazard ratio.				

The Prognostic Impact of TMB and TNM Stage on iCCA Patients

Next, we analyzed the effect of TMB on prognosis. The median TMB was 1.29 mutations/Mb (rang, 0.03–54.74 mutations/Mb). We further analyzed the prediction performance of TMB on OS. The AUC of ROC for TMB on 1-, 3-, and 5-year survival was 0.776, 0.685, and 0.621, respectively (Figure 4B). The threshold value of TMB were calculated by 3- year ROC curve analysis with the maximum of Youden index. We found that the AUC would achieve maximum when cut-off value of TMB was 1.29. Therefore, we defined 1.29 mutations/Mb as the cut-off value, classifying TMB >1.29 mutations/Mb as high group and TMB ≤1.29 mutations/Mb as low group. KM plotter of survival analysis showed that OS was significantly decreased in patients with high TMB compared to low TMB ($P < 0.001$; Figure 4A). We also explored the relationship between TMB and the prognostic mutant genes, indicating that TMB was moderately correlated with PIKFYVE (r=0.31) and RGS3 (r=0.34) (Supplementary Figure 3). In addition, we explored the prognostic role of TNM stage on iCCA. The results demonstrated that TNM stage had significant correlation with OS, but AUC of TNM stage on 1, 3, 5-year was 0.582, 0.641, and 0.628 respectively, which indicated a worse prognostic performance compared with MRS or TMB (Figure 4C and 4D).

Gene Set Enrichment Analysis

We also performed to visualize the enriched biological processes on the different MRS or TMB group by GSEA. The results indicated that patients in MRS high group were prone to have associations with innate immune response, negative regulation of cell death, positive regulation of immune system process, T cell

activation, MAPK signaling pathway and pathway in cancer (Figure 5A and 5B). In addition, TMB high group was enriched in positive regulation of immune system process and pathway in cancer. These results demonstrated that patients with high MRS or high TMB upregulated crosstalk involved in tumor signaling pathway and immune system (Figure 5C).

Construction and Evaluation of the Nomogram

To find the independent biomarker for prognosis, the univariate Cox regression was applied to analyze the association of factors (e.g., MRS, age, gender, TNM stage and TMB) with OS (as shown in Table 3). Furthermore, the results of multivariate Cox regression analysis demonstrated that TNM stage, MRS and TMB were regarded as independent risk predictors for iCCA (Table 2). We constructed a predictive nomogram based on the risk factors including MRS, TMB and TNM stage (Figure 6A). The C-index value for the nomogram was 0.721 (95% CI, 0.613-0.829). The calibration curve indicated that the observed and predicted values were satisfactorily consistent in predicting OS (Figure 6B).

Table 3
Univariable and multivariable analysis of overall survival.

	Univariable			Multivariable		
	HR	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**
Age, years						
< 65						
≥ 65	0.869	0.643-1.175	0.362			
Gender						
Female						
Male	0.686	0.787-1.439	0.686			
TNM Stage						
I			0.000			0.000
II	1.470	0.938-2.305	0.093	1.717	1.092-2.698	0.019
III	1.743	0.987-3.079	0.055	2.078	1.172-3.685	0.012
IV	2.716	1.799-4.101	0.000	2.963	1.957-4.485	0.000
TMB, mut/Mb						
Low						
High	1.967	1.454-2.660	0.000	1.500	1.085-2.073	0.014
MRS						
Low						
High	2.771	2.035-3.773	0.000	2.448	1.758-3.409	0.000
Abbreviations: CI, confidence interval; HR, hazard ratio.						

Discussion

In this study, we explored the role of mutation gene signature and TMB on survival in patients with iCCA. First, WES data of iCCA from two public database (ICGC and cBioportal) were acquired and frequent mutant genes were identified. Next, univariate, Lasso, and multivariate Cox regression analysis were used to further screen hub prognostic mutant signature and establish a mutation risk model for predicting prognosis. Afterward, the prognostic role of MRS was confirmed, PPI, GO, KEGG and GSEA analysis were conducted to reveal the potential cancer-related crosstalk involved. We also found that higher TMB was associated with poor prognosis. Furthermore, MRS, TMB and TNM stage were confirmed as independent

predictors for OS. We constructed a reliable nomogram model based on the risk factors for OS with a satisfactory performance.

Gene mutations are ubiquitous in tumorigenesis and development of iCCA. Previous studies have reported comprehensive molecular alterations in biliary tract cancers [18, 29]. In our study, the most relevant mutation was TP53 (26.7%), followed by TTN (20.7%), KRAS (19.1%), MUC2 (14.5%) and ARID1A (12.9%), which was consistent with Cao et al [31]. We also found some special mutation genes of iCCA in our study, such as IDH1 (7.5%), BAP1 (9.1%), PBRM1 (7.2%) and EPHA2 (7.9%). Jiao et al preformed exome sequencing on iCCA and found that frequent genes (such as BAP1, ARID1A and PBRM1) mutations in chromatin-remodeling pathway. Our results are consistent with previous reports [32]. We also attempt to link genetic alternations to prognosis of patients, and found that patients with some gene mutations exhibited a borderline significance in worse outcomes compared to wide-type patients. Remarkably, on the basis of these prognostic factors, we developed a mutation risk score to predict survival. The MRS of our model was calculated based on 6 hub prognostic mutant genes (CDC27, AAK1, TP53, RBM10, KRAS and IPO5). This MRS model shows high predictive accuracy for OS, and provides a reliable tool for the prognosis predictor, which will be further applied in the clinical practice. To shed light on the potential molecular mechanisms underlying iCCA, the functional pathway and GSEA analysis were performed. The result of functional enrichment pathway indicated these prognostic mutation genes were closely correlated to cancer associated signaling pathways, such as cancer development and immune related pathway. In addition, GSEA also showed that high MRS group enriched with immune related pathway.

In recent years, the extensive indications of immune checkpoint inhibitor (ICI) therapy have developed the treatment of patients with advanced-stage cancers. However, the satisfactory effect of ICI is limited to a minority of patients. TMB, a novel predictive biomarker, is considerable to predict clinical response to ICI and contribute to recognize patients who will obtain therapeutic benefit [19]. The higher the TMB, as generally believed, means more tumor antigen, which is beneficial to activating the body's immune function. Previous studies involving TMB mostly focused on the predictive capacity to the efficiency of ICIs, and showed robust correlation between higher TMB and better response to ICIs therapy. However, limited researches explored its prognostic value on iCCA. Many studies have indicated a relationship between TMB and survival in cancers. Owada-Ozaki et al from Japan found that higher TMB was correlated with shorter disease-free survival in NSCLC patients [33]. A study from China demonstrated that in HCC patients who had received radical resection, patients with higher TMB tend to have higher recurrence risk rate, and it was an independent risk factors of RFS [34]. We identified the median number of TMB in iCCAs was 1.25 (range 0.03–54.74). A large-scale examination of TMB on iCCA patients has been reported by Cao et al [31]. They analyzed the frequency and type of genetic aberrations in detail by comprehensive genomic profiling, and found the genomic heterogeneity between eastern and western patients of iCCA, but the relationship between TMB and prognosis was not mentioned. Besides, Tian et al investigated the comprehensive genomic features of Chinese patients with iCCA, and explored the relationship between TMB and some gene alternations [35]. It should be noted that TMB of their cohort was more than that in our study. The reason for this is that we only counted non-synonymous variants.

In consistent with previous findings on other tumors, our results showed that higher TMB was correlated with poor prognosis in patients. Therefore, we concluded that TMB had divergent survival difference. In addition, our results indicated that a prognostic model incorporating TMB is more than likely to improve prognostication and risk stratification in iCCA.

We explored the potential prognostic role of MRS and TMB on iCCA in this study, and found that the prognostic performance of the predictive model of TMB or MRS was better than that of TNM stage. Furthermore, the results of multivariate analysis indicated that TMB, MRS and TNM stage are independently prognostic factors in iCCA. Apart from new exploration on prognostic value of MRS and TMB, several drawbacks should be mentioned in our study. First of all, the mutation data of iCCA were extracted from public database, which only included the specimen performed WES. Targeted sequencing data was not considered in our study. more WES data from clinical patients should be incorporated to reduce selective bias. Secondly, the underlying mechanism behind the prognostic MRS and TMB in iCCA should be further investigated. Further experiments both in vitro and in vivo are required to support the present results. Finally, this study didn't prove the specific mutation genes whether led to the abnormal gene expression, which deserve further exploration.

Conclusion

In summary, this study demonstrated that mutation genes signature and TMB were associated with prognosis in patients with iCCA. The landscape of mutations was draw and the most commonly mutant genes were summarized. We further developed a risk model based on the prognostic mutation genes, and found that MRS of the model and TMB had divergent survival impacts in patients, which could be incorporated in prognostication on iCCA patients. More importantly, on basis of independent risk factors such as TNM stage, MRS and TMB, we constructed a reliable nomogram model for OS in iCCA.

Abbreviations

iCCA: intrahepatic cholangiocarcinoma; TMB: tumor mutation burden; MRS: mutation related signature; ICI: immune checkpoint inhibitor; WES: Whole exome sequencing; ICGC: International Cancer Genome Consortium; GO; Gene Ontology; KEGG: Kyoto Gene and Genome Encyclopedia; Lasso: least absolute shrinkage and selection operator regression; GSEA: Gene Set Enrichment Analysis; PPI: Protein-Protein Interaction; HR: Hazard ratio; BP: Biological process; MF: Molecular function; CC: Cellular component; KM: Kaplan-Meier; ROC: receiver operating characteristic; AUC: The area under the curve; CDC27: Cell Division Cycle 27; AAK1: AP2 Associated Kinase 1; TP53: Tumor Protein P53; RBM10: RNA Binding Motif Protein 10; IPO5: Importin 5; HCC: Hepatocellular carcinoma; NSCLC non-small cell lung carcinoma; AJCC: American Joint Committee on Cancer.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Availability of data and materials

All WES data and clinical information used in this study were acquired from ICGC portal (http://dcc.icgc.org/releases/current/Projects) (up to June 10, 2019) and cBioPortal (http://www.cbioportal.org) (up to June 10, 2019).

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China, No. 81572420, the Key Research and Development Plan of Shaanxi Province (No. 2017ZDXM-SF-055, No. 2017KW-060), the General Project of Natural Science Basic Research Plan of Shaanxi Province (No. 2020SF-070) and Clinical Research Fund of the First Affiliated Hospital of Xi'an Jiaotong University (XJTU1AF-CRF-2018-022). The funds mentioned were used for the design of the study, the collection, interpretation of data and bioinformatic analysis as well as the writing of the manuscript.

Authors’ contributions

DZ, ZG and RZ designed this work; RZ, QL, ZJ, JS and JZ performed the WES sequencing data bioinformatic analysis and statistical analysis; RZ and CC were responsible for interpretation, and drafted of the article; DZ and ZG revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank ICGC portal and cBioPortal database for sharing large amounts of data.

References

1. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, Cardinale V, Carpino G, Andersen JB, Braconi C et al: Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol 2020, 17(9):557-588.
2. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, Nellore V, Kongpetch S, Ng AWT, Ng LM et al: Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. Cancer Discov 2017, 7(10):1116-1135.
3. Mavros MN, Economopoulos KP, Alexiou VG, Pawlik TM: Treatment and Prognosis for Patients With Intrahepatic Cholangiocarcinoma: Systematic Review and Meta-analysis. JAMA Surg 2014, 149(6):565-574.
4. Buettner S, Galjart B, van Vugt JLA, Bagante F, Alexandrescu S, Marques HP, Lamelas J, Aldrighetti L, Gamblin TC, Maithel SK et al: Performance of prognostic scores and staging systems in predicting long-term survival outcomes after surgery for intrahepatic cholangiocarcinoma. J Surg Oncol 2017, 116(8):1085-1095.
5. Zhang H, Yang T, Wu M, Shen F: Intrahepatic cholangiocarcinoma: Epidemiology, risk factors, diagnosis and surgical management. Cancer Lett 2016, 379(2):198-205.
6. Weber SM, Ribero D, O'Reilly EM, Kokudo N, Miyazaki M, Pawlik TM: Intrahepatic cholangiocarcinoma: expert consensus statement. HPB (Oxford) 2015, 17(8):669-680.
7. Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ: Cholangiocarcinoma - evolving concepts and therapeutic strategies. Nat Rev Clin Oncol 2018, 15(2):95-111.
8. Chan KM, Tsai CY, Yeh CN, Yeh TS, Lee WC, Jan YY, Chen MF: Characterization of intrahepatic cholangiocarcinoma after curative resection: outcome, prognostic factor, and recurrence. BMC Gastroenterol 2018, 18(1):180.
9. Spolverato G, Kim Y, Alexandrescu S, Marques HP, Lamelas J, Aldrighetti L, Clark Gamblin T, Maithel SK, Pulitano C, Bauer TW et al: Management and Outcomes of Patients with Recurrent Intrahepatic Cholangiocarcinoma Following Previous Curative-Intent Surgical Resection. Ann Surg Oncol 2016, 23(1):235-243.
10. Zhang XF, Beal EW, Bagante F, Chakedis J, Weiss M, Popescu I, Marques HP, Aldrighetti L, Maithel SK, Pulitano C et al: Early versus late recurrence of intrahepatic cholangiocarcinoma after resection with curative intent. Br J Surg 2018, 105(7):848-856.
11. Croce CM: Oncogenes and cancer. N Engl J Med 2008, 358(5):502-511.
12. Ding L, Bailey MH, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, Gibbs DL, Weerasinghe A, Huang KL, Tokheim C et al: Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. Cell 2018, 173(2):305-320.e310.
13. Zhang C, Zheng Y, Li X, Hu X, Qi F, Luo J: Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. Ann Transl Med 2019, 7(18):427.
14. Smith JC, Sheltzer JM: Systematic identification of mutations and copy number alterations associated with cancer patient prognosis. eLife 2018, 7.

15. Buscail L, Bournet B, Cordelier P: Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. Nat Rev Gastroenterol Hepatol 2020, 17(3):153-168.
16. Chae H, Kim D, Yoo C, Kim KP, Jeong JH, Chang HM, Lee SS, Park DH, Song TJ, Hwang S et al: Therapeutic relevance of targeted sequencing in management of patients with advanced biliary tract cancer: DNA damage repair gene mutations as a predictive biomarker. Eur J Cancer 2019, 120:31-39.
17. Ma B, Meng H, Tian Y, Wang Y, Song T, Zhang T, Wu Q, Cui Y, Li H, Zhang W et al: Distinct clinical and prognostic implication of IDH1/2 mutation and other most frequent mutations in large duct and small duct subtypes of intrahepatic cholangiocarcinoma. BMC cancer 2020, 20(1):318.
18. Weinberg BA, Xiu J, Lindberg MR, Shields AF, Hwang JJ, Poorman K, Salem ME, Pishvaian MJ, Holcombe RF, Marshall JL et al: Molecular profiling of biliary cancers reveals distinct molecular alterations and potential therapeutic targets. J Gastrointest Oncol 2019, 10(4):652-662.
19. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A et al: Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019, 51(2):202-206.
20. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, Peters S: Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol 2019, 30(1):44-56.
21. Steuer CE, Ramalingam SS: Tumor Mutation Burden: Leading Immunotherapy to the Era of Precision Medicine? J Clin Oncol 2018, 36(7):631-632.
22. Wang Z, Duan J, Cai S, Han M, Dong H, Zhao J, Zhu B, Wang S, Zhuo M, Sun J et al: Assessment of Blood Tumor Mutational Burden as a Potential Biomarker for Immunotherapy in Patients With Non-Small Cell Lung Cancer With Use of a Next-Generation Sequencing Cancer Gene Panel. JAMA Oncol 2019, 5(5):696-702.
23. Felip E, Navarro A, Callejo A, Marti AM, Cedres S, Pardo N, Ros J, Assaf JD, Pedrola A, Viaplana C et al: Whole exome sequencing (WES) of non-small cell lung cancer (NSCLC) for tumor mutational burden (TMB) analysis and long-term benefit to immune checkpoint inhibitors (ICIs). J Clin Oncol 2019, 37:9071-9071.
24. Dummer R, Brase JC, Garrett J, Campbell CD, Gasal E, Squires M, Gusenleitner D, Santinami M, Atkinson V, Mandalà M et al: Adjuvant dabrafenib plus trametinib versus placebo in patients with resected, BRAF-mutant, stage III melanoma (COMBI-AD): exploratory biomarker analyses from a randomised, phase 3 trial. Lancet Oncol 2020, 21(3):358-372.
25. Greally M, Chou JF, Chatila WK, Margolis M, Capanu M, Hechtman JF, Tuvy Y, Kundra R, Daian F, Ladanyi M et al: Clinical and Molecular Predictors of Response to Immune Checkpoint Inhibitors in Patients with Advanced Esophagogastric Cancer. Clin Cancer Res 2019, 25(20):6160-6169.
26. Schrock AB, Ouyang C, Sandhu J, Sokol E, Jin D, Ross JS, Miller VA, Lim D, Amanam I, Chao J et al: Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. Ann Oncol 2019, 30(7):1096-1103.

27. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, Hama N, Hosoda F, Urushidate T, Ohashi S et al: Genomic spectra of biliary tract cancer. Nat Genet 2015, 47(9):1003-1010.
28. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, Nellore V, Kongpetch S, Ng AWT, Ng LM et al: Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. Cancer Discov 2017, 7(10):1116-1135.
29. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JS, Zhao X, Li Y, Li Q, Wang H et al: Mutational landscape of intrahepatic cholangiocarcinoma. Nat Commun 2014, 5:5696.
30. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, Schrock A, Campbell B, Shlien A, Chmielecki J et al: Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med 2017, 9(1):34.
31. Cao J, Hu J, Liu S, Meric-Bernstam F, Zhao H: Intrahepatic Cholangiocarcinoma: Genomic Heterogeneity Between Eastern and Western Patients. JCO Precis Oncol. 2020 Jun 1;4:PO.18.00414. doi: 10.1200/PO.18.00414. PMID: 32923885; PMCID: PMC7446410.
32. Jiao Y, Pawlik TM, Anders RA, Selaru FM, Streppel MM, Lucas DJ, Niknafs N, Guthrie VB, Maitra A, Argani P et al: Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas. Nat Genet 2013, 45(12):1470-1473.
33. Owada-Ozaki Y, Muto S, Takagi H, Inoue T, Watanabe Y, Fukuhara M, Yamaura T, Okabe N, Matsumura Y, Hasegawa T et al: Prognostic Impact of Tumor Mutation Burden in Patients With Completely Resected Non-Small Cell Lung Cancer: Brief Report. J Thorac Oncol 2018, 13(8):1217-1221.
34. Cai H, Zhang Y, Zhang H, Cui C, Li C, Lu S: Prognostic role of tumor mutation burden in hepatocellular carcinoma after radical hepatectomy. J Surg Oncol 2020, 121(6):1007-1014.
35. Tian W, Hu W, Shi X, Liu P, Ma X, Zhao W, Qu L, Zhang S, Shi W, Liu A et al: Comprehensive genomic profile of cholangiocarcinomas in China. Oncol Lett 2020, 19(4):3101-3110.

Figures

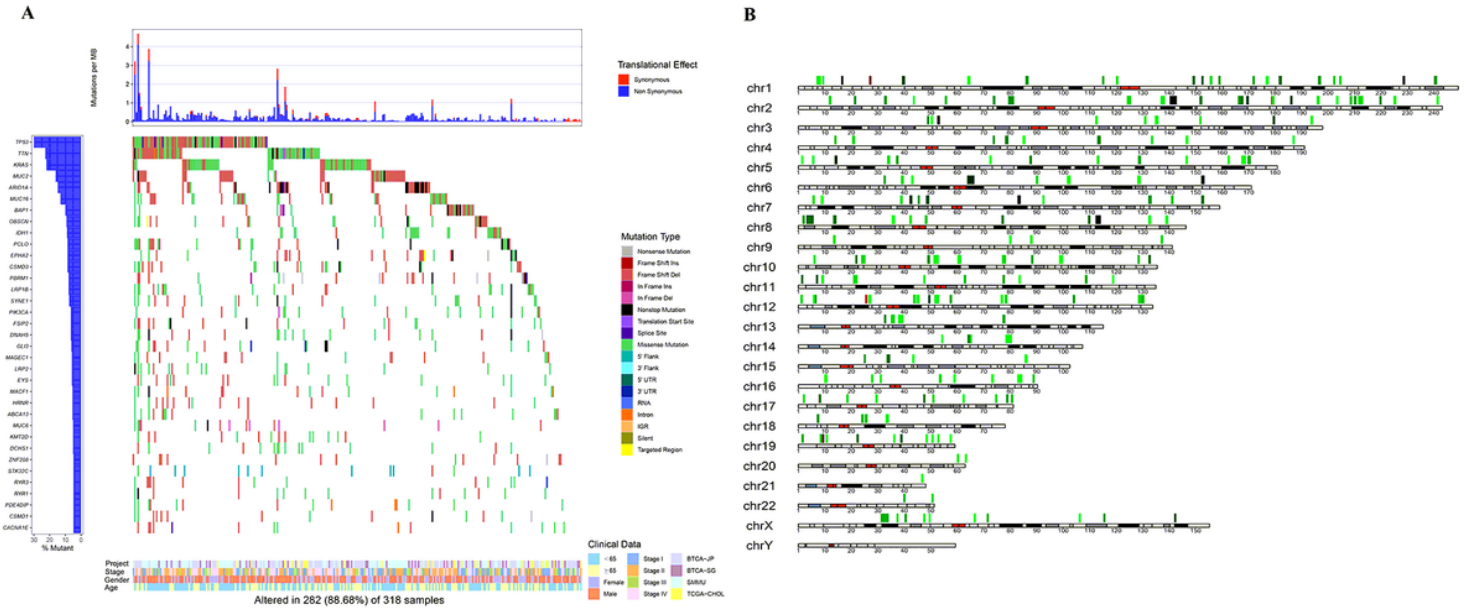

Figure 1

The mutational landscape of iCCA. (a) The top 35 frequent mutated genes in the waterfall plot; (b) The specific mutation location on chromosomes were mapped. Green stand for the low-frequency mutation sites, while red represented the high-frequency.

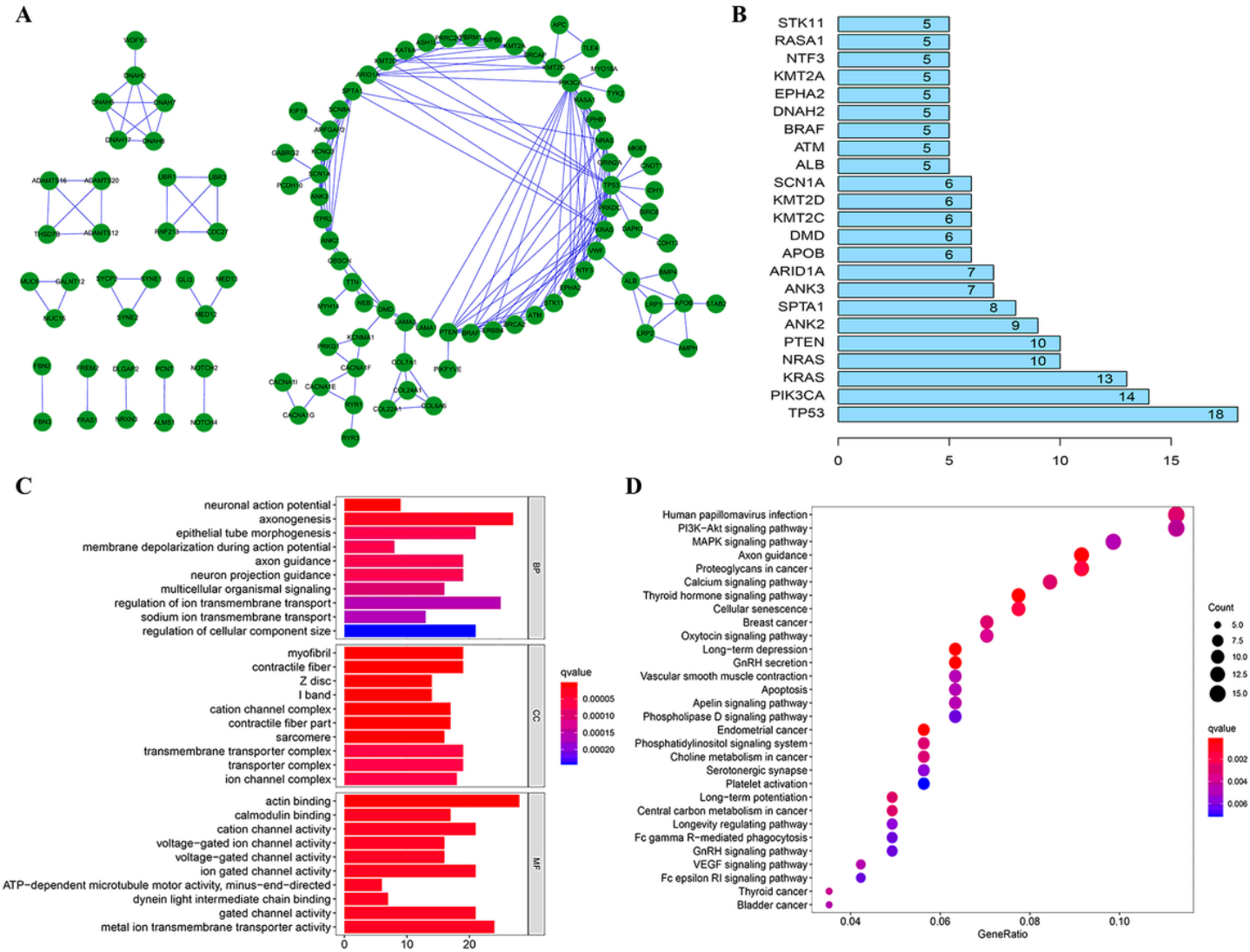

Figure 2

Protein to protein interaction (PPI) network and functional enrichment pathway of mutation genes. (a) PPI network of mutation genes; (b) the node with edge of >20 was considered as hub genes; (c) Gene Ontology analysis. The top 30 enriched terms in GO. Cellular Component (CC); Biological Process (BP); Molecular Function (MF); (d) KEGG analysis of mutation genes. 30 KEGG pathway were significantly enriched. The size of circle stands for gene numbers, and the color represents adj. P-value.

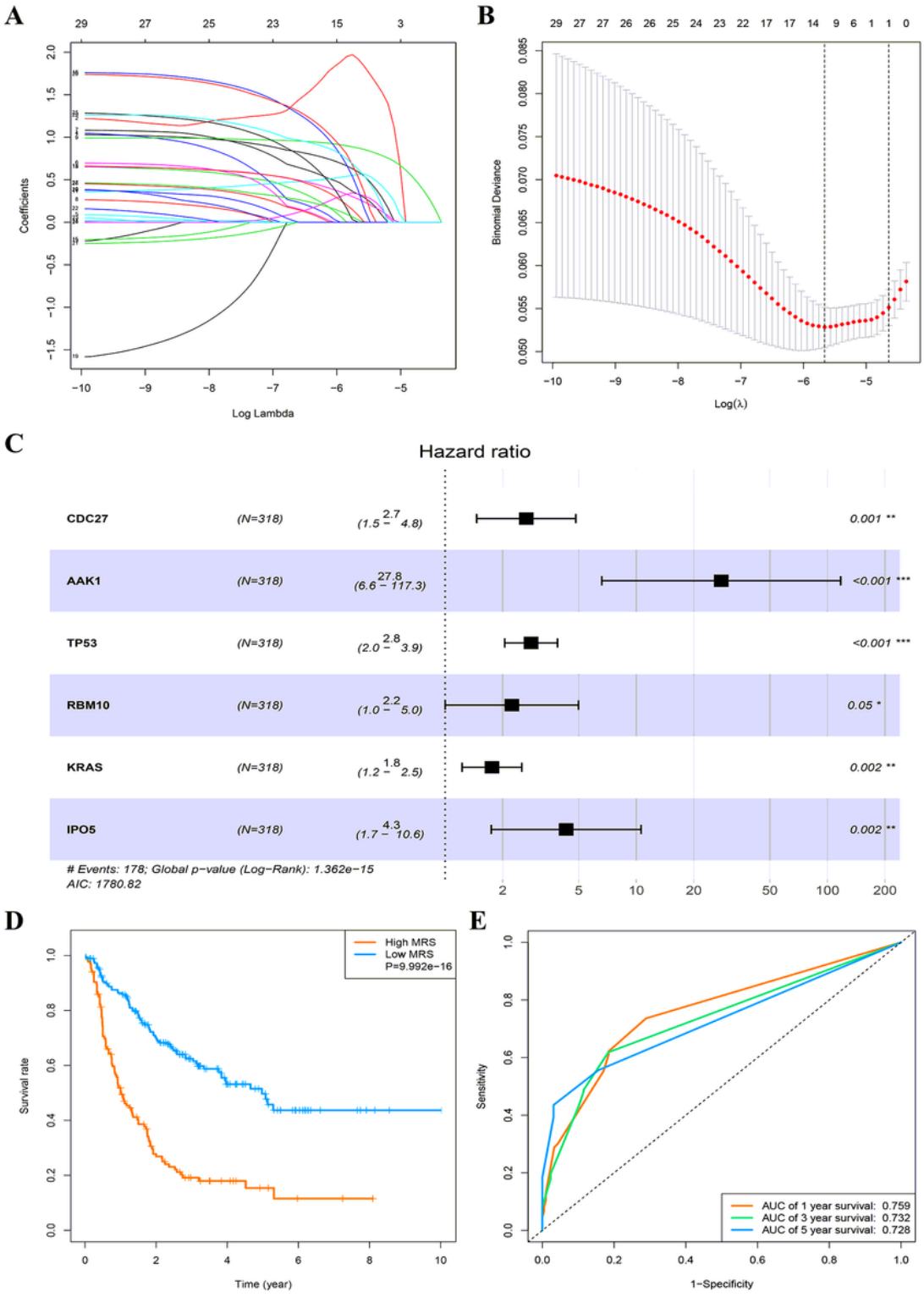

Figure 3

Screening of the optimal MRSs used for predictive model using Lasso regression method (a) Lasso coefficient profiles of mutation genes in iCCA cohort; (b) A coefficient profile plot was generated to find the optimal parameter (lambda); (c) Forest plots showed hazard ratio (HR) of selected prognostic mutation genes generated by multivariate Cox analysis; (d) KM plotter showed the difference between high- and low- MRS group, indicating that high MRS revealed poor survival outcomes; (e) Time dependent

ROC curves of MRS indicated 1, 3, 5-year AUC, showing the satisfactory predictive accuracy of MRS model.

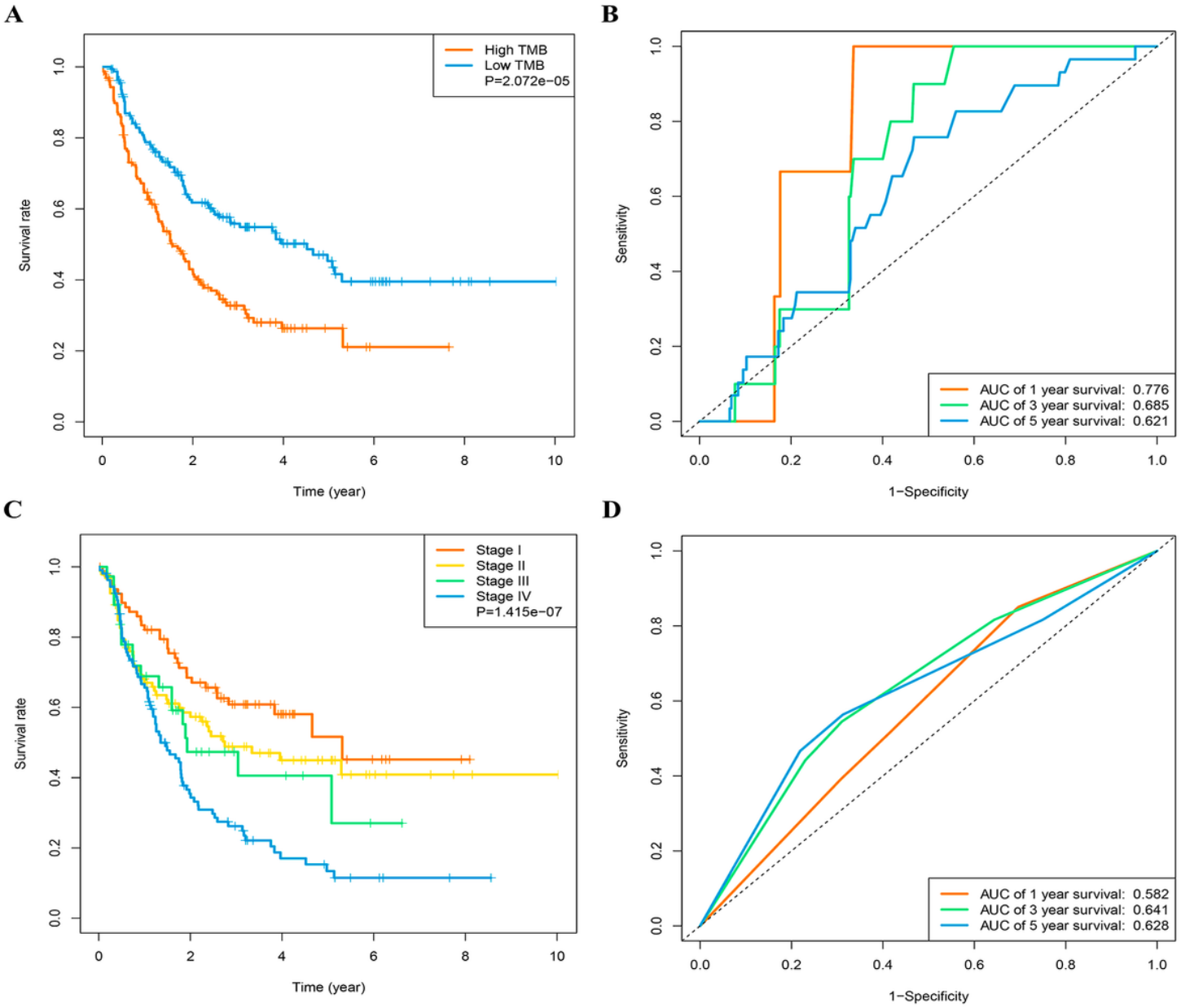

Figure 4

The prognostic impact of TMB on OS for iCCA patients. (a) Kaplan-Meier plot showed the survival difference between high- and low- TMB group, indicating that high TMB revealed poor survival outcomes; (b) Time dependent ROC curves of TMB indicated 1-year/3-year/5-year AUC, showing the satisfactory predictive accuracy of TMB; (c) Kaplan-Meier plot showed significant difference on TNM stage; (d) Time dependent ROC curves of TNM stage.

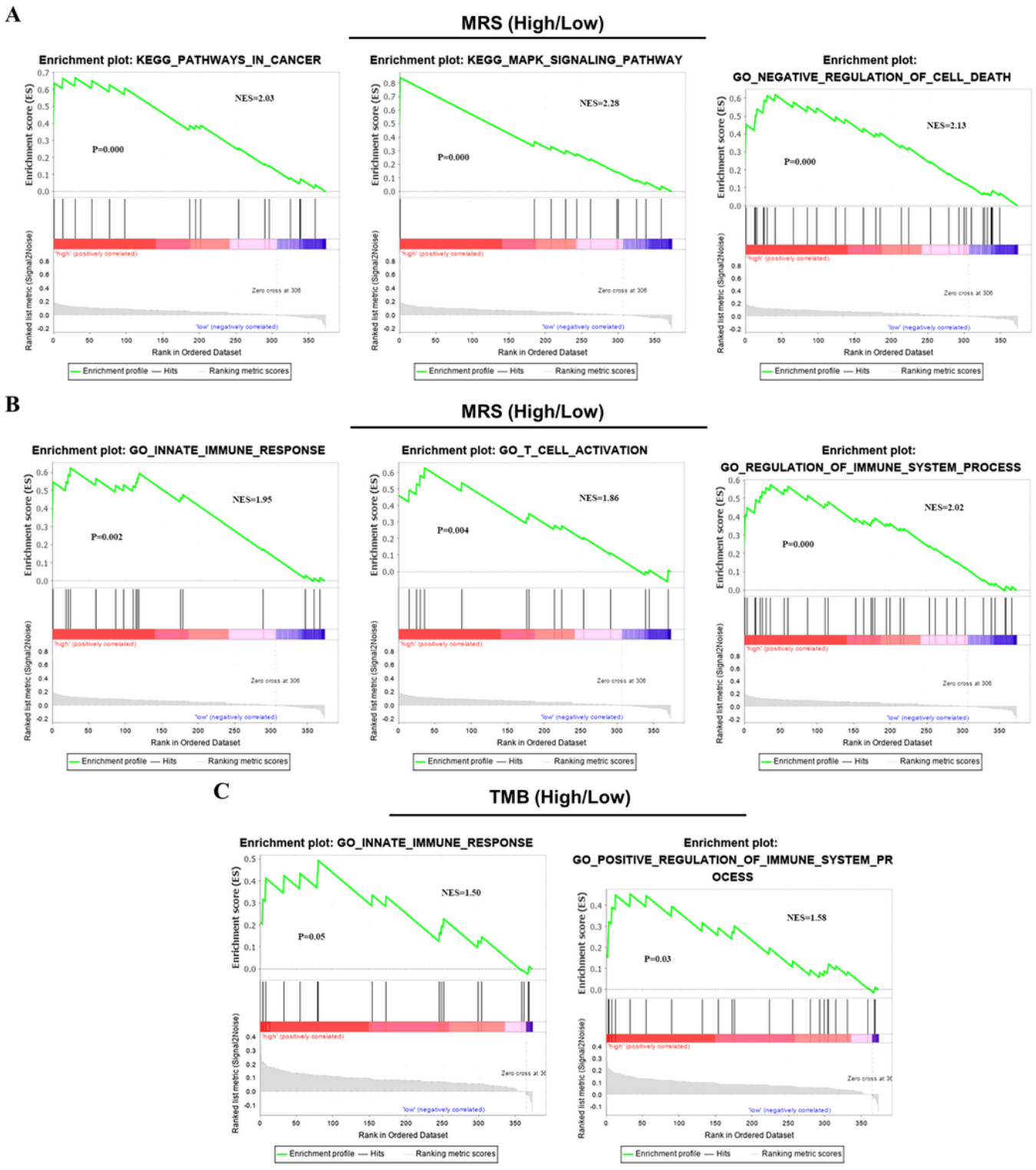

Figure 5

The functional analysis between different MRS or TMB groups by Gene Set Enrichment Analysis (GSEA). (a) (b) Representative KEGG pathways and GO pathways were analyzed in the low-MRS versus high-MRS groups; (c) Representative KEGG pathways and GO pathways were analyzed in the low-TMB versus high-TMB groups.

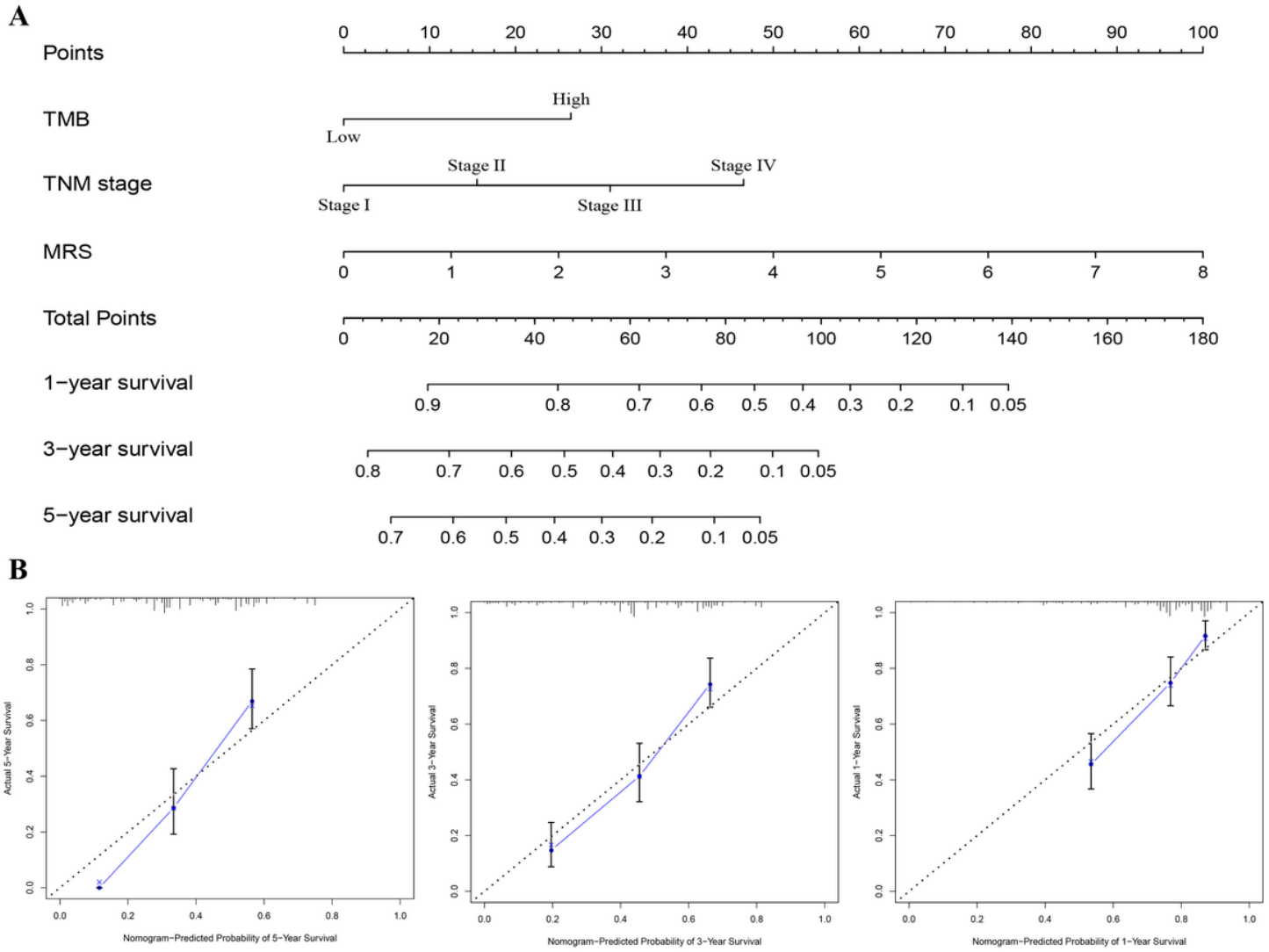

Figure 6

Building of prognostic nomogram for iCCA patients. (a) The predicted 1-, 3-, 5-year survival rates of iCCA based on nomogram using the MRS, TMB and TNM stage; (b) Calibration plots showed the concordance between predicted and actual observation and prediction in 1, 3, 5-year OS.

Supplementary Files

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

- SupplementaryFigure1.tif
- SupplementaryFigure2.tif
- SupplementaryFigure3.tif
- SupplementaryTable1.docx